DESIGN OF FUZZY LOGIC-BASED INTEGRATED ADAPTIVE DECAYED BRAIN EMOTIONAL LEARNING NETWORKS FOR ONLINE TIME SERIES PREDICTION

by

Houssen Salh Ali Milad

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
July 2021

DEDICATION

To my Lord, Allah; to my dear mother, Kadiga, and my honoured father, Salh; to my lovely wife, Nagia, and my beloved sons, Najmeden, Mohamed and Moneab; to my respected brothers and sisters; to my honest friends and colleagues; to all my teachers; and to those who wish me success.

Table of Contents

List of Tables

List of Figures

Abstract

The emotional neural network (ENN) is a new field in artificial intelligence systems (AISs). Although ENN was proven in control applications and successfully solved several control problems, it still suffers severe technical issues concerning prediction. This study aims to design a new model for an intelligent forecasting technique constructed by unified Adaptive Decayed Brain Emotional Learning (ADBEL) combined with a Neo-Fuzzy Neuron (NFN) network. In the literature, the ADBEL network is used to predict time series in online mode. Unlike other popular learning networks such as artificial neural networks (ANNs), the ADBEL network offers lower computational time, less complexity, and fast learning, making it ideal for time series prediction in online applications. This thesis aims to further enhance the ADBEL network's forecasting accuracy through three significant modifications in design. The first modification is its integration with a neo-fuzzy network in the orbitofrontal cortex section. The result is the Neo-Fuzzy integrated Adaptive Decayed Brain Emotional Learning (NF-ADBEL) network. The second modification is the integration with a neo-fuzzy network in two sections: the orbitofrontal cortex section and partially in the amygdala section. This modification leads to a new design: the Expanded Neo-Fuzzy integrated Adaptive Decayed Brain Emotional Learning (ENF-ADBEL) network. The third modification is to design a fuzzy logic-based parameter adjustment model for the ADBEL network, resulting in the F-ADBEL model. The F-ADBEL model can set the learning parameters (namely, alpha, beta, and gamma) of the online mode's ADBEL network. Root mean squared error and correlation coefficient criteria are used to evaluate the models. The chaotic time series, namely the Mackey-Glass, Lorenz, Rossler, Disturbance Storm Time Index, and the Narendra dynamic plant identification problem, were used for applications and validation. Stochastic series such as wind speed and wind power series are applied to validate the designed models mentioned above. Furthermore, we conducted a comparison between the developed models and the ADBEL and other models. The created models were tested in a MATLAB programming environment and showed superior performance compared to other state-of-the-art predictors.

List of Abbreviations and Symbols Used

List of Abbreviations (in the order of appearance in the thesis):

ENN ..... Emotional Neural Network

AISs ..... Artificial Intelligence Systems

ADBEL ..... Adaptive Decayed Brain Emotional Learning

NFN ..... Neo-Fuzzy Neuron

NF-ADBEL .... Neo-Fuzzy Adaptive Decayed Brain Emotional Learning

ENF-ADBEL ... Expanded Neo-Fuzzy Adaptive Decayed Brain Emotional Learning

F-ADBEL ..... Fuzzy Logic-based Parameters Adjustment Model for ADBEL

ANNs ..... Artificial Neural Networks

BELNN ..... Brain Emotional Learning Neural Network

BELIMs ..... Brain Emotional Learning Inspired Models

EMAI ..... Emotionally Motivated Artificial Intelligence

AI ..... Artificial Intelligence

BELBIC ..... Brain Emotional Learning-Based Intelligent Controller

PID ..... Proportional-Integral-Derivative Controller

BELPR ..... Brain Emotional Learning-based Pattern Recognizer

PCA-BEL ..... Principal Component Analysis and Brain Emotional Learning

NFBELPR ..... Neo-Fuzzy Brain Emotional Learning-based Pattern Recognizer

BELPM ..... Brain Emotional Learning-based Prediction Model

SD ..... Steepest Descent

LES ..... Least Square Estimator

D_{st} ..... Disturbance Storm Time Index

ANFIS ..... Adaptive Neuro-Fuzzy Inference System

NMSE ..... Normalized Mean Square Error

TH ..... Thalamus

SC ..... Sensory Cortex

AMY ..... Amygdala

OFC ..... OrbitoFrontal Cortex

REW ..... Reward Signal

BELFIS ..... Brain Emotional Learning-based Fuzzy Inference System

BELRFS ..... Brain Emotional Learning-based Recurrent Fuzzy System

ELiEC ..... Emotional Learning-inspired Ensemple Classifier

α ..... Learning Parameter of Amygdala

STM ..... short-term memory

LTM ..... long-term memory

R_o ..... Reinforcement Signal

a_i ..... Amygdala output

o_i ..... Orbitofrontal cortex output

w_j ..... weight combination

t ..... target

T ..... target

P_{t-j} ..... input data

$\acute{P}t$ ..... model output

m ..... maximum value

w_j ..... weights of Orbitofrontal cortex

v_j ..... weights of Amygdala

$\acute{E}_a$ ..... Amygdala output

E_a ..... Amygdala total output

E_o ..... Orbitofrontal cortex output

P_t ..... target

γ ..... Decay rate

β ..... Learning Parameter of Orbitofrontal cortex

ANFIS ..... Adaptive-Network-based Fuzzy Inference System

MLP ..... Multilayer Perceptron

LLNF ..... Locally Linear Neuro Fuzzy

M-G ..... Mackey-Glass

LZ ..... Lorenz

Ross ..... Rossler

NIP ..... Narendra Identification Plant

TS ..... Time Series

k ..... number of triangular membership function

i^{th} ..... number of neo-fuzzy neuron

$h_{ij}(x_i)$ ..... degree of membership function

c_j ..... center of membership function

d ..... uniform spacing between membership functions

GD ..... gradient descent

E_{AMY} ..... Amygdala final output

E_{OFC} ..... Orbitofrontal output

v_{th} ..... weight of max value sends from TH to AMY

LS ..... limbic system

e ..... error

A_g ..... aggregated fuzzy set

COG ..... Center Of Graph

RMSE ..... root mean squared error

COR ..... correlation coefficient

R^2 ..... correlation coefficient

PI ..... percentage improvement

zoomed ..... means a portion of data displayed

Acknowledgements

I would like to express my sincere thanks to my previous supervisor, Prof. Dr. Mohamed E. El-Hawary, who passed away. For his support and guidance for the most working parts of this thesis. I would like to thank my supervisor, Prof. Dr. Jason Gu, for his guidance, suggestions, and financial support during this work. I am also thankful to my supervisory committee member, Prof.Dr. William Phillips and Prof.Dr. Kamal EI-Sankary, to serve as my Ph.D. committee member. Their support and knowledge were a great help throughout my research, and it is my honour to have them review my work and provide me with valuable recommendations.

My special thanks go to the external examiner Prof.Dr. Adel Merabet provided his valuable time to review and improve the quality of this thesis. I would like to thank the departmental representative, Prof.Dr. Jack Ilow and the chair Dr.Elizabeth Cowley for their valuable time monitoring and arranging my Ph.D. defence. In addition, I would like to thank our collaborators for their support, guidance, and help, especially Dr. Umar Farooq, Dr. Ali Ridha Al-Roomi, and Dr. AbduAllah Algnkawey.

I particularly wish to emphasize my gratitude to the Libyan Ministry for Higher Education and to Dalhousie University for providing financial assistance in the form of scholarships, and I would like to thank Canada for granting me the opportunity to study here.

Finally, I owe deep gratitude to my sons Najmeden, Mohamed and Moneab, my wife Nagia, and my parents Salh and Kadiga for their support and encouragement all along the way.

Chapter 1

INTRODUCTION

Chaotic time series theories share some general characteristics, such as sensitive dependency on initial conditions and non-cyclical and bounded variations. However, the dependency characteristic of time series theories often renders the long-term prediction capacity of chaotic time series nearly impossible. To resolve this issue, if the predictor (model) can identify the chaotic behaviour, then short-term prediction is possible to some extent [1].

Many different techniques were applied to the chaotic times series prediction problem with varying degrees of success. The Artificial Neural Network (ANN) is probably the technique most often used. However, due to its structure and back-propagation training algorithm, there is no guarantee that the training processes will not land in a local minima position. Furthermore, there is an increase in time computational complexity because there is no optimal structure for the number of neurons, number of layers, or activation function suitable for the objective function. These drawbacks affect the reliability and accuracy of the prediction.

Brain Emotional Learning Neural Network (BELNN) has recently emerged as an alternative to classical artificial neural networks for approximating nonlinear functions. BELNN is inspired by both feed-forward neural networks and fast learning and has been applied for time series prediction techniques [2],[3],[4],[5].

1.1 Contributions

The thesis addresses the design of different ADBEL network models based on the integration of neo-fuzzy networks. Summarized below are the contributions which will be further elaborated in the subsequent chapters:

1. The design of a neo-fuzzy integrated adaptive decayed brain emotional learning (NF-ADBEL) neural network is proposed, and its applicability is demonstrated

for online time series prediction problems and other forecasting applications such as wind speed and wind power generation.

2. The design of an Expanded neo-fuzzy integrated adaptive decayed brain emotional learning (ENF-ADBEL) neural network is proposed, and its applicability is demonstrated for online time series prediction problems and other forecasting applications, such as wind speed and wind power generation.

3. A fuzzy-logic-based parameter-adjustment model design to use with the adaptive decayed brain emotional learning (F-ADBEL) network is proposed, and its applicability is demonstrated for online time series prediction problems and other forecasting applications such as wind speed and wind power generation.

4. A comparison is conducted between the designed proposed NF-ADBEL, ENF-ADBEL, F-ADBEL, and ADBEL models and other state-of-the-art models.

In addition to the above, other contributions include:

- Redesigning an ADBEL network for use as a benchmark to compare the proposed models' outcomes.

- Programming and simulating all the proposed models in MATLAB.

- Generating a comparison program in MATLAB to compare the results of the proposed models.

- Identifying time series data for Mackey-Glass, Lorenz, Rossler and Narendra identification plant in MATLAB programming and filing it in the attached Appendix of this thesis to aid the work of future researchers.

1.2 Thesis Outline

The thesis is organized into five main chapters. The current chapter presents an overview of chaotic time series prediction problems and Brain Emotional Learning networks, as well as the study's contributions. The remainder of the thesis is ordered as follows:

Chapter 2: Literature Review

This chapter briefly provides a literature review of Brain Emotional Learning models and various techniques to predict time series.

Chapter 3: Proposed Models

This chapter presents Brain Emotional Learning's proposed models to predict time series data and other nonlinear systems.

Chapter 4: Results and Discussions

This chapter presents the results of the proposed models and discusses their performance in comparison to other state-of-the-art predictors.

Chapter 5: Conclusions and Future Work

This chapter presents the conclusions of the thesis and makes suggestions for future work.

Chapter 2

LITERATURE REVIEW

2.1 Review of Brain Emotional Learning-inspired Models (BELiMs)

Brain emotional learning-inspired models (BELiMS) are computational models adapted to mimic the human brain. From the application's perspective, it can be categorized into three groups, as follows: Brain emotional learning-based prediction models, Brain emotional learning-based control models, and Brain emotional learning-based pattern recognizer models.

1. **Brain emotion learning-based controller models**

   B. Damas, in [6], presented an Emotionally Motivated Artificial Intelligence (EMAI) model that was applied to an artificial soccer team. This model was based on artificial intelligence (AI). EMAI is considered a first attempt to develop a decision-making agent using emotional features that incorporate emotional responses with rational responses. The results were reasonable.

   D. Juan, in [7], showed how emotions and behaviours can be integrated into an adaptive agent structure, using some emotion mechanisms to gain memories from experiences that act as bias mechanisms during decision-making. In terms of their ability to overcome the uncertainty and complexity that most control systems suffer, the Brain Emotional Learning-Based Intelligent Controller (BELBIC) is considered one of the most successful and superior control models [8] of all time. BELBIC has ben applied to different control systems, such as heating and air conditioning [9], Aerospace Launch Vehicle Control [10], Intelligent Modelling and Control of Washing Machines [11], and stepper motor trajectory tracking [12]. Further, BELBIC outperforms several existing controllers by its unique characteristics of simplicity, reliability, and stability.

   A BELBIC integrated with a neuro-fuzzy network controller was applied to track and control the output temperature in an electrically heated micro-heat

exchanger plant. The proposed model outcomes in [13] were compared with a PID controller, with the novel model showing a more remarkable ability to stabilize faster than the PID performance controller and with only minor distortion occurrence.

2. Brain emotionalal learning-based pattern classification models

A pattern classifier refers to a system that uses classification algorithms to classify new observations. These algorithms use either supervised or unsupervised techniques to achieve the task of classification. In supervised mode, the training data set is used to create a map between the characteristics in the data set and classes. Afterwards, the learning function is then validated using the test data set. In unsupervised cases, the data entered is aggregated, and new categories are created automatically based on the distance measurement. Numerous categorization algorithms exist, but not all algorithms perform well for all classification problems. Many algorithms and theories have been used in the literature to solve classification problems such as pattern classification [14], Pattern Recognition and Machine Learning [15], Deep Neural Networks for wireless localization in indoor and outdoor environments [16], visual-tactile fusion for object recognition [17], and optimizing classifier performance for Parkinson's disease detection [18].

It is generally well-known in the field that neural networks (NNs) developed on a biological basis have been widely used for solving problems of pattern recognition in traditional networks, such as MLP [19], chart pattern recognition [20], self-organizing spiking neural network [21], and deep neural networks. Although these networks are based on neuronal activity in mammals, they represent a small-scale model of the cerebral cortex. However, the recent trend of using models in deeper neuronal activity in the mammalian cerebral cortex for solving pattern recognition problems has been garnering increasing interest. Among these studies are: brain emotional learning-based pattern recognizer (BELPR) [22], gene expression microarray classification using PCA-BEL [23], supervised brain emotional learning [24], neural basis computational model of

the emotional brain for online visual object recognition [25], practical and emotional neural networks for pattern recognition problems [26], and neo-fuzzy-supported brain emotional learning-based pattern recognizer (NFBELPR) [27].

3. **Brain emotional learning-based prediction models**

   A brain emotional learning-based predictor is an intelligent digital prediction method characterized by easy learning, fast computation, lack of complexity, and fast response to system inputs and accurate prediction outcomes.

   Mahboobeh, in [2], proposed a new version of a brain emotional learning-based inspired model called the brain emotional learning-based prediction model (BELPM). The structure of the model is depicted in Figure 2.1. BELPM is based on different adaptive networks implemented in various regions of brain emotional learning (BEL), as shown in Figure 2.2, where the learning algorithms used are the steepest descent (SD) and least square estimator (LSE). The goal of this model is to enhance the prediction accuracy of the chaotic behaviour of nonlinear systems and to predict the disturbance storm time (Dst) index. The authors compared their model results to those of the adaptive neuro-fuzzy inference system (ANFIS), with the results indicating that BELPM provides better performance than ANFIS.

2.2 Motivations

Both adaptive decayed brain emotional learning (ADBEL) and neo-fuzzy neurons (NFN) share some characteristics and features. The main ones are listed in the next subsections.

2.2.1 ADBEL Features

1. Less computational time
2. Fast learning
3. Decreased spatial complexity
4. High accuracy

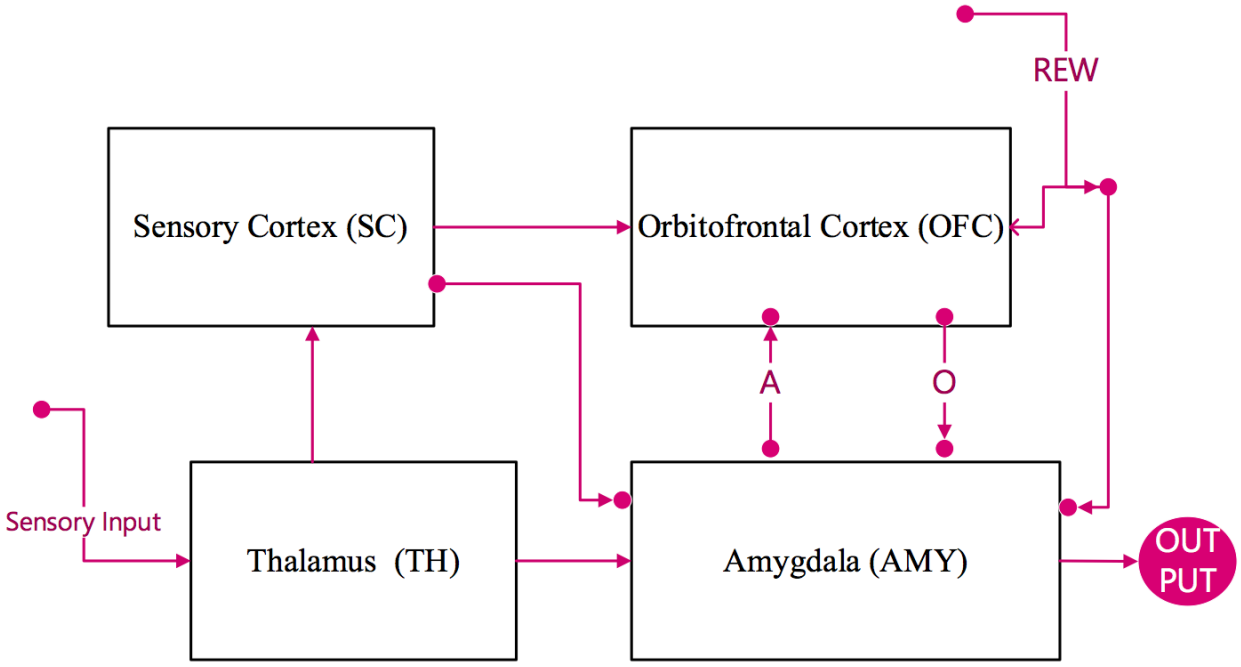

Figure 2.1: Structure of BELPM [2].

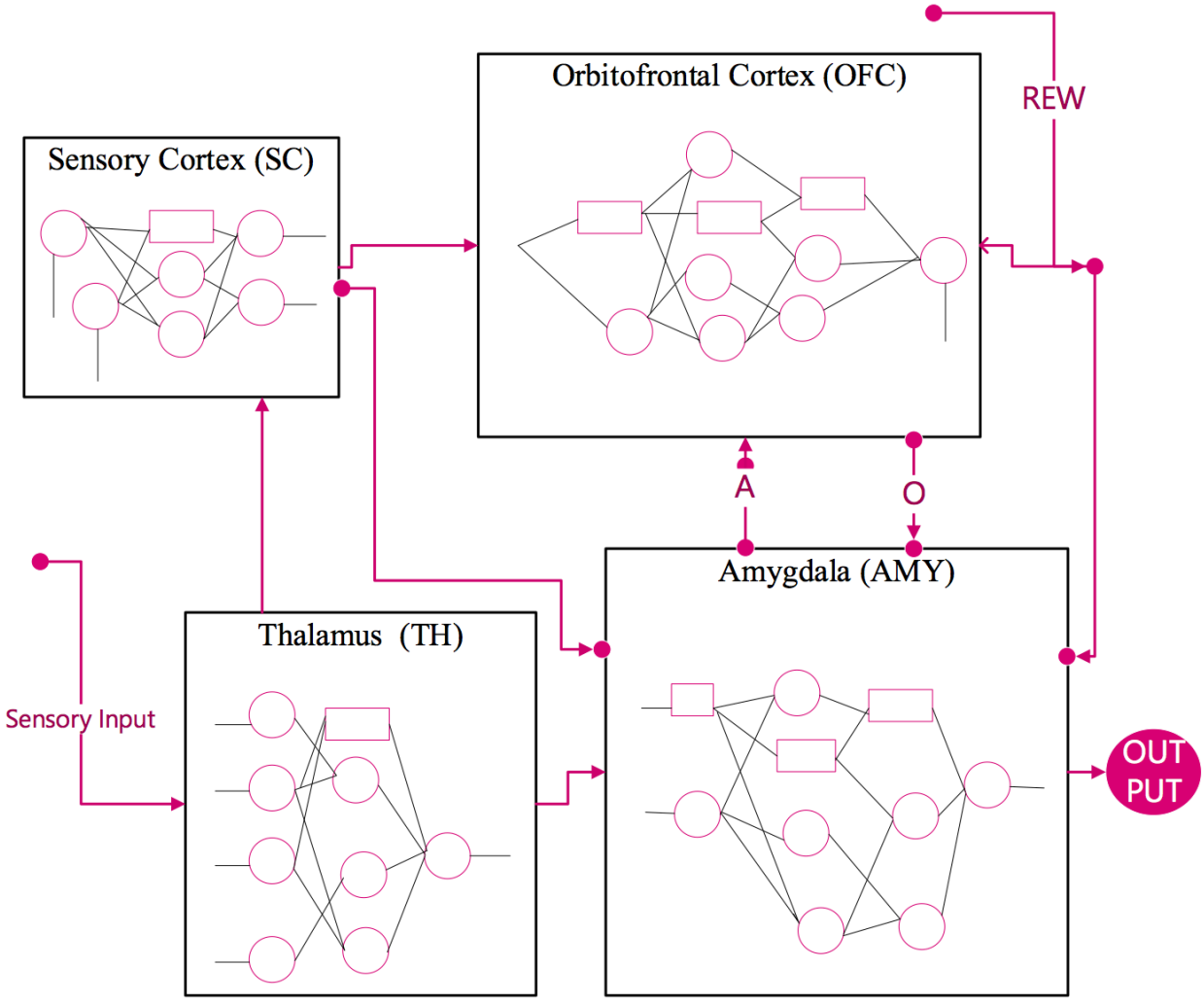

Figure 2.2: Assigning Adaptive Networks to Different Parts of BELPM [2].

2.2.2 Neo-fuzzy Features

1. Simplicity
2. Transparency
3. Accuracy
4. Low computational complexity

The neo-fuzzy-supported adaptive decayed brain emotional learning-based prediction model preserves the features of both networks, while simultaneously improving the performance of the proposed model (NF-ADBEL).

2.3 Review of the Concept of Emotion

Over the centuries, countless interdisciplinary scientists, whether medical scientists, neuroscientists, psychologists or philosophers, have sought to understand the mystery of emotions. They explored the effects of emotion on the human brain in terms of reaction and decision-making, developing argumentative hypotheses based on different grounds and mostly according to their own scientific field of study [28].

Exclusively noteworthy here are cognitive neuroscientists who have endeavoured to sketch and describe the emotional system and its processing mechanisms of operation. One of the earliest neuroscientists who inspired and studied the brain's emotional system in detail is James Papez. In 1937, Papez was able to sketch an explanatory and approximate diagram of the brain connection regions that are now known as the Papez circuit. Papez was eager to learn and study the human brain, and he found his object in [29], which considered the basis for understanding the mystery of the human mind. In 1907 and 1908, the neuroscientist Jacob discovered the visceral brain [29], which interpreted and understood the system of the inner brain. The system structure was based on a study conducted on degenerative diseases in apes, dogs and humans, a phenomenon known as the disease of deterioration and decay of the nerves. Degeneration is aggressive behaviour.

In 1937, Papez conducted a laboratory study on a disease that affects dogs called rabies or frenzy. This disease is characterized by particularly aggressive behaviour.

Papez noticed that this disease is directly related to the presence of deterioration and deformity of the hippocampus. He logically correlated that the deficiency of the hippocampus creates aggressive behaviour, and thus the hippocampus is thought to express emotions, as it is connected automatically to the nervous system.

However, Papez also noticed that, in other cases, stimuli such as the senses of smell, taste and pain are not solely related to the hippocampus. Rather, there are other parts of the brain working together as a center of control for emotions. Accordingly, Papez found the Papez circuit, naming this circuit as the cortical control of emotions.

In 1937, Paul D. MacLean was excited to learn about Papez's circuit and expanded his study to include research on Broca's area. Pierre Paul Broca was an anatomist who studied the problem of the inability to speak, or "aphasia." Broca found that it was caused by lesions in the left frontal cortex. His study is considered the first anatomy of the localization of brain function and has been named after him [30]. Broca's research discovered that the limbic lobe surrounding the posterior brain (the brainstem) was a common structure that occurred in all mammals. Papez's circuit linked the hypothalamus and the limbic lobe. In 1952, MacLean developed a new version of Papez's circuit that included not only the hippocampus and hypothalamus, but also the amygdala and septum [31].

The inclusion of the hippocampus, amygdala, and septum led to the labelling of the "visceral brain." Thus, the limbic lobe and visceral brain represent what is known today as the limbic system. MacLean was certain that the presence of the visceral brain in the limbic system gives a logical explanation for the process of emotion in the knowledge of external stimuli such as olfactory, visual, and audio [31], and that the visceral brain is therefore responsible for the senses and transposes the external stimuli to the limbic lobe.

From the above, it can be concluded that the limbic lobe works as a memory, while the visceral brain acts as a carrier of the external stimuli to the visceral lobe. Both of them then work together as an emotional control center. However, neuroscientists have rejected the philosophy of the limbic system, replacing the theory with the notions of fear and rewards, while preserving the structure of the Papez circuit. Neuroscientists have proposed that different parts of the brain are responsible for

different emotional behaviours [32]. Furthermore, fear is a common and instinctual emotional behaviour that excites both humans and animals. Fear conditioning is defined as learning fearful stimuli to avoid and/or predict penalties.

2.4 Description of Structure, Functions and Learning Algorithms of the Brain Emotional Learning Based-Prediction Model (BELPM)

2.4.1 BELM External Structure

The amygdala-orbitofrontal model is a relatively simple structure [33] and thus serves as the foundation model from which other models have been derived [2], [3], [34]. Although the external structure is almost the same as shown in Figure 2.1, there are some interior design variations. According to the goal, the interior design depends on whether the purpose is to predict, control, or recognize[35],[36],[37],[38]. Hence, the internal differences which we see in the existing models are based on training and learning rules as well as different theoretical theories and methods [39],[26], [24], [40], [41]. The target is to achieve a universal model that applies to multi-objectives, but this achievement has remained elusive.

2.4.2 BELM Internal Structure

In the literature, numerous proposal models have been derived from the amygdala orbitofrontal model. Including models as the Brain Emotional Learning-based Fuzzy Inference System (BELFIS) [5], the Brain Emotional Learning-based Recurrent Fuzzy System (BELRFS) [42], and the Emotional Learning-inspired Ensemble Classifier (ELiEC) [43], the latter for prediction and classification purposes.

2.4.3 BELM Terms

Several studies have been conducted to expose the "mystery" of learning feelings, including anatomical and behavioural aspects. Numerous studies [33],[44],[45] have concluded that, in terms of fear conditioning, the limbic system is primarily responsible for the process of learning emotions. Below, we will define the general concepts of the limbic system's names and other parts of the brain that may have some form of relationship with the main parts of the limbic system.

1. Thalamus (TH)

   The thalamus (TH) is a sensor for seeing, touching, and tasting. It is considered a relay station that directs these stimuli to the cortex's appropriate area in other regions of the brain. As emotions are very contingent on this area of the brain, the thalamus is considered a gateway to the limbic system and is responsible for providing the limbic system with all the available and received information in the form of emotional stimuli [33], and [46]. The thalamus sends the received data along a short path to the amygdala and then directly to the sensory cortex.

2. Sensory Cortex (SC)

   This region is responsible for receiving the information sent to it by the thalamus. The role of the sensory cortex is to forward the received information to the amygdala and orbitofrontal cortex [33], [47].

3. Amygdala (AMY)

   From a medical perspective, the amygdala is sometimes called the aggregation center. Experimentally, it has been shown that, when stimulated, the amygdala can produce feelings of anger, violence, fear, and anxiety. We can thus propose here that anger leads to violence and fear leads to anxiety. This part of the amygdala is sometimes presented by a plus sign. On the other hand, if the amygdala is destroyed, it is represented by a negative sign, whereby it can cause a very mellowing effect. This mellowing effect was noticed by the psychologist Dr. Kluver and the neurosurgeon Dr. Bucy, who subsequently termed it the Kluver-Bucy syndrome [48].

   The Kluver-Bucy syndrome appears when there is bilateral destruction of the amygdala. Symptoms of the syndrome include hyperorality, hyper-sexuality, and increased disinhibited behaviour that results in dismissing risk and thus engaging in dangerous and reckless pursuits. In the Kluver-Bucy syndrome, both sides (bilateral) of the amygdala are destroyed. On the other hand, if the amygdala is stimulated and produces fear and anxiety based on the type of stimuli, the result is anxiety disorders or an anxiety attack. Medication such as Benzodiazepine (benzos) is generally prescribed for extreme disorders.

Pharmacologically, benzos function similarly to alcohol. As is generally well-known, when people consume too much alcohol, they may experience symptoms of hyperorality and hyper-sexuality, as well as disinhibited behaviour.

It is worth mentioning that the two parts of the amygdala are connected by a curving structure around the thalamus, called the hippocampus. The hippocampus comprises the central region and is one of the most important parts of the limbic system [33], as it has access to other limbic system regions. Further, the hippocampus is mono-directional with the thalamus and sensory cortex, whereas it is dual-directional with the orbitofrontal cortex. Part of its multi-function includes memory that involves saving, storing, and releasing emotional stimuli [44], [45]. It also plays a key role in analysis and emotional learning and contributes to decision-making and emotional response, with the help of the orbitofrontal cortex and other sections, such as the hippocampus and reinforcement signals.

One of the most important characteristics of the amygdala's emotional learning is that it is permanent and monotonic [49]. Furthermore, the learning parameter of the amygdala, the so-called (α), is proportional to the strength of input stimuli and reaches the amygdala along a short path. This stimulus has all the information (based on the role of the orbitofrontal cortex) to inhibit an inappropriate response from the amygdala. Although this characteristic may be useful in solving some problems, such as predicting space weather when measuring storm intensity is important, this feature may help determine the level of the solar storm by reading the peaks and valleys.

4. Hippocampus

The hippocampus plays a key role in forming new memories. Specifically, the hippocampus helps to adapt short-term memory (STM) into long-term memory (LTM). This is important to the present research because, whether STM or LTM, memories can evoke emotions as well. On the other hand, if the hippocampus is damaged, there is difficulty forming new memories and whatever is experienced fades away. Thus, if the hippocampus were destroyed, new memories cannot form, but the brain still has the old memories intact. In other

words, the short-term function is still fine, which is why the amygdala has a monotonic memory.

5. Hypothalamus

   "Hypo" means "below," so the hypothalamus is positioned below the thalamus. The hypothalamus is actually a very tiny structure that makes up less than 1% of the total volume of the brain. However, despite its small size, the hypothalamus plays a significant role in regulating numerous functions in the brain as well as the human body in general. In terms of emotions and limbic system structure, the hypothalamus regulates the Autonomic Nervous System (ANS) [50], which controls the "fight or flight" versus "rest and digest" impulses. The hypothalamus does this by controlling the endocrine system and triggering the release of hormones into the bloodstream. Some of the hormones that are triggered release epinephrine, commonly known as adrenaline. When adrenaline is pumping through the body's veins, it is actually being regulated by the hypothalamus. The hypothalamus is also involved in regulating other basic drivers, like hunger, thirst, sleep, and sexual desire.

   The researchers in [51] investigated the relationship between the orbitofrontal cortex (OFC) and the hypothalamus. Their results found that the OFC has heterogeneous activities in the region itself and has a homogeneously dense connection to the hypothalamus in the two major parts of the lateral and medial, in terms of controlling autonomic functions. By using the resting-state function connectively, the researchers in [51] were able to divide the cerebral cortex into fictitious functional areas. The functional connectivity was examined between these areas in the OFC and lateral/medial hypothalamus. Specifically, [51] noticed that the functional double dissociation in the OFC, in that the lateral OFC was more connected to the lateral hypothalamus and the medial OFC was more connected to the medial hypothalamus. As well, [51] demonstrated a fundamental heterogeneous of the OFC region and recommended a potential neural basis of OFC-hypothalamic functional interaction.

6. OrbitoFrontal Cortex (OFC)

   The orbitofrontal cortex OFC is located near the amygdala and is a key part of

the limbic system. It has a mono-direction connection with the sensory cortex and a bi-directional connection with the amygdala. Its functions are processing stimuli, analyzing emotional stimuli, and evaluating reinforcement signals and emotional learning to prevent an inappropriate response from the amygdala.

7. Cerebral Cortex & Prefrontal Cortex

   By observing the structure of the brain, we can see it has front and back parts. The area of the brain known as the cerebral cortex plays a role in emotion. However, there are different ways to divide the cerebral cortex and organize it. In relation to the emotion concept, it is best to view the brain in terms of hemispheres. Based on this perspective, the brain has two hemispheres: one on the right side of the brain, and the other on the left side.

   Research indicates that the hemispheres are the sites of different functions. For instance, a positive emotion evokes more electrical activity on the left side of the brain than on the right side, whereas negative emotions tend to elicit more activity in the right hemisphere. These results are based on an experiment conducted with a number of participants who watched movies divided into two categories: those that evoked pleasure and those that evoked disgust. The researchers videotaped the facial expressions of the participants as well as recording their brain activities using an electroencephalogram (EEG) recording. The EEG measures the electrical activity of the brain. The research study found that the pleasurable movies increased activity in the left hemisphere of the participants' brains because these movies were associated with positive emotions like happiness, thankfulness, joy, and enthusiasm. On the other hand, the scary (negative) movies increased activity in the right hemisphere of the participants' brains because scary movies are associated with negative feelings like fear, isolation, timidity, avoidance, and even depression. Note that activity was found on both sides of the participants' brains during both movies, but that the positive emotions involved more activity on the left side and the negative emotions showed more activity on the right. The researchers also noticed that the left hemisphere was more active in certain situations, such as social interaction, whereas the right hemisphere was more active in situations that were defined

as isolated or lonely.

Another way to look at the cerebral cortex is by dividing it into functional divisions of three parts: front, middle, and back. In this case, the most important part, in terms of emotions, is the front section, known as the prefrontal cortex. The prefrontal cortex is the area immediately behind the forehead. This part of the brain is responsible for many high-order functions, including language, information processing, and decision-making.

8. Autonomic Nervous System (ANS)

   The best example for understanding the ANS is roller coaster-riding, which typically combines the emotions of fear and excitement. Accompanying the feelings of fear and excitement are an increase in heart rate and faster breathing. These physiological changes are not under your conscious control, i.e., you do not command yourself to start breathing faster, it just happens automatically. Physiological changes occur automatically through connections in the nervous system. The branch of the nervous system responsible for these automatic reactions is the Autonomic Nervous System (ANS). The ANS has two branches: the sympathetic nervous system and the parasympathetic nervous system, each of which causes different changes. Some changes may also occur in different organs and parts of our body.

   A number of researchers refer to the sympathetic nervous system as performing actions involving "fight or flight," while others refer to the parasympathetic nervous system as "rest and digest." Therefore, the sympathetic nervous system causes changes in our body consistent with the feelings and changes that one might experience when afraid of something (like riding a roller coaster). The fear response engenders multiple changes in our body automatically through the involvement of the sympathetic nervous system. The most important organs related to the ANS in terms of emotional response are the eyes, salivary glands, heart, lungs, gastrointestinal (GI) tract, liver, and kidneys. The latter organ is located next to the adrenal gland, which is responsible for releasing hormones like adrenalin.

Chapter 3

PROPOSED MODELS

3.1 Review of BEL Functions and Learning Algorithms

Brain emotional learning (BEL) networks are computational models that mimic the method employed by the mammalian brain in processing emotional stimuli, as described by LeDoux [52],[53]. According to LeDoux, emotional triggers are processed faster than non-emotional stimuli due to the existence of shorter paths in the part of the brain known as the emotional brain, which has a psychological aspect as well [54]. The emotional brain components responsible for processing the emotional stimuli include the thalamus, the sensory cortex, the amygdala, the orbitofrontal cortex, and the hippocampus. The process is initiated after the emotional stimulation is received by the thalamus, which submits imprecise information regarding the amygdala's stimulus. The stimulus is also propagated to the amygdala through the sensory cortex, which forms a longer path than the thalamus and amygdala. The sensory cortex passes the stimulus information onto the orbitofrontal cortex, prompting an emotional response in the amygdala. The amygdala also submits this emotional response to the orbitofrontal cortex, which evaluates the response and rectifies it with the help of the hippocampus. This process is depicted in Figure 3.1.

The above emotional process in the mammalian brain forms the basis for the brain's computational models. The first such model appeared in [33]. In the model, the thalamus's imprecise information to the amygdala represents the maximum value of the stimulus. The rectifying operation of the orbitofrontal cortex is achieved using the suppression operator, i.e., the BEL model's output is the emotional response of the amygdala minus the production from the orbitofrontal cortex. A reinforcement signal R_o is also computed for the model to learn, which is given as [4]:

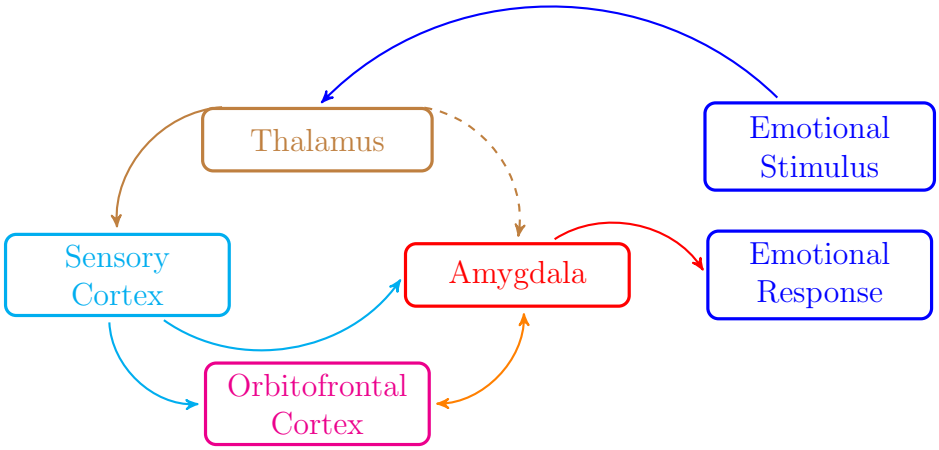

Figure 3.1: Routes of Limbic System.

$$R_o = \begin{cases} (\sum a_i - rew)^+ - \sum(o_i) \ , & \text{if } (rew \neq 0) \\ (\sum a_i - \sum o_i)^+ \end{cases} \tag{3.1}$$

otherwise, $\sum a_i$ represents the amygdala output, while $\sum o_i$ is the orbitofrontal cortex output. This reinforcement signal is used to adjust the weights of the BEL model to improve its response. Thus, through a history of input rewards and punishment signals, the model is made to learn the desired response to the emotional stimuli. However, it is not clear how the value is assigned to signal in the learning process. In other BEL models [38], [8] this signal is computed as the weighted combination of a set of reinforcement factors related to a process [4] :

$$rwe = \sum_j w_j r_j \tag{3.2}$$

The special way of computing the signal, as in Eq (3.2), makes the model less efficient in learning opposite behaviours, but it works well for a specific problem. The BEL models with (rew) given by Eq. (3.2) have shown great success in various real-time applications, including home appliances [11], robotics [55],[56], electrical drives [57],[34] and other industrial systems [9],[36]. These models have also been used for time series prediction problems [37],[38] where their performance in predicting the

peak points is excellent. However, their performance is poor at valley points in the time series data. To improve the performance of these BEL models, supervised BEL models are proposed in [24],[4]. These models employ the pattern-target samples in a supervised way for their learning and assign target values to signal during the learning process [4]:

$$rwe = t \tag{3.3}$$

The benefit of using target-pattern samples is that the model can be adjusted to follow peak or valley points in the time series data. The shortcoming is that the model can only yield good results for recent inputs, and the performance is degraded in instances of distant examples. A decay rate is added to the BEL model to overcome this weakness, which also has a neurobiological basis [39]. The resulting BEL model is called "ADBEL" and has shown good performance for online time series forecasting, as reported in [24],[4].

3.2 Review of Adaptive Decayed Brain Emotional Learning (ADBEL) Network

The ADBEL network was proposed in [24],[4] for time series prediction in a supervised way. It differs from other BEL models in that it can be used online, and no prior training is required before using the network. The ADBEL network trains itself in online mode, using an internal reward signal. As opposed to other BEL models, the ADBEL network utilizes a decay rate that has a neurobiological basis and serves to improve the network's performance. It is well-known that other artificial intelligence techniques, such as ANNs, suffer from the bias dilemma problem, whereby a model cannot track changes in a non-stationary environment. To do so, the parameters and structure of the ANNs model must be time-varying. In other words, the ratio between stability and plasticity must be coordinated by the forgetting factor, ignoring any knowledge that becomes invalid due to noise or non-stationary media. A decay rate that solved this problem was added to the BEL model, creating the ADBEL model. A schematic representation of the network is given in Figure 3.2.

The ADBEL network has four inputs $(p_{t-4}, p_{t-3}, p_{t-2}, p_{t-1})$ and one output $(\hat{p_t})$, as seen in Figure 3.2.

The ADBEL network has four inputs $(p_{t-4}, p_{t-3}, p_{t-2}, p_{t-1})$ and one output $(\hat{p_t})$, as seen in Figure 3.2.

In referring to [4], the number of ADBEL network's inputs, in general, could be n-inputs where $j = 1, 2, ...n,$. Therefore, Eq. 3.4, in general, can be formed as: $\hat{p_t} = f(p_{t-n}, ..., p_{t-2}, p_{t-1})$. However, Lotfi in [4] considered only four inputs following each four sequence samples as a pattern, using the fifth as its target. Therefore, in the present work, we assessed the same arrangement for a fair comparison of our designed model to the existing model in [4].

The four inputs are the time series values at the previous four-time instants, while the output is the predicted value of the time series at the current time instant. This mapping is given as:

$$\hat{p_t} = f(p_{t-4}, p_{t-3}, p_{t-2}, p_{t-1}) \tag{3.4}$$

The ADBEL network's functioning is such that after the inputs are presented to the network, a maximum value is computed by the thalamus (m), which then submits it to the amygdala. The inputs are then transferred to the sensory cortex, which introduces these inputs to the amygdala and orbitofrontal cortex through the weights v_j's and w_j's, respectively. After receiving the five inputs, the amygdala produces two outputs, E_a and $\acute{E_a}$, as:

$$E_a = \acute{E_a} + v_{th} \times m \tag{3.5}$$

$$\acute{E_a} = \sum_{j=1}^{4} (v_j \times p_{t-j}) \tag{3.6}$$

$$m = max_j \times (p_{t-j}) \tag{3.7}$$

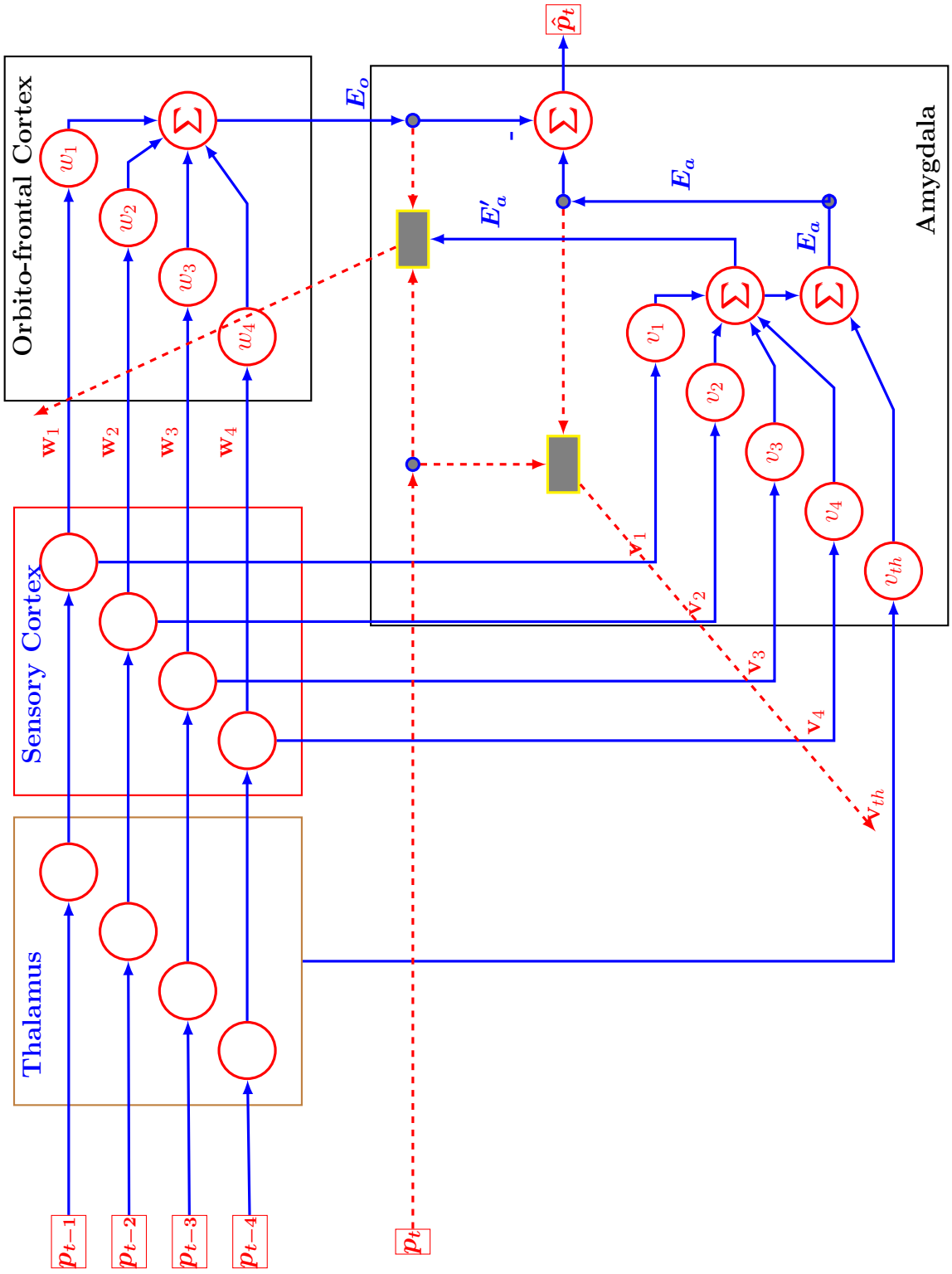

Figure 3.2: Schematic Drawing of ADBEL Network.

$$\forall j = 1, 2, ..., 4$$

Similarly, the orbitofrontal cortex produces an output, E_o as:

$$E_o = \sum_{j=1}^{4} (w_j \times p_{t-j}) \tag{3.8}$$

The final output from the ADBEL network, i.e., the predicted value of the time series at the current timestamp, is found to be:

$$\hat{p_t} = E_a - E_o \tag{3.9}$$

The steps shown in Eqs. (3.5 - 3.9) are termed as prediction steps in [4]. After the prediction stage, the ADBEL network is trained with the help of signals $E_a, \acute{E_a}, E_o, p_t$ and constant parameters α, β, γ . The output from the amygdala E_a, in conjunction with the current time series value p_t and decay rate γ, is used to adjust the amygdala weights in the following way:

$$\begin{cases} v_j(t+1) = (1-\gamma)v_j(t) + \alpha \ max(p_t - E_a, 0)p_{t-j} \\ v_{th}(t+1) = (1-\gamma)v_{th}(t) + \alpha \ max(p_t - E_a, 0)m \quad , \quad \forall j = 1, 2, ..., 4 \end{cases} \tag{3.10}$$

In adjusting the weights of the orbitofrontal cortex, an internal reward signal R_o is first computed as:

$$R_o = \begin{cases} max(\acute{E_a} - p_t, 0) - E_o \ , & \text{if } (p_t \neq 0) \\ max(\acute{E_a} - E_o, 0) \ , & \text{otherwise} \end{cases} \tag{3.11}$$

Based on this reward signal, the weights of the orbitofrontal cortex are updated as:

$$w_j(t+1) = w_j(t) + (\beta \ R_o \ p_{t-j}) \tag{3.12}$$

$$\forall j = 1, 2, ..., 4$$

As can be observed from the prediction and learning stages of the ADBEL network,

the ABDEL approach presents a simple method to time series prediction in an online mode. Unlike the other popular networks used for forecasting, such as ANN with a back-error propagation algorithm or ANFIS, neither can predict time series with shorter update intervals due to their computational complexity. It is also worth noting that three constant positive parameters (α, β, γ) are used in the network's learning stage and need to be tuned for the network's best performance. The range of these parameters is reported in [4] as:

$$\begin{cases} \alpha \leq 1, \\ \beta \leq 1, \\ 0 \leq \gamma \leq 0.2 \end{cases} \tag{3.13}$$

The parameters alpha, beta and gamma were assigned their values as in ref [4] as initial values. They were then selected after extensive experimentation to yield the best possible prediction performance in each case. All the network weights were initialized as zeros instead of being assigned random values. This aids in helping run the simulations only once; furthermore, in so doing, no averaging of the results is required, as the networks will yield the same performance every time. In section 3.7, we present a proposed method for the automatic tuning of these parameters.

3.3 The Proposed Modified Models of Adaptive Decayed Brain Emotional Learning (ADBEL) Network

This work considers three significant modifications of the Adaptive Decayed Brain Emotional Learning (ADBEL) network.

First of all, the ADBEL network was built in MATLAB programming using the same learning parameters as in [4]. Then three significant modifications to the ADBEL network were done in this thesis as follows:

The first modification is an integration of a neo-fuzzy network with the ADBEL network to yield a novel NF-ADBEL network with improved forecasting performance, as demonstrated in [58].

Both the ADBEL and the neo-fuzzy networks [59],[60] share essential features of

simplicity, accuracy, and less computational complexity, all of which are desirable for online forecasting problems. Thus, it is natural to investigate a hybrid-forecasting model based on these two networks. The proposed NF-ADBEL network [58] simulated in a MATLAB (R2014a) programming environment to forecast several chaotic time series in an online mode, including Mackey-Glass, Lorenz, Rossler, Narendra and the disturbance storm time index, as well as stochastic non-linear systems such as wind speed and wind power series. Comparing the prediction performance of both networks in terms of root mean square error and correlation coefficient criteria reveals the superiority of the proposed NF-ADBEL network in online forecasting problems.

Note that comparisons of the NF-ADBEL network with popular networks such as ANFIS and LLNF will not be performed in this work, as the ADBEL network has already shown better performance than these networks [24][4] concerning online time series forecasting problems. Also, both the ADBEL and the proposed NF-ADBEL networks do not require prior training to perform predictions. In contrast, the other networks mentioned above need to be trained before they can be deployed for prediction purposes.

The second modification is an Expanded integration of a neo-fuzzy network partially in the amygdala section. The NF-ADBEL network yields a novel ENF-ADBEL network with enhanced forecasting performance, as proposed in [61].

The ENF-ADBEL network [61] simulated in a MATLAB (R2014a) programming environment can forecast several chaotic time series in an online mode, including Mackey-Glass, Lorenz, Rossler, Narendra and the disturbance storm time index, as well as stochastic non-linear systems such as wind speed and wind power series. A comparison of the prediction performance of both networks (ENF-ADBEL and NF-ADBEL) in terms of root mean square error and correlation coefficient criteria reveals the superiority of the proposed ENF-ADBEL network in online forecasting problems.

The third modification of ABDEL in this work is the proposed Fuzzy Logic-Based Parameter Adjustment Model for the Adaptive Decayed Brain Emotional Learning network demonstrated in [62], named herein as the F-ADBEL model.

ADBEL is a computationally fast neural network that has shown promising results for online time series prediction problems. However, it is not clear how to set the network's free parameters, namely alpha, beta and gamma. In this study, we modelled

a fuzzy logic-based model for adjusting these ADBEL network parameters in an online mode, which will not jeopardize the network's simplicity or quick response. The proposed model uses prediction error and reward signals as inputs and produces network parameters as outputs. Furthermore, the fuzzy logic-based model employs a small rule base to vary the network parameters to maintain the ADBEL network's complexity at a minimum. A few chaotic time series (namely, the Mackey Glass, Lorenz, Rossler, and the disturbance storm time index) are used to verify the validity of the fuzzy logic-based parameter adjustment model for the ADBEL network using a MATLAB programming environment.

3.4 Review of Neo-Fuzzy Network

A neo-fuzzy network is a multi-input, single-output network which employs nonlinear synapses to generate the mapping between input and output data. An n-input neo-fuzzy network is depicted in Figure 3.3. The model functions such that each input presented to the network is fuzzified using (k) triangular membership functions. Each degree of belongingness thus computed is further weighted, and all such weighted degrees are summed to generate the output. Mathematically, the process is given by:

$$y_{nf} = \sum_{i=1}^{n} f_i(x_i) \tag{3.14}$$

where $f_i(x_i)$ is the response of the i^{th} neo-fuzzy neuron in the final network output and is given as:

$$f_i(x_i) = \sum_{j=1}^{k} h_{ij}(x_i) w_{ij} \tag{3.15}$$

where $h_{ij}(x_i)$ is the degree of membership of i^{th} input over j^{th} membership function and w_{ij} is the corresponding weight which needs to be determined for mapping input-output data. Please note that the membership functions are fixed, complementary, and equally spaced to cover the discourse universe. The overlapping parts of

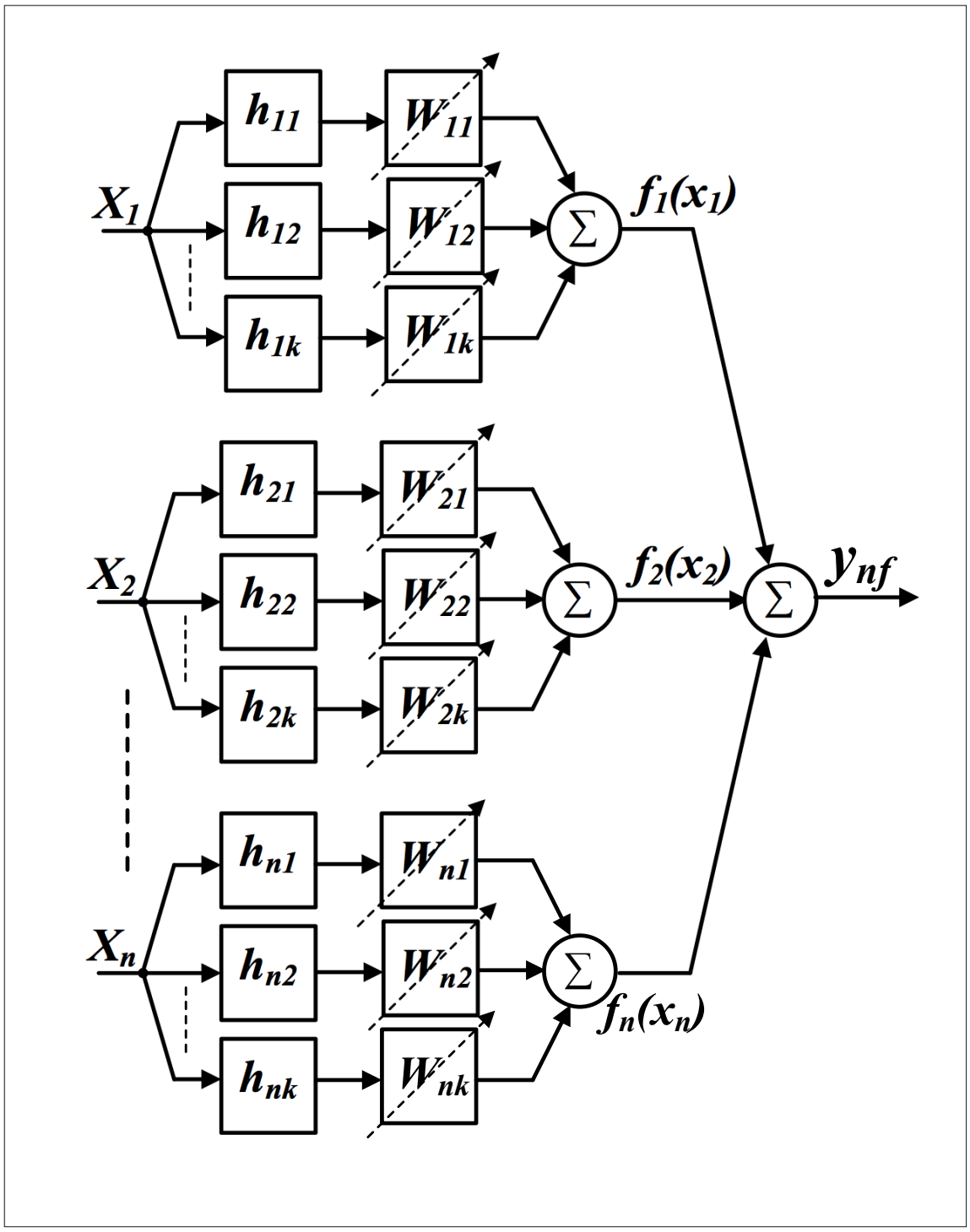

Figure 3.3: Neo-Fuzzy Network.

neighbouring membership functions, therefore, can be described as:

$$\begin{cases} h_{ij+1}(x_i) = \frac{x_i - c_j}{c_{j+1} - c_j} \quad , & x_i \in [c_j \ , c_{j+1}] \\ h_{ij} = 1 - h_{ij+1} \quad , & \forall j = 1, 2, ..., k-1 \end{cases} \tag{3.16}$$

where c_j is the center of the j^{th} membership function and could be computed using the knowledge of uniform spacing d between the membership functions as:

$$\begin{cases} c_{j+1} = c_j + d \quad , & \forall j = 1, 2, ..., k-1 \\ c_1 = x_{min} \quad , c_k = x_{max} \quad , & \\ d = \frac{c_k - c_1}{k-1} \quad , & \end{cases} \tag{3.17}$$

It is worth noting that the unknown parameters in the neo-fuzzy network are the membership degrees' weights. In contrast, the membership functions are fixed for all the neurons defining the network compared to a classical neuro-fuzzy system, where membership functions are adjusted during the learning process. Computing the adjustable weights by the error function is defined as:

$$E = \frac{1}{2}(y_{nf} - y_d)^2 \tag{3.18}$$

The minimization of this quadratic error function through the gradient descendent method yields the following parameter adjustment rule:

$$w_{ij}(t+1) = w_{ij}(t) + \beta(y_{nf}(t) - y_d(t))h_{ij}(x_i) \tag{3.19}$$

where β is a positive constant and is defined as the neo-fuzzy network's learning rate, the proposed model mimics emotional learning by integrating a neo-fuzzy neuron. The learning algorithm of the proposed model is based on gradient descent (GD). The proposed model aims to enhance the prediction accuracy in existing computational models that use brain emotional learning processing.

3.5 Neo-Fuzzy Integrated ADBEL Network

Inspired by the standard features offered by ADBEL and neo-fuzzy networks, this work considers a hybrid model called NF-ADBEL to improve the forecasting accuracy of the ADBEL network. Although neo-fuzzy neurons can replace all the neurons in various sections of the ADBEL network, we have only considered replacing the neurons in the orbitofrontal cortex section of the ADBEL network with neo-fuzzy neurons. Therefore, the hybrid network is still used in online mode for time series prediction. The resulting network is shown in Figure 3.4.

The functioning of the NF-ADBEL network is similar to that of the ADBEL network in a broad sense. For example, the output response of the NF-ADBEL network represents the difference between the amygdala and orbitofrontal cortex outputs after the input stimuli are fed to these sections through the thalamus and sensory cortex. The difference lies in the construction of the orbitofrontal cortex section and its learning, as can be observed by comparing Figures 3.2 and 3.4 (the dashed lines in these figures represent the learning of the network). The neo-fuzzy neuron for the orbitofrontal cortex is realized with three triangular membership functions. The universe of discourse is selected to be $[0, 1]$ for all inputs, which is, in fact, the normalized limit for the time series data points. Thus, the output of the proposed integrated NF-ADBEL network is given as:

$$\hat{P}_t = \sum_{j=1}^{4} \sum_{k=1}^{3} (v_j \times P_{t-j} \; - \; w_{jk} \times h_{jk}) + v_{th} \times m \tag{3.20}$$

The unknown weights of the amygdala and orbitofrontal cortex in Eq. (3.20) are adjusted in an online manner, using the laws in Eqs. 3.10 and 3.19, respectively. Please note that the proposed NF-ADBEL network does not have any knowledge about the time series, as is the case with the ADBEL network. Previous works on neo-fuzzy networks consider training the network with the time series data and then deploying the trained network to do future predictions [63],[64]. However, in this work, no prior training of neo-fuzzy network is assumed

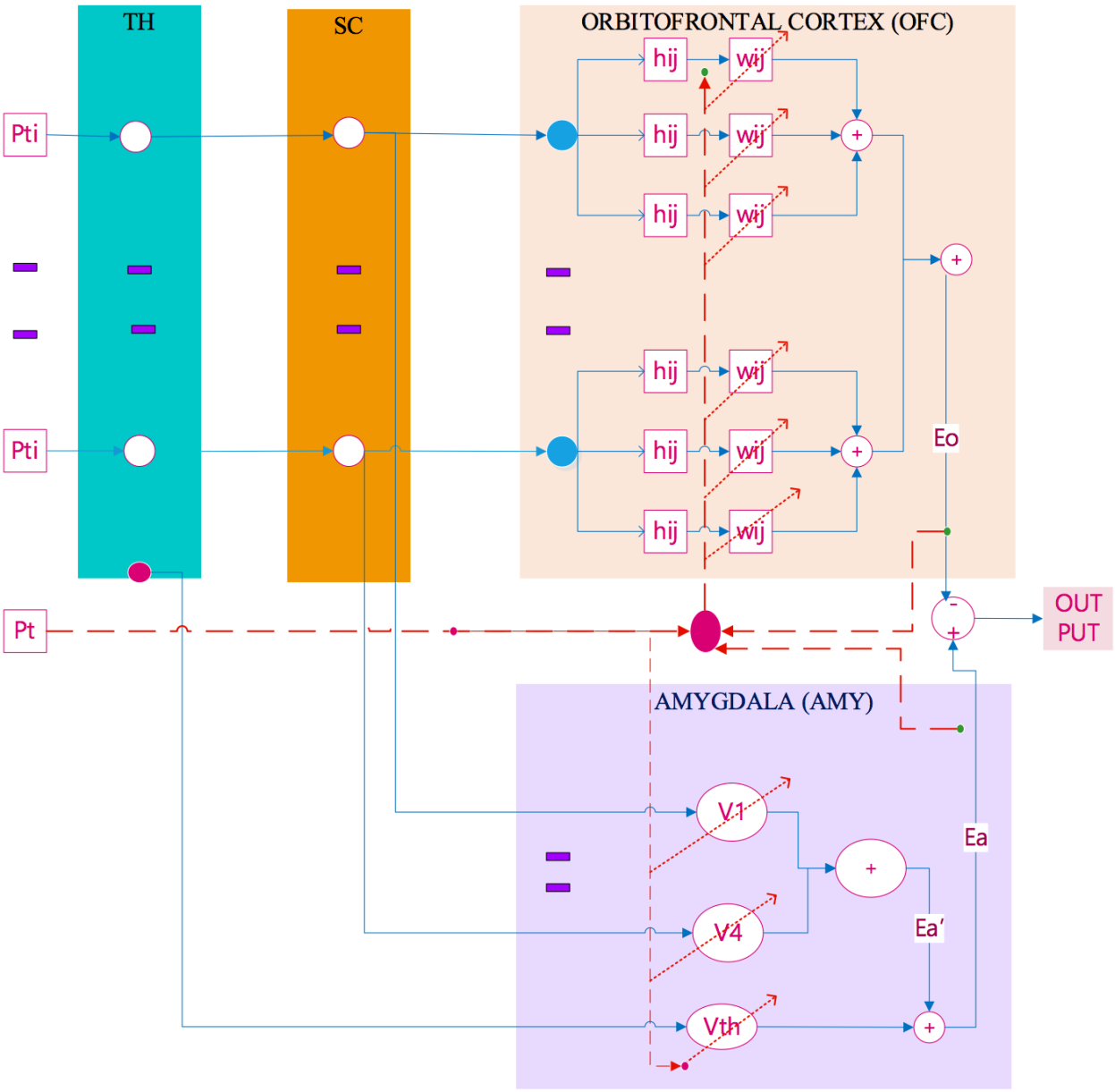

Figure 3.4: Proposed Neo-Fuzzy Integrated ADBEL Network.

3.6 Expanded Neo-Fuzzy Integrated ADBEL Network

A neo-fuzzy adaptive decayed brain emotional learning (NF-ADBEL) network is proposed for the online time series predicting problems, as demonstrated in [58]. The NF-ADBEL network offers required features for online prediction for shorter update intervals, such as fast learning, accuracy, simplicity, and lower computational complexity. The neo-fuzzy neuron in NF-ADBEL was integrated only in the orbitofrontal cortex (OFC) portion of the ADBEL network. In this work, we propose a new modification network that aims to investigate and further enhance the performance of the NF-ADBEL network by the partial integration of a neo-fuzzy neuron network into the amygdala section (AMY).

As depicted in Figures 3.2 and 3.4, the amygdala has two outputs. One of its responses is based on imprecise information received from the thalamus, while the other output process is treated by neo-fuzzy implementation. The result is the Expanded Nf-amygdala partial integration (ENF-ADBEL), as shown in Figure 3.5. The modified network is still simple and meets the required features for online prediction problems. A few of the chaotic and stochastic non-linear systems, namely the Mackey Glass, Lorenz, Rossler, disturbance storm time index, Narendra dynamic plant identification problem and wind speed and wind power series, are used to validate and evaluate the performance of the proposed network in terms of the root mean square error and correlation coefficient criterions using a MATLAB programming environment. ENF-ADBEL outcomes reveal the superior performance, fast learning, quick response, and ability to deploy the proposed model to predict time series applications in an online mode. The proposed network promises adequate performance in terms of stochastic problems. Generally speaking, the functioning of the ENF-ADBEL network is similar to that of the NF-ADBEL network.

As in [58], the ADBEL network is structured as four inputs, as given by (3.21) and one output as given by (3.22). This mapping is given as:

$$P = \begin{pmatrix} p_1 & p_2 & \cdots & p_j \end{pmatrix}^T \tag{3.21}$$

$$\forall j = 1, 2, ..., 4$$

where j^{th} indicates network inputs.

$$\hat{P}_t = f(P) \tag{3.22}$$

After the inputs are presented to the network through the TH, it computes the maximum values for the given stimulus's input data and sends them along a short path to AMY. At the same time, TH dispatches the presented data to SC, which sends the information to the AMY and OFC sections along the associated weights (V's and W) belonging to each one, respectively.

$$V = \begin{pmatrix} v_1 & v_2 & \cdots & v_j \end{pmatrix} \tag{3.23}$$

$$W = \begin{pmatrix} w_1 & w_2 & \cdots & w_j \end{pmatrix} \tag{3.24}$$

$$\forall j = 1, 2, ..., 4$$

AMY produces two outputs, E_{AMY} and $\acute{E}_a$, as:

$$\acute{E}_a = V \times P \tag{3.25}$$

$$m = max_j \times P \tag{3.26}$$

$$E_{AMY} = \acute{E}_a + V_{th} \times m \tag{3.27}$$

Similarly, OFC produces one output, E_{OFC}, as:

$$E_{OFC} = W \times P \tag{3.28}$$

The output of ADBEL is given by:

$$\hat{P}_t = E_{AMY} - E_{OFC} \tag{3.29}$$

The steps shown in ((3.27)-(3.29)) are termed as prediction steps in [58]. After the prediction stage, the ADBEL network is trained with the help of signals $E_{AMY}, \acute{E}_a, E_{OFC}, p_t$ and constant parameters α, β, γ . The output from the AMY E_a in combination with the current time series value $p_t = target\ (T_t)$ and decay rate γ, is used to adjust the amygdala weights in the following way:

$$\begin{cases} V(t+1) = (1-\gamma) \times V(t) + \alpha \times max(T_t - E_{AMY}, 0) \times P^T \\ V_{th}(t+1) = (1-\gamma) \times V_{th}(t) + \alpha \times max(T_t - E_{AMY}, 0) \times m, \end{cases} \tag{3.30}$$

To adjust the weights of the OFC, an internal reward signal R_o is first computed as:

$$R_o = \begin{cases} max(\acute{E}_a - T_t, 0) - E_{OFC}\ , & \text{if } (T_t \neq 0) \\ max(\acute{E}_a - E_{OFC}, 0)\ , & \text{otherwise} \end{cases} \tag{3.31}$$

Based on this reward signal, the weights of the OFC are updated as:

$$W(t+1) = W(t) + \beta \times R_o \times P^T \tag{3.32}$$

To study the ENF-ADBEL network's performance, the working principle of the proposed network remains the same as that of the ADBEL network. The difference lies in the definition of weight entries and the application of those entries in the AMY and OFC sections' learning rules. OFC outputs in the ENF-ADBEL network are computed using (3.24) with a different set of neo-fuzzy weights w_{ij} and corresponding degrees of membership functions h_{ij} as:

$$E_{OFC,enf} = W_{ij} \times H_{ij} \tag{3.33}$$

$$W_{ij} = \begin{pmatrix} w_{11} & w_{12} & w_{13} & w_{21} & w_{22} & w_{23} \cdots & w_{ij} \end{pmatrix} \tag{3.34}$$

$$H_{ij} = \begin{pmatrix} h_{11} & h_{12} & h_{13} & h_{21} & h_{22} & h_{23} \cdots & h_{ij} \end{pmatrix}^T \tag{3.35}$$

Similarly, AMY outputs in the ENF-ADBEL network are computed using (3.23) with a different set of neo-fuzzy weights v_{ij} and corresponding degrees of membership functions h_{ij} as:

$$\acute{E}_a = V_{ij} \times H_{ij} \tag{3.36}$$

$$E_{AMY,enf} = \acute{E}_a + V_{th} \times m \tag{3.37}$$

$$V_{ij} = \begin{pmatrix} v_{11} & v_{12} & v_{13} & v_{21} & v_{22} & v_{23} \cdots & v_{ij} \end{pmatrix} \tag{3.38}$$

leads to :

$$\hat{P_{enf}} = E_{AMY,enf} - E_{OFC,enf} \tag{3.39}$$

The minimization of the quadratic error function in [58] through the gradient descendent method yields the new following parameter adjustment rules:

$$W_{ij}(t+1) = W_{ij}(t) + \beta(\hat{P_{enf}}(t) - T(t))H_{ij}^T(x_i) \tag{3.40}$$

$$V_{ij}(t+1) = (1-\gamma)V_{ij}(t) + \alpha\ max(T(t) - E_{AMY,enf}(t), 0)H_{ij}^T(x_i) \tag{3.41}$$

$$V_{th}(t+1) = (1-\gamma)V_{th}(t) + \alpha\ max(T(t) - E_{AMY,enf}(t), 0)m \tag{3.42}$$

where α and β and γ are positive constants and are defined as the learning rates of the neo-fuzzy network, the proposed model mimics the emotional learning by integrating a neo-fuzzy neuron. The gradient descent (GD) method is employed for the learning algorithm of the proposed model. The proposed model aims to enhance the prediction accuracy in existing computational models that use brain emotional learning processing.

Broadly speaking, the ENF-ADBEL network's functioning is similar to that of the NF-ADBEL and ADBEL networks. The neo-fuzzy neurons for the ENF-ADBEL network are realized with three triangular membership functions, and the universe of discourse is selected to be $[0, 1]$, as in [58]. Thus, the output of the proposed integrated ENF-ADBEL network be given as:

$$\hat{P}_t = \sum_{j=1}^{4} \sum_{k=1}^{3} (V_{jk} \times H_{jk} \ - \ W_{jk} \times H_{jk}) + V_{th} \times m \tag{3.43}$$

The unknown weights of the amygdala and orbitofrontal cortex in Eq. (3.43) are adjusted in an online manner: using the laws in Eqs. 3.40, 3.41 and 3.42, respectively. Please note that the proposed ENF-ADBEL network does not have any knowledge about the time series, as is the case with the NF-ADBEL and ADBEL networks. Previous works on neo-fuzzy networks and state-of-the-art consider training the network with the time series data and then deploying the trained network to do future predictions [65] [66]. However, in this work, no prior training of neo-fuzzy network is assumed.

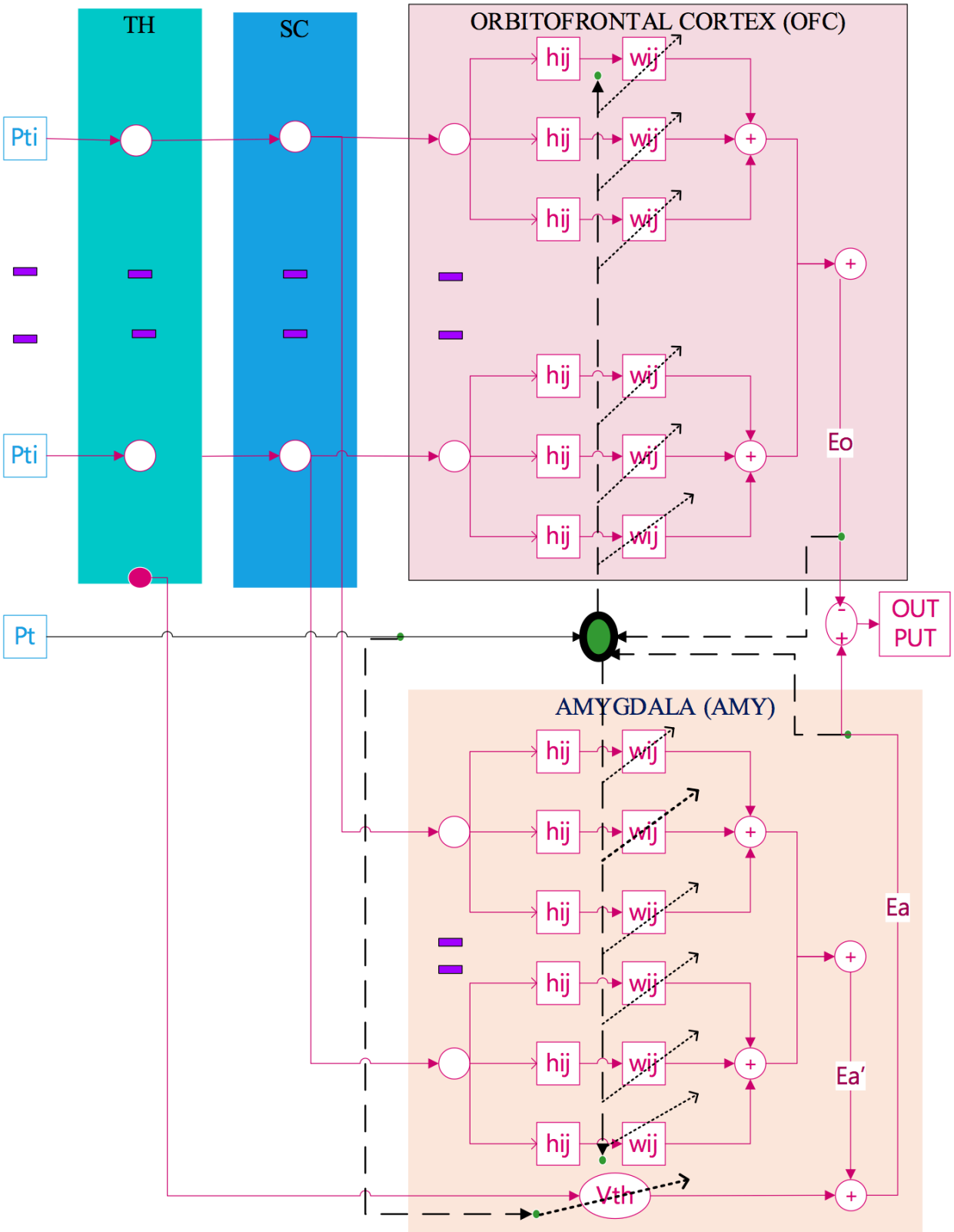

Figure 3.5: Proposed Expanded Neo-Fuzzy Integrated ADBEL Network.

3.7 Fuzzy Logic-based Parameter Adjuster Model

The target here is to build a model to adjust the free parameters of the ADBEL network. For this purpose, a simple fuzzy system is designed, as shown in Figure 3.6. The procedure takes as inputs the prediction error and reward signals and adjusts the ADBEL network parameters (namely, alpha, beta and gamma) in online mode. Since the ADBEL network is simple and produces the output quickly, the designed fuzzy system is also kept simple. The overall network remains straightforward, and the response of the network is not compromised. The resultant fuzzy integrated ADBEL network (F-ADBEL) is demonstrated in [62]. Figure 3.7 shows the proposed diagram. This network's validity is checked through MATLAB (R2014a) simulations for predicting chaotic time series, such as Mackey Glass, Lorenz, Rossler, the disturbance storm index, and its behaviour and initial conditions as shown in section 4.2.

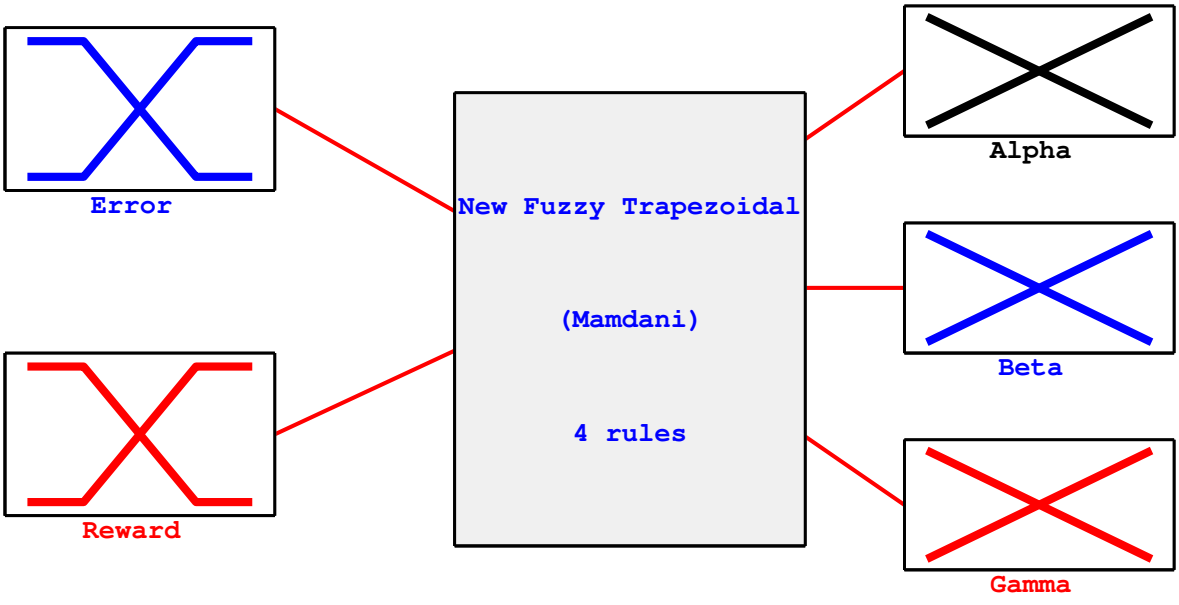

Figure 3.6: Block Diagram of Fuzzy Logic Designer.

Let us now discuss the basis for the design of the fuzzy parameter adjustment

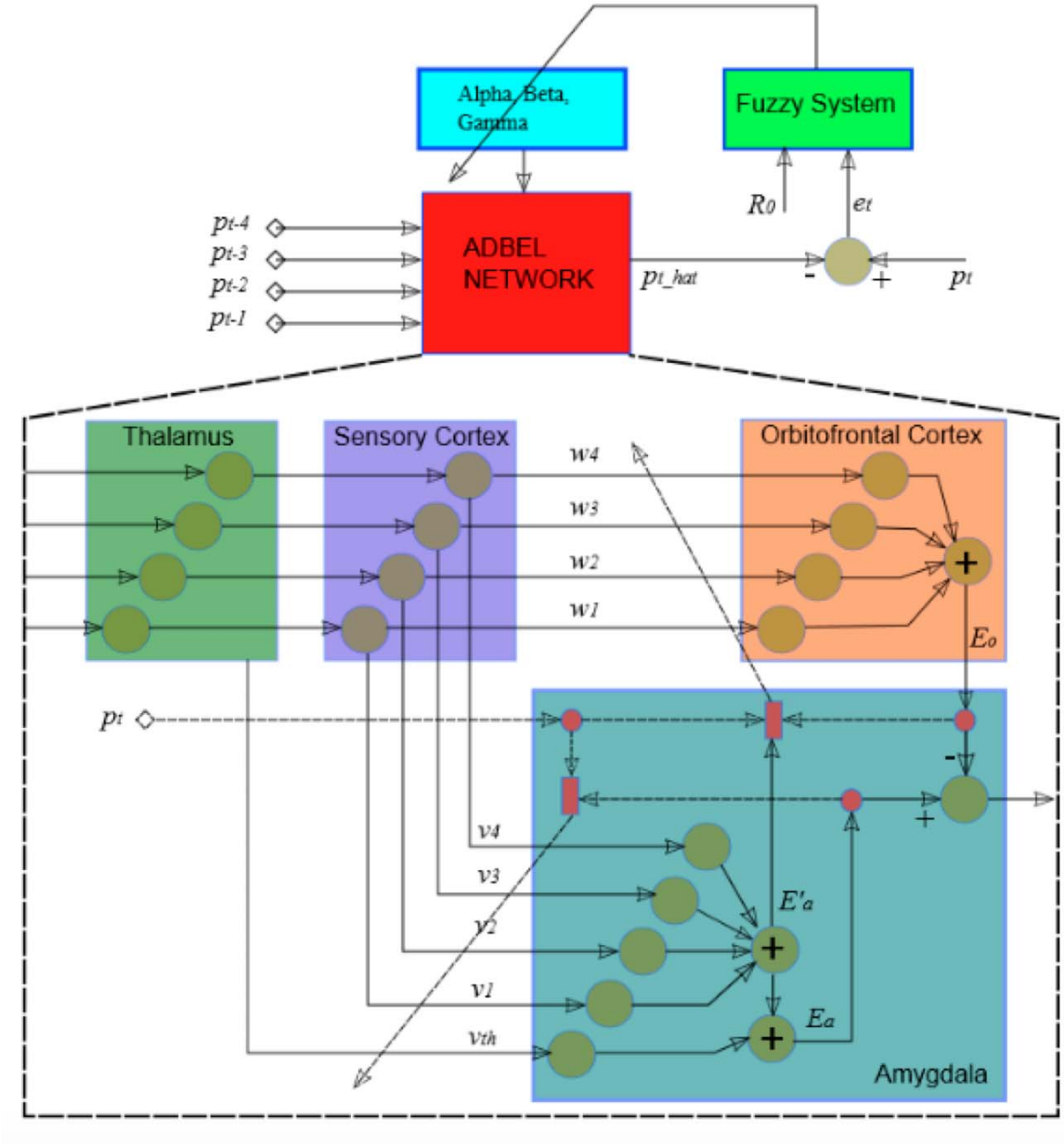

Figure 3.7: Proposed Fuzzy Integrated ADBEL Network.

model. In this section, we first define the prediction error as:

$$e_t = P_t - \hat{P}_t \tag{3.44}$$

where P_t is the time series value at the current timestamp and $\hat{P}_t$ is the predicted value as determined by Eq. (3.29). As discussed in [4], the time series values P_t for all indices are normalized in the range [0,1]. Further, the ADBEL network dynamics are such that they will not produce negative outputs when driven by the normalized time series values. Thus, both the input and predicted values will remain positive and will be lower-bounded by zero. However, the prediction error in Eq. (3.44) can be either positive or negative, depending on whether the predicted value $\hat{P}_t$ is less or greater than the desired value P_t, respectively.

Next, we discuss the case where the prediction error is positive. In this situation, the network output from Eq. (3.29) ($E_a - E_o$) has to be increased. This can be achieved by increasing E_a and/or decreasing E_o. Due to the special structure and learning of the ADBEL network, E_a will remain positive, whereas E_o can go negative. Thus, in order to increase E_a, parameter α will need to be increased, while parameter γ will need to be decreased, as can be inferred from Eqs. (3.27) and (3.10), respectively.

On the other hand, to decrease E_o, the reward signal R_o needs to be considered, as it can be either positive or negative. Given that the reward signal is positive, the parameter β needs to be decreased to lower the contribution of the reward signal in adjusting the orbitofrontal cortex weights, thereby reducing E_o, as can be seen from Eqs. (3.33) and (3.32), respectively. In this case, if the reward signal R_o is negative, the parameter β needs to be increased to lower E_o, which can again be followed from Eqs. (3.10) and (3.32). A similar analysis could be carried out when the prediction error is negative, which will require E_a to be reduced and E_o to be increased.

Based on this discussion, we select the prediction error e_t and reward signal R_o as the fuzzy model's two inputs. We also select the three ADBEL network parameters α, β, and γ as the fuzzy model's output. Since both the prediction error and reward signal can be either negative or positive, as previously discussed, we adopt two fuzzy sets, namely 'neg' and 'pos,' to describe their states, with the universe of discourse being the range [-0.1,0.1]. Further, to describe the states of the network parameters α, β, and γ as they need to be either decreased or increased, we select two fuzzy

sets, namely 'small' and 'large.' Here, the universe of discourse for the parameters α, and β is taken as [0,1]. In contrast, the parameter γ is set as the range [0, 0.05]. To describe these fuzzy sets, we define the following membership functions as shown in Figures 3.8, 3.9, 3.10, 3.11, and 3.12:

$$\mu_{\mathbf{neg}}(\mathbf{x}) = \begin{cases} 1\ , & x \leq -0.05\ , \\ -10x + 0.5\ , & -0.05 < x < 0.05 \\ 0\ , & x \leq 0.05 \end{cases} \tag{3.45}$$

$$\mu_{\mathbf{pos}}(\mathbf{x}) = \begin{cases} 0\ , & x \leq -0.05\ , \\ 10x + 0.5\ , & -0.05 < x < 0.05 \\ 1\ , & x \geq 0.05 \end{cases} \tag{3.46}$$

$$\mu_{\mathbf{small}}(\mathbf{y}) = -y + 1 \tag{3.47}$$

$$\mu_{\mathbf{large}}(\mathbf{y}) = y \tag{3.48}$$

$$\mu_{\mathbf{small}}(\mathbf{z}) = -20z + 1 \tag{3.49}$$

$$\mu_{\mathbf{large}}(\mathbf{z}) = 20z \tag{3.50}$$

where variable (x) represents the prediction error e_t or reward signal R_o, variable (y) represents the parameters α orβ, and variable(z) represents the parameter γ . Based on the aforementioned analysis, we now define the following rules using the

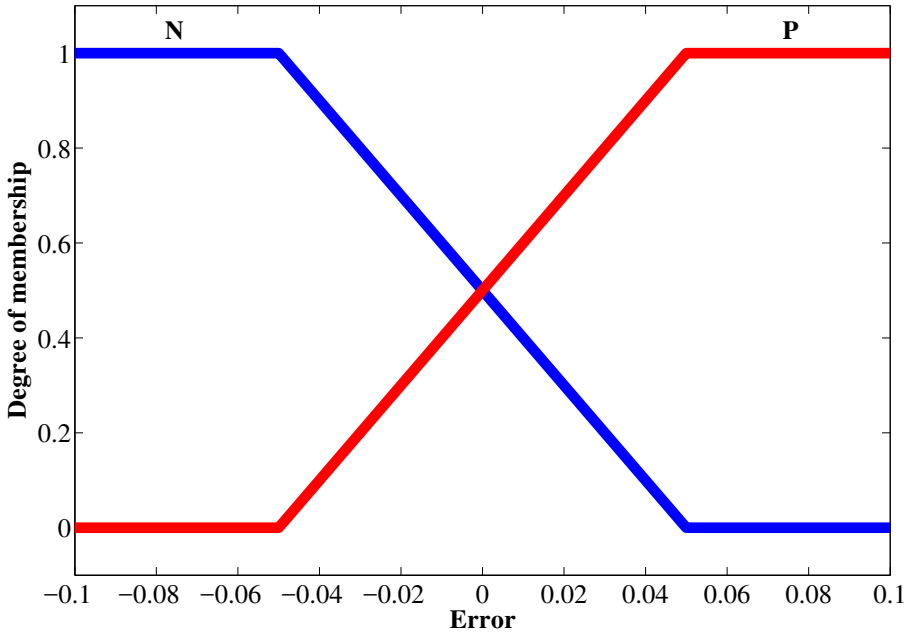

Figure 3.8: Membership Functions for Error Input Variable.

fuzzy sets:

1. R1: **IF** error is 'neg' and reward signal is 'neg' **THEN** α is 'small' and β is 'small' and γ is 'large'.

2. R2: **IF** error is 'neg' and reward signal is 'pos' **THEN** α is 'small' and β is 'large' and γ is 'large'.

3. R3: **IF** error is 'pos' and reward signal is 'neg' **THEN** α is 'large' and β is 'large' and γ is 'small'.

4. R4: **IF** error is 'pos' and reward signal is 'pos' **THEN** α is 'large' and β is 'small' and γ is 'small'.

After defining the fuzzy sets and the rule base, we employ the Mamdani fuzzy inference mechanism with (min) and (max) as the T-norm and T-conorm operators. Further, the centre of gravity method is used for defuzzification, which will produce the crisp values of network parameters from the corresponding aggregated fuzzy sets as:

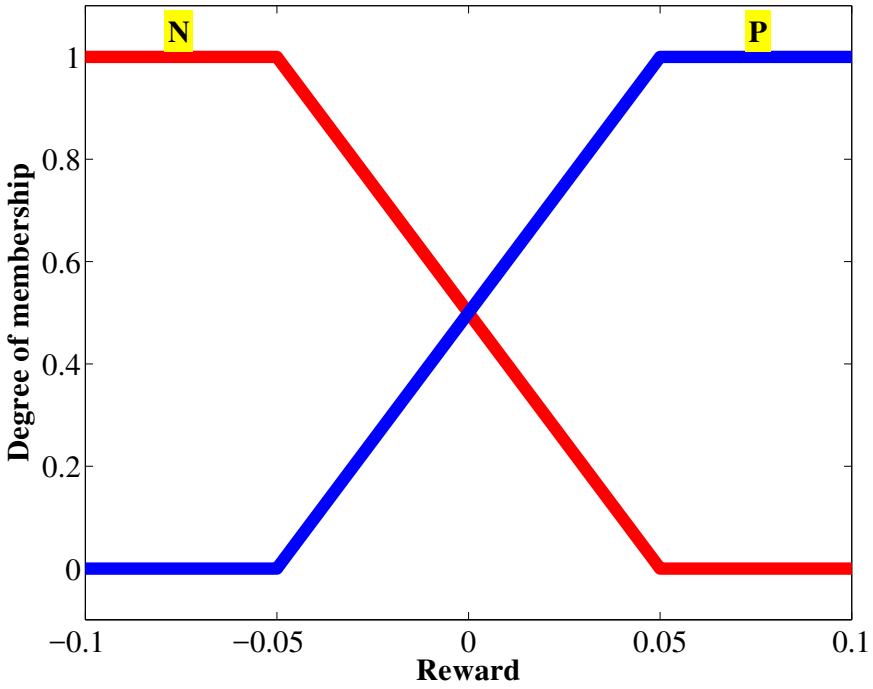

Figure 3.9: Membership Functions for Reward Sgnal Input Variable.

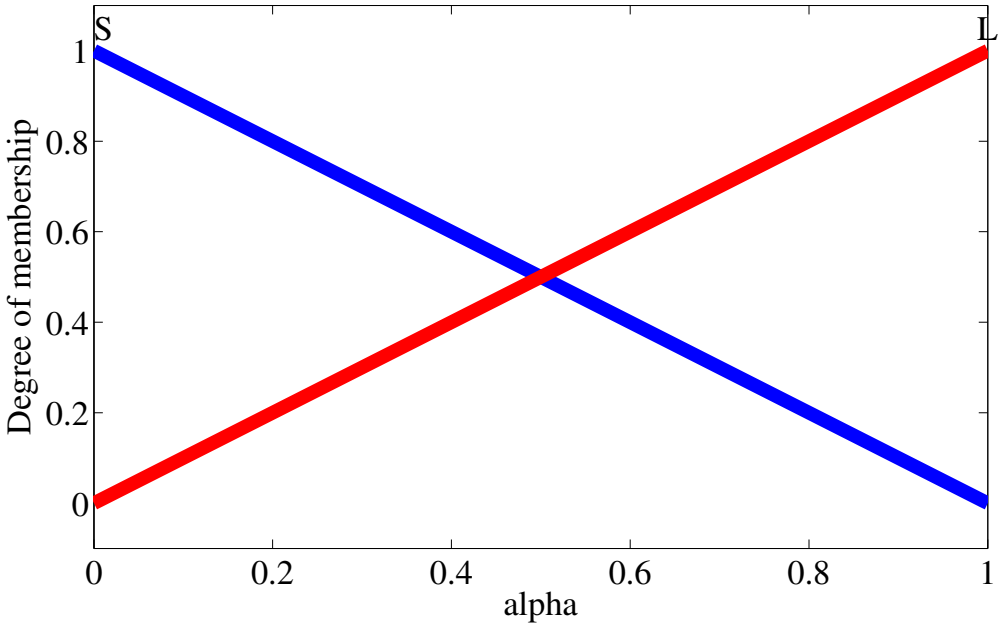

Figure 3.10: Membership Function for Alpha Output Parameter.

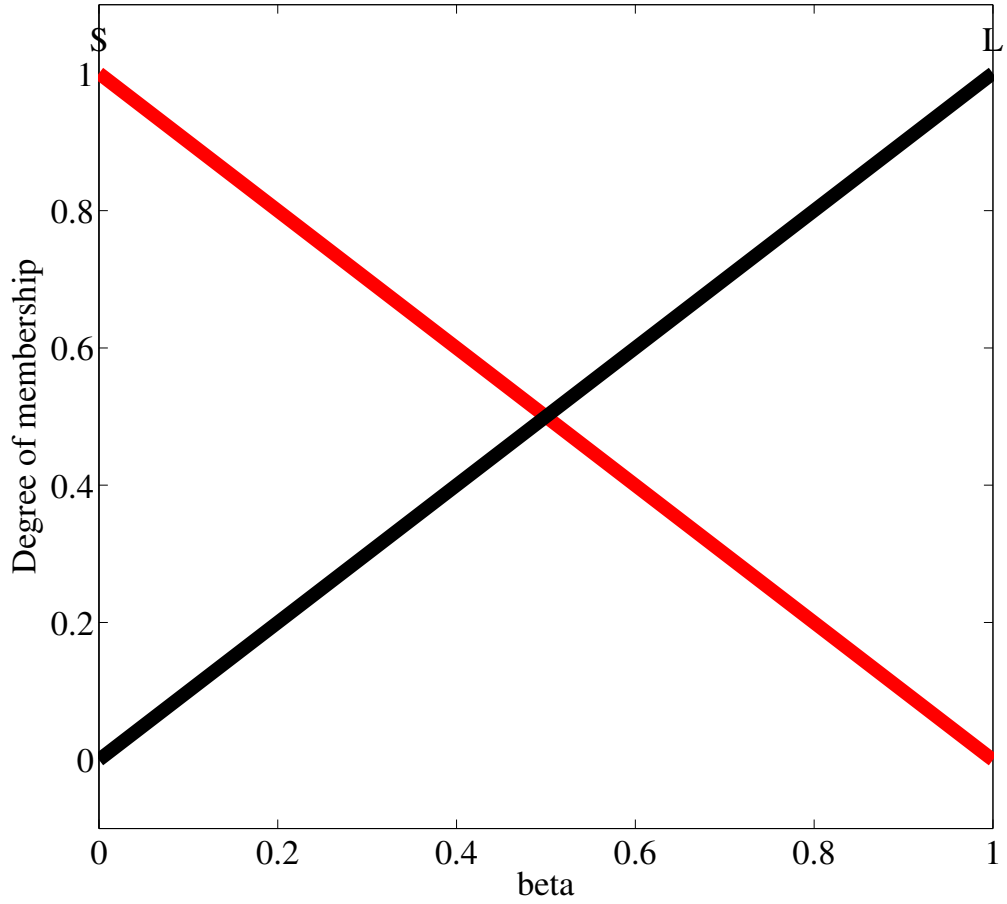

Figure 3.11: Membership Function for Beta Output Parameter.

