Addis Ababa University
Addis Ababa Institute of Technology
School of Electrical and Computer Engineering

Weather Forecasting using Deep Learning Algorithm for the Ethiopian Context

A thesis submitted to the School of Electrical and Computer Engineering In partial fulfilment of the requirements for the Degree of Master of Science in Computer Engineering

By Haftom Aregawi Hagos

Addis Ababa, Ethiopia

October, 2018

Addis Ababa University
Addis Ababa Institute of Technology
School of Electrical and Computer Engineering

Weather Forecasting using Deep Learning Algorithm for the Ethiopian Context

By

Haftom Aregawi Hagos

Advisor: Menore Tekeba

Addis Ababa University
Addis Ababa Institute of Technology
School of Electrical and Computer Engineering

Weather Forecasting using Deep Learning Algorithm for the Ethiopian Context

By

Haftom Aregawi Hagos

Approval by Board of Examiners

Dr. Yalemzewd Negash Dean, School of Electrical and Computer Engineering	______________________ signature
Menore Tekeba Advisor	______________________ signature
___________________________ Internal Examiner	______________________ signature
___________________________ External Examiner	______________________ signature

Declaration

I, the undersigned, certify that research work titled Weather Forecasting using Deep Learning algorithm for the Ethiopian Context is my own work. The work has not been presented elsewhere for assessment. Where material has been used from other sources, it has been properly acknowledged.

Haftom Aregawi Hagos ________________________

Name Signature

Place: Addis Ababa

____

Date of submission

This thesis has been submitted for examination with my approval as a university advisor.

Menore Tekeba ____________________________

Advisor's name signature

Acknowledgement

First and foremost I would like to thank my God for giving me such like opportunity, hope and strength to complete my research successfully, my parents and my colleagues for their supporting me continuously throughout the years of my study.

Second, I would like to express my gratitude and sincere thanks to my respected advisor Mr. Menore Tekeba for his professional guidance, advice, motivation, endurance and encouragements during his supervision period, Mr. Dawit Alemseged for his great contribution on my thesis, and my younger brother Mr. Alem Aregawi for his valuable comments giving me on my research work.

Finally, I would like to thanks Aksum University for allowing me to get this chance and the financial sponsorship supports, Addis Ababa University, especially school of Electrical and Computer Engineering (Computer Engineering chair staffs for their friendly relation, giving advice and encouragements.), for allowing me to study and for creating different learning opportunities from beginning up to now.

Table of Contents

List of Figures

List of Tables

Abstract

Weather forecasting is the application of science and technology to predict the state of the atmosphere for a future time and a given location. Now days, forecasting of accurate atmospheric conditions is the major challenge for the meteorologist and poor forecasting has significant impact on our daily lives. This brings the necessity to make research works on forecasting of the weather events with respect to Ethiopia.

A number of algorithms have been proposed for forecasting of atmospheric condition such as support vector machine, neural network, numerical, and statistical models. However, in this research the design and implementation of weather prediction for the Ethiopian context based on the forecasting ranges using deep neural network, support vector machine for regression, and numerical based regression is presented. Four and half years' time series daily and hourly, temperature, precipitation, humidity, visibility, dew point, air pressure, and wind historical recorded data is used from National Oceanic and Atmospheric Administrator (NOAA) to implement the system. Since making discussion on all-weather variables makes the report to long, forecasting of temperature and precipitation weather variables for Addis Ababa are only considered to be discussed and evaluated as a sample. And their results are examine based on percentage of Root Mean Square Error and time consumption. The same data records are applying for all algorithms; and the experimental result shows that, in forecasting of a big data, DBN provides a better performance relative to SVM and numerical regressions. In short range experiment we have achieved a forecasting accuracy of temperature 88.6%, 79.6%, and 52.5% using DBN, SVM, and Numerical algorithms respectively. However, if we apply a small dimension dataset as input values SVM and numerical regressions completely outperforms the DBN due to shallow training.

Keywords: *Weather, Deep Belief Network, Support Vector Machine, Restricted Boltzman Machine, Meteorology*

Abbreviations

ACE - Average Coverage Error

ANN – Artificial Neural Network

ARIMA - Auto-Regressive Integrated Moving Average Model

ARMA - Auto-Regressive and Moving Average Model

BP - Back propagation

BPN - Back Propagation Network

BPNN - Back-propagation Neural Network

CAeM – Commission for Aeronautical Meteorology

CD - Contrastive Divergence

CRPS - Continuous Rank Probability Score

CRWF - Cathedral Rocks wind farm

CT - Cloud Type

CTA - Cloud Total Amount

CTT - Cloud Top Temperature

DBN - Deep belief network

DNN - Deep Neural Network

DT - Difference of TB

EMOS - Ensemble model output statistics

ENN - Ensemble Neural Network

ERNN - Elnman Recurrent Neural Network

FC - Fuzzy clustering

GA - Genetic Algorithm

GRNN - General Regression Neural Network

GT - Gradient of the pixel TB

HFM - Hopfield Model

HKO - Hong Kong Observatory

ITCZ - Inter-tropical Convergence Zone

IS - Interval Sharpness

LM - Levenberg-Marquardt

MAR - missing at random

MATLAB – MATrix LABoratory

MAR - Missing At Random

MCAR - Missing Completely At Random

MLP - Multilayer Perceptron

MLPN - Multi-layer Perceptron Network

NMA - National Meteorology service Agency

NMSA - National Meteorological Services Agency

NMSE – Normalized Mean Square Error

MNAR - Missing Not At Random

NOAA – National Oceanic and Atmospheric Administrator

MSE - Mean square error

MSLP - Mean Sea Level Pressure

MWNN - Morlet Wavelet Neural Network

PDF - Probability density function

QR - Quantile regression

RBFN - Radial Basis Function Network

RMSE - Root Mean Square Error

SDE - Stochastic Differential Equations

SIWF - Shang chuan Island wind farm

SOM - Self-Organizing Map

SVM - Support Vector Machine

SVR - Support Vector Regression

TB - Top Brightness Temperature

WSF - Wind speed forecasting

WT - Wavelet transform

WVC - Water Vapor Content

Chapter One

1. Introduction

1.1. Introduction

Weather is the condition of the air on earth at a given time [1]. That is, a specific event or condition that happens over a period of time. Weather forecasting is the application of science and technology to predict the condition of the atmosphere from short-range to long-range and it is essential to help us to know about the future climate. Weather forecasting is the most significant attributes in the world. Because, an agricultural sector as well as numerous sectors such as, Industries, air aviation's, and so on depends on the atmospheric condition.

Before 1890, i.e. the Establishment of National Meteorological Services Agency (NMSE) [2], weather forecasting was an individual matter; that is predicting the atmospheric condition largely was dependent on individual's observation. Therefore, they were tried to observe the weather condition to predict the short-range condition of the atmosphere by try and error. However, because of different factors such as, temperature, wind, humidity, air pressure, cloud distribution and others that affects the atmosphere, their predictions were may be or may not be correct. Even now a day, those who are far apart from the meteorology technology information specifically in Ethiopia, they predict the atmospheric condition through their observations. For example, most of the time they predict related with rain fall based on temperature and wind; that is, if they observe high temperature and low wind speed, they guess there is a rain [3].

According to Commission for Aeronautical Meteorology (CAeM) Working Group on Advanced Techniques Applied to Aeronautical Meteorology researchers, modern weather forecasting starts in 1888 after the hydrodynamics H. Von Helmholtz was formulated the modern fundamental laws of atmospheric motion necessary to describe atmospheric motion [3].

In Ethiopia, Meteorological weather prediction was started at the end of 19th century in Addis Ababa by missionaries [4]. Therefore, Meteorological technology station was established in 1890 in Adamtilu and 1986[1896] in Gambela. Then, after the end of World War II, from 1946-1949 meteorology technology were carried out by the government for agriculture sector only to control locust. After that, due to the growing demands of meteorological information for safe operations of the air transport and as the other economic and social sectors began to realize the importance of meteorological services Ethiopian government officially established the National Meteorological Services Agency in December 31, 1980 under proclamation No 201 of 1980.

Now a days, especially in developing countries like Ethiopia, forecasting of weather and climate events are the major challenges for the meteorologist. Even so, a lot of researches had been done in this area using different mechanisms such as numerical method, statistical method, and neural network, but the final outcome was not up to the standard.

Therefore, the focus of this thesis is mainly on forecasting the weather with respect to Ethiopian context; which includes, how to forecast weather in Ethiopia based on short-range, medium-range, and long-range using deep learning algorithm, support vector machine, and numerical methods based on the four Ethiopian seasons such as, summer, winter, spring, and fall and rainfall regimes.

1.2. Background of Weather and Climate in Ethiopia

Ethiopia is located between approximately 4^0-15^0 latitude and 32^0-48^0 longitude [5] [6] [7] and her land area occupies around 1.12 million square kilometers in the horn of Africa, [6] with a total population of more than 90 million.

Due to high and rugged mountains; flat topped plateau, deep gorges, river valleys and plains geographical topography; Ethiopia has great geographical diversity from the horn of Africa [6][7]. Therefore, due to the aforementioned reasons, Ethiopia is unique from African countries. The country is broadly divided into highlands and lowlands with altitude range above 4620 meter above sea level and 120 meter below sea level [7]. However, according to Ethiopian National Meteorology service Agency (NMA), Ethiopia classified with altitude range 1500 meter above sea level as highlands which covers 45% of the country and with altitude range 1500 meter below sea level as low lands. Furthermore, in terms of climate, lifestyle, socio-economic activities, and so on; there is an essential difference in Ethiopia.

As described in Ethiopian NMA, Ethiopia has three major physio-graphical regions. Such as,

- ✓ The North, Central, and Southwestern Highlands and the associated Lowlands;
- ✓ The Southeastern Highlands and the associated Lowlands; and
- ✓ The Ethiopian Rift Valley.

1.2.1. Ethiopian Climate

Ethiopia is located in a tropical region lying between the equator and tropical cancer of the earth [8]. Therefore, Ethiopia generally has a **tropical climate** and one of the influences of Ethiopia's climate and weather is the seasonal migration of the Inter-tropical Convergence Zone (ITCZ); that is the position of the earth relative to the sun and by the complex geographical topography of the country [5].

Traditionally, based on altitude and temperature Ethiopia has five climatic zones [9]:

- ✓ **Wurch** which is a cold climate with altitude range 3000 meters above sea level.
- ✓ **Dega** (Cool Zone) which is a temperate atmospheric condition with altitude ranges approximately between 2500-3000 meters above sea level
- ✓ **Woina dega (**Sub-Tropical Zone**)** which is a warm climate with altitude ranges approximately between 1500-2500 meters above sea level
- ✓ **Kola** (Tropical Zone)which is a hot and arid type atmospheric condition with below 1500 meters altitude range; and
- ✓ **Bereha** which concludes hot and hyper-arid climates.

Generally, Ethiopia has four climatic Zones [7], which are classified by rainfall distribution. These are, with a distinct wet and a distinct dry season, with two wet and two dry seasons, with two wet seasons and one dry season in between, and with an undefined rainy season.

1.2.2. Ethiopian Seasons and Rainfall Regimes

The economy of Ethiopia depends on the agriculture [4]. Besides, the agriculture of Ethiopia is highly dependent on the atmospheric variable which is rainfall. Therefore, the Ethiopian climate is characterized by the variation of rainfall.

Based on rainfall distribution, there are four major seasons in Ethiopia:

- ✓ **Kiremt or Meher** (summer, June – September): is the main rainy season for the most country locution, except in the south and south-eastern part of the country.
- ✓ **Belg** (Autumn, September - November): Sometimes known as the harvest season
- ✓ **Bega** (winter, February – March): are the dry season with frost in morning especially in January.
- ✓ **Tseday** (spring, March - May): are the autumn season with occasional showers. May is the hottest month in Ethiopia.

As we discussed in the above, Ethiopian climate is highly influenced by Inter Tropical Convergence Zone (ITCZ) [10] and complex geographical location. Due to this reason, Ethiopia is classified into three climatic zones which have two, three, and four seasons in a year [7]. The first one is two seasons zone, which includes the western half of Ethiopia and has one rainy season (half of June – half of September) and one dry season (November - February). The second one is three season zones, which contains the central (including Addis Ababa) and most of the eastern part of the country with two rainy seasons known as Kiremt (main rainy season) and Belg (small rainy season), and one dry season known as Bega (Octeber - January). The last one is four seasons zone, which occupies south and south-eastern of Ethiopia with with two rainy seasons (very rainy-March - June and small rains-September - November) and two distinct dry seasons (July - August and December - February).