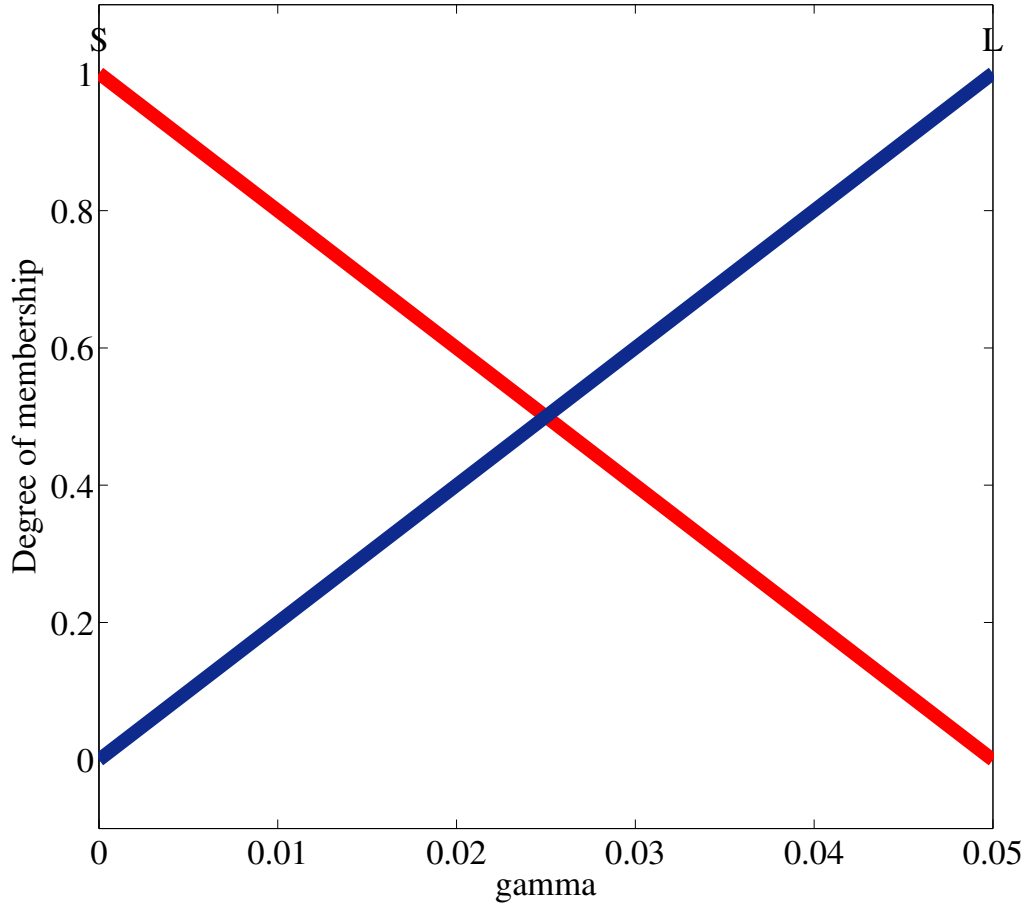

Figure 3.12: Membership Function for Gamma Output Parameter.

$$COG^g = \frac{\int_g \mu_{Ag}(g) g dg}{\int_g \mu_{Ag}(g) dg} \tag{3.51}$$

where A_g is the aggregated fuzzy set for the network parameter $g(\alpha, \beta, \gamma)$, as determined by the Mamdani fuzzy inference engine, and COG^g is the crisp value (non-fuzzy) of the network parameter $g(\alpha, \beta, \gamma)$. The variation of the network parameters α, β, γ, as computed by the proposed fuzzy model in response to the error and reward signals, is depicted in Figures 3.13, 3.14, and 3.15, respectively.

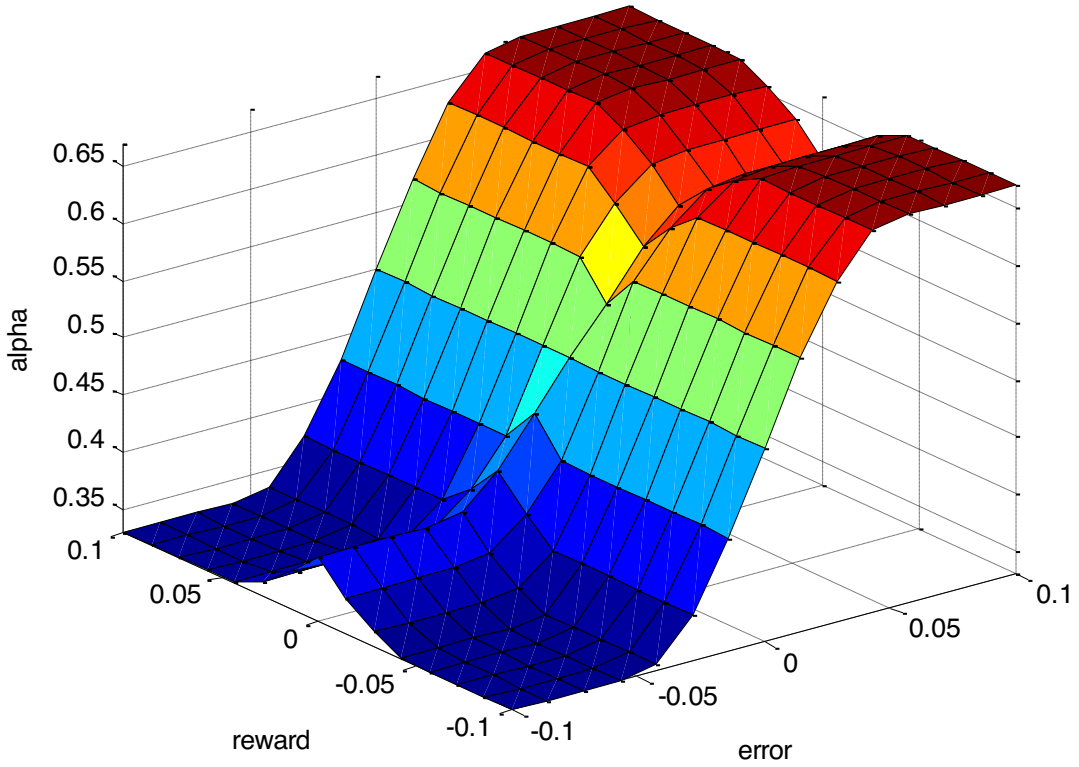

Figure 3.13: Surface for Fuzzy Integrated ADBEL Network, Alpha Parameter.

It is pertinent to mention here that the proposed fuzzy model for adjusting the network parameters is not complex. It uses a minimum number of fuzzy sets to cover the universe of discourse chosen for the parameters, which further results in a small rule base. Moreover, simple membership functions are employed in this work to

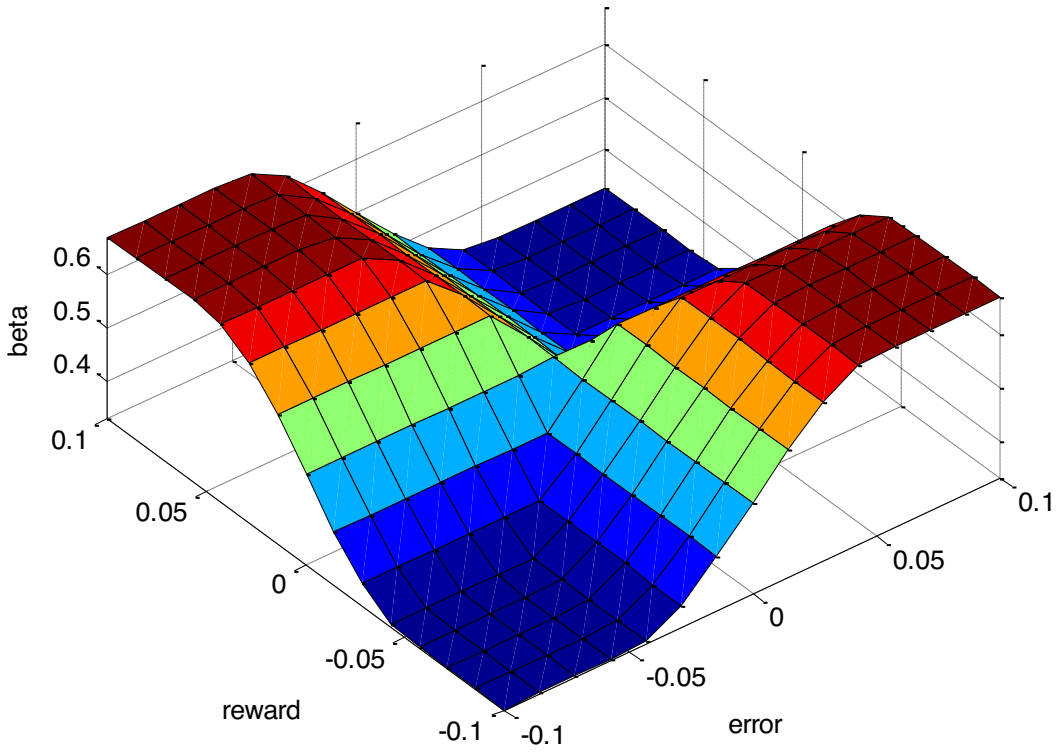

Figure 3.14: Surface for Fuzzy Integrated ADBEL Network, Beta Parameter.

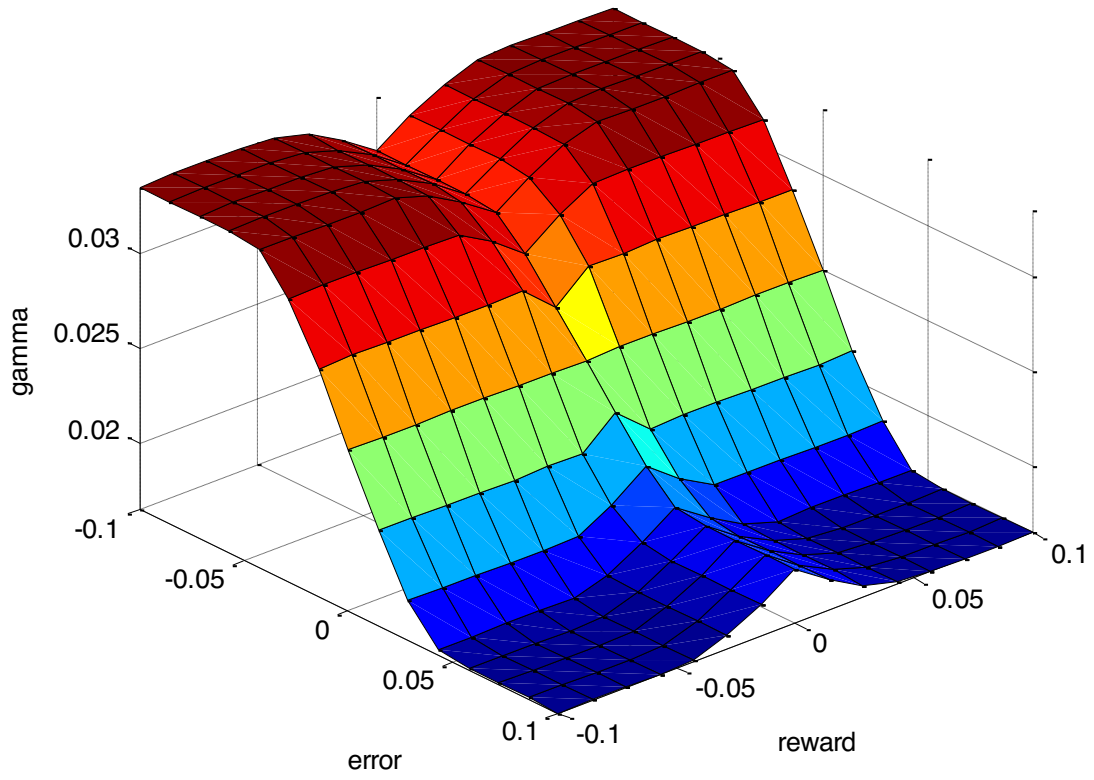

Figure 3.15: Surface for Fuzzy Integrated ADBEL Network, Gamma Parameter.

represent the fuzzy sets to keep the overall network simple. Thus, a fuzzy integrated ADBEL network still is used for online prediction of time series with sampling times in the order of a few seconds.

For the sake of completeness, a pseudo-code for the fuzzy integrated ADBEL network is presented here, where bold capital letters represent matrices, capital italic letters represent row vectors, small italic letters represent scalars, and n is the number of samples of time series:

Fuzzy integrated ADBEL online predictor

- Input data: $\mathbf{P_{in}} = (\ P_1^T \quad P_2^T \quad ... \quad P_n^T\ )$
- Output target: $P_t = (\ p_{t_1} \quad p_{t_2} \quad ... \quad p_{t_n}\ )$
- Predicted output: $\hat{P}_t = (\ \hat{p}_{t_1} \quad \hat{p}_{t_2} \quad ... \quad \hat{p}_{t_n}\ )$
- Adjustable orbitofrontal cortex weights: $W = (\ w_1 \quad w_2 \quad w_3 \quad w_4\ )$
- Adjustable amygdala weights: $V = (\ v_1 \quad v_2 \quad v_3 \quad v_4 \quad v_{th}\ )$

Fuzzy-ADBEL ($\mathbf{P_{in}}, P_t$)

1. **for** $j \leftarrow 1$ to n
2. ▷ prediction step
3. **do** $m \leftarrow max\ (P_j)$
4. $\acute{E}_a \leftarrow V.(P_j \quad 0)^T$
5. $E_a \leftarrow V.(P_j \quad m)^T$
6. $E_o \leftarrow W.P_j^T$
7. ▷ compute predicted output
8. $\hat{P}_{tj} = E_a - E_o$
9. ▷ learning step
10. $e \leftarrow p_{tj} - \hat{p}_{tj}$

11. $R_o \leftarrow \begin{cases} max(\acute{E}_a - p_t, 0) - E_o \ , & \text{if } (p_t \neq 0) \\ max(\acute{E}_a - E_o, 0) \ , & \text{otherwise} \end{cases}$

12. ▷ compute network parameters

13. $\alpha \leftarrow A_\alpha . \mu_{A_{\alpha n}}$

14. $\beta \leftarrow A_\beta . \mu_{A_{\beta n}}$

15. $\gamma \leftarrow A_\gamma . \mu_{A_{\gamma n}}$

16. ▷ update network weights

17. $W \leftarrow W + \beta R_o P_j$

18. $V \leftarrow (1 - \gamma)V + \alpha \ max(p_{t_j} - E_a, 0)(P_j \quad m)$

19. **end**

Chapter 4

RESULTS AND DISCUSSION

In this chapter, we discuss the performance results of each designed proposed model in different stages, as follows:

Stage 1: We discuss the performance of each designed proposed model (NF-ADBEL, ENF-ADBEL and F-ADBEL) and compare these performances to the existing ADBEL model. Please note that the current ADBEL model was also redesigned in this work and programmed in a MATLAB simulation for a fair comparison.

Stage 2: We compare the performances of the designed proposed models.

Stage 3: We compare the performances of the proposed models to other state-of-the-art predictors in the literature.

4.1 Performance of the Proposed Neo-Fuzzy Adaptive Decayed Brain Emotional Learning (NF-ADBEL) Model

The proposed neo-fuzzy integrated NF-ADBEL network is tested in a MATLAB (R2014a) programming environment for online forecasting of chaotic time series, including Mackey-Glass, Lorenz, Rossler, Narendra, and the disturbance storm time index.

The NF-ADBEL network's performance is accessed in terms of root mean squared error and correlation coefficient criteria. A comparison is also made with the ADBEL network driven by the near optimal set of parameters, where the percentage improvement index is used as a basis for comparison. These performance indices are defined as:

$$RMSE_m = \sqrt{\frac{1}{n_e - n_s} \sum_{i=n_s}^{n_e} e_{mi}^2} \tag{4.1}$$

$$COR_m = \frac{\sum_{i=n_s}^{n_e}(\hat{P}_{mti} - \overline{\hat{P}_{mt}})(P_{ti} - \overline{P_t})}{\sqrt{\sum_{i=n_s}^{n_e}(\hat{P}_{mti} - \overline{\hat{P}_{mt}})^2}\sqrt{\sum_{i=n_s}^{n_e}(P_{ti} - \overline{P}_t)^2}} \tag{4.2}$$

$$PI = \frac{PC_{m_1} - PC_{m_2}}{PC_{m_1}} \times 100 \tag{4.3}$$

where $RMSE_m$ is the root mean squared error, and subscript m can be either m_2 denoting the neo-fuzzy ADBEL network or m_1 representing ADBEL network without a neo-fuzzy neuron, n_e is the number of samples, n_s indicates the start of steady-state period, COR_m is the correlation coefficient obtained using m^{th} network, $\hat{P_{mti}}$ is the predicted value with m^{th} network, $\overline{\hat{P_{mt}}}$ is the mean of the predicted values with m^{th} network, P_{ti} is the target value, $\overline{P_t}$ is the mean of the target samples, PC_m is the performance criterion (which can be either $RMSE_m$ or COR_m), and index PIi s the percentage decrease with respect to the ADBEL network, if a low root mean squared error is achieved by the neo-fuzzy based ADBEL network. In this case, it could be treated as a percentage improvement concerning the neo-fuzzy integrated ADBEL network.

The time-series data is first normalized to the range [0, 1] by running the simulations. The normalized data is then organized in such a way that the first four samples form the inputs, while the fifth sample presents the output. By following this method, the size of input data for time series is set as $4 \times n_e$ while the size of the output data is set as $1 \times n_e$. Note that the number of samples n_e can differ for different time series, depending on availability. Also note that the total number of data used in this work can comply with this formula: $Total\ number\ data = n_e + number\ of\ a\ network\ inputs$.

After running the ADBEL and neo-fuzzy ADBEL algorithms, the predicted time series data is de-normalized. In this work, all the weights are initialized as zeros instead of randomly assigning them. This procedure also helps run the simulations only once; furthermore, no averaging of the results is required, as the networks will yield the same performance every time. The networks' learning parameters (namely, α, β, and γ) are selected after extensive experimentation to yield the best possible

prediction performance in each case.

4.1.1 Mackey-Glass Time Series Predicted by the Proposed NF-ADBEL Network

The Mackey-Glass system has been presented as a model of white blood cell production [67]. Let us first predict the time series data generated from a time-delayed Mackey-Glass nonlinear differential equation, which has been used as a benchmark by the researchers for validating their prediction algorithms [68],[69],[70]. The series could be defined as:

$$\dot{x}(t) = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t) \tag{4.4}$$

With the initial conditions as $x(t) = 0$, $t < 0$; $x(0) = 1.2$ and by setting the time delay as $\tau = 17$, Equation (4.4) is simulated in MATLAB to generate the time series data, which can be observed to be non-periodic and non-convergent. A total of $n_e = 1200$ data points, as shown in Appendix A, are generated for testing the networks.

By setting the learning parameters as $\alpha = 0.5$, $\beta = 0.2$ and $\gamma = 0.03$, an NF-ADBEL network is first deployed to predict this time series. The steady-state result over a pre-defined time window is depicted in Figures 4.1, 4.2, 4.3, respectively.

The same time series is also predicted with an ADBEL network using the learning parameters of $\alpha = 0.5$, $\beta = 0.8$, and $\gamma = 0.03$. The prediction error is recorded in both cases, with analysis showing that the transient period remains the same ≤ 5 s. Thus, the steady-state starting index is set as ≤ 5 s to compute the performance indices in both cases. The performance of the pre-designed ADBEL network for predicting Mackey-Glass is shown in Figures 4.4, 4.5, and 4.6, respectively.

Figures 4.8, and 4.10 displayed a portion of Figures-4.7, and 4.9 that show the zoomed view of the prediction and error prediction in steady-state as yielded by both the ADBEL and NF-ADBEL networks for predicting the Mackey-Glass time series. As can be seen, the NF-ADBEL network has performed better than the ADBEL network, showing lower peaks in the prediction error in the NF-ADBEL network. Furthermore, the root means squared error and correlation coefficient are

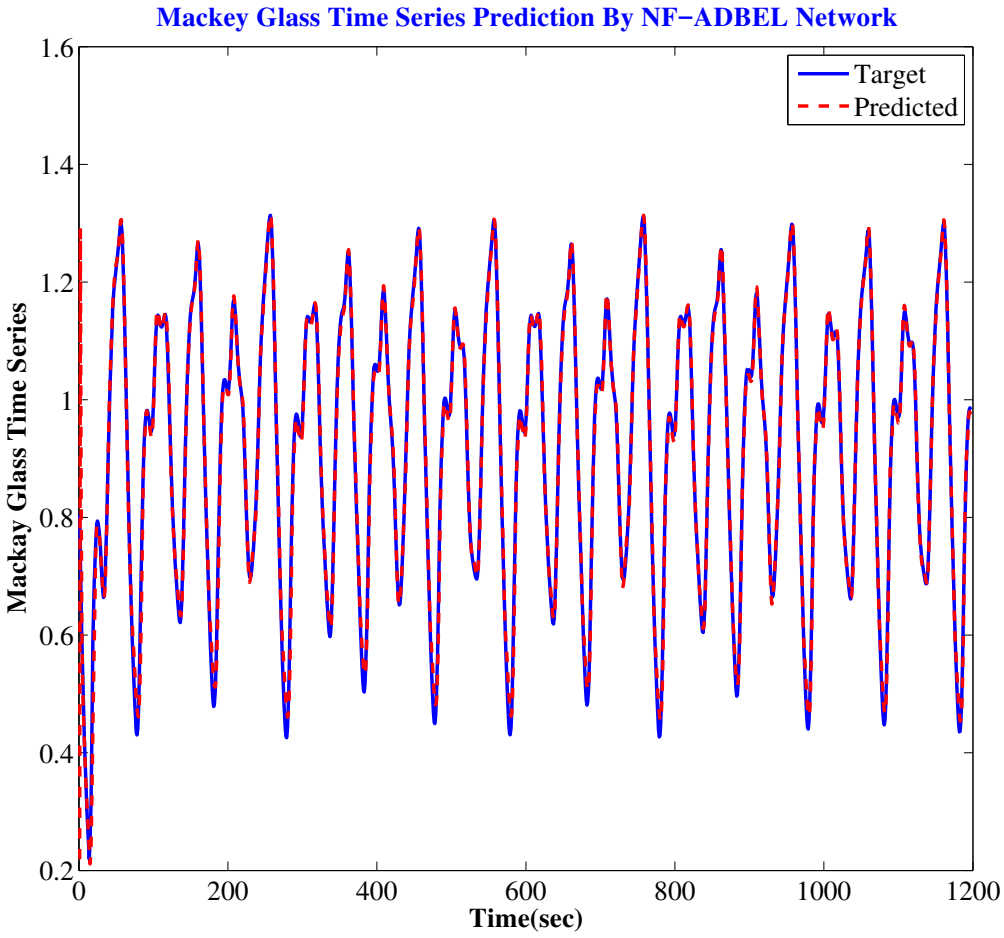

Figure 4.1: Mackey Glass Time Series as Predicted by NF-ADBEL Network.

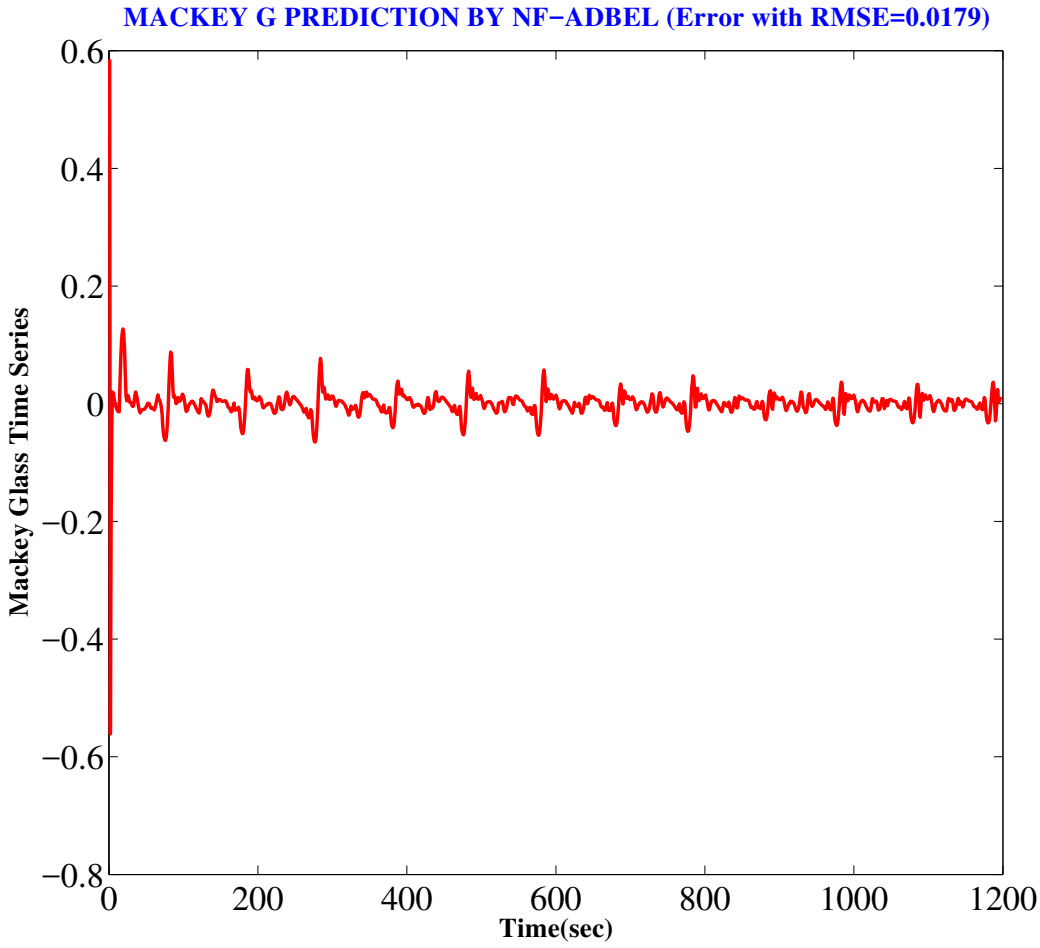

Figure 4.2: Error in Predicting Mackey-Glass Time Series by NF-ADBEL Network.

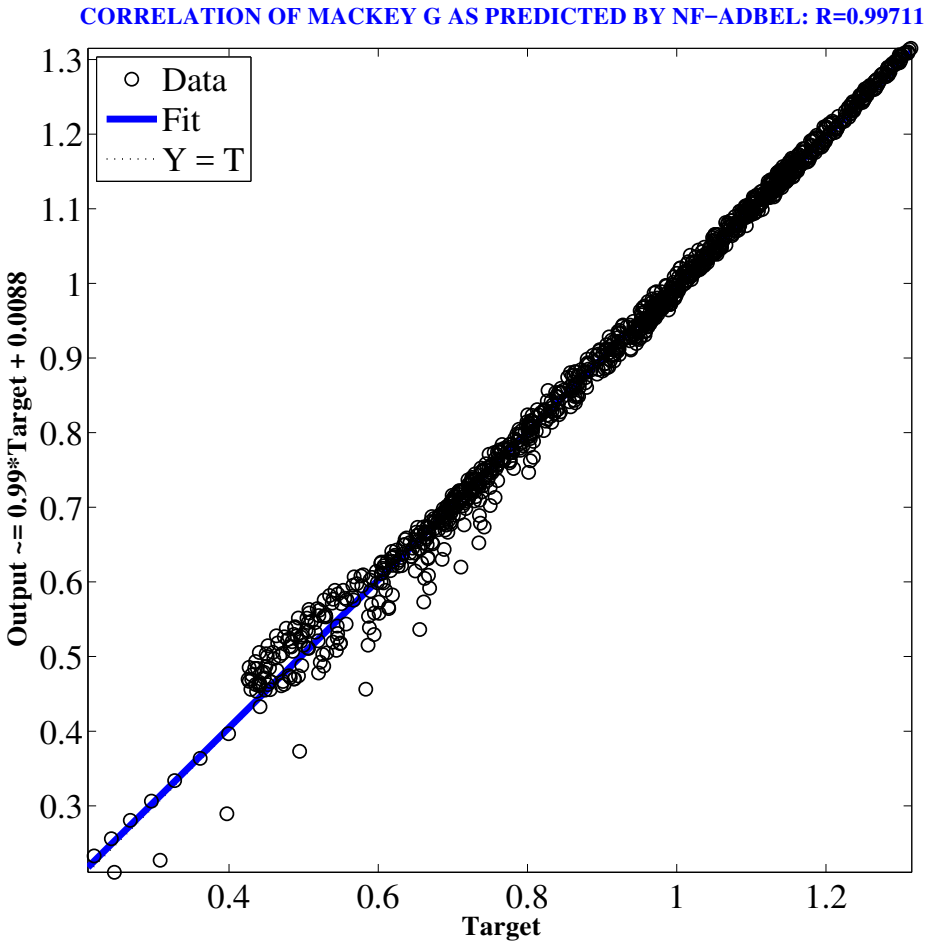

Figure 4.3: Correlation in Predicting Mackey-Glass Time Series by NF-ADBEL Network.

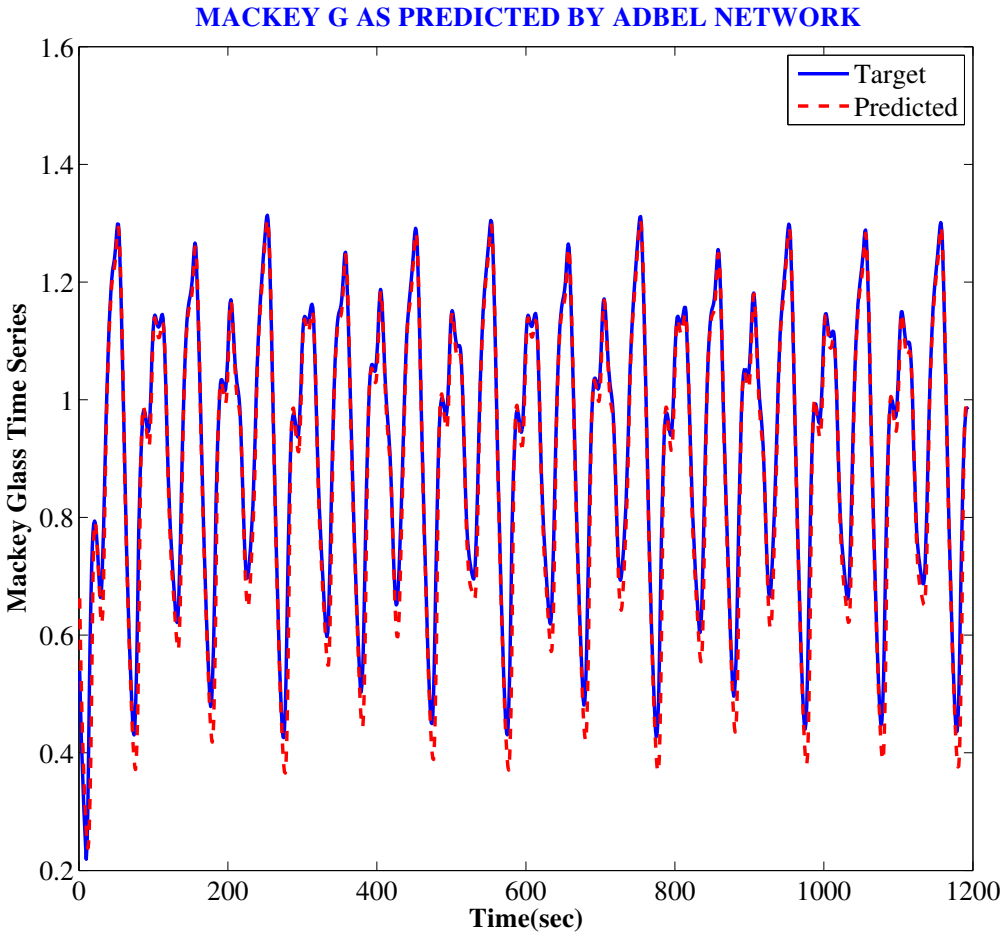

Figure 4.4: Mackey-Glass Time Series as Predicted by ADBEL Network.

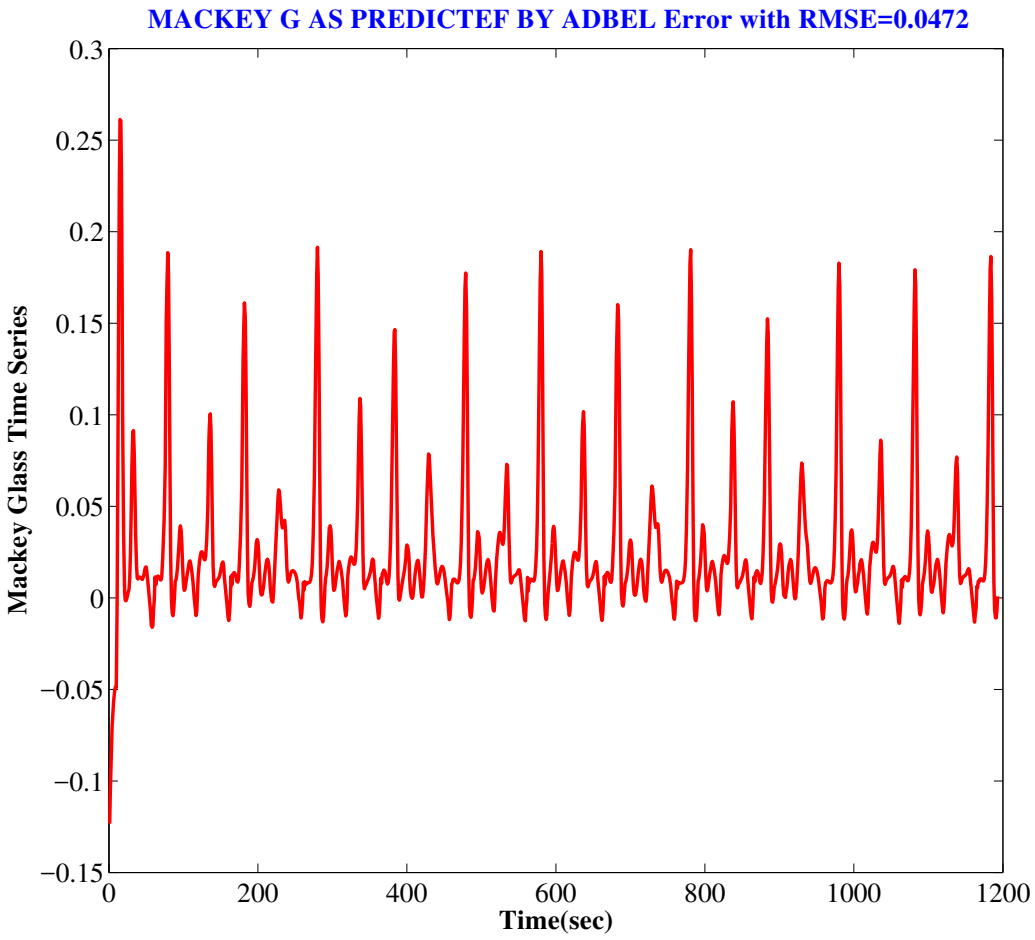

Figure 4.5: Error in Predicting Mackey-Glass Time Series by ADBEL Network.

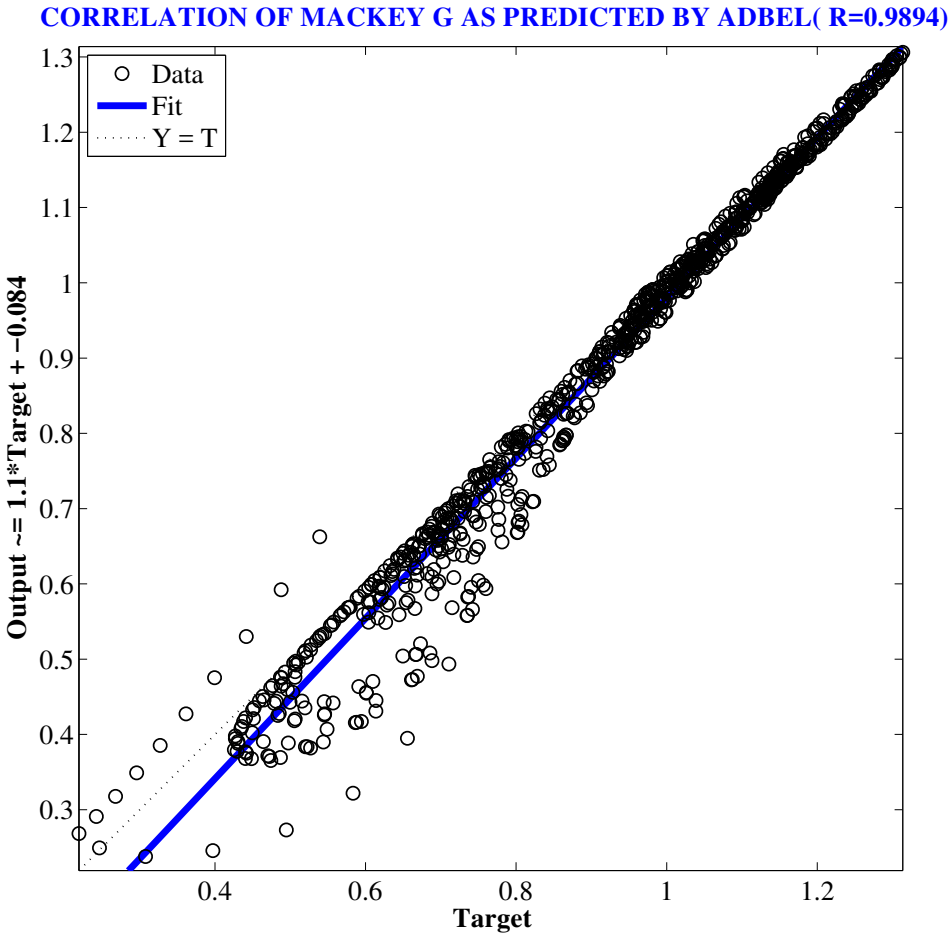

Figure 4.6: Correlation in Predicting Mackey-Glass Time Series by ADBEL Network.

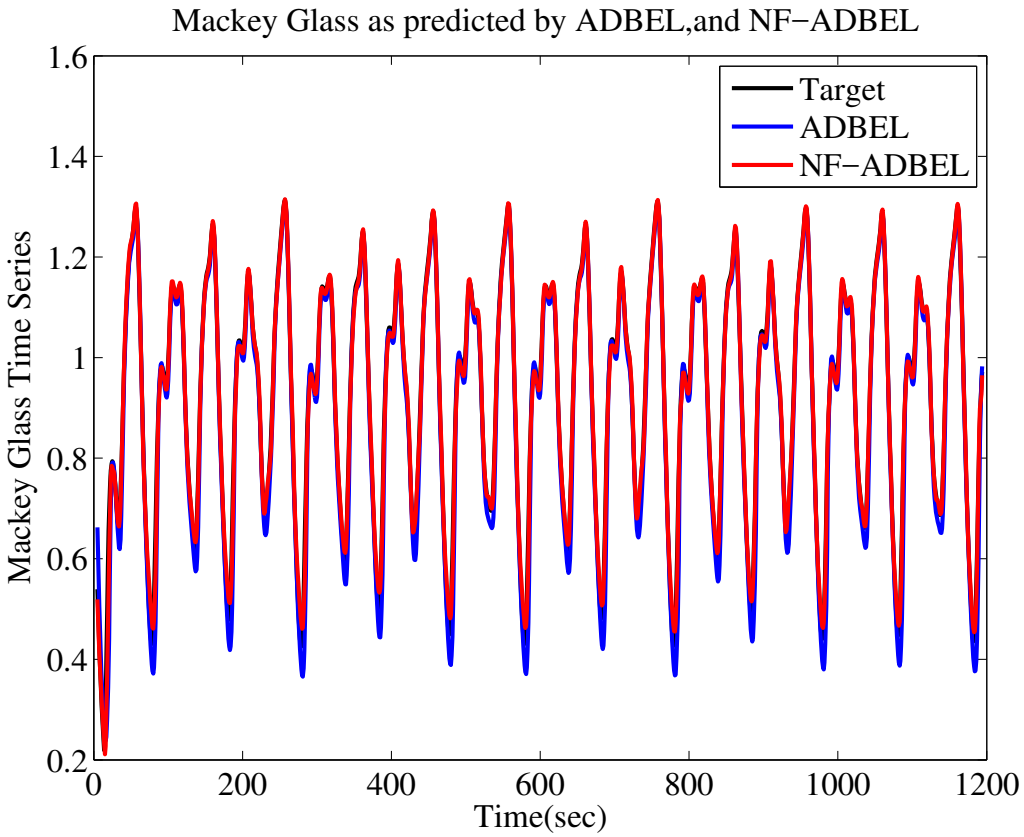

Figure 4.7: Mackey-Glass Time Series as Predicted by ADBEL and NF-ADBEL Networks.

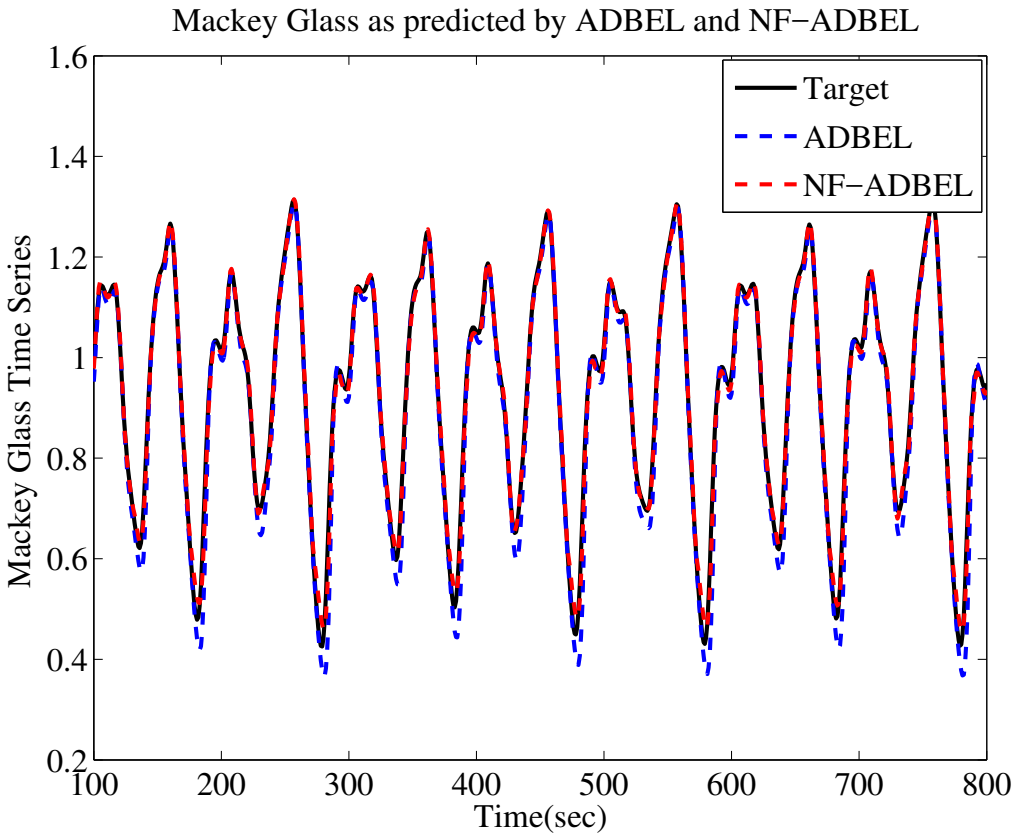

Figure 4.8: Mackey-Glass Time Series as Predicted by ADBEL and NF-ADBEL Networks (displayed portion of the Fig. 4.7).

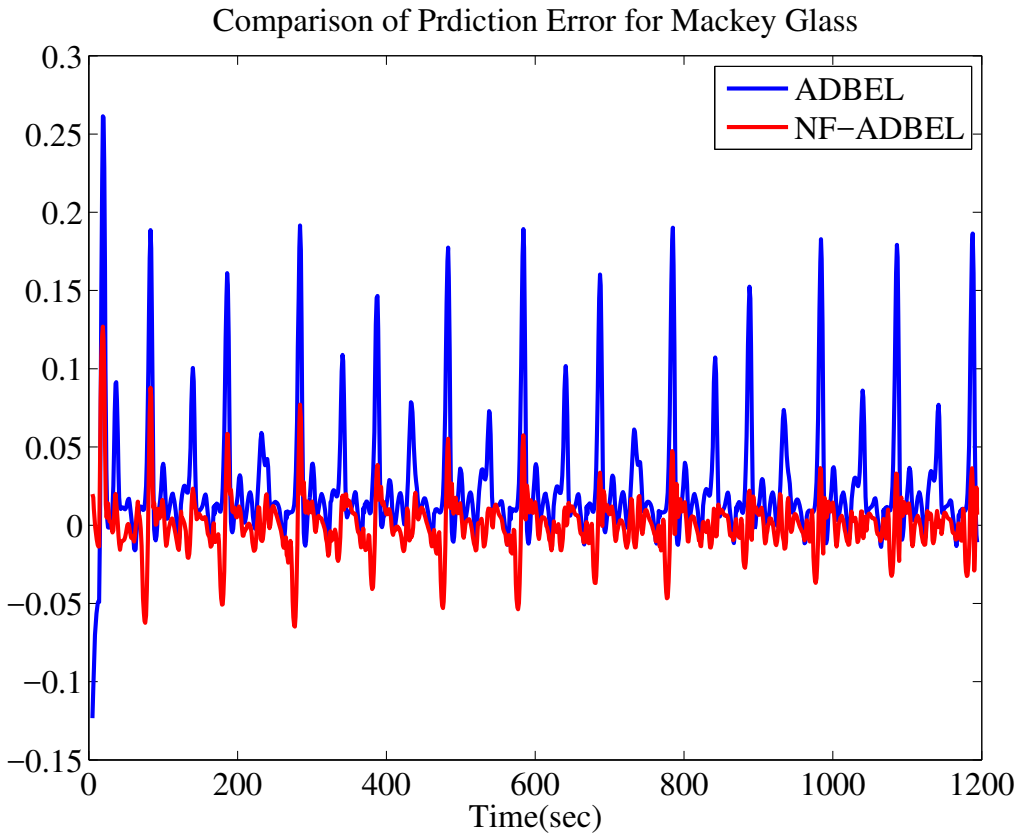

Figure 4.9: Error Comparison in Predicting Mackey-Glass Time Series by ADBEL and NF-ADBEL Networks.

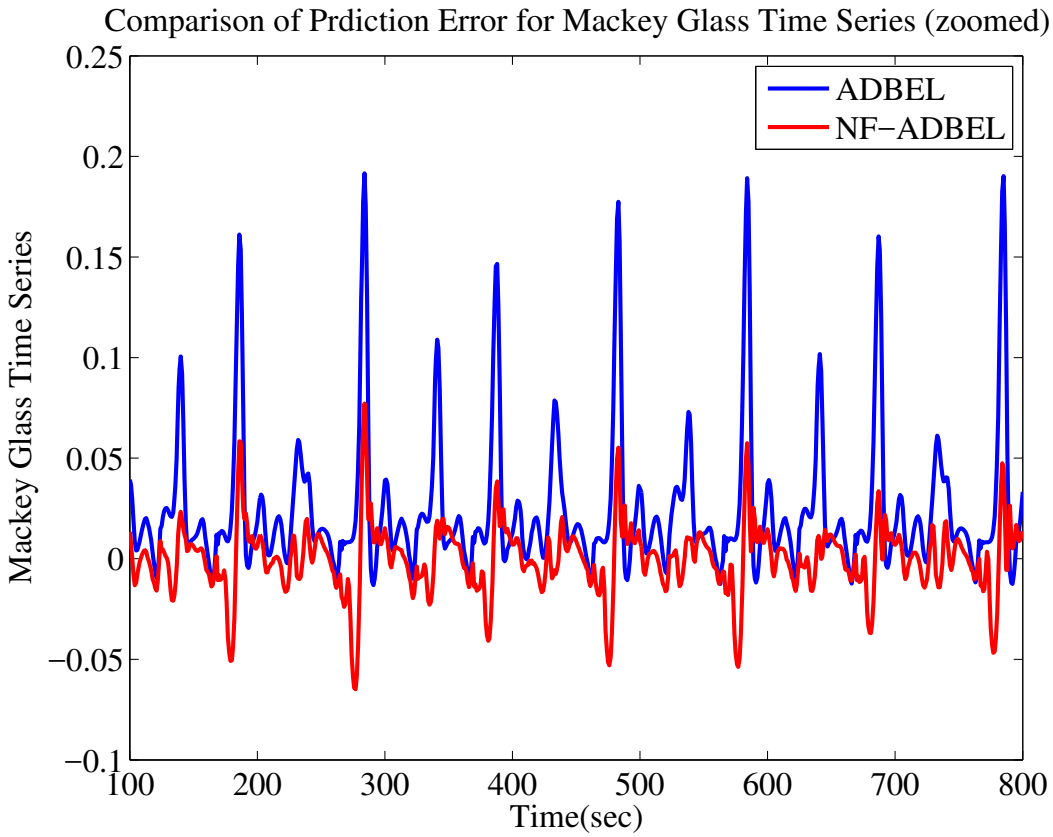

Figure 4.10: Error Comparison in Predicting Mackey-Glass Time Series by ADBEL and NF-ADBEL Networks (displayed portion of the Fig. 4.9).

also determined for both networks using the relations in Eqs. (4.1) and (4.2). The computed values are shown in Table 4.1.

Table 4.1: RMSE/COR/PI FOR MACKEY-GLASS TIME SERIES PREDICTION BY ADBEL AND NF-ADBEL NETWORKS

Time Series	Prediction Network	RMSE	COR	PI(%)
Mackey-Glass	ADBEL	0.04727	0.98952	
	NF-ADBEL	0.0180	0.99706	**61.92**

As can be seen, a lower root mean squared error and higher correlation coefficient are offered by the NF-ADBEL network for predicting the Mackey-Glass time series compared to the ADBEL network, as shown in Figures 4.8 and 4.10, 4.3 and 4.6, respectively. A significant amount of percentage improvement is also obtained, as expressed in Eq. (4.3) and shown in Table 4.1.

4.1.2 Lorenz Chaotic Time Series Predicted by Proposed NF-ADBEL Network

The Lorenz system was presented in 1963 by Lorenz in [71]. Neo-fuzzy-integrated ADBEL is simulated to predict the x-dynamics of the Lorenz chaotic time series. This series has also been used in various studies to verify the performance of prediction algorithms [72],[73],[74],[75]. The series is generated by [71] from the following coupled differential equations with $a = 10$, $b = 28$, and $c = 8/3$:

$$\begin{cases} \dot{x}(t) = a(y(t) - x(t)) \ , \\ \dot{y}(t) = x(t)(b - z(t)) - y(t) \ , \\ \dot{z}(t) = x(t)y(t) - cz(t) \end{cases} \tag{4.5}$$

For the generated Lorenz time series with $n_e = 16380$ data points (see Appendix A), we first evaluate the prediction performance of the NF-ADBEL network, as shown in Figure 4.12, with the learning parameters set as: $\alpha = 0.8$, $\beta = 0.2$, and $\gamma = 0.01$.

Figure 4.12 shows some data points of the Lorenz time series as predicted by the NF-ADBEL network. It can be observed in Figures 4.12, 4.14 and 4.15 that

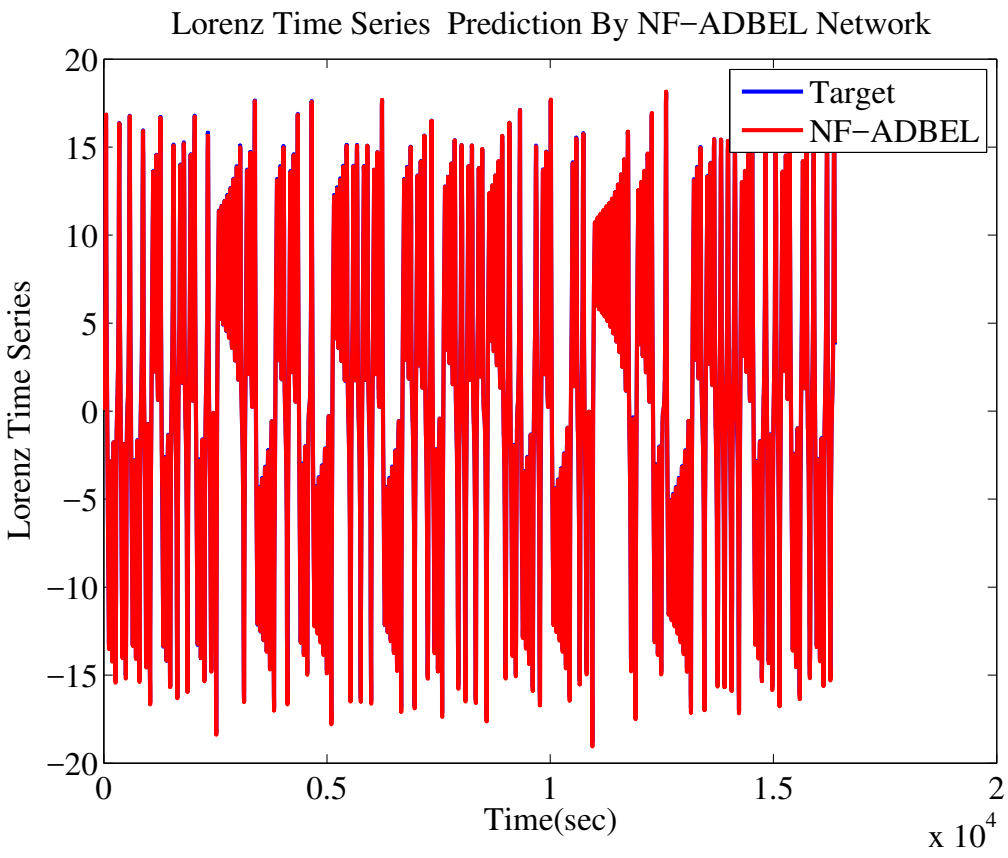

Figure 4.11: Lorenz x-Time Series as Predicted by NF-ADBEL Network .

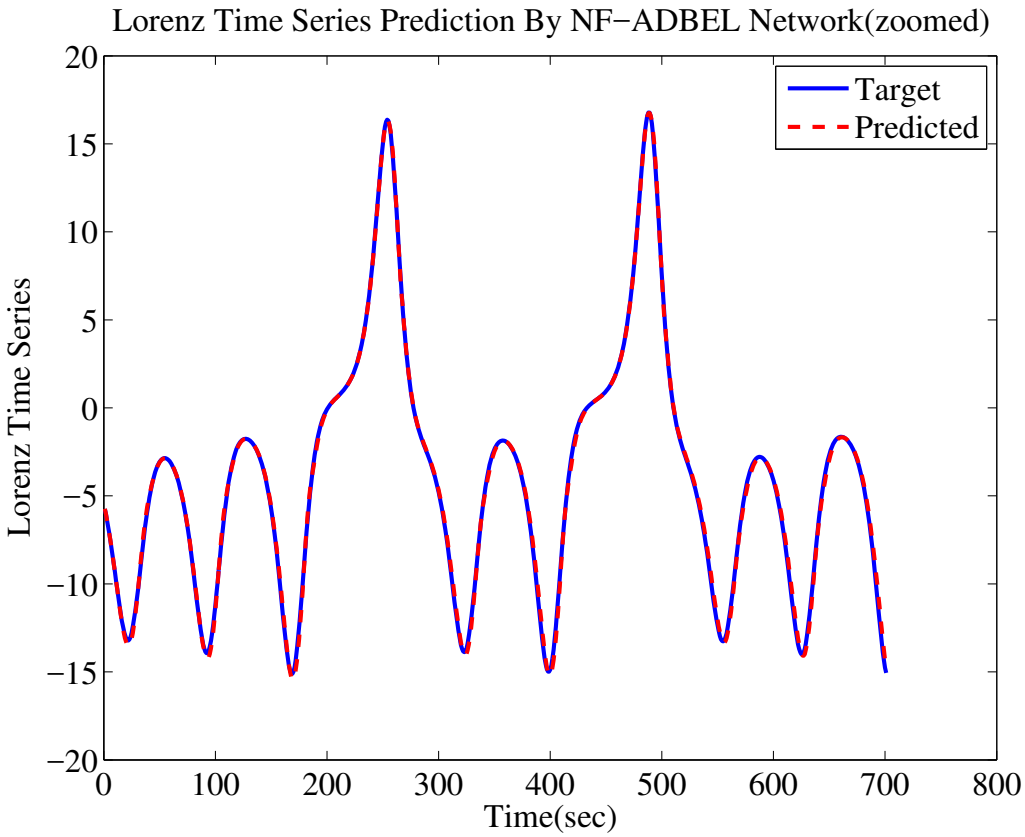

Figure 4.12: Lorenz x-Time Series as Predicted by NF-ADBEL Network (displayed portion of the Fig. 4.11).

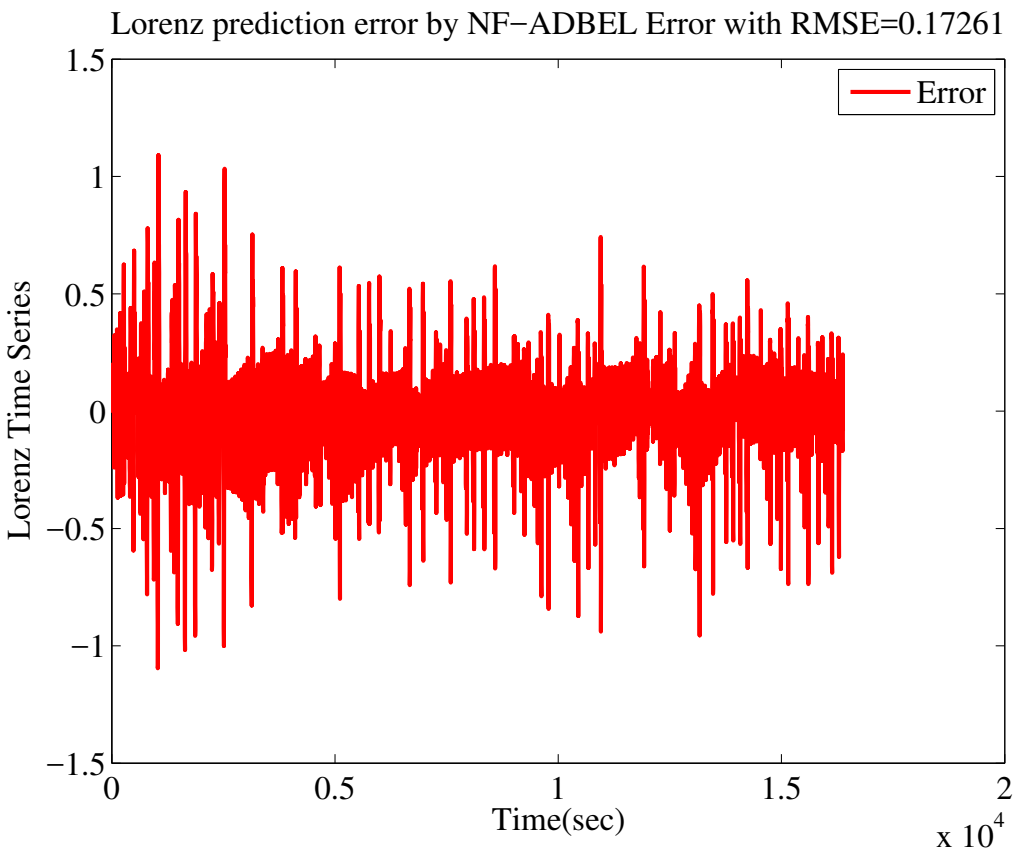

Figure 4.13: Error in Predicting Lorenz x-Time Series by NF-ADBEL Network.

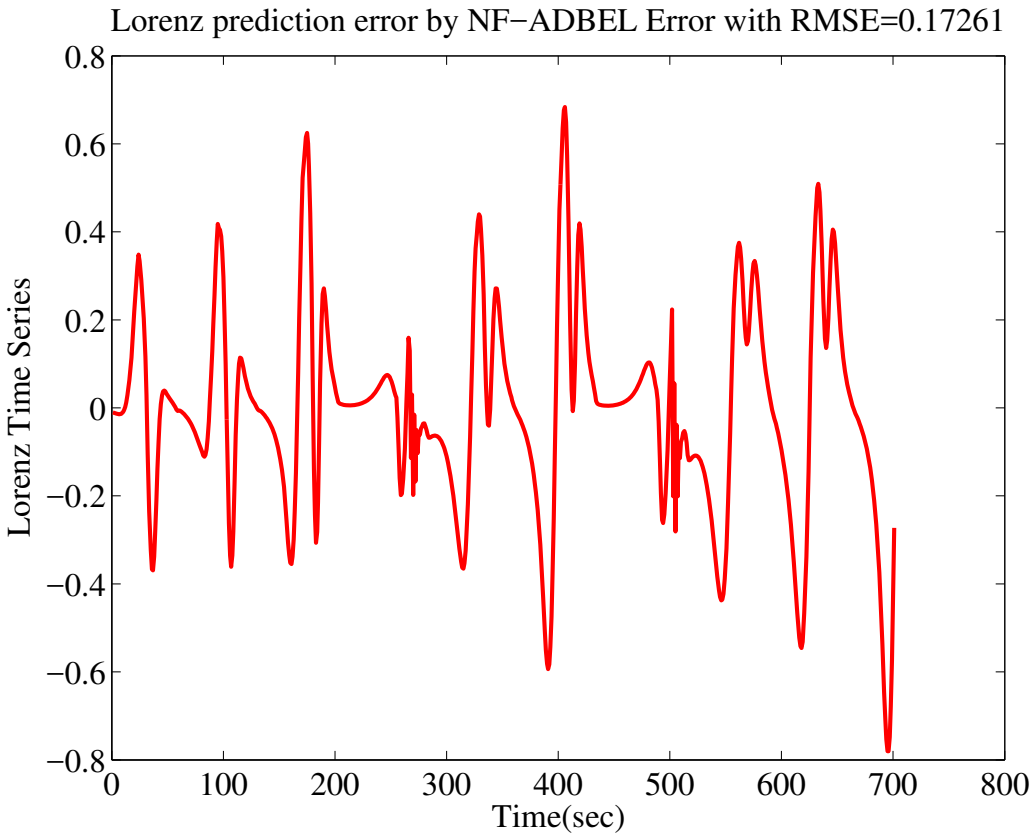

Figure 4.14: Error in Predicting Lorenz x-Time Series by NF-ADBEL Network (displayed portion of the Fig. 4.13)

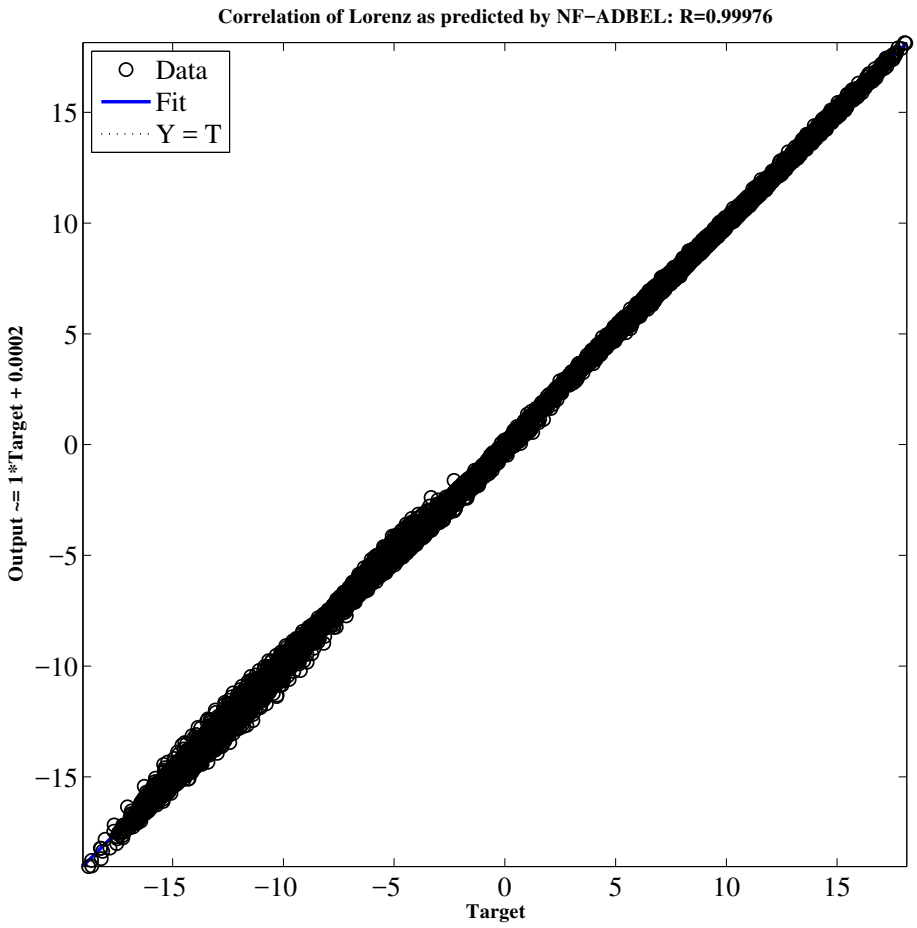

Figure 4.15: Correlation in Predicting Lorenz x-Time Series by NF-ADBEL Network.

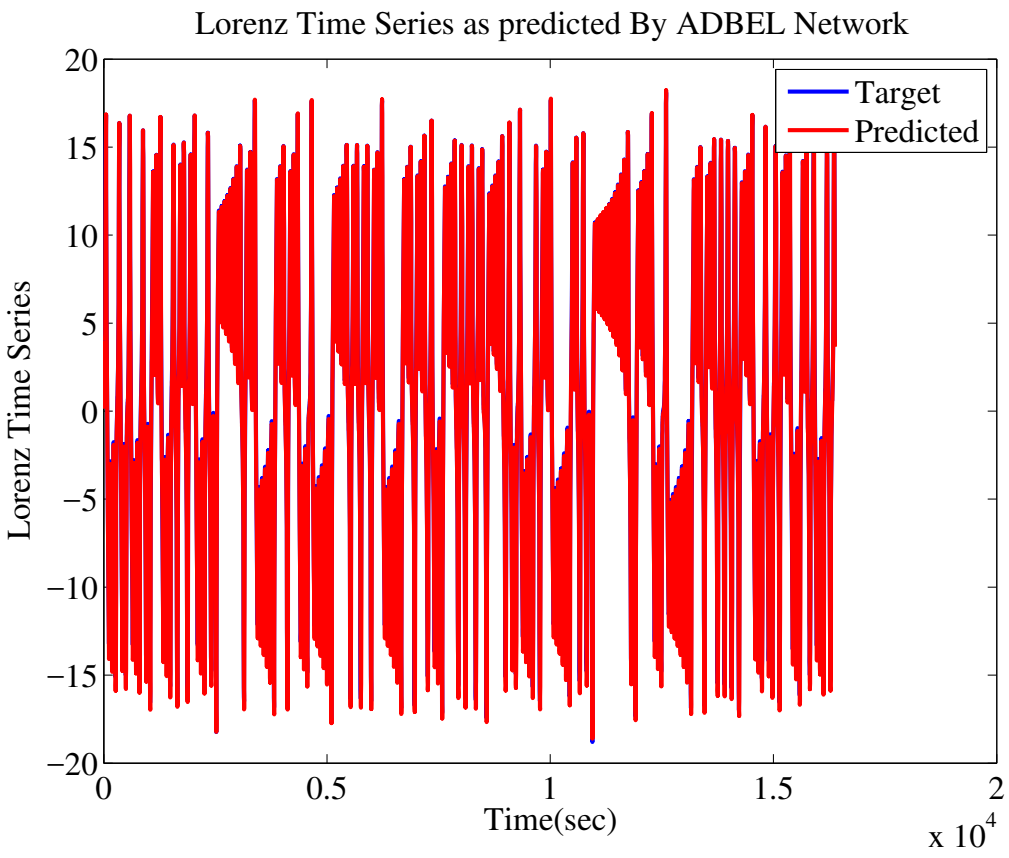

Figure 4.16: Lorenz x-Time Series as Predicted by ADBEL Network.

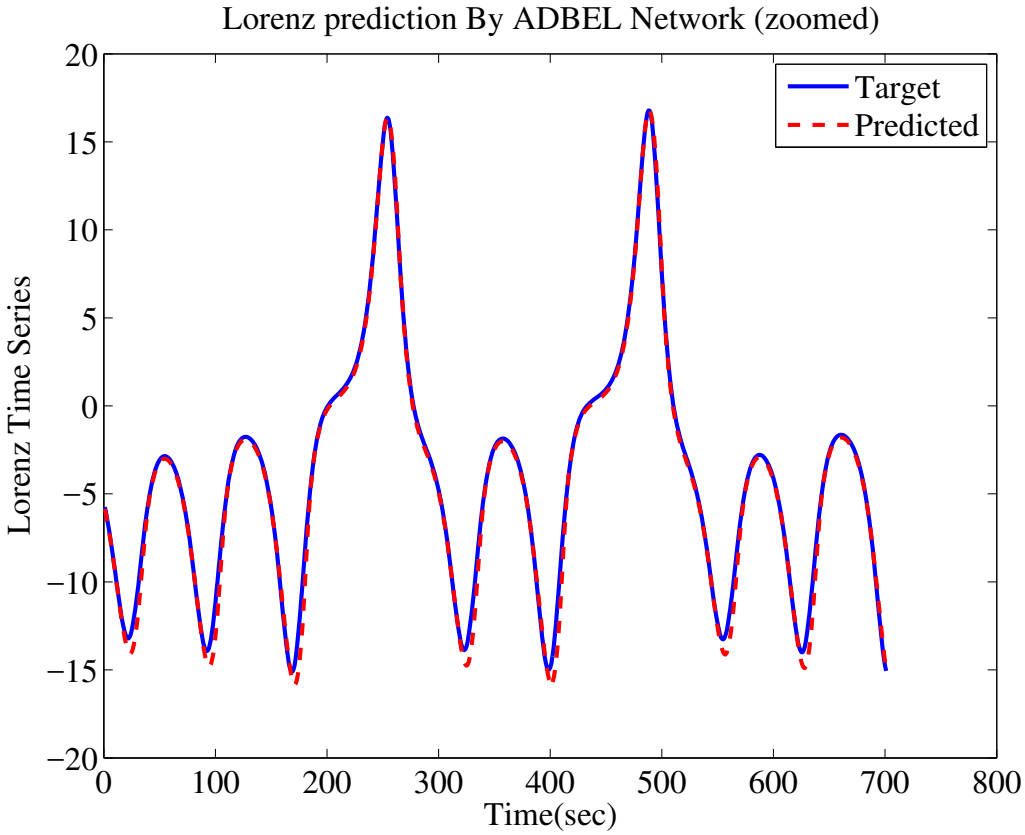

Figure 4.17: Lorenz x-Time Series as Predicted by ADBEL Network (displayed portion of the Fig. 4.16).

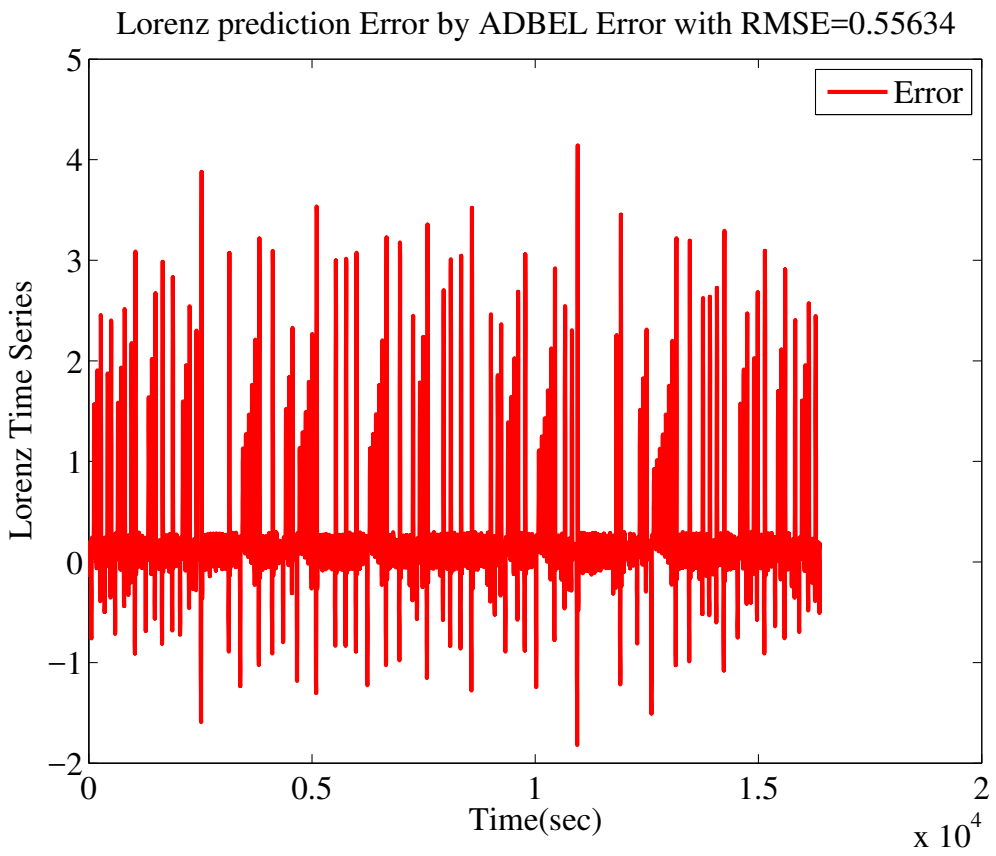

Figure 4.18: Error in Predicting Lorenz x-Time Series by ADBEL Network.

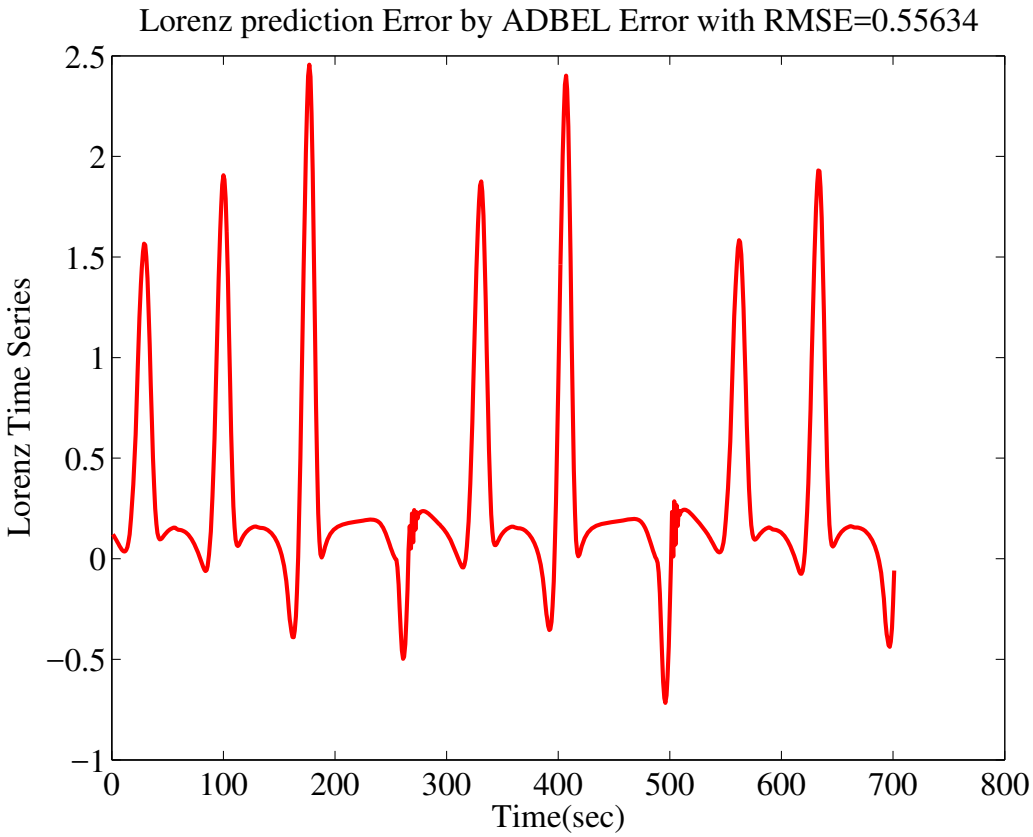

Figure 4.19: Error in Predicting Lorenz x-Time Series by ADBEL Network (displayed portion of the Fig. 4.18) .

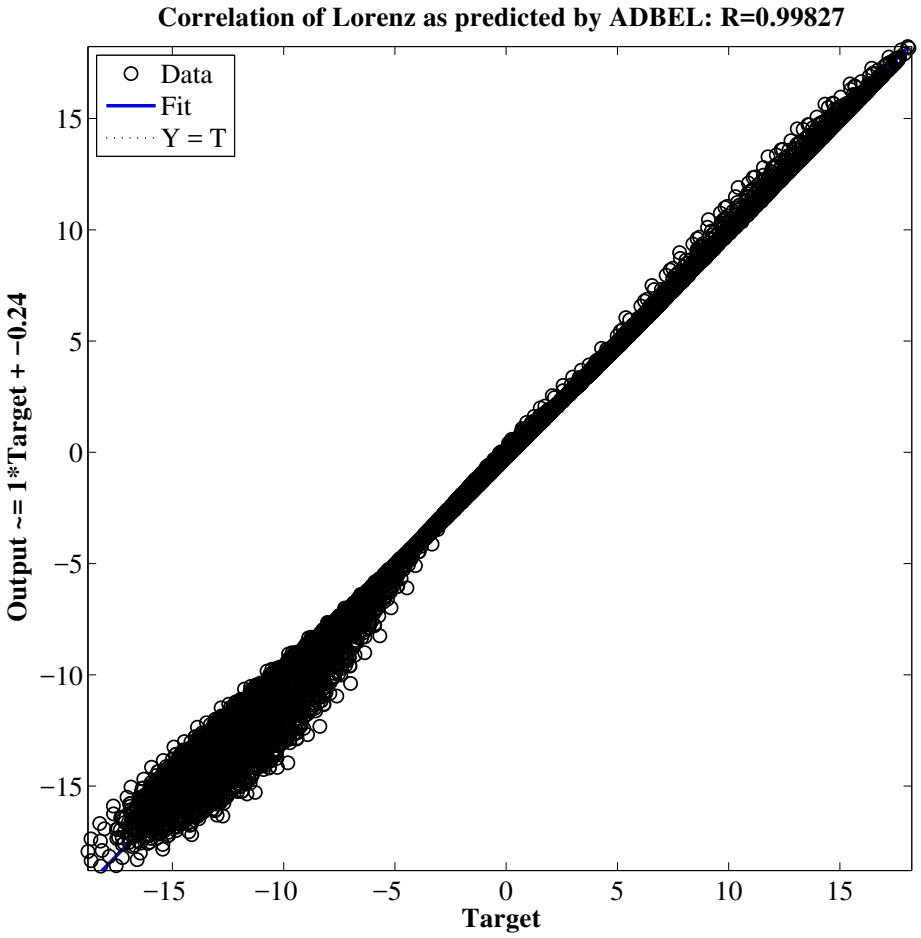

Figure 4.20: Correlation in Predicting Lorenz x-Time Series by ADBEL Network.

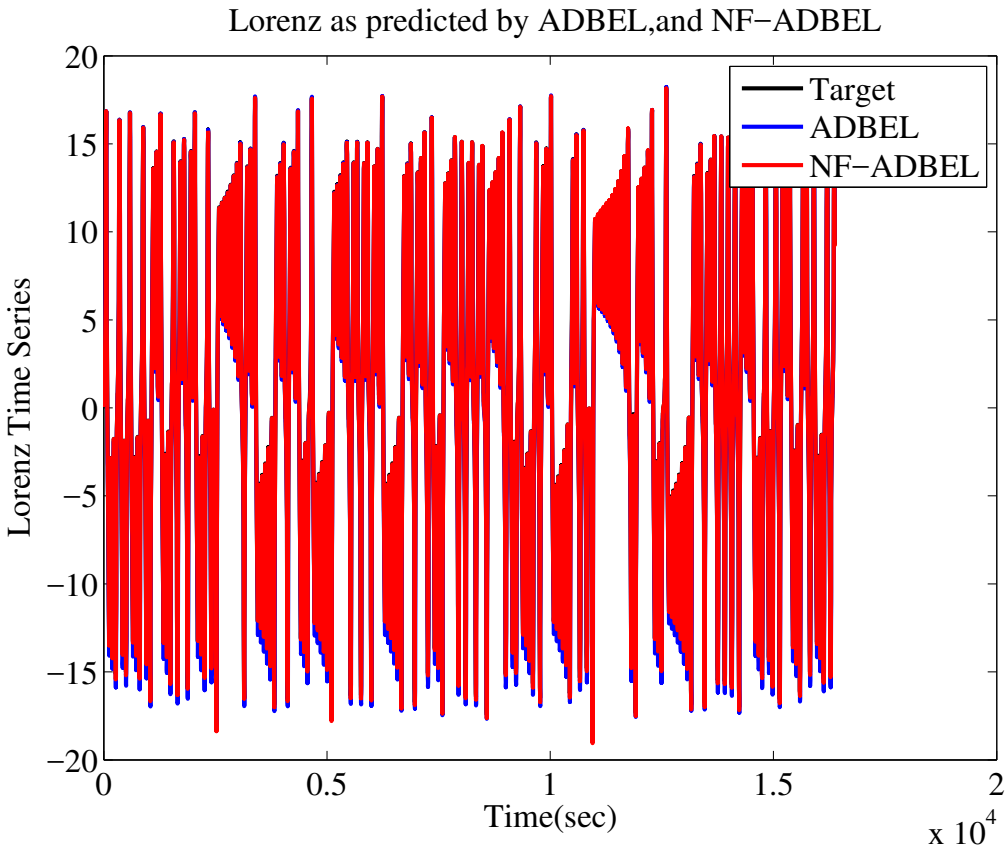

Figure 4.21: Lorenz x-Time Series as Predicted by ADBEL and NF-ADBEL Networks.

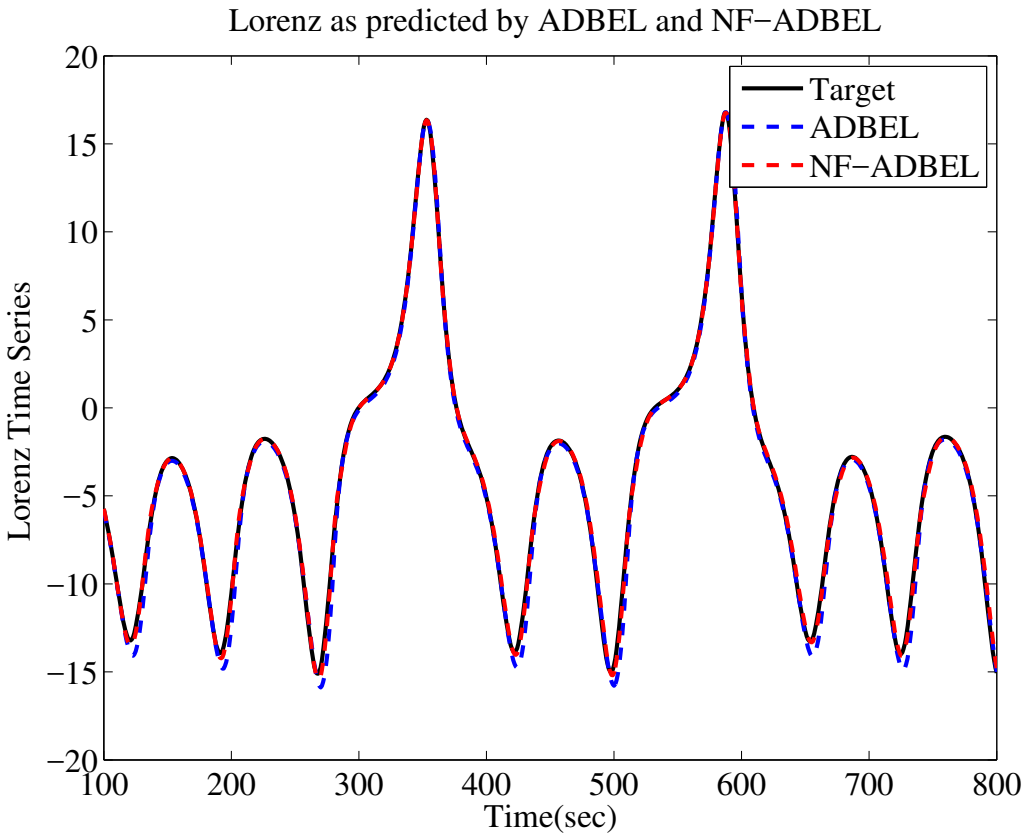

Figure 4.22: Lorenz x-Time Series as Predicted by ADBEL and NF-ADBEL Networks (displayed portion of Fig 4.21).

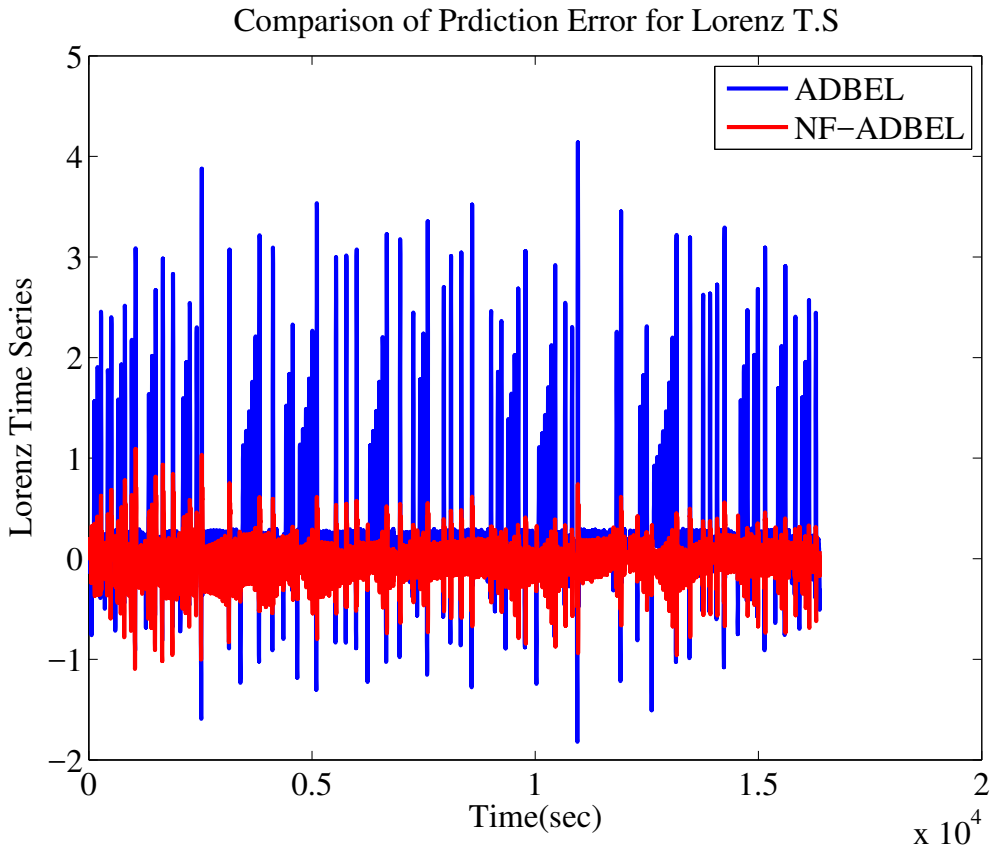

Figure 4.23: Comparison of Errors in Predicting Lorenz x-Time Series by ADBEL and NF-ADBEL Networks.

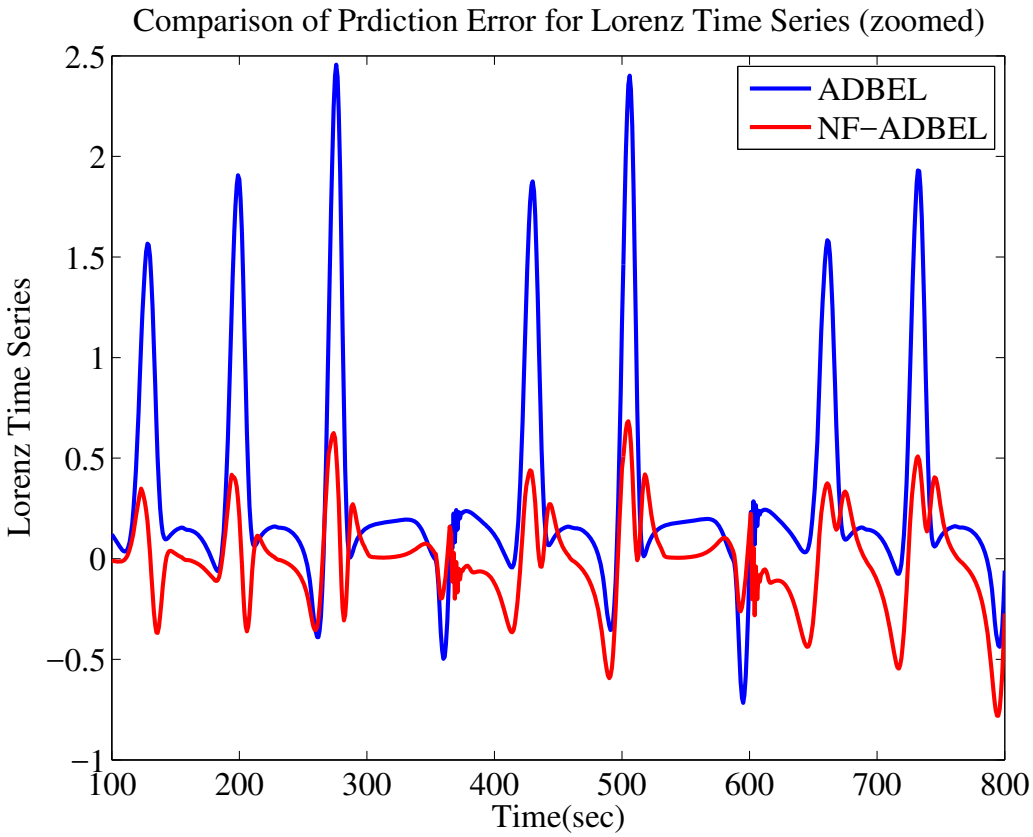

Figure 4.24: Comparison of Errors in Predicting Lorenz x-Time Series by ADBEL and NF-ADBEL Networks (displayed portion of Fig 4.23).

the prediction results of the NF-ADBEL network for the Lorenz time series performs better compared to results for the Mackey-Glass time series. Specifically, it is difficult to distinguish the predicted Lorenz time series from the original one.

For comparison purposes, the ADBEL network is also simulated to predict the Lorenz time series. The best learning parameters for the ADBEL network in predicting the Lorenz time series are the same as those for the NF-ADBEL network (i.e., $\alpha = 0.8$, $\beta = 0.2$, and $\gamma = 0.01$). By recording and analyzing the prediction error in both cases, it is found that the transient period is less than $5s$, and therefore the steady-state starting index is taken as $n_s = 5$.

The performance of the prediction results of the ADBEL network for the Lorenz time series perform are given in Figures 4.17, 4.19 and 4.20. A zoomed view of the prediction error as returned by both networks in steady-state is shown in Figures 4.22 and 4.24. The figures show that the proposed NF-ADBEL network has a lower error in predicting the Lorenz time series compared to the existing ADBEL network. The prediction performance in both cases is also analyzed in terms of root mean squared error in Eq. (4.1). It is also shown in Figures 4.14 and 4.19. The correlation coefficient for the Eq. (4.2) criteria is shown in Figures 4.15 and 4.20.

Table 4.2: RMSE/COR/PI FOR LORENZ X-TIME SERIES PREDICTION BY ADBEL AND NF-ADBEL NETWORKS

Time Series	Prediction Network	RMSE	COR	PI(%)
Lorenz	ADBEL	0.55635	0.99827	
	NF-ADBEL	0.1726	0.99976	**68.98**

The results for this analysis are included in Table 4.2 and show the superior performance of the NF-ADBEL network due to the lowered root mean squared error, higher correlation coefficient, and significant percentage improvement offered by this network.

4.1.3 Rossler Chaotic Time Series Predicted by the Proposed NF-ADBEL Network

We also used the neo-fuzzy-integrated ADBEL network to predict the Rossler chaotic time series. It has been used in the literature to evaluate the performance of prediction algorithms [76],[77]. The time series is generated through the following differential equations [78], with the constants selected to be $a = 0.15$, $b = 0.2$, and $c = 10$:

$$\begin{cases} \dot{x}(t) = -y(t) - z(t) \ , \\ \dot{y}(t) = x(t) + ay(t) \ , \\ \dot{z}(t) = b + z(t)(x(t) - c) \end{cases} \tag{4.6}$$

A total of $n_e = 8188$ samples are generated for the Rossler time series, as shown in Appendix A. To simulate the proposed NF-ADBEL network for predicting this time series, the learning parameters are selected to be $\alpha = 0.5$, $\beta = 0.5$, and $\gamma = 0.08$. The zoomed view of the predicted Rossler time series in steady-state is shown in Figures 4.26, 4.28 and 4.29.

Table 4.3: RMSE /COR/PI FOR ROSSLER TIME SERIES PREDICTION BY ADBEL AND NF-ADBEL NETWORKS

Time Series	Prediction Network	RMSE	COR	PI(%)
Rossler	ADBEL	1.5014	0.99224	
	NF-ADBEL	0.2999	0.99929	**80.03**

The ADBEL network that is driven by the parameters $\alpha = 0.8$, $\beta = 0.2$, and $\gamma = 0.05$ is also simulated to predict the Rossler time series in steady-state. It is shown in Figures 4.31, 4.33 and 4.34.