Generally, based on mean annual and mean monthly rainfall distribution, three rainfall regimes are commonly identified in Ethiopia [8]. These are, Mono-Modal (Single maxima), Bi-modal type-1 (Quasi-double maxima), and Bi-modal type-2 (Double maxima)

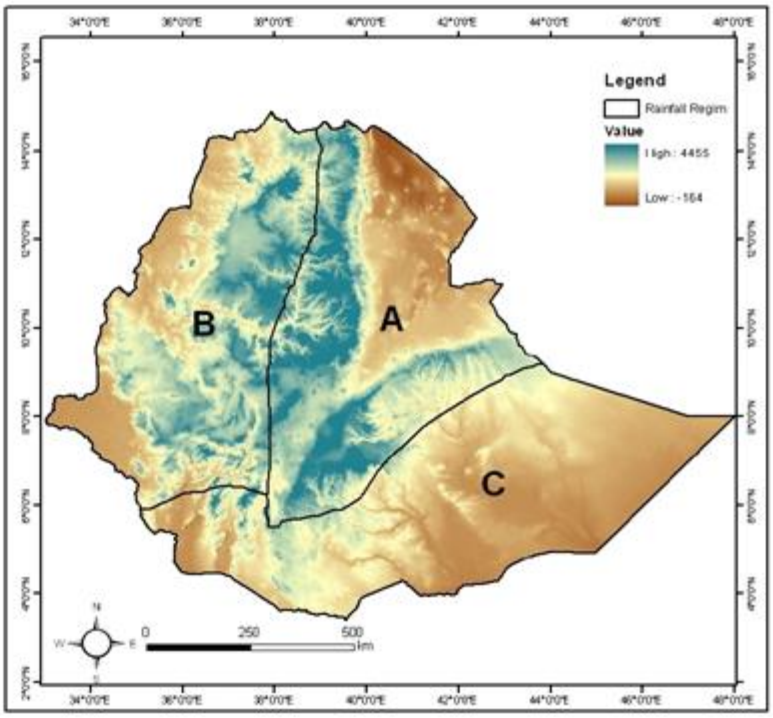


Fig1. 1 Topographic map of Ethiopia showing rainfall regimes [12]

The figure shows Ethiopia rainfall regime namely Mono-modal, Bi-modal type-1 and Bi-modal type-2 are designated by letter B, A and C, respectively.

Regime A: it occupies the central and the eastern part of the country with a bi-modal rain. It has three season zones with two rainy season which are Kiremt (long rainy season) and Belg (short rainy season), and one dry season known as Bega.

Regime B: is dominated by single maxima rainfall pattern that is from June to September. It concludes western part of the country that is from south-west to north-west. Besides, except northwards rainy is occur from February to November in the western and south-western of the country.

Regime C: is dominated by double maxima rainfall pattern in the south and south-eastern part of the country. This rainfall regime has four season zones which are two rainy seasons these are the main rainy season from February to May and short rainy season from October to November, and two dry periods from June to August and from December to February.

1.3. Statement of the problem

According to the Ethiopian National Meteorological Services Agency (**NMSA**), Ethiopia generally has a tropical climate and one of the influences of Ethiopia's climate and weather is the seasonal migration of the Inter-Tropical Convergence Zone (ITCZ) which affects the rainfall patterns in Ethiopia. Besides, Ethiopian meteorology uses static and dynamic methods to predict the atmospheric conditions and most of the time they try to predict short-range and medium range rainfall distribution and other weather events based on the Ethiopian seasons such as, summer, winter, spring, and fall.

The existing forecasting system does not consider what type of forecasting modeling is appropriate for Ethiopia and how to implement the system with the Ethiopian context. Therefore, the motivation behind this research is to implement more accurate weather forecasting system for Ethiopian context in contrast to the existing forecasting system. Since the existing forecasting method is not considering the Ethiopian context, it is difficult to predict the atmospheric condition accurately.

In order to solve the aforementioned problems it is better to design more appropriate forecasting method and implement which fits to the Ethiopian context and provides better accuracy in forecasting. Besides, to predict the drought that happens in our country, we can implement long-range forecast using the proposed algorithm.

1.4 Research Questions

- ✓ Does forecasting of the weather events using DBN network much better than SVM for regression (i.e. for forecasting) and numerical based regression?
- ✓ Does DBN network based forecasting of the atmospheric condition suitable for the Ethiopian Context?
- ✓ Is the variation of atmospheric condition solved using DBN and SVM for regression better than the existing forecasting method called numerical method regression? Which one is more appropriate for the Ethiopian context?

1.5. Objective

1.5.1. General Objective

The general objective of this thesis is to implement weather forecasting system using deep learning algorithm for the Ethiopian context which can forecast Ethiopian weather conditions with better accuracy than the existing system.

1.5.2. Specific objectives

- ✓ To select the appropriate types of weather forecasting models and methods for the Ethiopian context.
- ✓ To enhance short range, medium range, and long range weather prediction for Ethiopia.
- ✓ To enhance the overall prediction capability of weather forecasting in Ethiopia by introducing new weather forecasting methods.

1.6 Methodology

The methodologies used in this thesis are given below.

- ✓ **Literature survey-** Reviewing of supportive theoretical concepts and different literatures which are related to my work.
- ✓ **Select algorithm-** Selecting of appropriate algorithms from deep learning neural networks which has high efficient during forecasting the atmospheric conditions.
- ✓ **Propose the system-** By studying the existing system; propose the new system with new forecasting model which is suitable for Ethiopian context.
- ✓ **Design and Implement-** Designing and implementing of the proposed system using simulation software.

- ✓ **Test and Validate-** Testing and validating of the developed system and the result achieved from the proposed system respectively.

1.6. Scope

The scope of this thesis is to propose appropriate and suitable forecasting model and algorithm for the Ethiopian context by considering the four season and three rainfall regimes. The proposed system is evaluated using MATLAB program with the selected weather input variables namely temperature and precipitation. Besides, the result of the proposed system in this thesis is compared with the result of numerical based and support vector machine for regression methods.

1.7. Contribution

Weather and climate is a resource and considered as basic factors that affect directly or indirectly individuals, industries, and others of daily activities. Based on this reason, this thesis has a lot of contribution for our country regarding to individuals, agriculture, aviation, industries, and other socio-economic sectors as listed below.

- ✓ Its main contribution is, to enhance the forecasting mechanisms in Ethiopia by considering an appropriate model and optimization.
- ✓ It helps for selecting and implementing of appropriate forecasting algorithms for the Ethiopian context.

1.8. Thesis organization

The organization of this thesis has six chapters as it arranges in the following manner:

Chapter One- Introduction: Describes a general introduction about the thesis and the background with brief explanation of weather in Ethiopia.

Chapter Two- Literature Review: Reviews of different literatures which are related to the selected area.

Chapter Three- Design: Discusses all about the basic designing and modeling steps and methods of the proposed system in detail. Additionally, it concludes brief explanation of different bench marks as well as input variables.

Chapter Four- Implementation: It discusses the implementation of different algorithms using simulating software. Besides, the overall flow chart of the proposed system is discussed in detail here.

Chapter Five- Result and Discussion: In this chapter discussed the empirical result which has got from the chapter four. Additionally, it describes the explanation and comparison of the results of the different algorithm implemented in this thesis.

Chapter Six- Conclusion and Recommendation: In this chapter it describes the overall conclusion of the thesis followed by the feature works.

All in all, this thesis ends up with an appendix which has the source codes of the three implemented algorithms and references.

Chapter Two

2. Literature Review

2.1. Introduction

In modern world forecasting of atmospheric condition is the most challenging task for the meteorology departments of environment; and due to its highly non-linearity, dynamically, and complexity, it requires accurate computer modeling and simulation for reliable prediction. Besides, accurate weather prediction is very necessary in multiple areas such as, agriculture, airport, and other socio-economic industries.

Now a days, due to the aforementioned reasons there are various researchers made on weather forecasting using different techniques such as, ANN, Numerical based, and Statistical, etc. In addition to that, the researchers tried to improve existing weather forecasting with a specific parameter and some researchers worked to figure out the basic problems happen due to weather forecasting defect. However, it is difficult to cover all the researches which are worked by different researchers. Therefore, we compiled the most important researches by categorizing into three such as ANN, Numerical method, and Statistical method based which are related to my work here.

2.2 Artificial Neural Network Based Weather Forecasting

ANN models can be categorized into associating networks which are employed for data classification and prediction, feature extraction networks which are used for data dimension reduction, and non-adaptive networks which needs input values to learn the pattern of the inputs and reconstruct them when presented with incomplete data sets, based on their function.

In 2002, W. Taylor James *et al.* [11]: the researchers proposed “Neural Network Load Forecasting with Weather Ensemble Predictions” in order to investigate the use of weather ensemble predictions in the application of neural networks to load forecasting for lead times from 1 to 10 days ahead in contrast to the traditional weather forecasting using three weather parameters such as effective temperature, wind, and effective illumination. Finally, they conclude there is strong potential for the use of weather ensemble predictions in neural network load forecasting.

In May 2014, Malik Pooja [12]: The researcher proposed “An Effective Weather Forecasting Using Neural Network” to predict the atmospheric condition using the real data such as temperature, pressure, humidity, wind, direction which record in Rice Research center (Kaul) Haryana and these data are trained by Levenberg-Marquardt (LM) algorithm. Besides, he argues back propagation is the most important algorithm to train a neural network for weather forecasting and he concludes from the many back propagation (BP) algorithms, Levenberg BP has better learning rate.

In 2012, Abhishek Kumar *et al.* [13]: In order to develop an ANN model to forecast average monthly rainfall in the Udupi district of Karnataka, they proposed “A Rainfall Prediction Model using Artificial Neural Network” by considering the input data (Average humidity and wind speed) and output data (average rainfall). They have used Nntool and Nftool of the MATLAB to implement the proposed algorithm. Besides, the algorithms used were feed forward with BP, layer recurrent, and cascaded feed forward BP. From these tested algorithms, the researchers have achieved good performance using BP algorithms with its mean square error 3.6456. Finally, they observe from the three tested algorithms, as the number of neurons increases, the mean square error (MSE) decreases.

In Jan 2013, Kumar Neeraj *et al*. [14]: The researcher proposed “A Time Series ANN Approach for Weather Forecasting” to predict the atmospheric condition by considering only average maximum and minimum temperature. They uses 60 years of historical real data and they train these data’s using multilayer perceptron; and their result confirms based on Mean Square Error (MSE) functions. Finally, the researchers conclude that Multilayered Neural Network can be an effective tool in weather prediction. Even so, they were not comparing the new proposed model result with other weather forecasting methods.

July 2015, Narvekar Meera *et al.*[15]: The researchers proposed “Daily Weather Forecasting using Artificial Neural Network” to investigate a better approach for forecasting the atmospheric condition which compares with different ANN techniques like Ensemble Neural Network (ENN), BP network, Radial Basis Function Network (RBFN), General Regression Neural Network (GRNN), Genetic Algorithm (GA), Multilayer Perceptron (MLP), Fuzzy clustering (FC), etc. which are used for different types of forecasting by using real time data such as temperature, relative humidity, air pressure, wind speed and direction, cloud amount and height, and rainfall. Finally, the researchers checked their result using Mean Square Error (MSE) functions and they achieve good performance in the new proposed system relatively to the other weather forecasting techniques; and they highly recommended us it is better to use BP networks to predict the daily atmospheric condition.

In 2009, Hung N. Q. *et al.*[16]: The researchers proposed "an Artificial Neural Network Model for Rainfall Forecasting in Bangkok, Thailand" in order to improve real time rainfall forecasting performance and flood management in Bangkok, Thailand. At the beginning they implement a generalization feed-forward ANN model using a hyperbolic tangent transfer function with the combination of meteorological parameters such as relative humidity, air pressure, wet bulb temperature, and cloudiness and they achieved satisfactory result. Additionally, to reduce the complexity of the network and training time, they implement sensitivity analysis. Relatively to the model of simple persistent method, the ANN forecasts have superiority in predicting short range atmospheric condition.

In 20 May 2004, Maqsood Imran *et al.*[17]: Based on a real data of temperature, wind speed, and relative humidity of the weather parameters, the researchers proposed "An ensemble of neural networks for weather forecasting" to forecast and analysis the atmospheric condition of the Southern Saskatchewn, Canada. The new proposed system is developed based on the predictive models such as Multi-layer Perceptron Network (MLPN), Elnman Recurrent Neural Network (ERNN), Radial Basis Function Network (RBFN), Hopfield Model (HFM) predictive models, and regression techniques, examination of the applicability of the ANN approach for weather forecasting, comparison of the proposed predictive models with regression models, and performance quantification of the developed models using statistical measures. After that, the researchers observe from their empirical results, MLPN and ERNN networks did equally well in forecasting the given weather parameters, HFM is relatively less accurate, and RBFN is relatively more accurate and reliable. In comparison, the empirical result shows that the ensemble networks can be trained effectively without excessively compromising the performance. Due to this reason, the researchers conclude that the ensemble of neural network model produced the most accurate weather forecast in contrast to the predictive models which are mentioned in the above.

In June 2013, Ranjan Nayak Deepak [18]: The researchers proposed "A Survey on Rainfall Prediction using Artificial Neural Network" in order to decide and select the forecasting methods during we design a new system of rainfall prediction and to make the prediction of rainfall more accurate for the future. Therefore, the researcher provides a survey of various literatures of some methodology uses by different researchers to make use of ANN for rainfall prediction for the last 25 years; and they discusses the concept of basic ANNs. Besides, they describe some common neural network models used by different researchers such as Back Propagation Network (BPN), Radial Basis Function Network (RBFN), Support Vector

Machine (SVM), and Self Organizing Map (SOM). In contrast to the ANN weather prediction techniques, the researchers who are uses the traditional weather predicting methods which are statistical and numerical methods got a result below the satisfactory level. And also, from the survey they make sure, most of the researchers used BPN for rainfall prediction and got satisfactory results. Finally, they conclude that the forecasting techniques that use MLP, BPN, SOM, and SVM are suitable to predict rainfall than other forecasting techniques such as statistical and numerical methods.

In 12–19 July 2016, Chen Kai *et al*. [19]: The researcher proposed "Short-Term Precipitation Occurrence Prediction for Strong Convective Weather Using Fy2-G Satellite Data: A Case Study of Shenzhen, South China" in order to address the short term precipitation occurrence prediction challenges happen in Shenzhen. They use a real data of Top Brightness Temperature (TB), Cloud Top Temperature (CTT), Gradient of the pixel TB (GT), Difference of TB (DT), Cloud Total Amount (CTA), Cloud Type (CT), Middle and Upper Tropospheric Water Vapor Content (WVC), and from ground automatic station, i.e., Wind Speed, Humidity, Temperature, Air Pressure. They use support vector machine to predict short-term precipitation and they achieve better performance relatively to numerical simulation modeling. Besides, they argued numerical simulation model still remains low reliability in contrast to machine learning modeling.

In August 10-13, 2015, Grover Aditya, *et al.* [20]: the researchers proposed "A Deep Hybrid Model for Weather Forecasting" to explore a new directions with forecasting weather as a data intensive challenge that involves inferences across space and time and to model a weather forecasting to predict the atmosphere condition by considering the joint influence of key weather elements such as wind speed, atmospheric pressure, temperature, and dew point. Their approach to building the weather model was governed by temporal mining, spatial interpolation, and Inter-variable interactions guidelines. Finally, by applying deep belief algorithm they achieve satisfactory performance in contrast to the NOAA benchmarks.

N.K. Liu James *et al.*[21]: The researchers proposed "Deep Neural Network Based Feature Representation for Weather Forecasting" in order to investigate how to apply DNN to solve time series problems using real data's of the meteorology and to explore the use of DNN to deal with the large volume of weather data. The researchers use 30 years of real data which are temperature, dew point, Mean Sea Level Pressure (MSLP), and wind speeds observed by the

Hong Kong Observatory (HKO). Their experiments performance checked using Normalized Mean Square Error (NMSE). And they achieve good performance relatively to the SVR.

In May 12, 2016**,** Gao Bing *et al.*[22]: the researcher proposed "Using Data Mining Technique to Predict Seasonal Climate Change" to analyze atmospheric condition data in Vancouver Island based on the temperature data only. They use ARIMA (i.e. time series model) to build the model and they test its performance using Dickey-Fuller method. According to the researchers result, ARIMA allow for the inclusion of information from the past observations of a series. Although, they are not sure weather it is sufficient or not when they use different atmospheric parameters at the same time.

In Nov, 1999, Palmer T.N. [23]: The researcher proposed "Predicting uncertainty in forecasts of weather and climate" in order to deal the problem of forecasting uncertainty in weather and climate prediction. He tries to solve using perfect deterministic forecast model using probability density function (pdf) with Liouville equation. But, due to a knowledge gap problem, he does not solve this equation. Finally, he addressed the problem by developing ensemble prediction and he concludes ensemble prediction is a perfect application for parallel computing!

In Apr, 2016, Yadav Rohit Kumar, *et al*. [24]: The researchers are proposed "A Weather Forecasting Model uses the Data Mining Technique" to design efficient weather prediction system in contrast to the traditional algorithm ID3. In order to identify different patterns of the atmosphere they use hidden Markov model and to predict and extract the weather condition observation they use K-means clustering. Finally, they apply data mining technique to analyze the data and extract the valuable patterns from the data and they achieve high performance in contrast with the traditional algorithm (i.e. ID3) based weather forecasting system.

In 25 August 2016, Wang H.Z., *et al.*[25]: researchers proposed to design and implement a novel deep learning approach which is "Deep belief network based deterministic and probabilistic wind speed forecasting approach" originally to enhance the performance and prediction efficiency of the wind speed forecasting (WSF) using wind speed parameter. The new forecasting system proposed approach is a hybrid of Wavelet transform (WT), deep belief network (DBN) and spine quantile regression (QR). The performance and efficiency of their empirical result calculated by using Errors for probabilistic performance techniques such as Average Coverage Error (ACE), Interval Sharpness (IS), and Continuous Rank Probability Score (CRPS). In addition to that, to fully validate the effectiveness of the proposed algorithm, the results are compared with the Auto-Regressive and Moving Average Model (ARMA), the well-tuned Back-propagation Neural Network (BPNN), and the Morlet Wavelet Neural Network (MWNN) and they tested and benchmarked at two wind farms, the Shang chuan Island wind farm (SIWF) in Guangdong Province, China, and the Cathedral Rocks wind farm (CRWF) in Australia. Quantitatively, considering SIWF, the average of average coverage error using the proposed approach is 0.6742%, which has been improved by 23.96%, 26.79%, and 26.12% compared to ARMA, BPNN and MWNN. After all, they conclude non-linear features of the wind and applying uncertainty principles are used to improve forecast accuracy and reliability.

2.3. Statistical Based Weather Forecasting

In 17 February 2016, Taillardat Maxime *et al.* [26]: The researchers proposed "Calibrated Ensemble Forecasts Using Quantile Regression Forests and Ensemble Model Output Statistics" for post processing ensembles by using surface temperature and wind speed variables. The researchers achieve satisfactory result on both surface temperature and wind speed in contrast to ensemble model output statistics (EMOS).

2.4. Numerical Based Weather Forecasting

In 2015, B. Iversen Emil *et al.* [27]: The researchers proposed "Short-term probabilistic forecasting of wind speed using stochastic differential equations" modeling framework in order to provide reliable wind power forecasts in continuous time. They use a real wind speed dataset to test and validate the new proposed model (i.e. stochastic differential equations (SDE) based model). Finally, they apply an autocorrelation for the standardized Residuals resulting from their model and they achieve satisfactory result. However, they did not compare their result with different predicting models.

2.5. Summary

In this chapter mainly focuses on assessing of several research papers which are related to the proposed works. Since it is difficult to cover all the papers which are worked by different researchers, we compiled the most important researches by categorizing into three such as ANN, Numerical method, and Statistical method based weather forecasting models.

Chapter Three

3. Design of the System

3.1. Introduction

The new proposed system generally categorizes into three phases specifically preprocessing phase, design and train phase, and predicting phase. The preprocessing phase mainly focuses on selecting and gathering weather parameters, handling missing values, segmentation, and normalizes the dataset. The design and train phase provides designing of the new proposed network and train the network using the normalized dataset that we have got from the preprocessing phase. It also provides calculating the performances of the proposed networks, comparing the result of deep belief network and support vector machine, and testing and validating. The predicting phase uses the trained neural network in order to determine the prediction of weather condition.

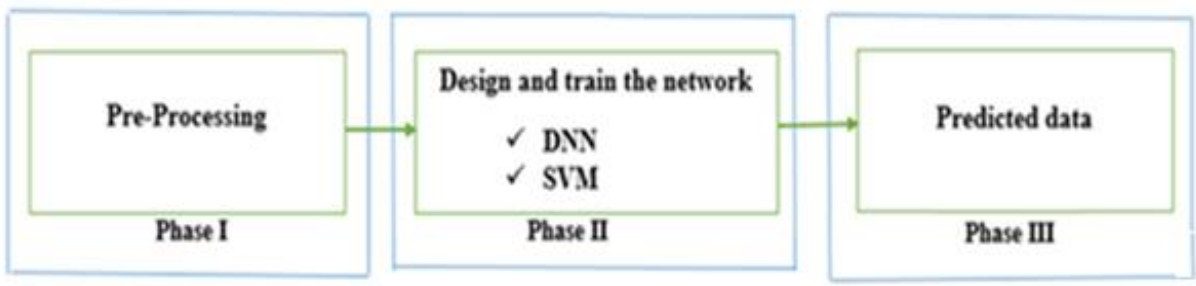


Fig 3. 1 Block Diagram of the proposed system

3.2. Preprocessing

Generally real world data are often incomplete, noisy, and inconsistent. Therefore, to resolve such issues we should use data preprocessing technique. Data preprocessing is a data mining technique which refers to cleaning, integration, transformation, and reduction of training dataset [28].

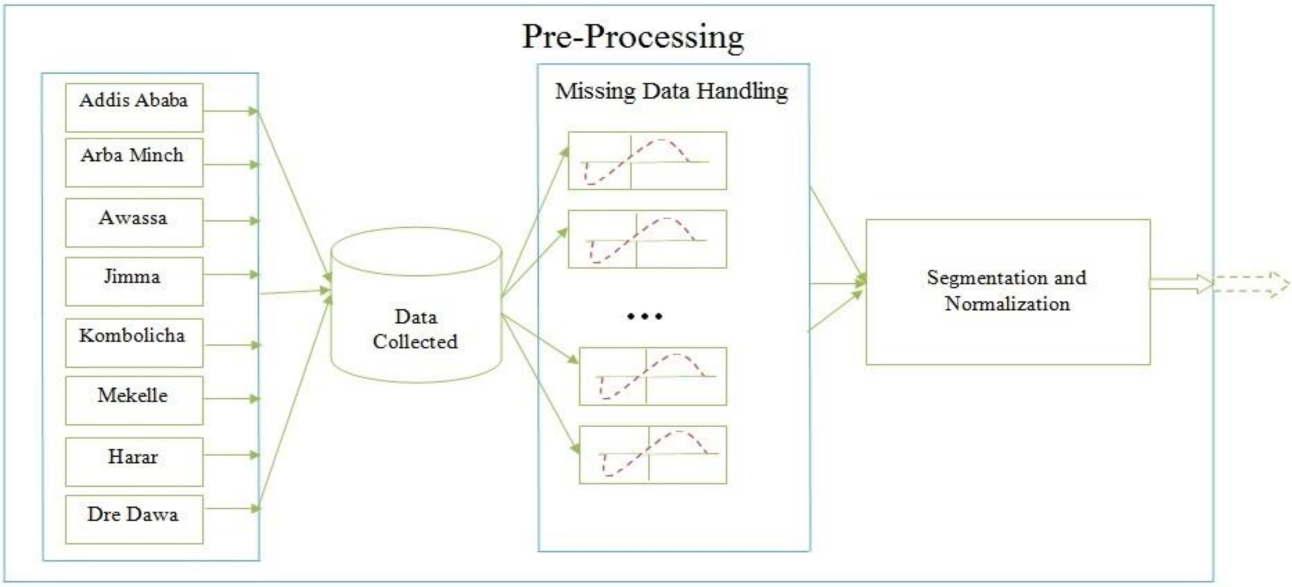


Fig 3. 2 Block Diagram of preprocessing phase

3.2.1. Selection of weather variables

Weather variable selection is the most important method in designing a new model or network to predict the atmospheric condition. Therefore, in order to design and implementing weather forecasting with the Ethiopian context we selected different atmospheric variables such as maximum and minimum of temperature, humidity, air pressure, wind direction, dew point, precipitation, and visibility weather variables.