A comparison of the two networks in terms of the prediction error is displayed in Figures 4.36 and 4.38. Note that the prediction error for both networks is shown for a finite duration in steady-state. The transient period happens to be the same as in the case of the Mackey-Glass and Lorenz time series, i.e., $n_s = 5s$. Furthermore, it can be seen that the amplitude of error signal for the NF-ADBEL network is lower

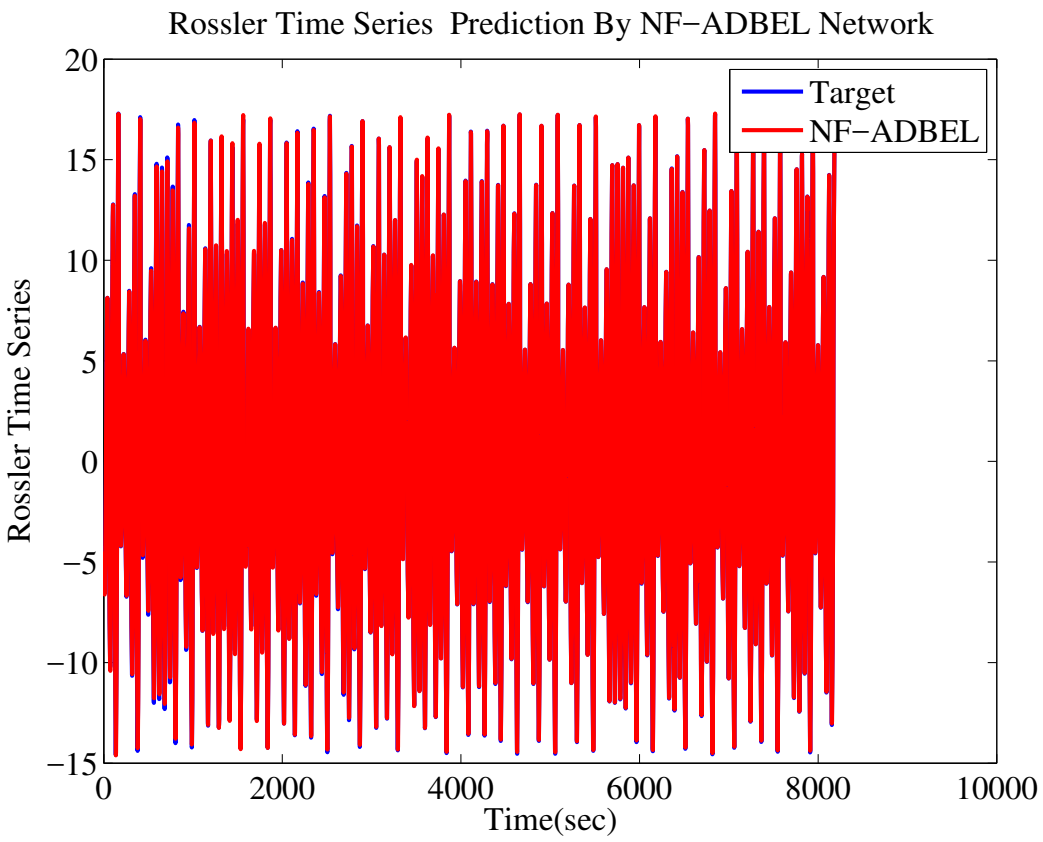

Figure 4.25: Rossler Time Series as Predicted by NF-ADBEL Network .

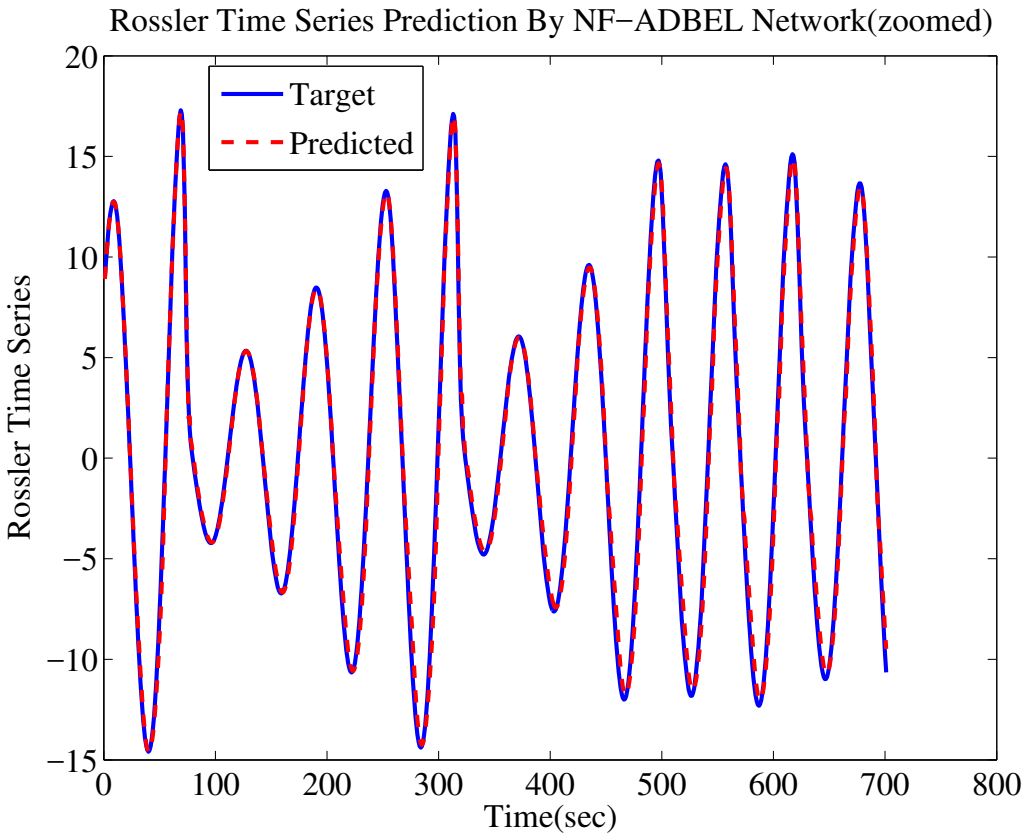

Figure 4.26: Rossler Time Series as Predicted by NF-ADBEL Network (displayed portion of Fig 4.25).

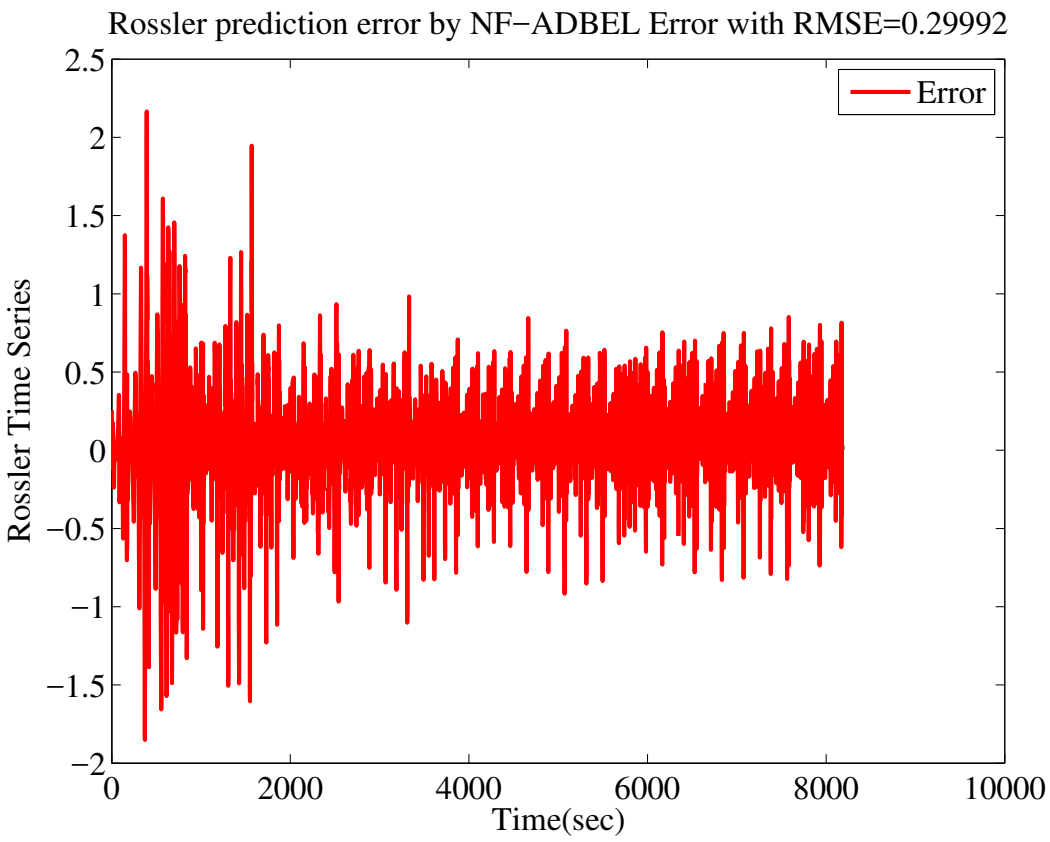

Figure 4.27: Error in Predicting Rossler Time Series by NF-ADBEL Network.

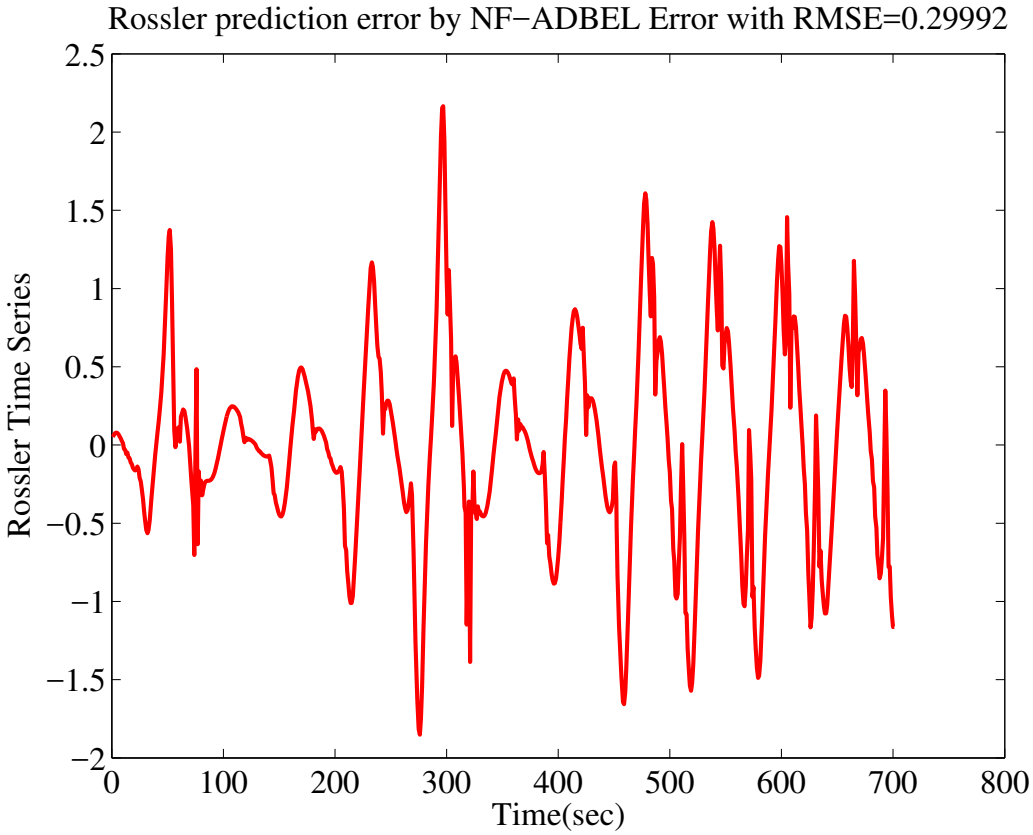

Figure 4.28: Error in Predicting Rossler Time Series by NF-ADBEL Network (displayed a portion of Fig. 4.27).

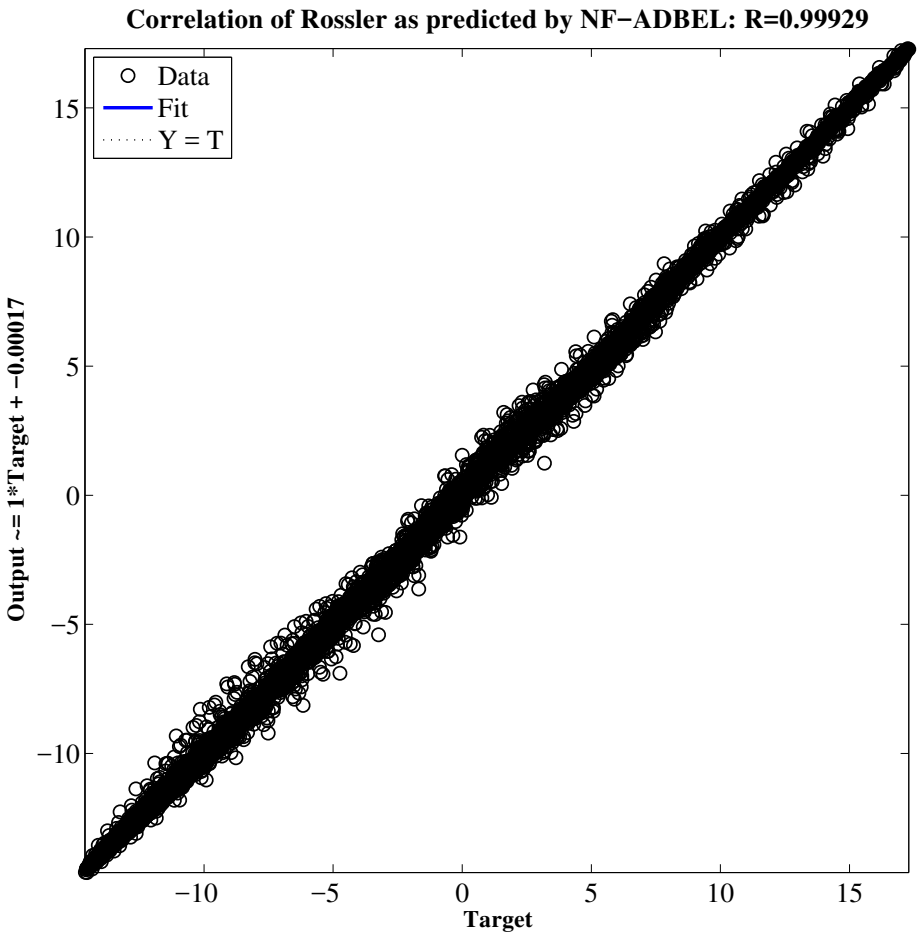

Figure 4.29: Correlation in Predicting Rossler Time Series by NF-ADBEL Network.

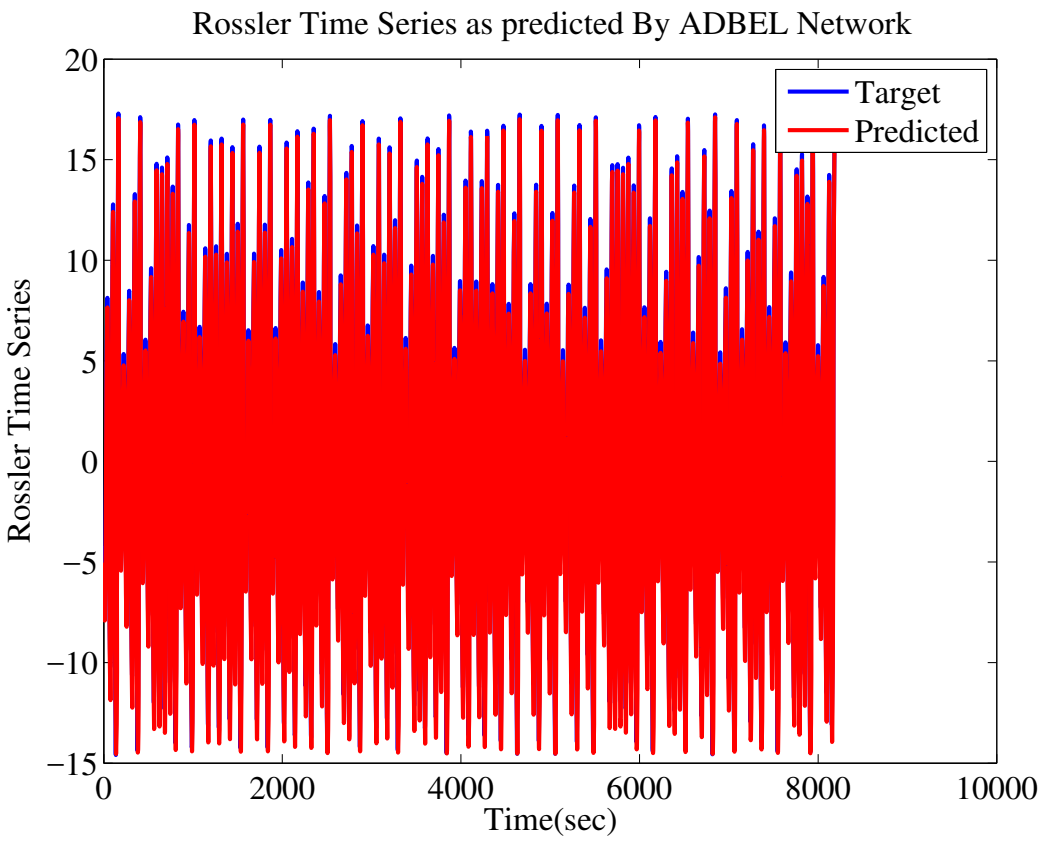

Figure 4.30: Rossler Time Series as Predicted by ADBEL Network.

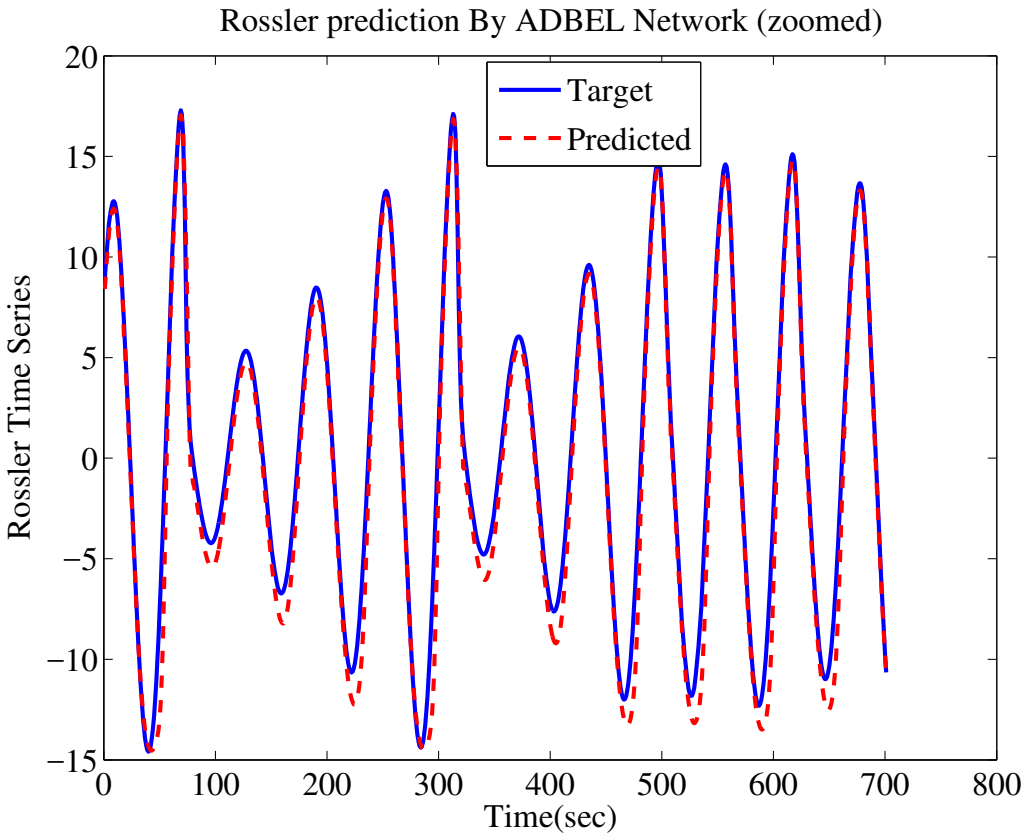

Figure 4.31: Rossler Time Series as Predicted by ADBEL Network(displayed a portion of Fig. 4.30).

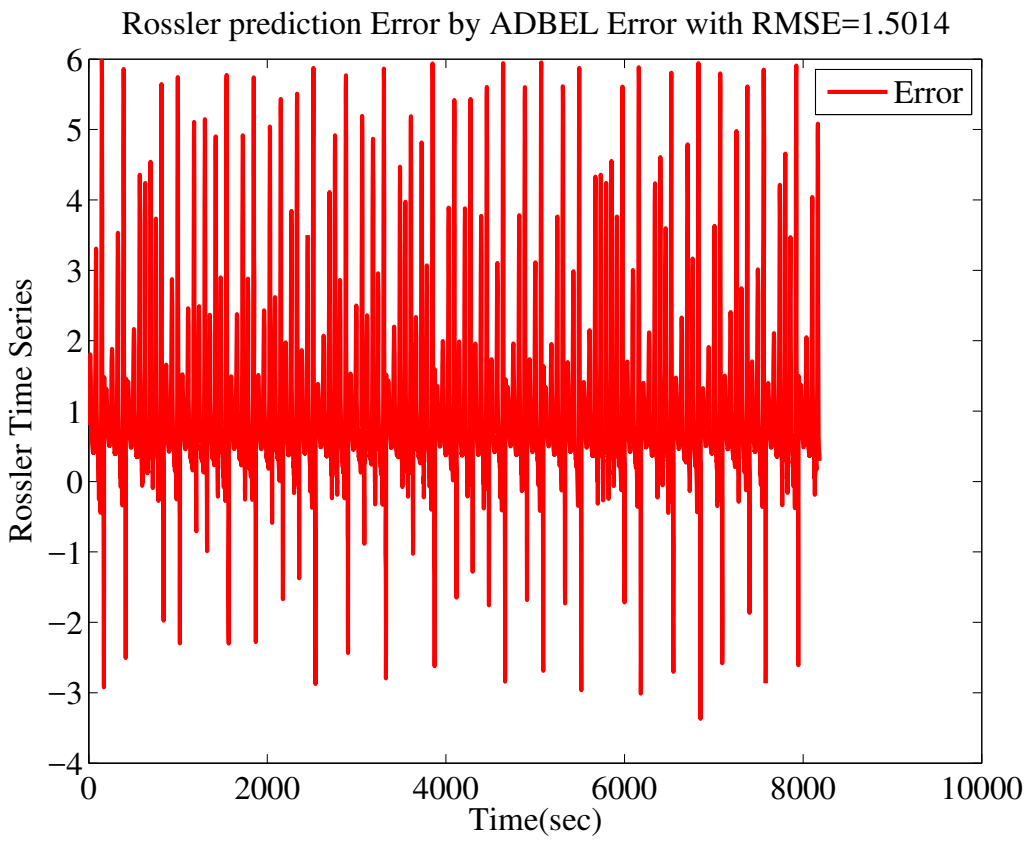

Figure 4.32: Error in Predicting Rossler Time Series by ADBEL Network.

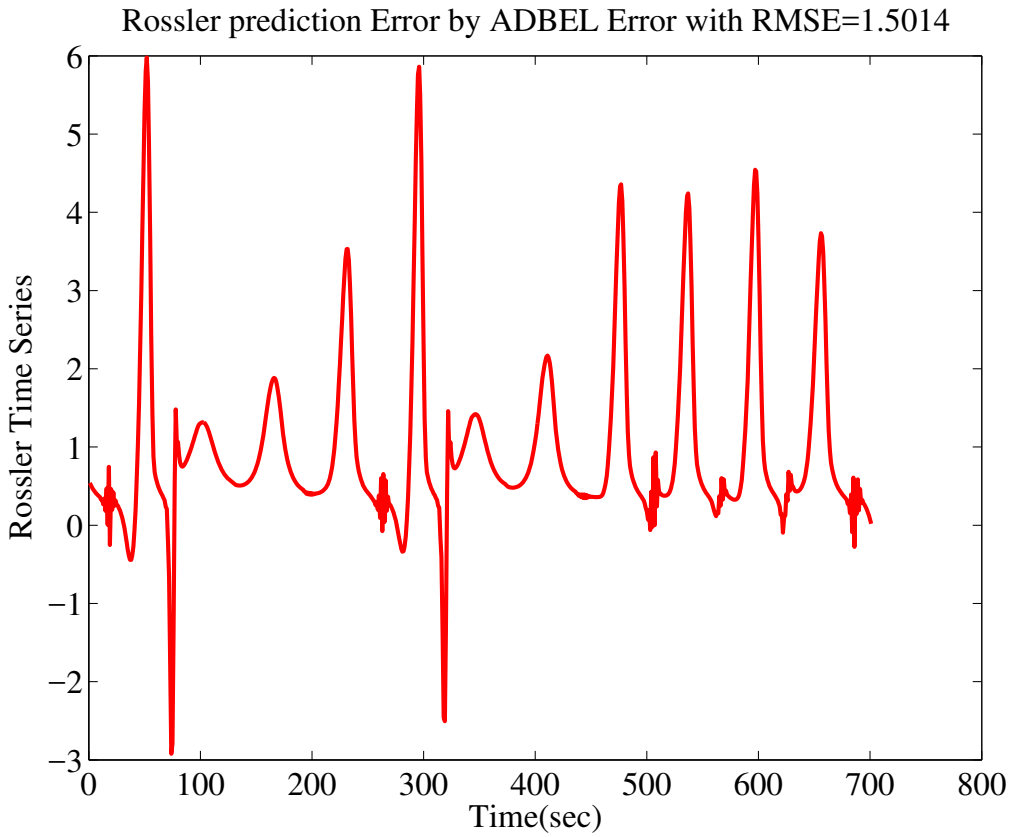

Figure 4.33: Error in Predicting Rossler Time Series by ADBEL Network (displayed a portion of Fig. 4.32).

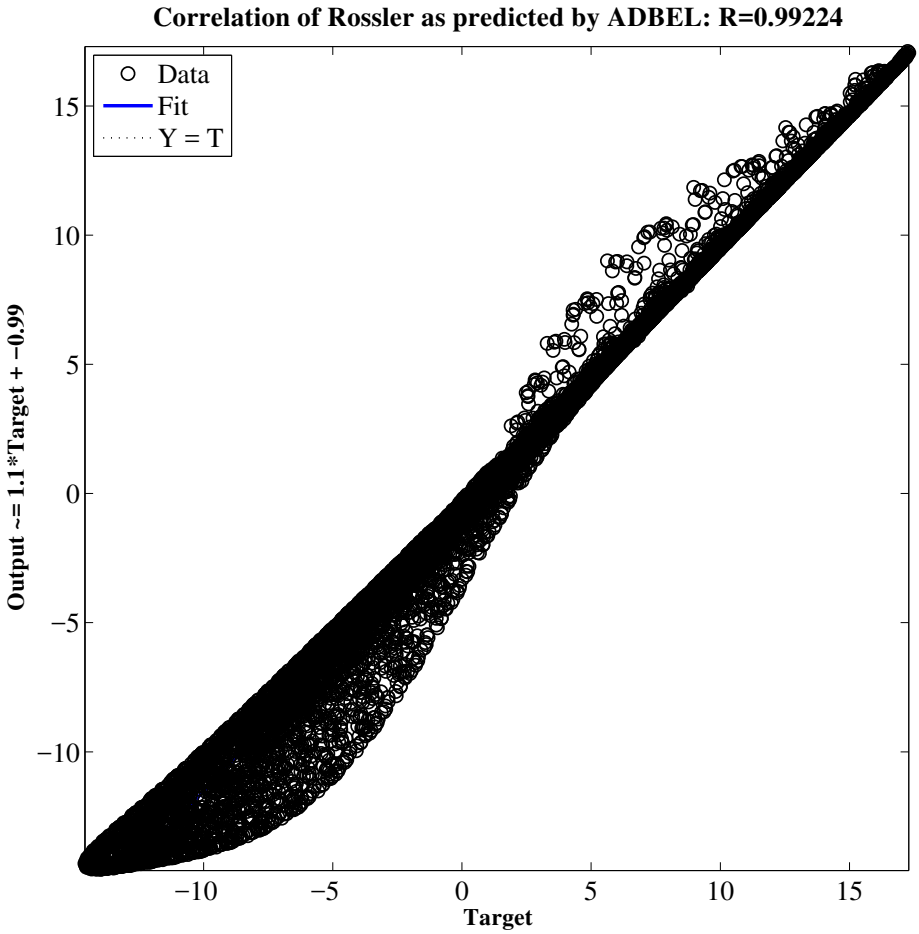

Figure 4.34: Correlation in Predicting Rossler Time Series by ADBEL Network.

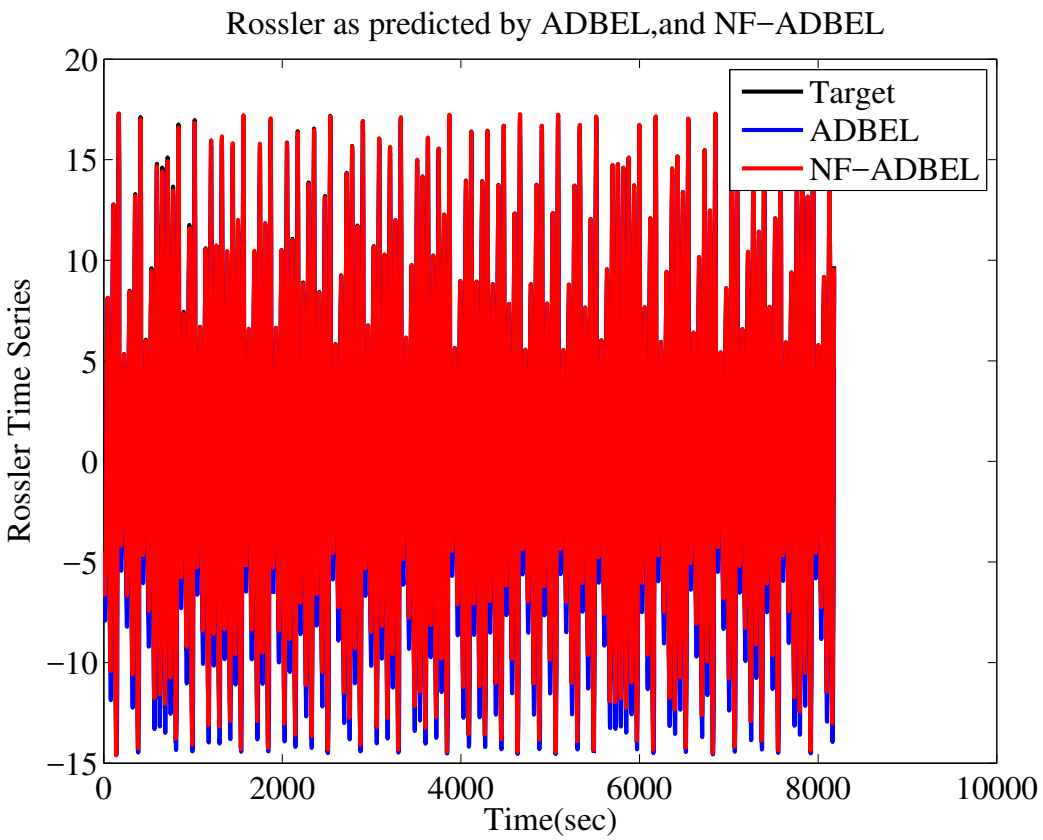

Figure 4.35: Rossler Time Series as Predicted by ADBEL and NF-ADBEL Networks.

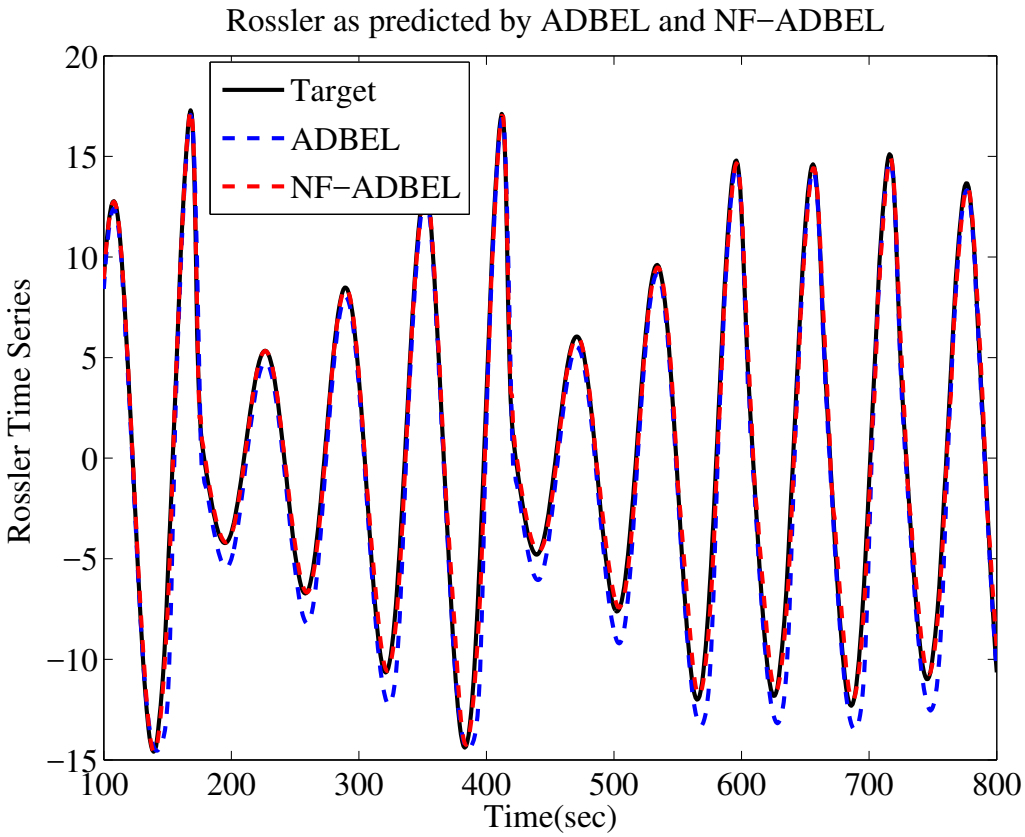

Figure 4.36: Rossler Time Series as Predicted by ADBEL and NF-ADBEL Networks (displayed a portion of Fig. 4.35).

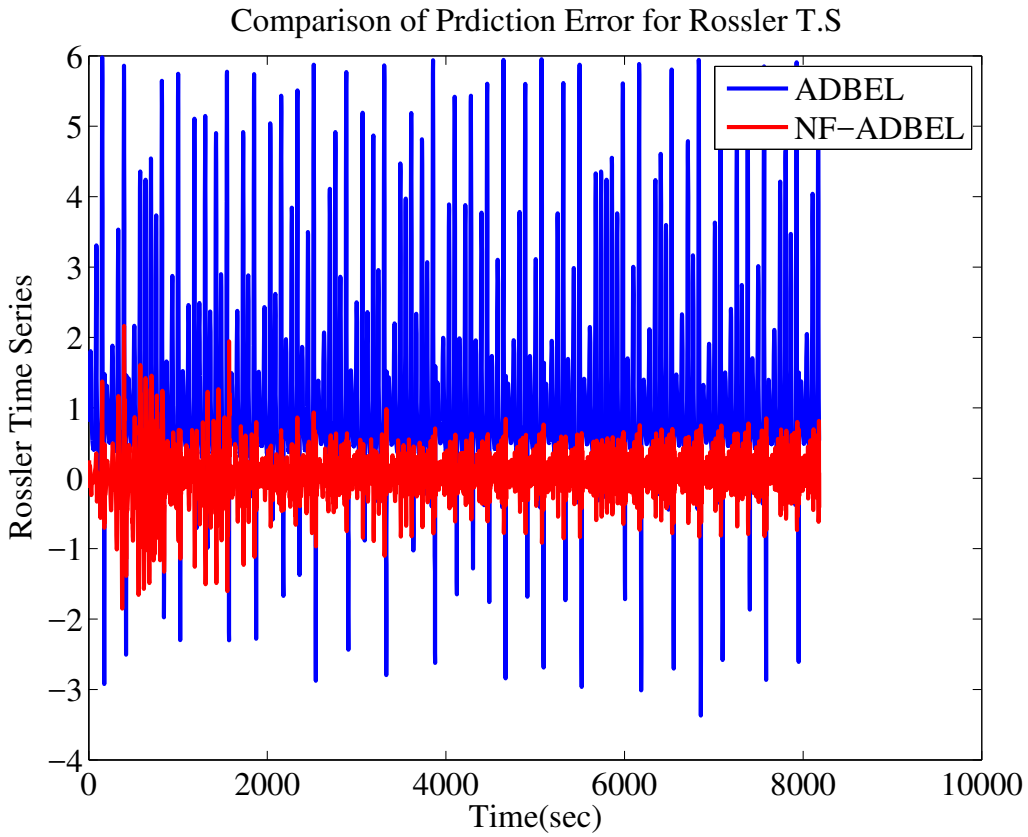

Figure 4.37: Comparison of Errors in Predicting Rossler Time Series by ADBEL and NF-ADBEL Networks.

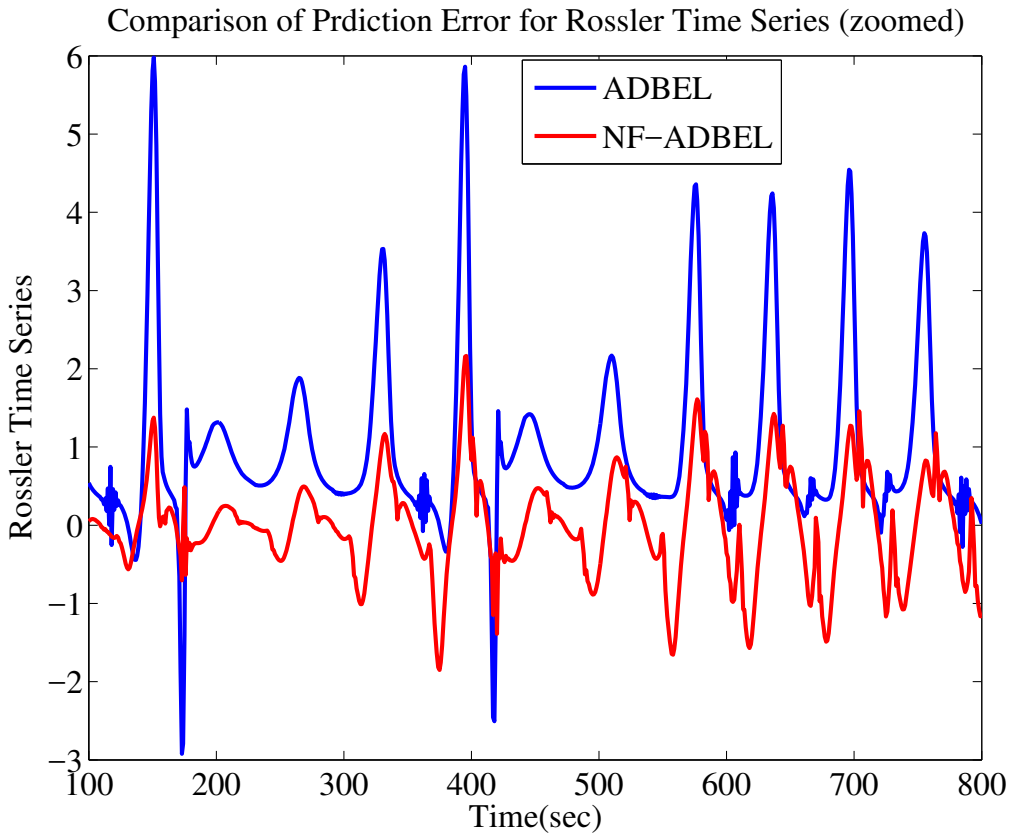

Figure 4.38: Comparison of Errors in Predicting Rossler Time Series by ADBEL and NF-ADBEL Networks(displayed a portion of Fig. 4.37).

compared to the ADBEL network, as shown in Figures 4.36 and 4.38. The figures illustrate the better prediction accuracy of the NF-ADBEL network.

Analysis of the predicted results for the Rossler time series in terms of the root mean squared error and correlation coefficient criteria are shown in Figures 4.29 and 4.34, respectively, and also listed in Table 4.3. As can be seen, they reveal the superiority of the NF-ADBEL over the ADBEL network. Finally, a reasonable amount of percentage improvement is yielded by the NF-ADBEL network to predict the Rossler time series, as shown in Table 4.3.

4.1.4 Disturbance Storm Time Index (D_{st}) Predicted by the Proposed NF-ADBEL Network

Precise forecasting of space weather, especially solar storms, has become increasingly urgent because of the destructive effects these storms can have on infrastructures such as satellites, telecommunication, and power grids [49]. In 1989, a solar storm hit North America that caused severe damage to power plants, networks, power stations, telecommunications, and space-based communications devices. In particular, the event caused significant damage to Quebec, Canada, by paralyzing and disrupting the electrical grid. Therefore, we must be aware of the likelihood for these types of natural disasters to occur again.

Several studies have already predicted solar storms and other space-related terrestrial disasters. Among these studies, ANN and ANFIS were found to be reasonably able to predict the occurrence and strength of solar storms [79],[80]. Space weather forecasting is one of our main incentives for providing a novelty version of a brain emotional learning model in a standardized and developed version to satisfy the need for accurate prediction.

The neo-fuzzy integrated ADBEL network is proposed to predict the disturbance storm time index, which is an hourly indicator of geomagnetic storms. This index's negative values are vital, as they indicate the weakening of Earth's magnetic field. Such an event can lead to geomagnetic storms, which can disrupt radio communications, damage satellites, and cause power system outages, all of which were seen in the Hydro-Qubec transmission grid during the 1989 storm. In March 1989, the province was plunged into darkness for more than nine hours [81]. Several models based on

differential equations and neural networks have been proposed in the literature for predicting the disturbance storm time index [82],[83],[84],[85].

Recently, the ADBEL network also proposed predicting this important index [4], which has been modified in the present work to yield a new NF-ADBEL network [58]. Here, we simulate the NF-ADBEL network to predict the disturbance storm time index D_{st} time series for a few months when considerable geomagnetic activity was observed. The data for these months, as seen in Appendix A, have been downloaded from the website World Data Center (WDC) in [86], "WDC for Geomagnetism, Kyoto." With the learning parameters set as $\alpha = 0.3$, $\beta = 0.3$ and $\gamma = 0.01$, the NF-ADBEL network is deployed to predict the D_{st} index for April 2000, July 2000, March 2001, October 2003, and July 2004. The number of samples is $n_e = 716$ for April 2000 and $n_e = 740$ for all the other months. Please note that the above parameters ($\alpha = 0.3$, $\beta = 0.3$ and $\gamma = 0.01$) are kept the same for the different D_{st} index data for the mentioned months.

The predicted results provided by the NF-ADBEL network in all these cases are shown in the following figures:

- Figures 4.39, 4.40 and 4.41 show the results of the proposed NF-ADBEL networks in terms of D_{st} index data for April 2000.
- Figures 4.47, 4.48 and 4.49 show the results of the proposed NF-ADBEL networks in terms of D_{st} index data for July 2000.
- Figures 4.55, 4.56 and 4.57 show the results of the proposed NF-ADBEL networks in terms of D_{st} index data for March 2001.
- Figures 4.63, 4.64 and 4.65 show the results of the proposed NF-ADBEL networks in terms of D_{st} index data for October 2003.
- Figures 4.71, 4.72 and 4.73 show the results of the proposed NF-ADBEL networks in terms of D_{st} index data for July 2004.

The transient period of the NF-ADBEL network for all the reported cases is found to be $n_s = 10\ hrs$, which then becomes the steady-state starting index. It can be observed that, despite the high initial transients, the NF-ADBEL network can follow

the D_{st} time series in steady-state and the important valley points are also well-predicted, which points towards the possible occurrence of geomagnetic storms. An existing ADBEL network is used to predict the D_{st} time series. For this purpose, the learning parameters of the ADBEL network are assigned the values of $\alpha = 0.8$, $\beta = 0.2$, and $\gamma = 0.01$. The predicted results provided by the ADBEL network in all these cases are shown in the following figures:

- Figures 4.42, 4.43 and 4.44 show the results of the ADBEL networks in terms of D_{st} index data for April 2000.
- Figures 4.50, 4.51 and 4.52 show the results of the ADBEL networks in terms of D_{st} index data for July 2000.
- Figures 4.58, 4.59 and 4.60 show the results of the ADBEL networks in terms of D_{st} index data for March 2001.
- Figures 4.66, 4.67 and 4.68 show the results of the ADBEL networks in terms of D_{st} index data for October 2003.
- Figures 4.74, 4.75 and 4.76 show the results of the ADBEL networks in terms of D_{st} index data for July 2004.

The results of this comparison for all these cases are displayed in figures below:

- Figures 4.45 and 4.46 show the results of the proposed NF-ADBEL networks compared to ADBEL outcomes in terms of D_{st} index data for April 2000.
- Figures 4.53 and 4.54 show the results of the proposed NF-ADBEL networks compared to ADBEL outcomes in terms of D_{st} index data for July 2000.
- Figures 4.61 and 4.62 show the results of the proposed NF-ADBEL networks compared to ADBEL outcomes in terms of D_{st} index data for March, 2001.
- Figures 4.69 and 4.70 show the results of the proposed NF-ADBEL networks compared to ADBEL outcomes in terms of D_{st} index data for October 2003.
- Figures 4.77 and 4.78 show the results of the proposed NF-ADBEL networks compared to ADBEL outcomes in terms of D_{st} index data for July 2004.

It can be observed that the ADBEL network has a long transient period compared to the NF-ADBEL network, which is found to be 110 hrs and acts as the steady-state starting index for the ADBEL-based D_{st} prediction analysis, i.e., $n_s = 110$ hrs.

It can also be seen that the initial transients in the ADBEL network have high amplitudes compared to the NF- ADBEL network. Further, the response of the NF-ADBEL network in steady-state is also better than the ADBEL network. This is supported by the lower root mean squared error and the higher correlation coefficient obtained by the NF-ADBEL network for predicting the D_{st} time series.

Table 4.4 lists performance indices for both the ADBEL and NF-ADBEL networks. As can be seen, a fair amount of percentage improvement can be obtained by deploying the proposed NF-ADBEL network in predicting the D_{st} time series.

Table 4.4: RMSE/COR/PI FOR Disturbance Storm D_{st} Time SERIES PREDICTION BY ADBEL AND NF-ADBEL NETWORKS.

Time Series	Prediction Network	RMSE	COR	PI(%)
D_{st} (Apr 2000)	ADBEL	14.8037	0.94155	
	NF-ADBEL	9.055	0.9706	**38.75**
D_{st} (Jul 2000)	ADBEL	21.20552	0.90623	
	NF-ADBEL	11.003	0.96731	**48.12**
D_{st} (Mar 2001)	ADBEL	21.3536	0.92396	
	NF-ADBEL	13.0485	0.96437	**38.92**
D_{st} (Oct 2003)	ADBEL	26.9453	0.8998	
	NF-ADBEL	13.98	0.96724	**48.10**
D_{st} (Jul 2004)	ADBEL	14.0643	0.95254	
	NF-ADBEL	7.205	0.9821	**48.79**

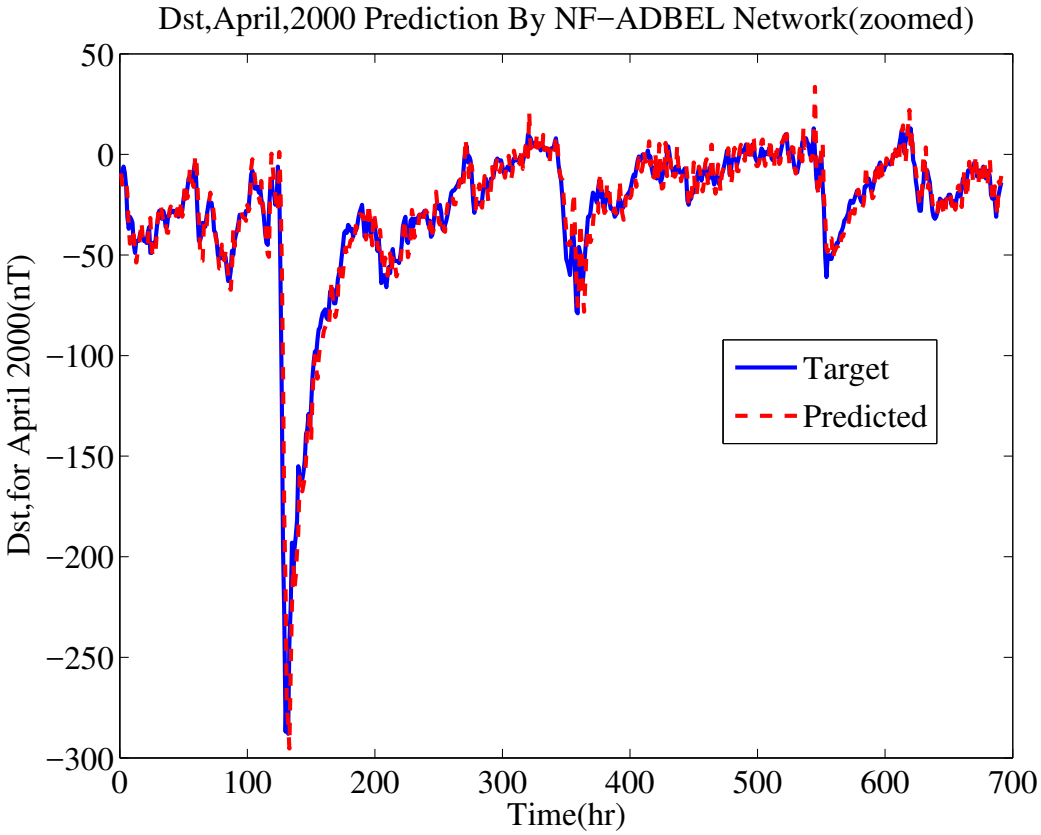

Figure 4.39: Disturbance Storm Time Index D_{st} for April 2000 as Predicted by NF-ADBEL Network.

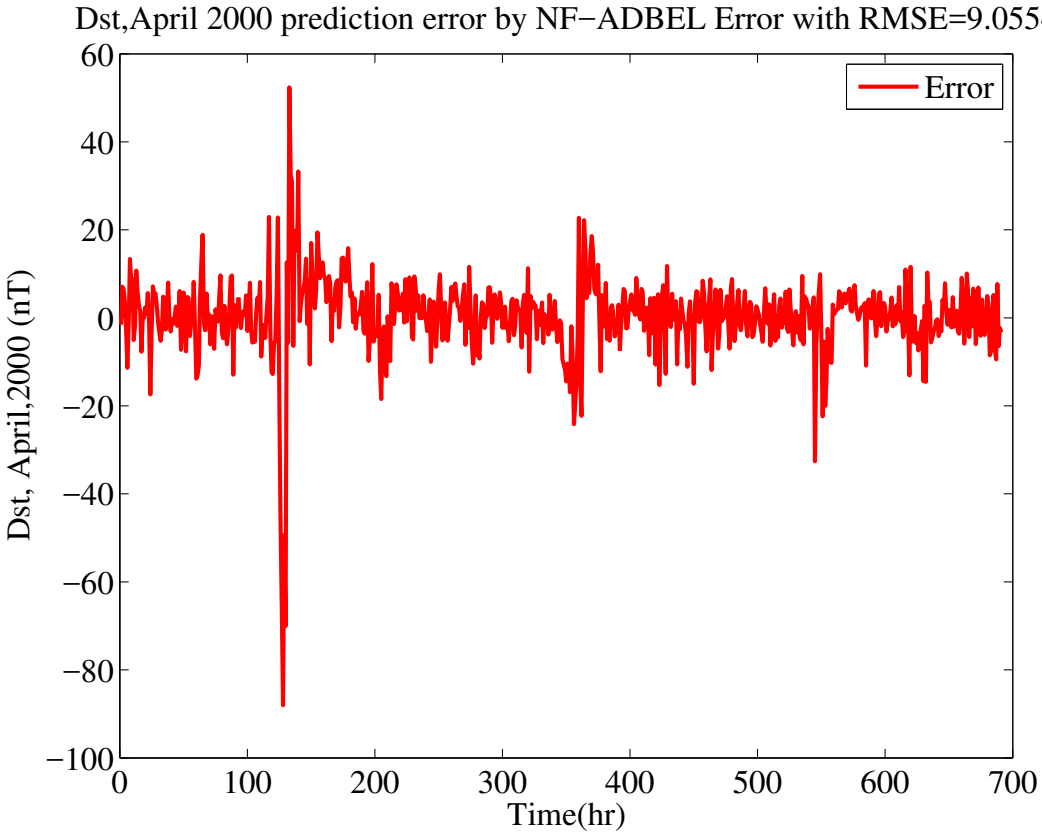

Figure 4.40: Error in Predicting Disturbance Storm Time Index D_{st} for April 2000 by NF-ADBEL Network.

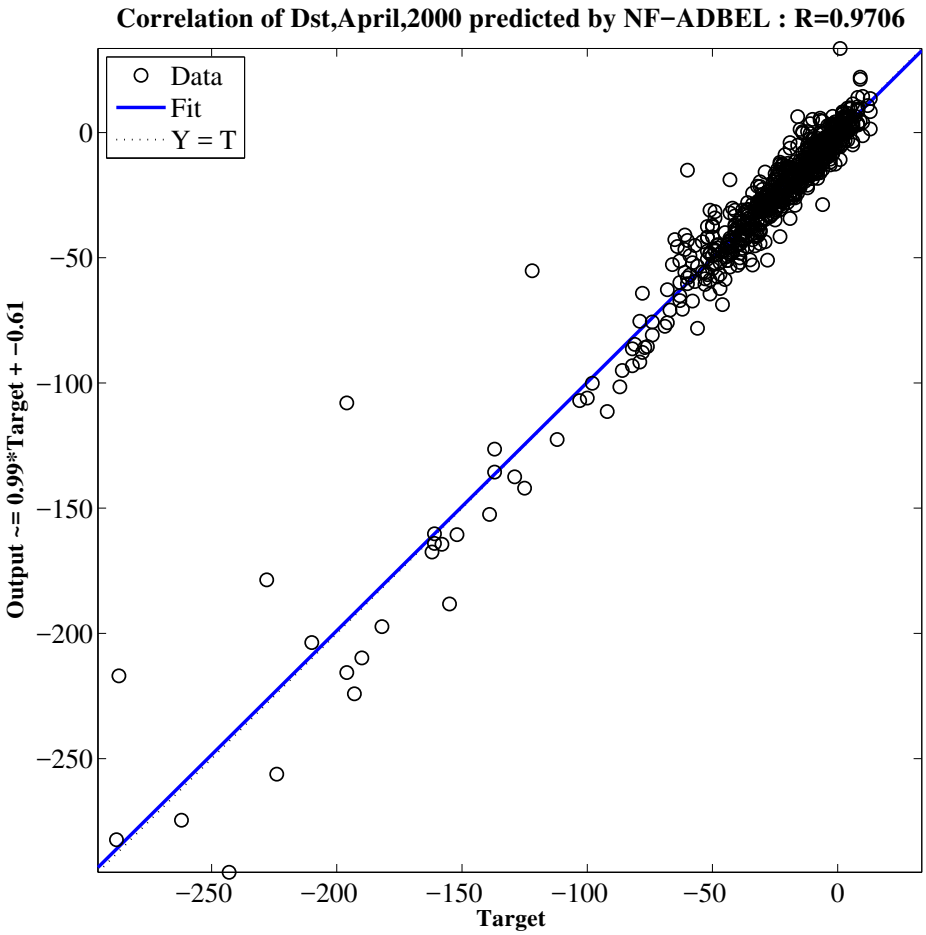

Figure 4.41: Correlation in Predicting Disturbance Storm Time Index D_{st} for April 2000 by NF-ADBEL Network.

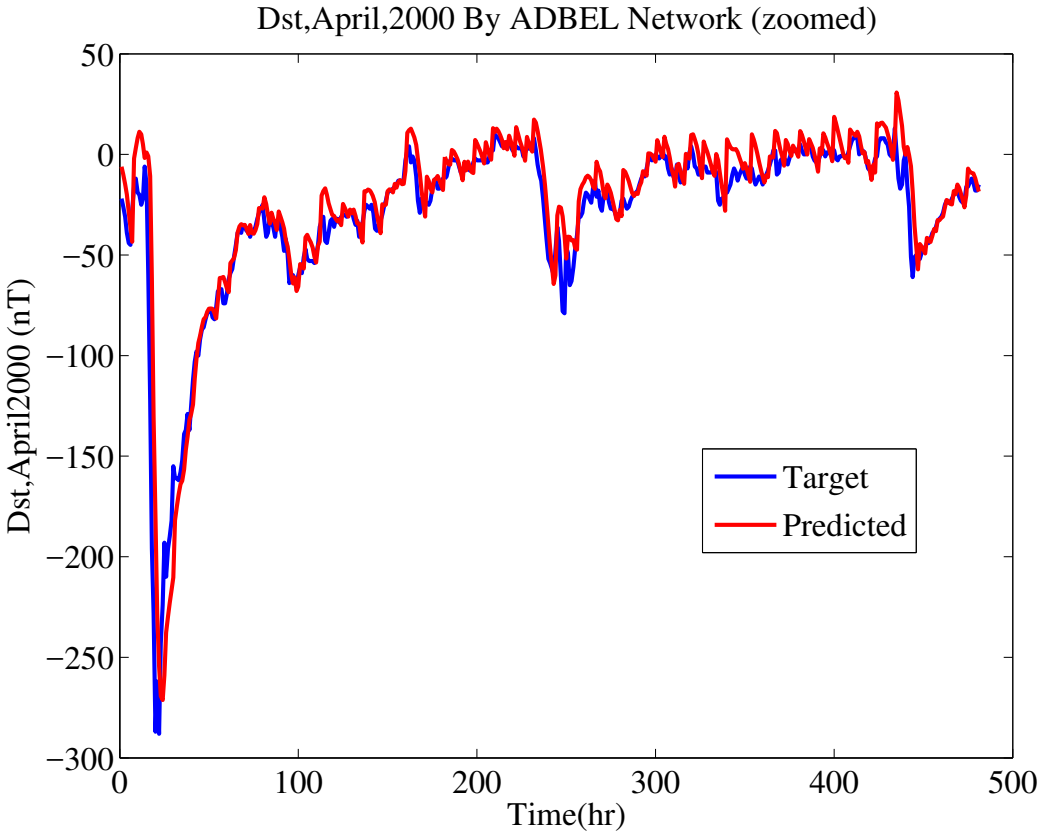

Figure 4.42: Disturbance Storm Time Index D_{st} for April 2000 as Predicted by ADBEL Network.

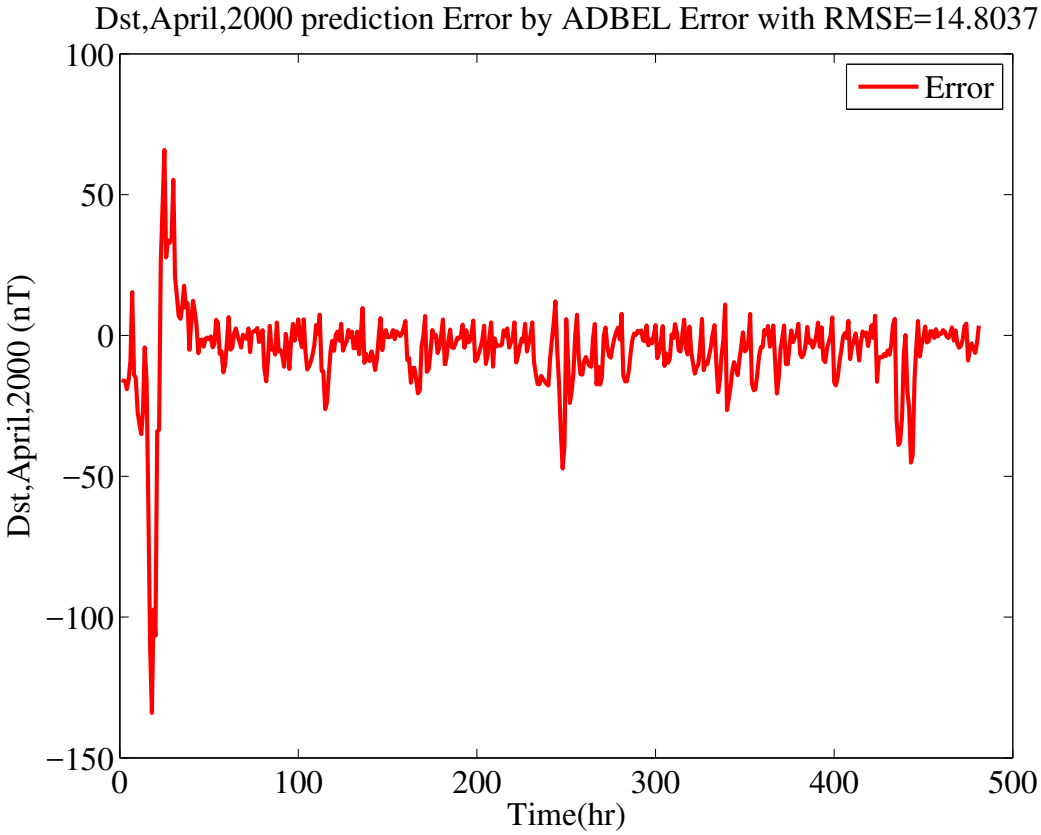

Figure 4.43: Error in Predicting Disturbance Storm Time Index D_{st} for April 2000 by ADBEL Network.

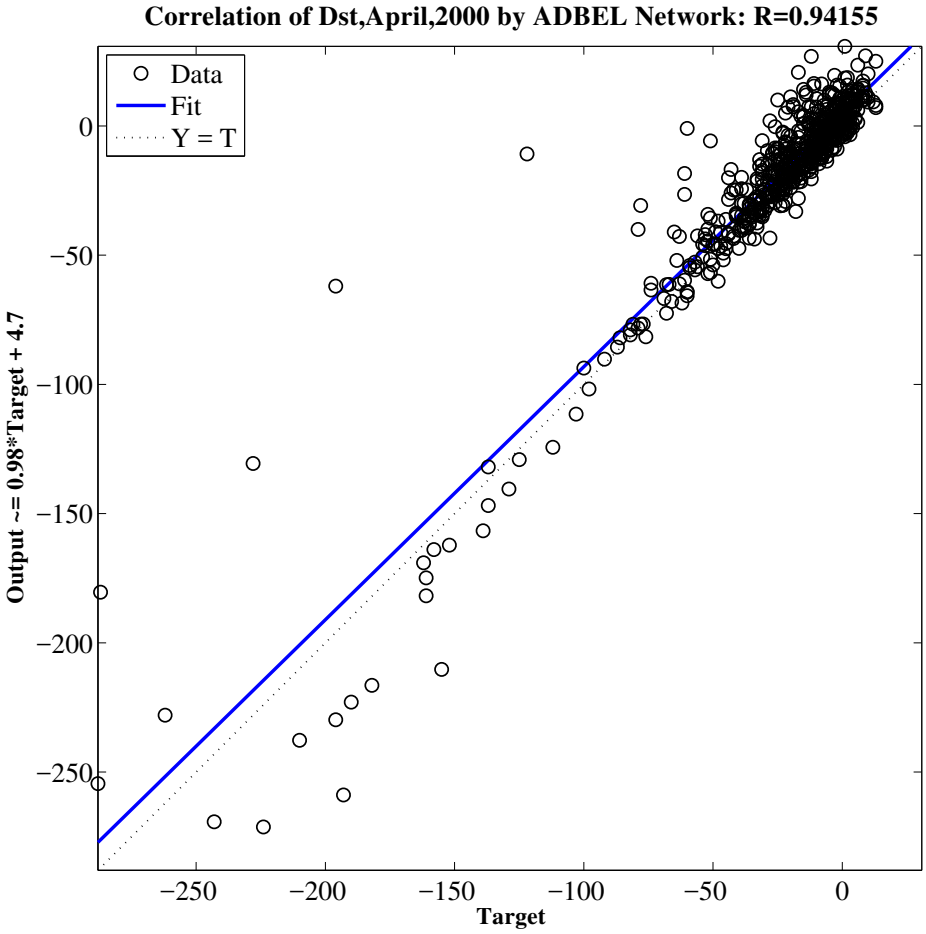

Figure 4.44: Correlation in Predicting Disturbance Storm Time Index D_{st} for April 2000 by ADBEL Network.

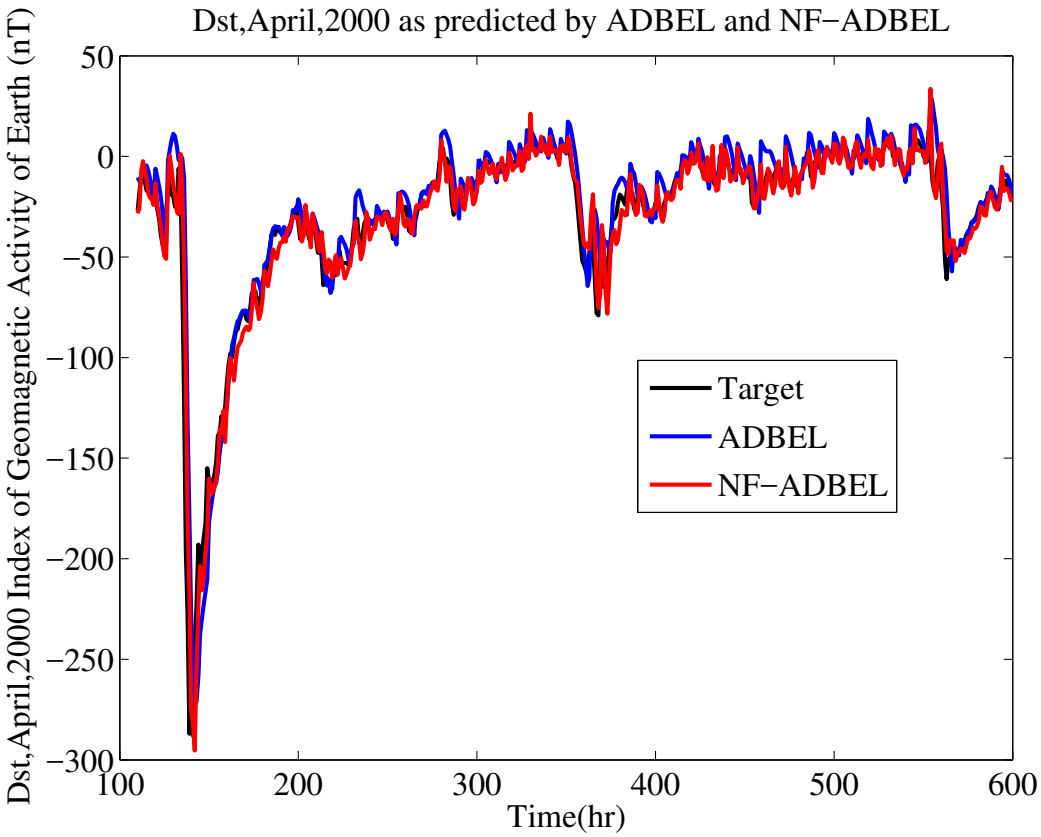

Figure 4.45: Disturbance Storm Time Index D_{st} for April 2000 as Predicted by ADBEL and NF-ADBEL Networks.

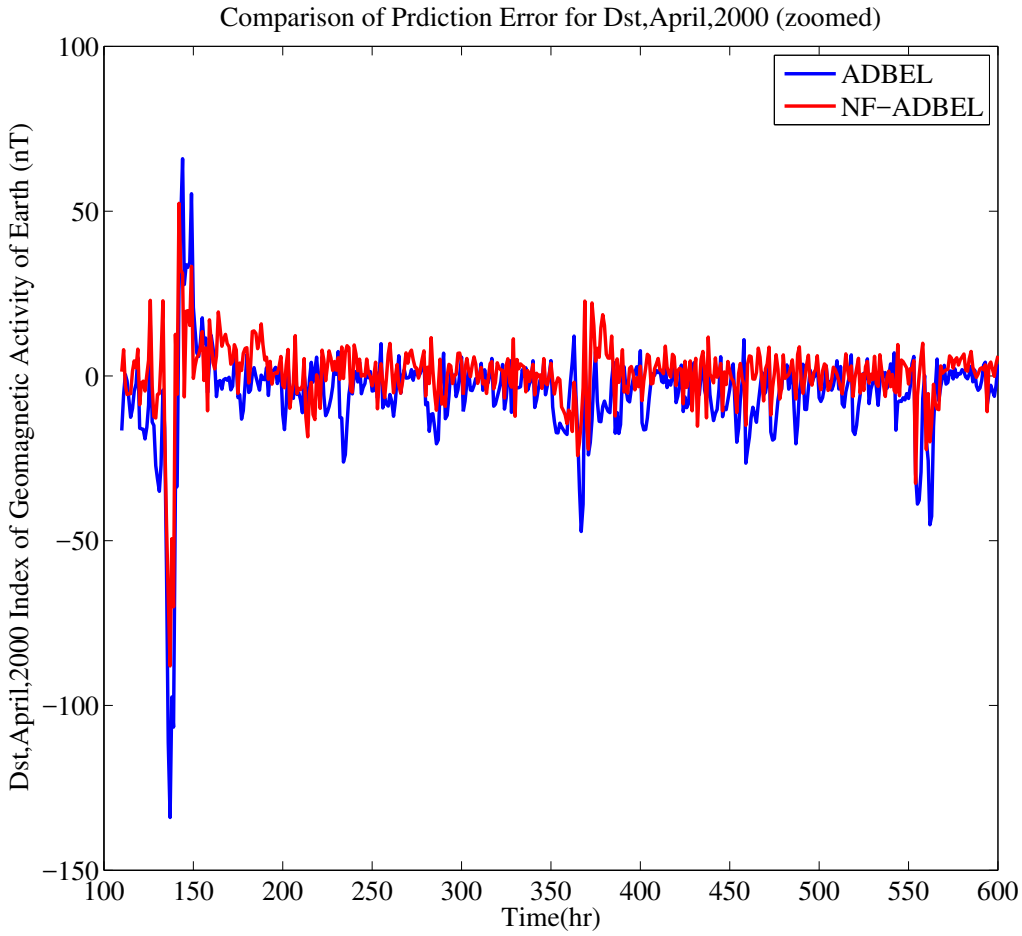

Figure 4.46: Comparison of Errors in Predicting Disturbance Storm Time Index D_{st} for April 2000 by ADBEL and NF-ADBEL Networks.

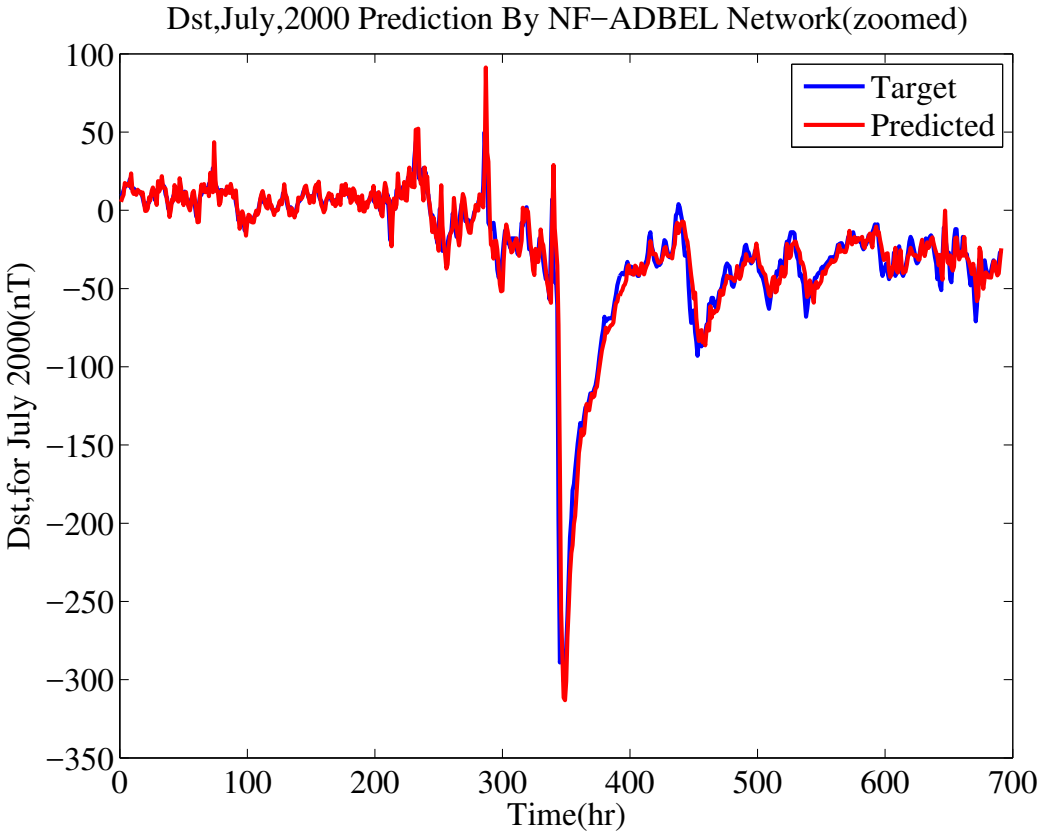

Figure 4.47: Disturbance Storm Time Index D_{st} for July 2000 as Predicted by NF-ADBEL Network.

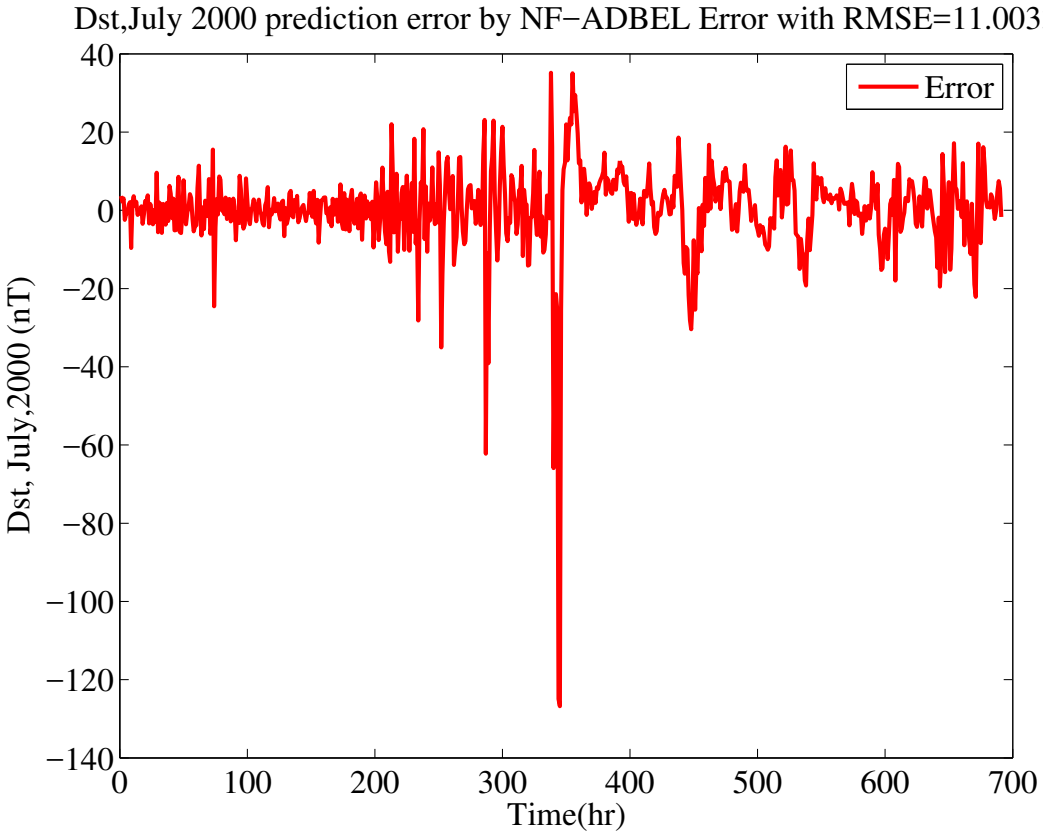

Figure 4.48: Error in Predicting Disturbance Storm Time Index D_{st} for July 2000 by NF-ADBEL Network.

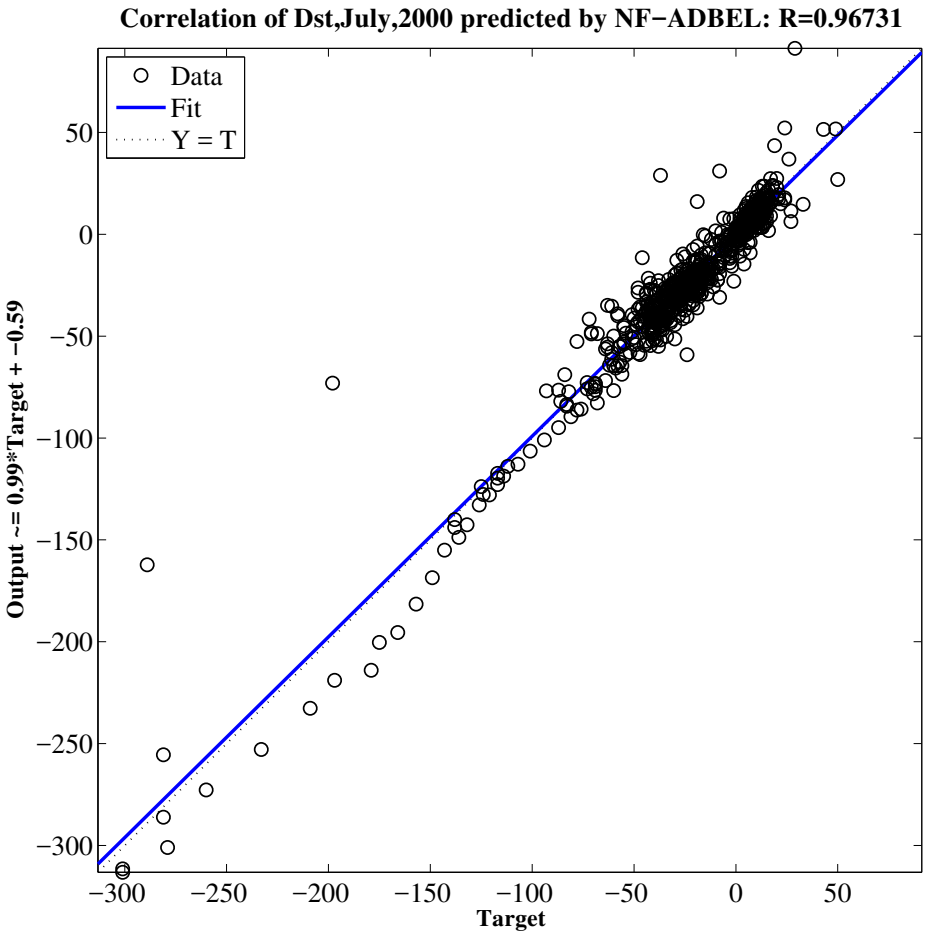

Figure 4.49: Correlation in Predicting Disturbance Storm Time Index D_{st} for July 2000 by NF-ADBEL Network.

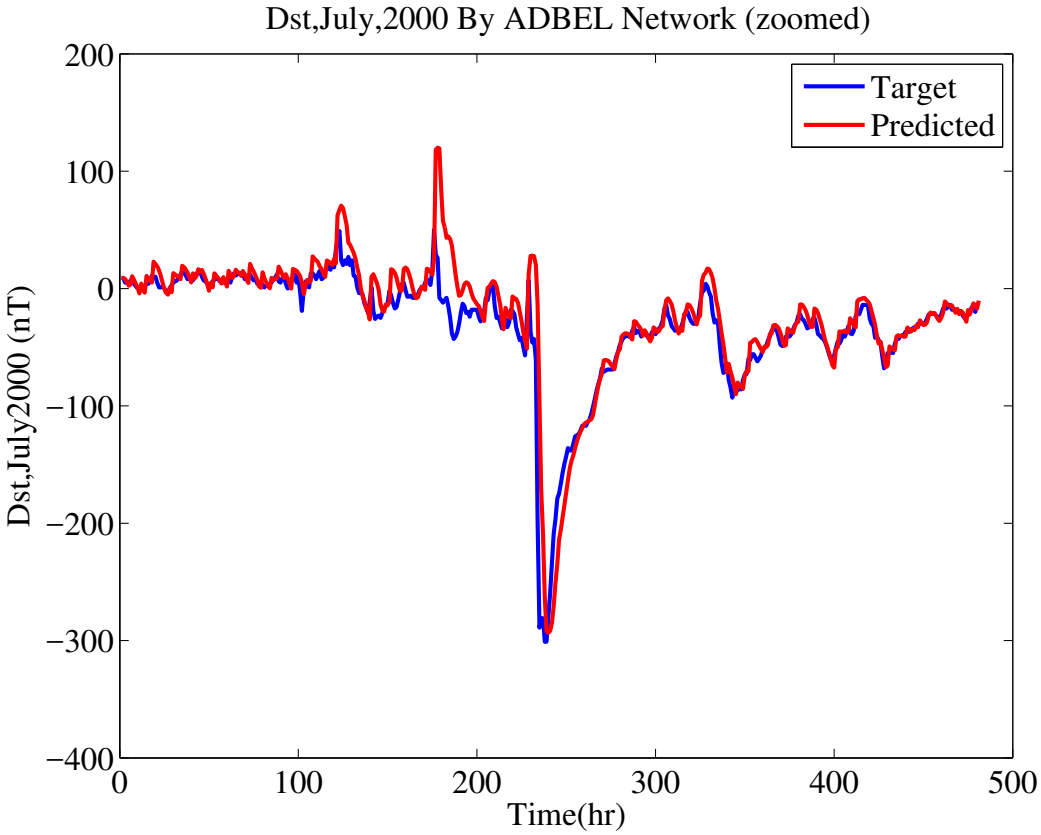

Figure 4.50: Disturbance Storm Time Index D_{st} for July 2000 as Predicted by ADBEL Network.

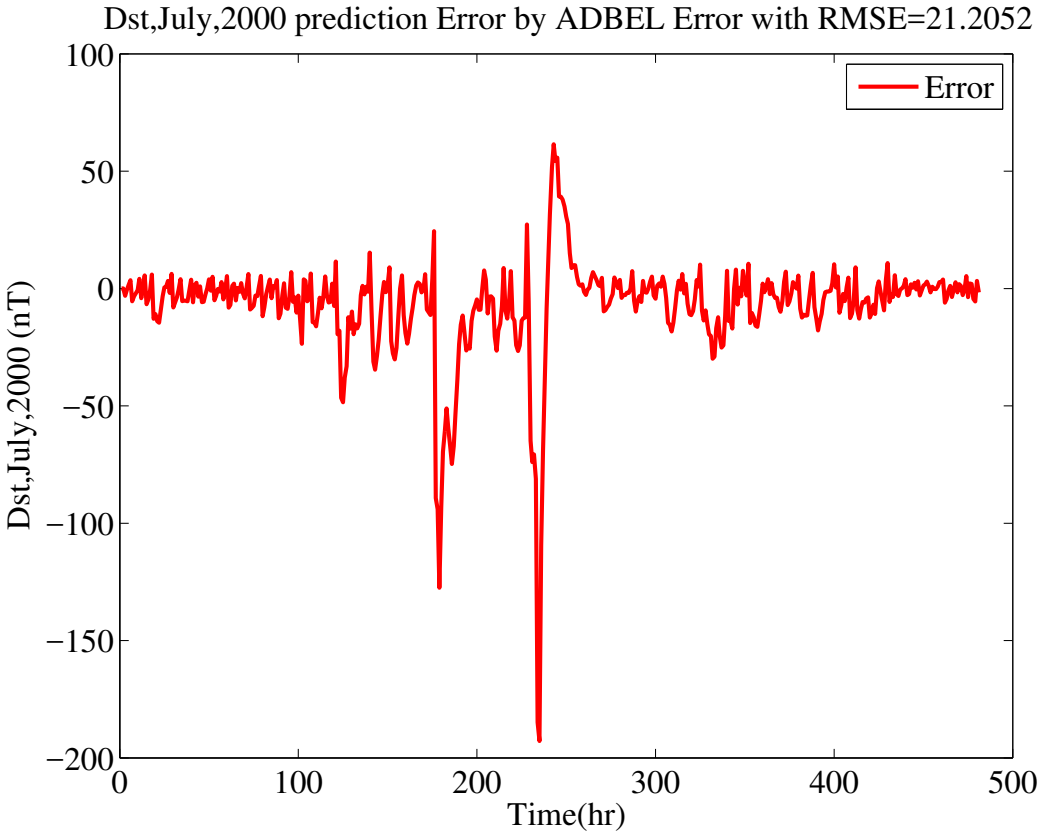

Figure 4.51: Error in Predicting Disturbance Storm Time Index D_{st} for July 2000 by ADBEL Network.

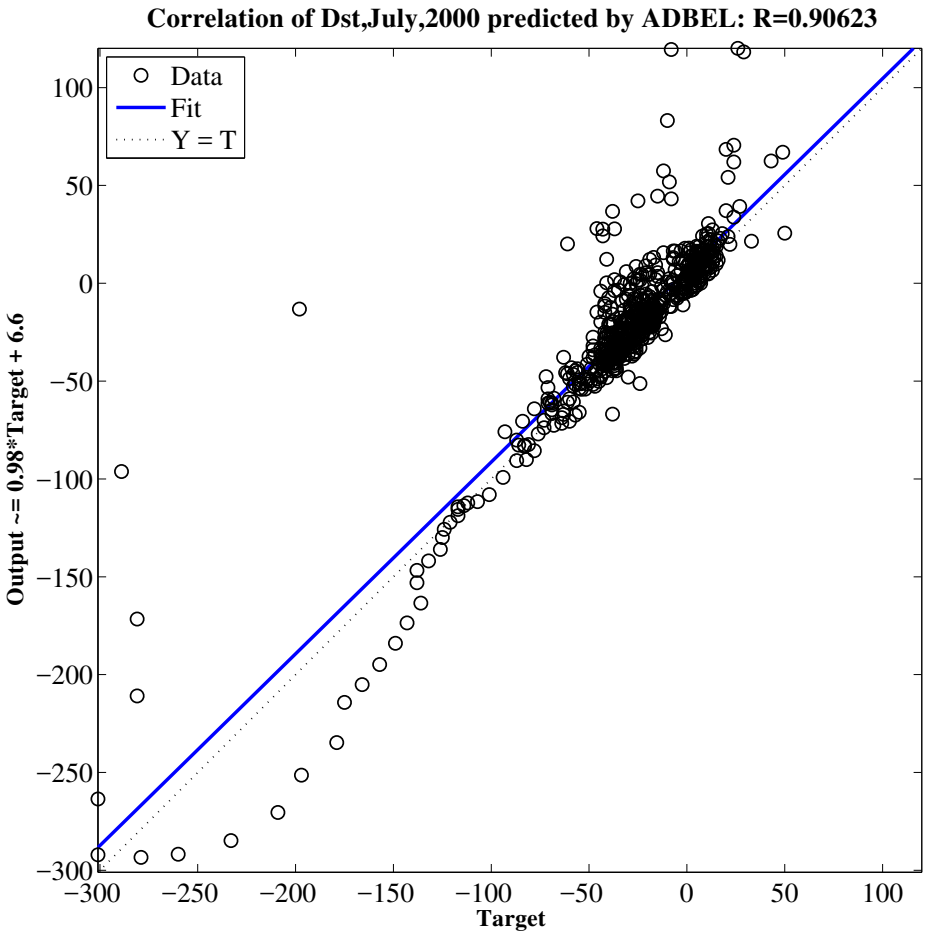

Figure 4.52: Correlation in Predicting Disturbance Storm Time Index D_{st} for July 2000 by ADBEL Network.

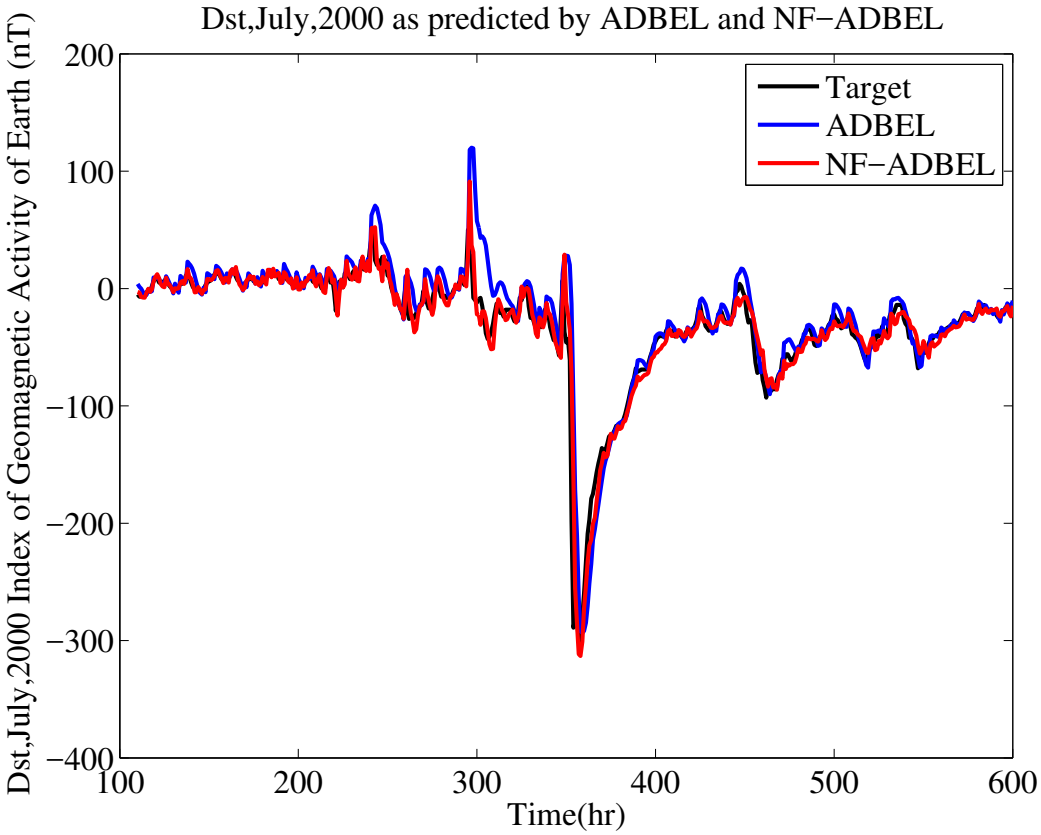

Figure 4.53: Disturbance Storm Time Index D_{st} for July 2000 as Predicted by ADBEL and NF-ADBEL Networks.

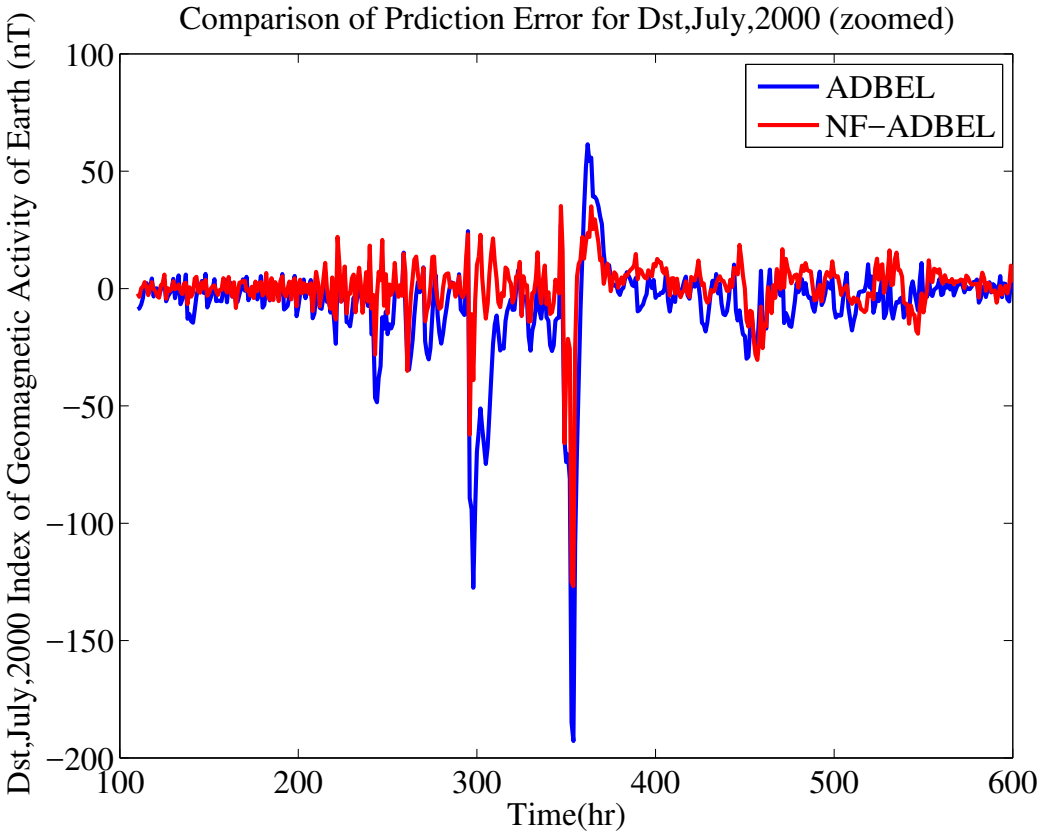

Figure 4.54: Comparison of Errors in Predicting Disturbance Storm Time Index D_{st} for July 2000 by ADBEL and NF-ADBEL Networks.

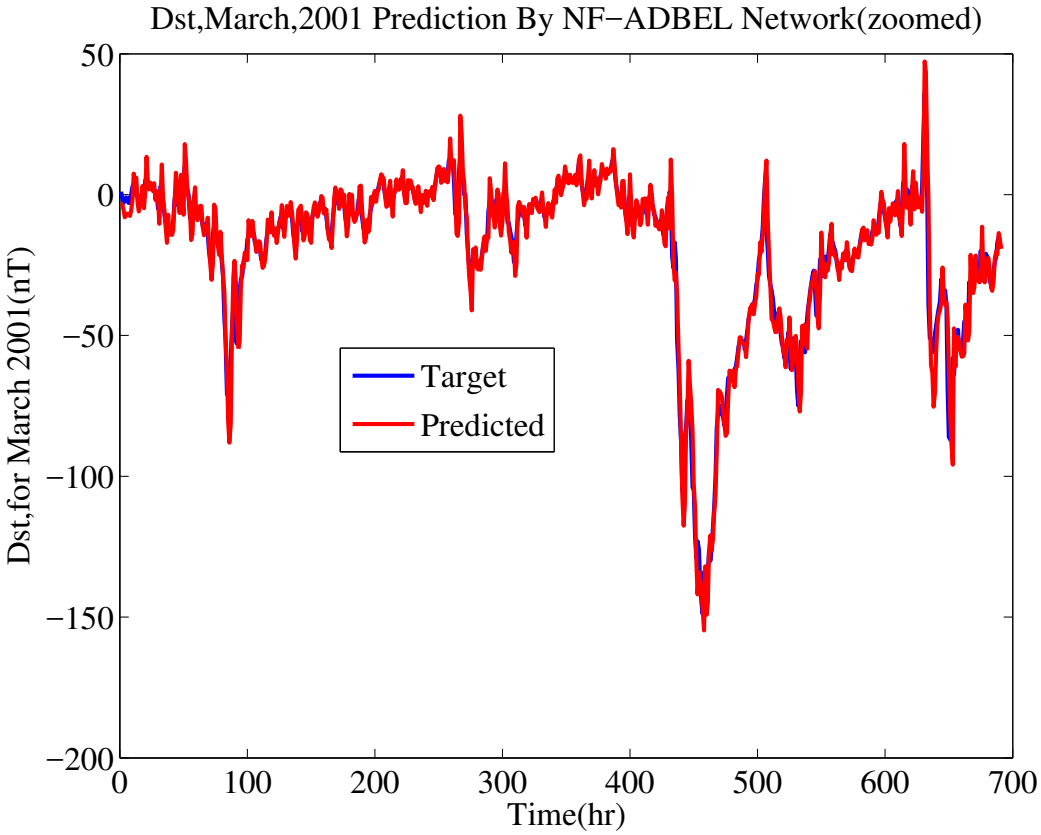

Figure 4.55: Disturbance Storm Time Index D_{st} for March 2001 as Predicted by NF-ADBEL Network.

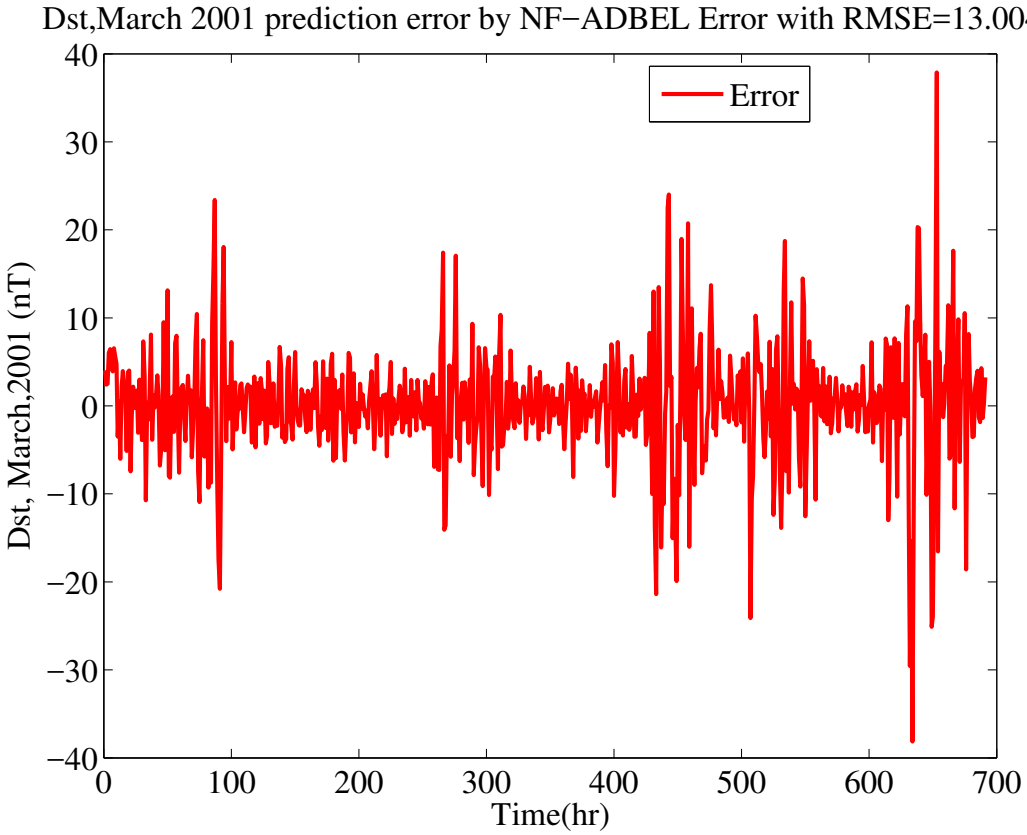

Figure 4.56: Error in Predicting Disturbance Storm Time Index D_{st} for March 2001 by NF-ADBEL Network.

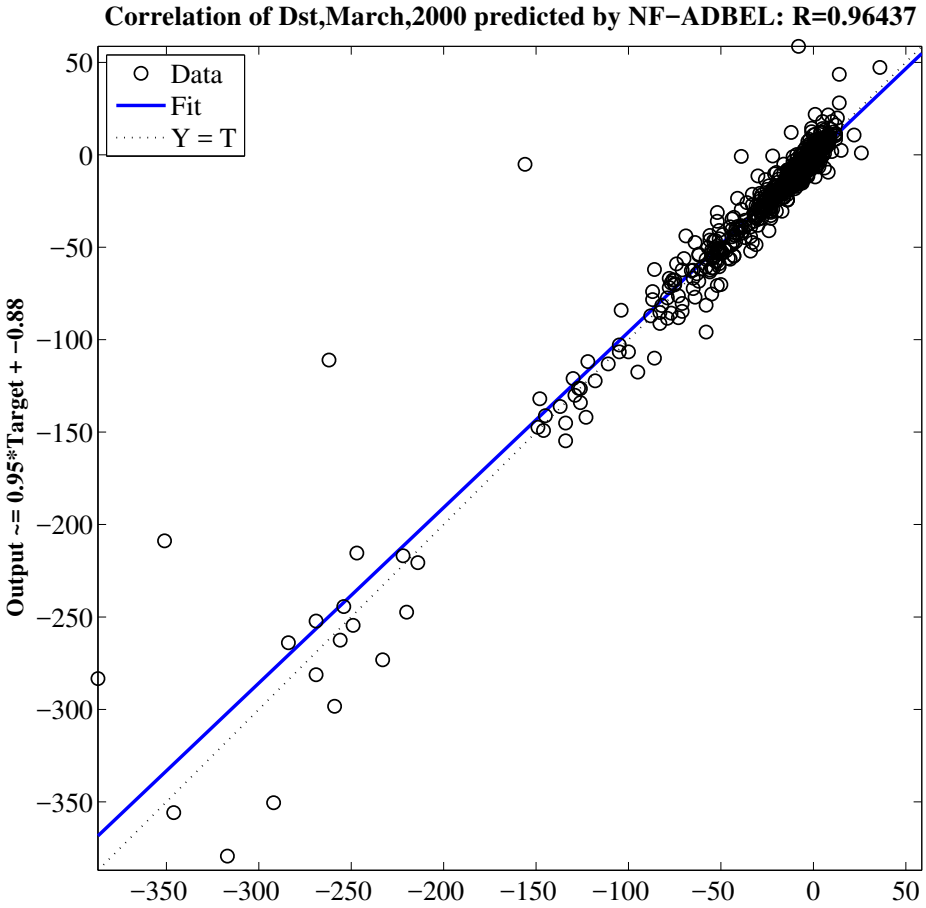

Figure 4.57: Correlation in Predicting Disturbance Storm Time Index D_{st} for March 2001 by NF-ADBEL Network.

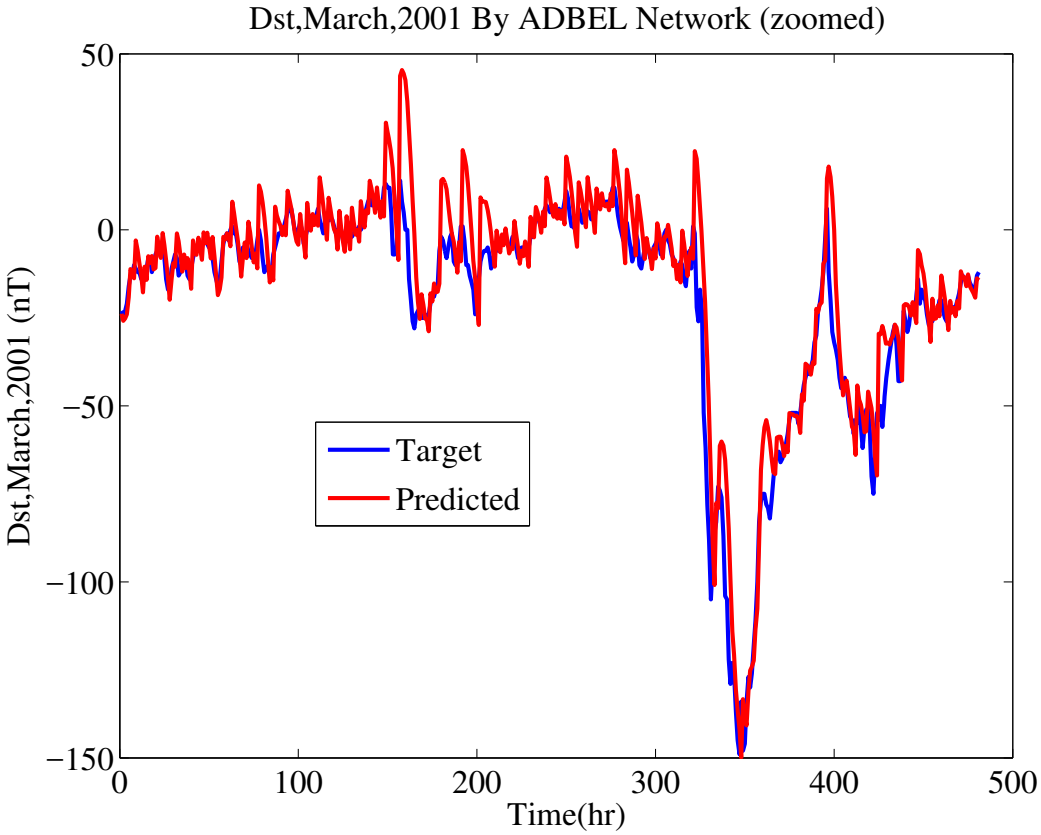

Figure 4.58: Disturbance Storm Time Index D_{st} for March 2001 as Predicted by ADBEL Network.

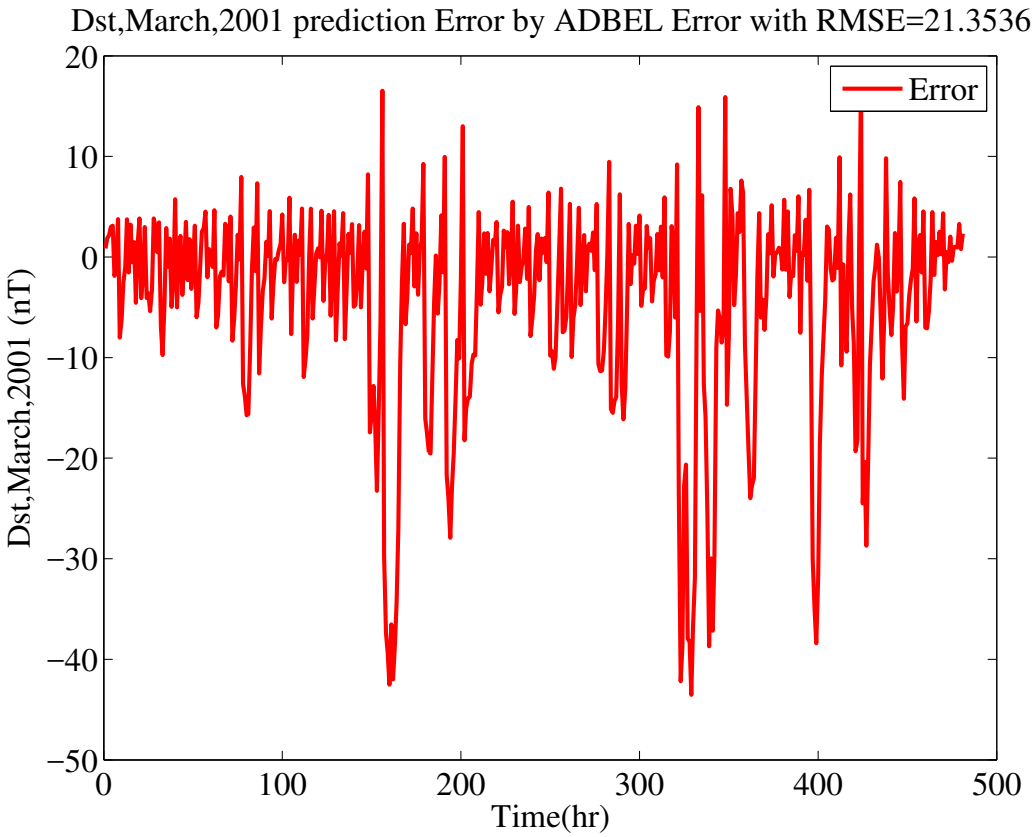

Figure 4.59: Error in Predicting Disturbance Storm Time Index D_{st} for March 2001 by ADBEL Network.

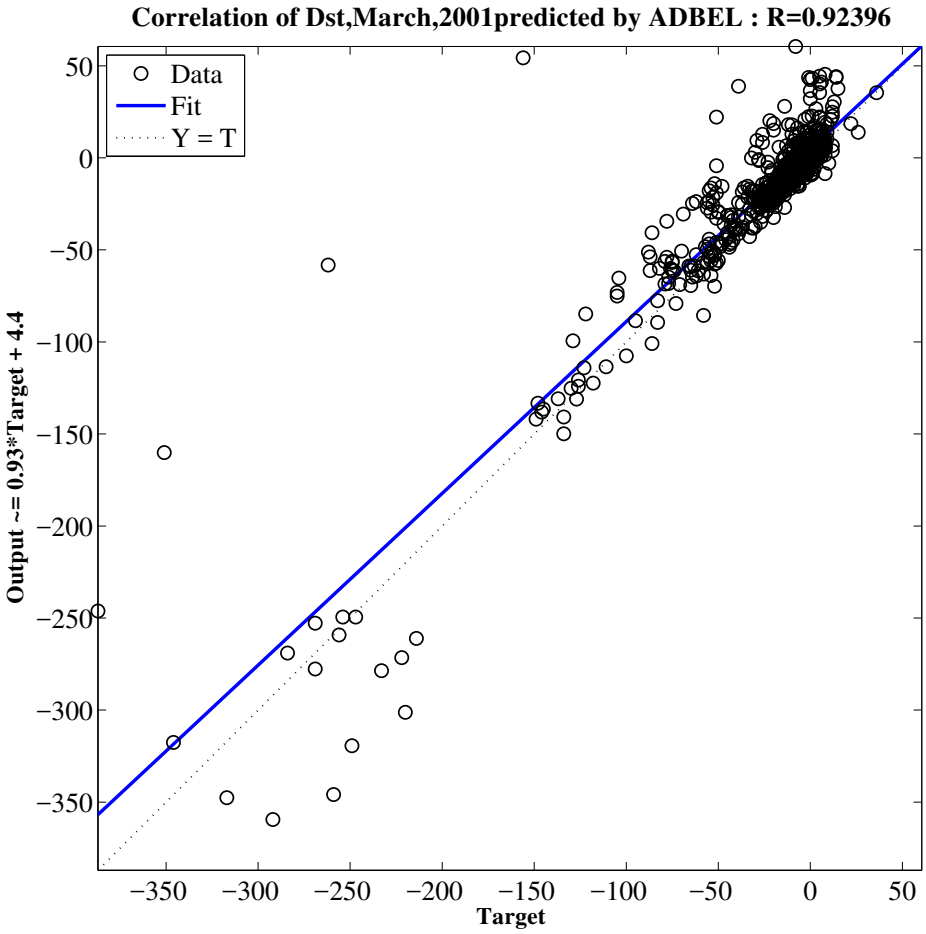

Figure 4.60: Correlation in Predicting Disturbance Storm Time Index D_{st} for March 2001 by ADBEL Network.

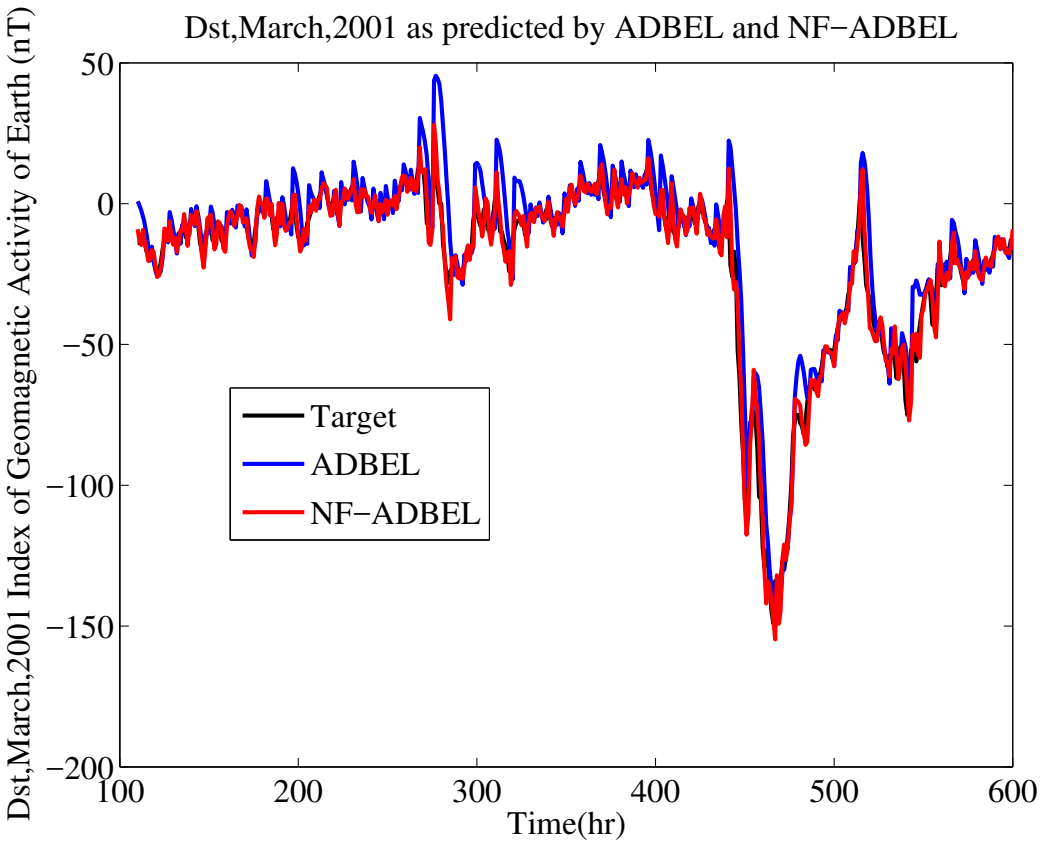

Figure 4.61: Disturbance Storm Time Index D_{st} for March 2001 as Predicted by ADBEL and NF-ADBEL Networks.

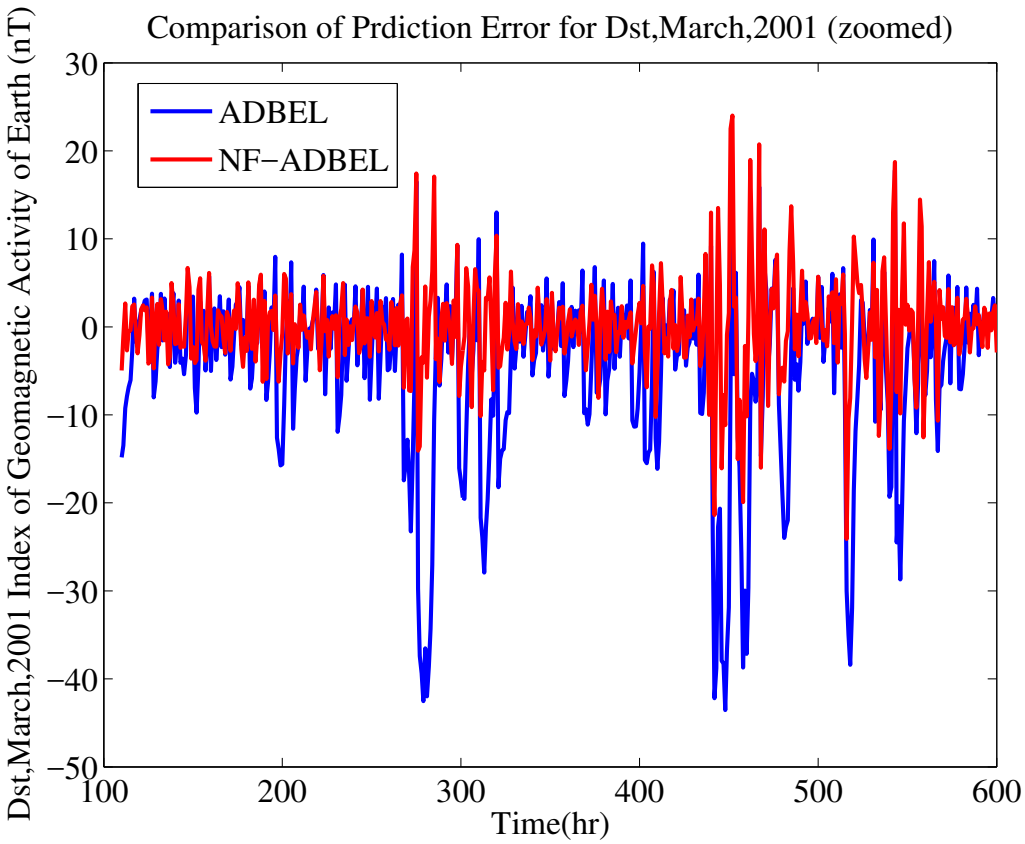

Figure 4.62: Comparison of Errors in Predicting Disturbance Storm Time Index D_{st} for March 2001 by ADBEL and NF-ADBEL Networks.

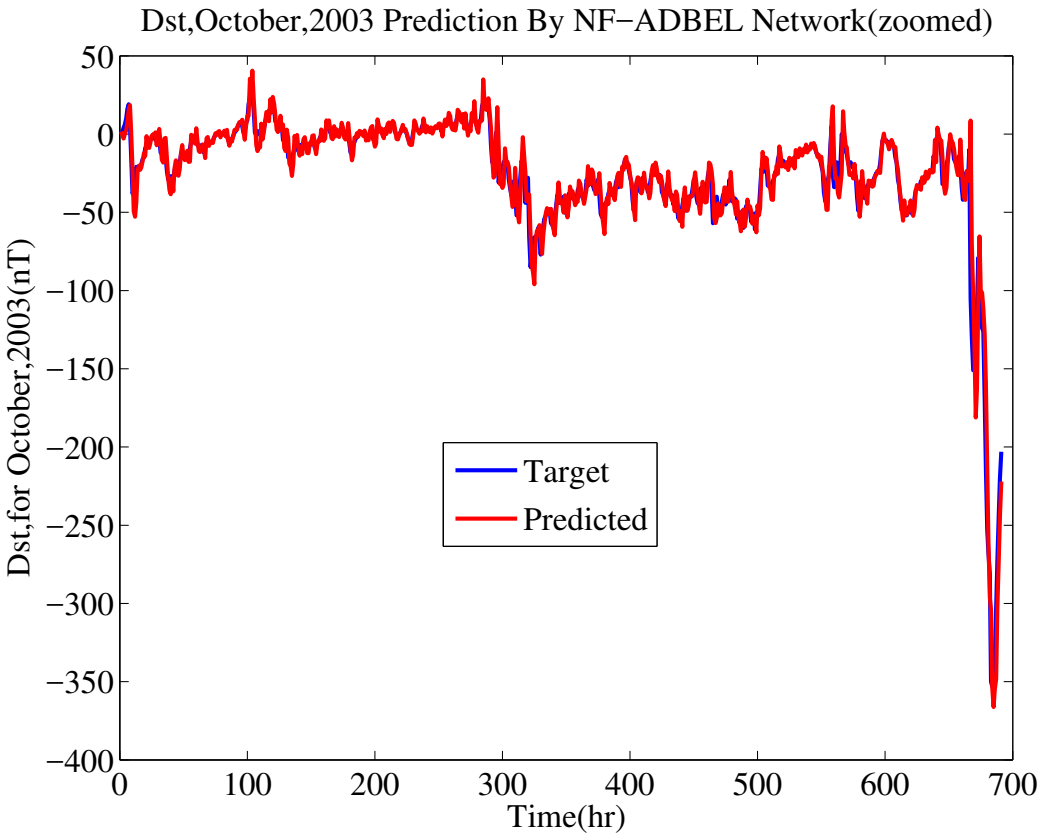

Figure 4.63: Disturbance Storm Time Index D_{st} for October 2003 as Predicted by NF-ADBEL Network.

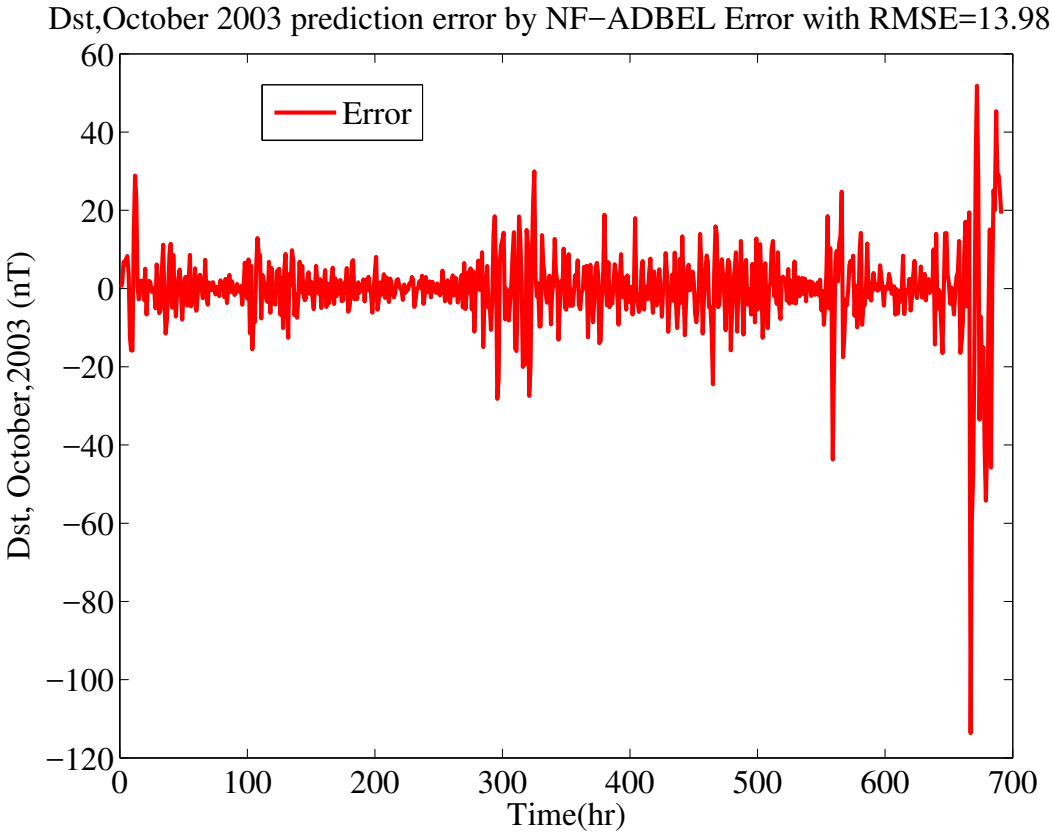

Figure 4.64: Error in Predicting Disturbance Storm Time Index D_{st} for October 2003 by NF-ADBEL Network.

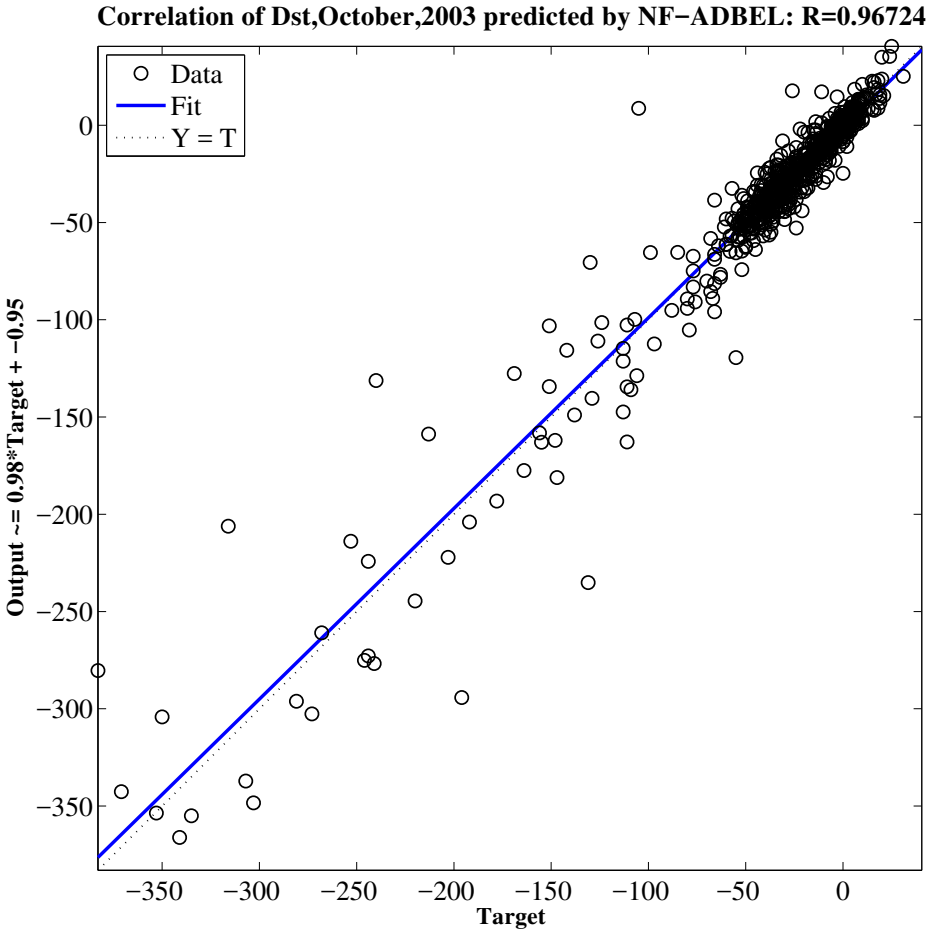

Figure 4.65: Correlation in Predicting Disturbance Storm Time Index D_{st} for October 2003 by NF-ADBEL Network.

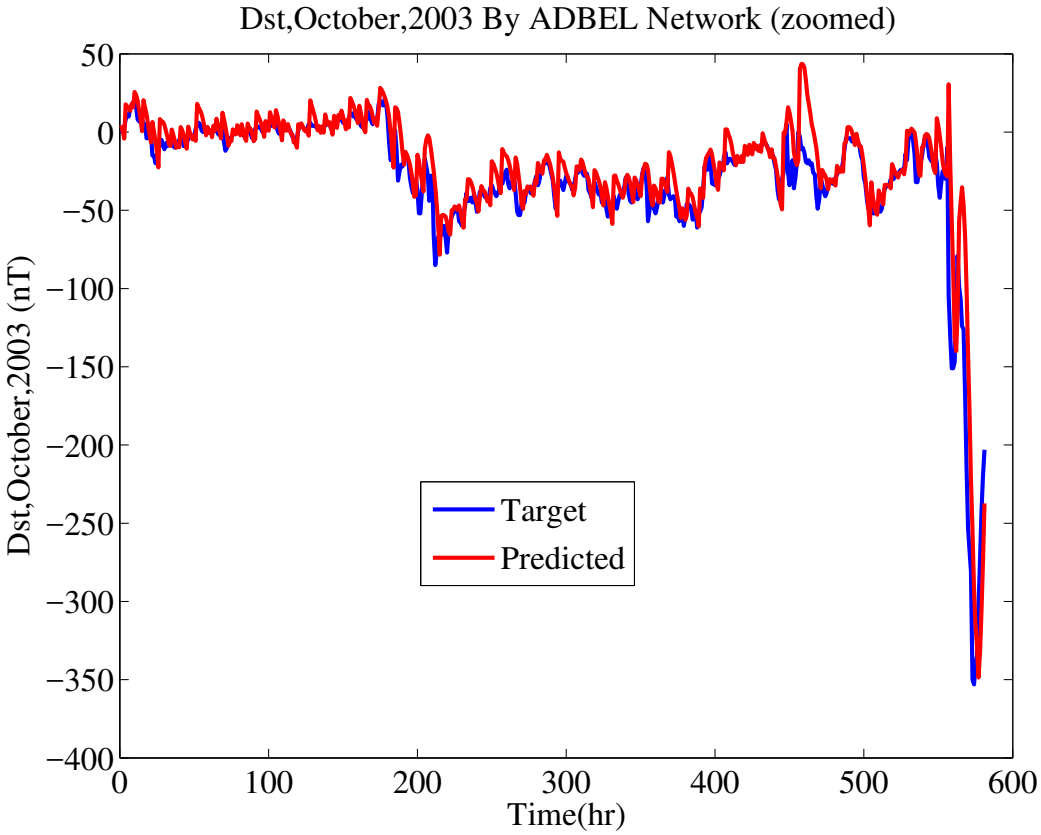

Figure 4.66: Disturbance Storm Time Index D_{st} for October 2003 as Predicted by ADBEL Network.

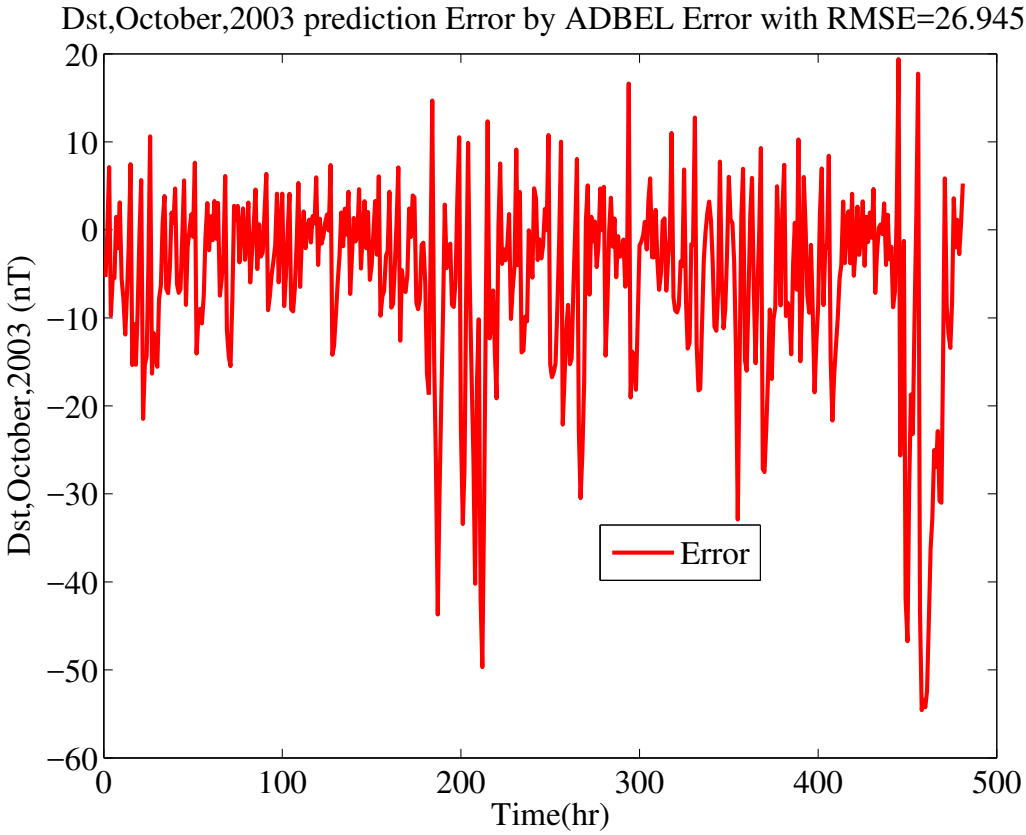

Figure 4.67: Error in Predicting Disturbance Storm Time Index D_{st} for October 2003 by ADBEL Network.

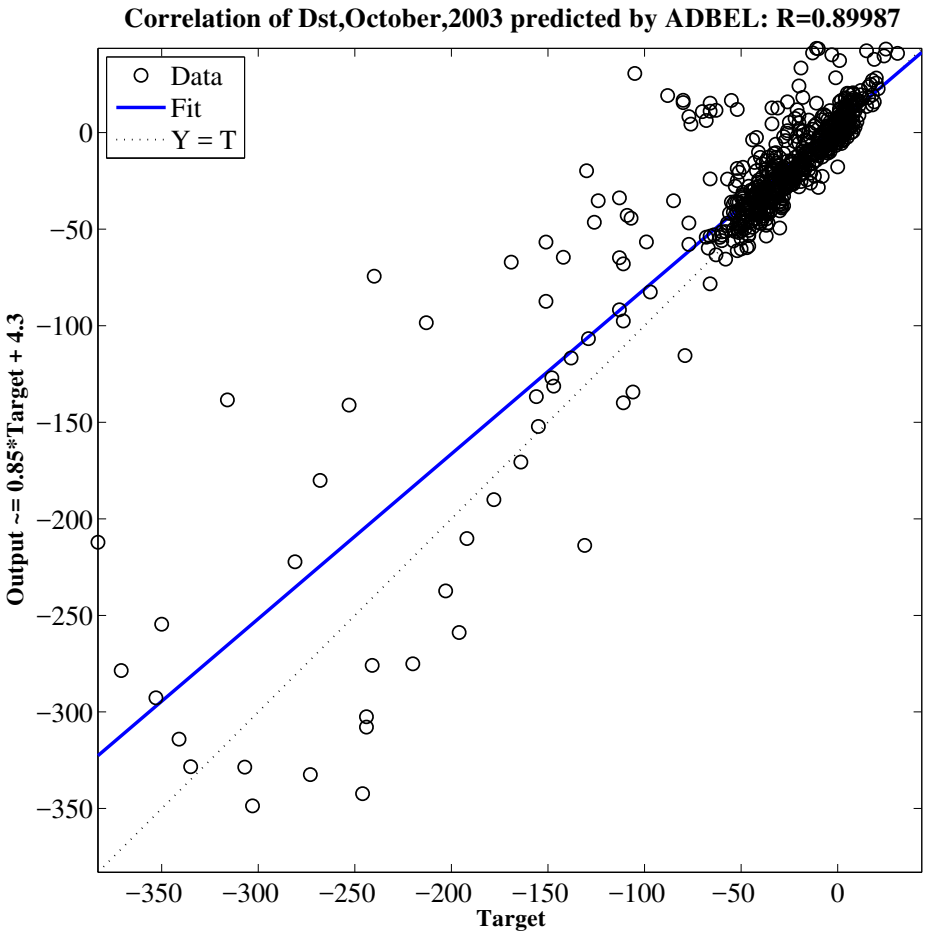

Figure 4.68: Correlation in Predicting Disturbance Storm Time Index D_{st} for October 2003 by ADBEL Network.

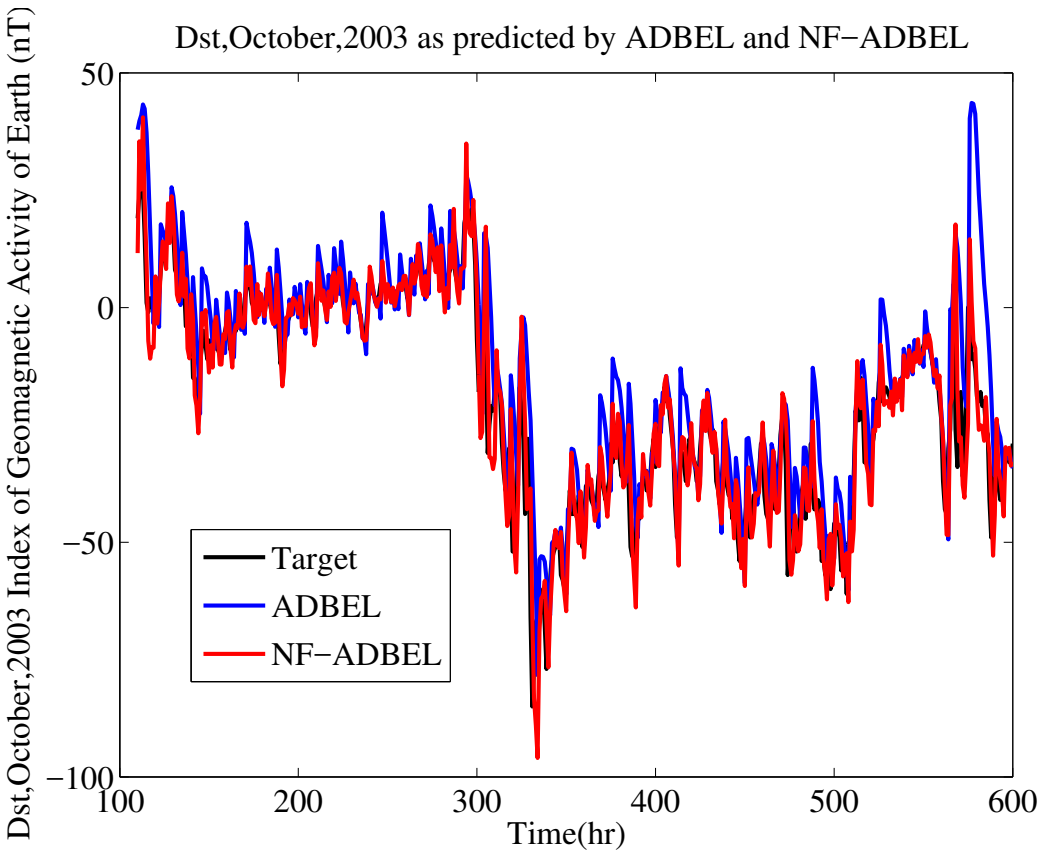

Figure 4.69: Disturbance Storm Time Index D_{st} for October 2003 as Predicted by ADBEL and NF-ADBEL Networks.

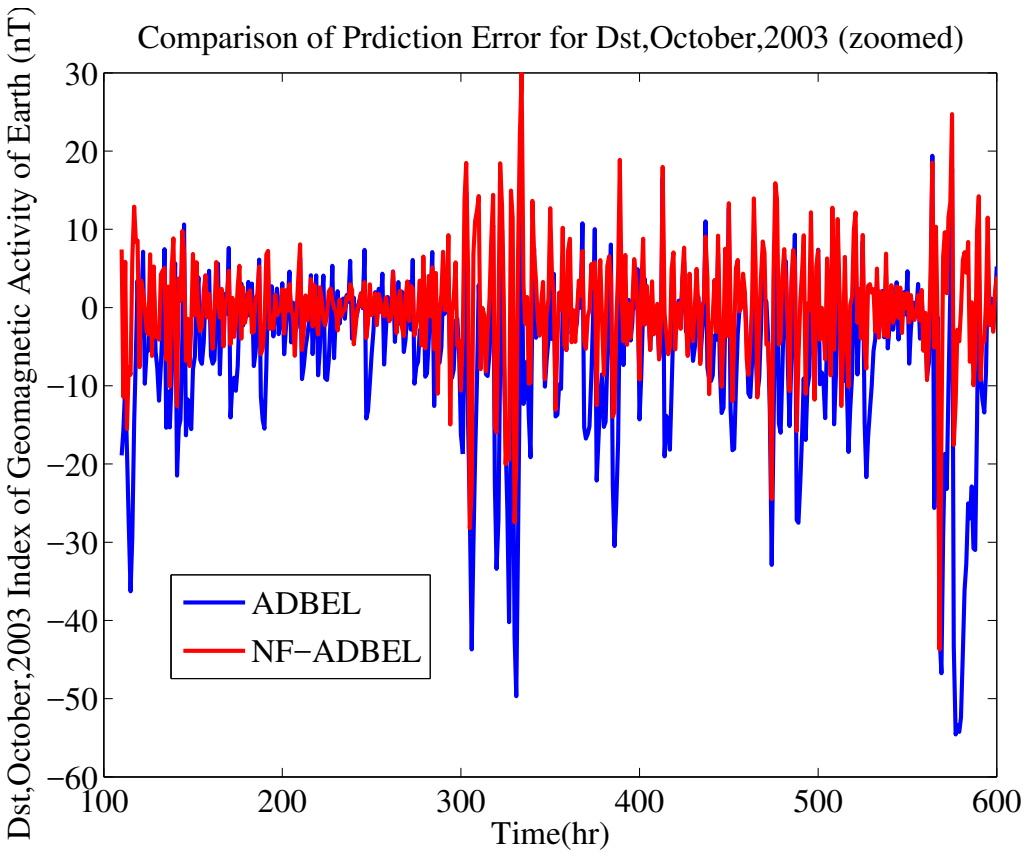

Figure 4.70: Comparison of Errors in Predicting Disturbance Storm Time Index D_{st} for October 2003 by ADBEL and NF-ADBEL Networks.

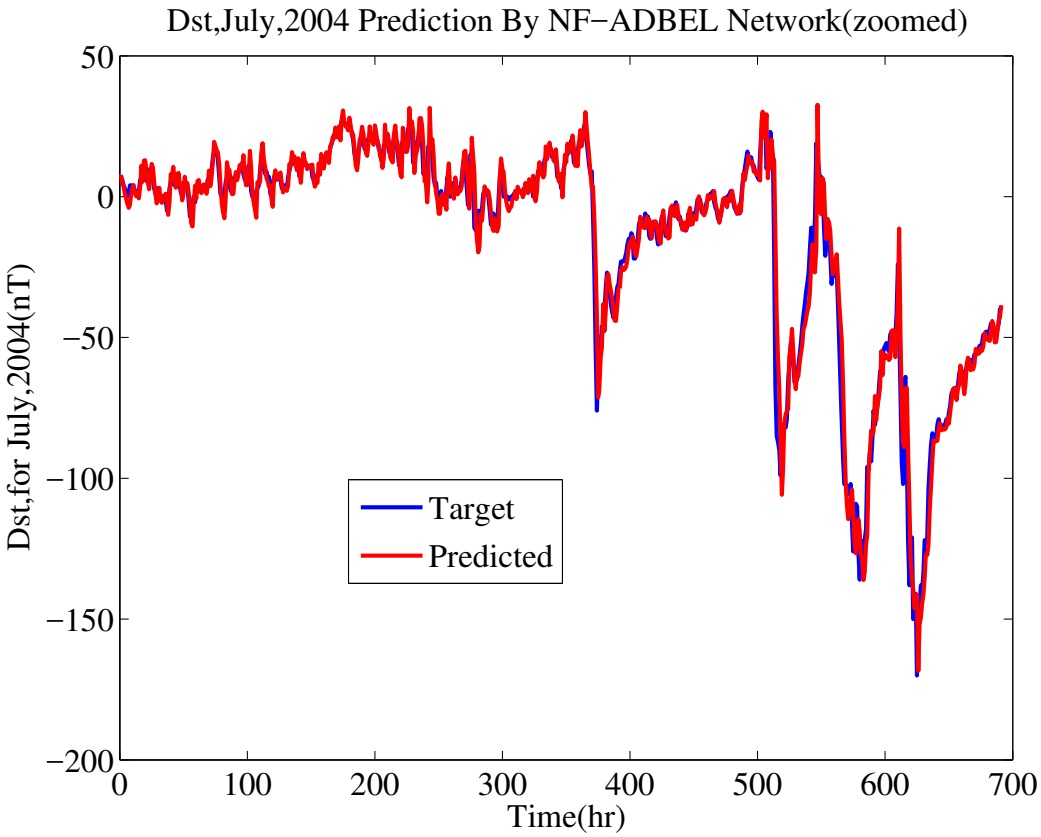

Figure 4.71: Disturbance Storm Time Index D_{st} for July 2004 as Predicted by NF-ADBEL Network.

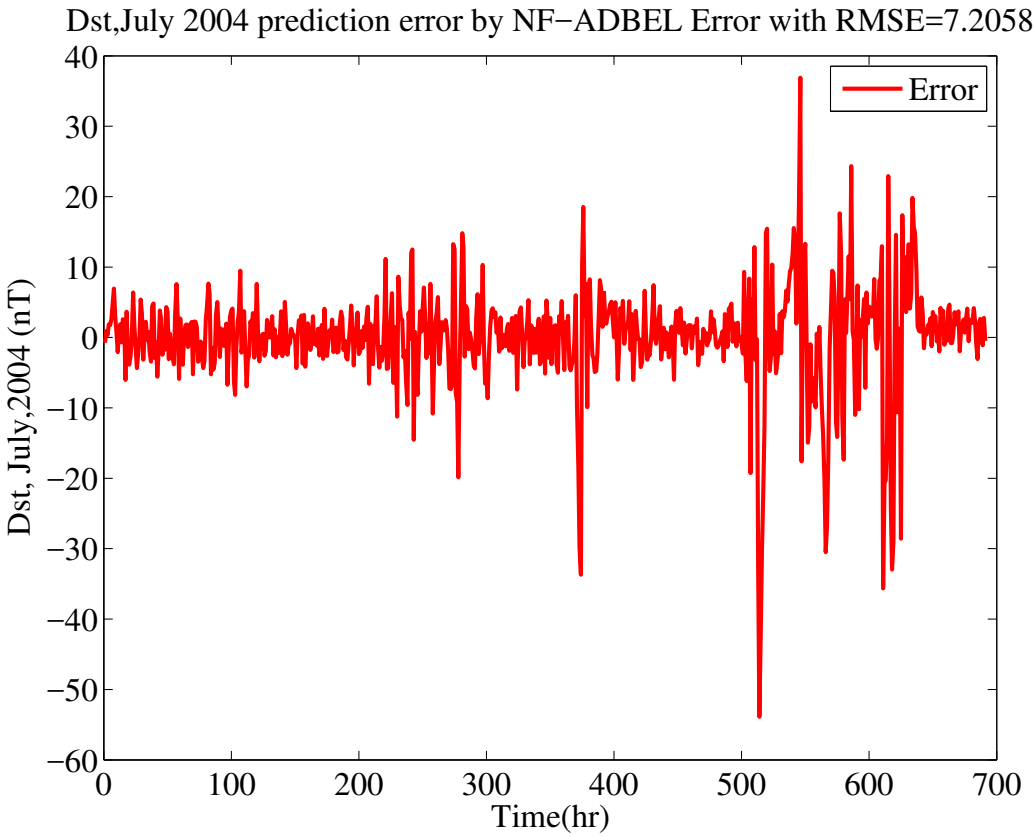

Figure 4.72: Error in Predicting Disturbance Storm Time Index D_{st} for July 2004 by NF-ADBEL Network.

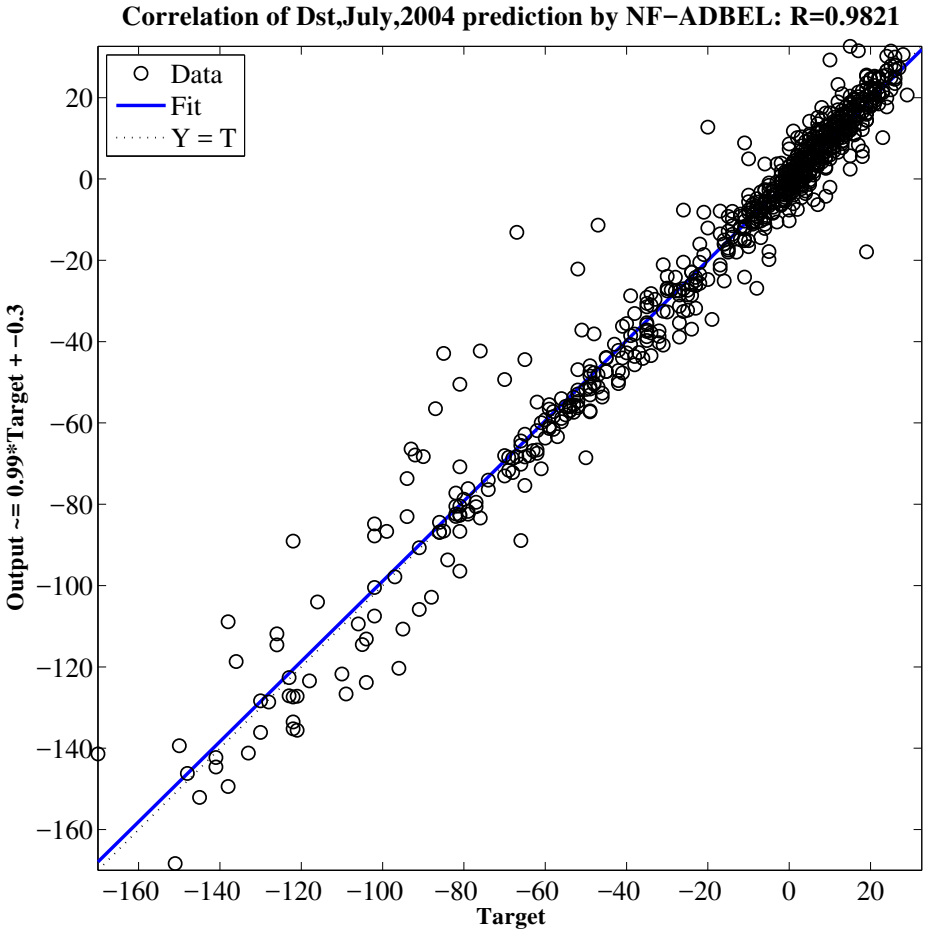

Figure 4.73: Correlation in Predicting Disturbance Storm Time Index D_{st} for July 2004 by NF-ADBEL Network.

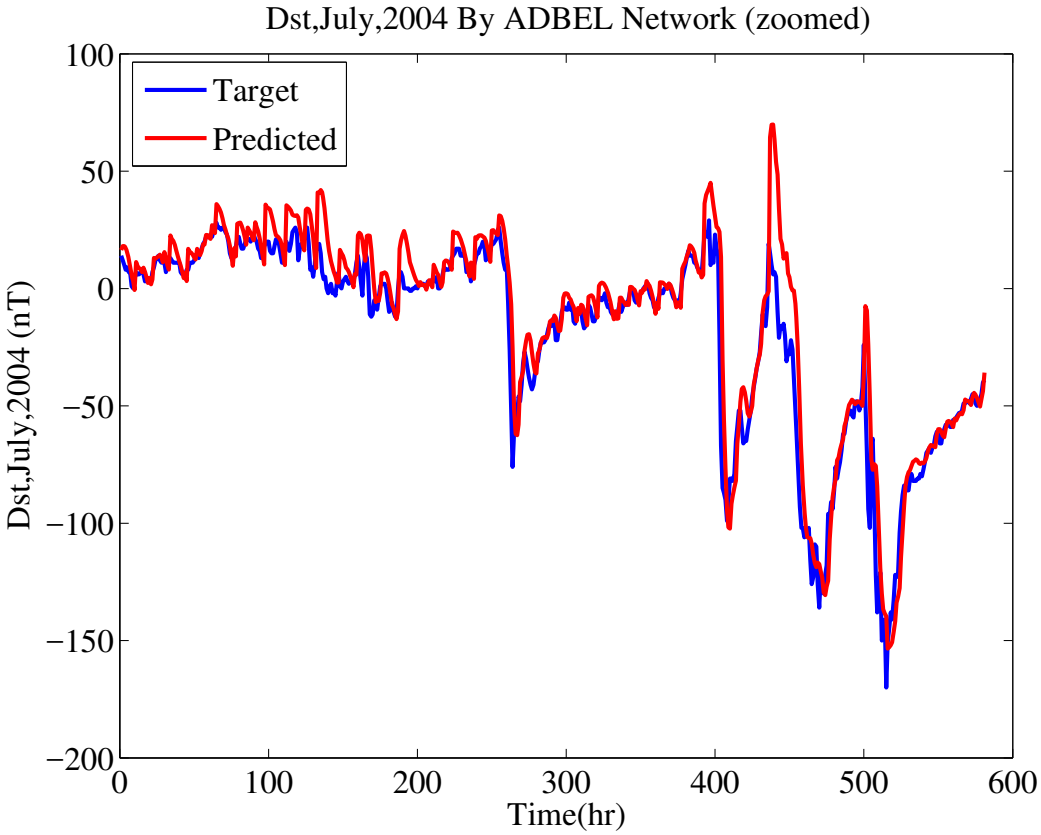

Figure 4.74: Disturbance Storm Time Index D_{st} for July 2004 as Predicted by ADBEL Network.

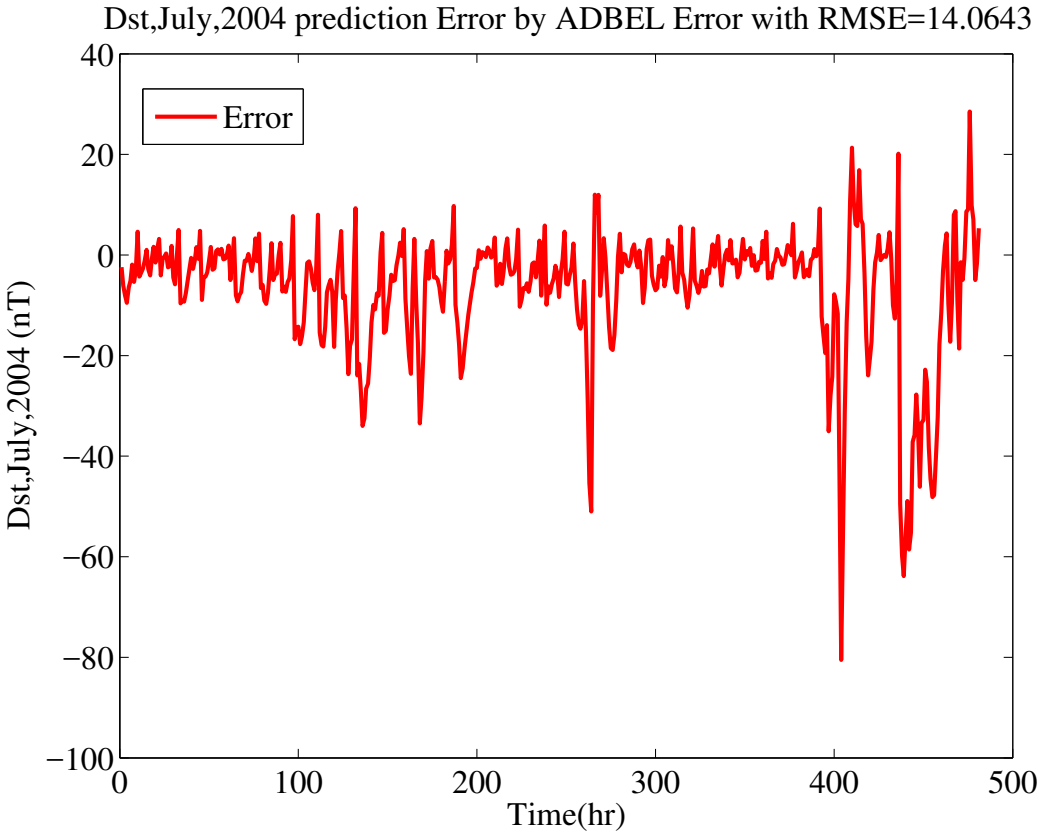

Figure 4.75: Error in Predicting Disturbance Storm Time Index D_{st} for July 2004 by ADBEL Network.

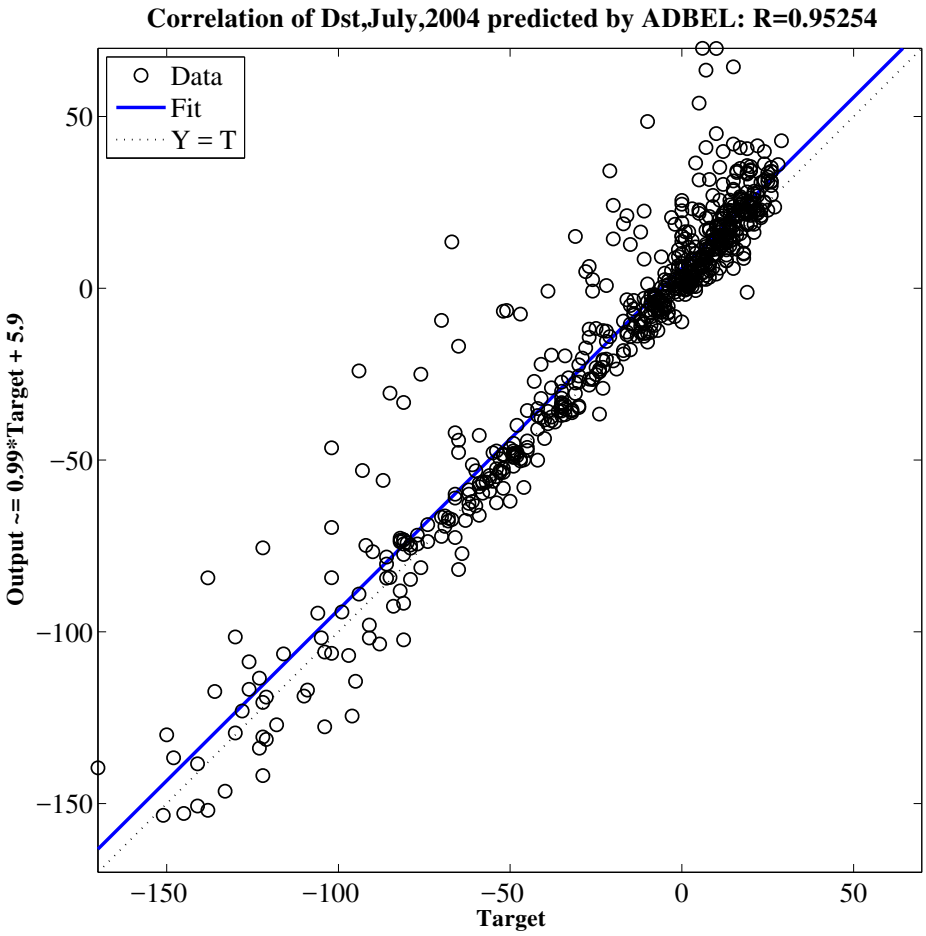

Figure 4.76: Correlation in Predicting Disturbance Storm Time Index D_{st} for July 2004 by ADBEL Network.

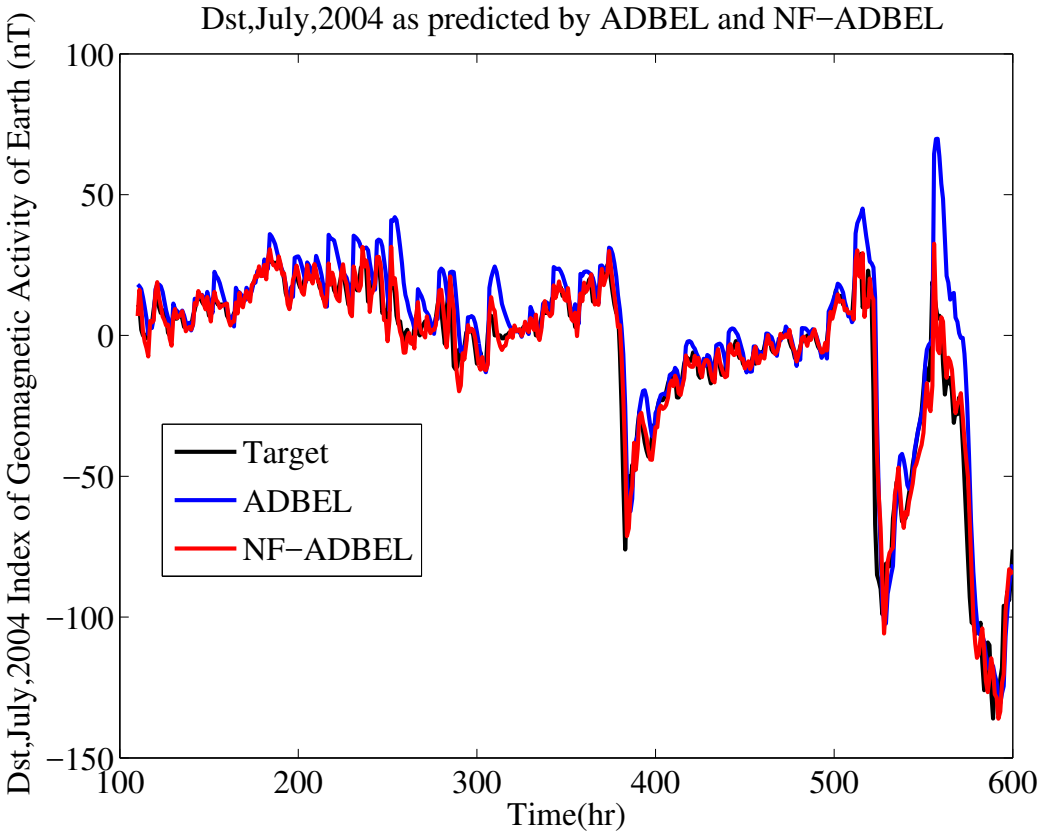

Figure 4.77: Disturbance Storm Time Index D_{st} for July 2004 as Predicted by ADBEL and NF-ADBEL Networks.

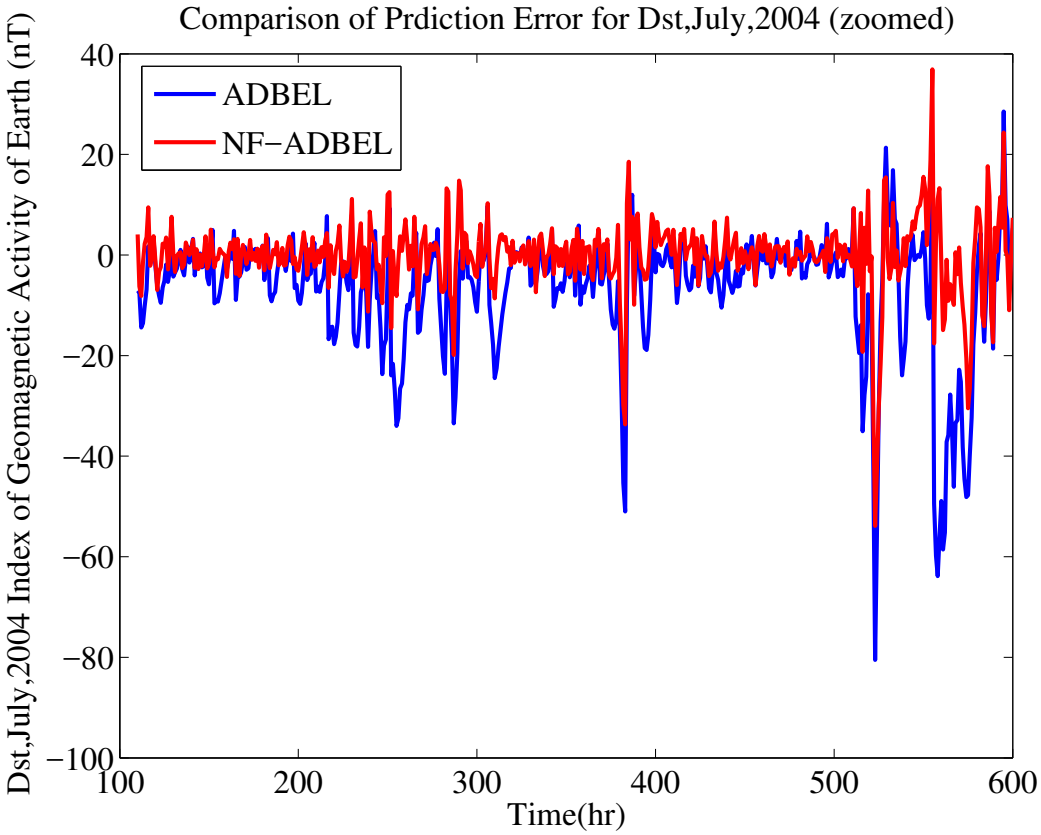

Figure 4.78: Comparison of Errors in Predicting Disturbance Storm Time Index D_{st} for July 2004 by ADBEL and NF-ADBEL Networks.

4.1.5 Narendra Dynamic Plant Predicted by the Proposed NF-ADBEL Network

We have also simulated the NF-ADBEL network for the online identification of the Narendra dynamic plant, which is described by the following discrete equation [87]:

$$y(t+1) = \frac{y(t)}{1+y^2(t)} + f(t) \tag{4.7}$$

$$f(t) = \begin{cases} sin^3(\frac{\pi t}{250}) & , t \leq 500 \\ 0.8\ sin(\frac{\pi t}{250}) + 0.2\ sin(\frac{\pi t}{25}) & , t > 500 \end{cases} \tag{4.8}$$

With an initial condition of $y(1) = 0.5$, a total of $n_e = 1996$ samples are generated according to Eqs. (4.7) and (4.8), respectively. The NF-ADBEL network is first employed to identify the dynamic plant using the learning parameters of $\alpha = 0.3$, $\beta = 0.5$, and $\gamma = 0.01$.

A zoomed view of the identification result in steady-state is shown in Figure 4.80, where the steady-state starting index is assumed to be $n_s = 5sec$. As can be seen, the NF-ADBEL network can identify the dynamic plant, as shown in Figures 4.80, 4.82, and 4.83, respectively.

The Narendra plant is also identified using the ADBEL network (as depicted in Figures 4.85, 4.87 and 4.88, respectively) to compare its identification performance with the NF-ADBEL network. The simulation is run with the learning parameters for the ADBEL network being set as $\alpha = 0.5$, $\beta = 0.5$, and $\gamma = 0.01$. The identification error is presented in Table 4.5.

The temporary period for the ADBEL network is the same as that for the NF-ADBEL network. However, the NF-ADBEL network shows better performance than the ADBEL network, owing to the lesser identification error being offered by this network during steady-state, as shown in Figures 4.90, and 4.92. A lower root mean squared error, higher correlation coefficient, and sufficient percentage improvement as yielded by the NF-ADBEL network validate its superior performance over the ADBEL network in identifying the Narendra plant, as presented in Table 4.5.

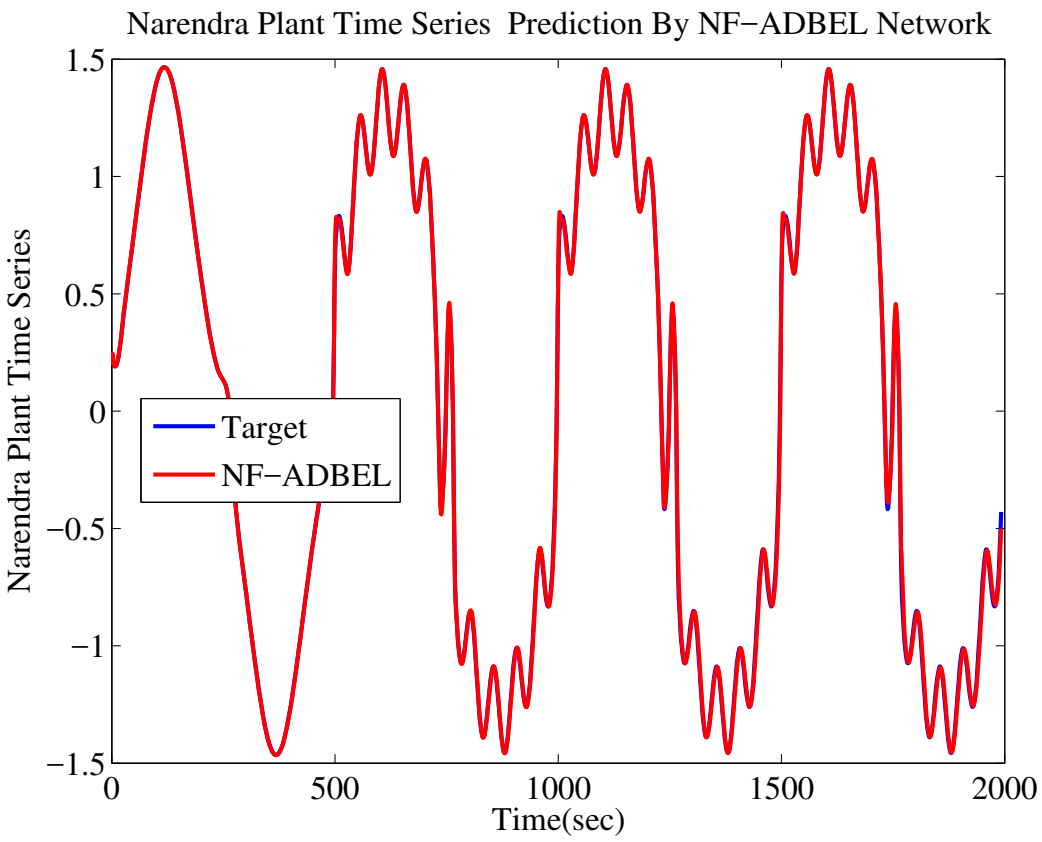

Figure 4.79: Narendra Plant as Predicted by NF-ADBEL Network.

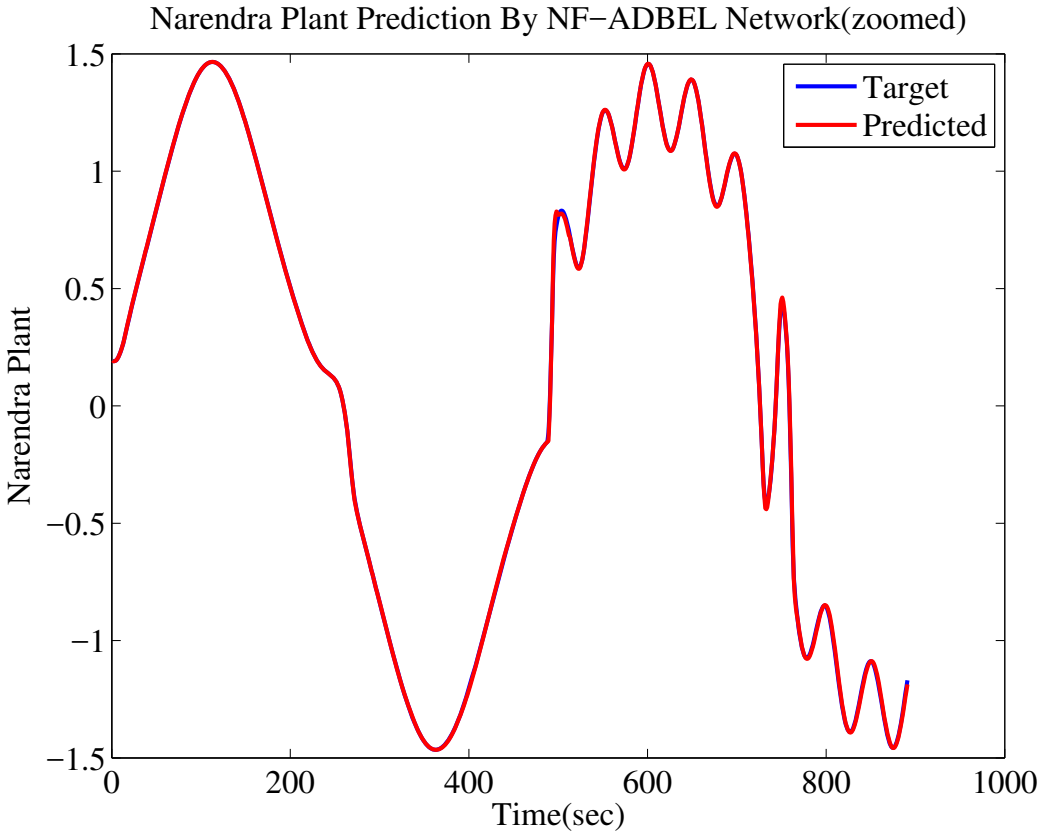

Figure 4.80: Narendra Plant as Predicted by NF-ADBEL Network (displayed a portion of Fig. 4.79).

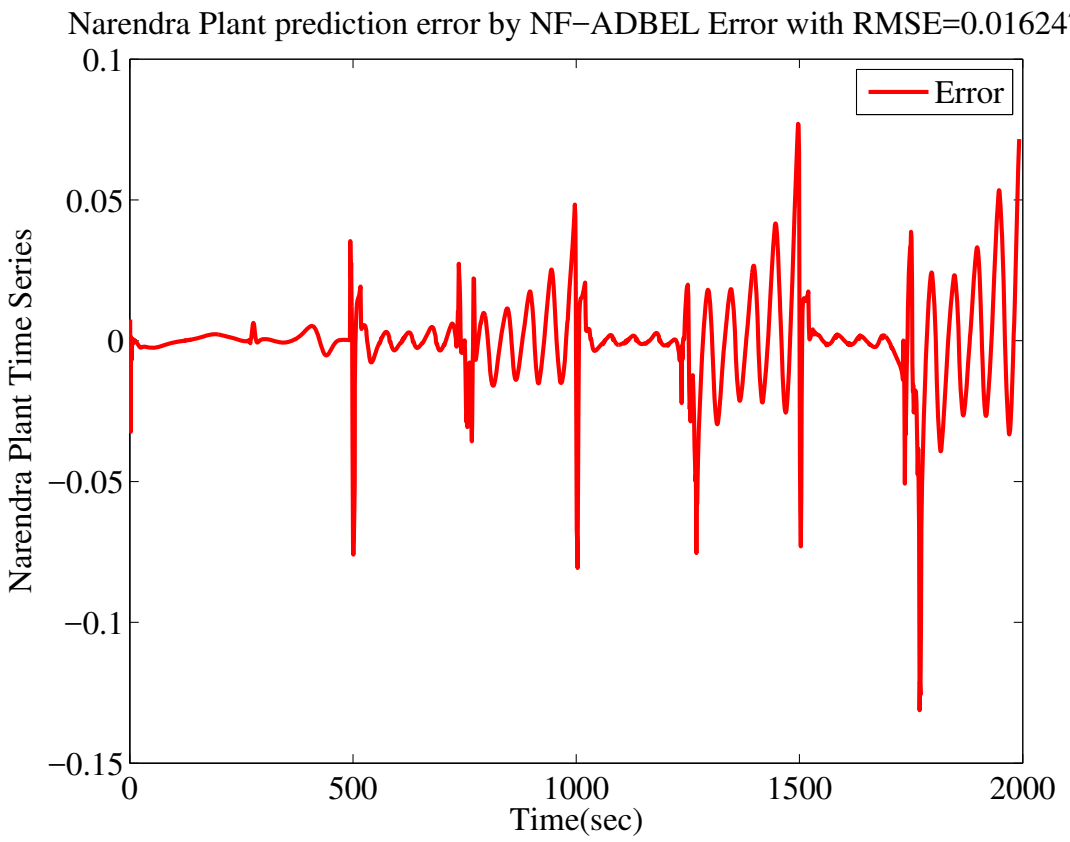

Figure 4.81: Error in Predicting Narendra Plant by NF-ADBEL Network.

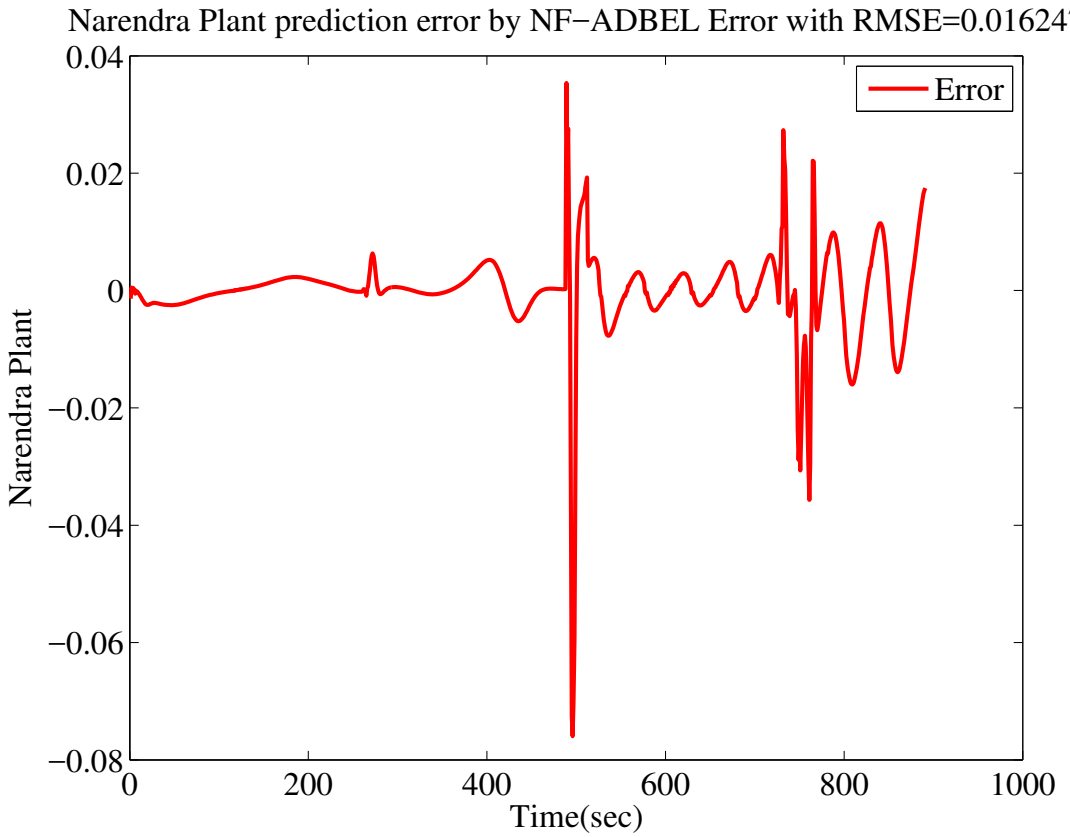

Figure 4.82: Error in Predicting Narendra Plant by NF-ADBEL Network(displayed a portion of Fig.4.81).

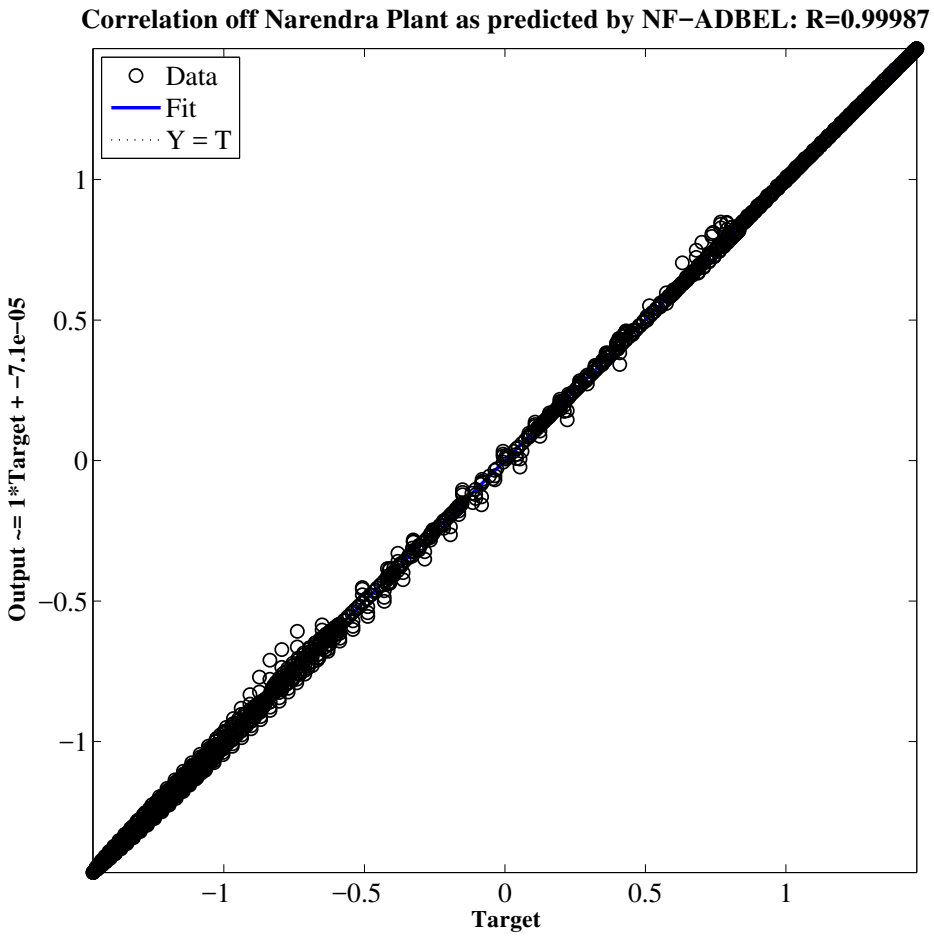

Figure 4.83: Correlation in Predicting Narendra Plant by NF-ADBEL Network.

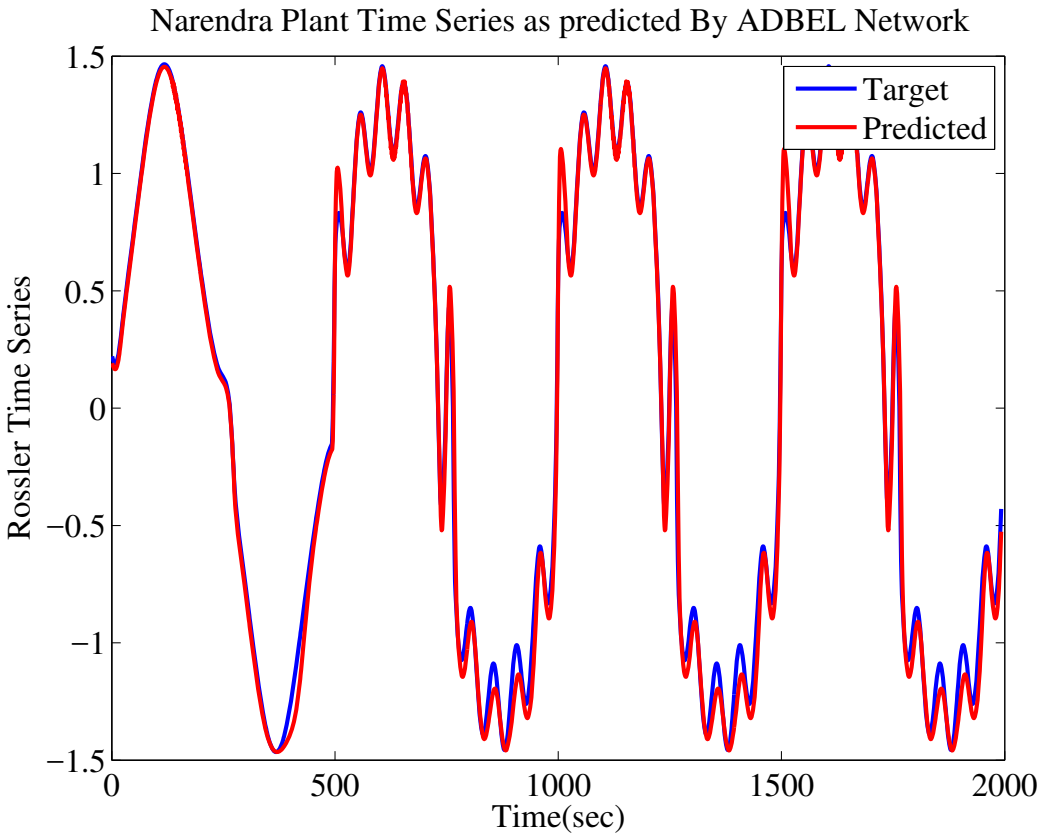

Figure 4.84: Narendra Plant as Predicted by ADBEL Network.

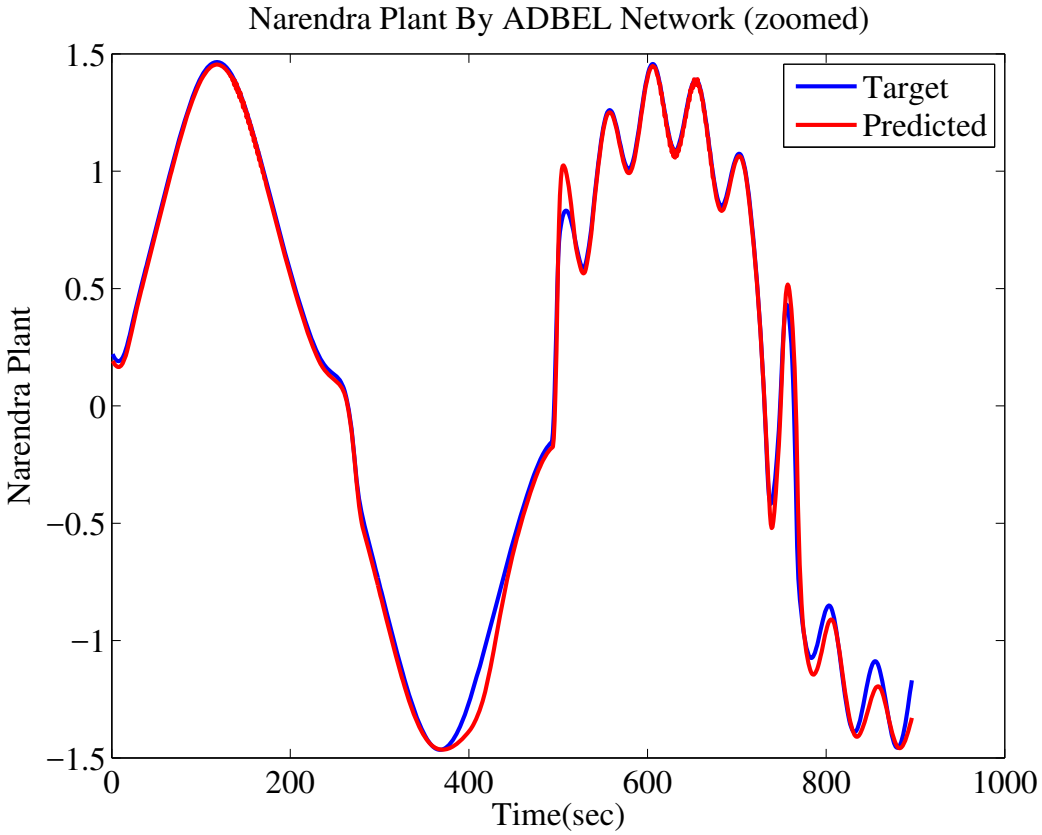

Figure 4.85: Narendra Plant as Predicted by ADBEL Network (displayed a portion of Fig. 4.84).

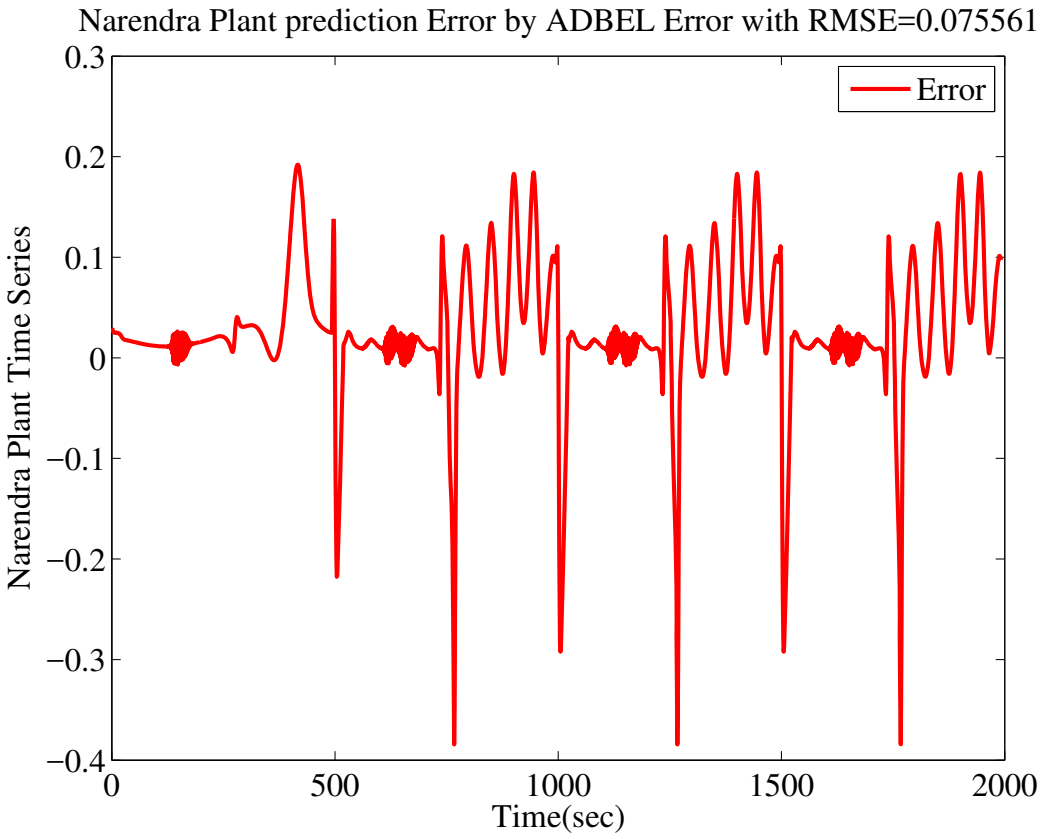

Figure 4.86: Error in Predicting Narendra Pant by ADBEL Network.

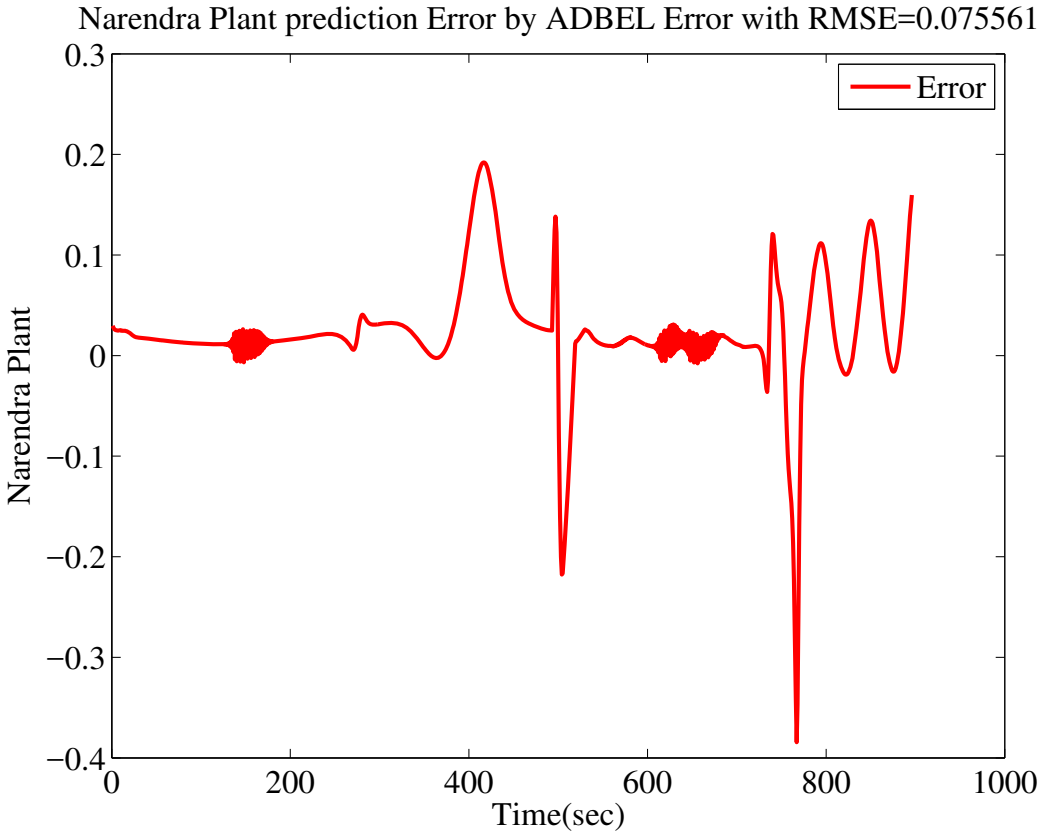

Figure 4.87: Error in Predicting Narendra Pant by ADBEL Network (displayed a portion of Fig. 4.86).