3.2.2. Data collection

Based on seasonal weather condition and geographical topology of Ethiopia, we uses four and half year's daily and hourly (i.e. Sep 2012 – Feb 2017) real time series datasets of eight towns such as Addis Ababa, Jimma, Mekelle, Kombolicha, Harar, Arbaminch, Awasa, and Dredawa from [29].

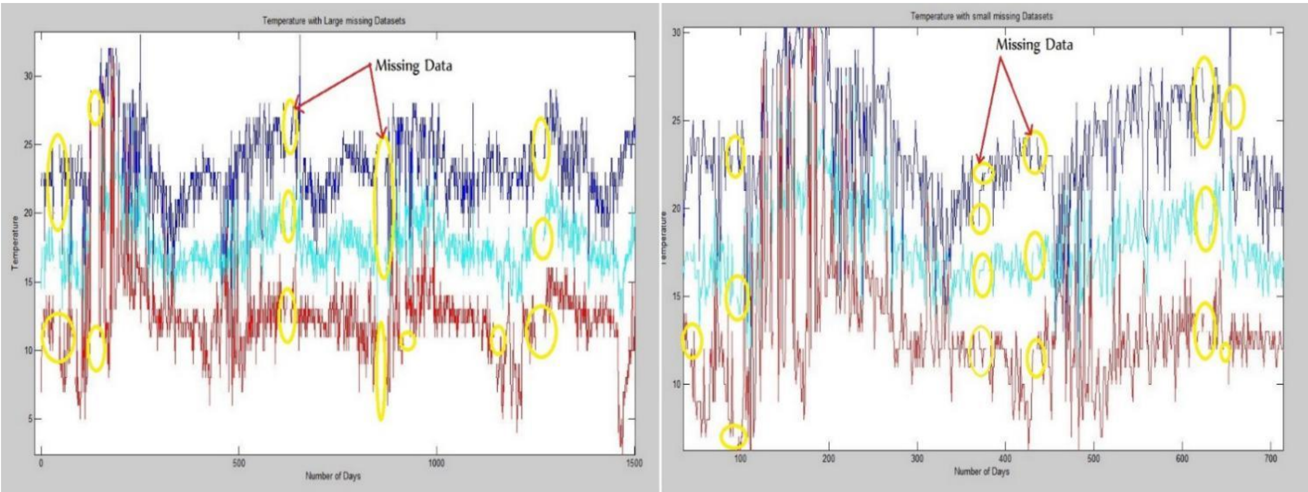
3.2.3. Handling missing data

The data collected from different stations are not complete, reliable, and consistent. Due to this reason, it is difficult to forecast the weather condition using these incomplete datasets [30].Besides, some variables air pressure are not fully recorded. In such cases, we provided statistical and numerical methods to fix these missing datasets. In [31] missing data mechanism is classified into three categories such as missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). And these mechanisms represent the statistical relationship between observations (variables) and the probability of missing data.

Based on the aforementioned mechanisms various researchers made on handling a missing datasets using different methods such as data deletion, imputation, and augmentation. Even so, it is difficult to handle a missing data using all methods that we have read on the literature. Therefore, we tested and handled the missing data set using cubic spline interpolation and linear interpolation methods based on small and large missing data sets. But, the effect of numerical interpolation is not considerd.

3.2.3.1. Handling Small Missing Datasets

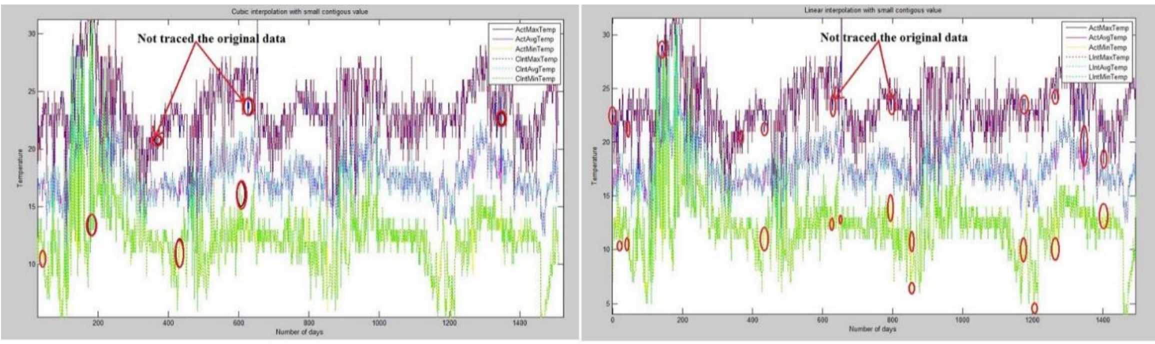
In [32] depicted, in contrast to the other polynomial interpolation techniques, Cubic Spline interpolation is the most effective on handling a small amount of missing datasets i.e. not more than four contiguous values. Furthermore, they point out Cubic Spline interpolation fails when large contiguous values are missing. Besides, the researcher discussed when large amount of data is missing in the time series data, it is better to use linear interpolation. Therefore, we use Addis Ababa's temperature daily datasets as sample to ensure their argument.



(a) (b)

Fig 3. 3 (a) and (b) Actual maximum, minimum, and average temperature of Addis Ababa with random missing large and small contiguous values respectively

The Figure 3.3 (a) and (b) indicates the graph of temperature datasets with small and large contiguous missing values respectively. The missing datasets on both graphs are taken randomly.



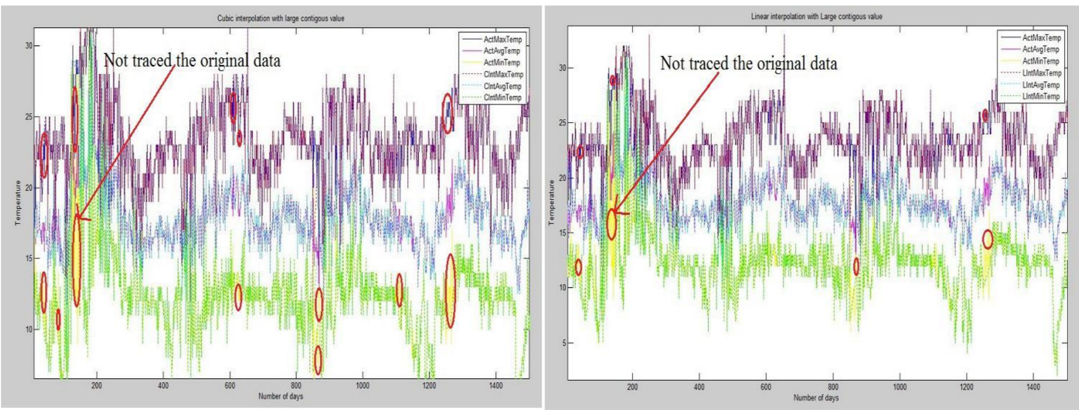
(a) (b)

Fig 3. 4 (a) and (b) Actual maximum, minimum, and average temperature of Addis Ababa using cubic spline interpolation and linear interpolation with small missing values respectively

The figure 3.4 (a) and (b) shows Cubic Spline interpolated and Linear interpolated against the original daily temperature datasets over a period of five years. As we have seen the interpolated data from these figures, Cubic Spline interpolation traces the original data very closely except in three places during interpolated the maximum temperature and four during interpolated the minimum temperature as indicated in the plot. However, in contrast to Cubic Spline interpolation, linear interpolation fails almost in more than ten places during interpolated both maximum and minimum temperature as indicated in the plot.

3.2.3.2. Handling Large Missing Datasets

Figure 3.5 (a) and (b) indicates cubic spline and linear interpolations against the original temperature data records of the graph of large amount of data missing that indicated in the figure 3.3 (b).



(a) (b)

Fig 3. 5 (a) and (b) Actual maximum, minimum, and average temperature of Addis Ababa using cubic spline interpolation and linear interpolation with large missing values respectively

As we have seen the result from the two graphs (i.e. from fig 3.5 (a) and (b)), the cubic spline interpolation fails when a large data is missing. In contrast, linear interpolation traces the original data closely. Therefore, when large amount of data is missing, it is better to handle using linear interpolation method. Handling of a small contiguous values using cubic spline and linear interpolation is 89.75% and 76.2% of accuracy is respectively achieved. However, during handling of large contiguous missing values 71.45% and 80.33% of accuracy using cubic spline and linear interpolation respectively.

3.2.4. Data segmentation

In this section the time-series data is divided into a sequence of discrete segments based on Ethiopian seasons such as Kiremt or Meher (Summer, June-August), Belg (Autumn, September-November), Bega (Winter, December-February),Tsedey (Spring, March-May) and rain fall regimes such as Regime A (Bi-Modal type 1 or quasi-double maxima), Regime B (Mono-Modal or single maxima), and Regime C (Bi-Modal type 2 or double-maxima).

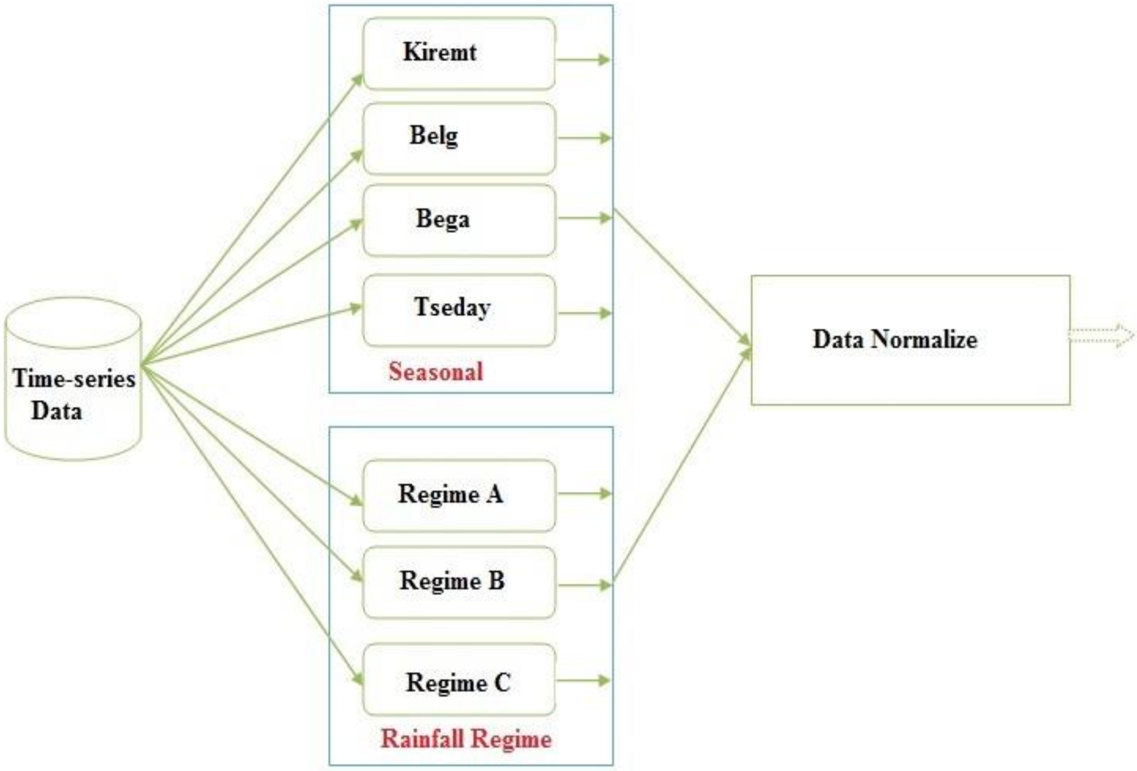


Fig 3. 6 Block diagram of Time-Series data segmentation

As in [2] discussed, the economy of Ethiopia is highly dependent on agriculture and in [1] mentioned, temperature and rainfall variables are the main factor that affect Ethiopian agriculture. Therefore, the block diagram of time-series data segmentation shows the classification of Ethiopian weather based on temperature, rainfall distribution, and other weather variables.

3.2.5. Data normalization

Normalization refers to normalizing the data dimensions so that they are of approximately the same scale [33]. Input data normalization is fundamental in a training process, which is a crucial to obtain a good result from the network. In [34] [35] [36] depicted, the tow originates of data normalization such as the normalization which originate from linear algebra and statistics. Additionally, in [36] discussed, for BPNN model, input data normalization depends on the selected activation function. For example, if we use a standard sigmoid function, each input should be normalized between 0 and 1. However, if we use hyperbolic tangent, the input should be normalized between -1 to 1, otherwise many hidden neurons will have the fixed value - 1, 1 or 0.