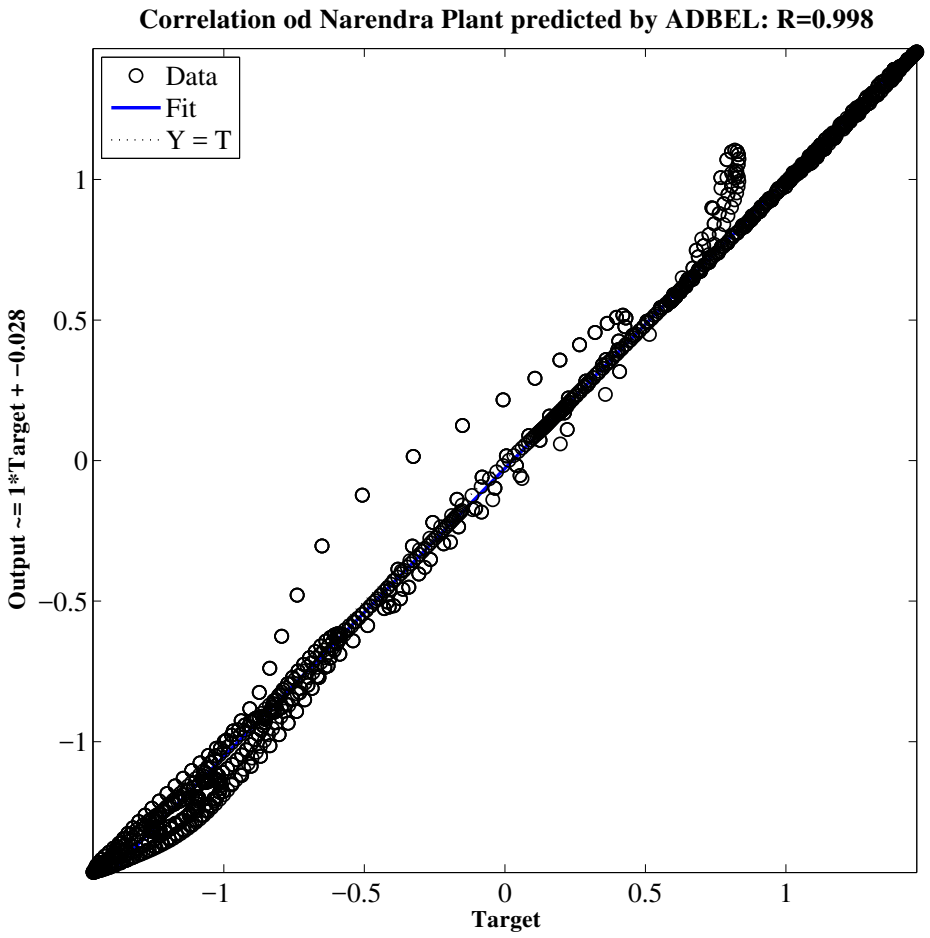

Figure 4.88: Correlation in Predicting Narendra Plant by ADBEL Network.

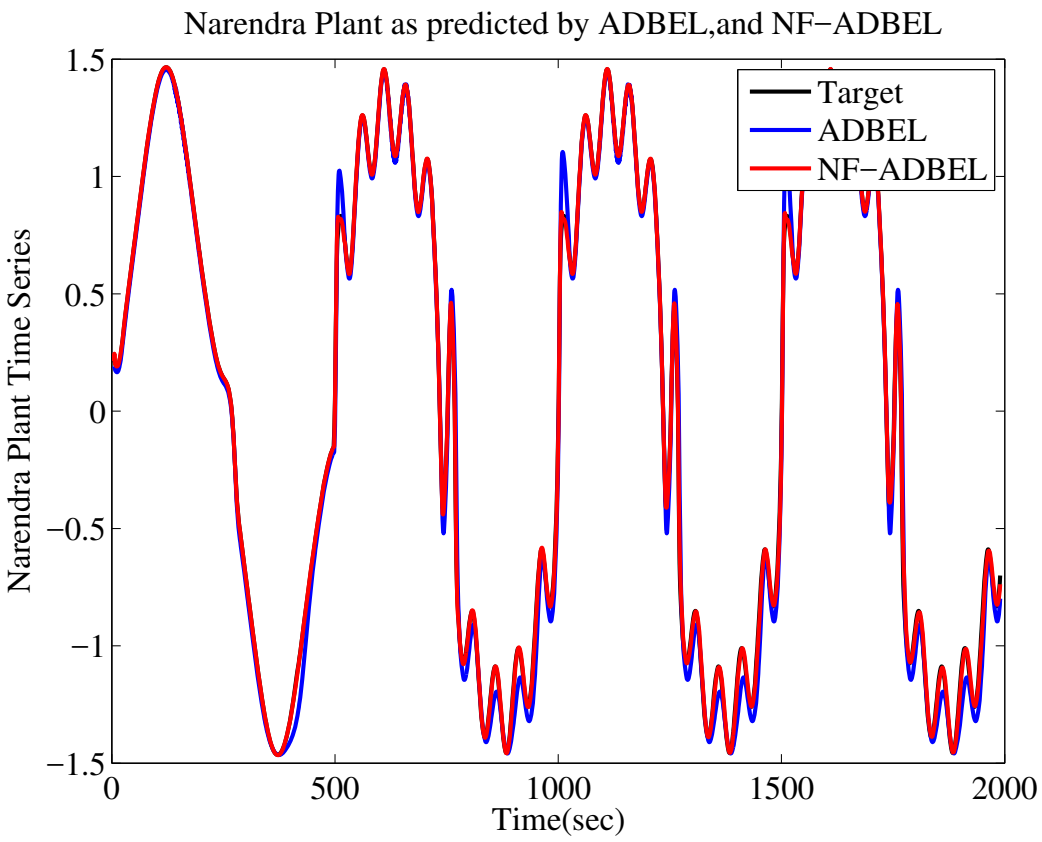

Figure 4.89: Narendra Plant as Predicted by ADBEL and NF-ADBEL Networks.

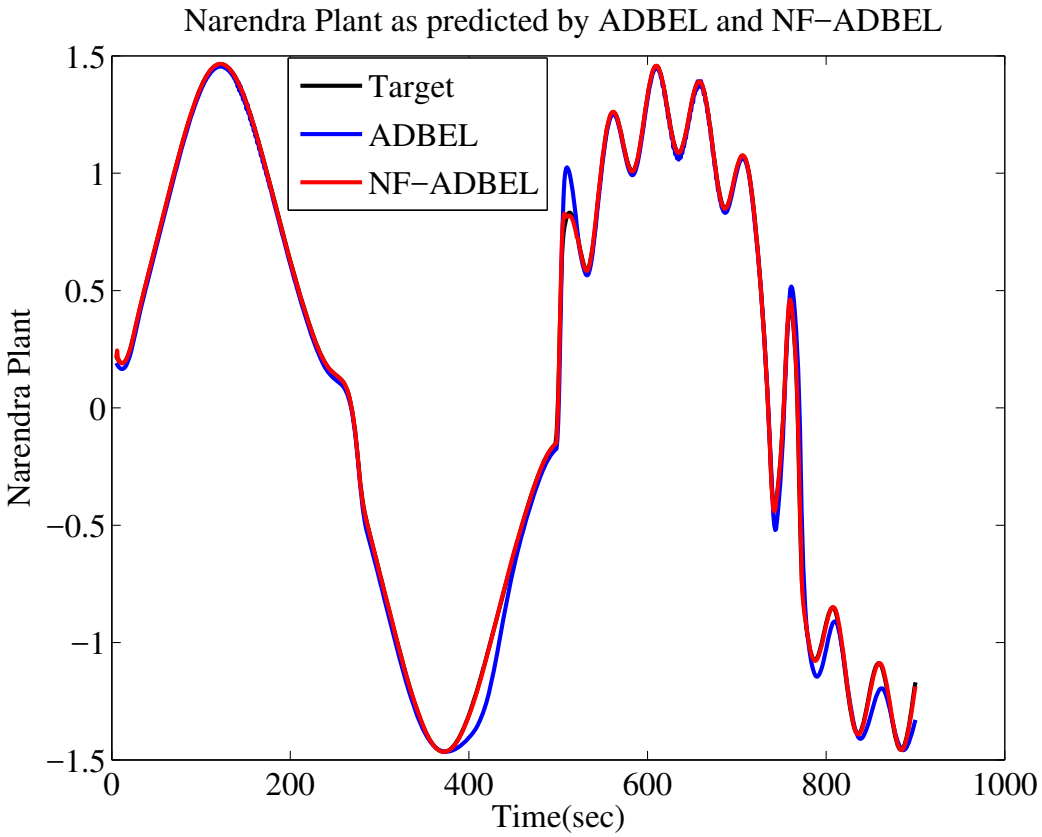

Figure 4.90: Narendra Plant as Predicted by ADBEL and NF-ADBEL Networks(displayed a portion of Fig. 4.89).

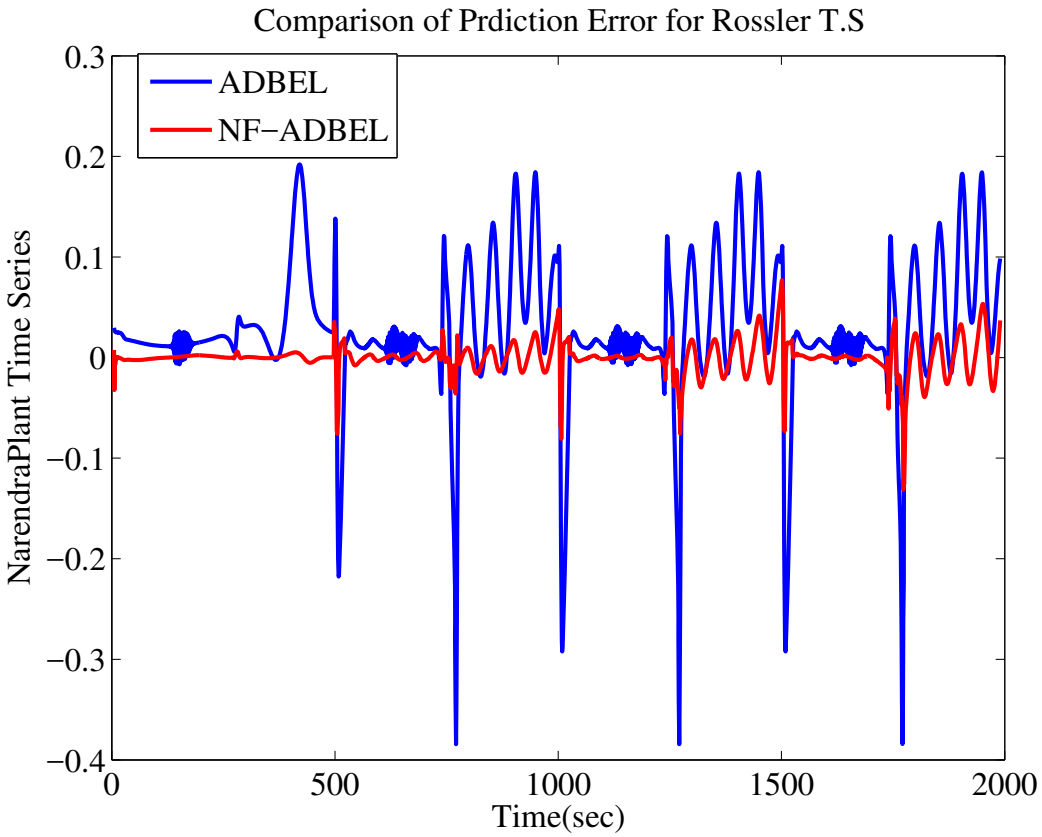

Figure 4.91: Comparison of Errors in Predicting Narendra Plant by ADBEL and NF-ADBEL Networks.

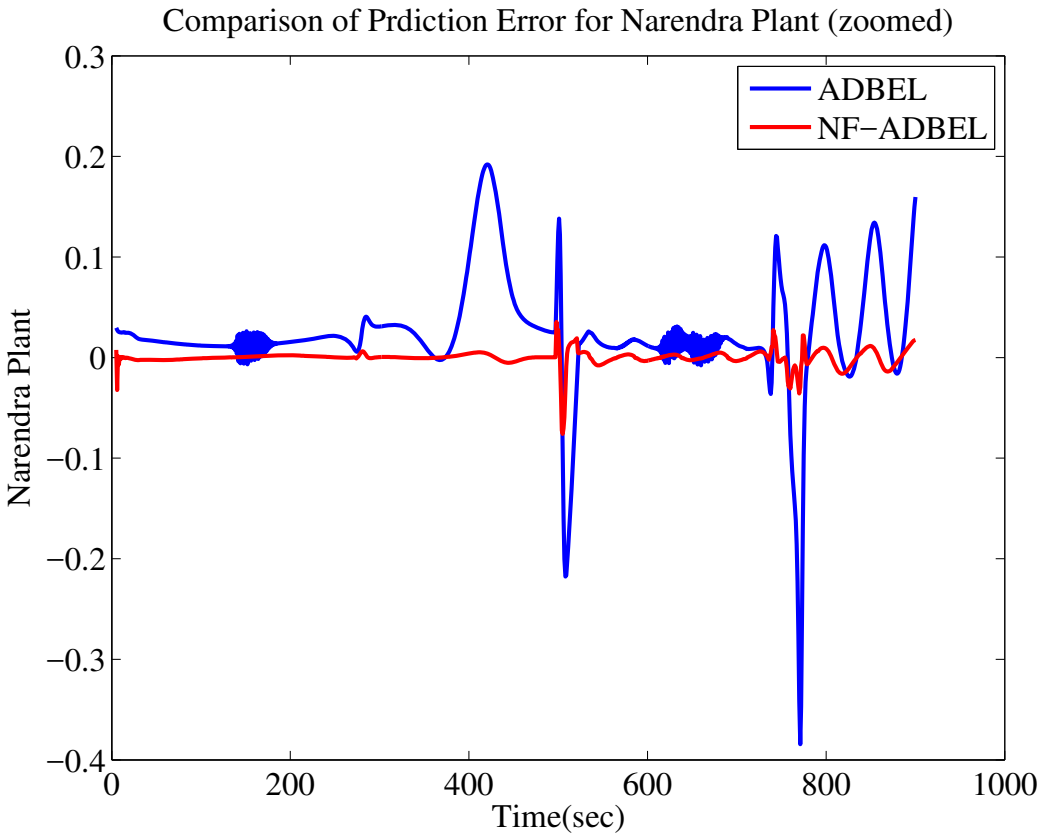

Figure 4.92: Comparison of Errors in Predicting Narendra Plant by ADBEL and NF-ADBEL Networks(displayed a portion of Fig. 4.91).

Table 4.5: RMSE /COR/PI For Narendra Dynamic Plant Identification Prediction by ADBEL and NF-ADBEL Networks

Time Series	Prediction Network	RMSE	COR	PI(%)
Narendra Plant	ADBEL	0.07556	0.998	
	NF-ADBEL	0.0162	0.99989	**78.49**

4.1.6 Conclusions

The design of a neo-fuzzy integrated ADBEL network (NF-ADBEL) is presented in this work for the time series prediction in an online mode. The integration of the neo-fuzzy network is only considered in the orbitofrontal cortex section to retain the simplicity and quickness of the proposed NF-ADBEL network. The selection of a few membership functions for implementing the neo-fuzzy neurons further helps keep the computational complexity of the proposed network at a minimum. Simulations are carried out in a MATLAB programming environment to predict several chaotic time series, including Mackey-Glass, Lorenz, Rossler, and disturbance storm time index. Simulations are also conducted to identify a dynamic Narendra plant model by deploying the proposed NF-ADBEL network. The proposed NF-ADBEL network's performance is evaluated using root mean squared error and correlation coefficient criteria. Also, the ADBEL network was redesigned and simulated to predict the exact time-series with near optimal parameters. A percentage improvement index is defined to compare the performance of both networks. Simulation results show the superiority of the proposed NF-ADBEL network, with a lower root mean squared error and higher correlation coefficient being obtained compared to the ADBEL network. A fair amount of percentage improvement is also observed in all cases when the NF-ADBEL network is used to predict the time series data.

4.2 Performance of the Proposed Fuzzy Logic-Based Parameter Adjuster Model for ADBEL (F-ADBEL) Network

Please note that in this study we are not seeking to improve the accuracy of ADBEL. Rather, we are trying to demonstrate the capability of the designed F-ADBEL to adjust the parameters α, β, and γ for different types of prediction applications.

The proposed fuzzy integrated F-ADBEL network is tested in a MATLAB (R2014a) programming environment for online forecasting of chaotic time series that includes Mackey-Glass, Lorenz, Rossler and disturbance storm index. The performance of the proposed F-ADBEL network is accessed in terms of root mean squared error, correlation coefficient and percentage improvement criteria, as defined in Eqs. (4.1), (4.2) and (4.3), respectively.

The time-series data are first normalized to the range [0,1]. The normalized data are then arranged such that the first four samples form the inputs, while the fifth sample presents the output. By following this method, the size of the input data for all the time series is set as $4 \times n_e$, while the size of the output data is set as $1 \times n_e$. After running the ADBEL and proposed F-ADBEL algorithms, the predicted time series data are de-normalized. Note that in this study, all network weights are initialized as zeros instead of randomly assigning them. This setting also helps run the simulations only once, and no averaging of the results is required, as the networks will yield the same performance every time.

4.2.1 Mackey-Glass Chaotic Time Series Predicted by the Proposed F-ADBEL Network

Let us first predict the time series data generated from a time-delayed Mackey-Glass nonlinear differential equation, as defined in Eq. (4.4). This equation is simulated in MATLAB to create the time series data, which can be observed as non-periodic and non-convergent. The fuzzy integrated F-ADBEL network is deployed to predict this time series, as shown in Figures 4.93, 4.94 and 4.95, in terms of root mean squared error and correlation coefficient, with the results depicted in Table 4.6. The variations of the amygdala weights and orbitofrontal cortex weights are presented in Figures 4.96 and 4.97, respectively.

Table 4.6: RMSE /COR/PI for Mackey-Glass Prediction by ADBEL and F-ADBEL Networks

Time Series	Prediction Network	RMSE	COR	PI(%)
Mackey-Glass	ADBEL	0.047	0.9895	
	F-ADBEL	0.037	0.9918	**21.27**

Figure 4.97 illustrates that the orbitofrontal cortex weights converge after the initial transient period. On the other hand, the amygdala weights, as shown in Figure 4.96, display more significant variation than orbitofrontal cortex weights in steady-state. This supports the hypothesis that the orbitofrontal cortex is the more stable portion of the emotional brain and can correct the amygdala's response, which reacts quickly to emotional stimuli. Note that the amygdala weights are lower-bounded by zero, as was discussed during the development of the fuzzy parameter adjuster in section 3.7.

The variations of the network parameters α, β, and γ produced by the proposed fuzzy model F-ADBEL are also shown in Figures 4.99, 4.100 and 4.101, respectively. As can be observed, the variations in parameter β are more stable, which can also describe the changes in orbitofrontal cortex weights. On the other hand, the fuzzy model continues adjusting the different network parameters (α, and γ) throughout the entire horizon, which also explains the variation of the amygdala weights. Figure 4.98 shows the interpretation of the reward signal. It could assume either positive or negative values and support the concept discussed during the fuzzy parameter adjuster development in section 3.7.

The same time series is also predicted with the ADBEL network using the learning parameters of $\alpha = 0.5$, $\beta = 0.8$ and $\gamma = 0.03$. As illustrated in Figure 4.102, the prediction errors are recorded in both cases, with analysis showing that the transient period remains the same: ≤ 5 s. Thus, the steady-state starting index is set to compute the performance indices in both cases. The performances of the designed F-ADBE and ADBEL networks for predicting Mackey-Glass are shown in Table 4.6.

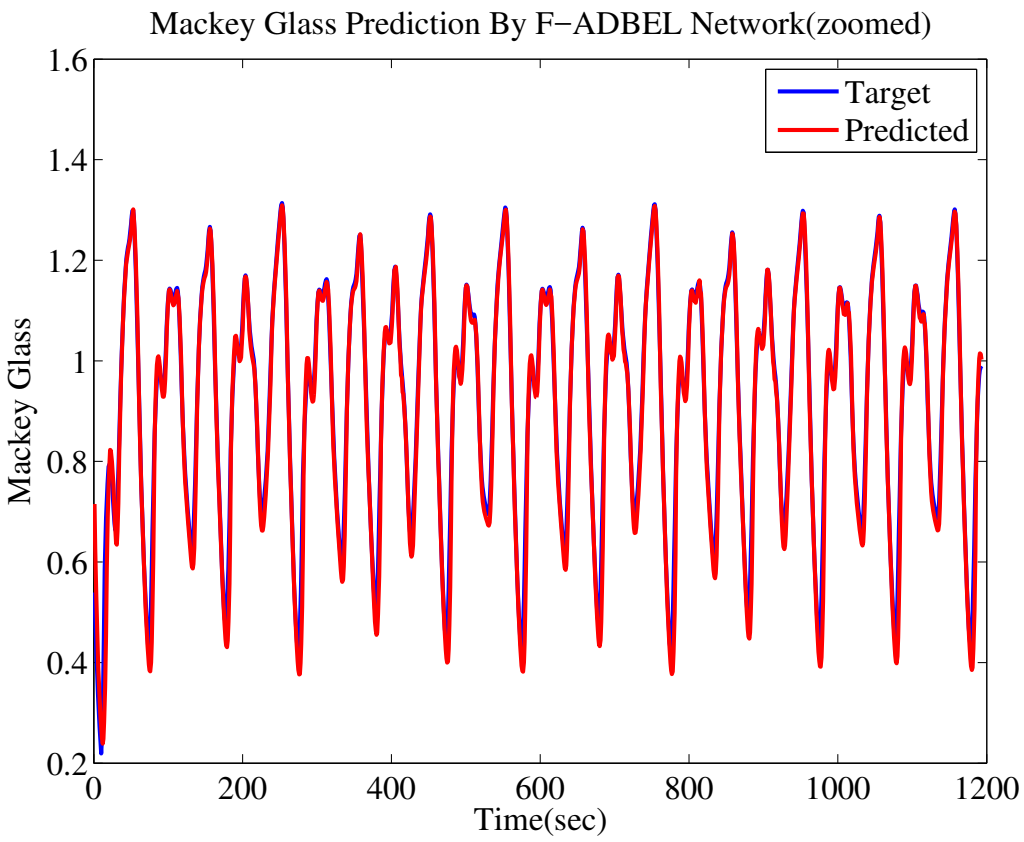

Figure 4.93: Mackey-Glass as Predicted by the Proposed F-ADBEL Network.

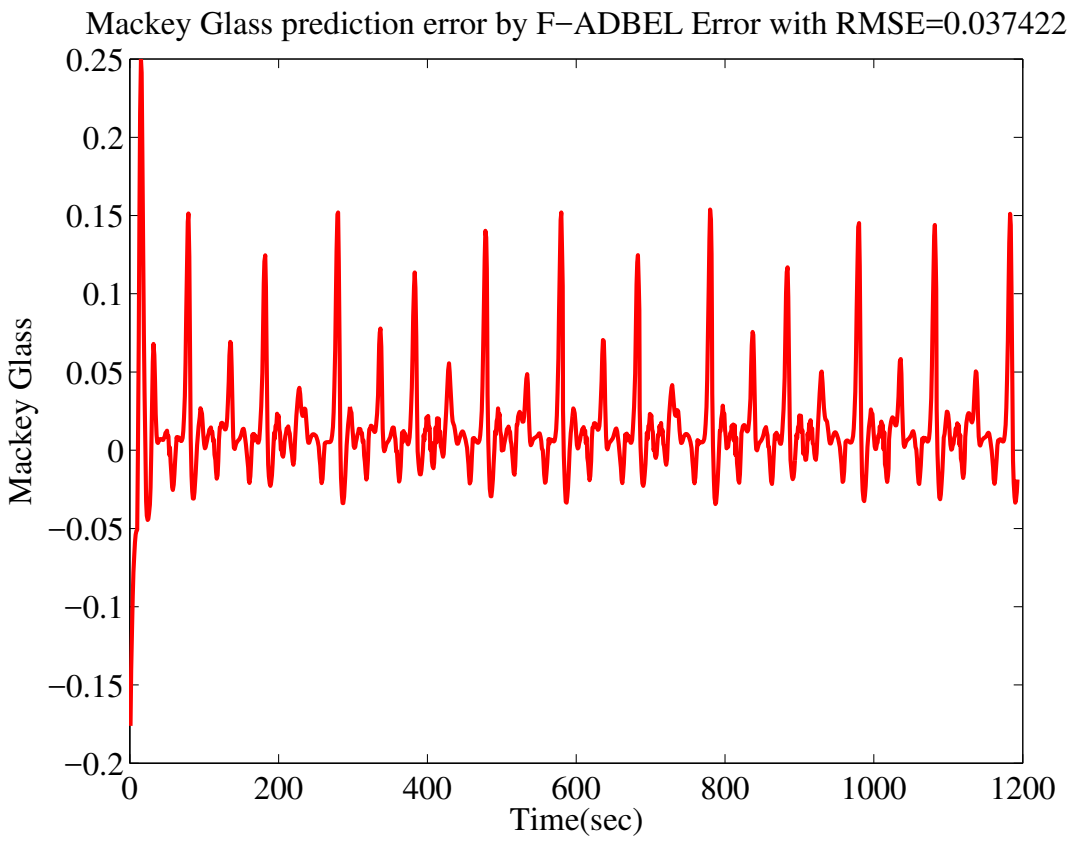

Figure 4.94: Error in Predicting Mackey-Glass by the Proposed F-ADBEL Network.

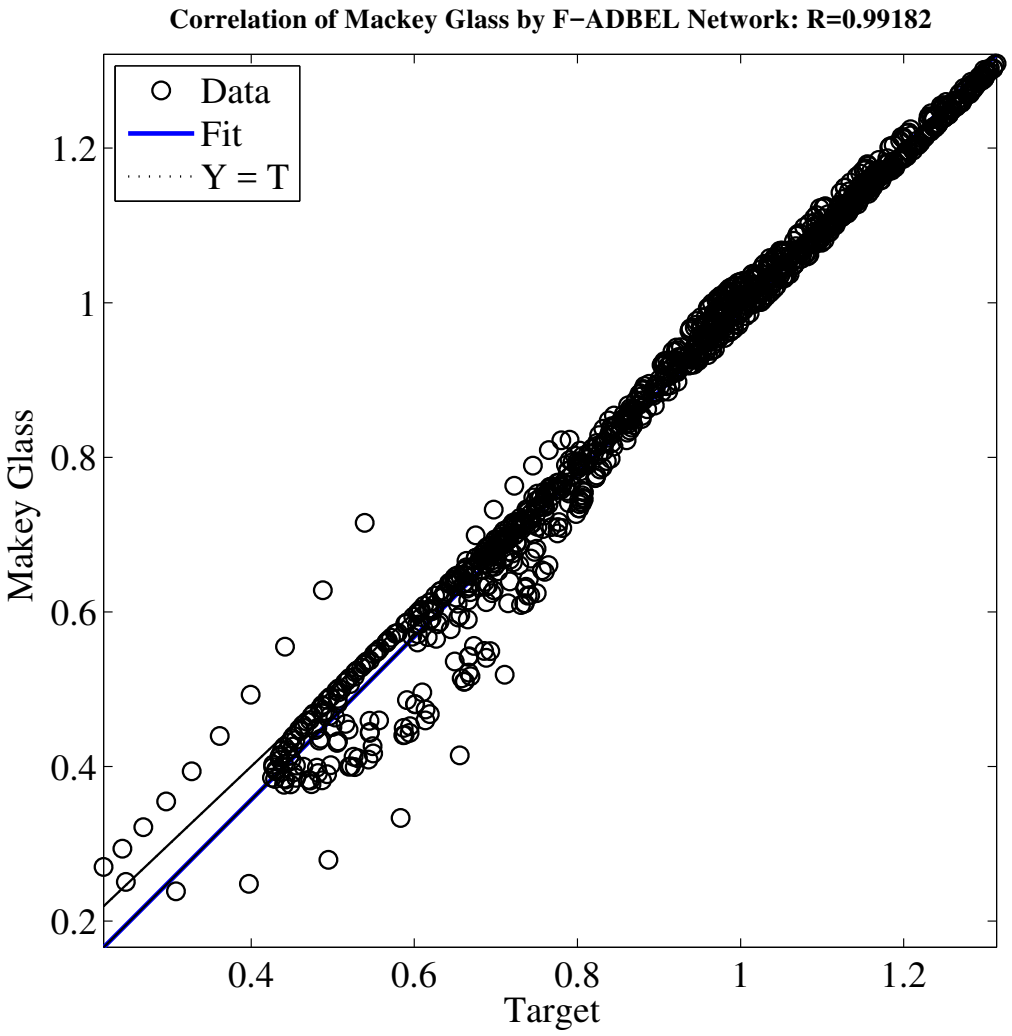

Figure 4.95: Correlation in Predicting Mackey-Glass by the Proposed F-ADBEL Network.

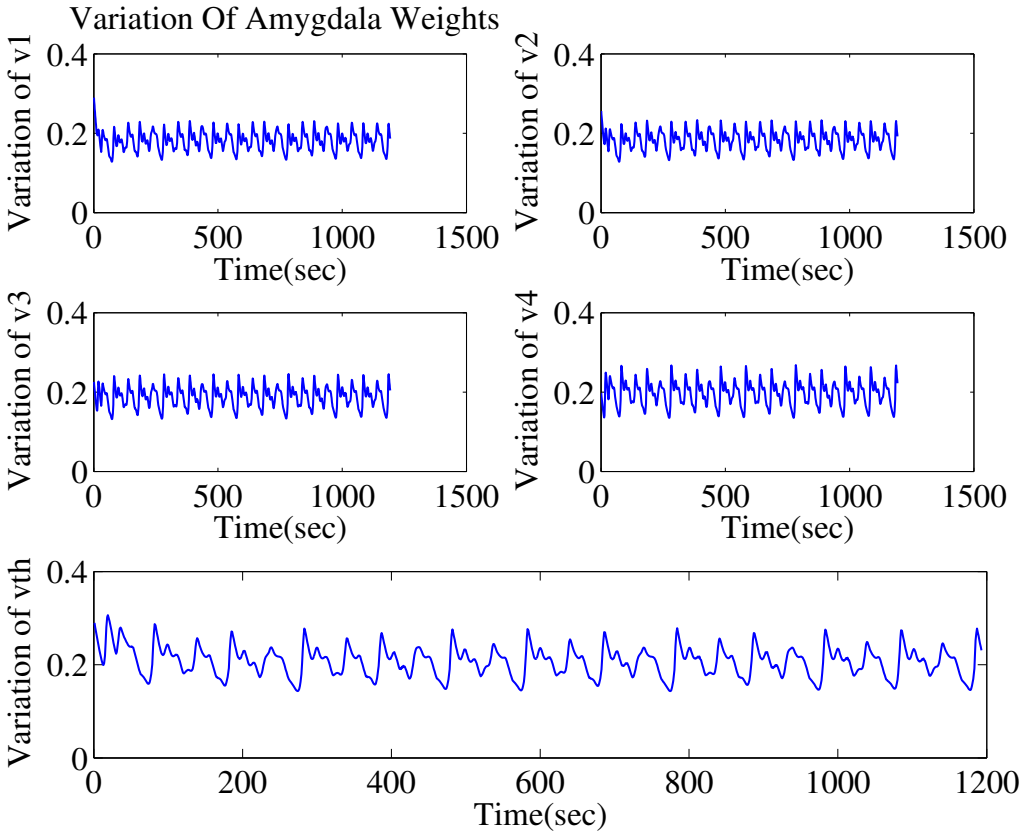

Figure 4.96: Variations in Amygdala Weights During Mackey-Glass Predictions by the Proposed F-ADBEL Network

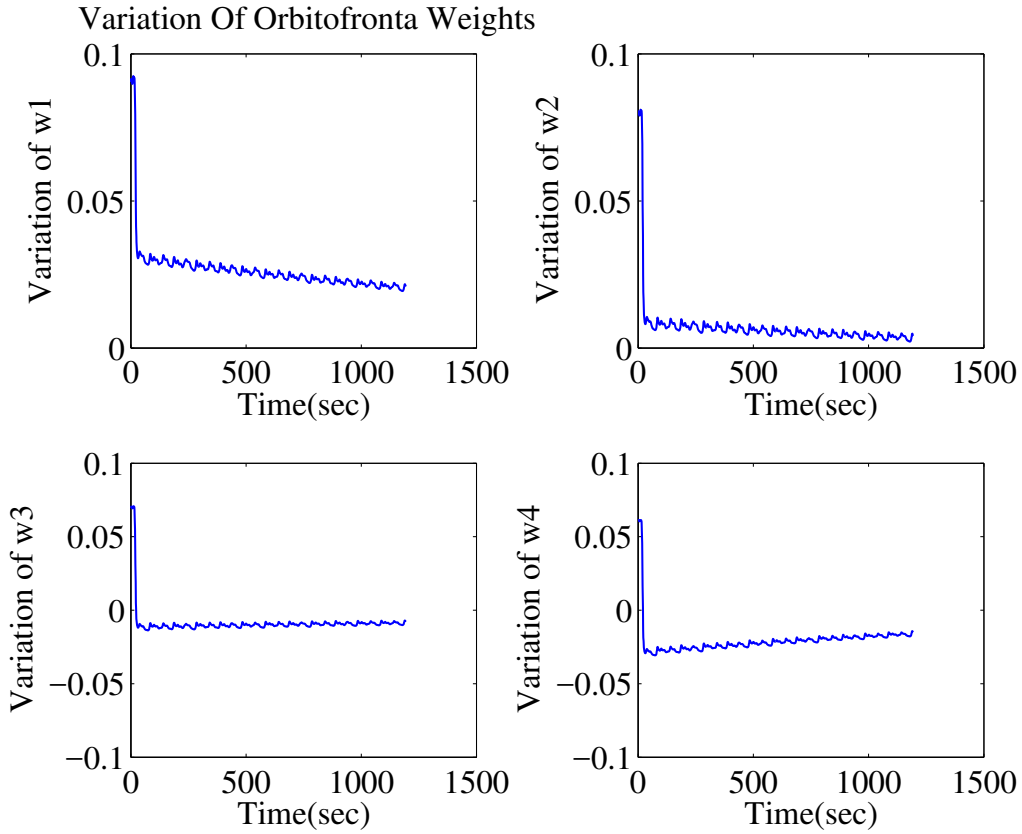

Figure 4.97: Variations in Orbitofrontal Cortex Weights During Mackey-Glass Predictions by the Proposed F-ADBEL Network.

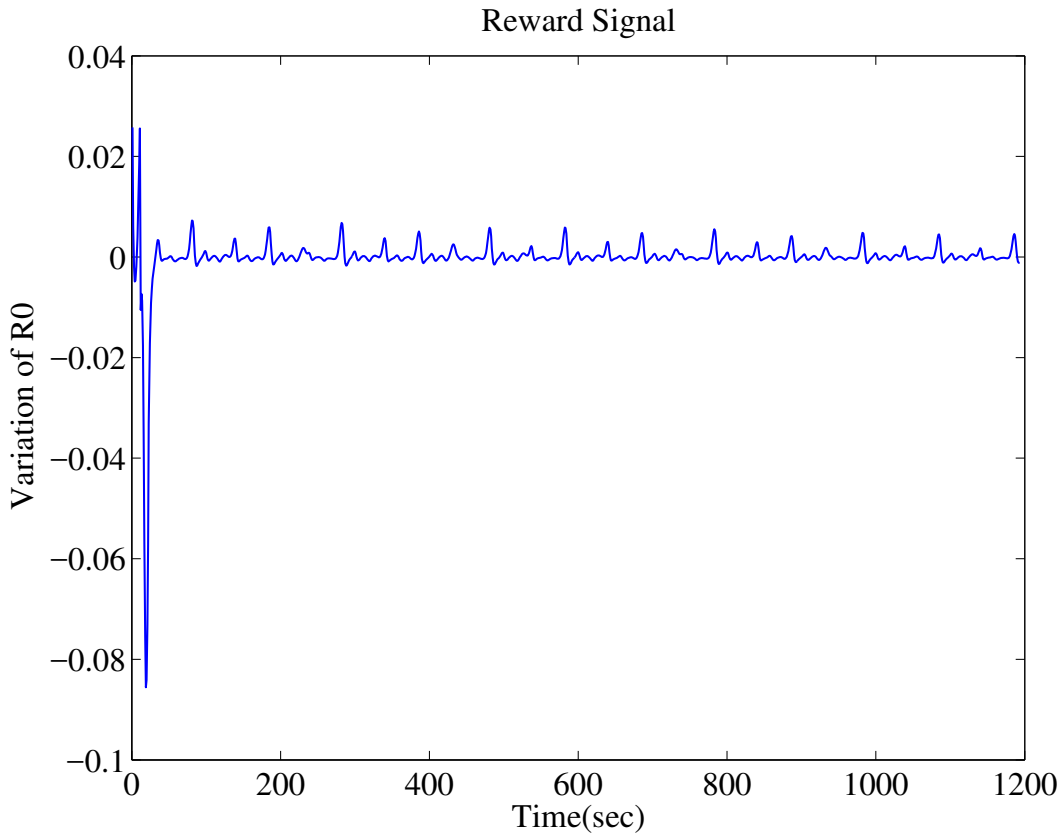

Figure 4.98: Variations in Reward Signal During Mackey-Glass Predictions by the Proposed F-ADBEL Network.

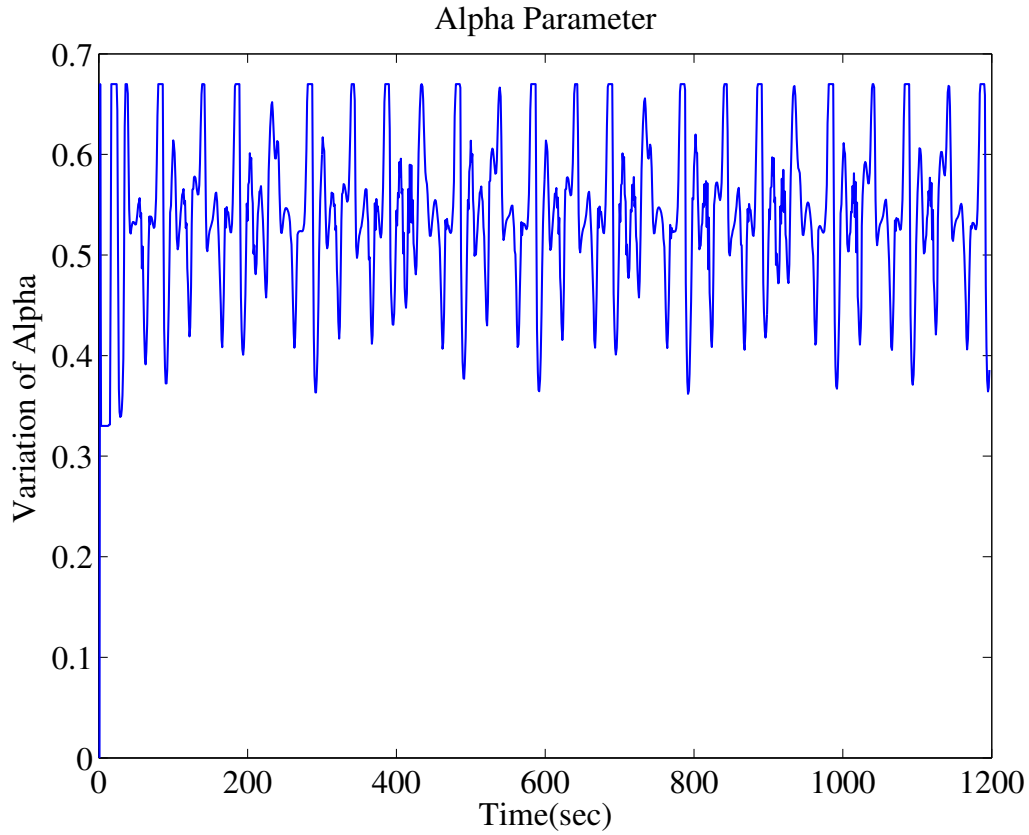

Figure 4.99: Variations in Alpha Parameter During Mackey-Glass Predictions by the Proposed F-ADBEL Network.

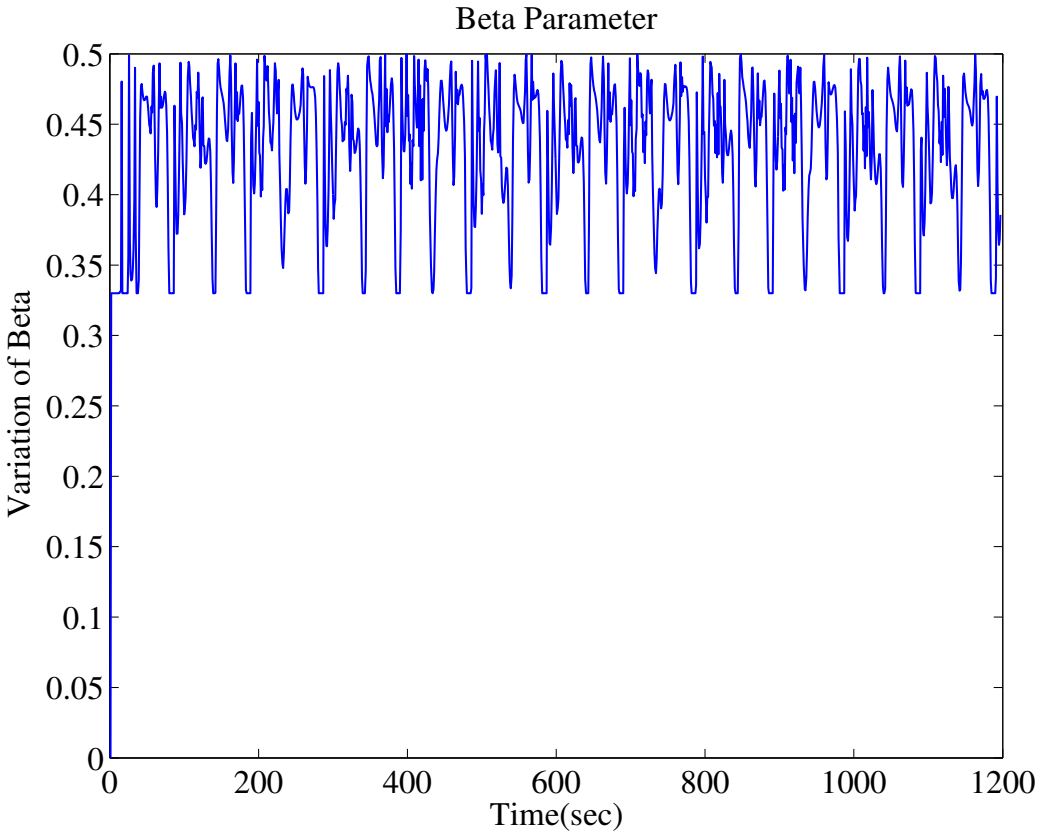

Figure 4.100: Variations in Beta Parameter During Mackey-Glass Prediction by the Proposed F-ADBEL Network.

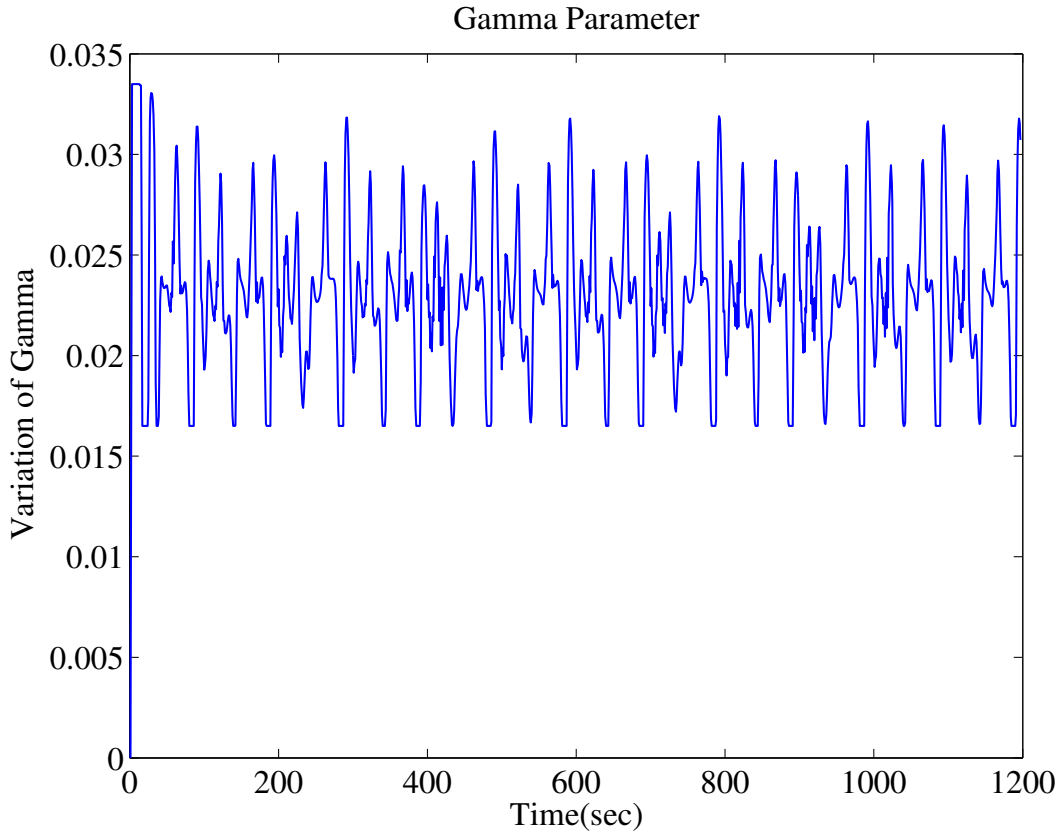

Figure 4.101: Variations in Gamma Parameter During Mackey-Glass Prediction by the Proposed F-ADBEL Network.

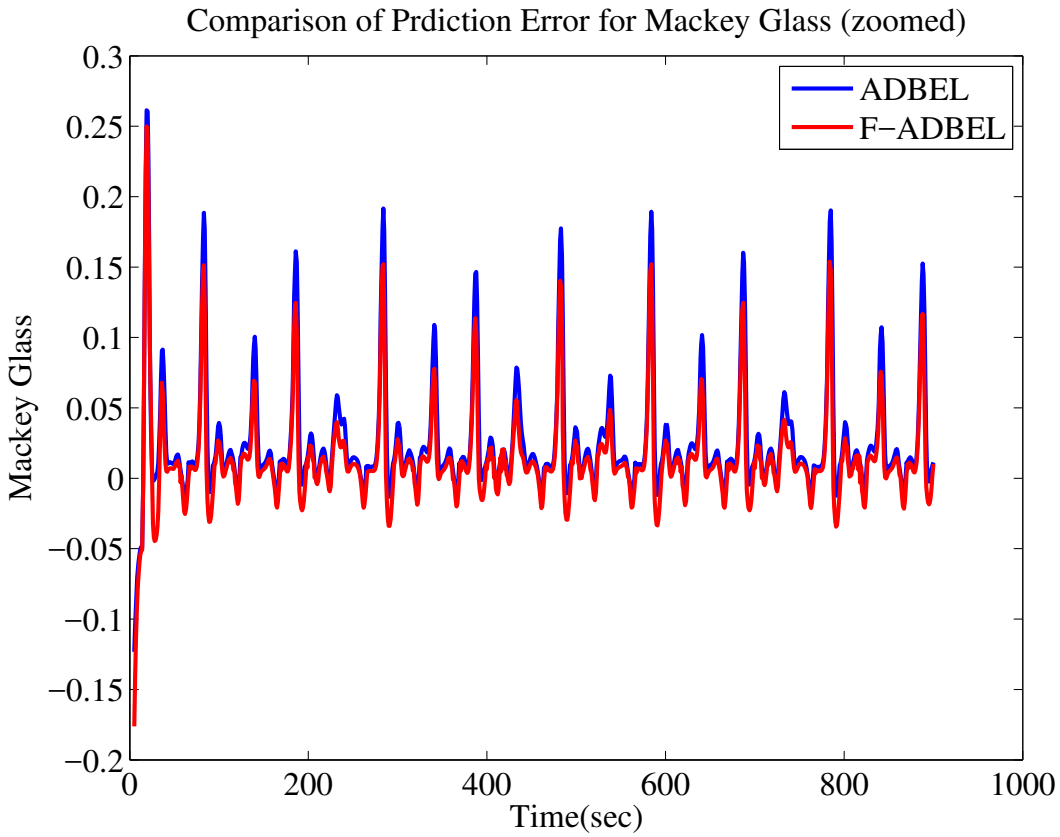

Figure 4.102: Comparison of Errors in Predicting Mackey-Glass by the ADBEL and Proposed F-ADBEL Networks.

4.2.2 Lorenz Chaotic Time Series Predicted by the Proposed F-ADBEL Network

F-ADBEL is also simulated to predict the x-dynamics of the Lorenz chaotic time series generated from the coupled differential Eq. (4.5). The proposed F-ADBEL network is deployed to predict the Lorenz time series, as shown in Figures 4.103, 4.104 and 4.105, in terms of root mean squared error and correlation coefficient. The results are presented in Table 4.7, and the variations in the amygdala weights and orbitofrontal cortex weights are presented in Figures 4.106 and 4.107, respectively.

Figure 4.107 illustrates that the orbitofrontal cortex weights converge after the initial transient period. On the other hand, the amygdala weights, as shown in Figure 4.106, display more significant variations than the orbitofrontal cortex weights in steady-state. This supports the hypothesis that the orbitofrontal cortex is the more stable portion of the emotional brain and can correct the response of the amygdala, which reacts quickly to emotional stimuli. Note that the amygdala weights are lower-bounded by zero, as was discussed during the development of the fuzzy parameter adjuster in section 3.7.

Variations in the network parameters α, β and γ produced by the proposed fuzzy model F-ADBEL are also shown in Figures 4.109, 4.110 and 4.111, respectively. As can be seen, the variations in the parameter β are minimal after the first few time-series samples, which can also be used to describe the changes in orbitofrontal cortex weights. On the other hand, the fuzzy model continues adjusting the different network parameters (α and γ) throughout the entire horizon, which also explains the variations in the amygdala weights.

Figure 4.108 shows the interpretation of the reward signal, which could become either positive or negative in value and support the concept discussed during the fuzzy parameter adjuster development in section 3.7. The same time series is also predicted with the ADBEL network, using the learning parameters: $\alpha = 0.8$, $\beta = 0.2$, and $\gamma = 0.01$. As illustrated in Figure 4.112, the prediction errors are recorded in both cases, with analysis showing that the transient period remains the same: $\leq 5\ s$. Thus, the steady-state starting index is set to compute the performance indices in both cases.

The performances of the designed F-ADBEL and ADBEL networks for predicting

Lorenz are shown in Table 4.7. As can be seen, the redesigned ADBEL network performed better in prediction error in terms of root mean squared error and high correlation compared to the proposed F-ADBEL network for the Lorenz time series. This result is likely because adjusting the F-ADBEL network's parameters is not optimal. The run-time of the simulation recorded that the ADBEL network performed in 1 second, while the F-ADBEL performed in 23 seconds.

Table 4.7: RMSE /COR/PI for Lorenz Time Series Prediction by ADBEL and F-ADBEL Networks

Time Series	Prediction Network	RMSE	COR
Lorenz	ADBEL	0.55	0.9980
	F-ADBEL	0.85	0.9978

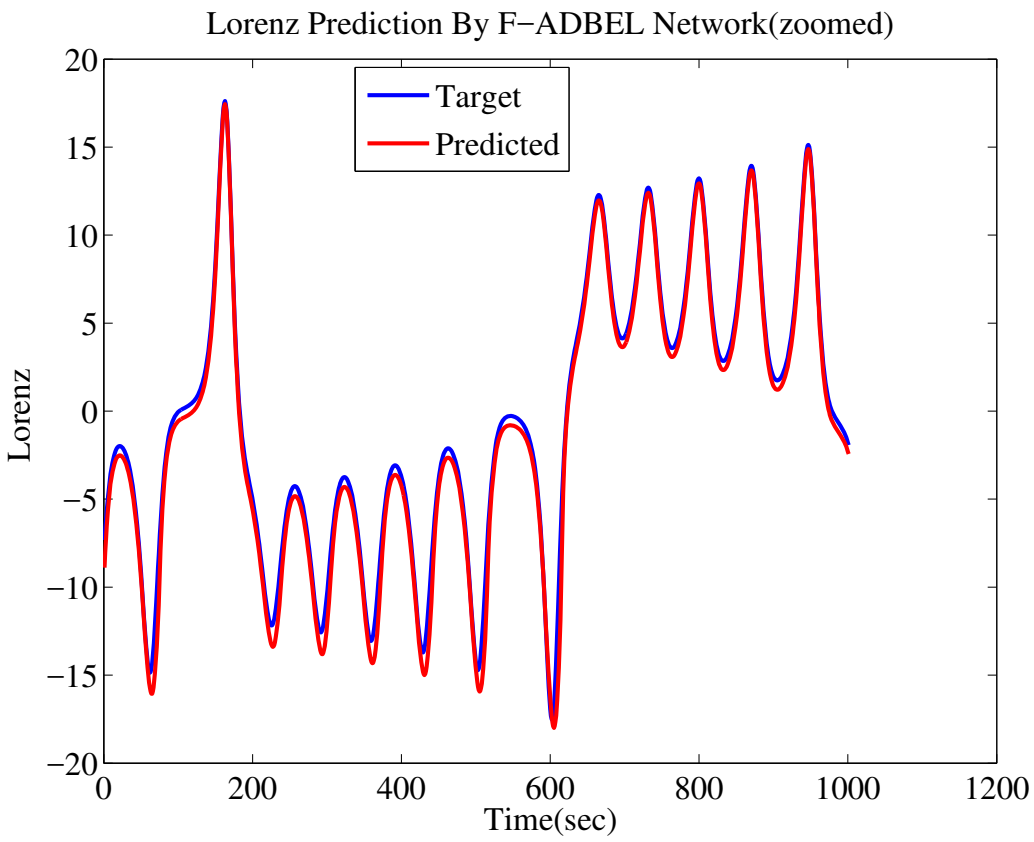

Figure 4.103: Lorenz Time Series as Predicted by the Proposed F-ADBEL Network.

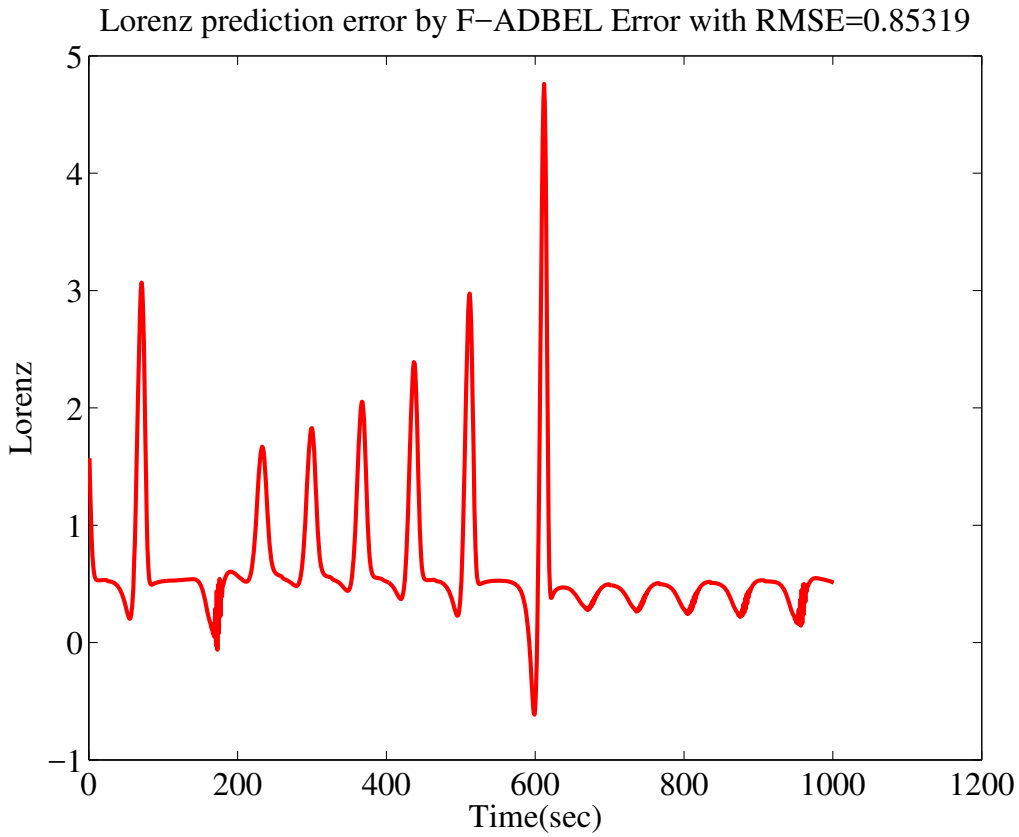

Figure 4.104: Error in Predicting Lorenz Time Series by the Proposed F-ADBEL Network.

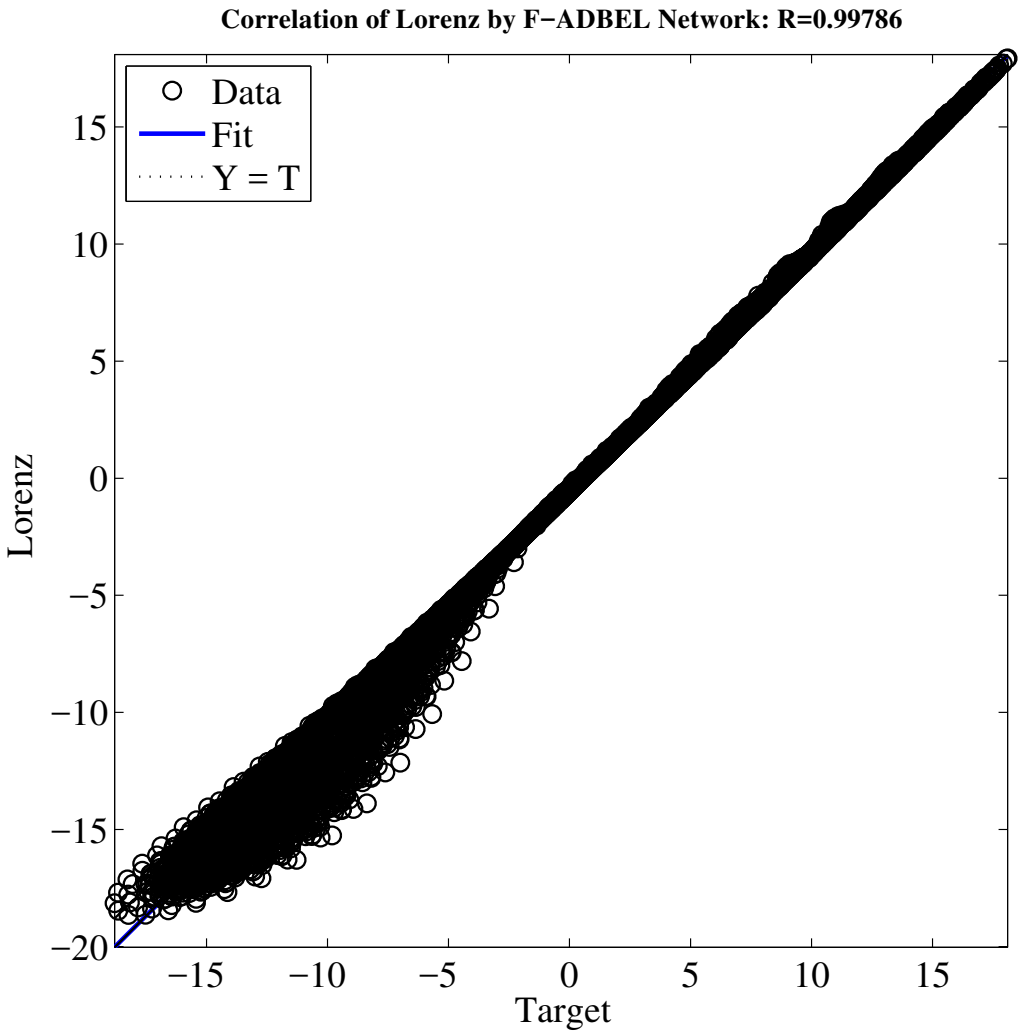

Figure 4.105: Correlation in Predicting Lorenz Time Series by the Proposed F-ADBEL Network.

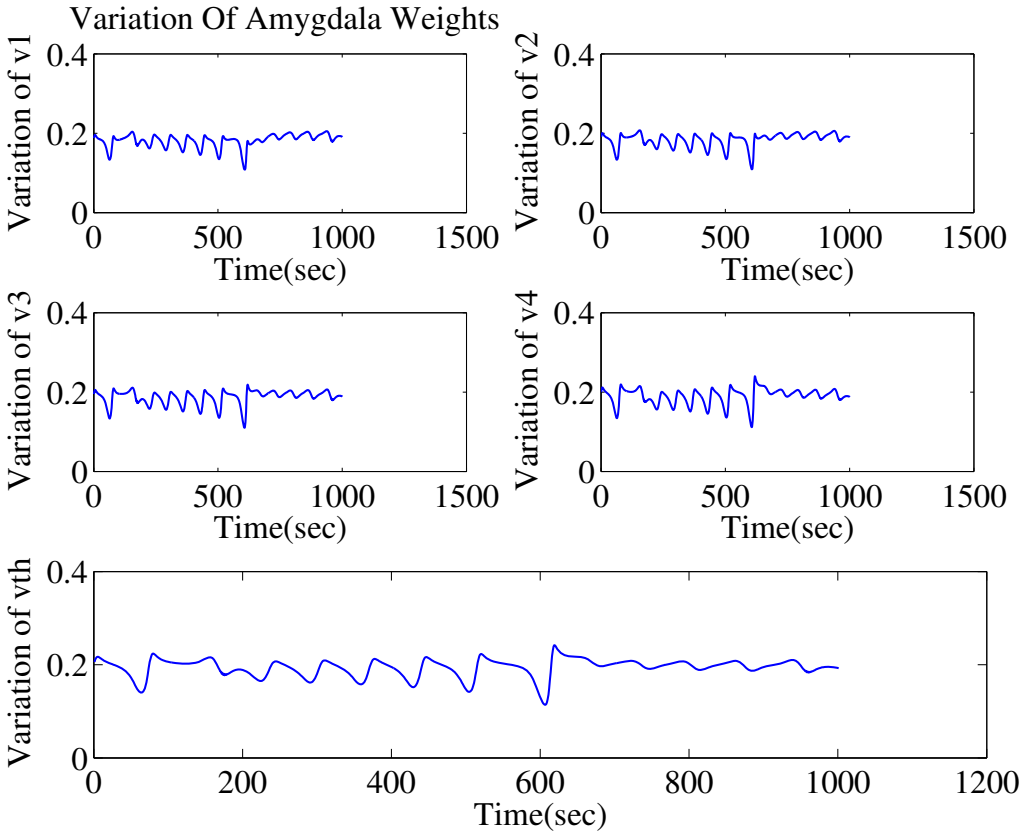

Figure 4.106: Variations in Amygdala Weights During Lorenz Time Series Prediction by the Proposed F-ADBEL Network.

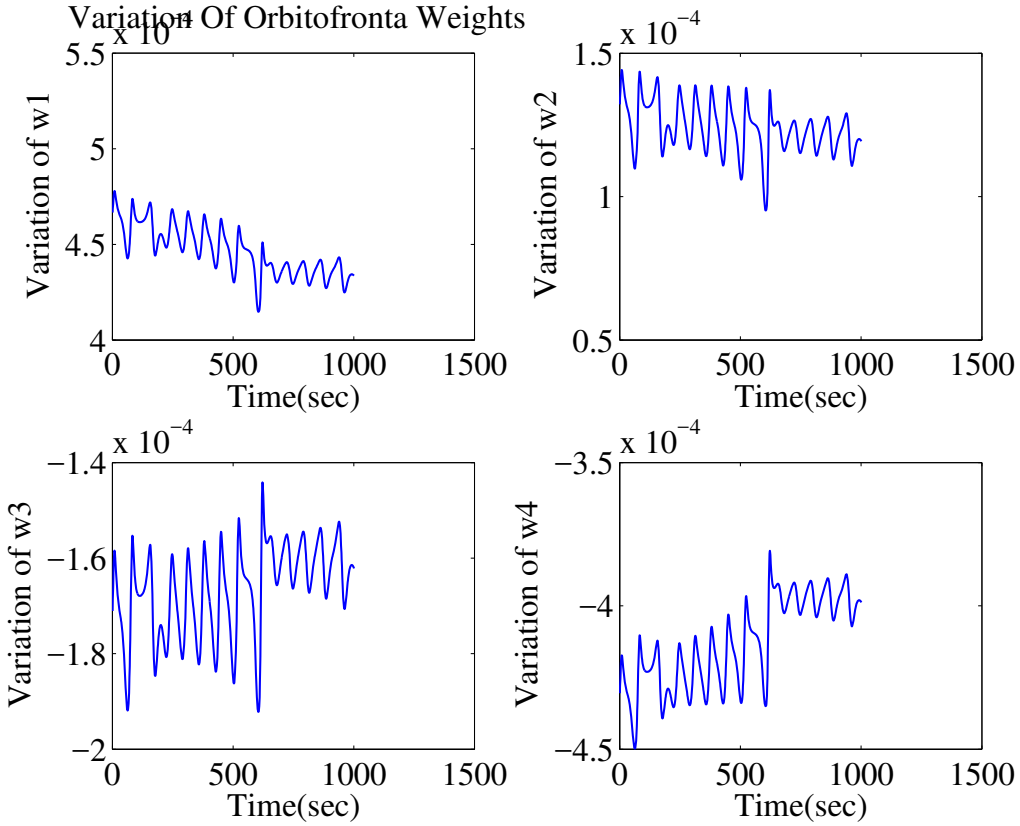

Figure 4.107: Variations in Orbitofrontal Cortex Weights During Lorenz Time Series Prediction by the Proposed F-ADBEL Network.

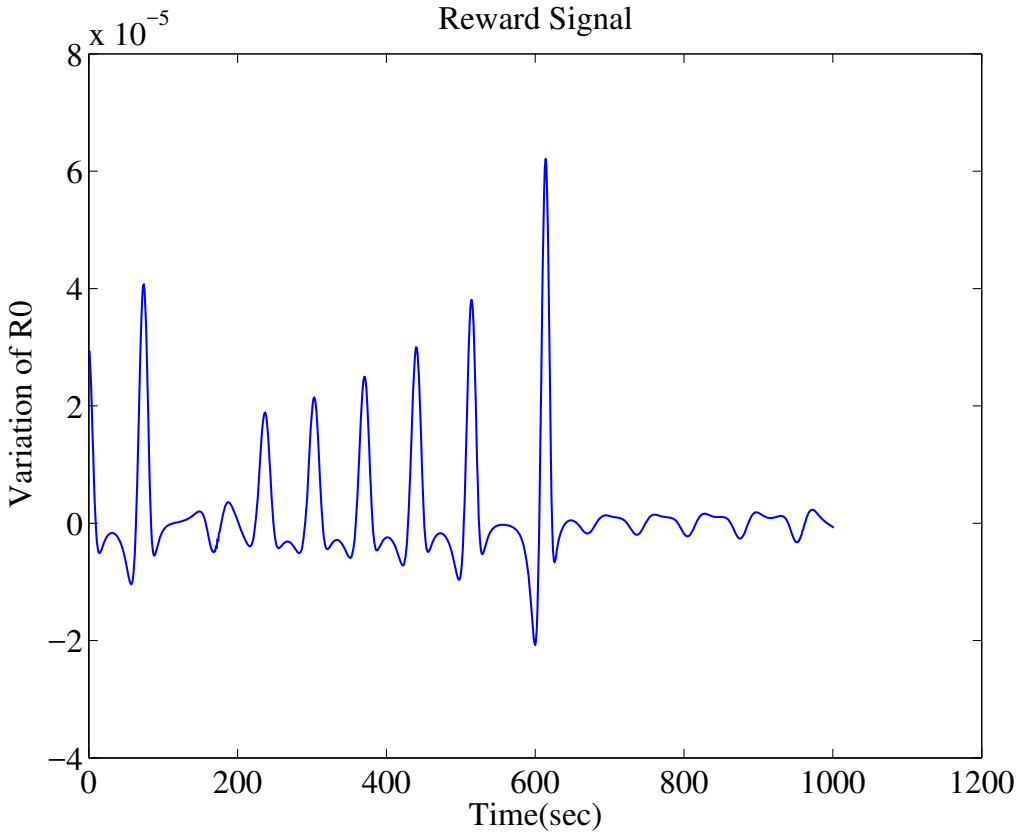

Figure 4.108: Variations in Reward Signal During Lorenz Time Series Prediction by the Proposed F-ADBEL Network.

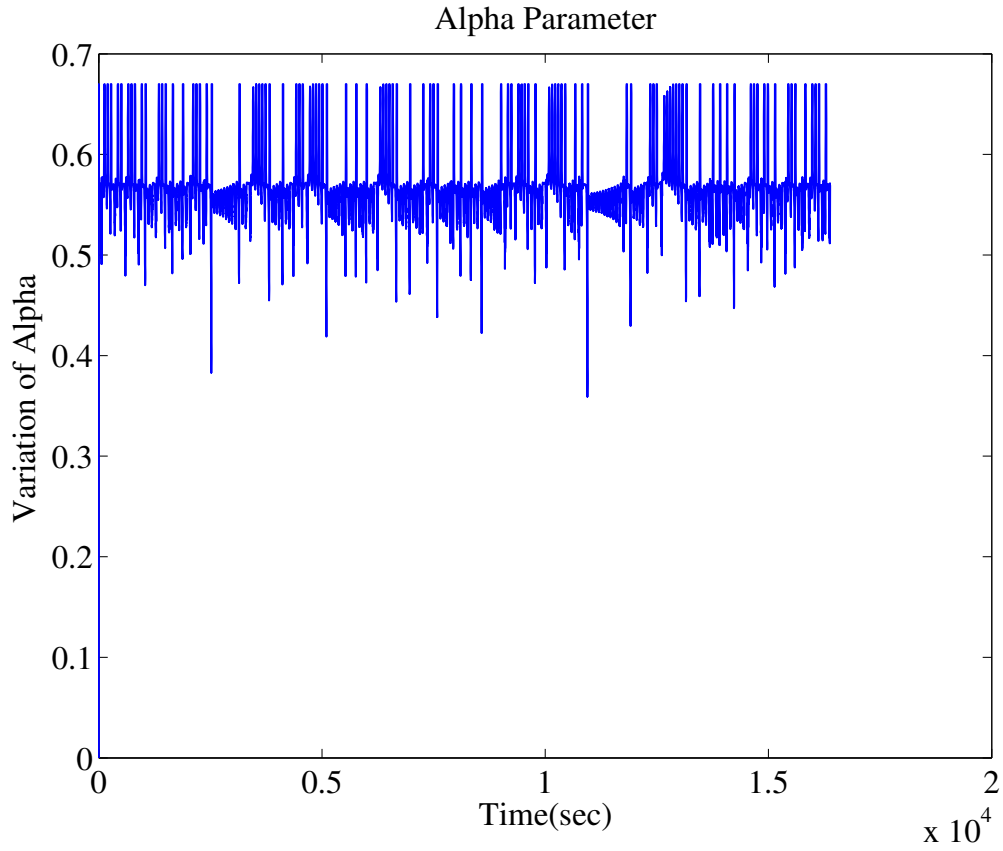

Figure 4.109: Variations in Alpha Parameter During Lorenz Time Series Prediction by the Proposed F-ADBEL Network.

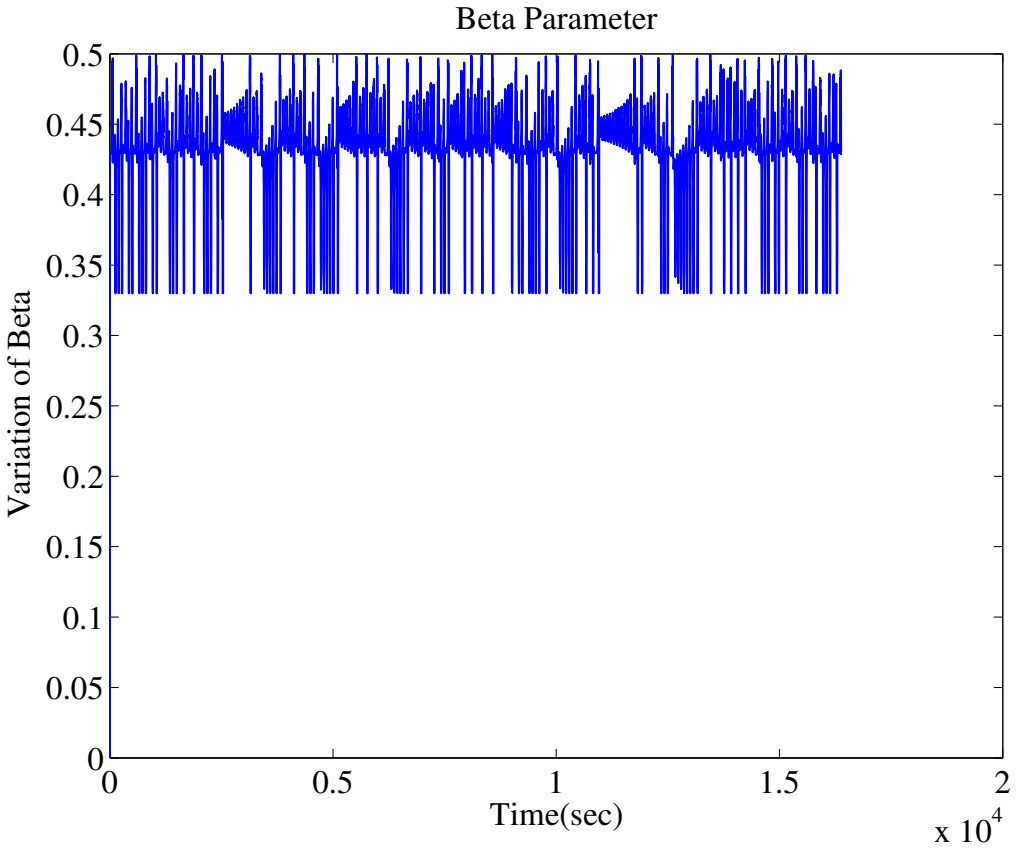

Figure 4.110: Variations in Beta Parameter During Lorenz Time Series Prediction by the Proposed F-ADBEL Network.

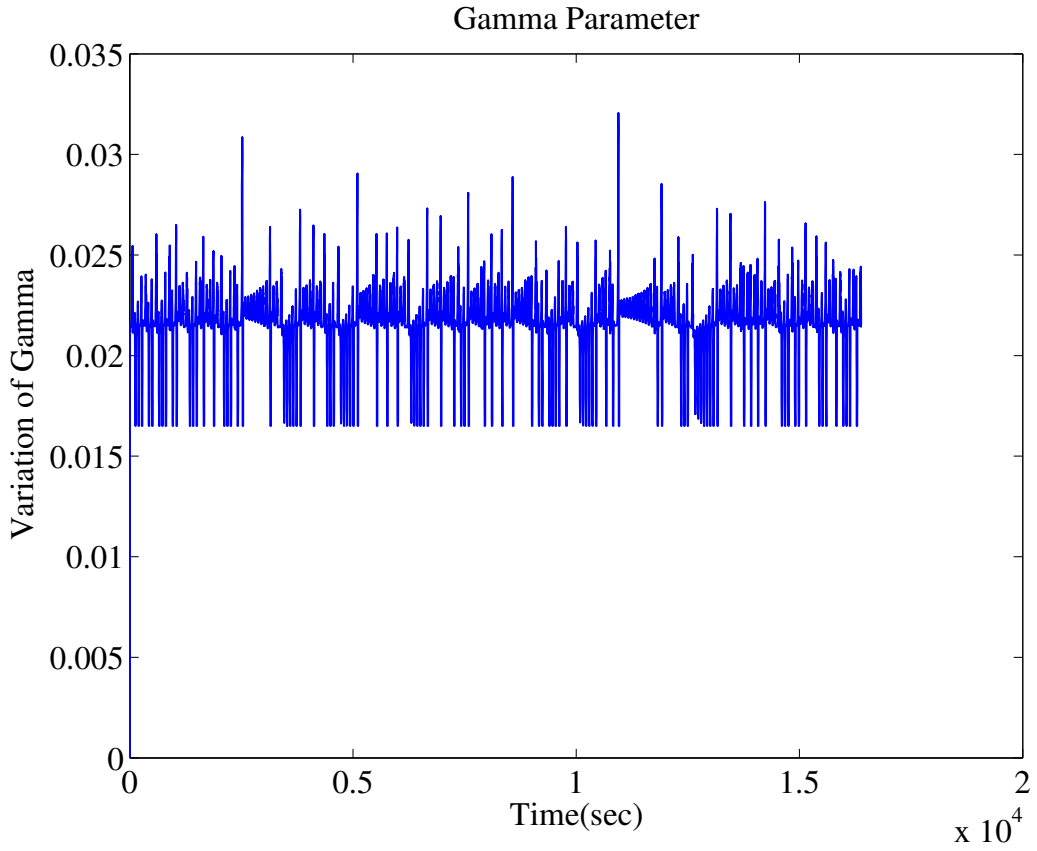

Figure 4.111: Variations in Gamma Parameter During Lorenz Time Series Prediction by the Proposed F-ADBEL Network.

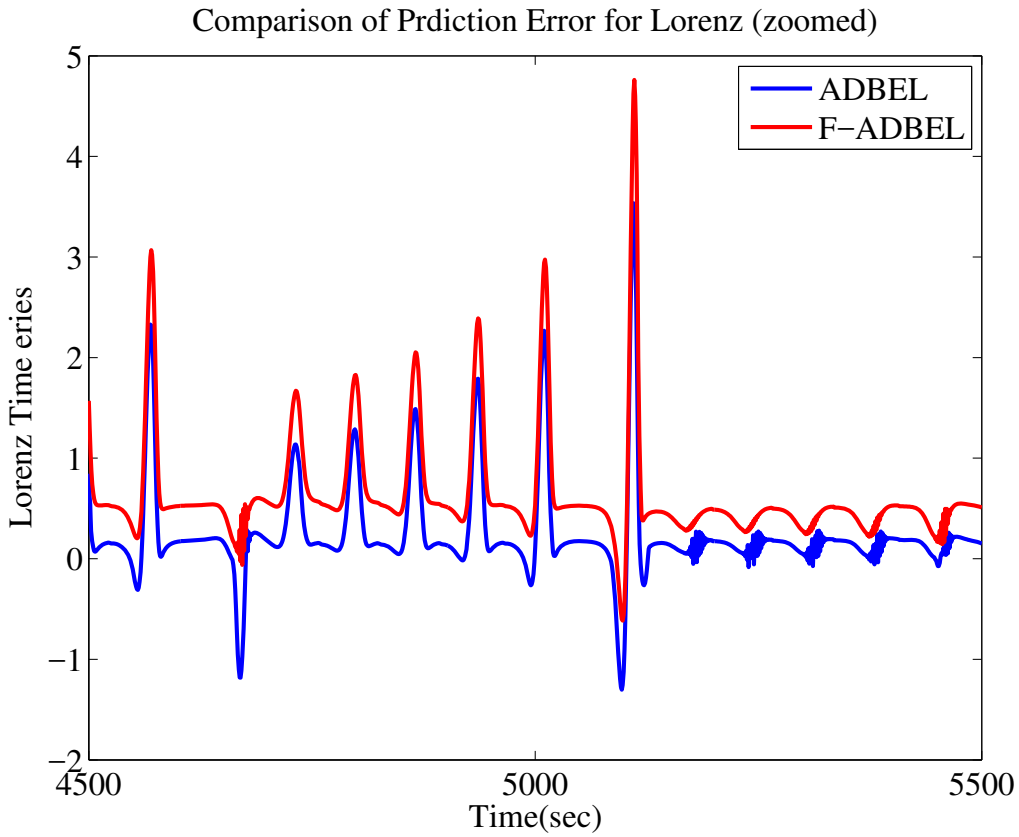

Figure 4.112: Comparison of Errors in Predicting Lorenz Time Series by ADBEL and Proposed F-ADBEL Networks.

4.2.3 Rossler Chaotic Time Series Predicted by the Proposed F-ADBEL Network

F-ADBEL is also simulated to predict the Rossler chaotic time series. The series is generated from the coupled differential Eq. (4.6). The proposed F-ADBEL network is deployed to predict the Rossler time series, as shown in Figures 4.113, 4.114 and 4.115, in terms of root mean squared error and correlation coefficient. The results are presented in Table 4.8, and thee variations in the amygdala weights and orbitofrontal cortex weights are presented in the zoomed Figures 4.116 and 4.117, respectively.

Figure 4.117 illustrates that the orbitofrontal cortex weights converge after the initial transient period. On the other hand, the amygdala weights, as shown in Figure 4.116, display almost the same variation in orbitofrontal cortex weights in steady-state. This supports the hypothesis that the orbitofrontal cortex is the more stable portion of the emotional brain and can correct the amygdala's response, which reacts quickly to emotional stimuli. Note that the amygdala weights are lower-bounded by zero, as was discussed during the development of the fuzzy parameter adjuster in section 3.7. The variations in the network parameters α, β, and γ produced by the proposed fuzzy model F-ADBEL are also shown in Figures 4.119, 4.120 and 4.121, respectively. As can be observed, the variations in the parameter β are minimal after the first few time-series samples, which can also be used to describe the changes in orbitofrontal cortex weights.

On the other hand, the fuzzy model continues adjusting the different network parameters (α and γ) throughout the entire horizon, which also explains the variation in the amygdala weights. Figure 4.118 shows the interpretation of the reward signal, which could become positive or negative in value and support the concept discussed during the fuzzy parameter adjuster development in section 3.7.

The same time series is also predicted with the ADBEL network using the learning parameters: $\alpha = 0.8$, $\beta = 0.2$, and $\gamma = 0.05$. As shown in Figure 4.122, the prediction errors are recorded in both cases, with analysis indicating that the transient period remains the same: $\leq 5\ s$. Thus, the steady-state starting index is set to compute the performance indices in both cases.

The performances of the designed F-ADBEL and ADBEL networks for predicting Rossler are presented in Table 4.8. As can be seen, the designed F-ADBEL network

had a better prediction error in terms of root mean squared error and high correlation compared to the ADBEL network in the case of the Rossler time series. Moreover, the simulation's run-time also recorded that the ADBEL network performed in 0.28 seconds, while F-ADBEL performed in 12 seconds.

Table 4.8: RMSE /COR/PI for Rossler Time Series Prediction by the ADBEL and F-ADBEL Networks

Time Series	Prediction Network	RMSE	COR	PI(%)
Rossler	ADBEL	1.5	0.9922	
	F-ADBEL	1.24	0.9926	**17.33**

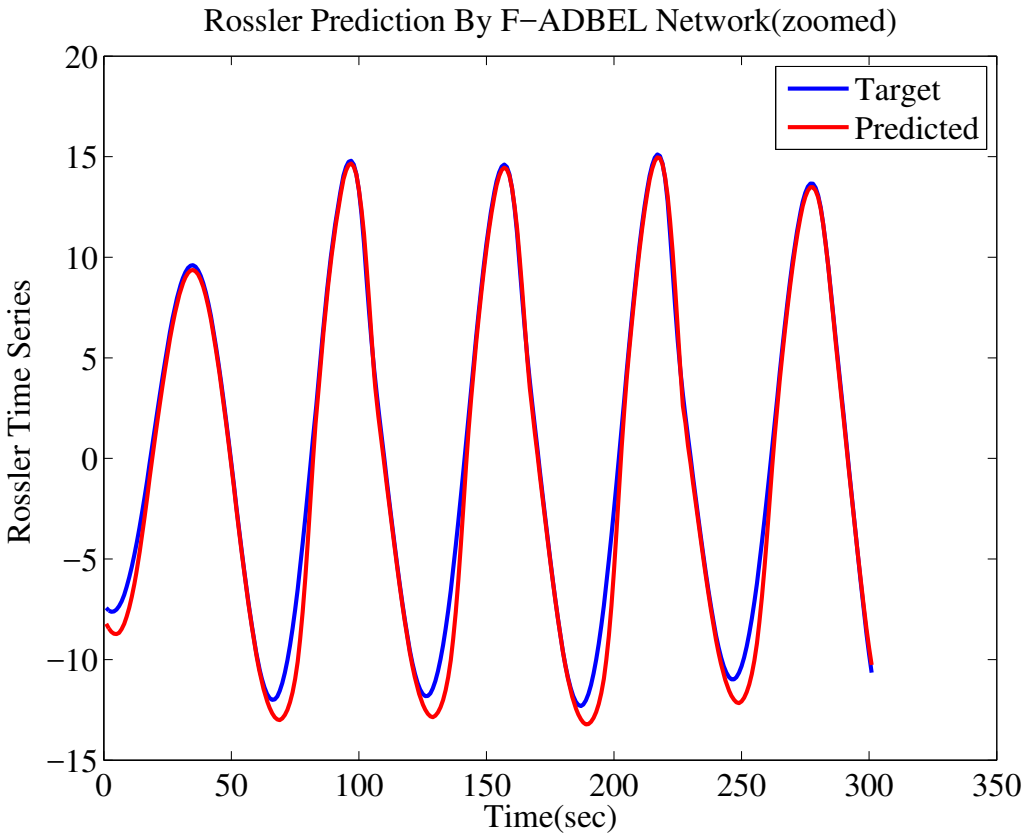

Figure 4.113: Rossler Time Series as Predicted by the Proposed F-ADBEL Network.

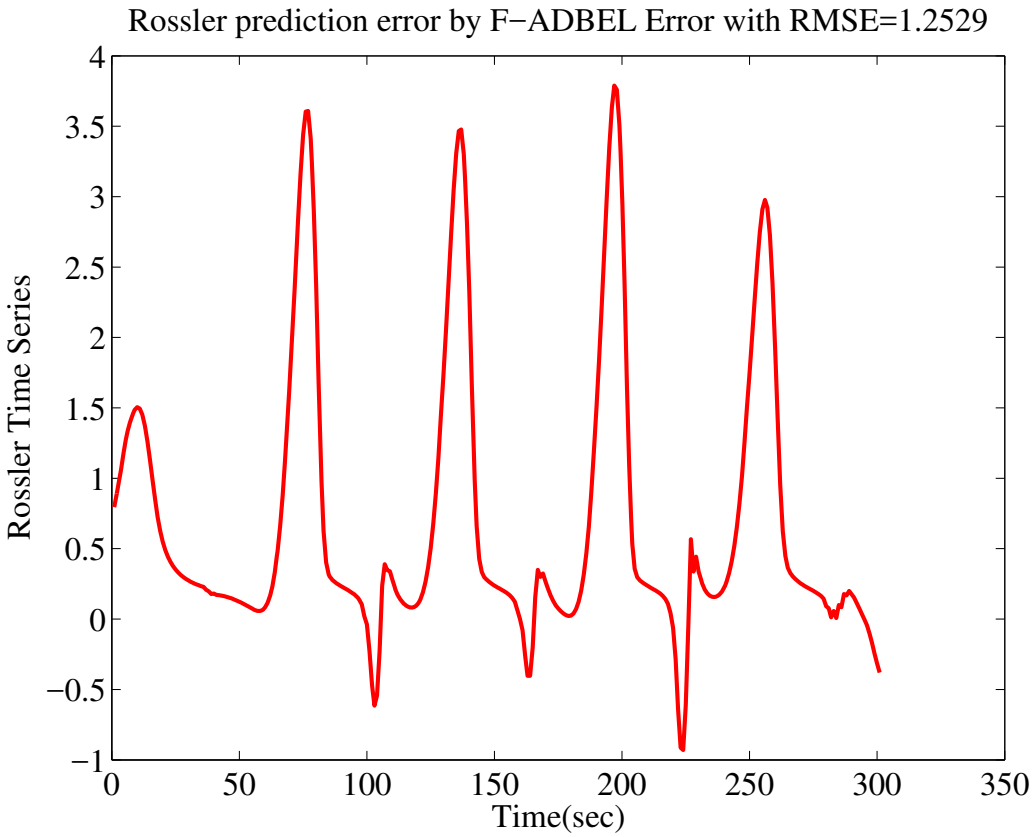

Figure 4.114: Error in Predicting Rossler Time Series by the Proposed F-ADBEL Network.

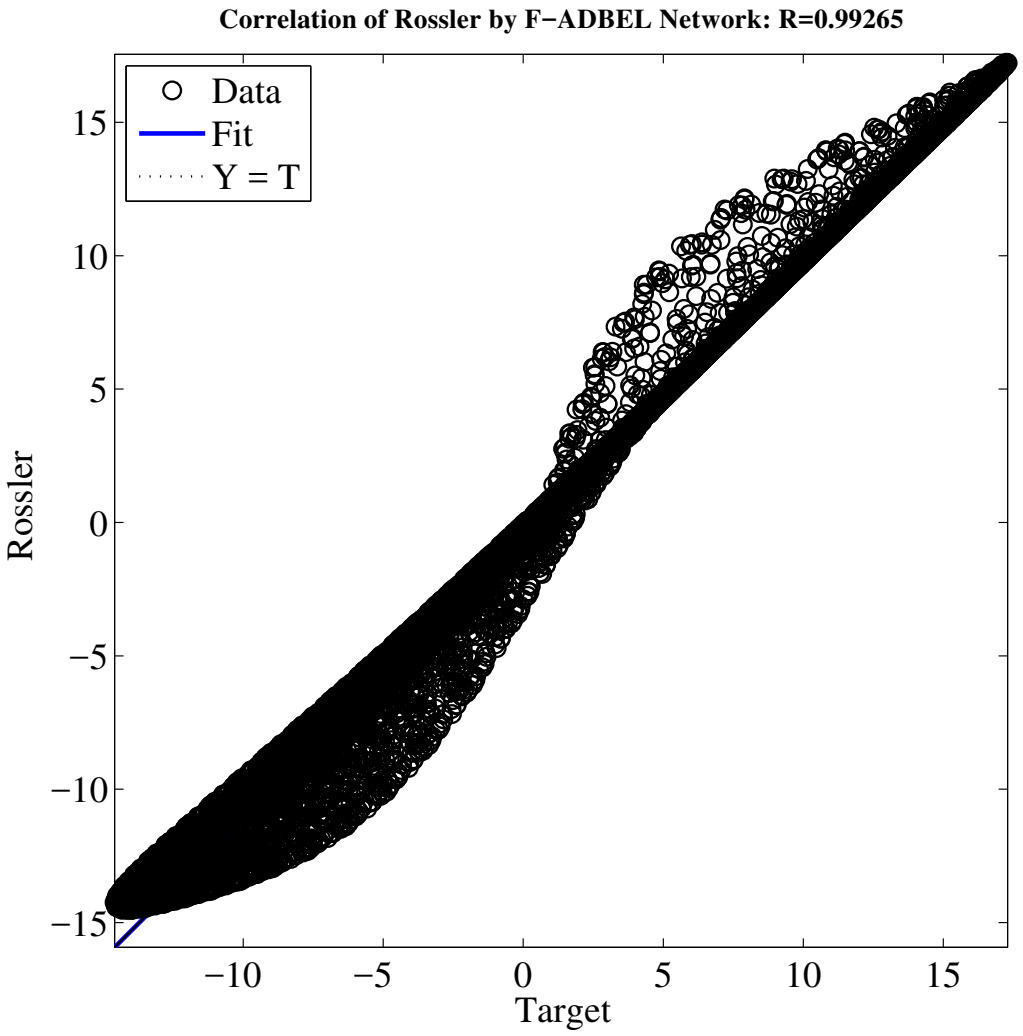

Figure 4.115: Correlation in Predicting Rossler Time Series by the Proposed F-ADBEL Network.

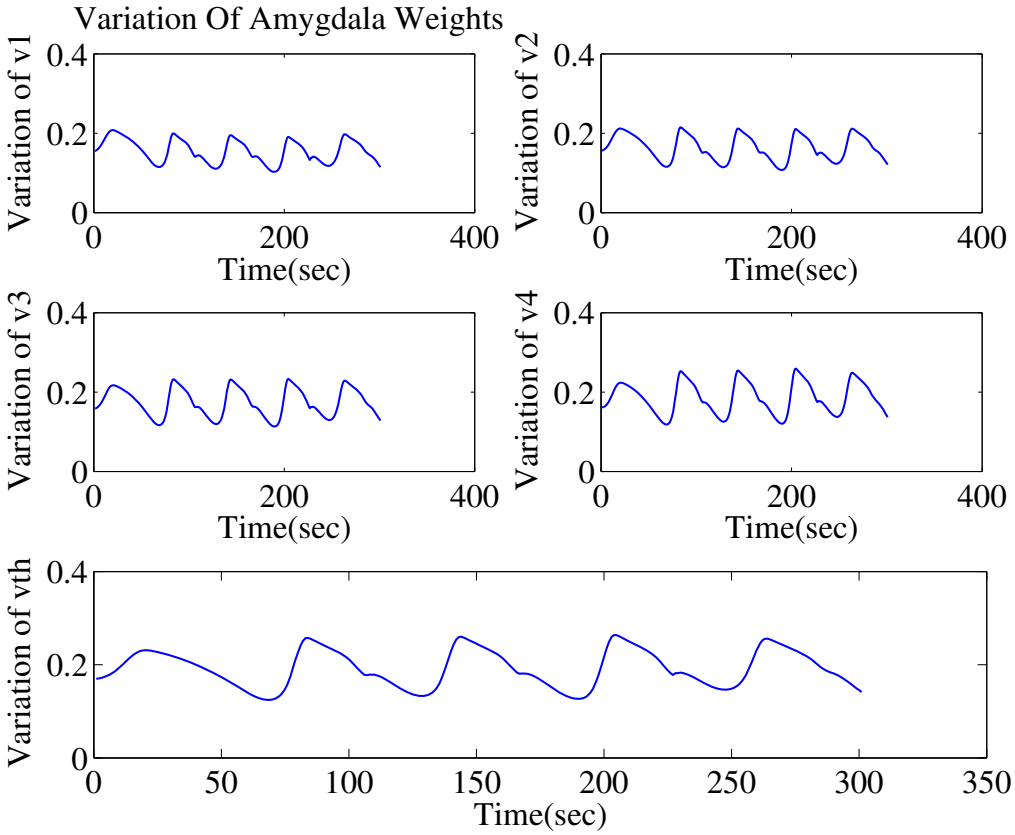

Figure 4.116: Variations in Amygdala Weights During Rossler Time Series Prediction by the Proposed F-ADBEL Network.

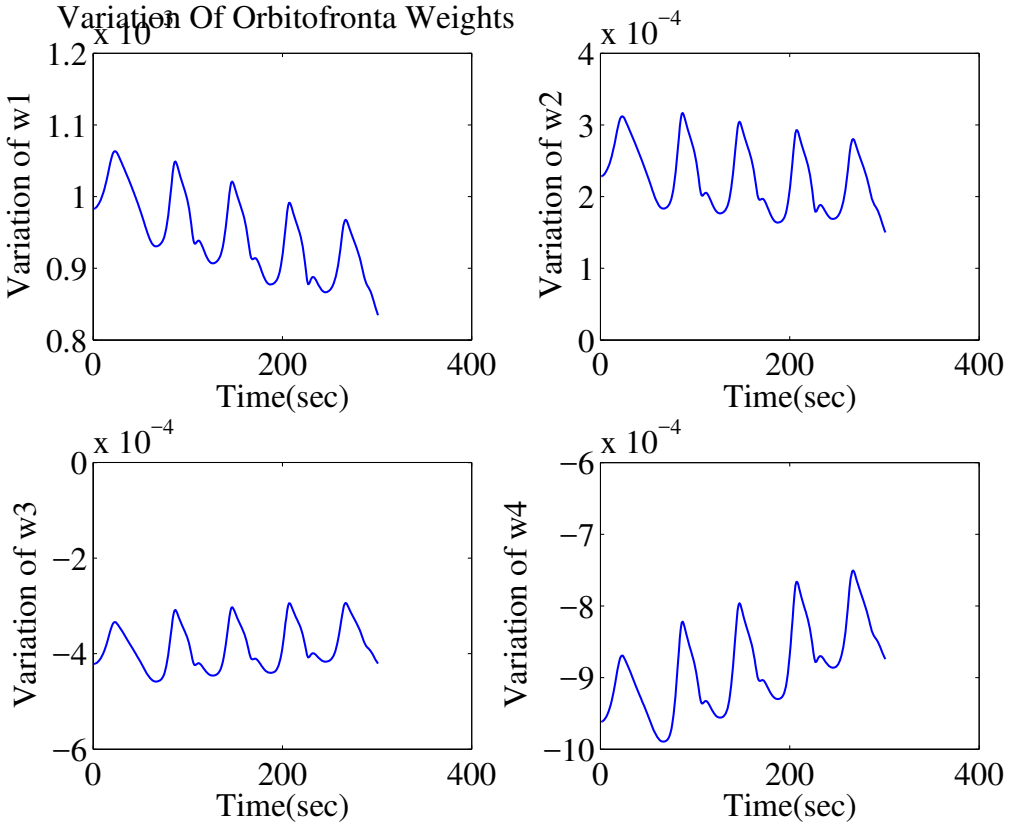

Figure 4.117: Variations in Orbitofrontal Cortex Weights During Rossler Time Series Prediction by the Proposed F-ADBEL Network.

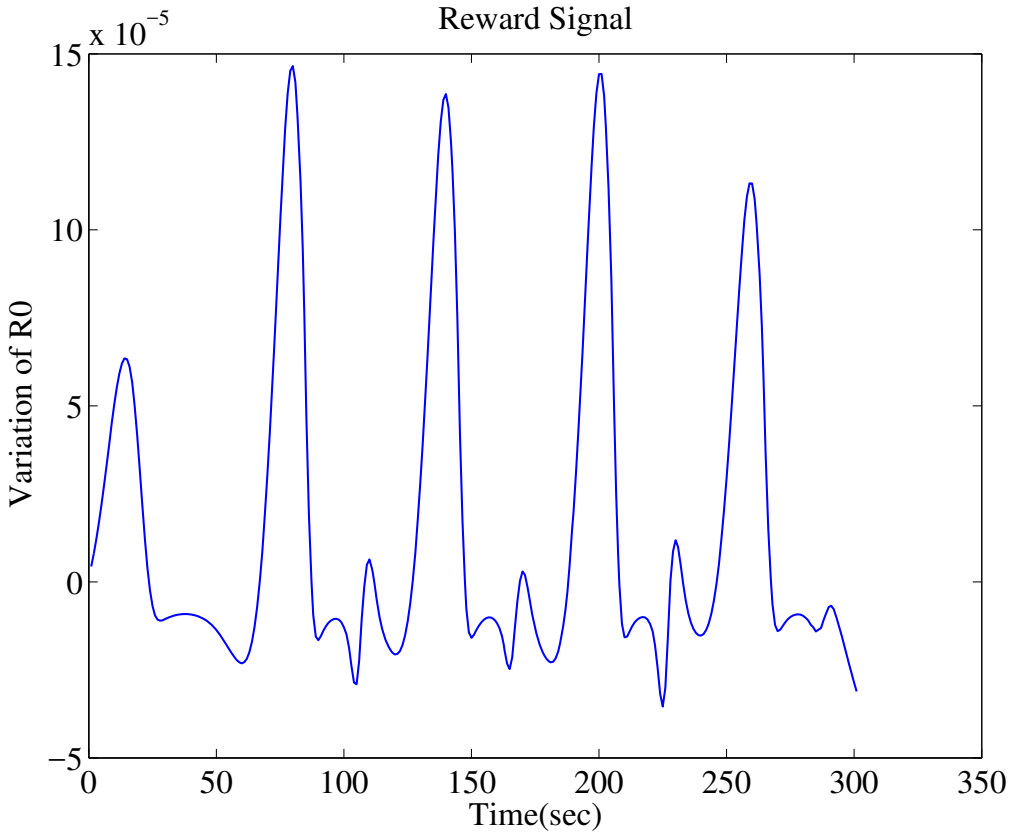

Figure 4.118: Variations in Reward Signal During Rossler Time Series Prediction by the Proposed F-ADBEL Network.

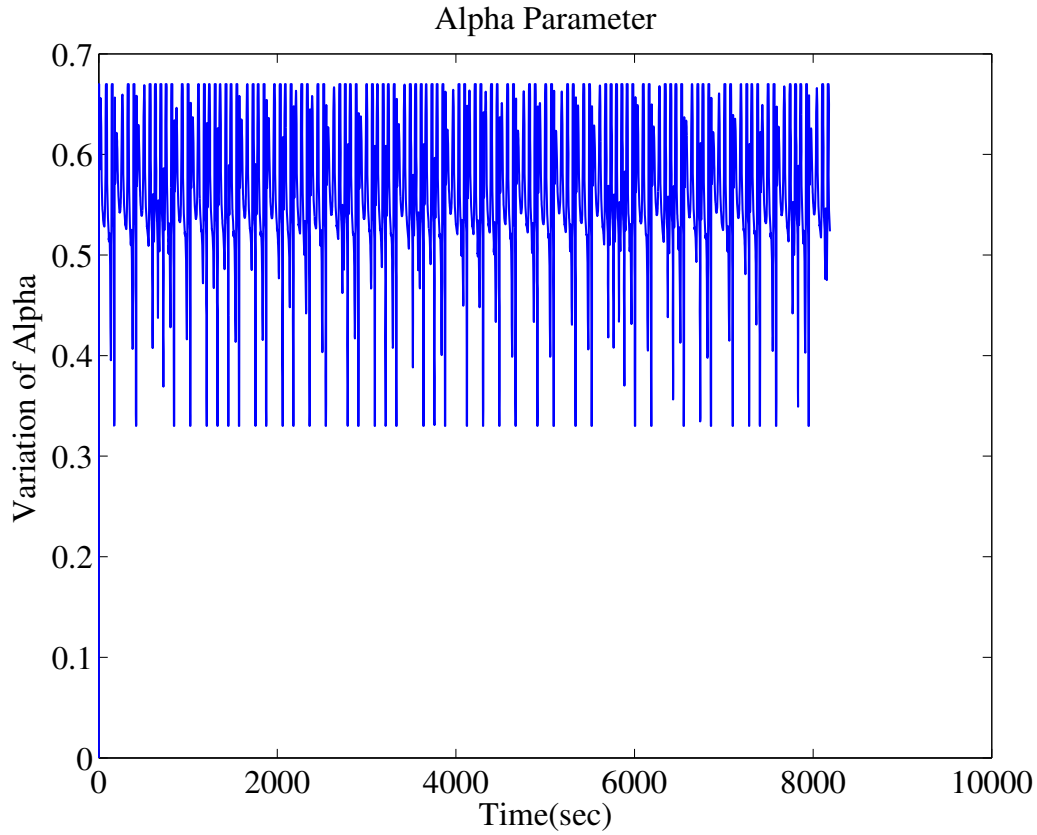

Figure 4.119: Variations in Alpha Parameter During Rossler Time Series Prediction by the Proposed F-ADBEL Network.

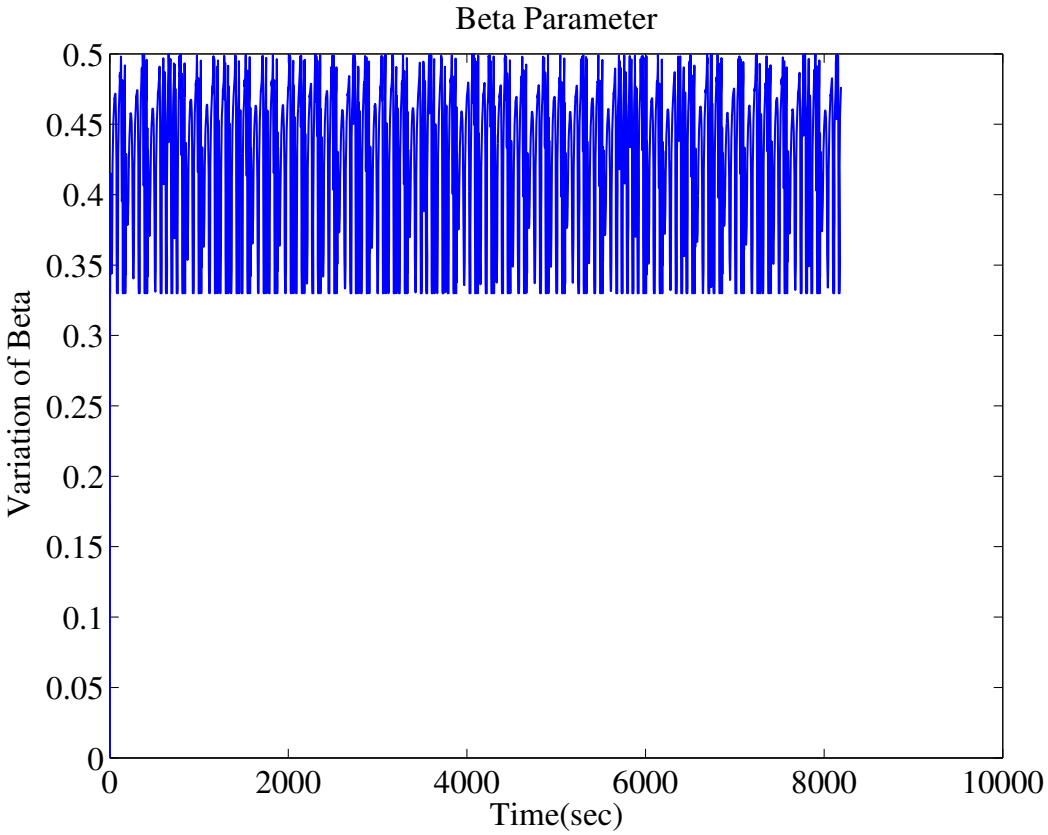

Figure 4.120: Variations in Beta Parameter During Rossler Time Series Prediction by the Proposed F-ADBEL Network.

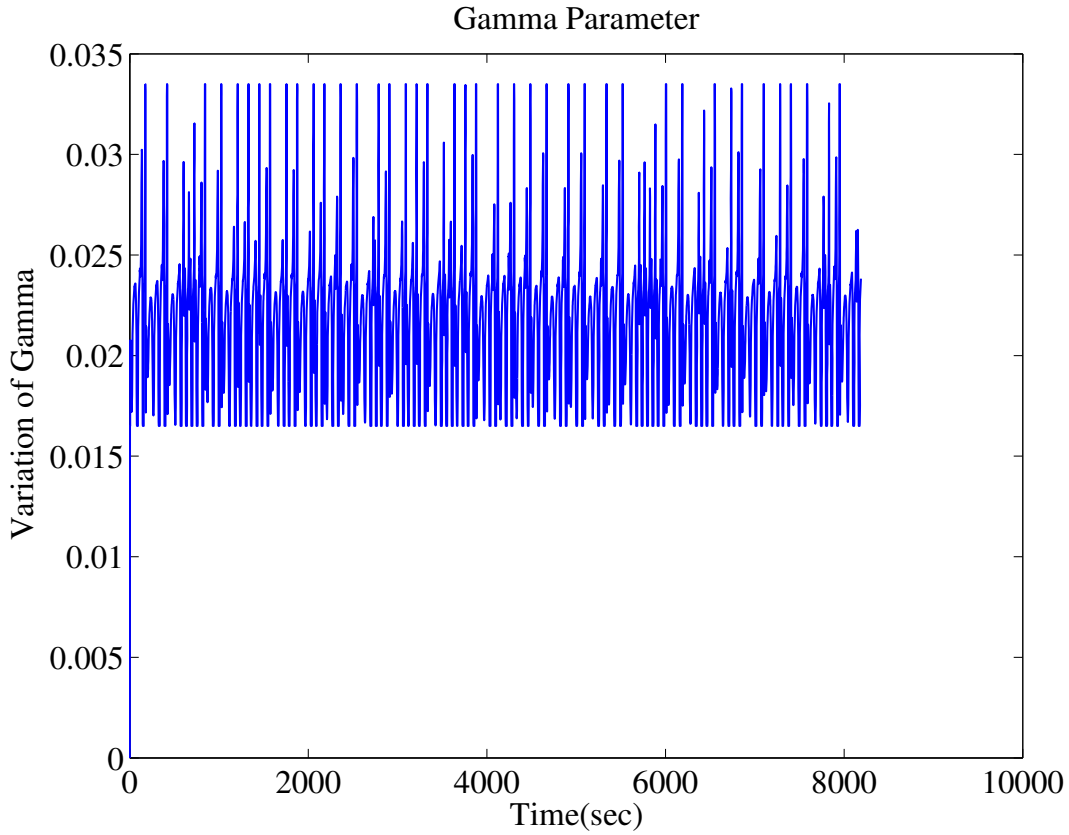

Figure 4.121: Variations in Gamma Parameter During Rossler Time Series Prediction the Proposed F-ADBEL Network.

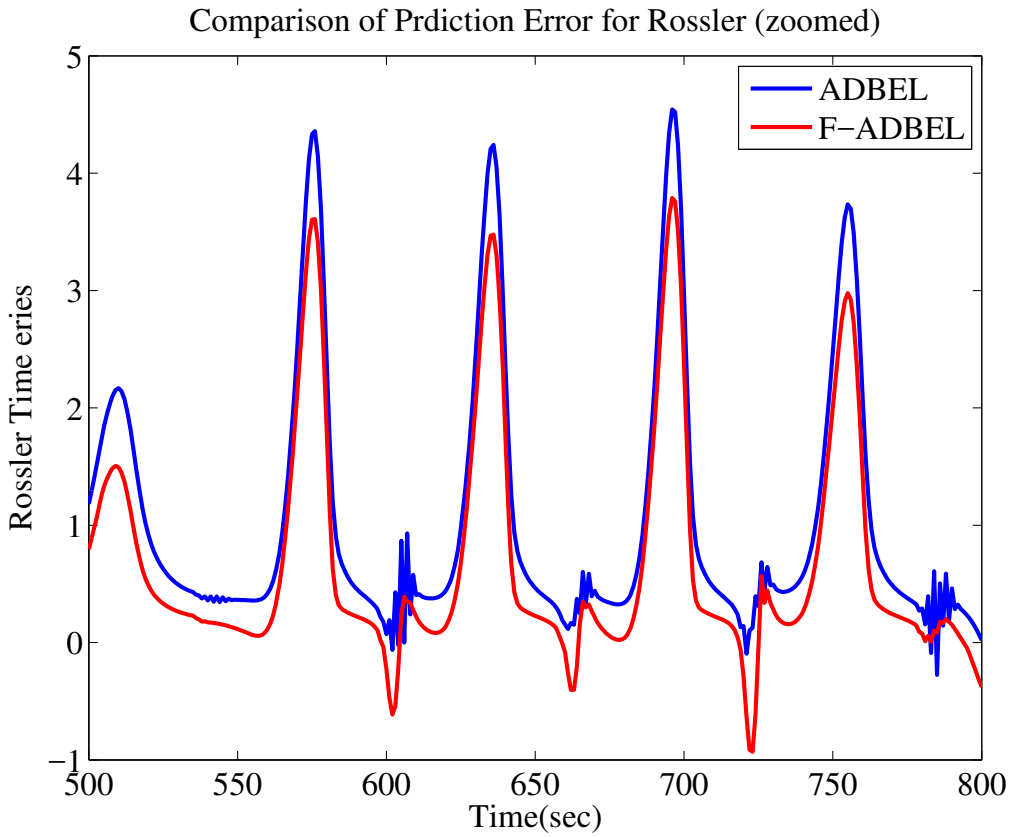

Figure 4.122: Comparison of Errors in Predicting Rossler Time Series by ADBEL and Proposed F-ADBEL Networks.

4.2.4 Disturbance Storm Time Index (D_{st}) Predicted by the Proposed F-ADBEL Network

The proposed F-ADBEL network is used to predict the disturbance storm (D_{st}) index, which is an indicator of solar storms. This time series's negative values are vital, as they indicate the weakening of the earth's magnetic field, leading to solar storms. In this section, we will apply the proposed F-ADBEL only for one category of (D_{st}) index, which is (D_{st}), July 2000, as an application to test the proposed model.

The proposed F-ADBEL network is deployed to predict (D_{st}), July 2000, as shown in Figures 4.123, 4.124 and 4.125, in terms of root mean squared error and correlation coefficient. The results are presented in Table 4.9, and the variations in the amygdala weights and orbitofrontal cortex weights are presented in zoomed Figures 4.126 and 4.127, respectively.

Figure 4.127 illustrates that the orbitofrontal cortex weights converge after the initial transient period. On the other hand, the amygdala weights, as shown in Figure 4.126, display almost the same variations in orbitofrontal cortex weights in steady-state. This supports the hypothesis that the orbitofrontal cortex is the more stable portion of the emotional brain and can correct the amygdala's response, which reacts quickly to emotional stimuli. Note that the amygdala weights are lower-bounded by zero, as was discussed during the development of the fuzzy parameter adjuster in section 3.7.

The variations in the network parameters α, β, and γ produced by the proposed fuzzy model F-ADBEL are also shown in Figures 4.129, 4.130 and 4.131, respectively. As can be observed, the variations in the parameter β are minimal after the first few time-series samples, which can also be used to describe the changes in orbitofrontal cortex weights. On the other hand, the fuzzy model continues adjusting the different network parameters (α and γ) throughout the entire horizon, which also explains the variations in the amygdala weights. Figure 4.128 shows the interpretation of the reward signal, which could become positive or negative in value and support the concept discussed during the fuzzy parameter adjuster development in section 3.7.

The same time series is also predicted with the ADBEL network using the learning parameters: $\alpha = 0.8$, $\beta = 0.2$, and $\gamma = 0.0D$. As illustrated in Figure 4.132, the prediction errors are recorded in both cases, with analysis showing that the short

period is $\leq 27\ hr$ for F-ADBEL and $110hr$ for the ADBEL network.

The performances of the designed F-ADBEL and ADBEL networks for predicting (D_{st}), July 2000, are given in Table 4.9. As can be seen, the designed F-ADBEL network performed better in prediction error in terms of root mean squared error and high correlation compared to the ADBEL network. The simulation's run-time also records that the ADBEL network performed in 1.28 seconds, while F-ADBEL performed in 2.11 seconds.

Table 4.9: RMSE /COR/PI for (D_{st}) July 2000 Prediction by ADBEL and F-ADBEL Networks

Time Series	Prediction Network	RMSE	COR	PI(%)
(D_{st}) July 2000	ADBEL	21.2	0.9062	
	F-ADBEL	15.13	0.9449	**28..63**

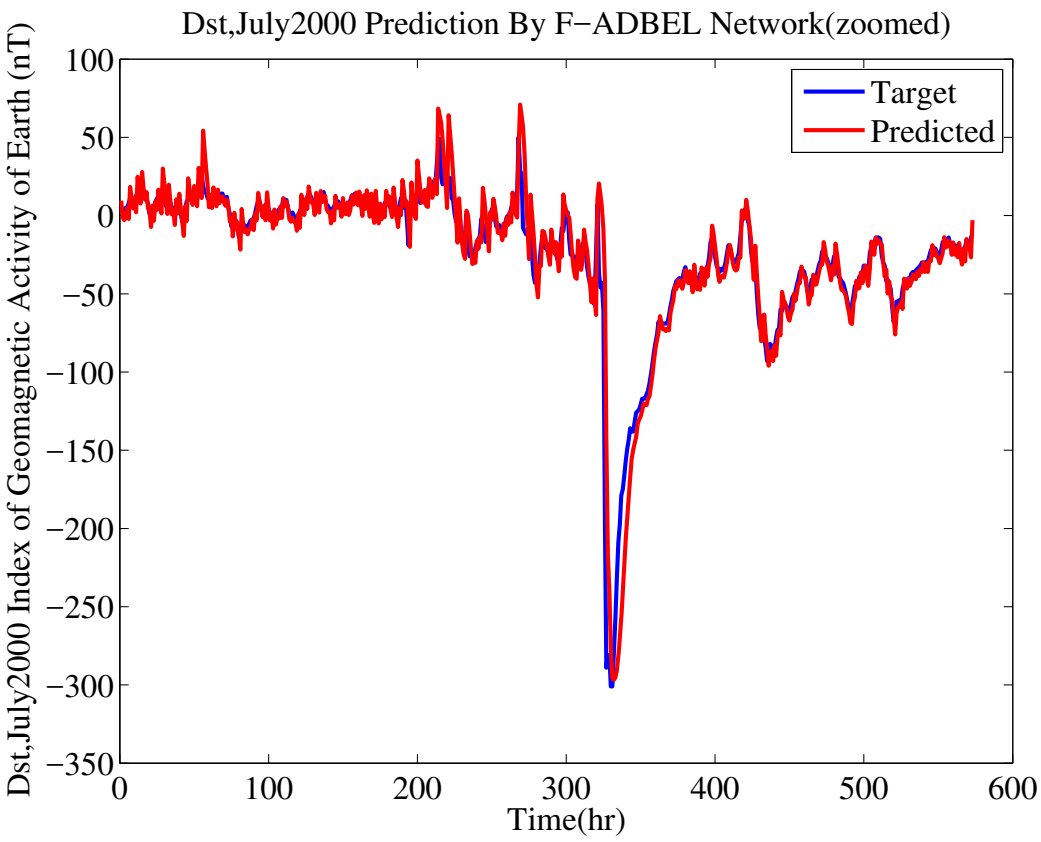

Figure 4.123: (D_{st}) July 2000 as Predicted by the Proposed F-ADBEL Network.

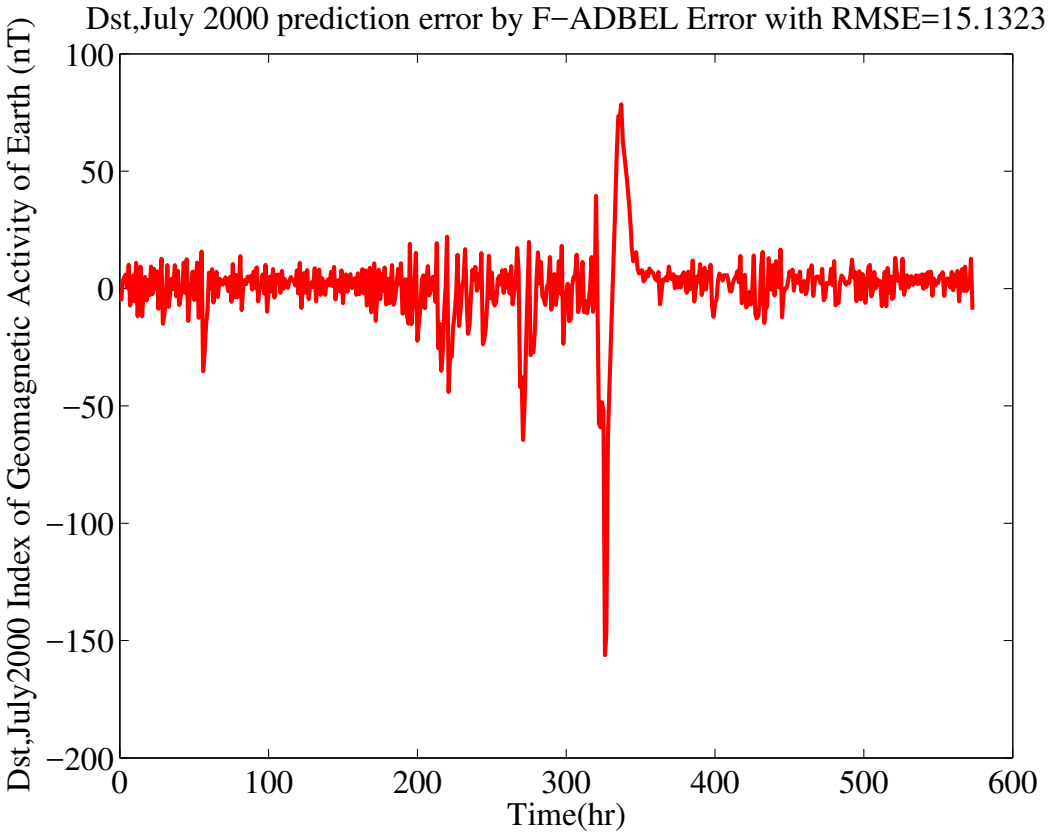

Figure 4.124: Error in Predicting (D_{st}) July 2000 by the Proposed F-ADBEL Network.

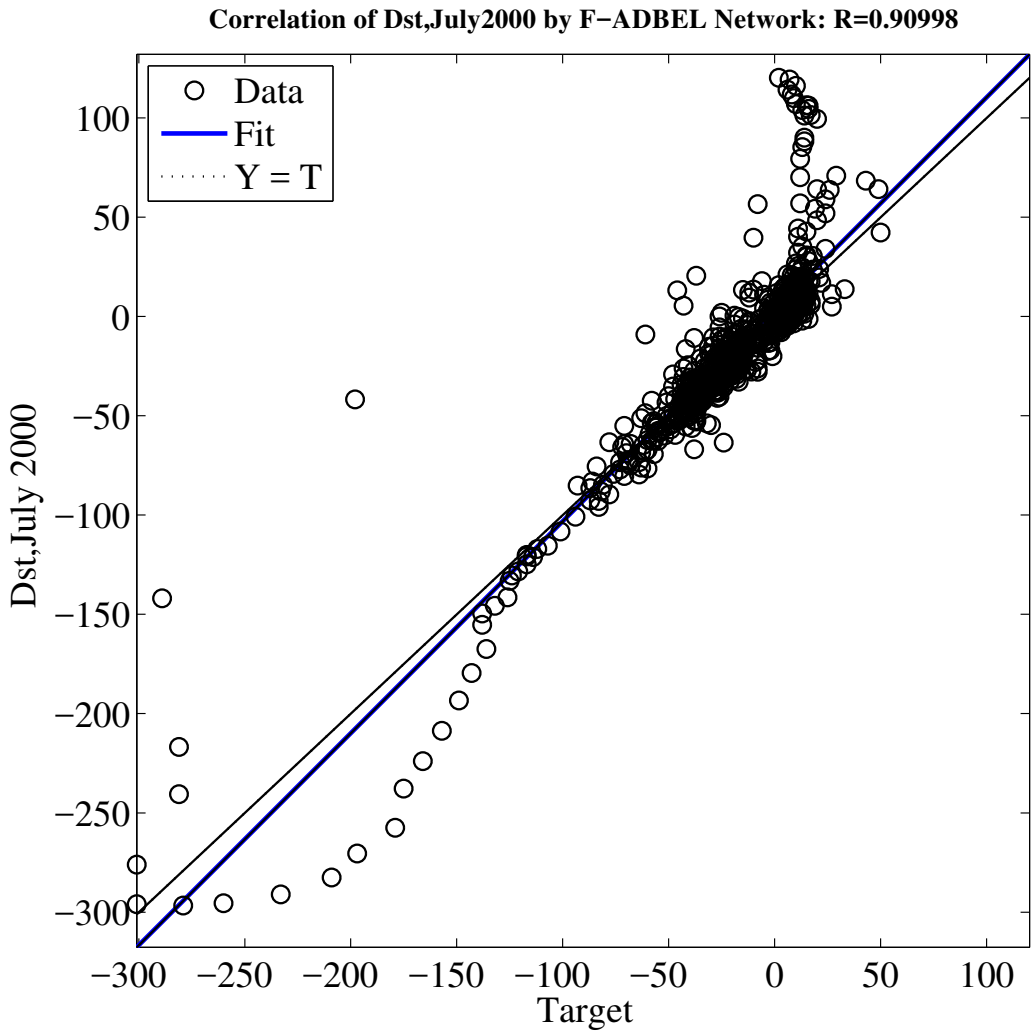

Figure 4.125: Correlation in Predicting (D_{st}) July 2000 by the Proposed F-ADBEL Network.

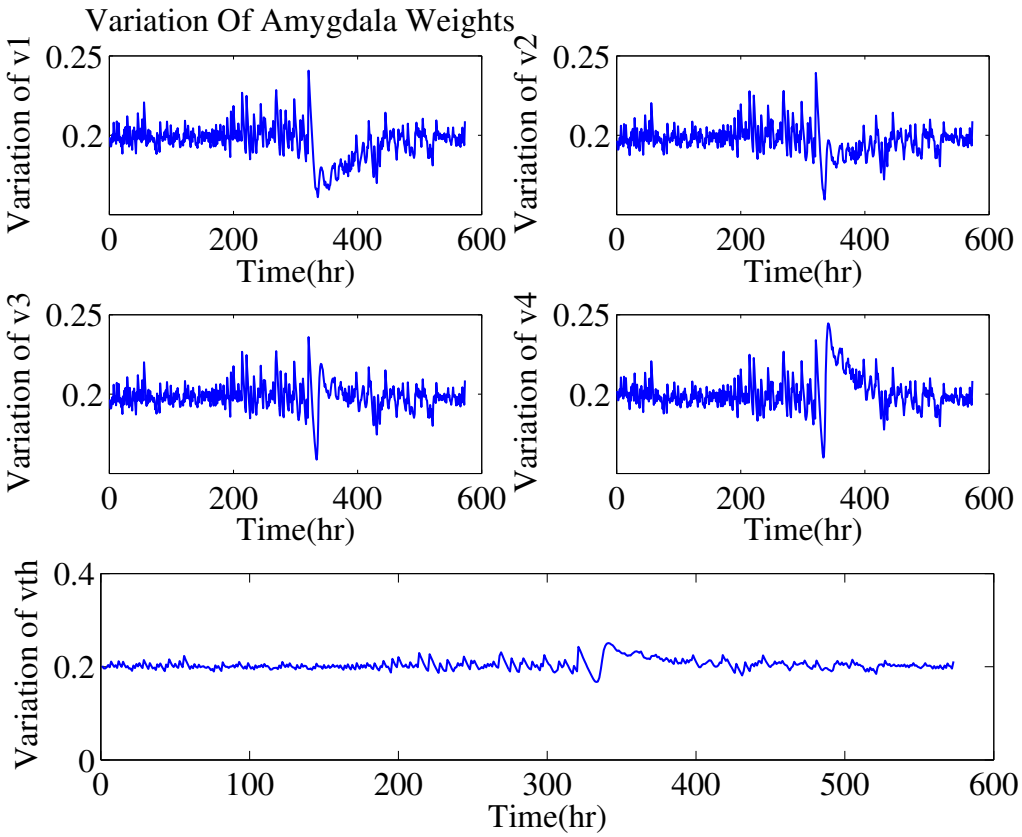

Figure 4.126: Variations in Amygdala Weights During (D_{st}) July 2000 Prediction by the Proposed F-ADBEL Network.

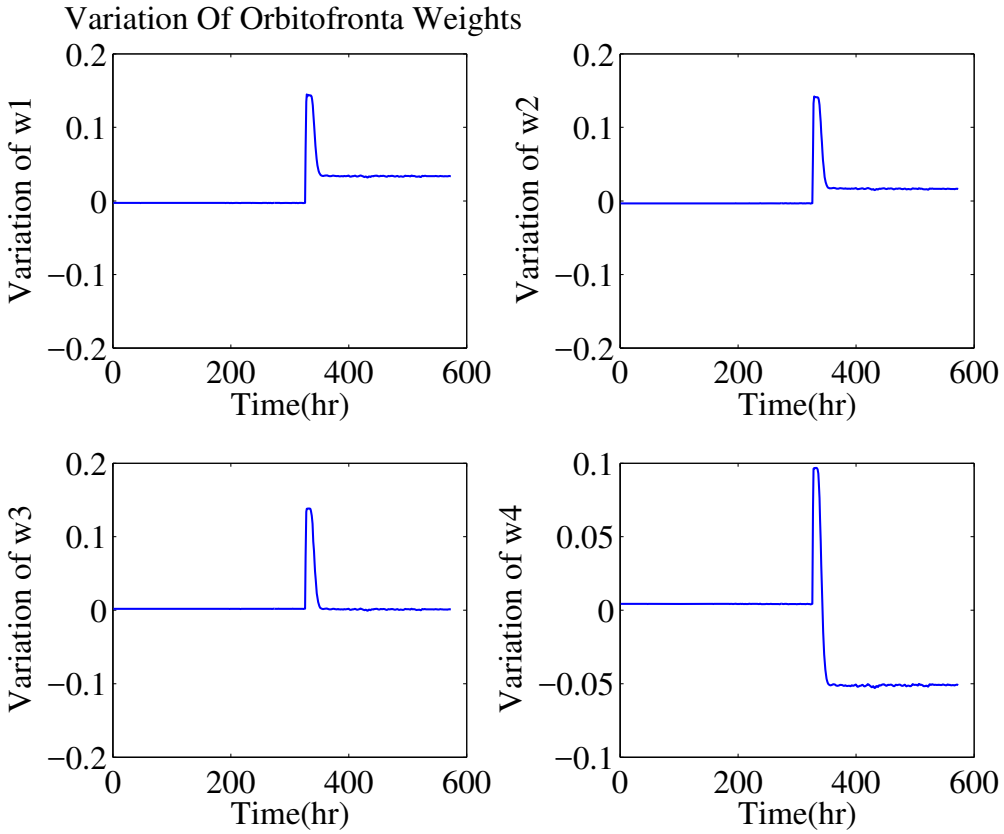

Figure 4.127: Variations in Orbitofrontal Cortex Weights During (D_{st}) July 2000 Prediction by the Proposed F-ADBEL Network.

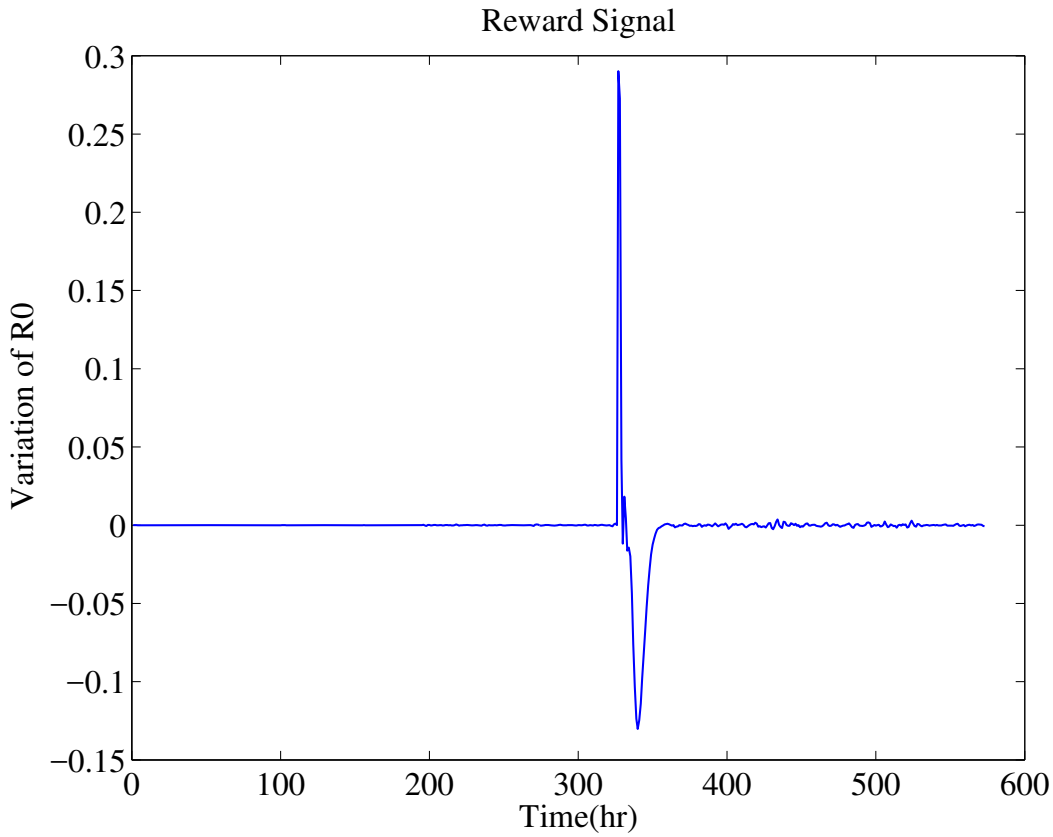

Figure 4.128: Variations in Reward Signal During (D_{st}) July 2000 Prediction by the Proposed F-ADBEL Network.

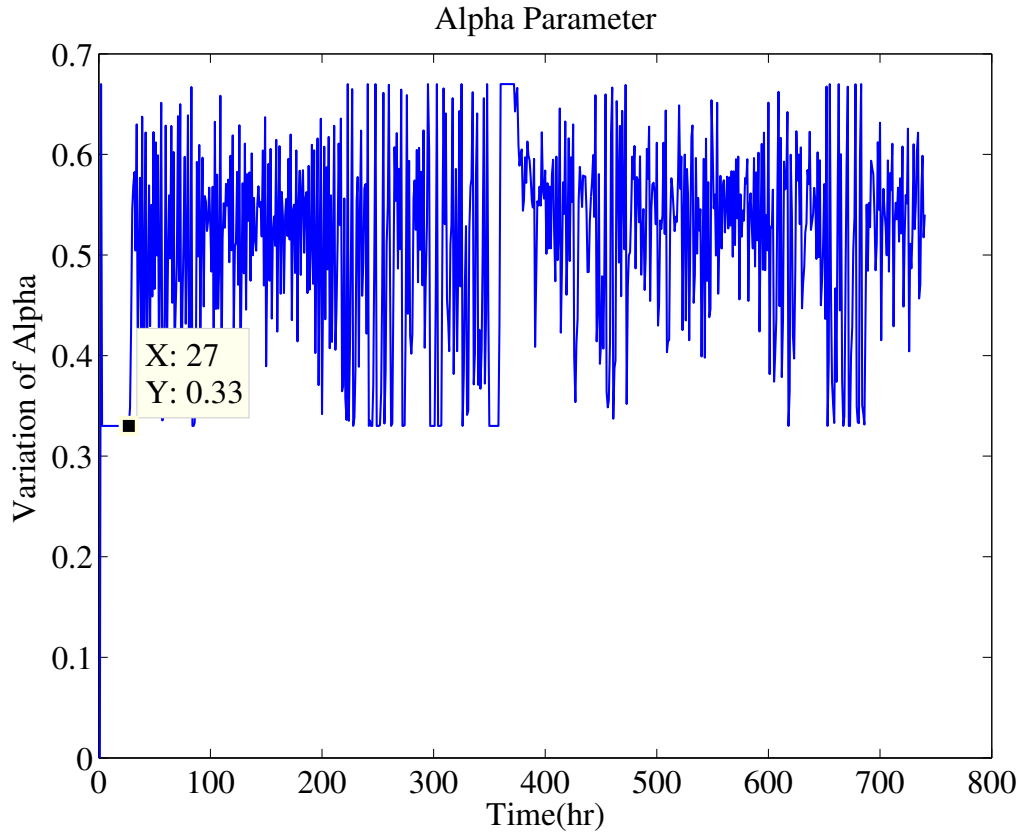

Figure 4.129: Variations in Alpha Parameter During (D_{st}) July 2000 Prediction by the Proposed F-ADBEL Network.

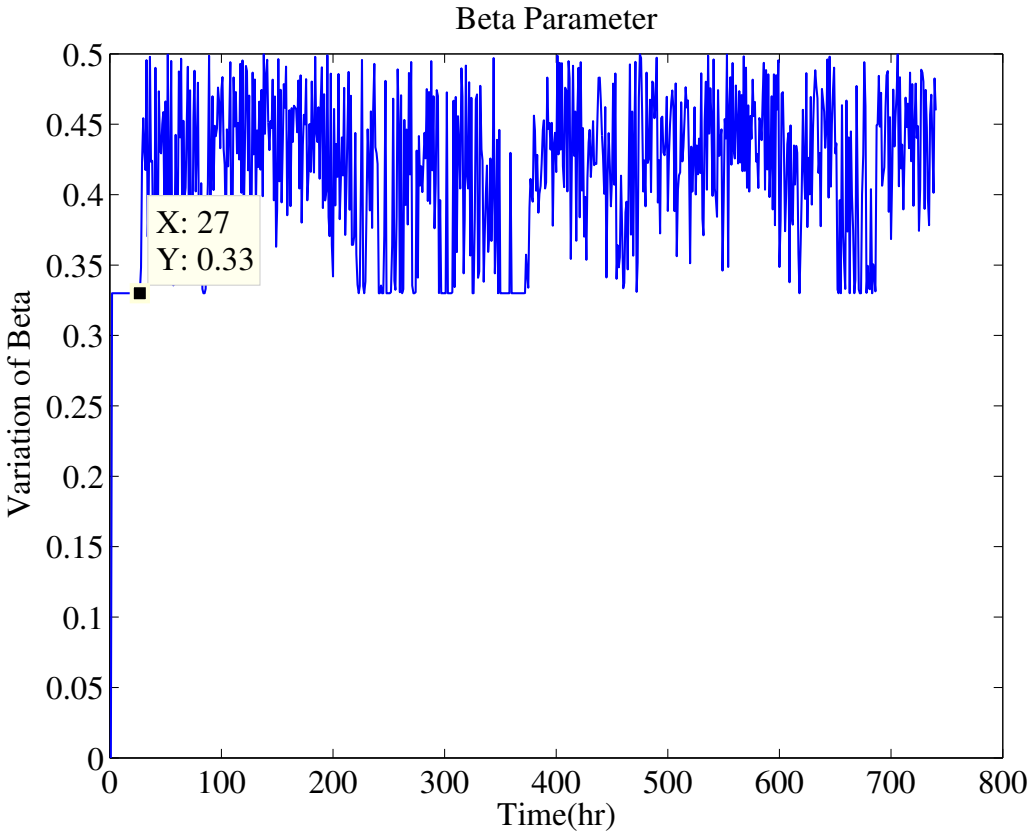

Figure 4.130: Variations in Beta Parameter During (D_{st}) July 2000Prediction by the Proposed F-ADBEL Network.

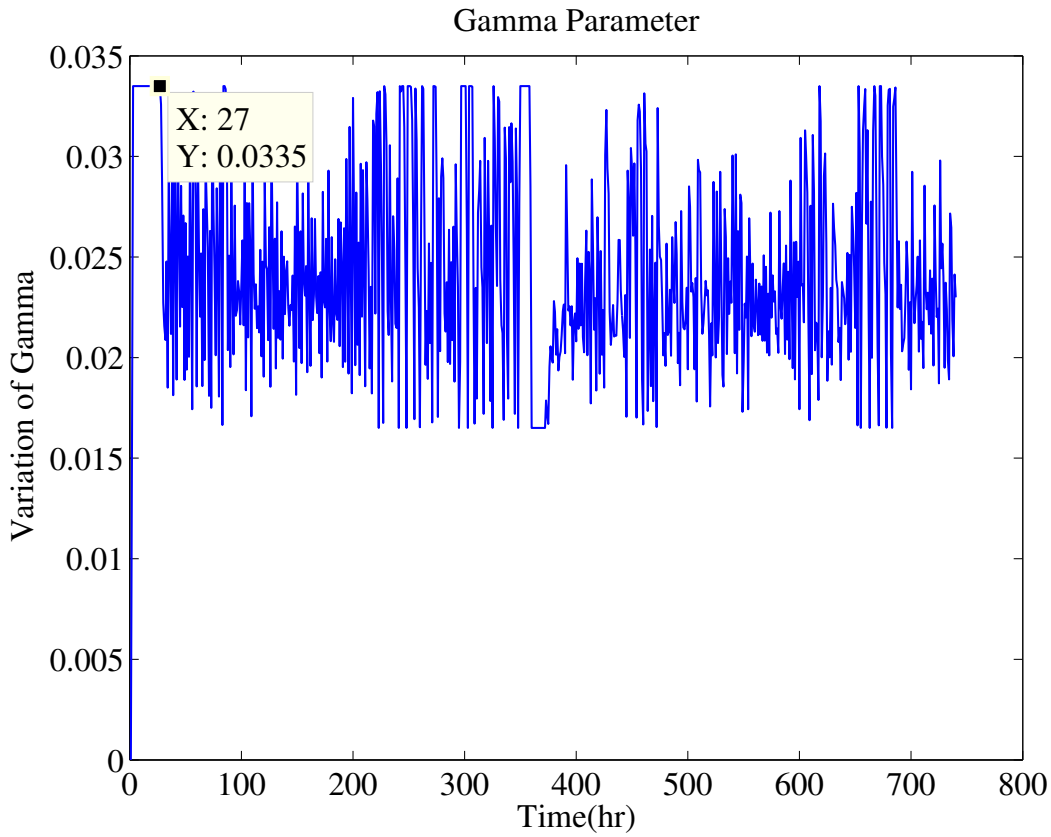

Figure 4.131: Variations in Gamma Parameter During (D_{st}) July 2000 Prediction by the Proposed F-ADBEL Network.

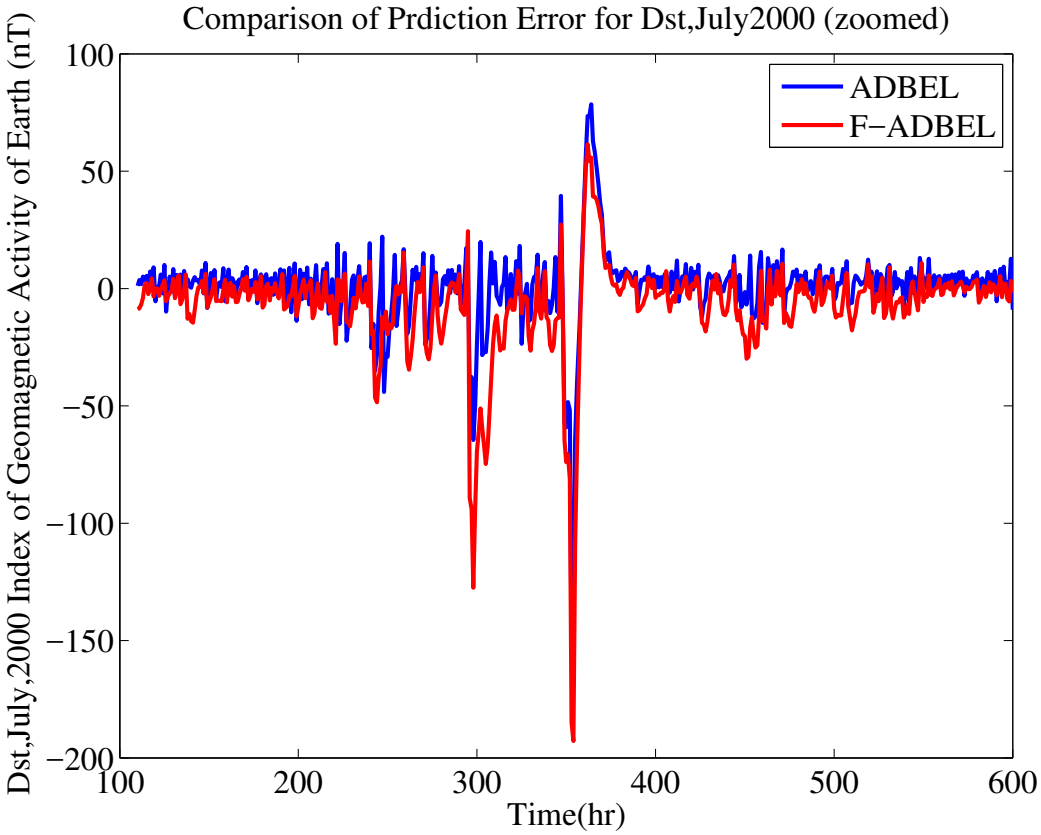

Figure 4.132: Comparison of Errors in Predicting (D_{st}) July 2000 by ADBEL and Proposed F-ADBEL Networks.

4.2.5 Conclusions

In this work, a fuzzy logic-based model is presented to vary the parameters of the ADBEL network. The proposed model is simple and uses only a few fuzzy sets and a smaller rule base to adjust the online parameters. Thus, the fuzzy integrated ADBEL network (F-ADBEL) can be used for the online forecasting of time series, with shorter update intervals on the order of a few seconds. The resultant fuzzy ADBEL network F-ADBEL is used to predict some popular chaotic time series, including Mackey-Glass, Lorenz, Rossler and disturbance storm time. The results indicate that the proposed fuzzy adjuster is able to adjust the ADBEL network's parameters. A comparison to ADBEL is discussed in different applications, demonstrating that the proposed F-ADBEL has been successfully accomplished in this work.

4.3 Performance of the Proposed Expanded Neo-Fuzzy Adaptive Decayed Brain Emotional Learning (ENF-ADBEL) Model

The proposed Expanded neo-fuzzy integrated ADBEL (ENF-ADBEL) network is tested in a MATLAB (R2014a) programming environment for online forecasting. Some chaotic time series, including Mackey-Glass, Lorenz, disturbance storm time index, wind speed and wind power generation, are used to validate the model.

The ENF-ADBEL network's performance is accessed in terms of root mean squared error and correlation coefficient criteria. Because we have already demonstrated the superiority of the proposed NF-ADBEL network compared to the ADBEL network in section 4.1, the comparisons we make here involve the NF-ADBEL network driven by near optimal set of parameters. The proposed ENF-ADBEL network's performance is evaluated in terms of root mean squared error and correlation coefficient criteria. Furthermore, because the NF-ADBEL network was deployed and simulated to predict the exact time-series with near optimal parameters, the percentage improvement index is used as a basis for comparison.

4.3.1 Mackey-Glass Time Series as Predicted by the Proposed ENF-ADBEL Network

Let us first predict the time series data generated from a time-delayed Mackey-Glass nonlinear differential equation, as defined in Eq. (4.4). This equation is simulated in MATLAB to create the time series data, which can be observed as non-periodic and non-convergent. The ENF-ADBEL network is deployed to predict the time series.

By setting the learning parameters as $\alpha = 0.5$, $\beta = 0.5$ and $\gamma = 0.07$, the ENF-ADBEL network is first deployed to predict this time series. The steady-state is 5 seconds, resulting in the pre-defined time window depicted in Figure 4.133.

The same time series is also predicted for the NF-ADBEL network, using the learning parameters as $\alpha = 0.5$, $\beta = 0.2$, and $\gamma = 0.03$. The prediction errors are recorded in both cases, with analysis showing that the transient period remains the same: $\leq 5\ s$. Thus, the steady-state starting index is set at $5 seconds$ in both cases to compute the performance indices.

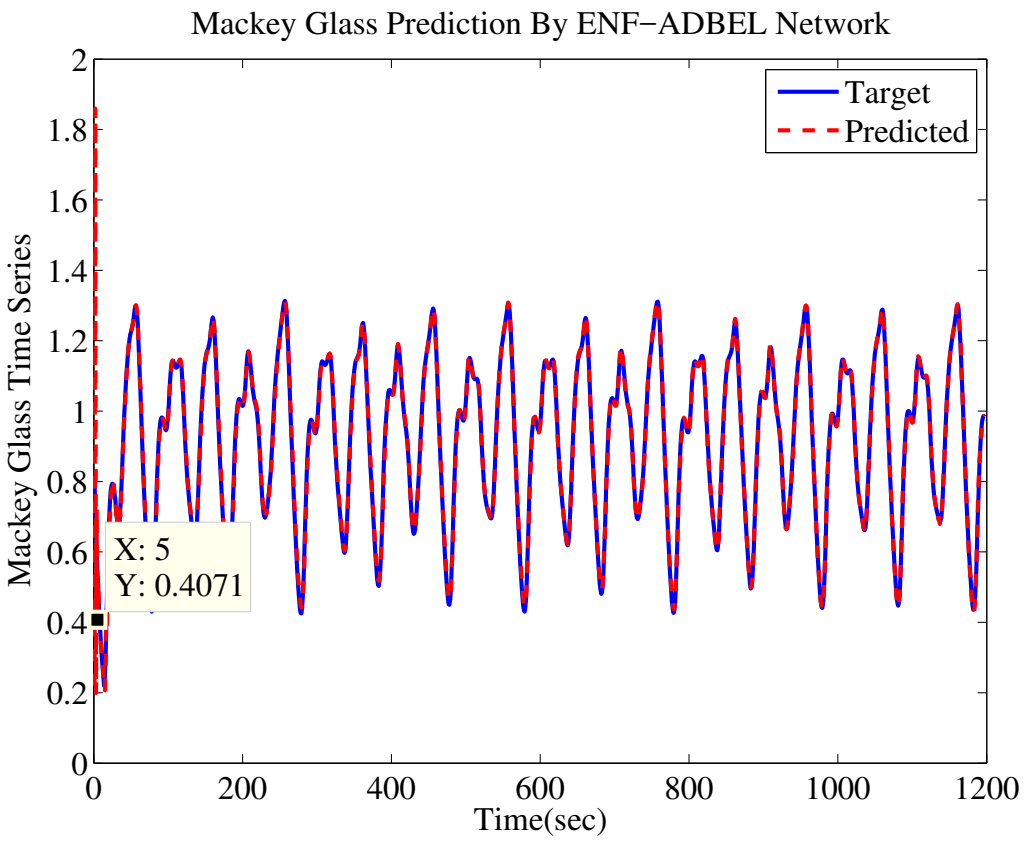

Figure 4.133: Mackey-Glass Time Series as Predicted by the ENF-ADBEL Network.

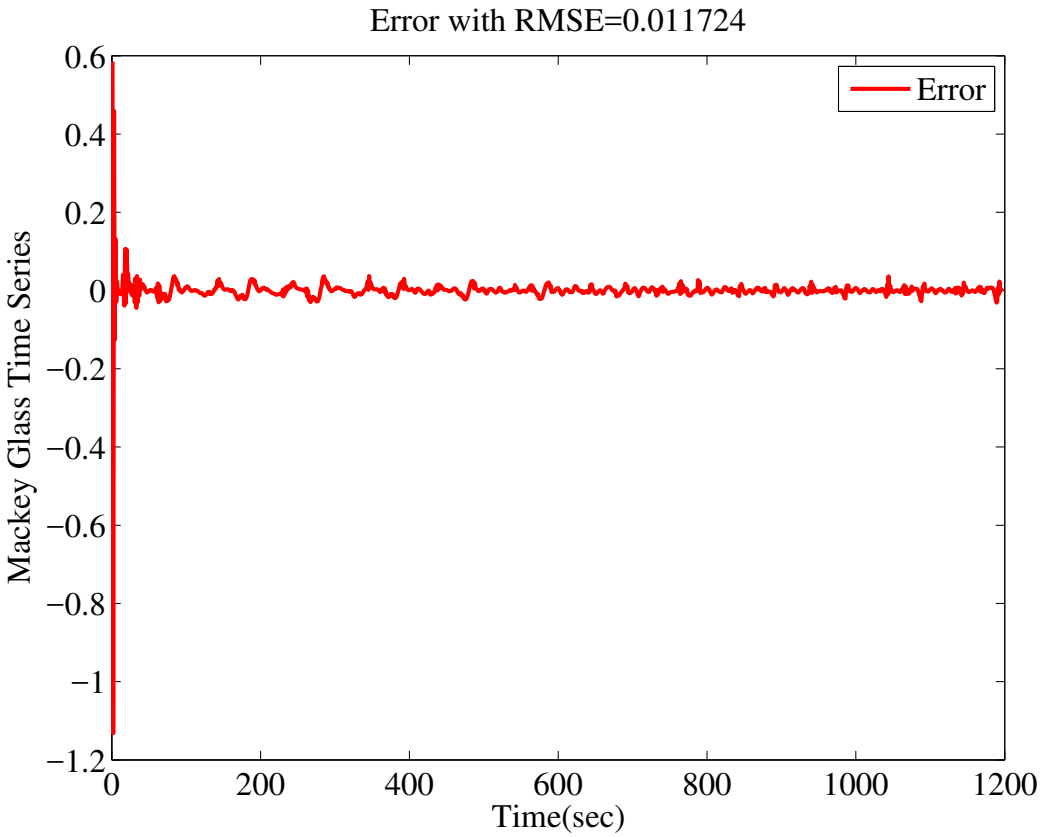

Figure 4.134: Error in Predicting Mackey-Glass Time Series by the Proposed ENF-ADBEL Network.

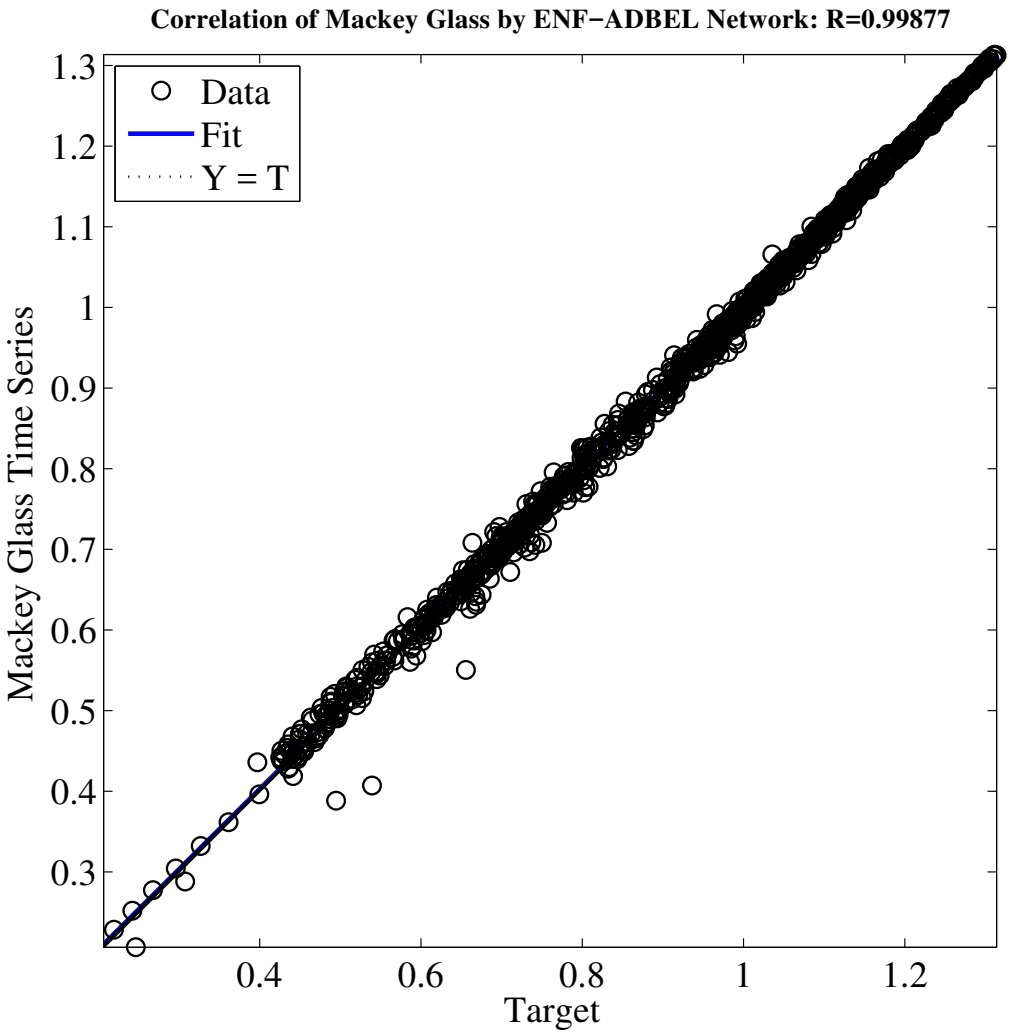

Figure 4.135: Correlation in Predicting Mackey-Glass Time Series by the Proposed ENF-ADBEL Network.

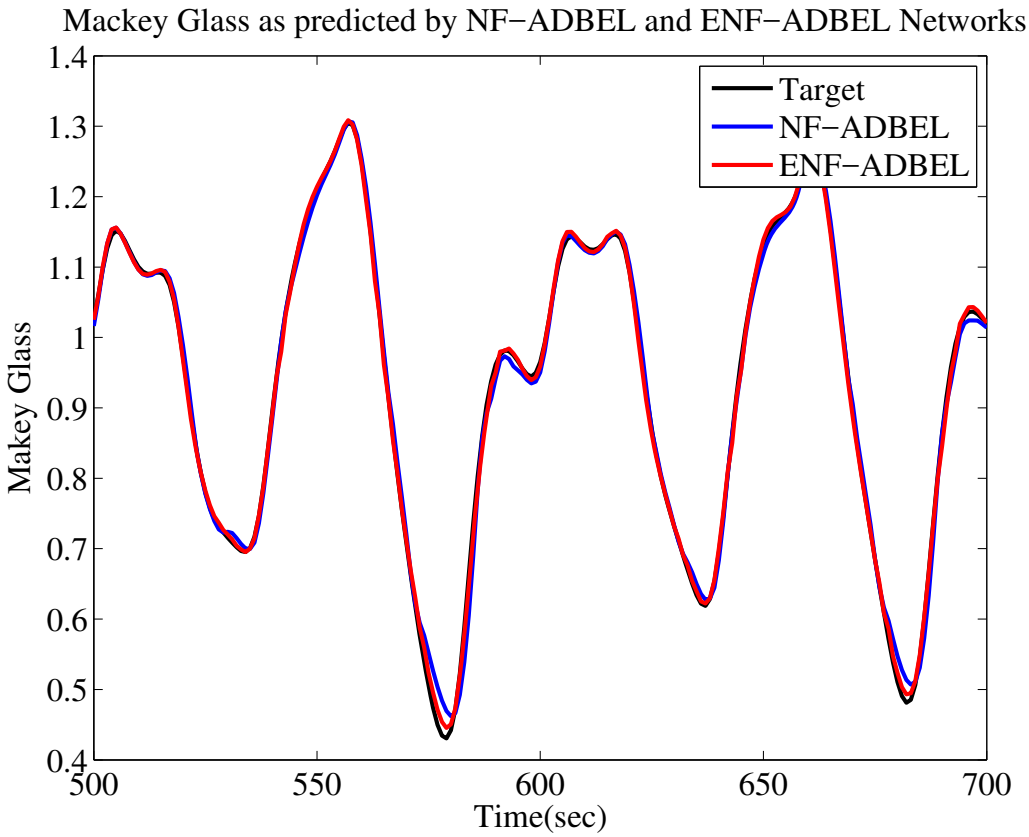

Figure 4.136: Mackey-Glass Time Series as Predicted by the ENF-ADBEL and NF-ADBEL Networks.

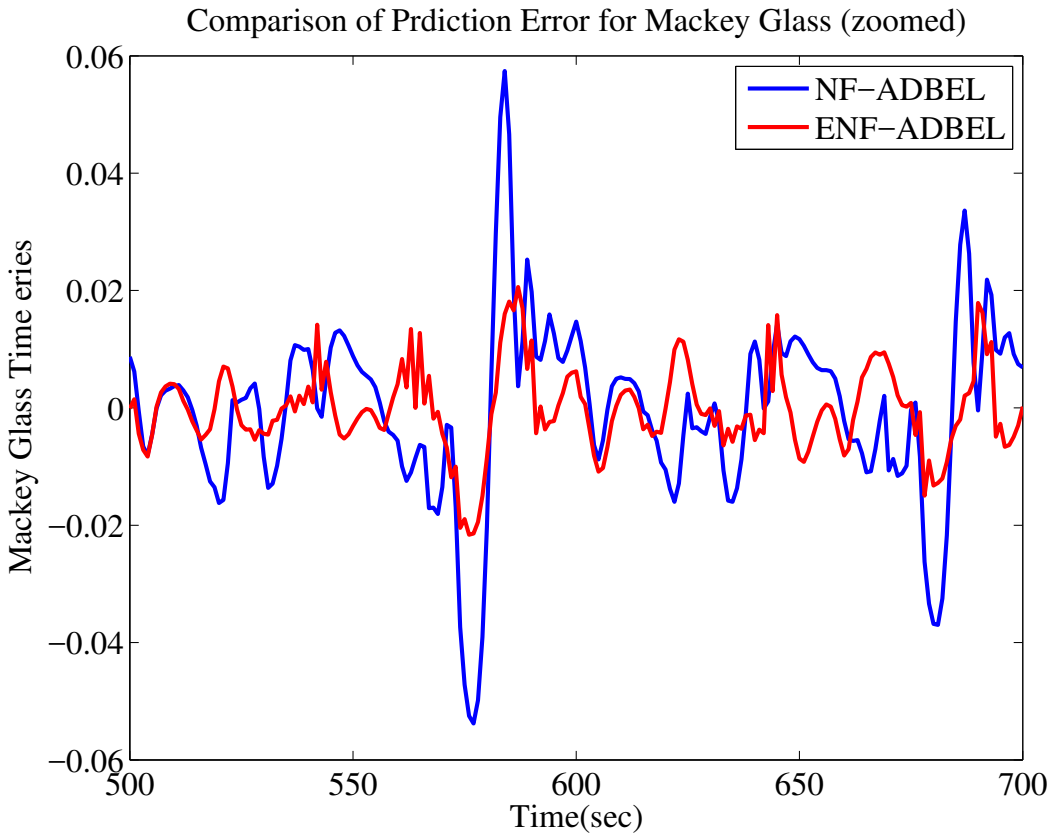

Figure 4.137: Error Comparison in Predicting Mackey-Glass Time Series as Predicted by the ENF-ADBEL and NF-ADBEL Networks.

Figure 4.136 and Figure 4.137 show the prediction and error comparison in steady-state as yielded by the ENF-ADBEL and NF-ADBEL networks for predicting the Mackey-Glass time series. The ENF-ADBEL network gives better results than the NF-ADBEL networks, showing lower peaks in the prediction error in the ENF-ADBEL network. Further, the root mean squared error and correlation coefficient are also determined for both networks using the relations in (4.1) and (4.2). The computed values are presented in Table 4.10.

A lower root mean squared error and a higher correlation coefficient are offered by the ENF-ADBEL network for predicting the Mackey-Glass time series as compared to NF-ADBEL networks, as illustrated in Figure 4.133, Figure 4.134 and Figure 4.135, respectively. A fairly significant amount of percentage improvement is also obtained, as expressed in (4.3). Please note that the results in Table 4.10 were obtained by the proposed network where no prior training data is assumed.

Table 4.10: RMSE & R^2 for Mackey-Glass Time Series Prediction by the ENF-ADBEL and NF-ADBEL Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Mackey-Glass	ENF-ADBEL	0.011	99.87	38.88
	NF-ADBEL [58]	0.018	99.71	

Furthermore, a multilayer perceptron (MLP) neural network is used for the same Mackey-Glass data points to compare the proposed ENF-ADBEL. According to [88], MLP is considered the most widely used neural network in time series data forecasting. In MLP, we structured the network for ten hidden layers, as shown in Figure 4.138. We used the GD method, with the data divided as follows: 70% as trained data, 15% as validated data, and 15% as tested data. The results, presented in Table 4.11, show that ENF-ADBEL performed significantly better and had better outcomes.

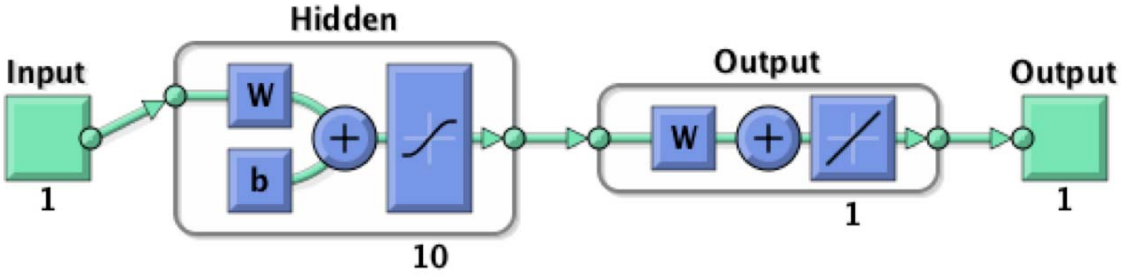

Figure 4.138: Multilayer Perceptron Neural Network (MLP).

Table 4.11: RMSE & R^2 for Mackey-Glass Time Series Prediction by the ENF-ADBEL and MLP Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Mackey-Glass	ENF-ADBEL	0.011	99.89	66.66
	MLP	0.033	98.28	

4.3.2 Lorenz Time Series as Predicted by the Proposed ENF-ADBEL Network

For the generated Lorenz time series having $n_e = 16380$ data points using eq (4.5), we first evaluate the prediction performance of the ENF-ADBEL network, as shown in fig 4.139, with the learning parameters set as: $\alpha = 0.5$, $\beta = 0.3$, and $\gamma = 0.04$.

Table 4.12: RMSE & R^2 for Lorenz Time Series as Predicted by ENF-ADBEL and NF-ADBEL Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Lorenz	ENF-ADBEL	0.13054	99.95	24.36
	NF-ADBEL [58]	0.1726	99.97	

Figure 4.139 shows the Lorenz time series as predicted by the ENF-ADBEL network. It can be observed that the prediction results of the ENF-ADBEL network for the Lorenz time series perform better when compared to the results for the Mackey-Glass time series. Specifically, it is difficult to distinguish the predicted Lorenz time series from the target data. The run-time for the proposed network is 1.42 seconds.

For the purpose of comparison, the NF-ADBEL network is also simulated to predict the Lorenz time series. The best learning parameters for the NF-ADBEL network in predicting the Lorenz time series are found to be $\alpha = 0.8$, $\beta = 0.2$, and $\gamma = 0.01$), with a run-time of 1.54 second. By recording and analyzing the prediction error in both cases, as shown in Figure 4.142, it was found that the transient period is less than $5s$, so the steady-state starting index is taken as $n_s = 5$.

A zoomed view of the prediction error as returned by both networks in steady-state is shown in Figure 4.143. From the figure, it is evident that the proposed ENF-ADBEL network has a lower error in predicting the Lorenz time series compared to the NF-ADBEL network. The prediction performance is also analyzed in terms of root mean squared error, as shown in Figure 4.140, and in terms of the correlation coefficient, as shown in Figure 4.141. The results for this analysis are included in Table 4.12, showing the superior performance of the ENF-ADBEL network. This is due to the lowered root mean squared error, higher correlation coefficient, and significant percentage improvement offered by the proposed network.

Moreover, a multilayer perceptron MLP neural network was applied to the same Lorenz data that was used to validate the proposed ENF-ADBEL. The comparison results are depicted in Table 4.13, showing that the ENF-ADBEL network performed with better accuracy.

Table 4.13: RMSE & R^2 for Lorenz Time Series Prediction by ENF-ADBEL and MLP Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Lorenz	ENF-ADBEL	0.1305	99.98	69.76
	MLP	0.43	99.96	

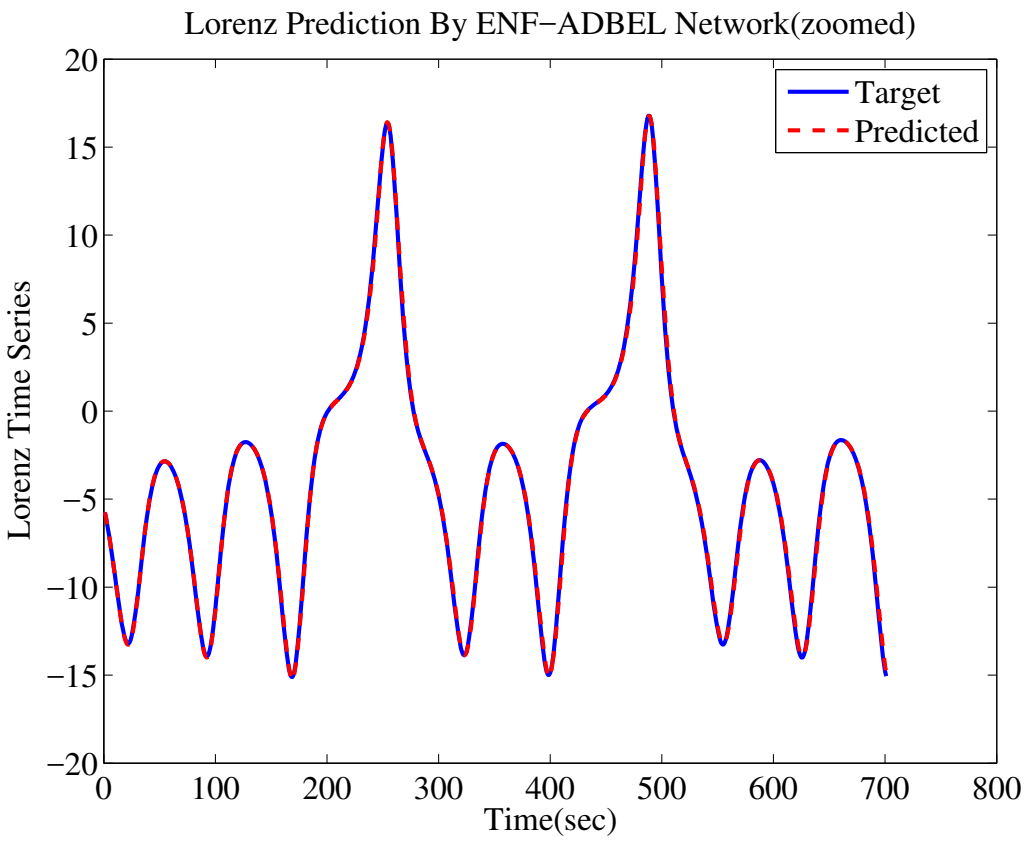

Figure 4.139: Lorenz Time Series as Predicted by the ENF-ADBEL Network.

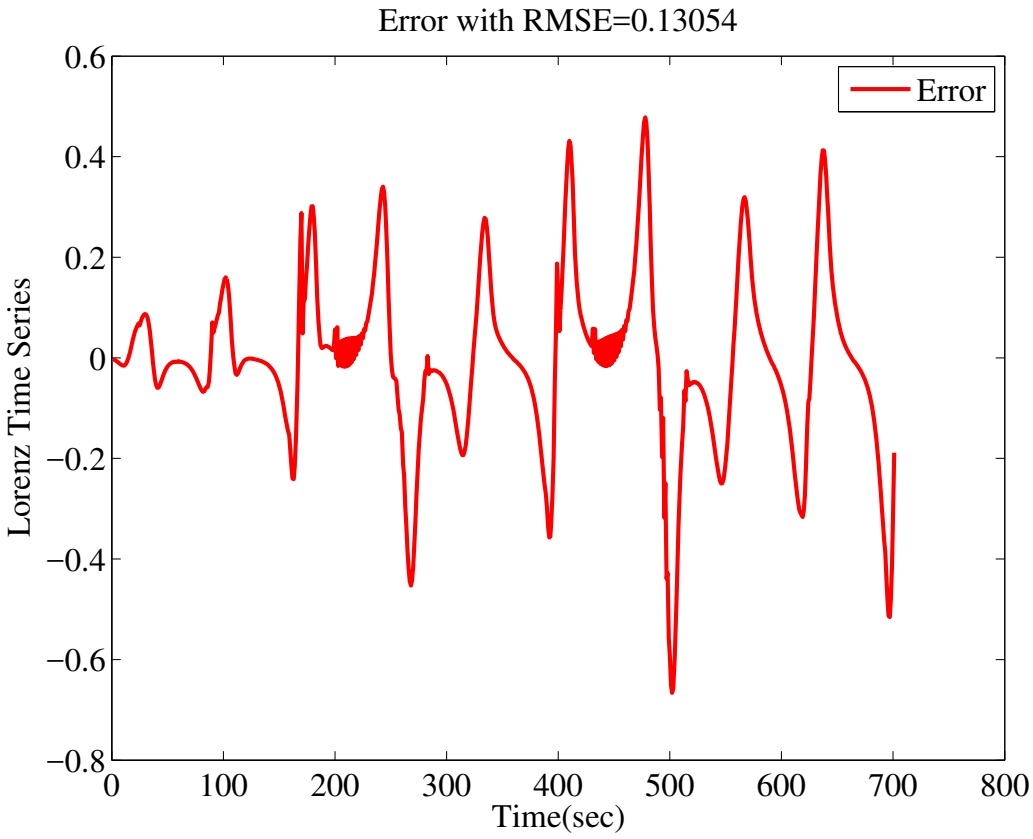

Figure 4.140: Error in Predicting Lorenz Time Series by the ENF-ADBEL Network.

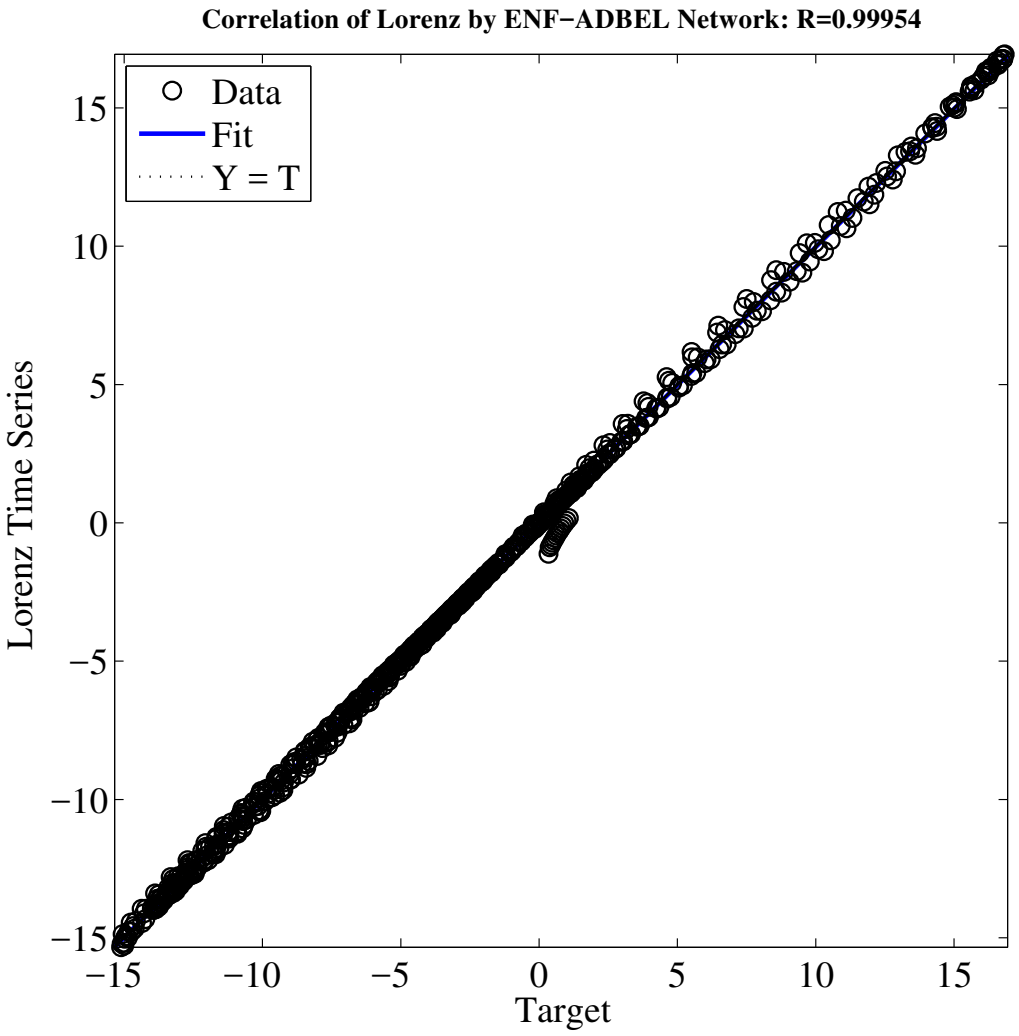

Figure 4.141: Correlation in Predicting Lorenz Time Series by the ENF-ADBEL Network.

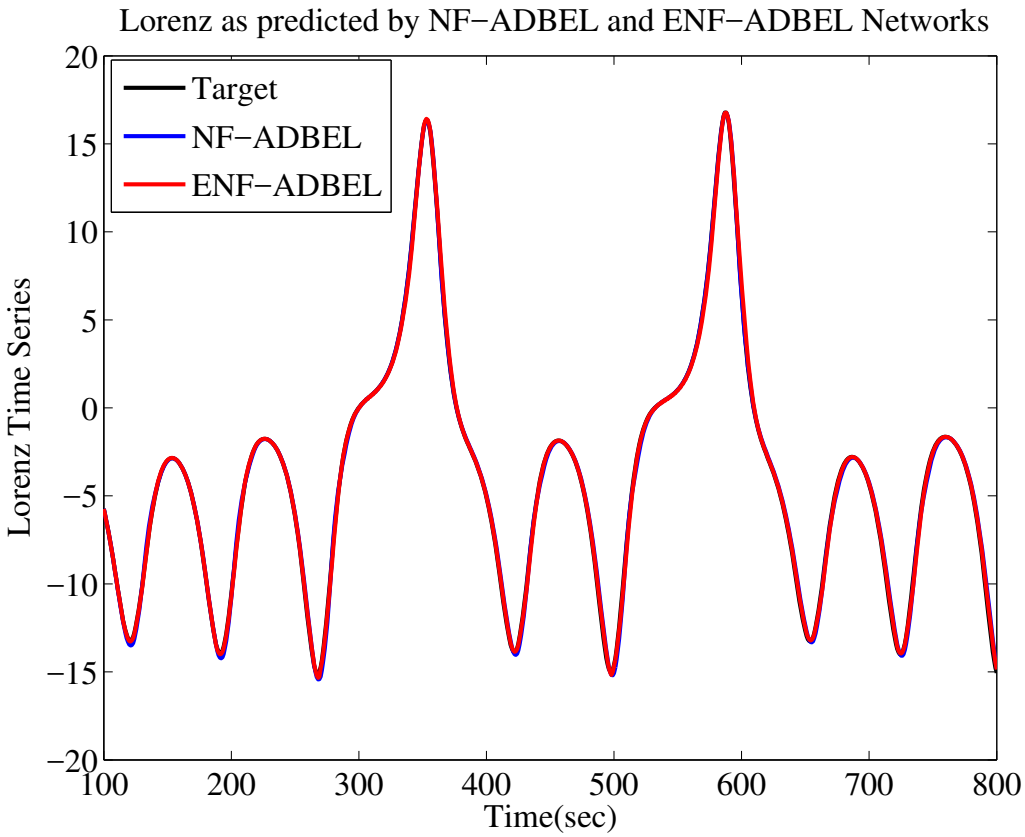

Figure 4.142: Lorenz Time Series as Predicted by the ENF-ADBEL and NF-ADBEL Networks.

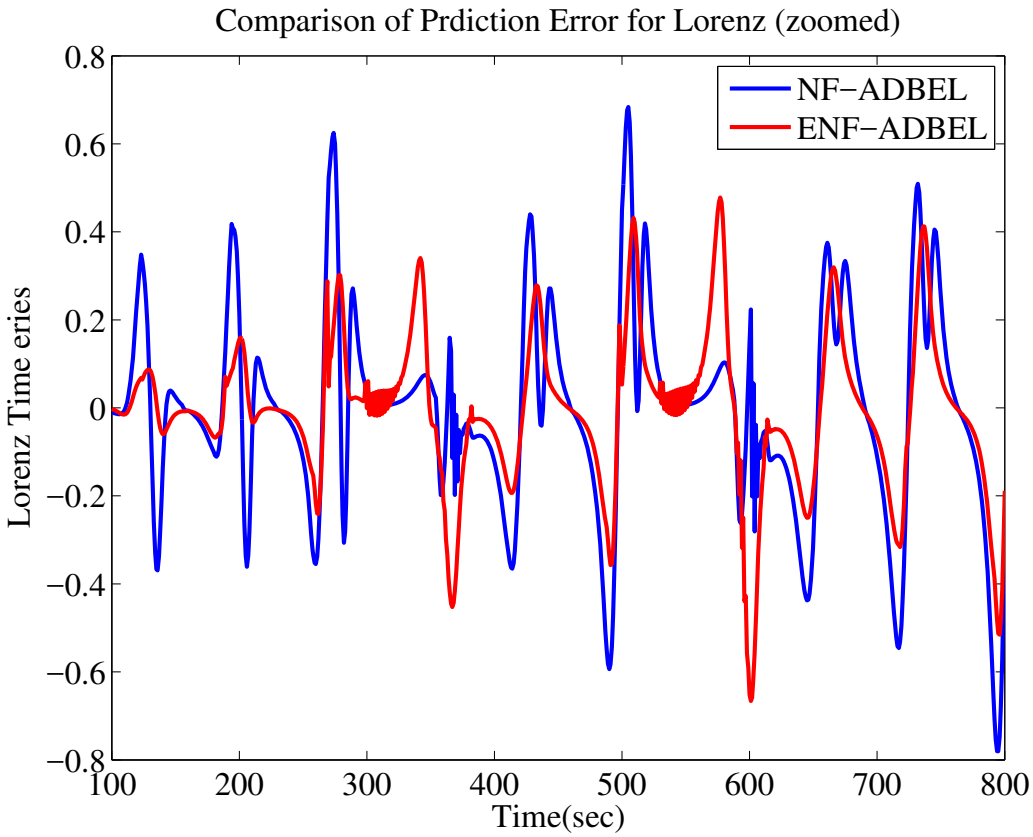

Figure 4.143: Error Comparison in Predicting Lorenz Time Series as Predicted by the ENF-ADBEL and NF-ADBEL Networks.

4.3.3 Disturbance Storm Time Index as Predicted by the Proposed ENF-ADBEL Network

The ENF-ADBEL network is proposed here to predict the disturbance storm time index, which is an hourly indicator of geomagnetic storms. The negative values in this index are vital, as they indicate the weakening of Earth's magnetic field. This event can lead to geomagnetic storms, which can disrupt radio communications, damage satellites, and cause power system outages, all of which were seen in the Hydro-Qubec transmission grid during the 1989 storm. In March 1989, the entire province was plunged into darkness for more than nine hours [81]. A number of models based on differential equations and neural networks have been proposed in the literature for predicting the disturbance storm time index [82],[83],[84],[85].

Recently, the ADBEL network also proposed predicting this important index [4], and NF-ADBEL in [58], which has been modified in the present work to yield a new ENF-ADBEL network. Here, we simulate the ENF-ADBEL network to predict the disturbance storm time index D_{st} time series for April 2000 when considerable geomagnetic activity was observed. The data for this month have been downloaded from the website World Data Center (WDC) [86], "WDC for Geomagnetism, Kyoto." With the learning parameters set as $\alpha = 0.15$, $\beta = 0.38$ and $\gamma = 0.25$, the ENF-ADBEL network is deployed to predict the D_{st} index for April 2000. The number of samples is $n_e = 716$ for the month.

The predicted results provided by the ENF-ADBEL network are shown in Figures 4.144, 4.145, and 4.146. The transient period of the ENF-ADBEL network is found to be $n_s = 10\ hrs$, which then becomes the steady-state starting index. It can be observed that, despite the high initial transients, the ENF-ADBEL network is able to follow the D_{st} time series in steady-state. The important valley points are also well-predicted, which actually points towards the possible occurrence of geomagnetic storms. The run-time for the proposed ENF-ADBEL network is 1.09 seconds.

In order to draw a comparison, an existing NF-ADBEL network is used to predict the D_{st} time series. For this purpose, the learning parameters of the NF-ADBEL network are assigned the values of $\alpha = 0.3$, $\beta = 0.3$, and $\gamma = 0.01$. The results of this comparison in terms of the prediction error are displayed in Figures 4.147 and 4.148, and the run-time for the proposed ENF-ADBEL network is 1.107 seconds. It

can be observed that the ENF-ADBEL network gives a better performance than the NF-ADBEL, as shown in Table 4.14.

Table 4.14: RMSE & R^2 for D_{st} by ENF-ADBEL, NF-ADBEL and F-ADBEL Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Dst Apr 2000	ENF-ADBEL	6.86	98.31	24.44
	NF-ADBEL [58]	9.08	97.06	

for more fair comparison, a MLP used to predict the same data of Dst for April,2000. The comparison results displayed in Table4.15. Results showed that ENF-ADBEL network performed better accuracy.

Table 4.15: RMSE & R^2 for D_{st} by ENF-ADBEL and MLP Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Dst Apr 2000	ENF-ADBEL	6.86	98.31	7.5
	MLP	7.42	95.32	

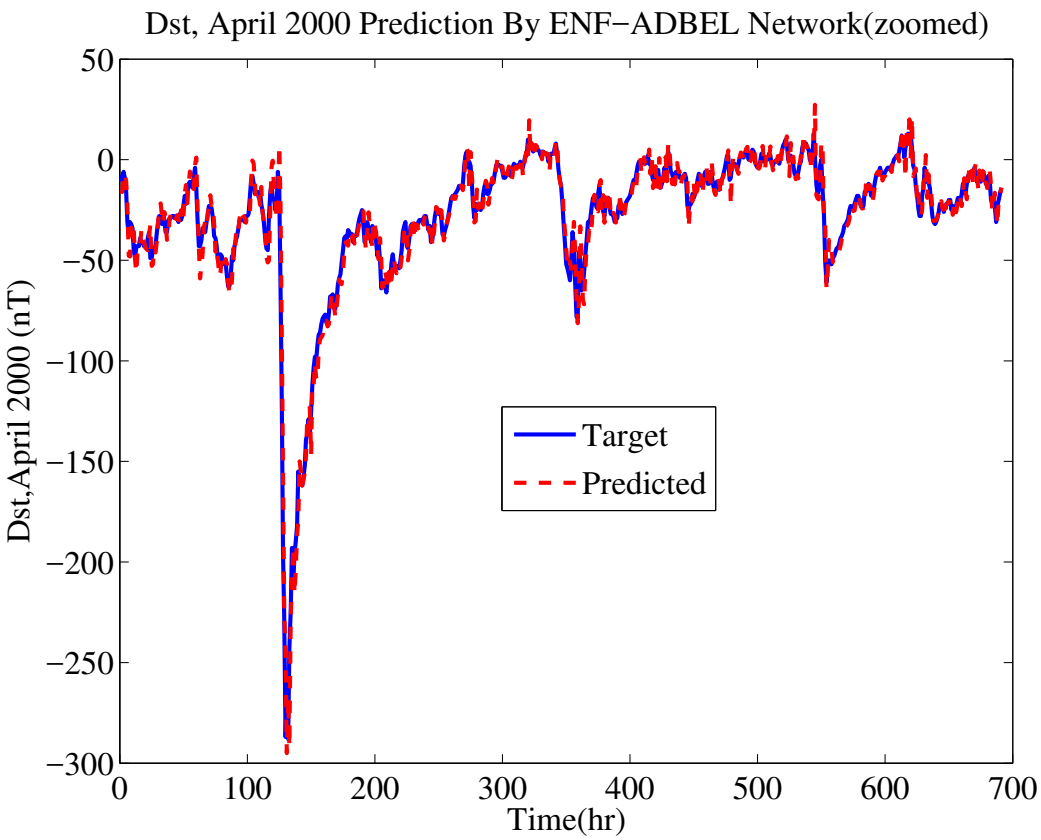

Figure 4.144: Dst April 2000 as Predicted by ENF-ADBEL Network.

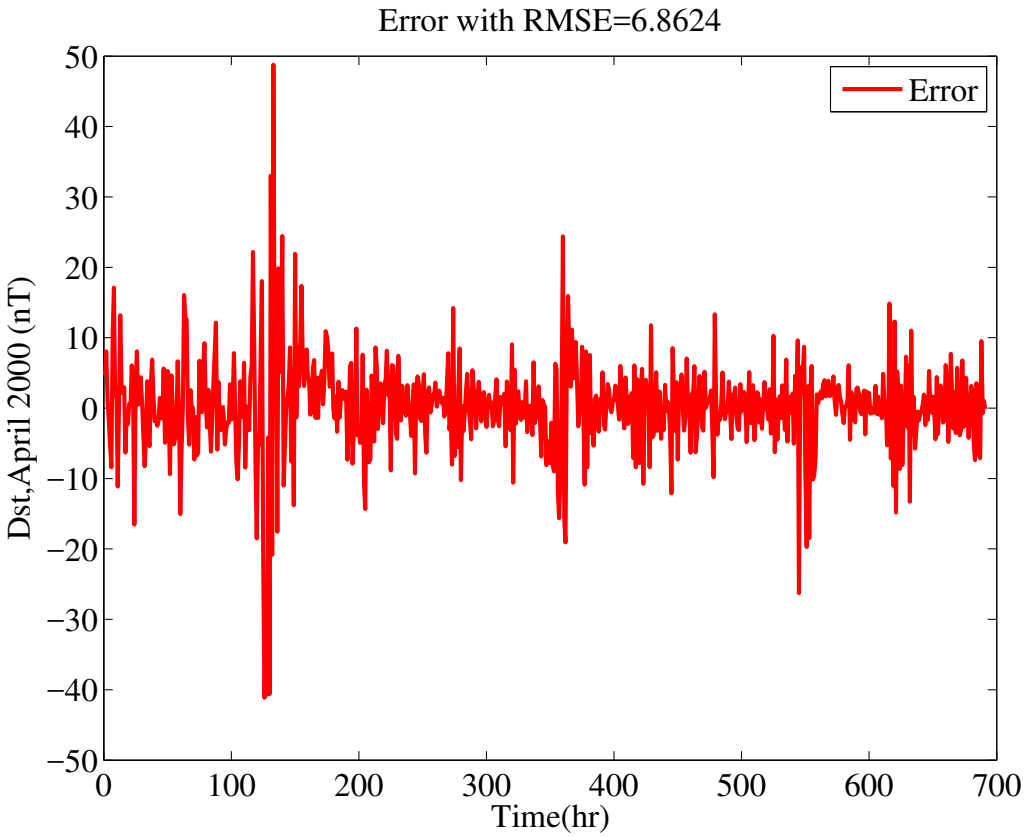

Figure 4.145: Error in Predicting D_{st} April 2000 by ENF-ADBEL Network.

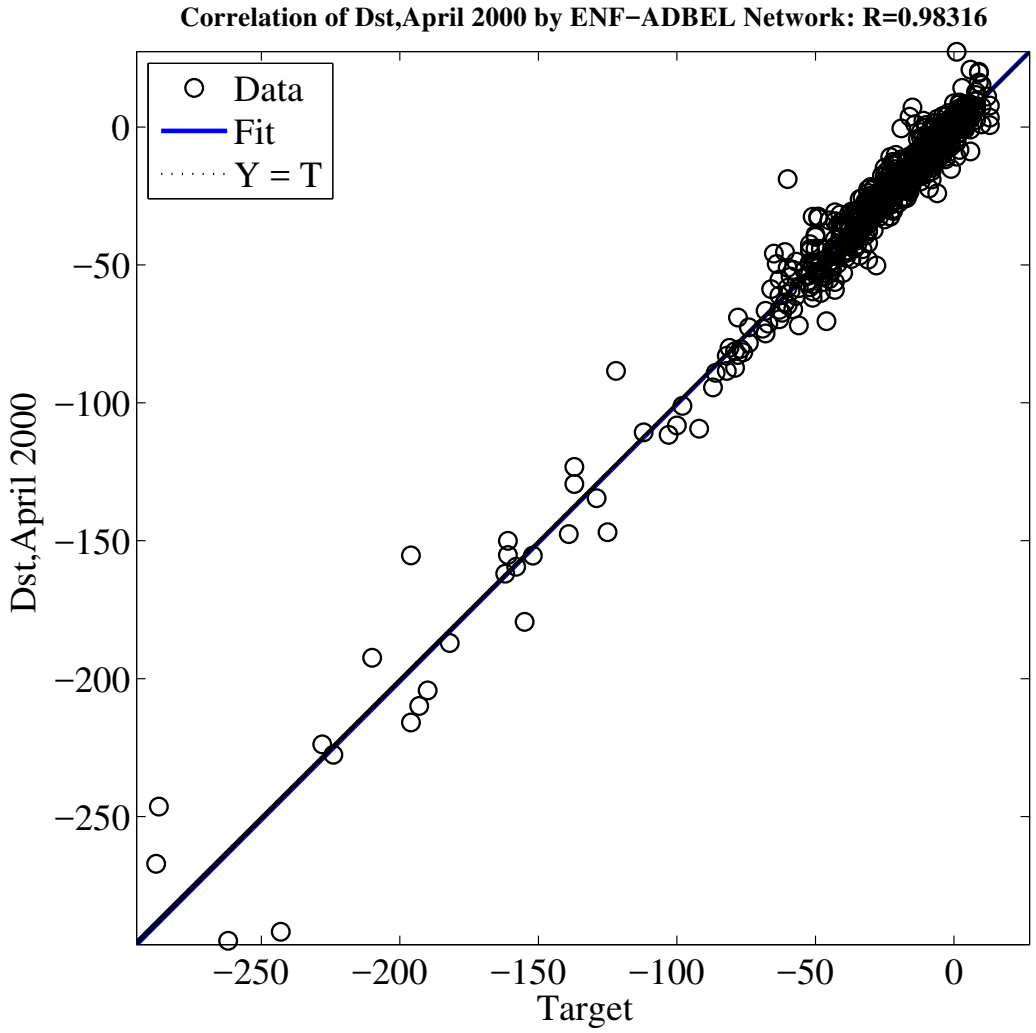

Figure 4.146: Correlation in Predicting D_{st} April 2000 by ENF-ADBEL Network.