Therefore, after the input data are segmented seasonally and based on regimes, we normalize them based on the following equation [37].

$$\tilde{a} = \frac{a}{||a||} \quad (1)$$

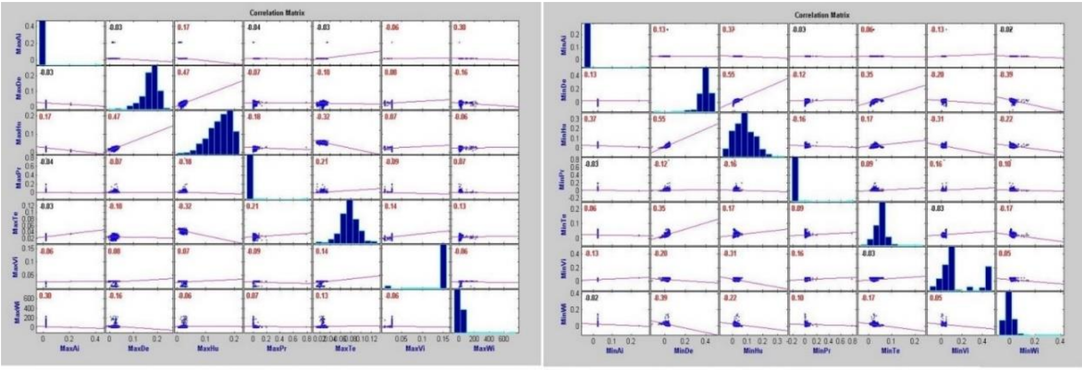
Where: $a = [a_1, a_2, a_3, \ldots, a_n]$, $||a||$ = norm a, $\tilde{a}$ is the normalized vector of "a"

$$||a|| = \sqrt{a1^2 + a2^2 + \cdots + an^2} \quad (2)$$

3.3. Correlation among input and output variables

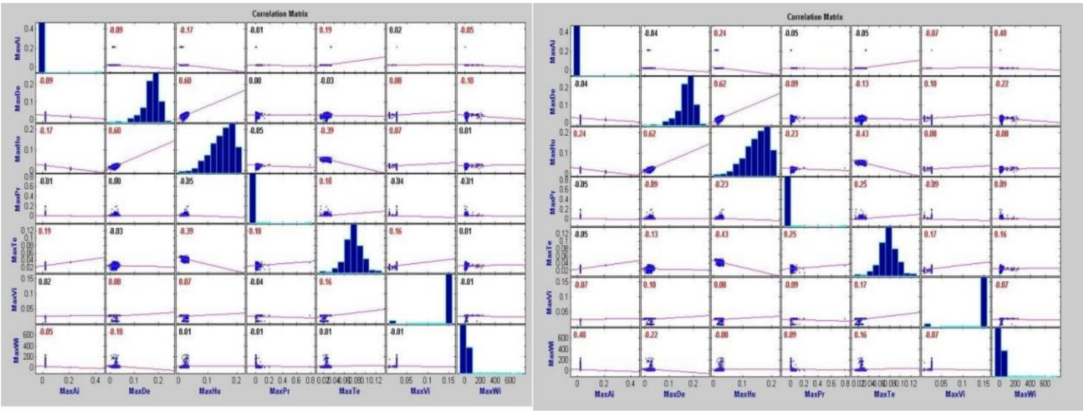
Basically, selecting input and output weather variables are the most crucial task during predicting the atmospheric condition. Therefore, in order to select the input variables we implement a correlation which is a bivariate analysis that measures the strengths of association between two variables and the direction of the relationship here.

In statistics, we measure four types of correlations [38] [39] [40]: Pearson correlation, Kendall rank correlation, Spearman correlation, and the Point-Biserial correlation. However, in this thesis the correlation among the predictor and predicting variables are implemented by comparing among Kendall, Pearson, and Spearman correlations.



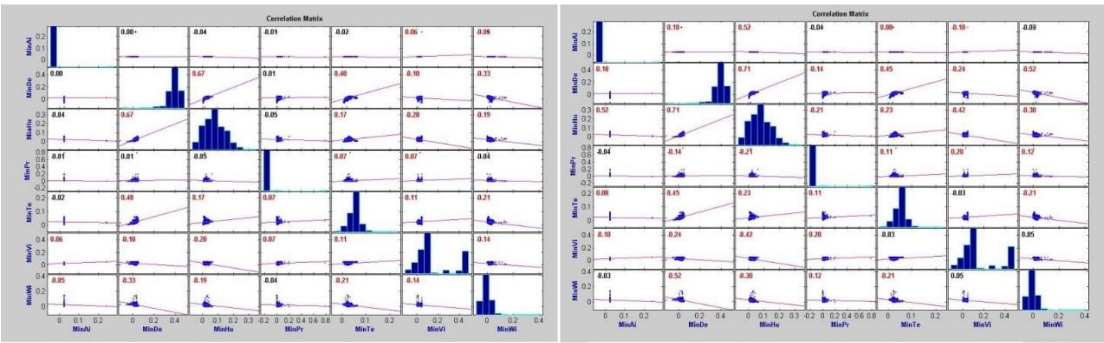
(a) (b)

Fig 3. 7 (a) and (b) Matrix correlation all weather variables using their daily maximum and minimum after normalized the actual values respectively (Kendall)



(a) (b)

Fig 3. 8 (a) and (b) Matrix correlation all weather variables using their daily maximum after normalized actual values (Pearson and Spearman respectively)



(a) (b)

Fig 3. 9 (a) and (b) Matrix correlation all weather variables using their daily minimum after normalized actual values (Pearson and Spearman respectively)

By taking four and half years weather variables data recorded of Addis Ababa as sample, we performed correlation among each variable. For instance, as we have seen in the figures, both figure 3.7 (a) and (b) shows the Kendall matrix correlation among the variables using their maximum and minimum value respectively after normalized of the given data sets as a sample.

As we have seen in the above figures (figure 3.7, 3.8, and 3.9), Spearman correlation is more powerful than Kendall's and Pearson's statistical correlation. Therefore, the input and output variables are selected based on the Spearman's correlation result as follow:

Table 1 Selection of Input and Output variables based on Spearman's correlation

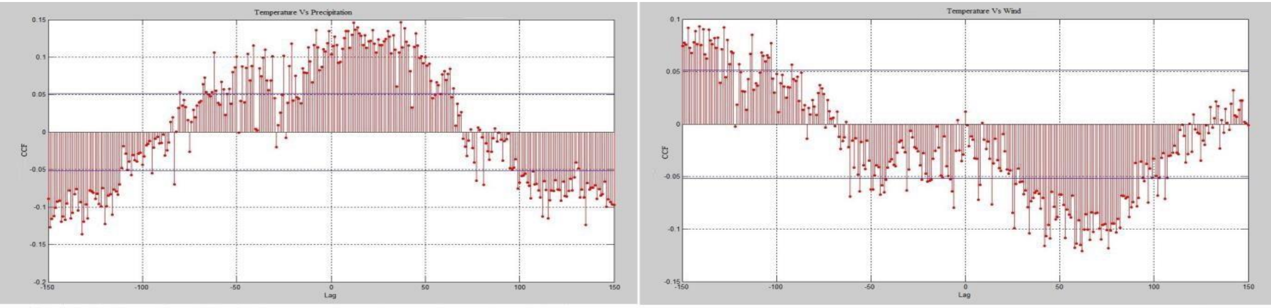
Output Variable		**Input Variables**
Temperature	Maximum	✓ Precipitation, Visibility, and wind.
	Minimum	✓ Air pressure, Dew point, Humidity, and Precipitation,
Precipitation	Maximum	✓ Temperature and Wind.
	Minimum	✓ Temperature, Wind, and Visibility.
Wind	Maximum	✓ Air pressure, Precipitation, and Temperature.
	Minimum	✓ Precipitation and Visibility.
Dew point	Maximum	✓ Visibility, Humidity, and Visibility.
	Minimum	✓ Humidity, Temperature, and Air pressure.
Humidity	Maximum	✓ Air pressure, Dew point, and Visibility.
	Minimum	✓ Air pressure, Dew point, and Temperature.
Visibility	Maximum	✓ Dew point, Humidity, and Temperature.
	Minimum	✓ Precipitation and Wind.
Air pressure	Maximum	✓ Humidity and Wind.
	Minimum	✓ Temperature, and Dew point.

3.3.1. Cross Correlation among the selected input and output variables

The above matrix correlation among the variables was performed for feature extraction as well as feature selection without lead/lag information. Due to this reason, the previous correlation does not capture the prediction capability of the predictor perfectly. Therefore, in this section, to select appropriate input variables, we implement cross correlation among the input and output variables by introducing lead/lag information. Besides, we use and implement 150 numbers of lags/leads (i.e. from -150 up to 150) and the default standard deviation (STD) (i.e. STD = 2) which corresponds to approximate 95% of confidence bounds.

3.3.2. Cross correlation between Temperature and Input Variables

From figure 3.10 – 3.13 indicates the cross correlation among maximum and minimum temperature and the previous (i.e. from matrix correlation) selected input variables.



(a) (b)

Fig 3. 10 (a) and (b) Cross-correlation between maximum temperature and precipitation and between maximum temperature and wind respectively

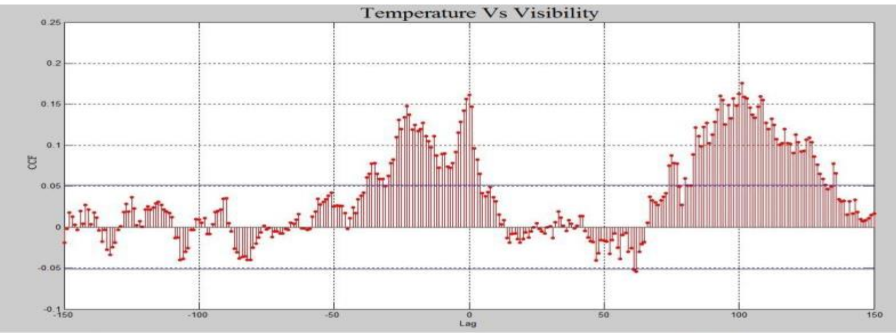
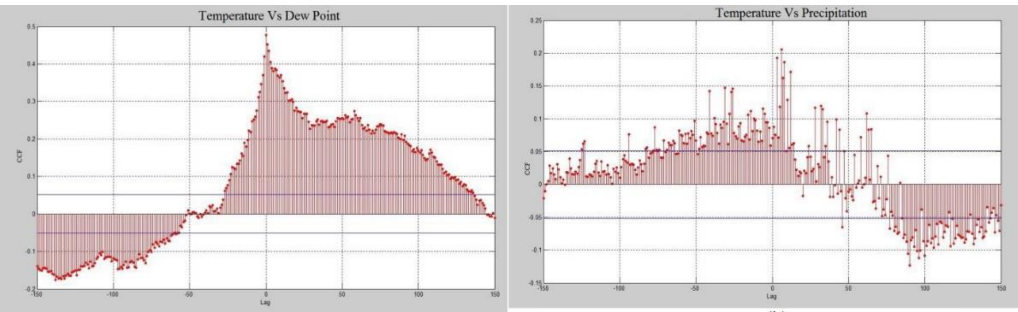


Fig 3. 11 Cross-correlation between maximum temperature and visibility



(a) (b)

Fig 3. 12 (a) and (b) Cross-correlation between minimum temperature and dew point and between minimum temperature and precipitation respectively

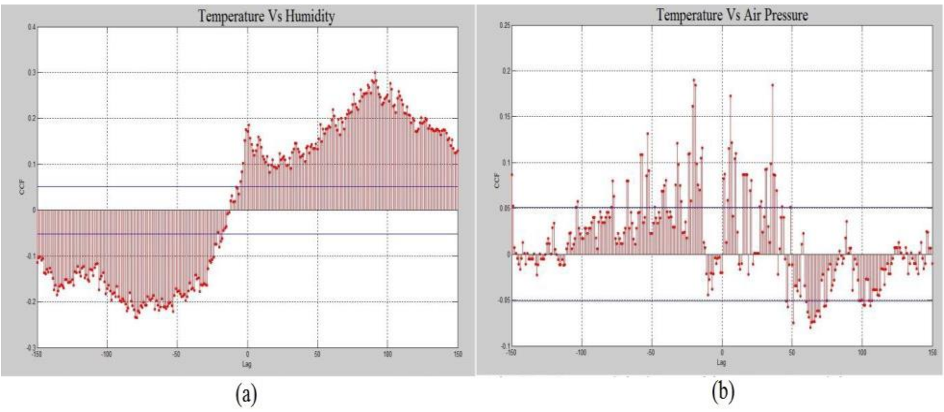


Fig 3. 13 (a) and (b) Cross-correlation between minimum temperature and humidity and between minimum temperature and air pressure respectively

As we have seen in the above figures, the cross-correlation graph among the minimum and maximum temperature and selected input variables except minimum temperature and dew point shows significant correlation at greater than zero lags. But, the figure 3.12 (a) indicates minimum dew point traces the temperature very closely. Therefore, the dew point information cannot predict the future values of the temperature.