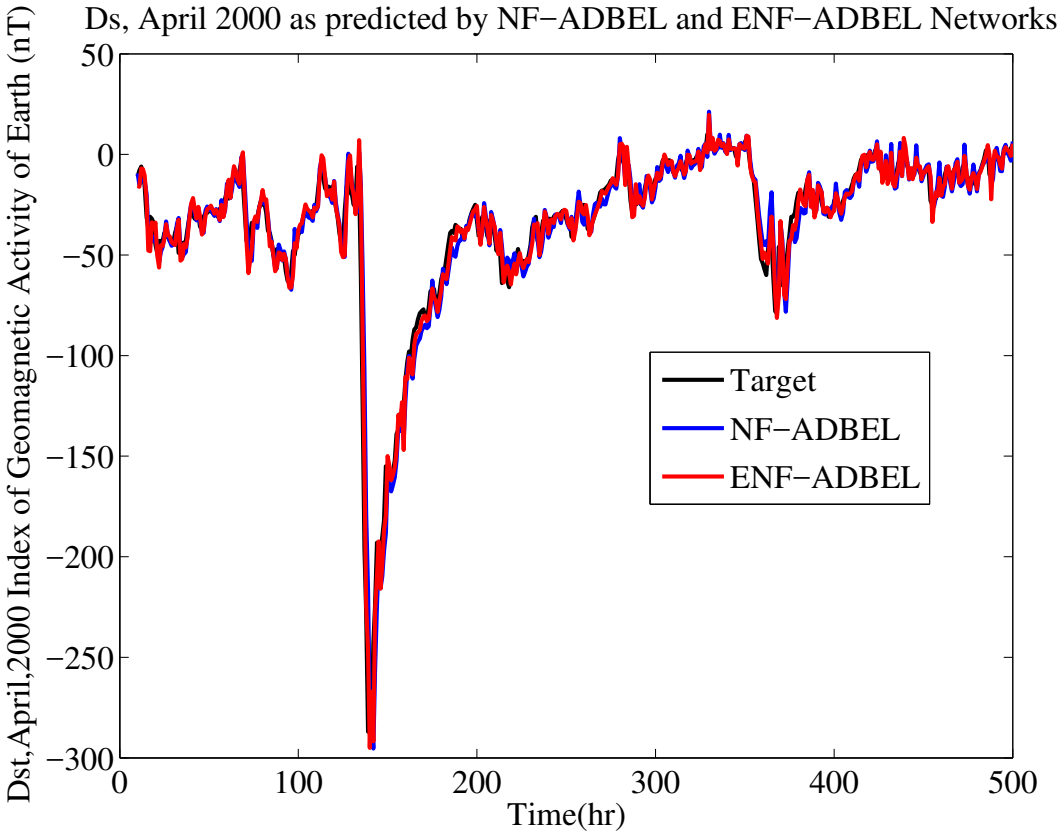

Figure 4.147: D_{st} April 2000 as Predicted by ENF-ADBEL and NF-ADBEL Networks.

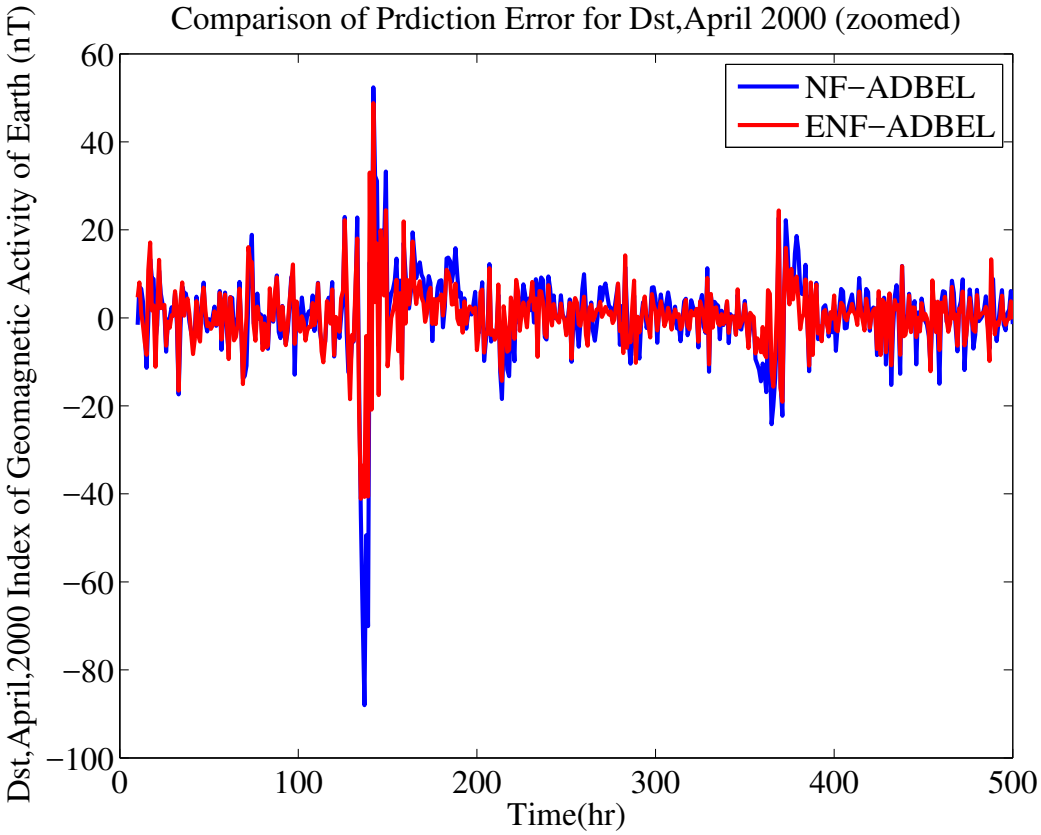

Figure 4.148: Error Comparison in Predicting D_{st} April 2000 as Predicted by ENF-ADBEL and NF-ADBEL Networks.

4.3.4 Wind Speed as Predicted by the Proposed ENF-ADBEL Network

The conventional ways of electricity generation are continuously polluting the environment. Renewable energy resources have the potential of both overcoming the problem of air pollution as well as meeting the load demand. Amongst various renewable energy resources, wind energy offers a viable way to harness electricity owing to its cost-effectiveness and sustainable nature [89]. The available wind power depends on the wind speed. Due to the randomly fluctuating characteristics of wind speed, wind power's prediction results may change rapidly. Accurate wind speed prediction can significantly improve power quality, security, supply-demand balancing and, in general, wind generation management in the smart grid [90]. Therefore, applying wind speed prediction techniques offering the best forecasting accuracy over time scales is required [91].

Approaches for wind speed prediction are classified into two categories: physical methods and statistical methods. Physical methods are referred to as meteorological forecasting to the wind speed, with physical models consisting of numerical approximation models that define the state of the atmosphere [92]. For physical methods, the Numerical Weather Prediction (NWP) technique is generally used. In this approach, meteorological parameters such as temperature, air density, humidity, air pressure, and surface roughness are taken into consideration. While physical methods can predict wind speed more accurately [93], physical models are usually unable to provide reliable wind speed forecasts, especially in complex landscape regions, due to shortcomings in horizontal resolution, topography parameterizations, and initial and boundary conditions.

Furthermore, in terms of short-time forecasting, physical methods are not always suitable because they need a long computational time. To reduce this drawback, hybridization of physical and statistical models [94] may be used. Physical method forecasts have been shown to outperform statistical methods after a 3-6h look-ahead time, while statistical approaches are reliable for short-term prediction, i.e., less than 6h [94]. Based on the time horizons [95], methods for forecasting wind speed can be divided into four categories: concise (from a few seconds to 30min ahead); short-term (from 30min to 6h ahead); medium-term (from 6h to 24h ahead); and long-term (from 24h to 1 week or more ahead). Note that there is no absolute limit to these periods.

The results of the physical models represent the first step towards forecasting wind speed. Hence, the physical model that predicts wind speed can be considered as an auxiliary input to the statistical models [96]. Statistical methods make predictions by finding relationships using historical wind speed data and, sometimes, involving other variables (e.g., wind direction, air pressure or temperature). The data utilized in this technique depend on the data available and recorded at a particular site or other nearby locations.

Statistical models can be divided into two categories: mathematical models and Artificial Intelligence (AI) approaches. Many mathematical methods have been applied to this theme, such as the AutoRegressive (AR) model, AutoRegressive Integrated Moving Average (ARIMA) model, and Kalman Filters (KF), etc. The mathematical models can be used at any stage in the modelling process and are often combined with various other methods [97].

The second approach is AI techniques, such as artificial neural networks (ANNs). In contrast to the mathematical models, the distinctive features of AI methods are their high learning capability and their ability to handle noisy and incomplete data. Furthermore, AI methods predict future time series data without any predefined mathematical models. AI systems have the ability to learn from previous data and attempt to mimic the behaviour for accurate prediction. However, the techniques mentioned above mainly rely on complicated mathematics approaches that require extensive meteorological and topographic data [98],[99],[100].

In this work, a proposed ENF-ADBEL is deployed to forecast wind speed. Hourly wind speed data for three months (January, February and March 2020) from a Canadian meteorological station located at Lunenburg, Nova Scotia, is used as an application of the proposed neural network [101].

The ENF-ADBEL network is first employed using the learning parameters $\alpha = 0.3$, $\beta = 0.015$, and $\gamma = 0.25$. As can be seen, the ENF-ADBEL network can predict wind speed for one hour ahead, as shown in Figures 4.149, 4.150 and 4.151, when the steady-state starting index is taken to be $n_s = 1hr$. To compare its identification performance with the NF-ADBEL network, the simulation is run with the learning parameters for the NF-ADBEL network set as $\alpha = 0.77$, $\beta = 0.04$, and $\gamma = 0.19$. Figures 4.152, 4.153, and 4.154 show the NF-ADBEL performance. The forecasting

error and correlation indices for both networks are presented in Table 4.16.

The temporary period for the ENF-ADBEL network is found to be the same as that of the NF-ADBEL network. However, the ENF-ADBEL network shows better performance compared to the NF-ADBEL networks, owing to the lesser forecasting error being offered by this network during steady-state, as can be seen from Figures 4.150, 4.153, and 4.156. In terms of run-time, the proposed ENF-ADBEL accomplished the performance in 1.17 seconds, while NF-ADBEL took 1.23 seconds. A lower root mean squared error, higher correlation coefficient and sufficient percentage improvement as yielded by the ENF-ADBEL network validates its adequate performance over the NF-ADBEL network in forecasting wind speed, as shown in Table 4.16.

Table 4.16: RMSE & R^2 for Wind Speed in ENF-ADBEL and NF-ADBEL Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Wind Speed	ENF-ADBEL	5.32	90.18	4.93
	NF-ADBEL	5.596	88.68	

Further, a multilayer perceptron (MLP) neural network was used for the same wind speed data to validate the proposed ENF-ADBEL. A comparison of the results is presented in Table 4.19. As can be seen, the ENF-ADBEL network performed with better accuracy.

Table 4.17: RMSE & R^2 for Wind Speed ENF-ADBEL and MLP Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
wind speed	ENF-ADBEL	5.28	89.08	2.76
	MLP	5.43	88.81	

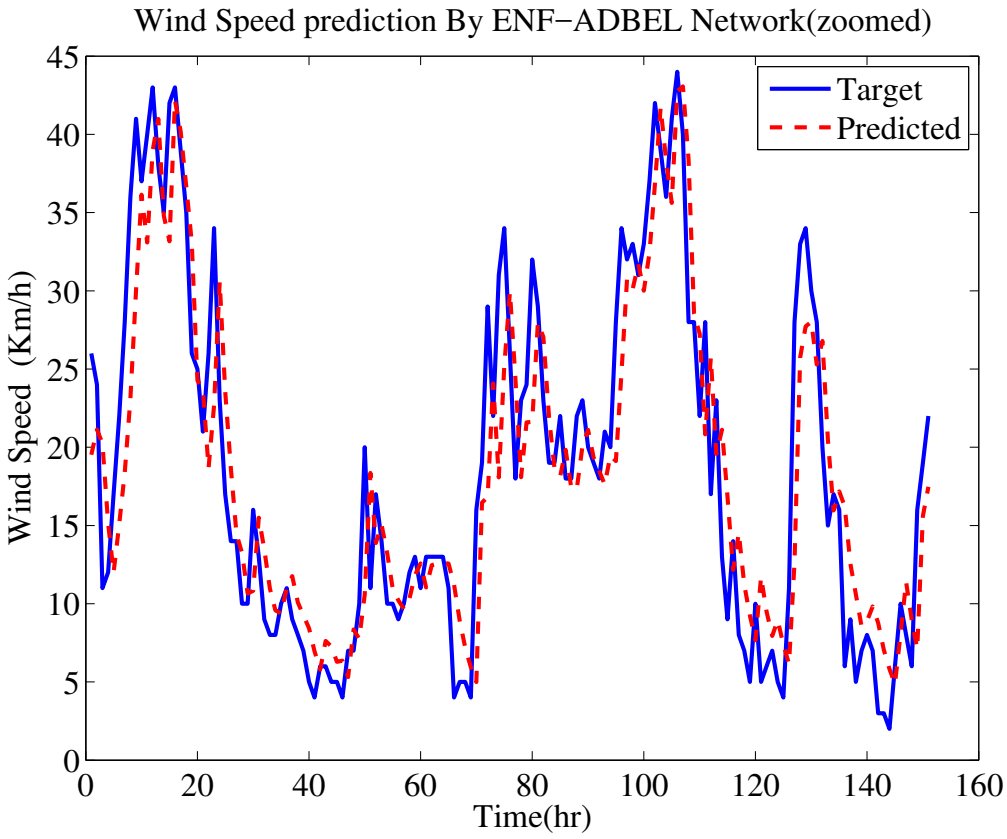

Figure 4.149: Wind Speed as Predicted by ENF-ADBEL Network.

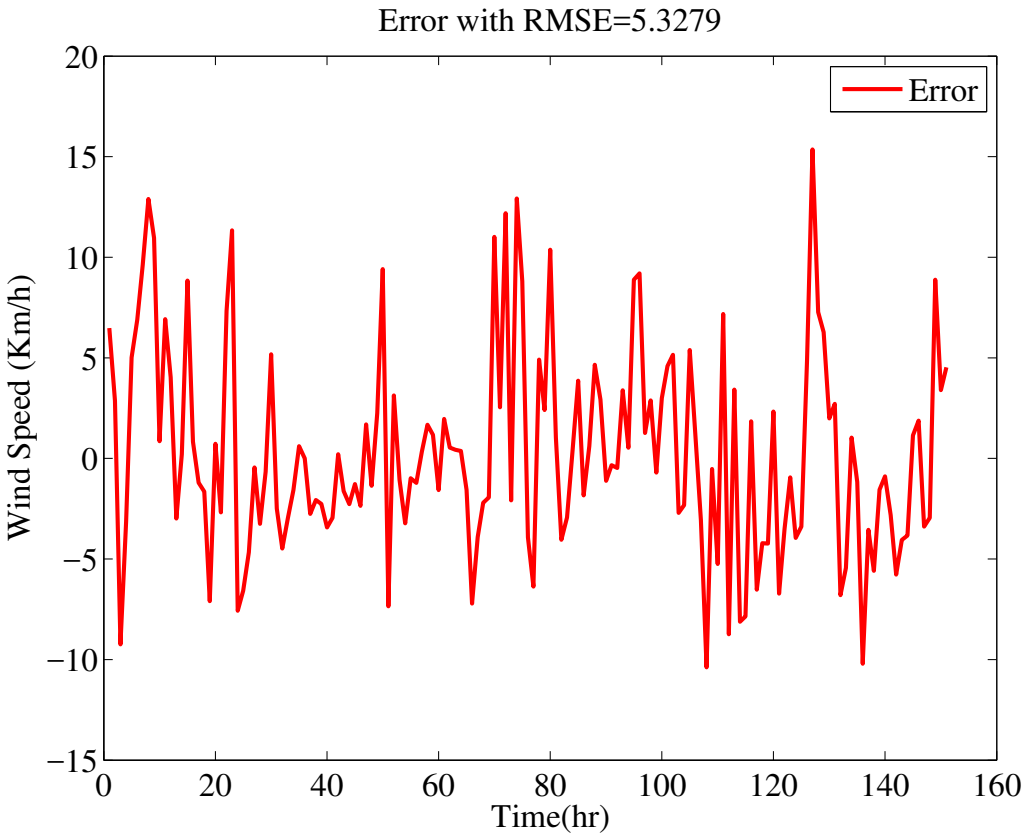

Figure 4.150: Error in Predicting Wind Speed by ENF-ADBEL Network.

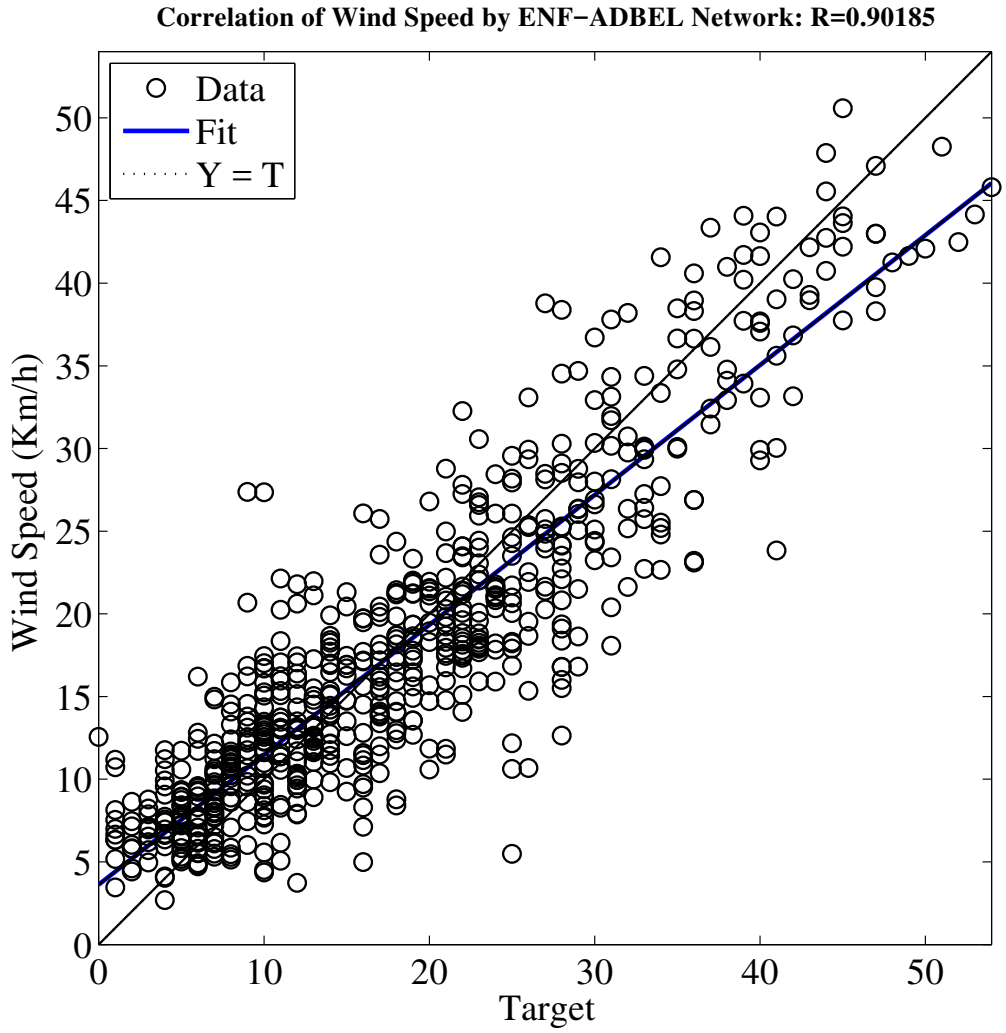

Figure 4.151: Correlation in Predicting Wind Speed by ENF-ADBEL Network.

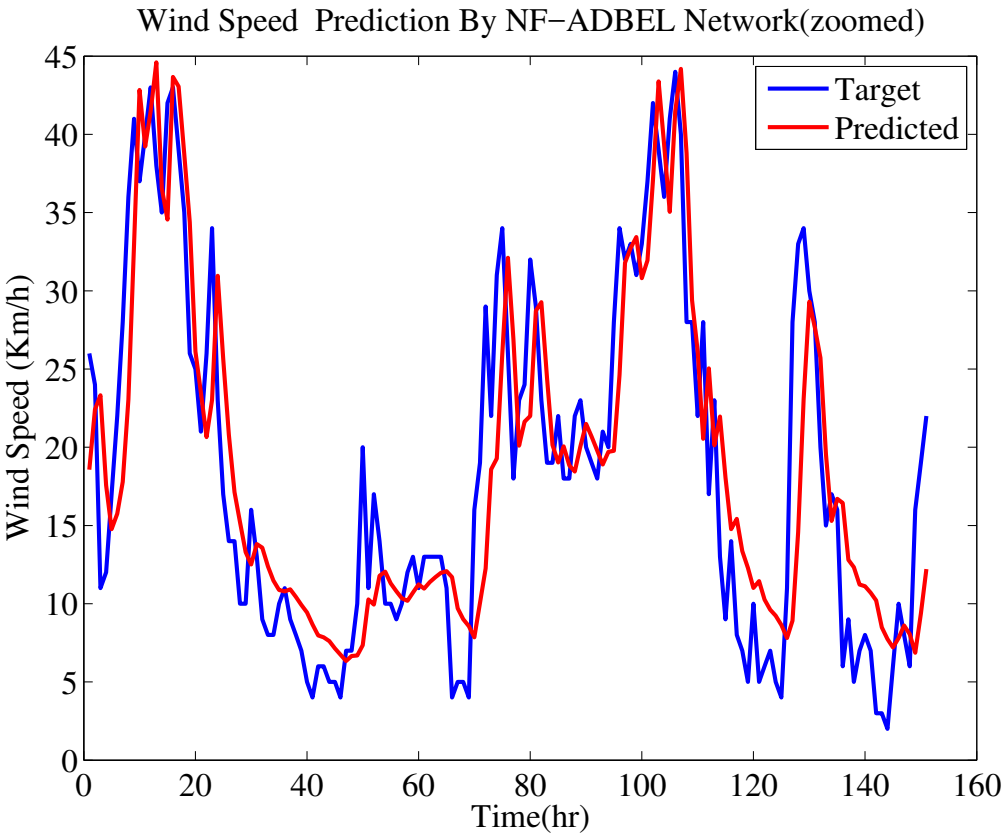

Figure 4.152: Wind Speed as Predicted by NF-ADBEL Network.

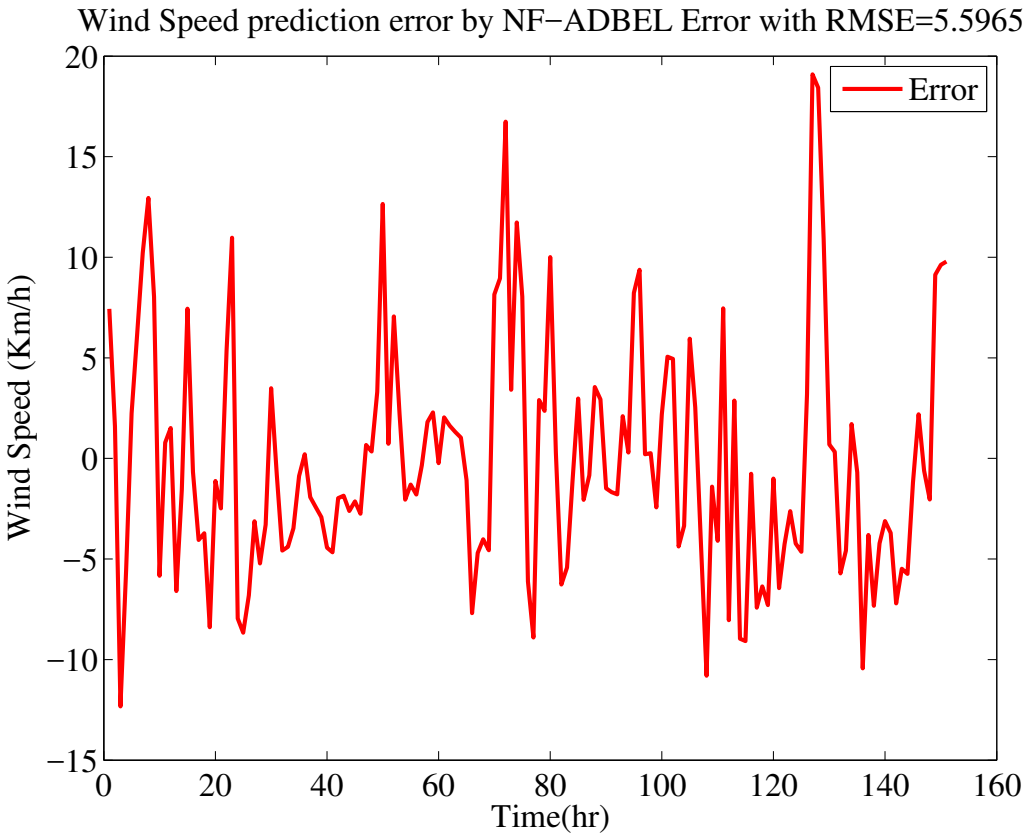

Figure 4.153: Error in Predicting Wind Speed by the Proposed NF-ADBEL Network.

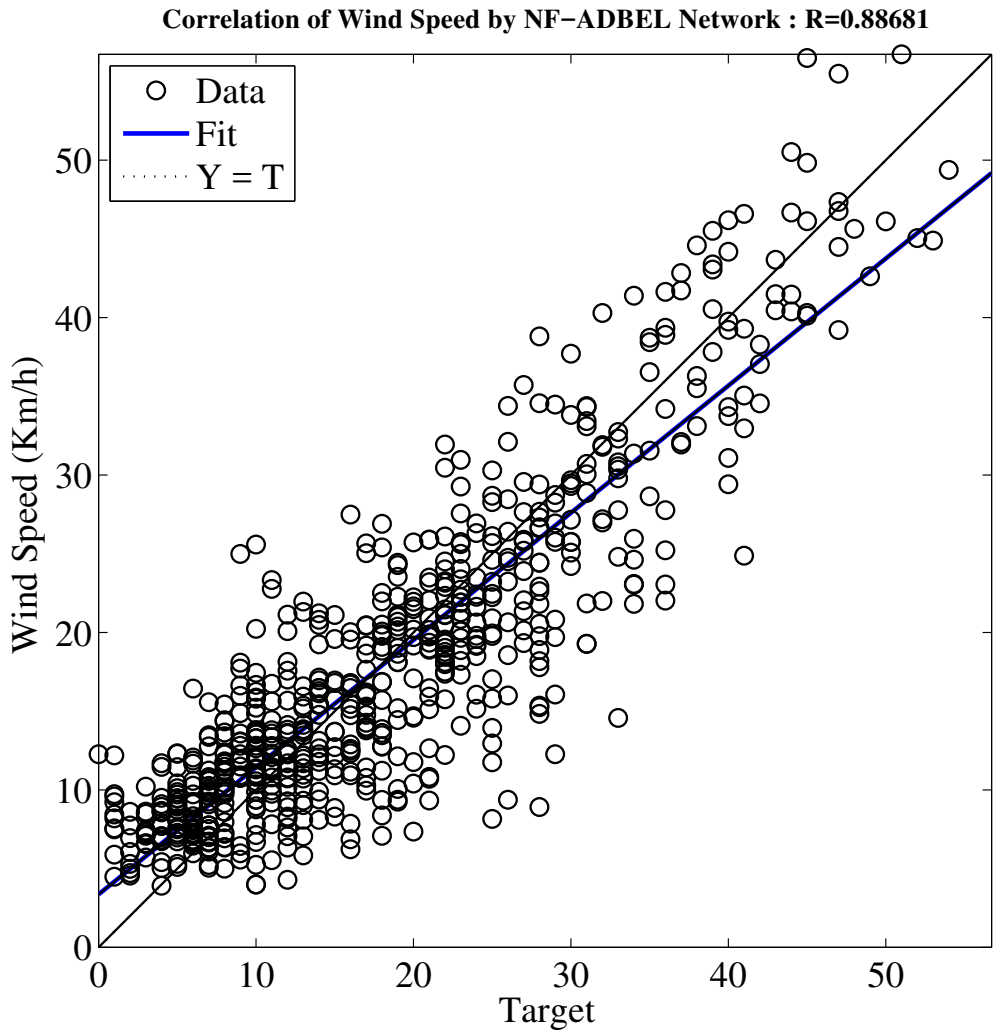

Figure 4.154: Correlation in Predicting Wind Speed by the Proposed NF-ADBEL Network.

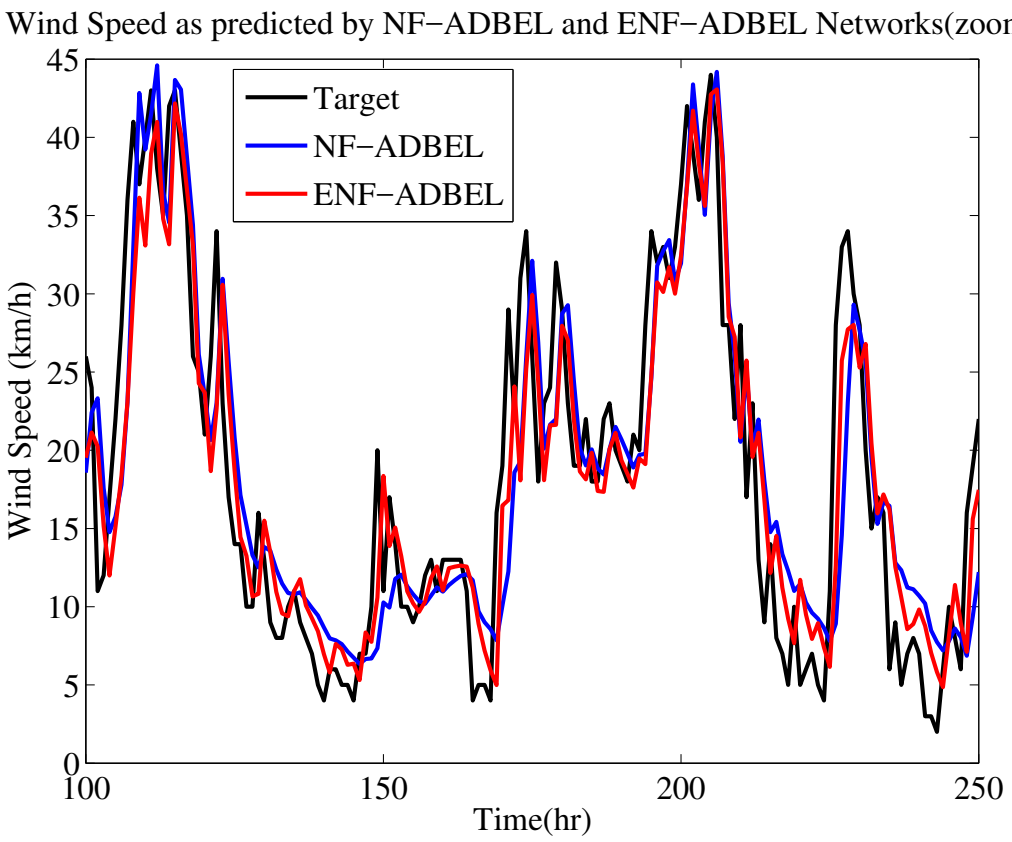

Figure 4.155: Wind Speed as Predicted by ENF-ADBEL and NF-ADBEL Networks.

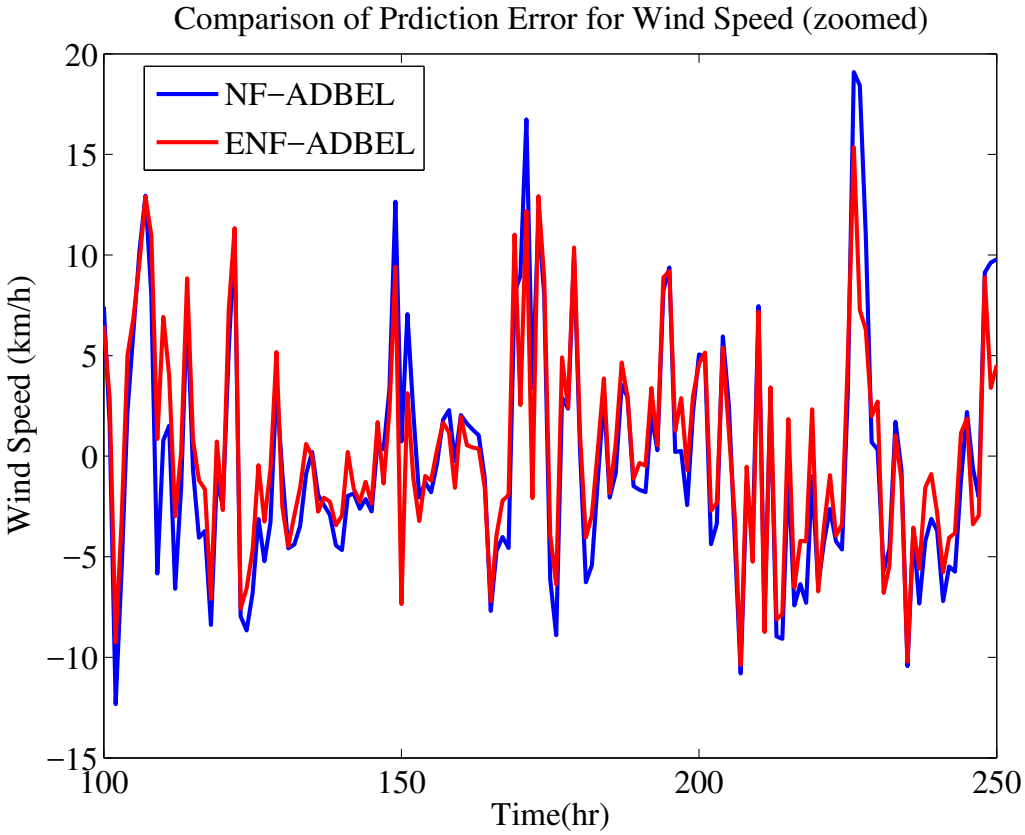

Figure 4.156: Error Comparison in Predicting Wind Speed as Predicted by ENF-ADBEL and NF-ADBEL Networks.

4.3.5 Wind Power as Predicted by the Proposed ENF-ADBEL Network

As an alternative source of clean power, wind power plays a crucial role in the secure management of the power system. However, wind power also has a major disadvantage in that it is dependent on the wind speed. Moreover, because wind speed is a highly stochastic and intermittent feature of wind energy, accurate forecasting models are needed that can provide information ahead of time to protect the power stability and predict the wind energy source's output before connecting to the grid.

In recent decades, extensive efforts have been made to develop efficient wind power-forecasting models at multiple scales. Accurate wind power forecasting can help to arrange generation plans, maintain grid stability, and provide a reliable basis for grid operation [102]. Different approaches are used for different time scales and different data sources [103]. For instance, statistical methods and learning methods are used based on the history of wind power data. Among statistical methods are the time series method [104], regression analysis method, and Kalman filter method [105]. Learning methods use deep long short-term memory (LSTM) [106],[107] to predict wind power and compare the results to the support vector machine (SVM) approach. Backpropagation neural networks (BPNNs) have shown that LSTM has more accurate power prediction.

In [108], seven-day-ahead hourly wind power data are employed as an application for the proposed ENF-ADBEL network. These wind power data indicate the amount available to the grid in Alberta, Canada, on a seven-day ahead basis by updating the energy statistics every six hours. The data indicate three levels of estimation wind power availability: minimum wind power expectation, most likely available wind power, and maximum wind power forecast.

Firstly, the proposed ENF-ADBEL network is simulated for predicting this wind power in terms of minimum wind power data. The learning parameters are selected as $\alpha = 0.47$, $\beta = 0.32$, and $\gamma = 0.14$. Figures 4.157, 4.158, and 4.159 show the performance of the proposed ENF-ADBEL network.

The NF-ADBEL network, which is driven by the parameters $\alpha = 0.4$, $\beta = 0.5$, and $\gamma = 0.13$, is also simulated to forecast a minimum wind power. Figures 4.160, 4.161, and 4.162 show the performance of the NF-ADBEL network. The results of the

proposed models in terms of low root mean square error and high correlation are presented in Table4.18. In terms of run-time, the proposed ENF-ADBEL accomplished the performance in 0.94 seconds, while NF-ADBEL took one second.

A comparison of both networks' prediction error performance is illustrated in Figures 4.163 and 4.164. As can be seen, the amplitude of the error signal for the ENF-ADBEL network is lower compared to the NF-ADBEL networks, as shown in Figure 4.164. This indicates that the ENF-ADBEL network offers better prediction accuracy. Analysis of the predicted results for minimum power forecasting data in terms of the root mean squared error and correlation coefficient criteria are shown in Figures 4.158 and 4.159, respectively. Again, the ENF-ADBEL network shows better performance than the NF-ADBEL. A reasonable amount of percentage improvement is yielded by the ENF-ADBEL network for predicting the minimum wind power, as can be seen in Table 4.18.

Table 4.18: RMSE, R^2 for Minimum Wind Power Predicted by ENF-ADBEL Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Wind power	ENF-ADBEL	16.28	99.70	15.03
	NF-ADBEL	19.16	99.59	

We also applied the MLP neural network for the same wind power data (minimum power) used to validate the proposed ENF-ADBEL. The comparison results are given in Table 4.19. As can be seen, the ENF-ADBEL network performed with better accuracy.

Table 4.19: RMSE, R^2 for Minimum Wind Power Predicted by ENF-ADBEL and MLP Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
min wind power	ENF-ADBEL	16.28	99.7	33.95
	MLP	24.65	98.15	

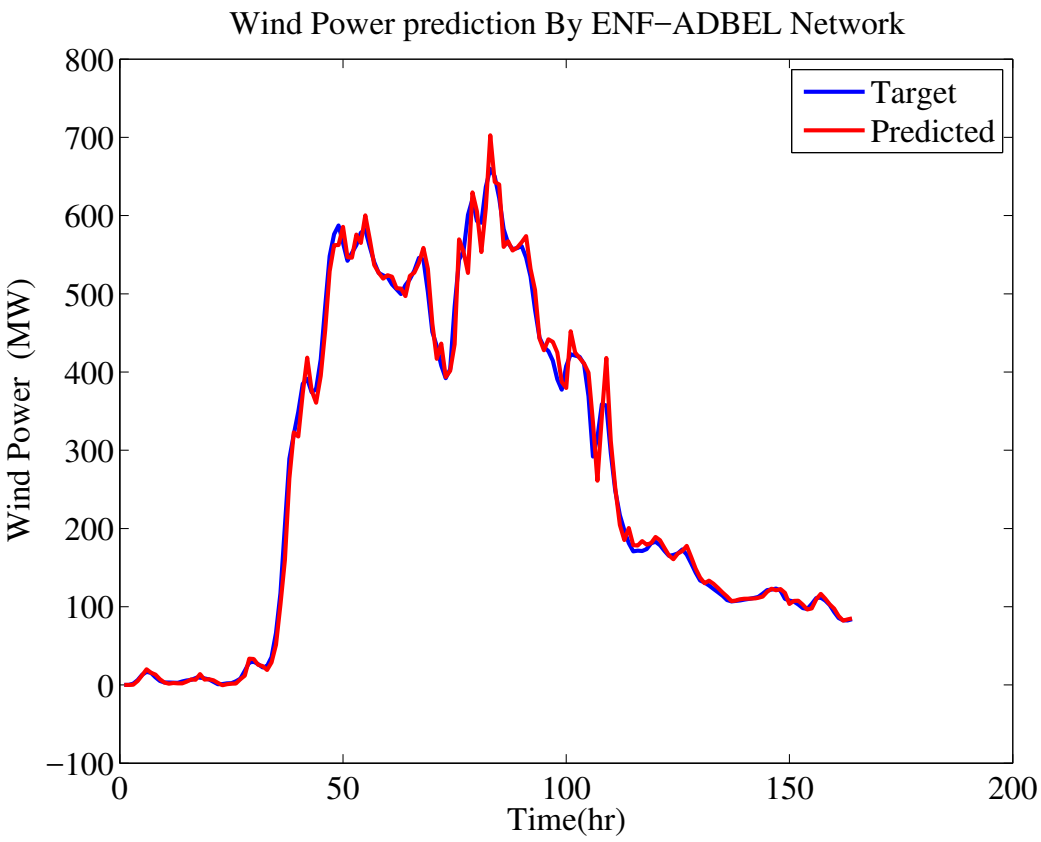

Figure 4.157: Minimum Wind Power as Predicted by ENF-ADBEL Network.

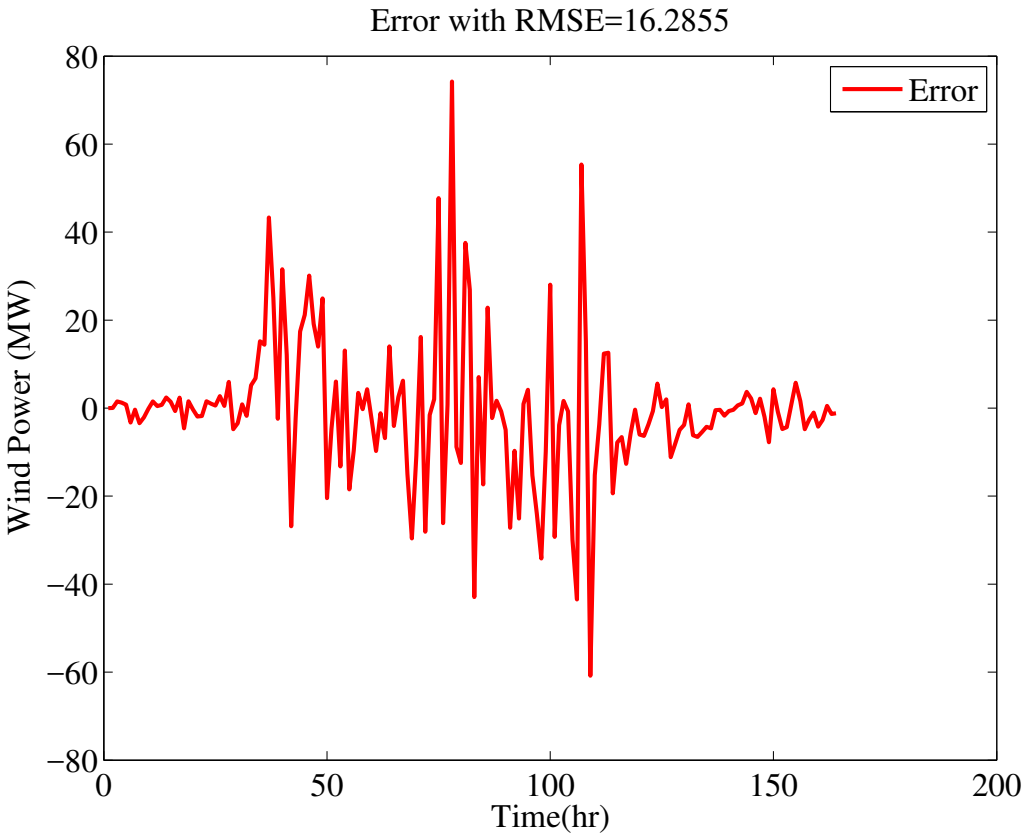

Figure 4.158: Error in Predicting Minimum Wind Power by ENF-ADBEL Network.

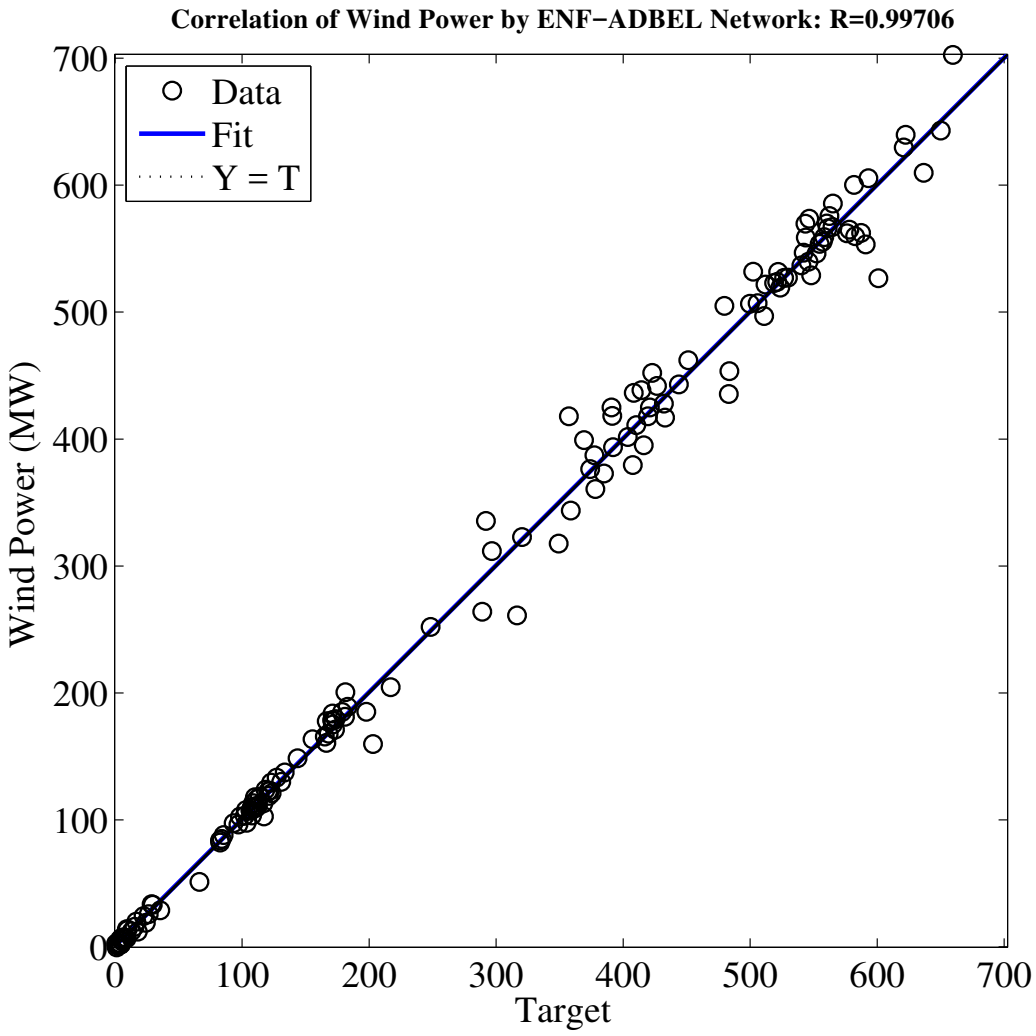

Figure 4.159: Correlation in Predicting Minimum Wind Power by ENF-ADBEL Network.

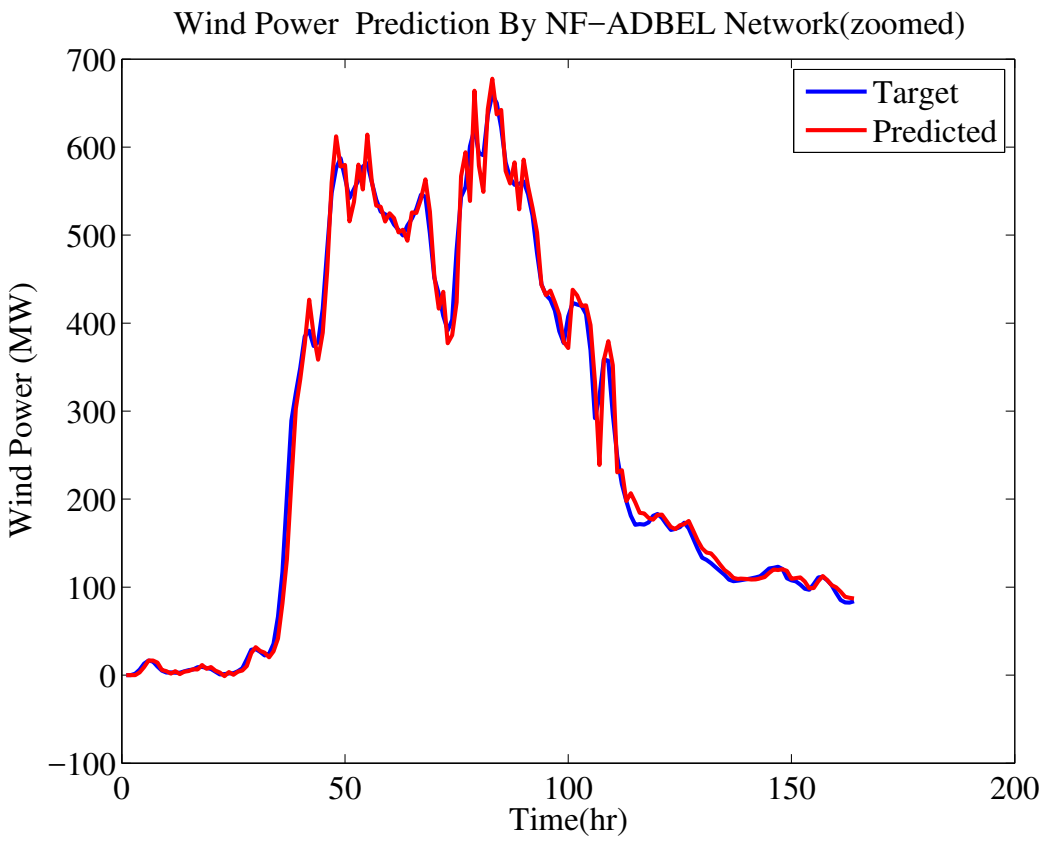

Figure 4.160: Minimum Wind Power as Predicted by NF-ADBEL Network.

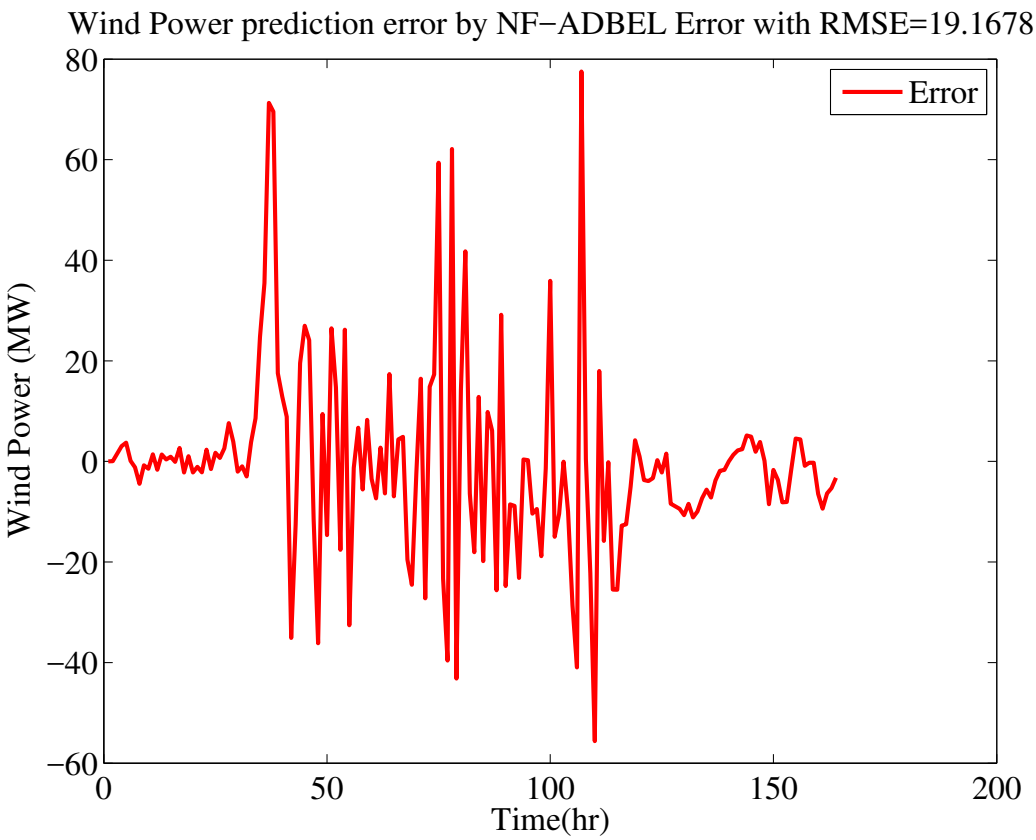

Figure 4.161: Error in Predicting Minimum Wind Power by NF-ADBEL Network.

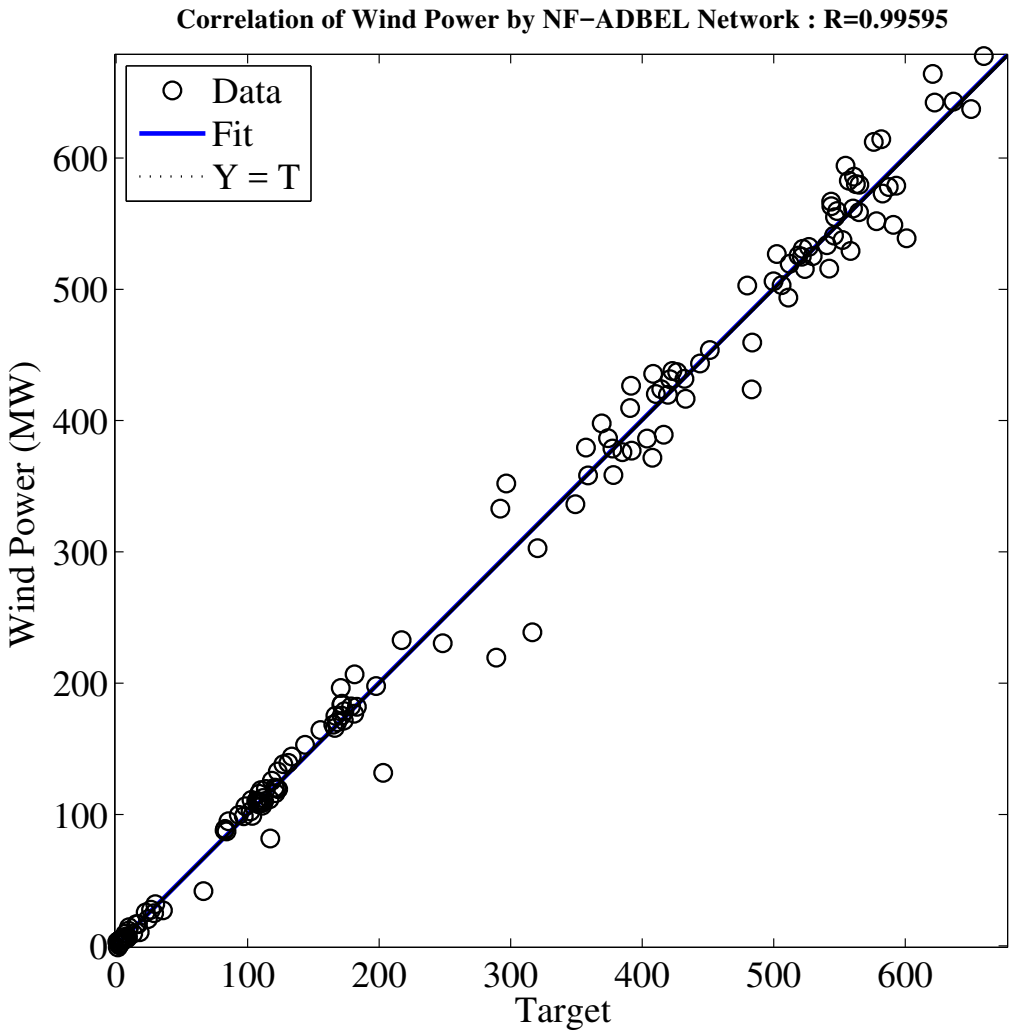

Figure 4.162: Correlation in Predicting Minimum Wind Power by NF-ADBEL Network.

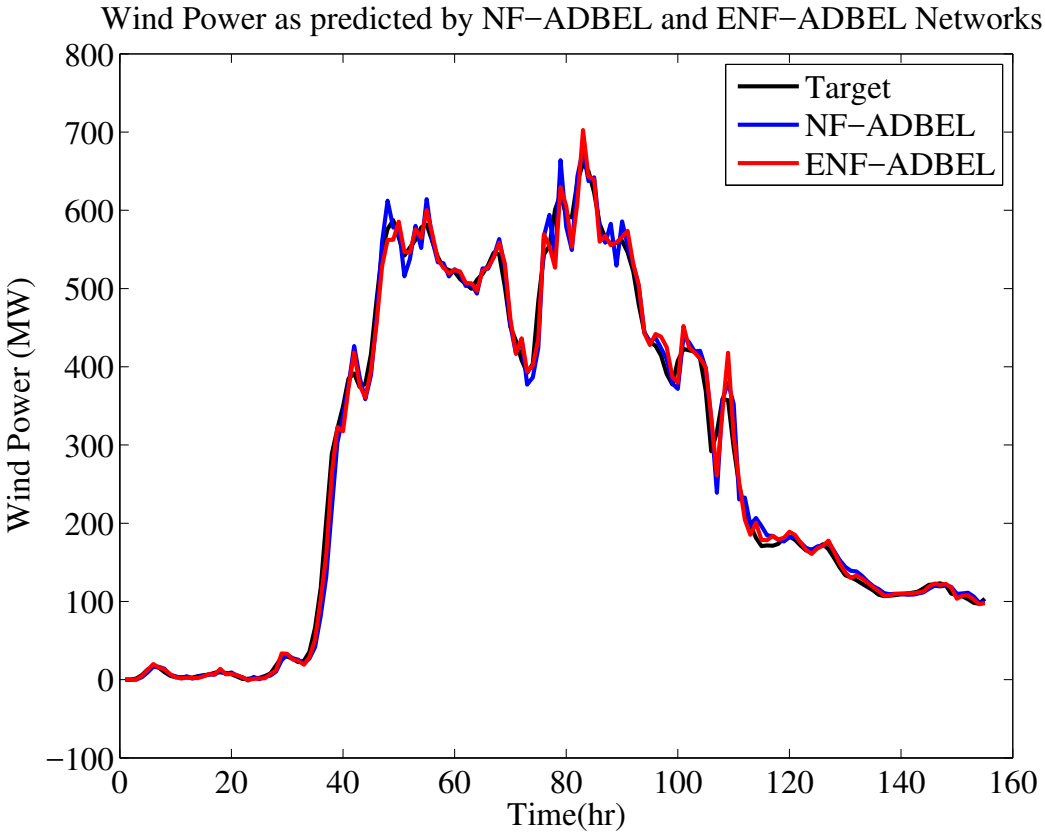

Figure 4.163: Minimum Wind Power as Predicted by ENF-ADBEL and NF-ADBEL Networks.

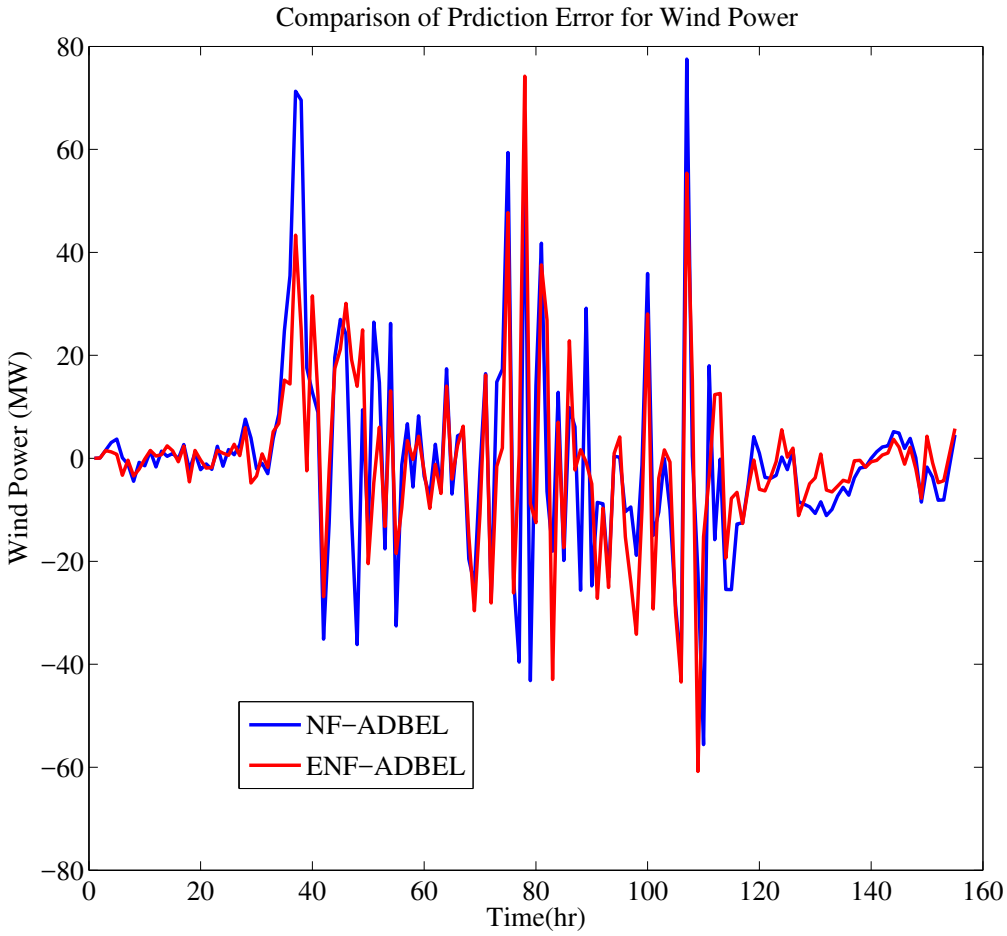

Figure 4.164: Error Comparison in Predicting Minimum Wind Power as Predicted by ENF-ADBEL and NF-ADBEL Networks.

Secondly, the proposed ENF-ADBEL network is deployed for predicting the most-likely wind power data, with the learning parameters selected as $\alpha = 0.4$, $\beta = 0.44$, and $\gamma = 0.24$. Figures 4.165, 4.166, and 4.167 show the performances of the proposed ENF-ADBEL network. Meanwhile, the NF-ADBEL network, utilizing the parameters $\alpha = 0.43$, $\beta = 0.49$ and $\gamma = 0.13$, is also simulated to forecast a most-likely wind power. Figures 4.168, 4.169, and 4.170 show the performance of the NF-ADBEL network. The results in terms of low root mean square error and high correlation are given in Table 4.20. In terms of run-time, the proposed ENF-ADBEL accomplished the performance in 2.25 seconds, while the NF-ADBEL took 1 second.

A comparison of both networks for prediction error is displayed in Figures 4.171 and 4.172. As can be seen, the amplitude of the error signal for the ENF-ADBEL network is lower compared to that of the NF-ADBEL, which shows significant fluctuations (Figure 4.172). Further, the ENF-ADBEL network obtained a lower prediction error and offered better performance accuracy. Analysis of the predicted results for most likely power forecasting data in terms of the root mean squared error and correlation coefficient criteria is given in Figures 4.166 and 4.167, respectively, with the ENF-ADBEL network showing better results than the NF-ADBEL network. Overall, a fair amount of percentage improvement is yielded by the ENF-ADBEL network for predicting the most likely wind power, as presented in Table 4.20.

Table 4.20: RMSE, R^2 for Most-Likely Wind Power Predicted by ENF-ADBEL, NF-ADBEL and F-ADBEL Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Wind power	ENF-ADBEL	17.34	99.87	12.82
	NF-ADBEL	19.89	99.83	

As well, we deployed an MLP neural network for the same wind power data (most-likely power) to validate the proposed ENF-ADBEL. The comparison results are given in Table 4.21, showing that the ENF-ADBEL network performed better accuracy.

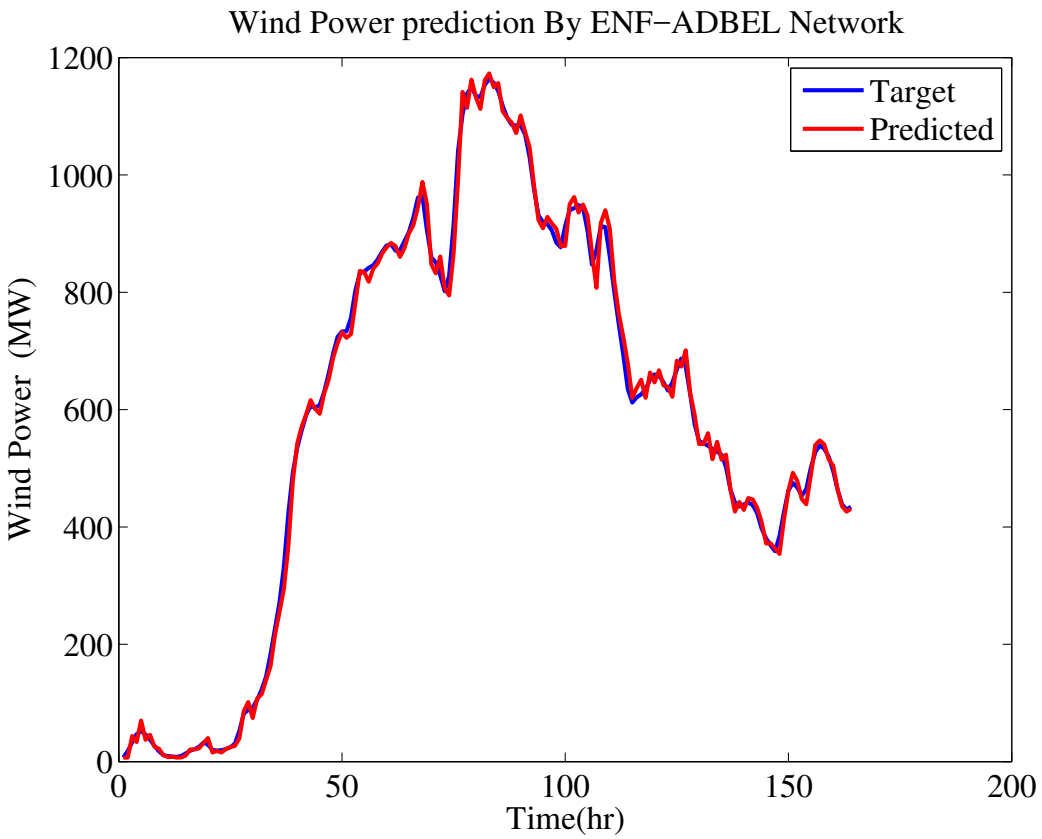

Figure 4.165: Most-Likely Wind Power as Predicted by ENF-ADBEL Network.

Table 4.21: RMSE, R^2 for Most-Likely Wind Power Predicted by ENF-ADBEL and MLP Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
most wind power	ENF-ADBEL	17.34	99.87	44.51
	MLP	31.25	99.79	

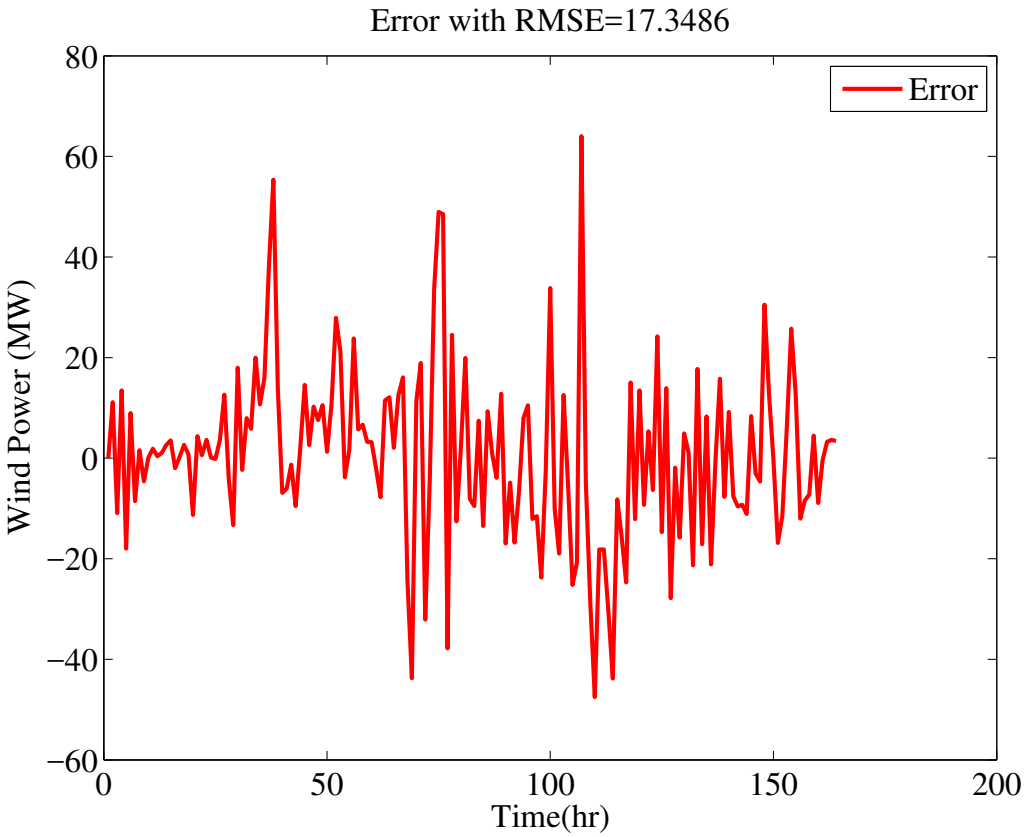

Figure 4.166: Error in Predicting Most-Likely Wind Power by Proposed ENF-ADBEL Network.

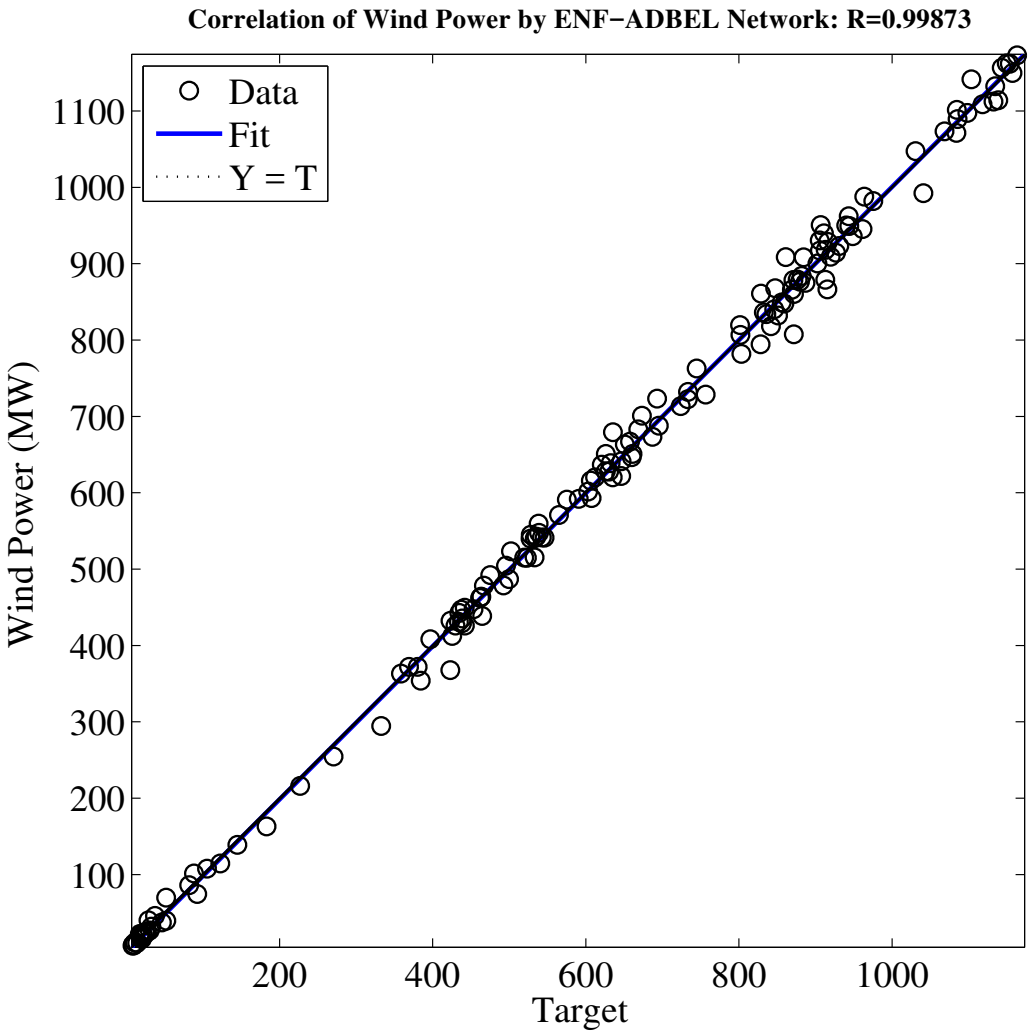

Figure 4.167: Correlation in Predicting Most-Likely Wind Power by ENF-ADBEL Network.

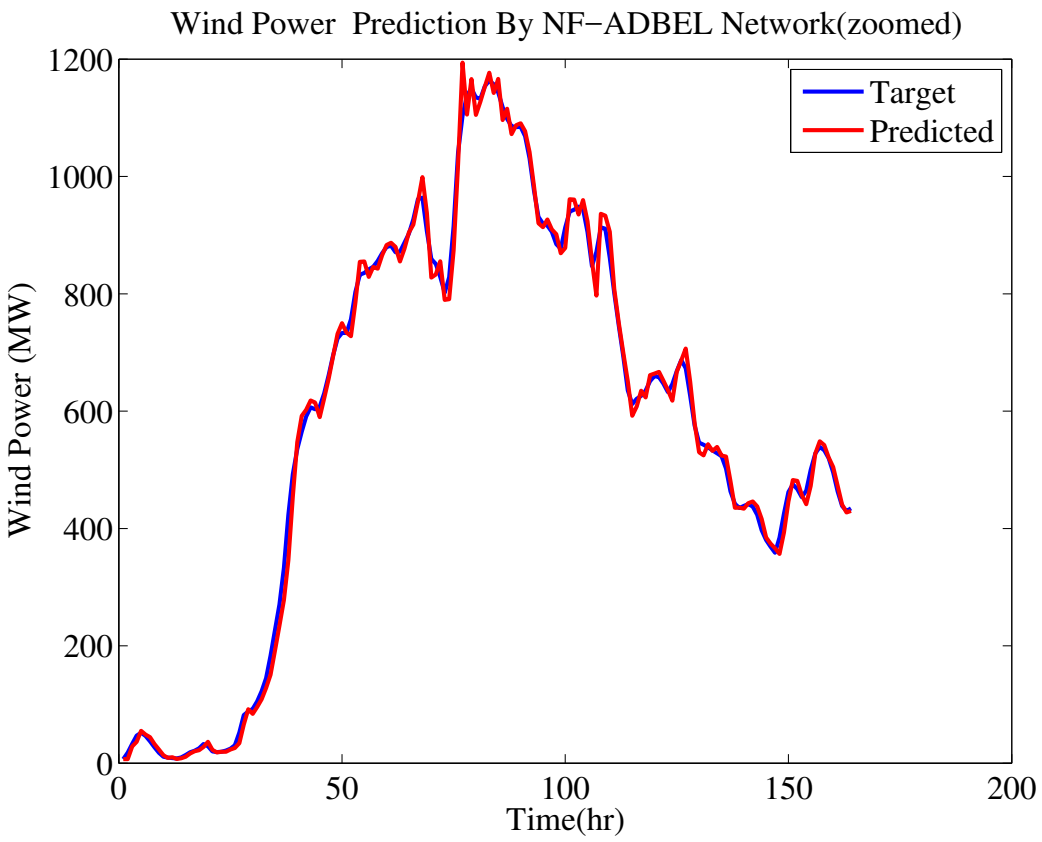

Figure 4.168: Most-Likely Wind Power as Predicted by NF-ADBEL Network.

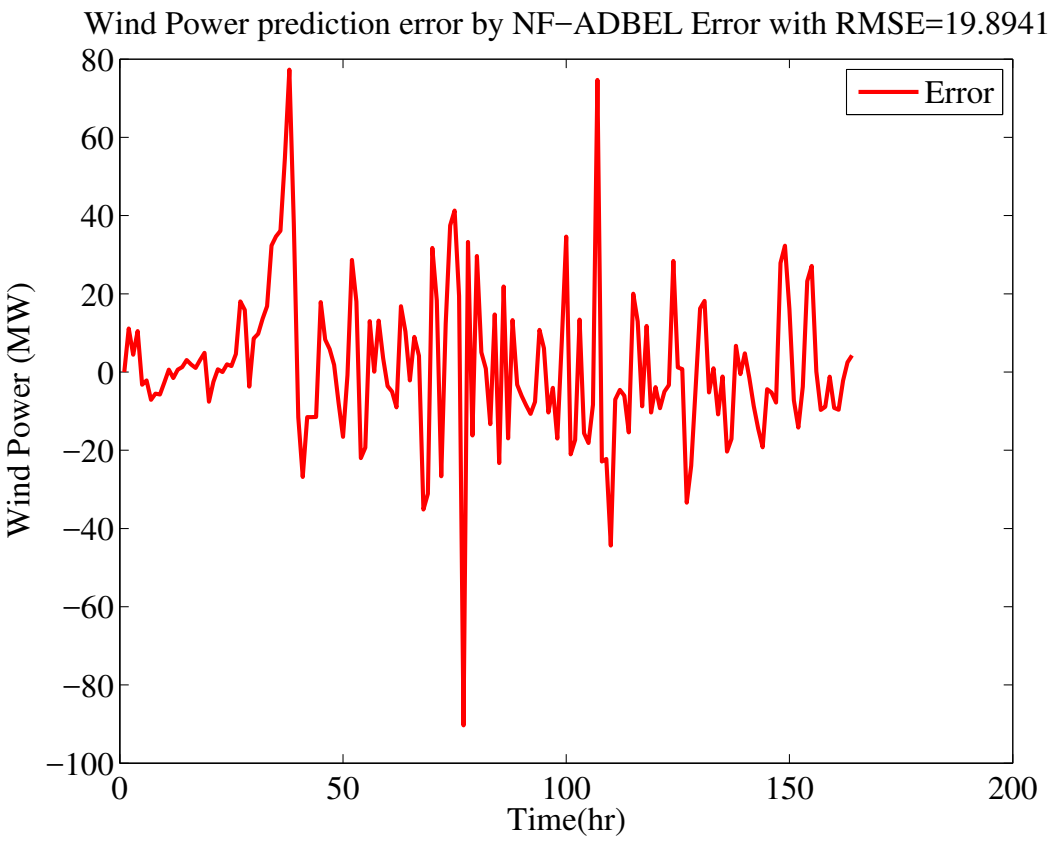

Figure 4.169: Error in Predicting Most-Likely Wind Power by NF-ADBEL Network

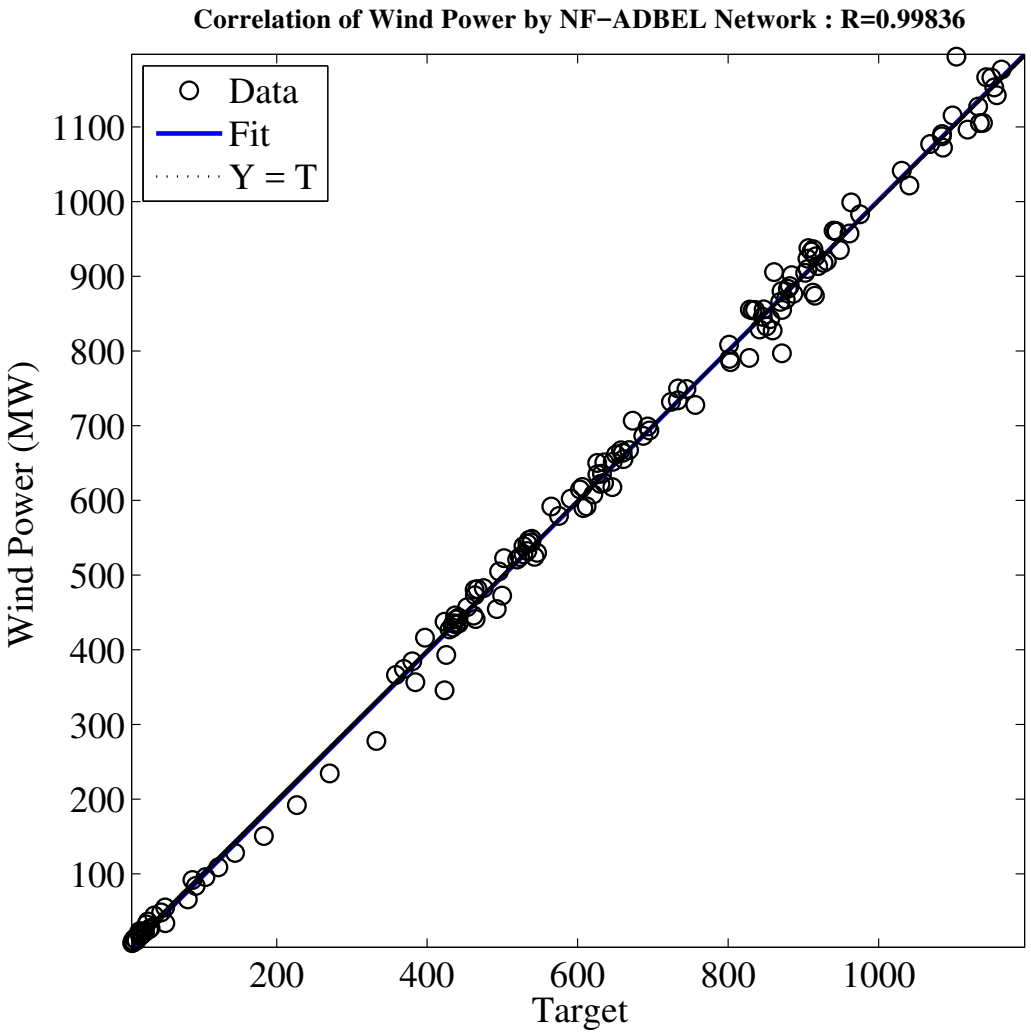

Figure 4.170: Correlation in Predicting Most-Likely Wind Power by NF-ADBEL Network.

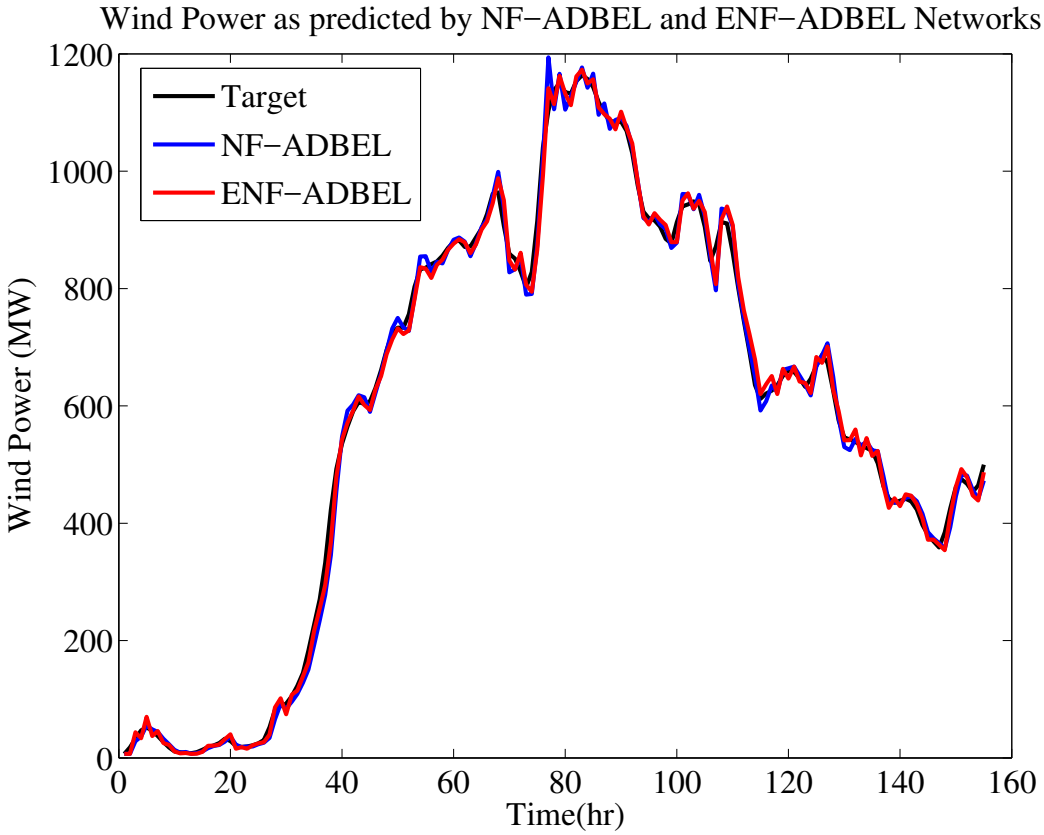

Figure 4.171: Most-Likely Wind Power as Predicted by ENF-ADBEL and NF-ADBEL Networks.

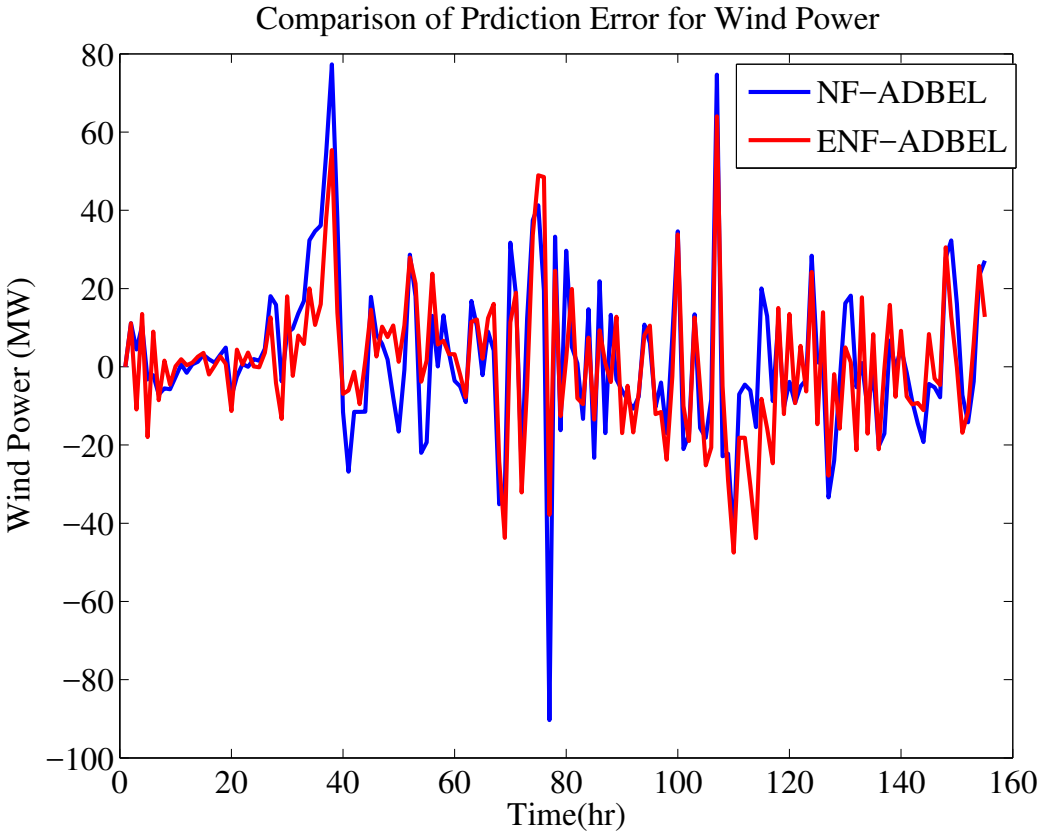

Figure 4.172: Error Comparison in Predicting Most-Likely Wind Power as Predicted by ENF-ADBEL and NF-ADBEL Networks.

Finally, the proposed ENF-ADBEL network is arranged for predicting data on maximum wind power, with the learning parameters selected as $\alpha = 0.23$, $\beta = 0.45$, and $\gamma = 0.09$. Figures 4.173, 4.174, and 4.175 show the performances of the proposed ENF-ADBEL network. Meanwhile, the NF-ADBEL network applied the parameters $\alpha = 0.2$, $\beta = 0.5$, and $\gamma = 0.18$ to simulate the forecast of maximum wind power. Figures 4.176, 4.177, and 4.178 illustrate the performance of that network. The results in terms of low root mean square error and high correlation are presented in Table 4.22. In terms of run-time, the proposed ENF-ADBEL accomplished the performance in 2 seconds, whereas NF-ADBEL took 2.15 seconds.

The prediction error for all networks is displayed in Figures 4.179 and 4.180. As can be seen, the amplitude of error signal for the ENF-ADBEL network is lower compared to that of the NF-ADBEL network, as depicted in Figure 4.180. The ENF-ADBEL network offers slightly better performance accuracy. Analysis of the predicted results for maximum power forecasting in terms of the root mean squared error and correlation coefficient criteria is provided in Figures 4.175 and 4.178, respectively. As can be seen, ENF-ADBEL shows better results than either the NF-ADBEL or F-ADBEL networks. Overall, a fair amount of percentage improvement is yielded by the ENF-ADBEL network for predicting the maximum wind power, as can be seen in Table 4.22. Therefore, the proposed ENF-ADBEL network illustrates the best fitting ability for multiple wind power series among all the implemented networks.

Table 4.22: RMSE, R^2 for Max Wind Power in ENF-ADBEL and NF-ADBEL Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Wind power	ENF-ADBEL	29.75	99.78	0.53
	NF-ADBEL	29.91	99.78	

Addition, we applied an MLP neural network for the same wind power data (max power) to validate the proposed ENF-ADBEL. The comparison results, as given in Table 4.23, show that the ENF-ADBEL network performed with better accuracy.

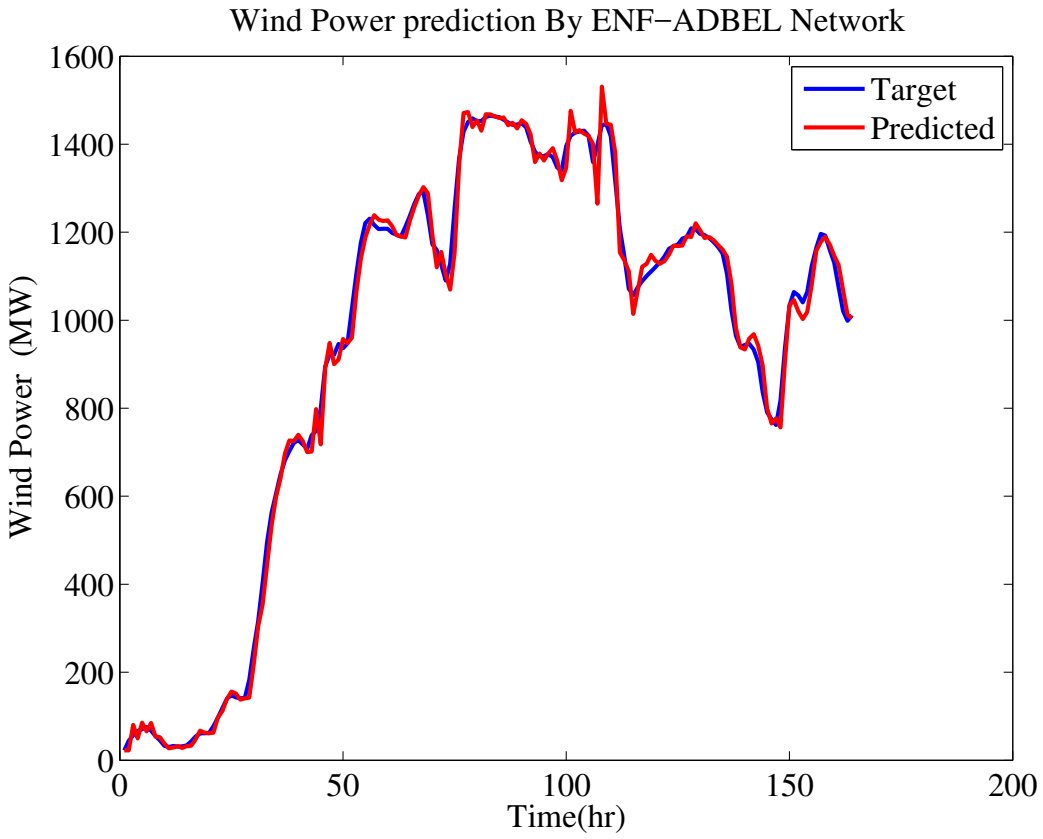

Figure 4.173: Maximum Wind Power as Predicted by ENF-ADBEL Network.