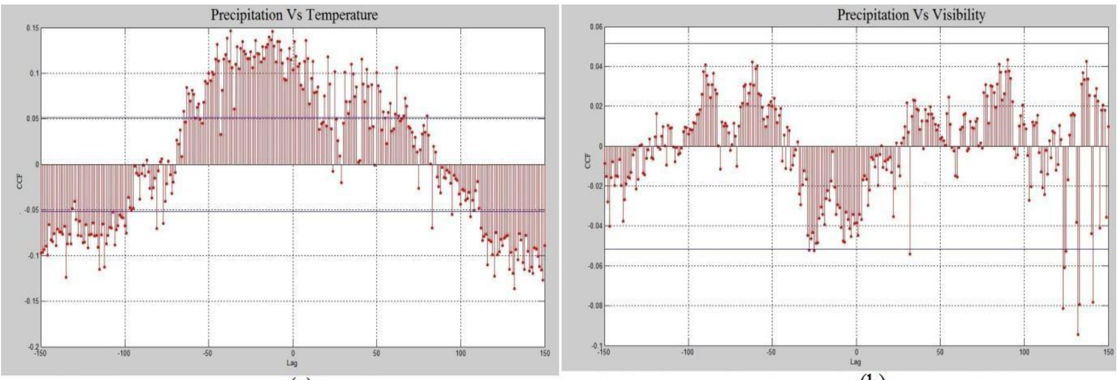
Therefore, based on the cross-correlation results achieve, we selected the input variables for maximum and minimum temperature as follow:

Table 2 Selected input variables to predict minimum and maximum temperature based on cross-correlation

Output Variable		**Input Variables**
Temperature	Maximum	✓ Precipitation, Visibility, Wind, and Gaussian noise
	Minimum	✓ Air pressure, Humidity, Precipitation, and Gaussian noise

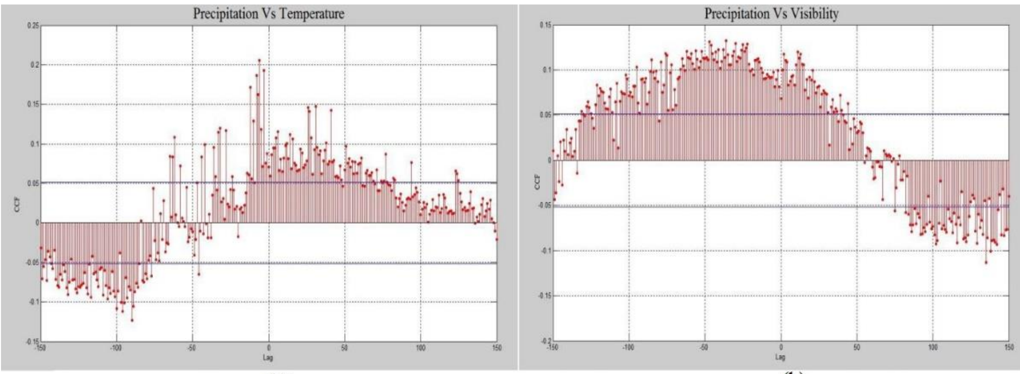
3.3.3. Cross correlation between Precipitation and Selected Input Variables

From figure 3.14 – 3.16 indicates the cross-correlation among maximum and minimum precipitation and the previous selected input variables.



(a) (b)

Fig 3. 14 (a) and (b) Cross-correlation between maximum precipitation and temperature and between maximum precipitation and humidity respectively



(a) (b)

Fig 3. 15 (a) and (b) Cross-correlation between minimum precipitation and temperature and between minimum precipitation and visibility respectively

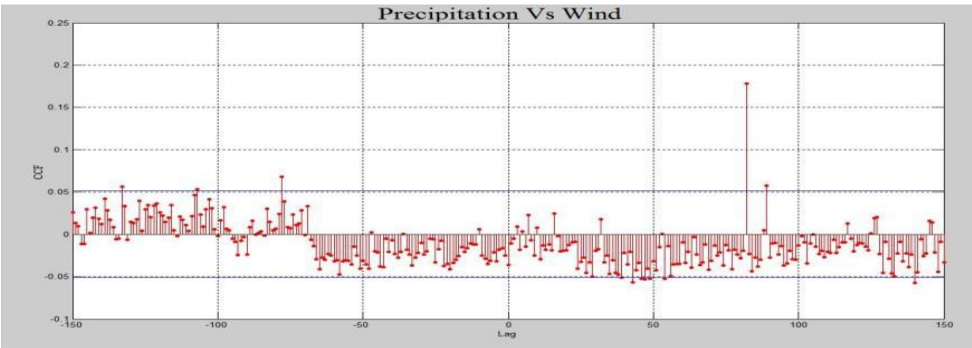


Fig 3. 16 Cross-correlation between minimum precipitation and wind

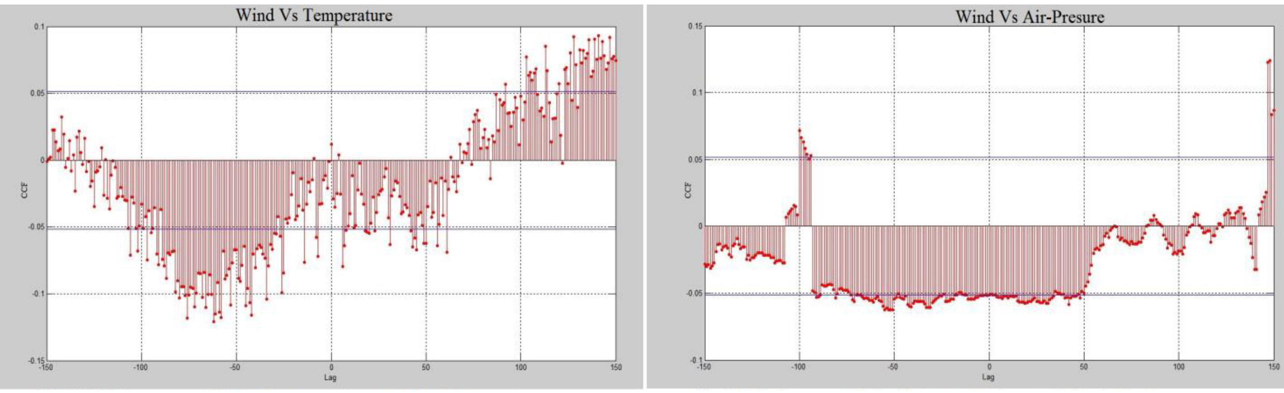
The above figures (figure 14, 15, and 16) indicate cross-correlation result among the minimum and maximum precipitation and the selected input variables. And they show significant correlation at greater than zero lags. Therefore, based on the cross-correlation result, we selected the input variables which are used to predict the future of maximum and minimum precipitation in the table below.

Table 3 Selected input variables to predict minimum and maximum Precipitation based on cross-correlation

Output Variable		**Input Variables**
Precipitation	Maximum	✓ Temperature, Wind, and Gaussian noise
	Minimum	✓ Temperature, Wind, Visibility, and Gaussian noise

3.3.4. Cross-correlation between Wind and Selected Input Variables

From figure 3.17 – 3.19 indicates the cross correlation among maximum and minimum wind and the selected input variables.



(a) (b)

Fig 3. 17 (a) and (b) Cross-correlation between maximum wind and temperature and between maximum wind and air-pressure respectively

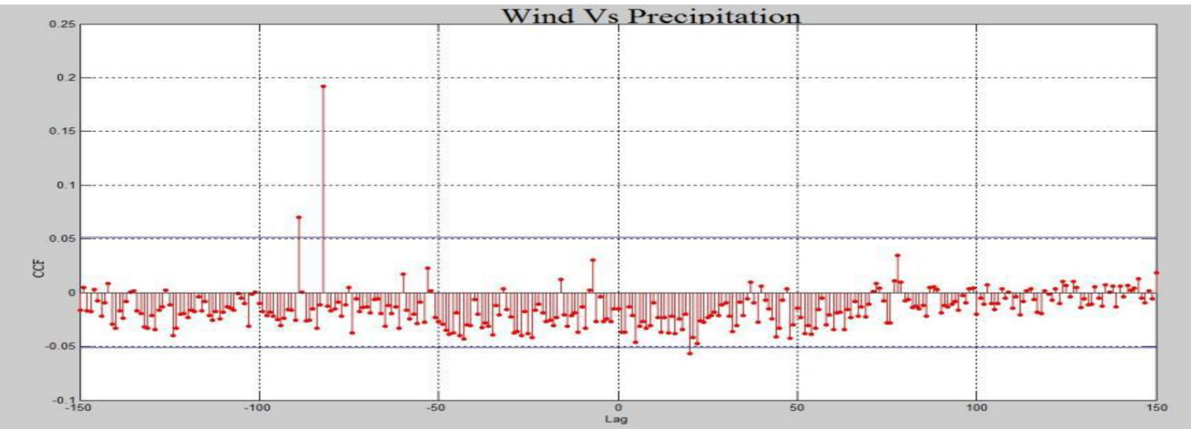
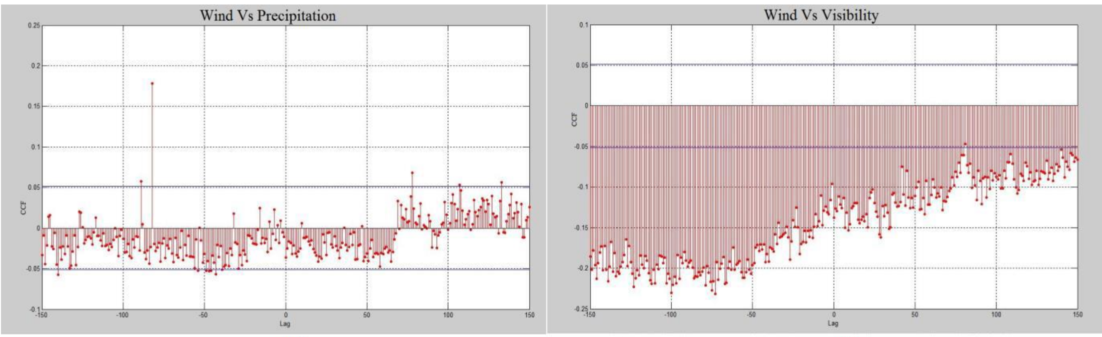


Fig 3. 18 Cross-correlation between maximum wind and precipitation



(a) (b)

Fig 3. 19 (a) and (b) Cross-correlation between minimum wind and precipitation and between minimum wind and air-visibility respectively

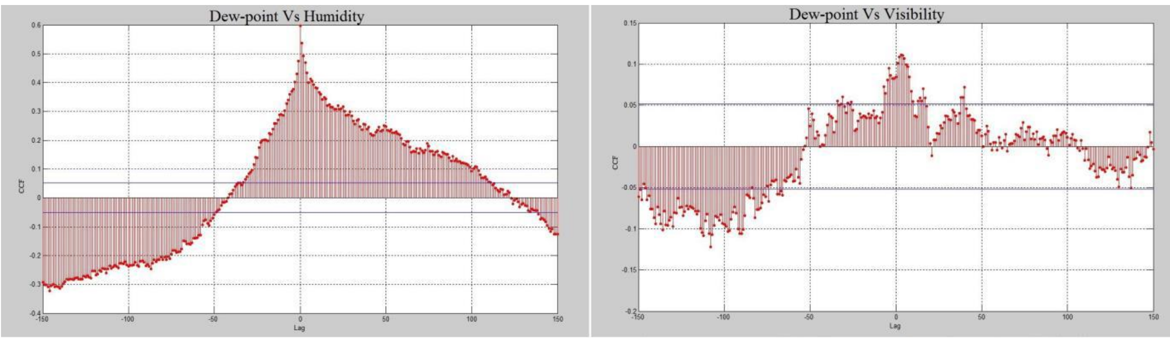
As we have seen the result of cross-correlation between maximum and minimum wind and the previous selected input variables graph, all have significant correlation at greater than zero lags. However, figure 3.19 (b) indicates, minimum wind and visibility have a pure negative correlation. Therefore, based on the cross-correlation result, we selected the input variables for maximum and minimum wind as follow:

Table 4 Selected input variables to predict minimum and maximum wind based on cross-correlation

Output Variable		**Input Variables**
Wind	Maximum	✓ Air pressure, Precipitation, Temperature, and Gaussian noise
	Minimum	✓ Precipitation, Visibility, and Gaussian noise

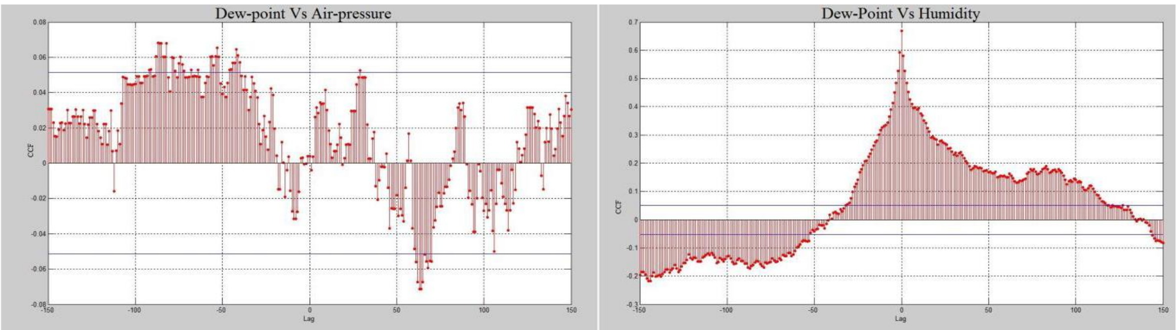
3.3.5. Cross-relation between Dew-point and Selected Input Variables

From figure 3.20 – 3.22 indicates the cross correlation among maximum and minimum Dew-point and the previous selected input variables.



(a) (b)

Fig 3. 20 (a) and (b) Cross-correlation between maximum dew-point and humidity and between maximum dew-point and visibility respectively



(a) (b)

Fig 3. 21 (a) and (b) Cross-correlation between minimum dew-point and air-pressure between minimum dew-point and humidity respectively

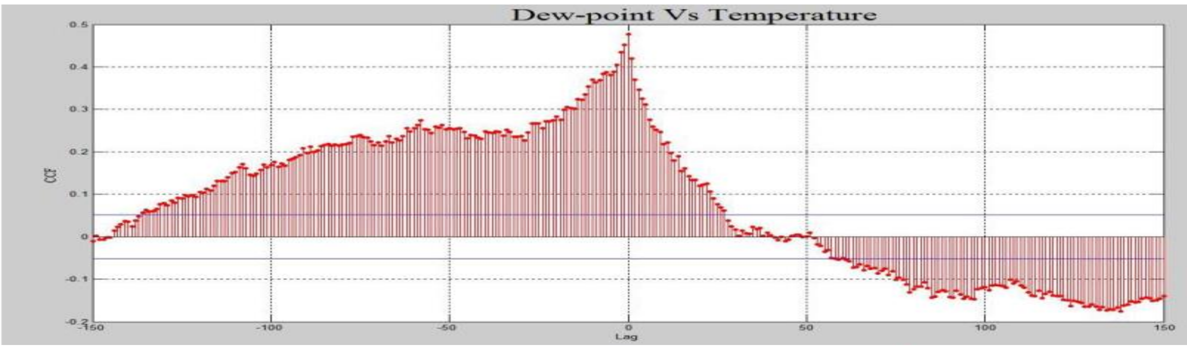


Fig 3. 22 Cross-correlation between minimum dew-point and air-pressure

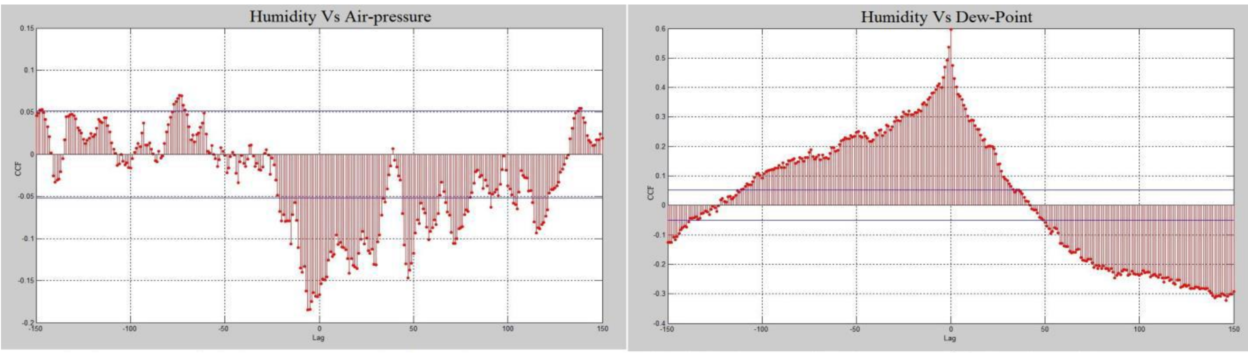
As we have seen the cross-correlation result in figure 3.20 (a), 3.21 (b), and 3.22, maximum dew-point and humidity, minimum dew-point and humidity, minimum dew-point and temperature respectively traces the dew-point very closely. Thus, the maximum and minimum humidity's and minimum temperature are not good for predicting the future of the dew-point. Therefore, based on the cross-correlation among dew-point and the selected input variables result, we put the predictor variables in the table below.

Table 5: Selected input variables to predict minimum and maximum dew-point based on cross-correlation

Output Variable		**Input Variables**
Dew-point	Maximum	✓ Visibility, Visibility, and Gaussian noise
	Minimum	✓ Air pressure, and Gaussian noise

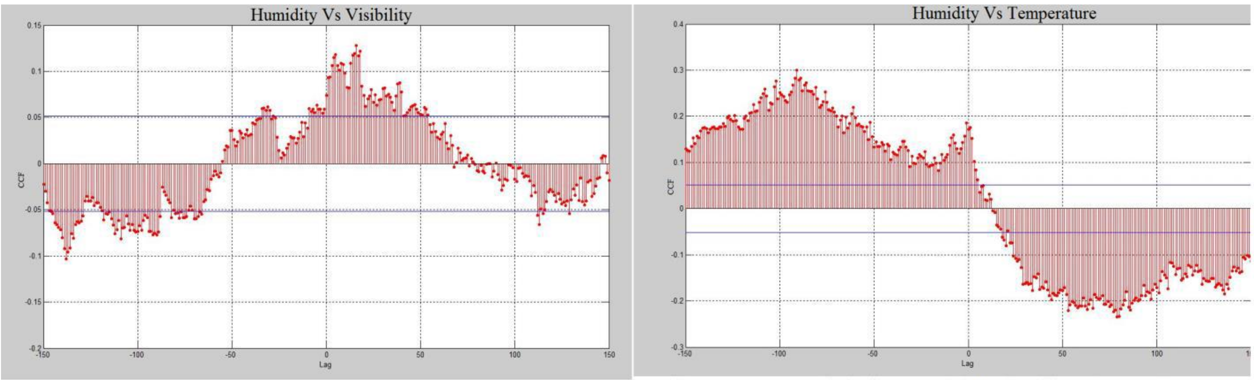
3.3.6. Cross-correlation between Humidity and the selected input variables

From figure 3.23 – 3.25 indicates the cross correlation among maximum and minimum humidity and the previous selected input variables.



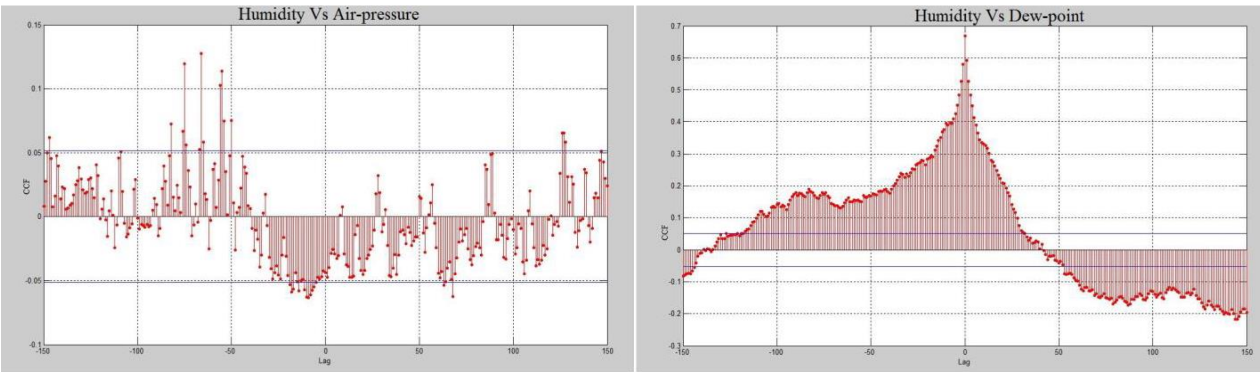
(a) (b)

Fig 3. 23 (a) and (b) Cross-correlation between maximum humidity and air-pressure and between maximum humidity and dew-point respectively



(a) (b)

Fig 3. 24 (a) and (b) Cross-correlation between maximum humidity and visibility and between maximum humidity and temperature respectively



(a) (b)

Fig 3. 25 (a) and (b) Cross-correlation between minimum humidity and air-pressure and between minimum humidity and dew-point respectively

The figure 3.23 (b) and 3.25 (b) indicates that the cross-correlation graph between the maximum and minimum humidity and dew-point respectively. As we have seen from the figures, humidity and dew-point shares very high correlation at zero lag. This indicates that, dew-point information traces very closely to the humidity. Due to this reason, dew-point information cannot predict the future value of humidity.

Therefore, based on the result of cross-correlation of selected input variables and the humidity, I put the appropriate input variables which are used to predict humidity in the next table.

Table 6 Selected input variables to predict minimum and maximum humidity based on cross-correlation

Output Variable		**Input Variables**
Humidity	Maximum	✓ Air pressure, Visibility, and Gaussian noise
	Minimum	✓ Air pressure, Temperature, and Gaussian noise

3.3.7. Cross-correlation between Visibility and the Selected Input Variables

From figure 3.26 – 3.28 shows the cross-correlation among maximum and minimum Visibility and the selected input variables.

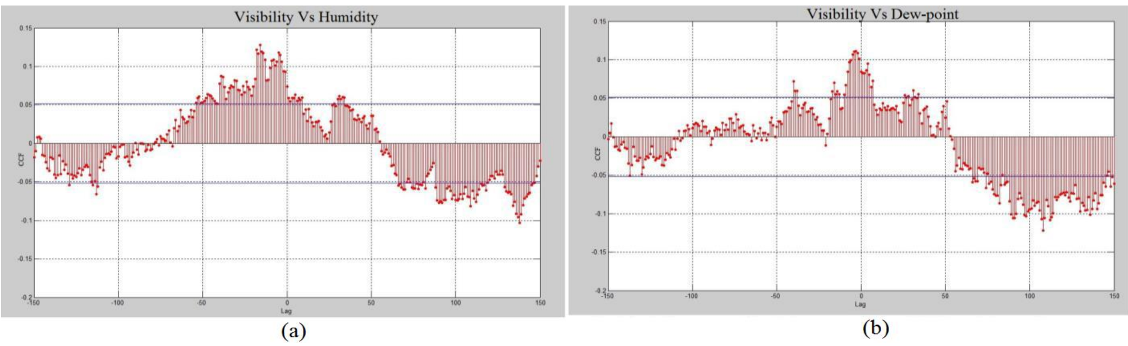


Fig 3. 26 (a) and (b) Cross-correlation between maximum visibility and humidity between maximum visibility and dew-point respectively