Table 4.23: RMSE, R^2 for Max Wind Power in ENF-ADBEL and MLP Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
max wind power	ENF-ADBEL	29.75	99.78	25.13
	MLP	39.74	99.25	

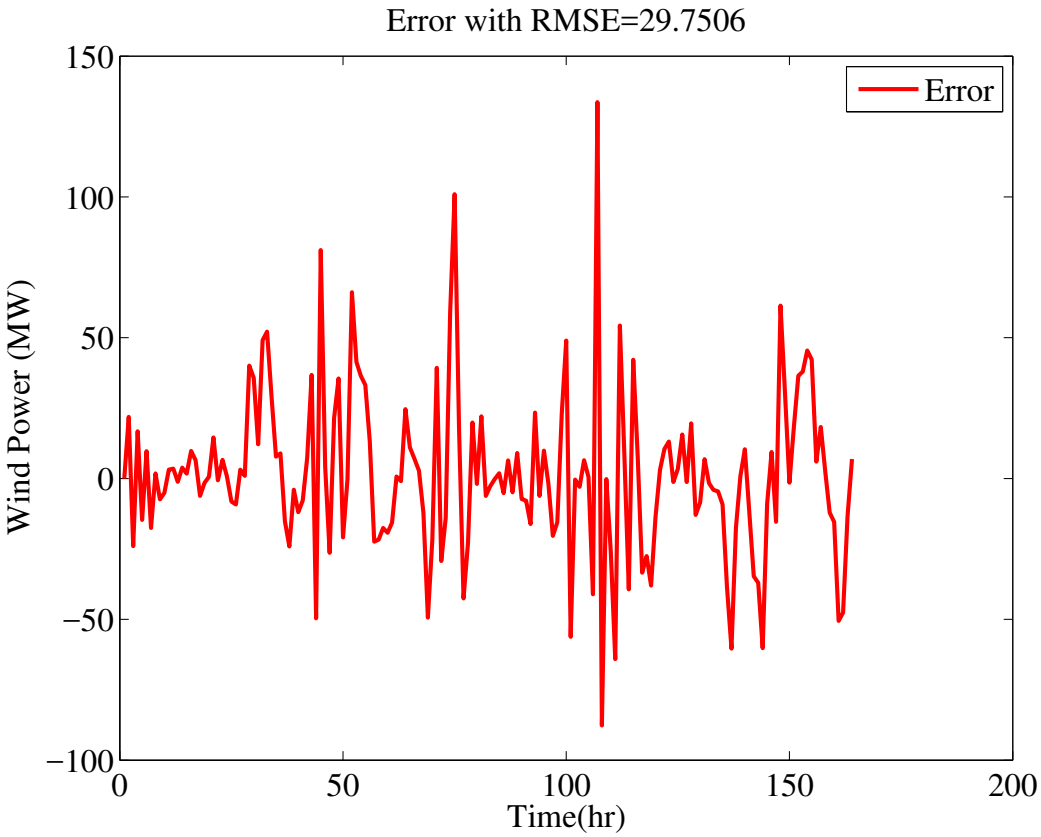

Figure 4.174: Error in Predicting Maximum Wind Power by ENF-ADBEL Network.

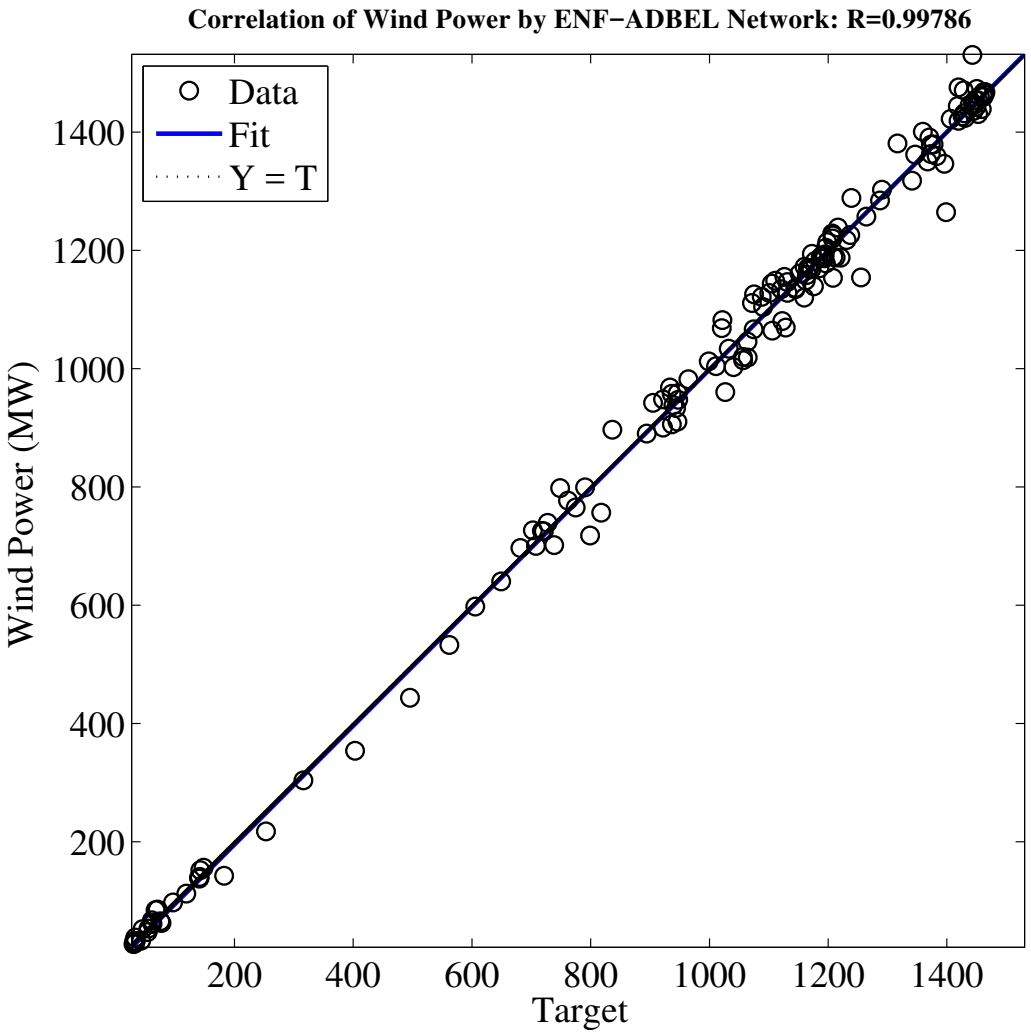

Figure 4.175: Correlation in Predicting Maximum Wind Power by ENF-ADBEL Network.

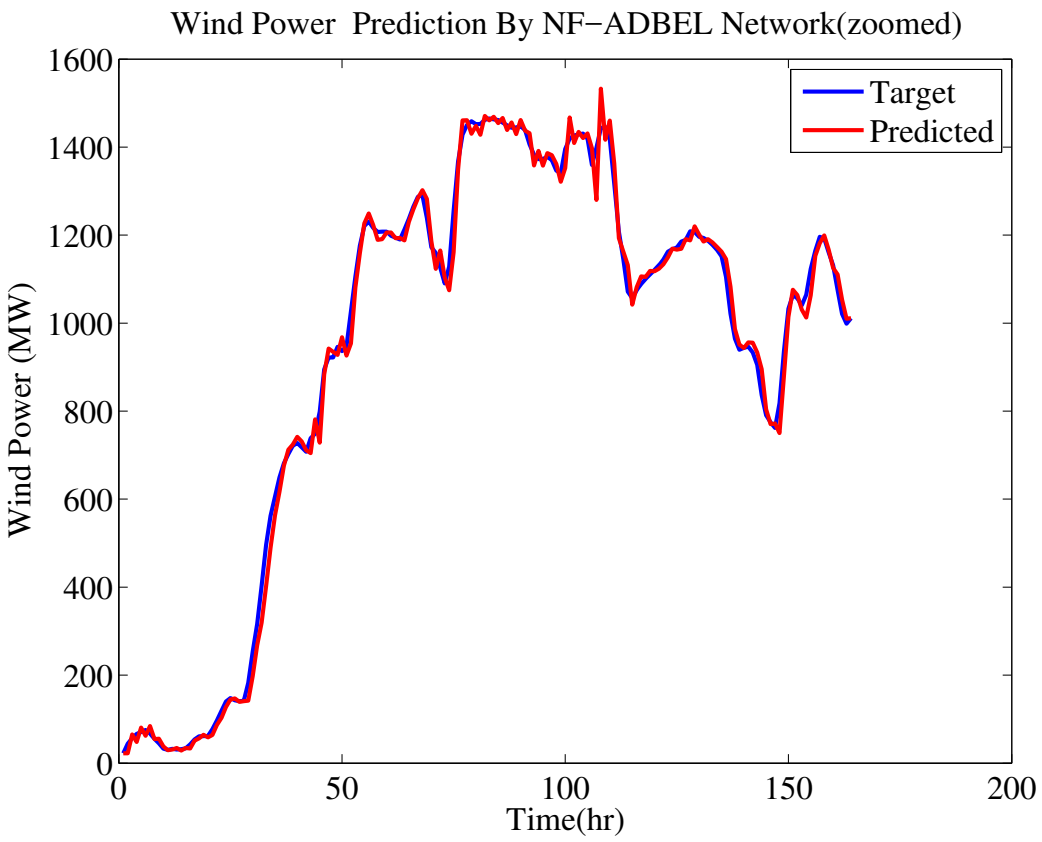

Figure 4.176: Maximum Wind Power as Predicted by NF-ADBEL Network.

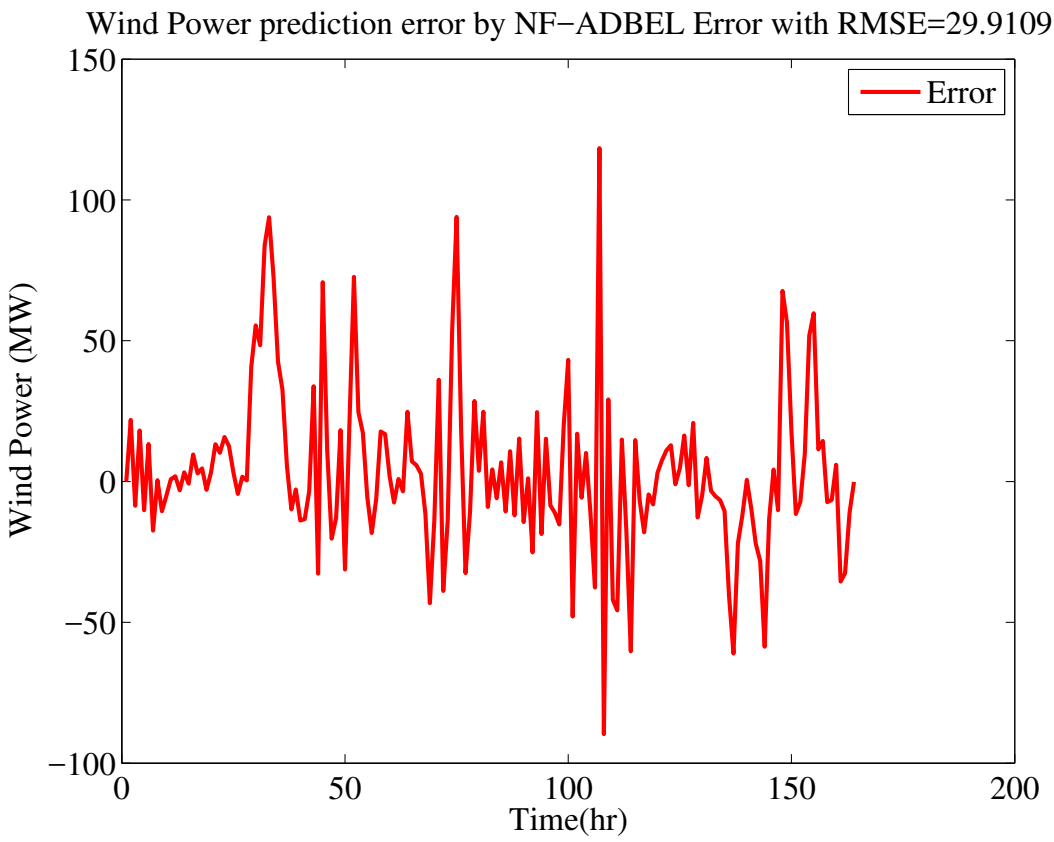

Figure 4.177: Error in Predicting Maximum Wind Power by NF-ADBEL Network.

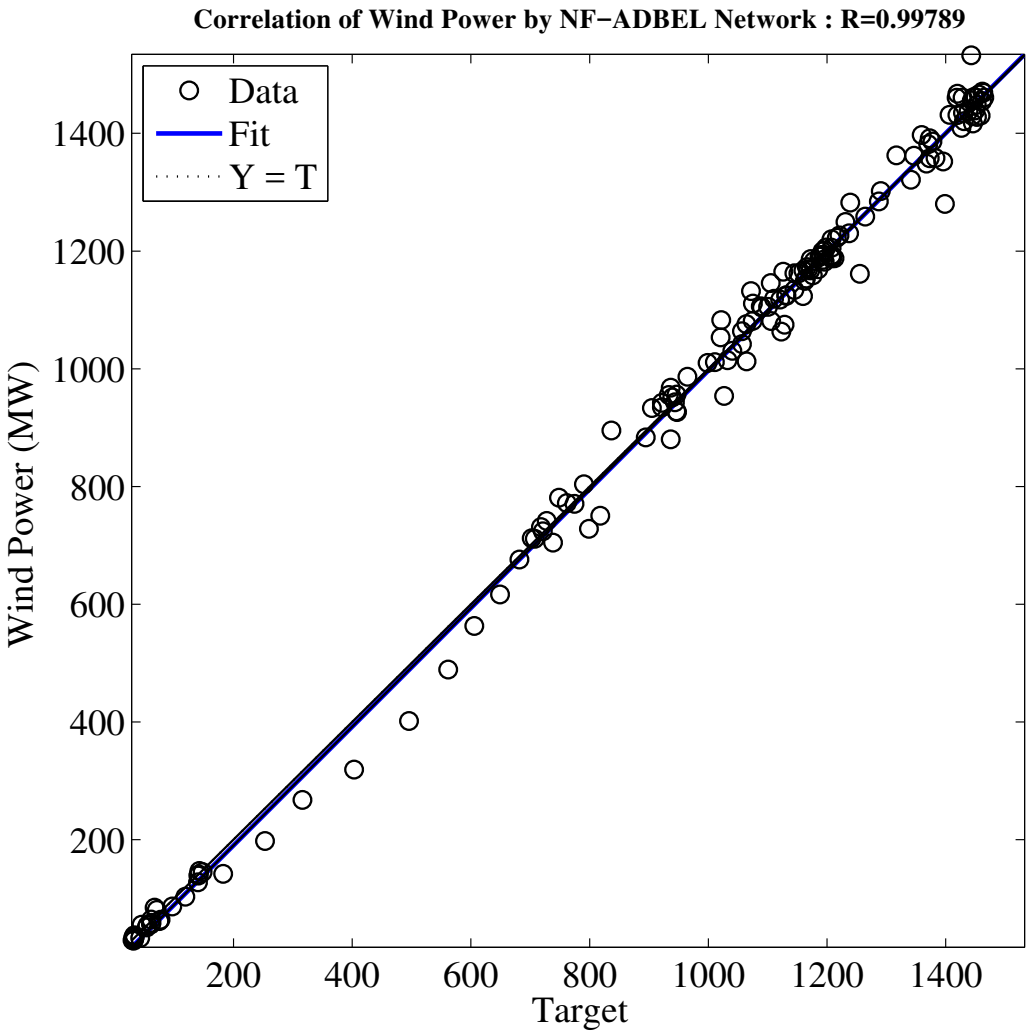

Figure 4.178: Correlation in Predicting Maximum Wind Power by NF-ADBEL Network.

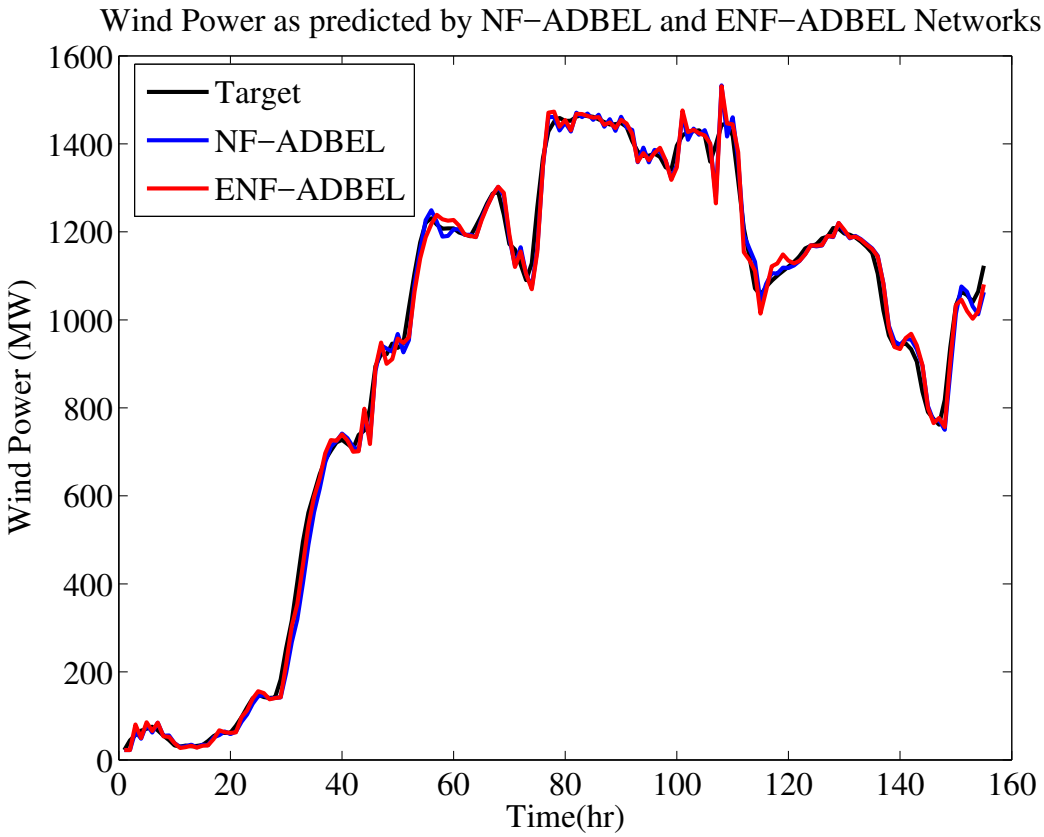

Figure 4.179: Maximum Wind Power as Predicted by ENF-ADBEL and NF-ADBEL Networks.

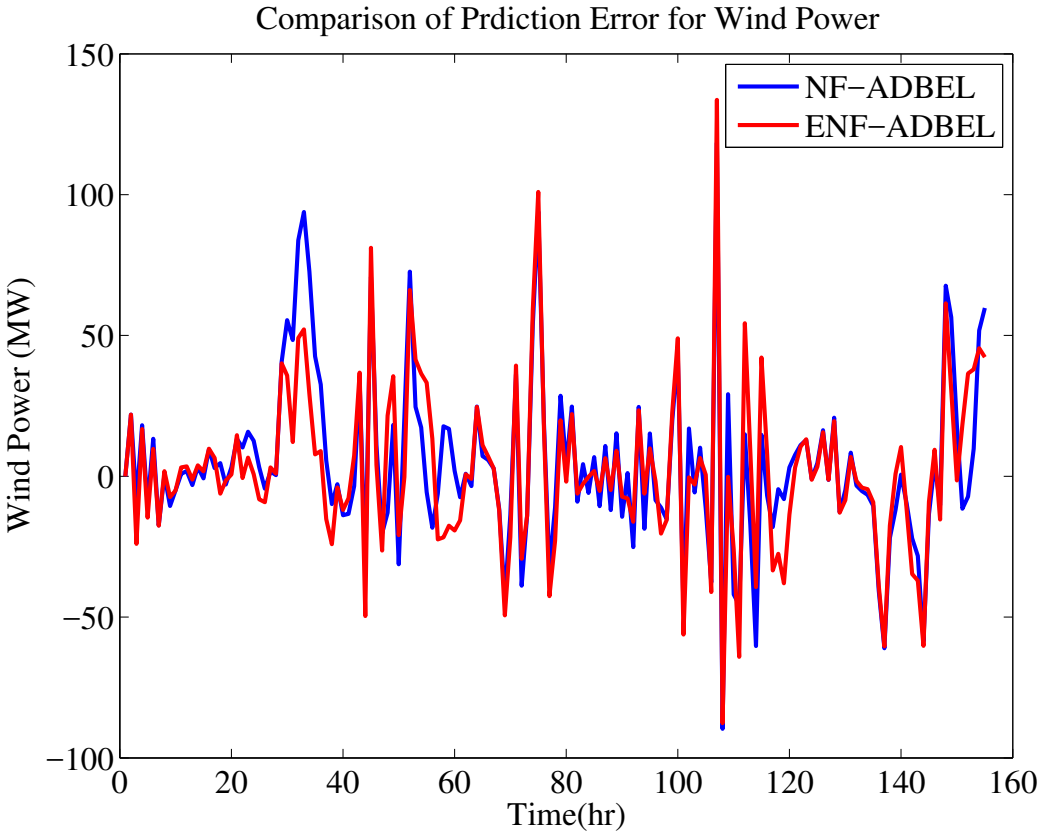

Figure 4.180: Error Comparison in Predicting Maximum Wind Power as Predicted by ENF-ADBEL and NF-ADBEL Networks.

4.3.6 Proposed Model's Performance Compared to State-of-the-Art Predictors

Thus far, we have achieved the objectives of the proposed models, as follows. First of all, the proposed NF-ADBEL network aimed to improve the accuracy of the ADBEL network. Our results show the superiority of the proposed NF-ADBEL network over the ADBEL in terms of high accuracy. Secondly, the proposed F-ADBEL model aimed to adjust the learning parameters α, β, and γ of the ADBEL network. The performance of the proposed F-ADBEL displays the ability to adjust the parameters and can be deployed for online prediction. Finally, the proposed ENF-ADBEL network aimed to enhance the accuracy of the NF-ADBEL network, and the ENF-ADBEL network performed with better accuracy than the NF-ADBEL network.

It should be mentioned that the proposed models are deployed to predict and forecast different online mode applications with no prior training and no prior knowledge of predicted data. Therefore, in this section, we aim to compare the proposed model's performance to other state-of-the-art predictors.

To compare our proposed model to other trained models, we assumed different factors of data as trained data by adjusting the steady-state depending on the data, as will be elaborated in this section. Note that the data we used might differ from the data used in other models in the literature. In each application, we employed the same steady-state to compare the performance of the proposed model.

4.3.6.1 Mackey-Glass Time Series as Predicted by Trained ENF-ADBEL Network

By setting the learning parameters as $\alpha = 0.5$, $\beta = 0.5$ and $\gamma = 0.07$, we first deployed the ENF-ADBEL network to predict this time series. The steady-state is 5 seconds, resulting in a pre-defined time window depicted in Figure 4.133. The same time series is also predicted with the NF-ADBEL and F-ADBEL networks using the learning parameters of $\alpha = 0.5$, $\beta = 0.2$, and $\gamma = 0.03$. In the case of F-ADBEL, the parameters are varying. The prediction errors are recorded in all the states, with analysis showing that the transient period remains the same ≤ 5 s. Thus, the steady-state starting index is set to compute the performance indices in all the cases.

Figures 4.181 and 4.182 show the prediction and error comparisons in steady-state

as yielded by the F-ADBEL, NF-ADBEL and ENF-ADBEL networks for predicting the Mackey-Glass time series. As seen, the ENF-ADBEL network performs better than the F-ADBEL and NF-ADBEL networks, showing lower peaks in the prediction error in the ENF-ADBEL network. As well, the root mean squared error and correlation coefficient were determined for all the networks, with the computed values presented in Table 4.24.

Table 4.24: RMSE, R^2 for Mackey-Glass Time Series Prediction by ENF-ADBEL, NF-ADBEL, and F-ADBEL Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Mackey-Glass	ENF-ADBEL	0.011	99.89	improved by
	NF-ADBEL	0.018	99.71	38.88
	F-ADBEL	0.0336	99.39	67.26

The authors in [109] proposed the short-term prediction of a backpropagation network (BP), based on a difference method (DMBP). In general, BP is widely used for prediction. The structure of the BP network is a multilayer feed-forward network, trained according to error backpropagation. DMBP structures the training layer as two sub-layers. The change degree layer reflects the absolute value, while the change trend layer reflects positive and negative data to overcome prediction errors.

The DMBP method was applied to a Mackey-Glass time series and the results were compared to other methods, such as BP network and support vector regression (SVR) machine. The autoregressive integrated moving average (ARIMA) in terms of RMSE was also given. The authors in [28] used 70% of data as training data and 30% of data as test data. The results of DMBP in [109] are shown in Table 4.25.

To compare the proposed ENF-ADBEL network to the methods in [109], we set the learning parameters as $\alpha = 0.54$, $\beta = 0.5$ and $\gamma = 0.07$. Further, the steady-state is selected as 50 seconds, which reflects 4% of the Mackey-Glass data points. The results are given in Table 4.25. As can be seen, the proposed ENF-ADBEL network had superior performance in terms of RMSE, as shown in Figure 4.183, compared to results in [109]. The run-time of 0.18 seconds reflects the simplicity and fast learning of the proposed ENF-ADBEL model, demonstrating that it can be deployed for online prediction.

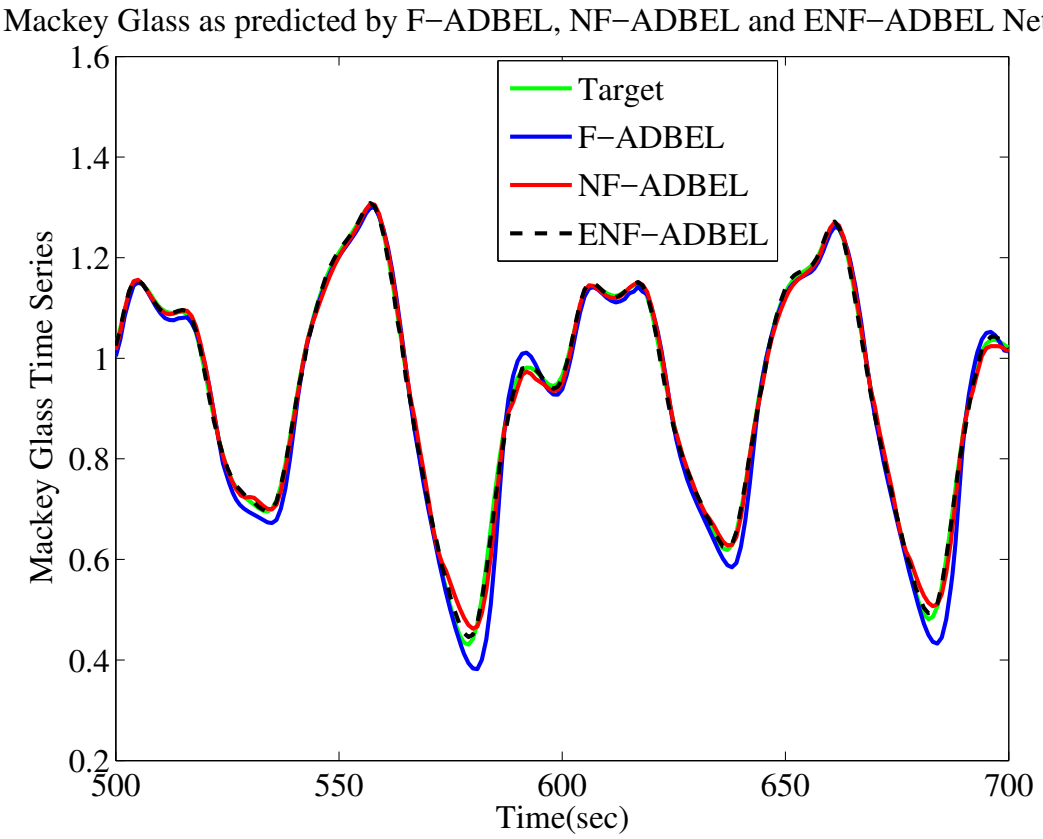

Figure 4.181: Mackey-Glass Time Series as Predicted by F-ADBEL, NF-ADBEL and ENF-ADBEL Networks.

Table 4.25: RMSE, R^2 for Mackey-Glass Time Series Prediction by ENF-ADBEL, BP, SVR, NARIMA, and DMBP Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Mackey-Glass	ENF-ADBEL	0.0094	99.87	improved by
	DMBP [109]	0.041	-	77.07
	NARIMA [109]	0.05	-	81.20
	SVR [109]	0.23	-	95.91
	BP [109]	0.291	-	96.76

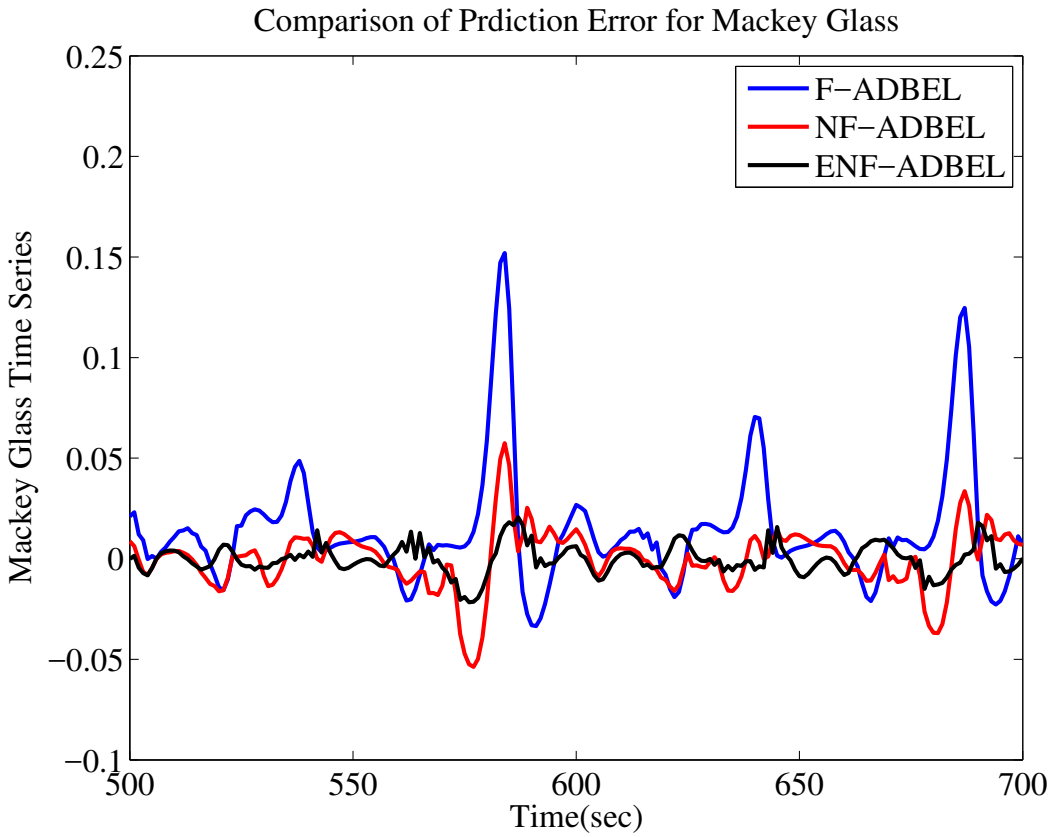

Figure 4.182: Error Comparison in Predicting Mackey-Glass Time Series as Predicted by F-ADBEL, NF-ADBEL and ENF-ADBEL Networks.

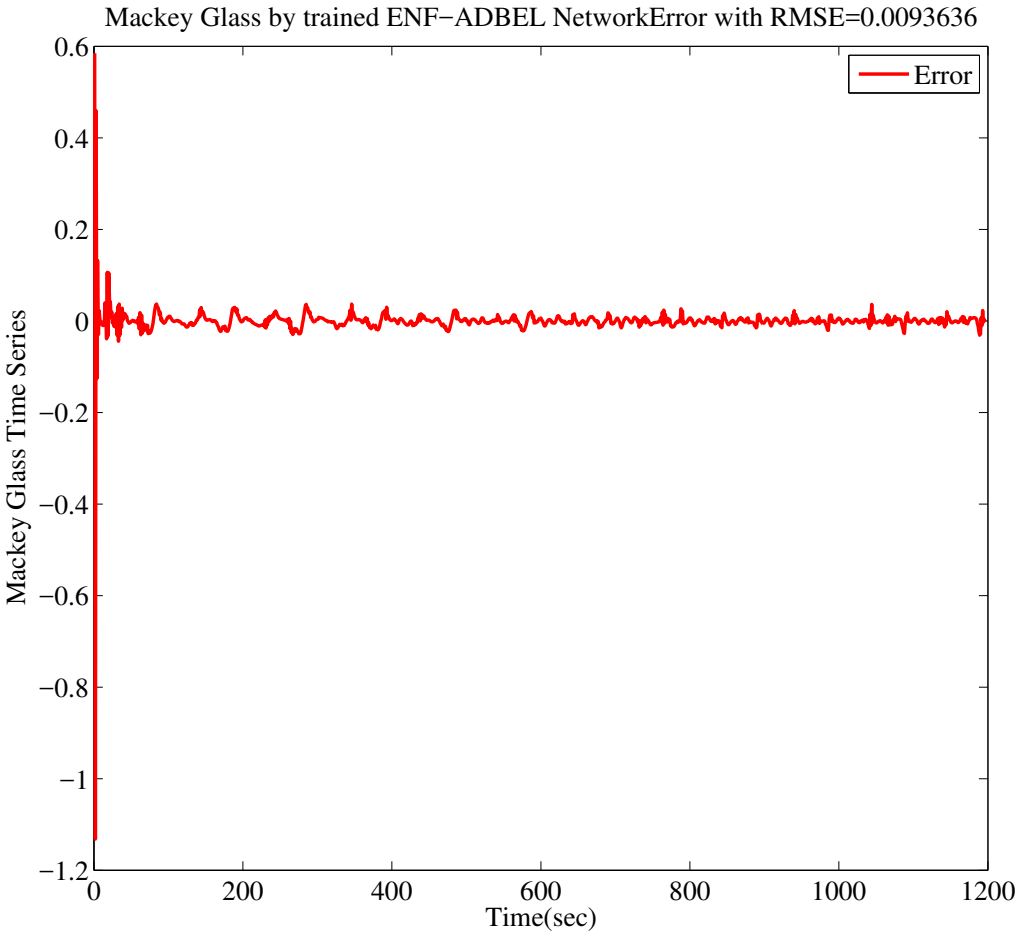

Figure 4.183: Error predicting Mackey-Glass Time Series for 4% Training Data as Predicted by ENF-ADBEL Network.

4.3.6.2 Lorenz Time Series as Predicted by Trained ENF-ADBEL Network

We first evaluate the prediction performance of the ENF-ADBEL network, with the learning parameters set as $\alpha = 0.5$, $\beta = 0.3$, and $\gamma = 0.04$. For comparison, the NF-ADBEL network is also simulated to predict the Lorenz time series and the F-ADBEL network. The best learning parameters for the NF-ADBEL network in predicting the Lorenz time series are found to be $\alpha = 0.8$, $\beta = 0.2$, and $\gamma = 0.01$), while for F-ADBEL the parameters are varying. The run-time is 1.42 seconds and 20.50 seconds, respectively.

By recording and analyzing the prediction error in all cases, we found that the transient period is less than $5s$, and therefore, the steady-state starting index is presumed as $n_s = 5$. A zoomed view of the prediction error as returned by all steady-state networks is shown in Figure 4.185. As can be see, the proposed ENF-ADBEL network has a lower error in predicting the Lorenz time series than the existing NF-ADBEL and F-ADBEL networks. The prediction performance in all cases is also analyzed in terms of root mean squared error and correlation coefficient criteria. The results for this analysis are included in Table 4.26 and show the superior performance of the ENF-ADBEL network due to the lowered root mean squared error, higher correlation coefficient, and significant percentage improvement offered by the network.

Table 4.26: RMSE, R^2 for Lorenz Time Series as Predicted by ENF-ADBEL, NF-ADBEL, and F-ADBEL Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Lorenz	ENF-ADBEL	0.13054	99.98	
	NF-ADBEL [58]	0.1726	99.97	24.36
	F-ADBEL	0.8564	99.78	84.75

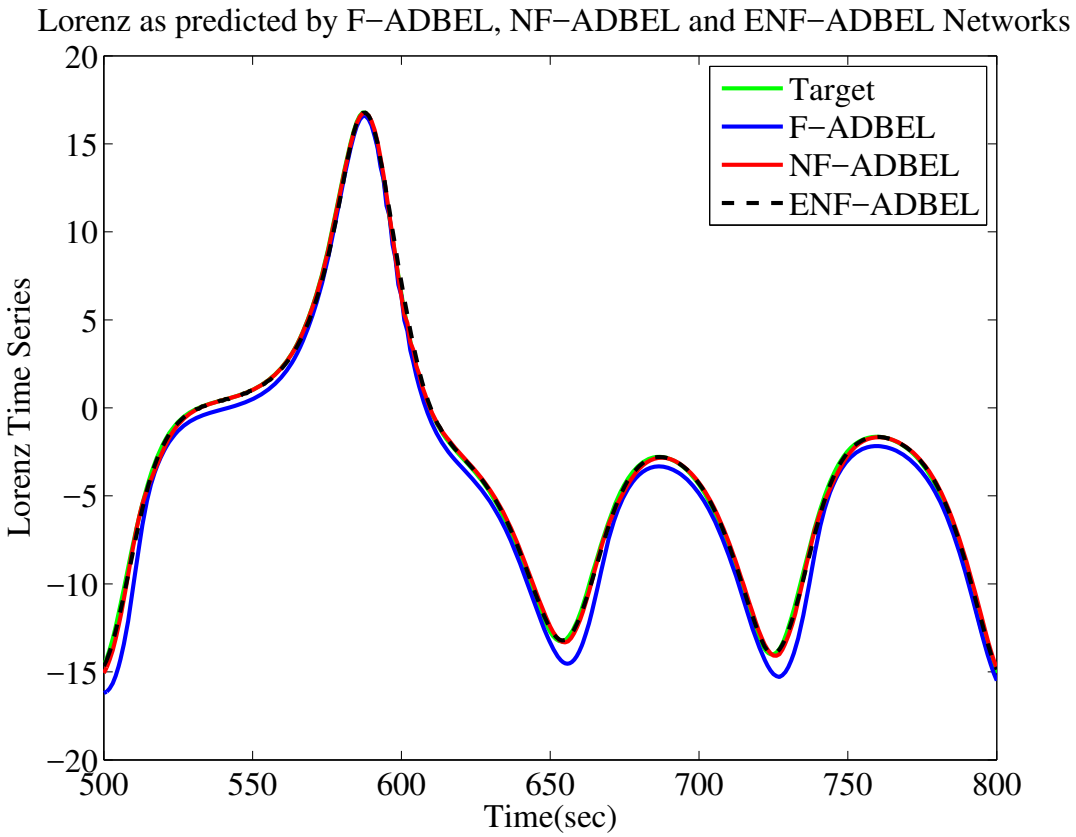

Figure 4.184: Lorenz Time Series as Predicted by F-ADBEL, NF-ADBEL and ENF-ADBEL Networks.

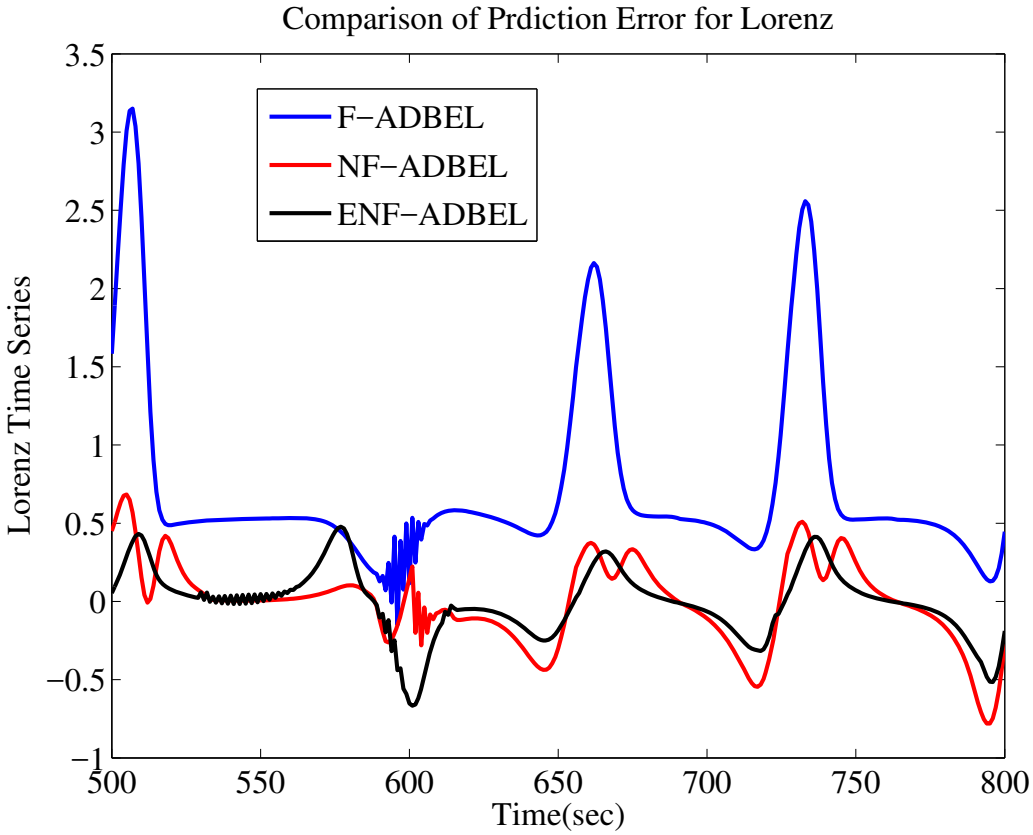

Figure 4.185: Error Comparison in Predicting Lorenz Time Series as Predicted by F-ADBEL, NF-ADBEL and ENF-ADBEL Networks.

Table 4.27: RMSE, R^2 for Lorenz Time Series Prediction by ENF-ADBEL, Naive LSTM, Multivariate Interpolated LSTM, and LSTM Approach Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Lorenz	ENF-ADBEL	0.0217	99.97	improved by
	LSTM approach[66]	0.0282	99.57	23.04
	Multivariate Interpolated LSTM[66]	0.0365	99.06	40.54
	Naive LSTM [66]	0.0463	98.89	53.13

The authors in [66] used long short-term memory (LSTM) recurrent neural network to predict the time series. However, the LSTM recurrent network has difficulty representing temporal and non-temporal inputs simultaneously in multivariate data. In [66], the authors proposed a hierarchal decomposition of univariate LSTMs and combined the resulting features in final feed-forward layers. The model selected is based on early stopping, with 20% validation data of 3,000 data samples and 1000 epochs. The condition was: if the validation performance did not improve within 100 epochs, stop training.

The authors in [66] then applied the proposed LSTM to the Lorenz time series and compared one-step-ahead prediction error in terms of lowest RMSE and high correlation coefficient to other approaches, such as Nave LSTM and multivariate interpolated LSTM. The LSTM approach gave better performance results compared to the other methods.

For a fair comparison to the ENF-ADBEL network, we used 75% data as a steady-state and tuned the parameters as $\alpha = 0.49$, $\beta = 0.42$, and $\gamma = .04$. The proposed ENF-ADBEL shows better results in terms of low prediction error and high correlation, as presented in Table 4.27 and Figure 4.186. According to [66], the used LSTM converged within several hours of training, while the proposed ENF-ADBEL performed convergence within a few seconds (run-time of 1.44 seconds). The proposed ENF-ADBEL thus gave the best performance and fastest response.

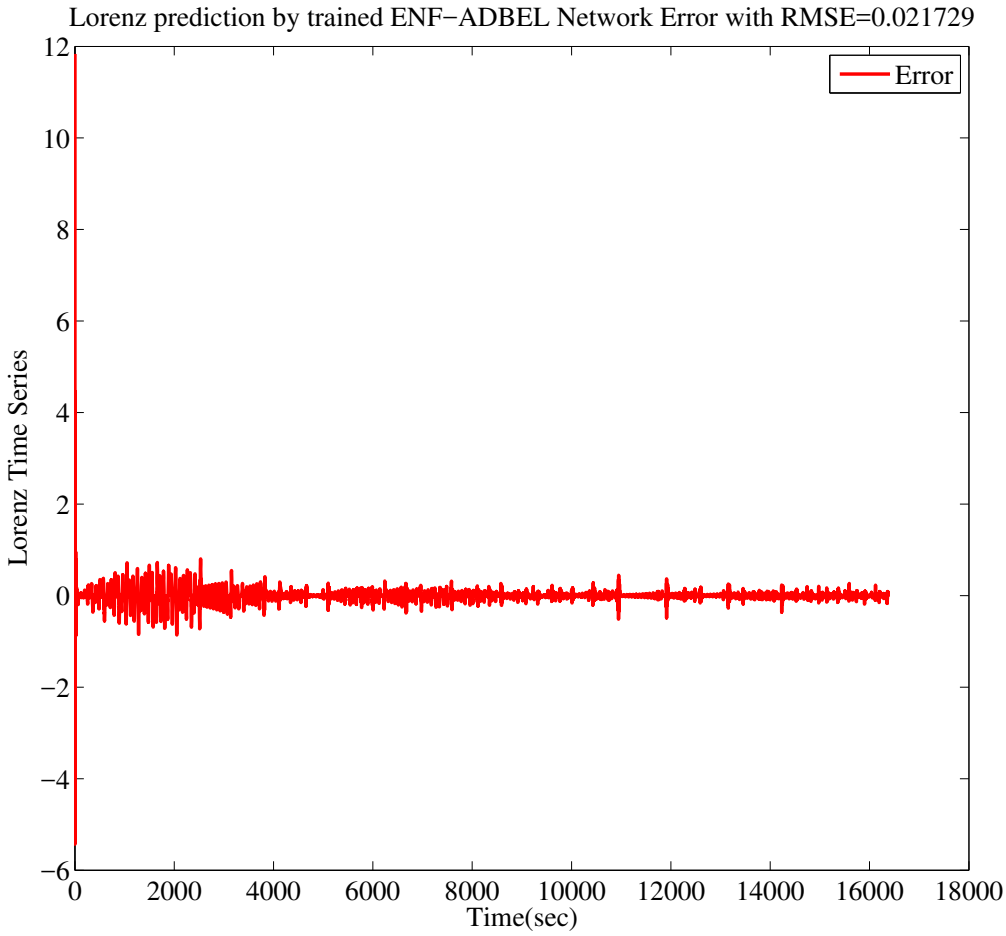

Figure 4.186: Error predicting Lorenz Time Series for 75% Training Data as Predicted by ENF-ADBEL Network.

4.3.6.3 Disturbance Storm Time Index (D_{st}) as Predicted by Trained ENF-ADBEL Network

Next, we simulate the ENF-ADBEL network to predict the disturbance storm time index D_{st} time series for April 2000, at a time when considerable geomagnetic activity was observed. The data for this month have been downloaded from the website World Data Center (WDC) [86], "WDC for Geomagnetism, Kyoto." With the learning parameters set as $\alpha = 0.15$, $\beta = 0.38$ and $\gamma = 0.25$, the ENF-ADBEL network is deployed to predict the D_{st} index for the month of April 2000. The number of samples is $n_e = 716$.

The transient period of the ENF-ADBEL network is $n_s = 10\ hrs$, which then becomes the steady-state starting index. It can be observed that, despite the high initial transients, the ENF-ADBEL network can follow the D_{st} time series in steady-state. The important valley points are also well-predicted, which points towards the possible occurrence of geomagnetic storms.

An existing NF-ADBEL network is used to predict the D_{st} time series. For this purpose, the learning parameters of the NF-ADBEL network are assigned the values of $\alpha = 0.3$, $\beta = 0.3$, and $\gamma = 0.01$. The results of this comparison in terms of the prediction error are displayed in Figures 4.187 and 4.188. As can be seen, the ENF-ADBEL network gives a better performance than the NF-ADBEL and F-ADBEL networks, as shown in Table 4.28.

The authors in [65] designed an NFCBEL predictor, which combines the type of emotional neural network and neo-fuzzy neurons. The NFCBEL is deployed to predict Dst between the years 2000 and 2006, inclusive. The NFCBEL was trained offline with 70% of the data in order to predict 30% of the data. The performance of the RMSE and correlation is presented in Table 4.29.

To compare the proposed model with the model in [65], we deploy the ENF-ADBEL to predict the Dst for April 2000 for 27% of 716 data points, using $ns = 200$. Further, the ENF-ADBEL network is assigned the values of $\alpha = 0.1$, $\beta = 0.3$, and $\gamma = 0.93$. Our results show that the ENF-ADBEL performed excellently in terms of RMSE and high correlation, as shown in Figure 4.189 and Table 4.29. The run-time was 0.988 seconds as compared to NFCBEL, which performed its outcomes after 50 iterations. Note that because the length of the data is different, the results in Table

4.29 reflect the proposed model's performance based on the available data (in this case, April 2000). Overall, the results reveal that the ENF-ADBEL network is a good candidate for online forecasting.

Table 4.28: RMSE, $R^2$for D_{st} by ENF-ADBEL, NF-ADBEL and F-ADBEL Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Dst Apr 2000	ENF-ADBEL	6.86	98.31	
	NF-ADBEL [58]	9.08	97.06	24.44
	F-ADBEL	13.137	94.36	47.78

Table 4.29: RMSE, R^2 for D_{st} Prediction by Trained ENF-ADBEL Network

Time Series	Prediction Network	RMSE	R^2(%)	Epochs
Dst Apr 2000	ENF-ADBEL	4.05	98.07	0
Dst [65]	NFCBEL	4.7649	98.047	50

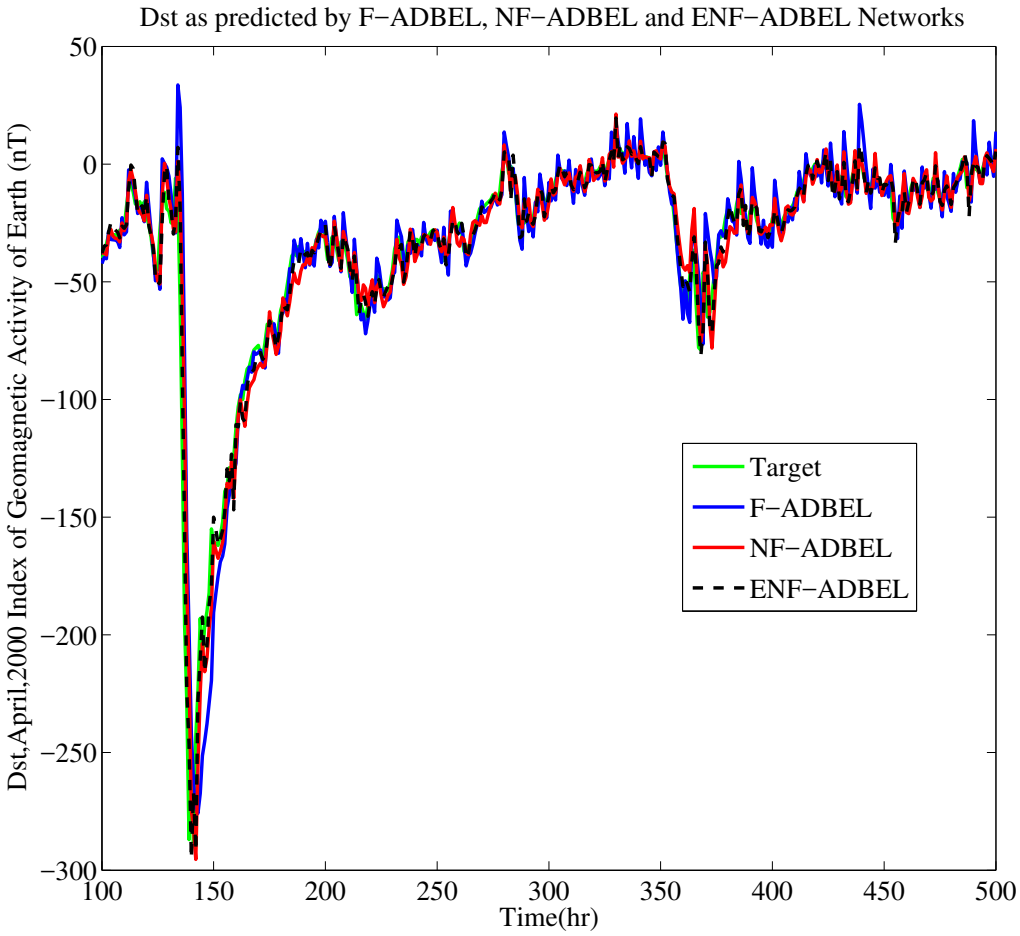

Figure 4.187: Dst April 2000 as Predicted by F-ADBEL, NF-ADBEL and ENF-ADBEL Networks.

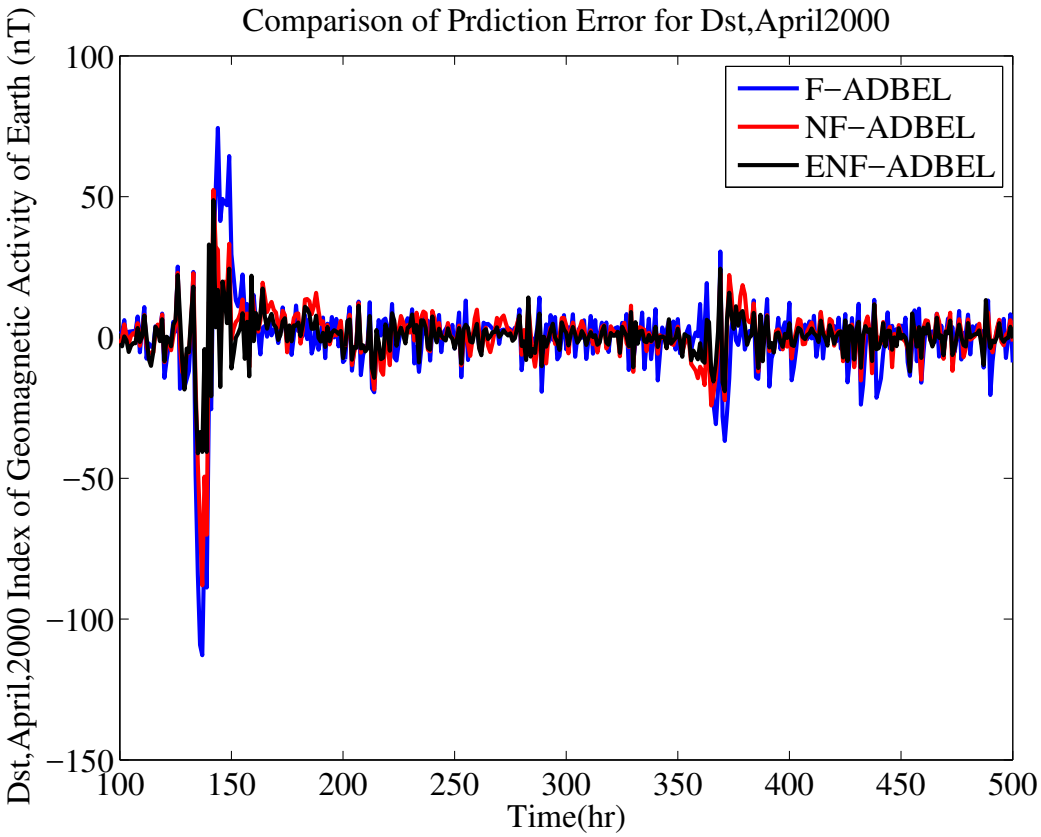

Figure 4.188: Error Comparison of Dst April 2000 as Predicted by F-ADBEL, NF-ADBEL and ENF-ADBEL Networks.

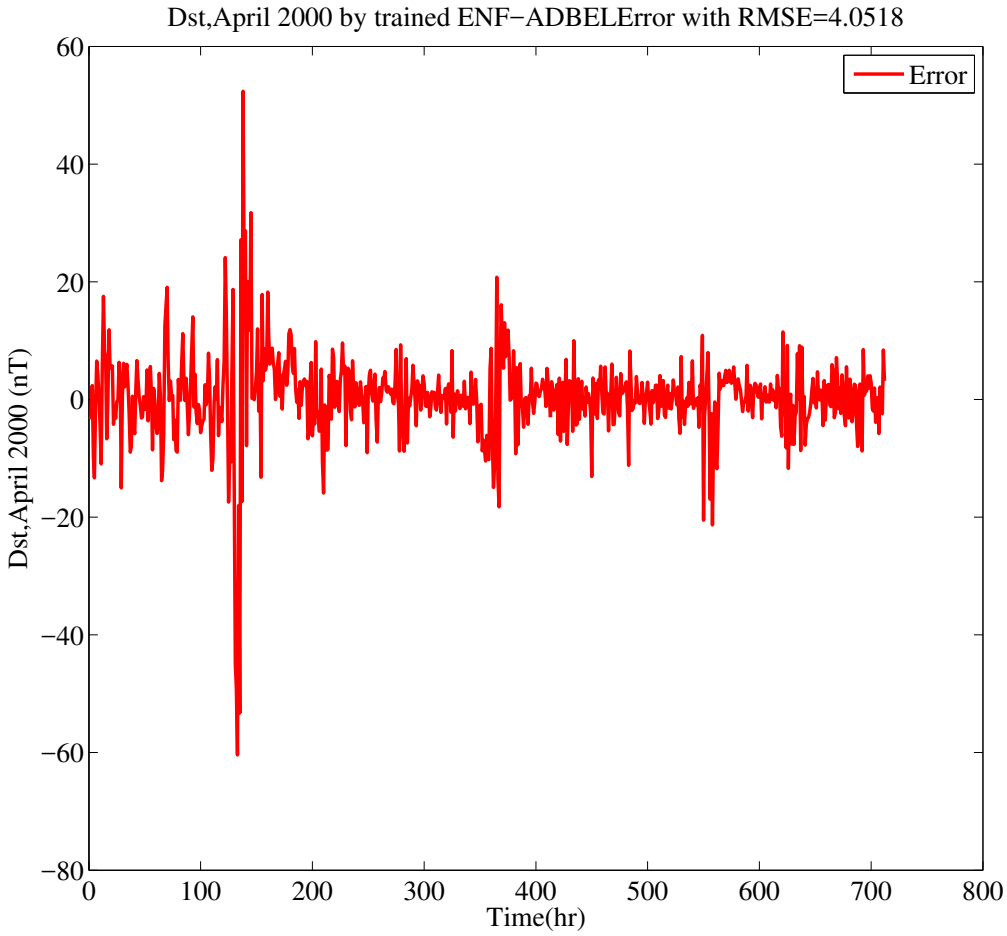

Figure 4.189: Error predicting Dst April 2000 for 27% Training Data as Predicted by ENF-ADBEL Network.

4.3.6.4 Wind Speed as Performance by ENF-ADBEL Network

The ENF-ADBEL network is first employed using the learning parameters $\alpha = 0.3$, $\beta = 0.015$, and $\gamma = 0.25$. As can be seen, the ENF-ADBEL network can predict wind speed 1h ahead. The steady-state starting index is taken as $n_s = 1hr$. To compare the model's performance with the ENF-ADBEL network, a simulation is run with the learning parameters for the NF-ADBEL network set as $\alpha = 0.77$, $\beta = 0.04$, and $\gamma = 0.19$. The F-ADBEL network has varying parameter values. Forecasting error and correlation indices are presented in Table 4.30.

Table 4.30: RMSE & R^2 for Wind Speed ENF-ADBEL, NF-ADBEL and F-ADBEL Networks

Time Series	Prediction Network	RMSE	R^2(%)	PI(%)
Wind Speed	ENF-ADBEL	5.28	89.08	
	NF-ADBEL	5.53	88.10	4.52
	F-ADBEL	6.81	86.01	28.97

The transient period for the NF-ADBEL network is the same as that for the ENF-ADBEL network. However, the ENF-ADBEL network shows better performance than the NF-ADBEL and F-ADBEL networks, owing to the lower forecasting error rate being offered by this network during steady-state, as can be seen from Figures 4.190 and 4.191. In terms of run-time, the proposed ENF-ADBEL accomplished the performance in 1,09 seconds, while the NF-ADBEL took 1.03 seconds and the F-ADBEL 3.7 seconds. The lower root mean squared error, higher correlation coefficient and sufficient percentage improvement yielded by the ENF-ADBEL network validates its performance over the NF-ADBEL and F-ADBEL networks in the prediction of wind speed, as shown in Table 4.30.

The authors in [110] built eight models to predict wind speed, namely BPNN, GA-BPNN, PSO-BPNN, LSTM, SVR, GA-SVR, Bagging and Boosting. The models were all supported by GA and PSO to help to find a global optimal. Additionally, the authors used a dataset from Open EI; the size of the data was 36,295 samples, and the timeline was from May 13, 2003, to Jan 20, 2004. The data were randomly split into a training set and a test set, with 70% used for training and 30% for testing.

We used the correlation coefficient criteria to compare the results from different models with those from the proposed ENF-ADBEL network. The results are presented in Table 4.31. as can be seen, the proposed model has high correlation and a fast processing time, whereas the other methods were trained and accomplished the performance within 1,000 iterations. Note that the data used in the proposed model were different in size and location than those used in the other models.

Table 4.31: R^2 for Wind Speed Prediction by ENF-ADBEL, BPNN, GA-BPNN, PSO-BPNN, LSTM, SRV, GA-SVR, Bagging and Adaboost Models

Time Series	Prediction Network	R^2(%)
Wind Speed	ENF-ADBEL	89.08
	BPNN [110]	88.76
	GA-BPNN [110]	88.47
	PSO-BPNN [110]	87.21
	LSTM [110]	87.99
	SRV [110]	88.35
	GA-SRV [110]	88.55
	Bgging [110]	88.63
	Adaboost [110]	88.55

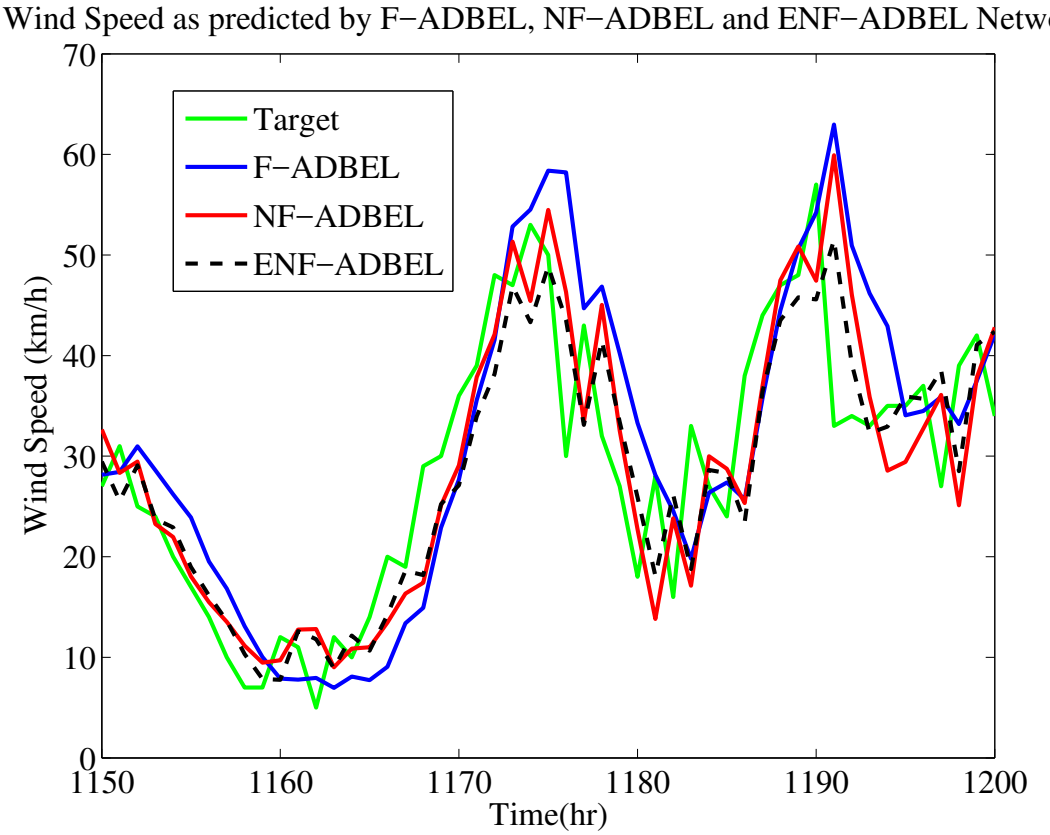

Figure 4.190: Wind Speed as Predicted by F-ADBEL, NF-ADBEL and ENF-ADBEL Networks.

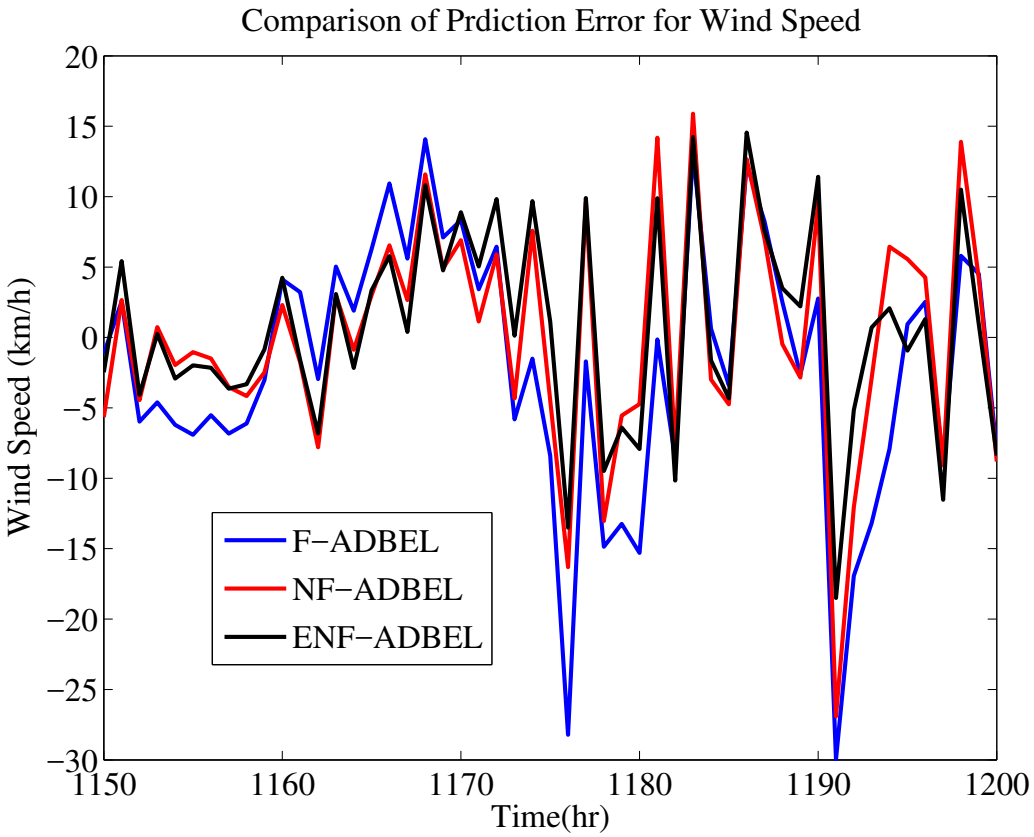

Figure 4.191: Error Comparison in Predicting Wind Speed as Predicted by F-ADBEL, NF-ADBEL and ENF-ADBEL Networks.

4.3.7 CONCLUSIONS

1. **Highlight 1:** In the proposed ENF-ADBEL model, neo-fuzzy neurons are applied in the orbitofrontal cortex section and partially in amygdala section of Adaptive decayed brain emotional learning network. Partial integration is done purposefully as amygdala section has two outputs, the one which relies on the imprecise information is set free from neo-fuzzy integration to keep the computational principle of limbic system.

2. **Highlight 2:** The integration of neo-fuzzy network in the amygdala section does not increase the computational complexity of the resulting proposed model to a noticeable extent.

3. **Highlight 3:** Learning parameters, α, β, and γ play a crucial rule in the performance of proposed ENF- ADBEL performance. Currently, an exhaustive search is done to find the near optimal parameters.

4. **Highlight 4:** Proposed ENF-ADBEL model has no prior knowledge of the time-series data which implies that no prior training is required.

5. **Highlight 5:** Comparison of the proposed predictor with F-ADBEL [62], NF-ADBEL [58] and others, reveals its superiority for time series prediction problems with shorter update intervals.

6. H**ighlight 6:** It is known that the size of data can affect the performance of predictors. The proposed model can also be deployed where size of the data is considerably large.

In this work, we presented a novel design for a hybrid model of a neo-fuzzy adaptive decayed brain emotional learning network, intending to enhance the prediction accuracy of NF-ADBEL for on-line time series prediction. We called the resulting prediction network the Expanded Neo-Fuzzy Adaptive Decayed Brain Emotional Learning (ENF-ADBEL) model. The proposed model integrates the neo-fuzzy neurons in the orbitofrontal cortex (OFC) section while partially implementing the amygdala (AMY) section. The proposed model combines competitive emotional neural networks with neo-fuzzy neurons to yield an effective ENF-ADBEL predictor, which offers features

such as low computational complexity and fast learning. The low complexity results from fewer membership functions in neo-fuzzy neuron networks, while fast learning is inherited from employing the mammalian brain's emotion-processing mechanism.

Furthermore, the proposed ENF-ADBEL network, which is implemented in the MATLAB programming environment, has been deployed to predict several chaotic time series, including Mackey-Glass, Lorenz, Rossler and disturbance storm time index. Simulations have also been conducted in this work to predict stochastic problems, namely, wind speed and wind power series. To keep computational complexity at a minimum, we only used three neo-fuzzy neurons membership functions to process each feature in all OFC and AMY sections of ENF-ADBEL.

The proposed model was deployed for on-line time series prediction with no prior training. The performance of the model was also evaluated in terms of RMSE and R^2. As well, the NF-ADBEL network and F-ADBEL were simulated to forecast the same time series with near optimal parameters.

Additionally, a comparison of the proposed ENF-ADBEL with other state-of-the-art methods was made. A percentage improvement index was defined to compare the proposed model's performance with those of the NF-ADBEL and F-ADBEL networks. Comparing the proposed model with the state-of-the-art predictors NF-ADBEL and F-ADBEL reveal its superior performance, as the model offers the lowest RMSE and a higher R^2. A substantial amount of percentage improvement is also observed in wind speed and wind power forecasting. Finally, based on the simulation results, the proposed model demonstrated the best performance.

Chapter 5

CONCLUSIONS AND FUTURE WORK

This chapter highlights the contributions of the thesis, and also presents suggestions for possible future work directions.

5.1 Contributions

1. The first contribution of this thesis was the design of a Neo-fuzzy Integrated Adaptive Decayed Brain Emotional Learning (NF-ADBEL) network. The NF-ADBEL network demonstrated the ability and capability for online time series prediction problems and other forecasting applications, such as wind speed and wind power generation. It also enhanced the ADBEL network accuracy.

2. The work's second contribution was the design of the Expanded Neo-fuzzy Integrated Adaptive Decayed Brain Emotional Learning (ENF-ADBEL) network. The ENF-ADBEL network demonstrated the ability and capability for online time series prediction problems and other forecasting applications, such as wind speed and wind power generation. It also enhanced the NF-ADBEL network accuracy.

3. The third contribution of this thesis was the design of the Fuzzy-logic-based Parameter-adjustment Model to use with the Adaptive Decayed Brain Emotional Learning (F-ADBEL) network. The F-ADBEL demonstrated the ability and capability to tune the ADBEL parameters in online mode. NF-ADBEL enhanced the ADBEL network's accuracy and can be used for online time series prediction problems and other forecasting applications such as wind speed and wind power generation.

4. The fourth contribution of this study was the series of comparisons conducted between the designed proposed NF-ADBEL, ENF-ADBEL, F-ADBEL, and ADBEL models and other state-of-the-art models.

In addition to the above, other contributions, include the following:

- Redesigning/representing the ADBEL network and using it as a benchmark to compare the proposed models' outcomes.
- Programming and simulating all the proposed models in MATLAB.
- Generating a comparison program in MATLAB to compare the results between the proposed models.
- Generating the time series data for Mackey-Glass, Lorenz, Rossler and Narendra identification plant using MATLAB programming, and filing it in an attached appendix in this thesis to support future research work.

5.2 Future Work Directions

The design of the proposed fuzzy logic-based parameter adjustment (F-ADBEL) model, is an alternative way for finding the parameters and may not be optimal. A possible future direction could be studying the optimality and stability of the parameter adjustment model.

Similarly, in the design of the proposed neo-fuzzy adaptive decayed brain emotional learning (NF-ADBEL) network, the tuned parameters may not be optimal. A Possible future direction could be studying the optimal tuned parameters further.

Finally, with regards to the design of the proposed Expanded neo-fuzzy adaptive decayed brain emotional learning (ENF-ADBEL) network, the tuned parameters may not be optimal. So, one possible future direction could be studying ways to find and implement optimal tuned parameters.

As well, more investigation could be conducted on the integration of the neo-fuzzy network in Thalamus and sensory cortex sections.

Bibliography

[1] J. Doyne Farmer and John J. Sidorowich. Predicting chaotic time series. *Phys. Rev. Lett.*, 59:845–848, Aug 1987.

[2] M. Parsapoor, U. Bilstrup, and B. Svensson. A brain emotional learning-based prediction model for the prediction of geomagnetic storms. In *2014 Federated Conference on Computer Science and Information Systems*, pages 35–42, Sept 2014.

[3] E. Lotfi and M. R. Akbarzadeh-T. Emotional brain-inspired adaptive fuzzy decayed learning for online prediction problems. In *2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–7, July 2013.

[4] Ehsan Lotfi and M.-R. Akbarzadeh-T. Adaptive brain emotional decayed learning for online prediction of geomagnetic activity indices. *Neurocomputing*, 126:188–196, 2014.

[5] M. Parsapoor and U. Bilstrup. Brain emotional learning based fuzzy inference system (belfis) for solar activity forecasting. In *2012 IEEE 24th International Conference on Tools with Artificial Intelligence*, volume 1, pages 532–539, Nov 2012.

[6] B. D. Damas and L. Custdio. Emotion-based decision and learning using associative memory and statistical estimation. *J. Informatica (Slovenia)*, 27(2):145–156, 2004.

[7] Juan D. Velásquez. When robots weep: Emotional memories and decision-making. In *Proceedings of the Fifteenth National/Tenth Conference on Artificial Intelligence/Innovative Applications of Artificial Intelligence*, 1998.

[8] Caro Lucas, Danial Shahmirzadi, and Nima Sheikholeslami. Introducing belbic: Brain emotional learning based intelligent controller. *Intelligent Automation & Soft Computing*, 10(1):11–21, 2004.

[9] N. Sheikholeslami, D. Shahmirzadi, E. Semsar, C. Lucas, and M. J. Yazdanpanah. Applying brain emotional learning algorithm for multivariable control of hvac systems. *J. Intell. Fuzzy Syst.*, 17(1):35–46, January 2006.

[10] Ali Reza Mehrabian, Caro Lucas, and Jafar Roshanian. Aerospace launch vehicle control: an intelligent adaptive approach. *Aerospace Science and Technology*, 10(2):149–155, 2006.

[11] Caro Lucas, Rasoul M. Milasi, and Babak N. Araabi. Intelligent modeling and control of washing machine using locally linear neuro-fuzzy (llnf) modeling and modified brain emotional learning based intelligent controller (belbic). *Asian Journal of Control*, 8(4):393–400, 2006.

[12] A. M. Yazdani S. Buyamin S. Mahmoudzadeh Z. Ibrahim and M. F. Rahmat. Ibrain emotional learning based intelligent controller for stepper motor trajectory tracking. *International Journal of Physical Sciences*, 7(15):2364–2386, 2012.

[13] Hossein Rouhani Mahdi Jalili Babak N. Araabi Wolfgang Eppler and Caro Lucas. Brain emotional learning based intelligent controller applied to neurofuzzy model of micro-heat exchanger. *Expert Systems with Applications*, 32(3"):911–918, 2007.

[14] D.G.Stork R.O.Duda, P.E.Hart. *Pattern Classification*. Hoboken, NJ, USA: Wiley,, 2001.

[15] C.M.Bishop. *Pattern Recognition and Machine Learning*. NewYork,NY, USA: Springer,, 2006.

[16] Wei Zhang ; Kan Liu ; Weidong Zhang ;Youmei Zhang and Jason Gu. Deep neural networks for wireless localization in indoor and outdoor environments. *Neurocomputing*, 194(Supplement C):279–287, 2016.

[17] H. Liu ; Y. Yu ;F. Sun and J. Gu. Visual-tactile fusion for object recognition. *IEEE Transactions on Automation Science and Engineering*, 14(2):996–1008, April 2017.

[18] N. M. Omar and M. E. El-Hawary. Optimizing classifier performance for parkinson's disease detection. In *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1–6, April 2017.

[19] D.T. Pham and S. Sagiroglu. Training multilayered perceptrons for pattern recognition: a comparative study of four training algorithms. *International Journal of Machine Tools and Manufacture*, 41(3):419–430, 2001.

[20] D T PHAM and A B Chan. Unsupervised adaptive resonance theory neural networks for control chart pattern recognition. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 215(1):59–67, 2001.

[21] D T PHAM, M S Packianather, and E Y A Charles. Control chart pattern clustering using a new self-organizing spiking neural network. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 222(10):1201–1211, 2008.

[22] Ehsan Lotfi and M. R. Akbarzadeh-T. Brain emotional learning-based pattern recognizer. *Cybern. Syst.*, 44(5):402–421, July 2013.

[23] Ehsan Lotfi and Azita Keshavarz. Gene expression microarray classification using pca-bel. *Computers in Biology and Medicine*, 54(Supplement C):180–187, 2014.

[24] E. Lotfi and M. R. Akbarzadeh-T. Supervised brain emotional learning. In *The 2012 International Joint Conference on Neural Networks (IJCNN)*, pages 1–6, June 2012.

[25] Ehsan Lotfi, Saeed Setayeshi, and Saeed Taimory. A neural basis computational model of emotional brain for online visual object recognition. *Applied Artificial Intelligence*, 28(8):814–834, 2014.

[26] Ehsan Lotfi and M.-R. Akbarzadeh-T. Practical emotional neural networks. *Neural Networks*, 59(Supplement C):61–72, 2014.

[27] M. U. Asad, U. Farooq, J. Gu, J. Amin, A. Sadaqat, M. E. El-Hawary, and J. Luo. Neo-fuzzy supported brain emotional learning based pattern recognizer for classification problems. *IEEE Access*, 5:6951–6968, 2017.

[28] M.S. Gazzaniga, R.B. Ivry, and G.R. Mangun. *cognitive Neuroscience: The Biology of the Mind*. W.W. Norton, 2009.

[29] Lazaros C. Triarhou. Centenary of christfried jakob's discovery of the visceral brain: An unheeded precedence in affective neuroscience. *Neuroscience & Biobehavioral Reviews*, 32(5):984 – 1000, 2008.

[30] C. Sagan. *Broca's brain: reflections on the romance of science*. Random House, 1979.

[31] John D. Newman and James C. Harris. The scientific contributions of paul d. maclean (1913-2007). 197(1):3–5, 1 2009.

[32] Joseph LeDoux. *The emotional brain: The mysterious underpinnings of emotional life*. Simon and Schuster, 1998.

[33] Jan MorAn Christian Balkenius. Emotional learning: A computational model of the amygdala. *Cybernetics and Systems*, 32(6):611–636, 2001.

[34] E. Daryabeigi, G. A. Markadeh, and C. Lucas. Interior permanent magnet synchronous motor (ipmsm), with a developed brain emotional learning based intelligent controller (belbic). In *2009 IEEE International Electric Machines and Drives Conference*, pages 1633–1640, May 2009.

[35] S. Valizadeh, M. R. Jamali, and C. Lucas. A particle-swarm-based approach for optimum design of belbic controller in avr system. In *2008 International Conference on Control, Automation and Systems*, pages 2679–2684, Oct 2008.

[36] Sima Seidi Khorramabadi, Mehrdad Boroushaki, and Caro Lucas. Emotional learning based intelligent controller for a {PWR} nuclear reactor core during load following operation. *Annals of Nuclear Energy*, 35(11):2051–2058, 2008.

[37] A Gholipour, CARO Lucas, and DANIAL Shahmirzadi. Predicting geomagnetic activity index by brain emotional learning. *WSEAS Transactions on Systems*, 3(1):296–299, 2004.

[38] T. Babaie, R. Karimizandi, and C. Lucas. Learning based brain emotional intelligence as a new aspect for development of an alarm system. *Soft Comput.*, 12(9):857–873, April 2008.

[39] Oliver Hardt, Karim Nader, and Lynn Nadel. Decay happens: the role of active forgetting in memory. *Trends in cognitive Sciences*, 17(3):111–120, 2013.

[40] A. Khashman. A modified backpropagation learning algorithm with added emotional coefficients. *Neural Networks, IEEE Transactions on*, 19(11):1896–1909, Nov 2008.

[41] Javad Abdi, Behzad Moshiri, Baher Abdulhai, and Ali Khaki Sedigh. Forecasting of short-term traffic-flow based on improved neurofuzzy models via emotional temporal difference learning algorithm. *Engineering Applications of Artificial Intelligence*, 25(5):1022 – 1042, 2012.

[42] M Parsapoor and Urban Bilstrup. Chaotic time series prediction using brain emotional learning-based recurrent fuzzy system (belrfs). *Int. J. Reasoning-based Intelligent Systems*, 5(2):113–126, 2013.

[43] M. Pasrapoor and U. Bilstrup. An emotional learning-inspired ensemble classifier (eliec). In *2013 Federated Conference on Computer Science and Information Systems*, pages 137–141, Sept 2013.

[44] B.Roozendaal B. Ferry and McGaugh J. Role of norepinephrine in mediating stress hormone regulation of long-term memory storage: a critical involvement of the amygdala. *Biological Psychiatry Journal*, 46(9):1140–1152, 1999.

[45] E.R. Kandel, J. Schwartz, and T. Jessell. *Principles of Neural Science, Fourth Edition*. McGraw-Hill Companies,Incorporated, 2000.

[46] K. Amunts K.and Kedo O. Kindler M.Pieperhoff P.Mohlberg H. Shah N.J. Habel U. Schneider F.and Zilles. Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. *Anatomy and Embryology*, 210(5):343–352, Dec 2005.

[47] JORGE L. ARMONY and JOSEPH E. LEDOUX. How the brain processes emotional information. *Annals of the New York Academy of Sciences*, 821(1):259–270, 1997.

[48] I. Gościński, S. Kwiatkowski, J. Polak, and M. Orlowiejska. The kluver-bucy syndrome. *Acta Neurochirurgica*, 139(4):303–306, Apr 1997.

[49] Ali Gholipour, Cameron Lucas, and D. Shahmirzadi. Purposeful prediction of space weather phenomena by simulated emotional learning. *International Journal of Modelling and Simulation*, 24:65–72, 01 2004.

[50] Gaetano Valenza and Enzo Pasquale Scilingo. *Autonomic Nervous System Dynamics for Mood and Emotional-State Recognition: Significant Advances in Data Acquisition, Signal Processing and Classification*. Springer Publishing Company, Incorporated, 2013.

[51] S. Hirose S. Osada T. Ogawa A. Tanaka M. Wada H. Yoshizawa Y. Konishi. Lateral-medial dissociation in orbitofrontal cortex-hypothalamus connectivity. *Frontiers in Human Neuroscience.10.244*, Oct 2016.

[52] Joseph Ledoux. Emotion and the limbic system concept. *Concepts in Neuroscience*, 2:169–199, 1991.

[53] Joseph E. LeDoux. Emotion circuits in the brain. *Annual Review of Neuroscience*, 23(1):155–184, 2000.

[54] Daniel Goleman. *Emotional Intelligence: Why It Can Matter More Than IQ*. Bantam Books, 2005.

[55] S. Jafarzadeh, R. Mirheidari, M. R. J. Motlagh, and M. Barkhordari. Designing pid and belbic controllers in path tracking and collision problem in automated highway systems. In *2008 10th International Conference on Control, Automation, Robotics and Vision*, pages 1562–1566, Dec 2008.

[56] M. A. Sharbafi, C. Lucas, and R. Daneshvar. Motion control of omnidirectional three-wheel robots by brain-emotional-learning-based intelligent controller. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(6):630–638, Nov 2010.

[57] G. R. Markadeh, E. Daryabeigi, C. Lucas, and M. A. Rahman. Speed and flux control of induction motors using emotional intelligent controller. *IEEE Transactions on Industry Applications*, 47(3):1126–1135, May 2011.

[58] H. S. A. Milad, U. Farooq, M. E. El-Hawary, and M. U. Asad. Neo-fuzzy integrated Adaptive Decayed Brain Emotional Learning network for online time series prediction. *IEEE Access*, 5:1037–1049, 2017.

[59] T. Miki H. Kusanagi T. Yamakawa E. Uchino. A neo fuzzy neuron and its applications to system identification and prediction of the system behavior. In *Proc. 2nd International Conference on Fuzzy Logic and Neural Networks,Japan*, pages 477–483, 1992.

[60] EIJI UCHINO and TAKESHI YAMAKAWA. System modeling by a neo-fuzzy-neuron with applications to acoustic and chaotic systems. *International Journal on Artificial Intelligence Tools*, 04:73–91, 1995.

[61] Houssen Salh Ali Milad and Jason Gu. Expanded neo-fuzzy adaptive decayed brain emotional learning network for online time series predication. *IEEE Access*, 9:65758–65770, 2021.

[62] H. S. A. Milad, U. Farooq, M. E. El-Hawary, V. E. Balas, and M. U. Asad. Fuzzy logic based parameter adjustment model for adaptive decayed brain emotional learning network with application to online time series prediction. In *2017 IEEE Electrical Power and Energy Conference (EPEC)*, pages 1–6, 2017.

[63] Yevgeniy V. Bodyanskiy, Oleksii K. Tyshchenko, and Daria S. Kopaliani. An extended neo-fuzzy neuron and its adaptive learning algorithm. *CoRR*, 2:21–26, 2016.

[64] D. Zurita, M. Delgado, J. A. Carino, J. A. Ortega, and G. Clerc. Industrial time series modelling by means of the neo-fuzzy neuron. *IEEE Access*, 4:6151–6160, September 2016.

[65] Umar Farooq, Jason Gu, Valentina Balas, Ghulam Abbas, Muhammad Usman Asad, and Marius Balas. A hybrid time series forecasting model for disturbance storm time index using a competitive brain emotional neural network and neo-fuzzy neurons. *Acta Polytechnica Hungarica*, 16:213–219, 07 2019.

[66] K. Ma and H. Leung. A novel LSTM approach for asynchronous multivariate time series prediction. In *2019 International Joint Conference on Neural Networks (IJCNN)*, pages 1–7, 2019.

[67] Michael C Mackey, Leon Glass, et al. Oscillation and chaos in physiological control systems. *Science*, 197(4300):287–289, 1977.

[68] B. B. Ustundag and A. Kulaglic. High-performance time series prediction with predictive error compensated wavelet neural networks. *IEEE Access*, 8:210532–210541, 2020.

[69] R. Volianskyi, O. Sadovoi, N. Volianska, and O. Sinkevych. Root methods for dynamic analysis of the one class chaotic systems. In *2019 IEEE 14th International Conference on Computer Sciences and Information Technologies (CSIT)*, volume 1, pages 117–121, 2019.

[70] D. ZHANG and M. Jiang. Hetero-dimensional multitask neuroevolution for chaotic time series prediction. *IEEE Access*, 8:123135–123150, 2020.

[71] Edward N Lorenz. Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, 20(2):130–141, 1963.

[72] J. Gao, H. Sultan, J. Hu, and W. W. Tung. Denoising nonlinear time series by adaptive filtering and wavelet shrinkage: A comparison. *IEEE Signal Processing Letters*, 17(3):237–240, March 2010.

[73] M. Han, S. ZHANG, M. Xu, T. Qiu, and N. Wang. Multivariate chaotic time series online prediction based on improved kernel recursive least squares algorithm. *IEEE Transactions on Cybernetics*, 49(4):1160–1172, 2019.

[74] Y. Liu, Y. Xu, J. Yang, and S. Jiang. A polarized random fourier feature kernel least-mean-square algorithm. *IEEE Access*, 7:50833–50838, 2019.

[75] J. Sun. Clustering multivariate time series based on riemannian manifold. *Electronics Letters*, 52(19):1607–1609, 2016.

[76] Q. Luo, X. Fang, Y. Sun, J. Ai, and C. Yang. Self-learning hot data prediction: Where echo state network meets NAND flash memories. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 67(3):939–950, 2020.

[77] J. Sunny, J. Schmitz, and L. ZHANG. Artificial neural network modelling of Rossler's and Chua's chaotic systems. In *2018 IEEE Canadian Conference on Electrical Computer Engineering (CCECE)*, pages 1–4, 2018.

[78] Shiquan Shao and Xin Gao. Synchronization in time-delayed fractional order chaotic Rossler systems. In *2008 International Conference on Communications, Circuits and Systems*, pages 652–654, 2008.

[79] Ali Gholipour, Babak Araabi, and Caro Lucas. Predicting chaotic time series using neural and neuro fuzzy models: A comparative study. *Neural Processing Letters*, 24:217–239, 12 2006.

[80] Ali Gholipour, Caro Lucas, Babak N Araabi, Masoud Mirmomeni, and Masoud Shafiee. Extracting the main patterns of natural time series for long-term neuro fuzzy prediction. *Neural Computing and Applications*, 16(4-5):383–393, 2007.

[81] J. G. Kappernman and V. D. Albertson. Bracing for the geomagnetic storms. *IEEE Spectrum*, 27(3):27–33, 1990.

[82] H. Lundstedt, H. Gleisner, and P. Wintoft. Operational forecasts of the geomagnetic Dst index. *Geophysical Research Letters*, 29(24):34–1–34–4, 2002.

[83] Giuseppe Pallocchia, E. Amata, Giuseppe Consolini, M. Marcucci, and Igor Bertello. Ann prediction of the dst index. *Memorie della Societa Astronomica Italiana Supplementi*, 9:120, 01 2006.

[84] M. Jawad, A. Rafique, I. Khosa, I. Ghous, J. Akhtar, and S. M. Ali. Improving disturbance storm time index prediction using linear and nonlinear parametric models: A comprehensive analysis. *IEEE Transactions on Plasma Science*, 47(2):1429–1444, 2019.

[85] Y. Gu, H. L. Wei, M. A. Balikhin, R. J. Boynton, and S. N. Walker. Machine learning enhanced narmax model for dst index forecasting. In *2019 25th International Conference on Automation and Computing (ICAC)*, pages 1–6, 2019.

[86] WDC. Geomagnetic Equatorial Dst index Home Page, April 2000, year = 2000, note =`http://wdc.kugi.kyoto-u.ac.jp/dst_final/200004/index.html."timeframe=\&year=2000\&month=4`, .

[87] K. S. Narendra and K. Parthasarathy. Identification and control of dynamical systems using neural networks. *IEEE Transactions on Neural Networks*, 1(1):4–27, 1990.

[88] D. Huang, C. ZHANG, Q. Li, H. Han, D. Huang, T. Li, and C. Wang. Prediction of solar photovoltaic power generation based on MLP and LSTM neural networks. In *2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2)*, pages 2744–2748, 2020.

[89] A. U. Haque, P. Mandal, J. Meng, M. E. Kaye, and Liuchen Chang. A new strategy for wind speed forecasting using hybrid intelligent models. In *2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1–4, 2012.

[90] Lazar Lazic, Goran Pejanovic, and Momcilo Zivkovic. Wind forecasts for wind power generation using the eta model. *Renewable Energy*, 35:1236–1243, 11 2009.

[91] M. Negnevitsky, P. Johnson, and S. Santoso. Short term wind power forecasting using hybrid intelligent systems. In *2007 IEEE Power Engineering Society General Meeting*, pages 1–4, 2007.

[92] Heping Liu, Jing Shi, and Ergin Erdem. Prediction of wind speed time series using modified Taylor Kriging method. *Fuel and Energy Abstracts*, 35:4870–4879, 12 2010.

[93] Kuilin Chen and Jie yu. Short-term wind speed prediction using an unscented Kalman filter based state-space support vector regression approach. *Applied Energy*, 113:690–705, 01 2014.

[94] Federico Cassola and Massimiliano Burlando. Wind speed and wind energy forecast through Kalman filtering of numerical weather prediction model output. *Applied Energy*, 99:154–166, 11 2012.

[95] Saurabh Soman, Hamidreza Zareipour, O.P. Malik, and Paras Mandal. A review of wind power and wind speed forecasting methods with different time horizons. pages 1 – 8, 10 2010.

[96] Ma Lei, Luan Shiyan, Jiang Chuanwen, Liu Hongling, and ZHANG Yan. A review on the forecasting of wind speed and generated power. *Renewable and Sustainable Energy Reviews*, 13:915–920, 05 2009.

[97] Xin Zhao, Haikun Wei, Chenxi Li, and Kanjian ZHANG. A hybrid nonlinear forecasting strategy for short-term wind speed. *Energies*, 13:1596, 04 2020.

[98] Tomonobu Senjyu, Atsushi Yona, Naomitsu Urasaki, and Toshihisa Funabashi. Application of recurrent neural network to long-term-ahead generating power forecasting for wind power generator. pages 1260 – 1265, 12 2006.

[99] Atsushi Yona, Tomonobu Senjyu, Naomitsu Urasaki, and Toshihisa Funabashi. Application of recurrent neural network to 3-hours-ahead generating power forecasting for wind power generators. *Ieej Transactions on Power and Energy*, 129:591–597, 01 2009.

[100] Erick Lopez, Carlos Valle, Hctor Allende-Cid, and Hector Allende. *Comparison of Recurrent Neural Networks for Wind Power Forecasting*, pages 25–34. 06 2020.

[101] climate.weather.gc.ca. Climate Data hourly data, 2020. `https://climate.weather.gc.ca/climate_data/hourly_data_e.htmlhlyRange=2002-07-127C2021-01-12\&dlyRange=2005-11-017C2021-01-1\&mlyRange=2007-07-017C2007-07-01\&StationID=31829\&Prov=NS\&urlExtension=_e.html\&searchType=stnProv\&optLimit=yearRange\&StartYear=2020\&EndYear=2021\&selRowPerPage=25\&Line=24\&Month=1\&Month=2\&Month=3\&lstProvince=NS\&timeframe=1\&Year=2020`,.

[102] Yuan-Kang Wu, Shih-Ming Chang, and Paras Mandal. Grid connected wind power plants: A survey of the integration requirements in modern grid codes. *IEEE Transactions on Industry Applications*, PP:1–1, 08 2019.

[103] Yusheng Xue, C. Yu, J. Zhao, K. Li, X. Liu, Q. Wu, and Guangya Yang. A review on short-term and ultra-short-term wind power prediction. *Dianli Xitong Zidonghua/Automation of Electric Power Systems*, 39:141–151, 03 2015.

[104] Chen Ye, Gengyin Li, and Ming Zhou. A combined prediction method of wind farm power. In *2010 5th International Conference on Critical Infrastructure (CRIS)*, pages 1–5, 2010.

[105] U. Solanki, G. P. Prajapat, and P. Jha. Unscented kalman filter based mechanical parameter estimation of wind power systems. In *2019 8th International Conference on Power Systems (ICPS)*, pages 1–6, 2019.

[106] M. Negnevitsky, P. Mandal, and A. K. Srivastava. Machine learning applications for load, price and wind power prediction in power systems. In *2009 15th International Conference on Intelligent System Applications to Power Systems*, pages 1–6, 2009.

[107] Qu Xiaoyun, Kang Xiaoning, ZHANG Chao, Jiang Shuai, and Ma Xiuda. Short-term prediction of wind power based on deep long short-term memory. pages 1148–1152, 10 2016.

[108] AESO. "alberta 12 hour wind power forecast updated as of 1/28/2021, 2021. `https://www.aeso.ca/grid/forecasting/Wind-power-forecasting`,.

[109] Hu Menghui and Liu Yian. Short-term prediction of BP neural network based on difference method. In *2020 19th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)*, pages 58–61, 10 2020.

[110] Y. Long and R. ZHANG. Short-term wind speed prediction with ensemble algorithm. In *2020 Chinese Automation Congress (CAC)*, pages 6192–6196, 2020.

Appendices

Appendices are attached in separate file.

Appendix A

Time Series Data