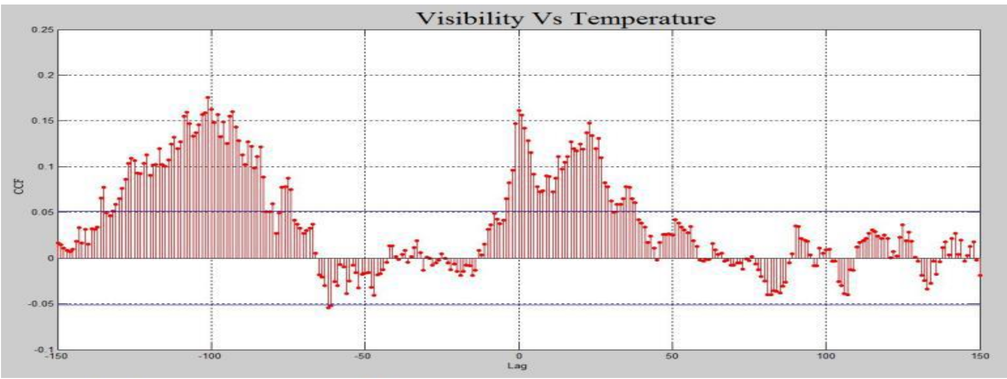
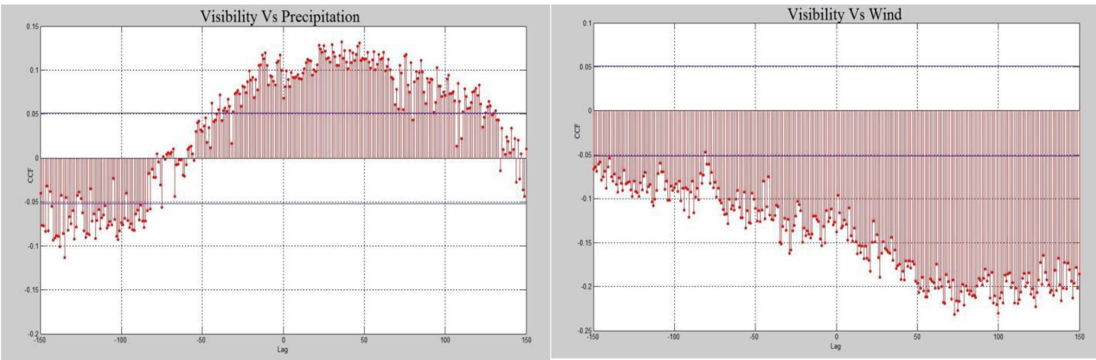


Fig 3. 27 Cross-correlation between maximum visibility and temperature



(a) (b)

Fig 3. 28 (a) and (b) Cross-correlation between minimum visibility and precipitation between minimum visibility and wind respectively

The above figures (figure 26, 27, and 28) indicate the cross-correlation graph among the visibility and the input variables. Since all input variables information does not trace very closely the visibility information, we selected all previous selected input variables and we put them in the table below. Besides, as we have seen in figure 3.28 (b), the visibility and wind have pure negative correlation.

Table 7 Selected input variables to predict minimum and maximum visibility based on cross-correlation

Output Variable		**Input Variables**
Visibility	Maximum	✓ Dew point, Humidity, Temperature, and Gaussian noise
	Minimum	✓ Precipitation, Wind, and Gaussian noise

3.4. Summary

To conclude this chapter, the proposed system generally categorizes into preprocessing phase which mainly focuses on selecting and gathering weather parameters, handling missing values, segmentation, and normalizes the dataset, design and train phase which provides designing of the proposed network and train the network, and predicting phase which focuses on determining of the prediction of weather condition. For handling the missing data numerical interpolation namely cubic spline and linear interpolations are provided. Besides, in order to know the strength among the selected input and output variables statistical matrix correlations such as Pearson, Kendall, and spearman are presented. However, the matrix correlations are not capable for fully forecast purpose. Then, to figure out this problem, cross correlation with lag/lead information among the predicting and predictor are presented.

Chapter Four

4. Implementing

4.1. Introduction

In this chapter we explain a brief over view of all the mathematical background of the algorithms including the overall flow chart of the proposed system. Furthermore, we also presented here all about the pseudo code of the algorithms.

4.2. Overall Flow chart of DBN and SVM networks

The figure 4.1 shows the overall flowchart of the proposed system and we summarize it as follows:

Step1: collect the weather dataset

Step 2: handle the missing data's. If the missing data has more than four contiguous values, handle it using linear interpolation method, otherwise use cubic spline interpolation method. However, if data is not missing, go to step three.

Step 3: if data is not missing, segment the datasets based on Ethiopian season and rainfall regimes.

Step 4: select input and output variables by applying correlation among the variables.

Step 5: Normalize the data dimensions

Step 6: feed the normalized data to the designed network which are DBN and SVM.

Step 7: calculate the error. If the output error is much greater than the expected error, return to step 6, otherwise, test and validate the system.

Step 8: finally, predict the atmospheric condition based on the forecasting ranges such as short range, medium range, and long range.

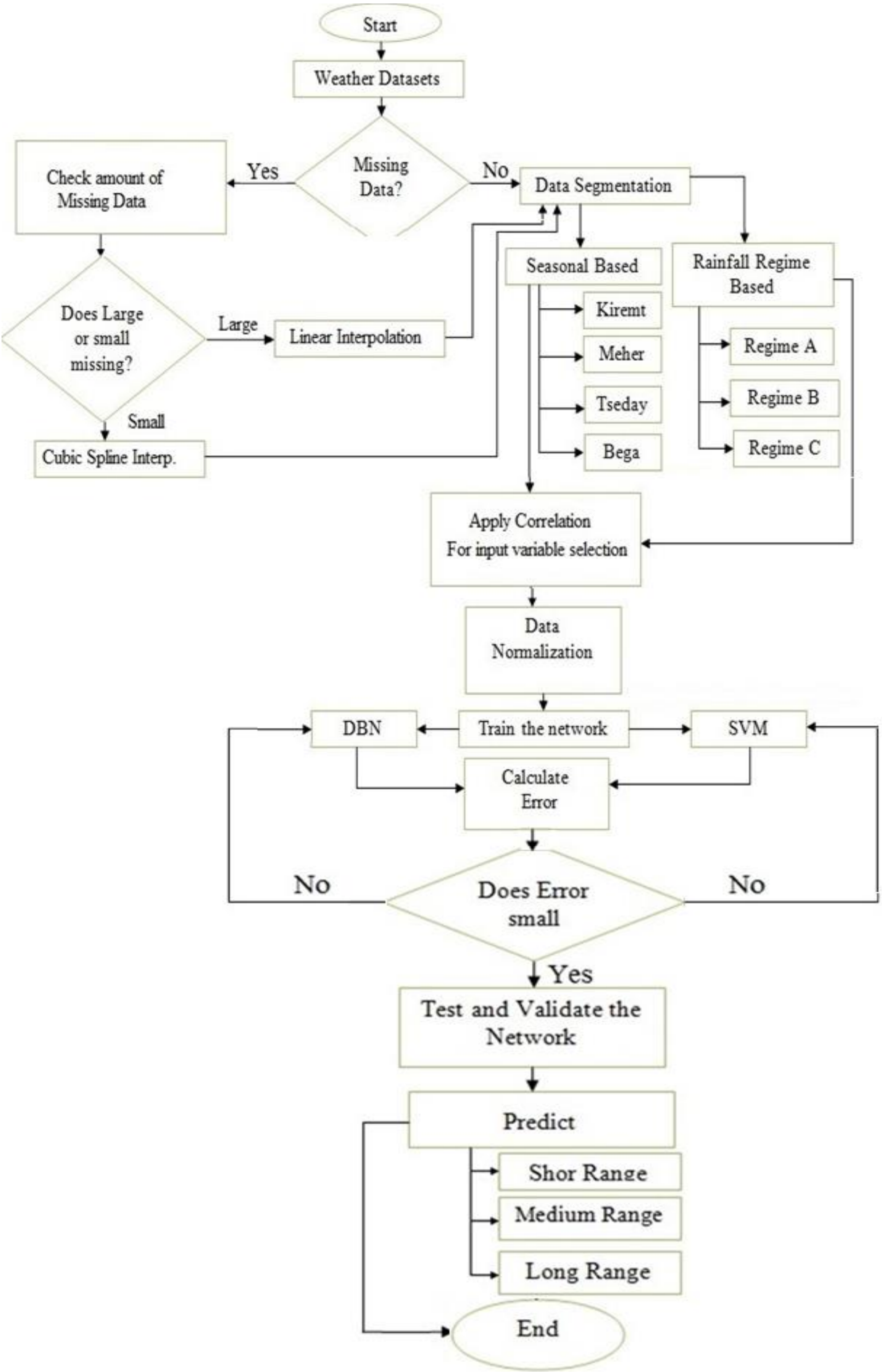


Fig 4. 1 over all flow chart of the proposed system

4.3. Deep Neural Network Algorithm

4.3.1. Restricted Boltzmann Machine (RBM)

RBM is a two-layer stochastic network such as visible layer **V** and hidden layer **h** [41] [42]. And also, there are only connections between input and hidden layers and no connections among units in the same layer; see figure 4.2.

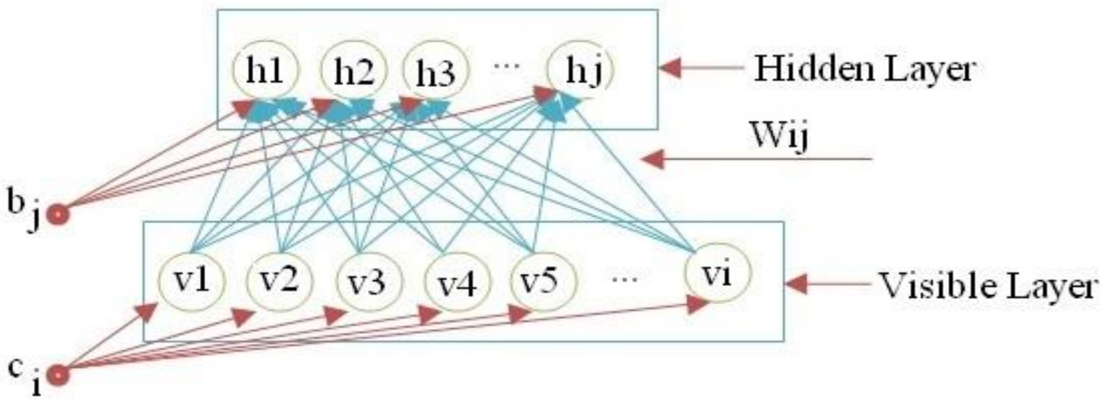


Fig 4. 2 The structure of the restricted Boltzmann machine.

Besides, it is an undirected bipartite graphical model with its probability distribution governed by an energy function E, parameterized by λ ={W; b; c} [43]:

E (v, h; ʎ) = - $\sum_{i\,\varepsilon\, visible} CiVi$ - $\sum_{j\,\varepsilon\, hidden} bjhj$ - $\sum_{i,j} Vihjwij$ ……………………… (3)

Where vi and hj are the binary state of visible unit i and hidden unit j respectively; ci and bj are their corresponding biases; wij is the weight between visible and hidden unit. The joint probability for configuration v; h is formulated as follows:

P(v, h; ʎ) = $\frac{\exp(-E(v,h;ʎ)}{Z}$ ………………………………………………………………………………… (4)

Where: Z is the partition function, which ensures the function a valid probability function. The conditional probability in this RBM is thus,

P(v_i\ h; ʎ) = sigm (C_i + $\sum_j wijhj$) ………………………………………………………………………… (5)

P(h_j\ v; ʎ) = sigm (b_i + $\sum_i wijvi$) ………………………………………………………………………… (6)

Where sigm is the sigmoid function; Parameter learning of this model is usually done by the ML learning. The log probability that the network assigns to the data is given by

$\text{Log } p(v; \Lambda) = \log\sum_{h} p(v, h; \Lambda)$ .... (7)

Taking the partial derivative of log p(v; λ) with respect to the parameters {W; c; b} yields the following update rule:

$W_{ij} = W_{ij} + \langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{model}$ ……………………..………………………… (8)

$c_i = c_i + \langle v_i \rangle_{data} - \langle v_i \rangle_{model}$ …………………………..…………………..……… (9)

$b_j = b_j + \langle h_j \rangle_{data} - \langle h_j \rangle_{model}$ ……………………………..……………………… (10)

Where<> denotes the expectation. <>data means empirical expectation, and <>model denotes expectation with respect to model distribution.

RBMs are usually trained by using the contrastive divergence (CD) learning procedure, which is introduced in [44]. To avoid the difficulty in computing the log likelihood gradient, the CD method approximately follows the gradient of a different function. CD has been applied effectively to various problems, using Gibbs sampling or hybrid Monte Carlo as the transition

4.1.3.2. Deep Belief Network

DBN is a multilayer, stochastic generative model that is created by training a stack of RBMs, each of which is trained by using the hidden activities of the previous RBM as its training data. Each time anew RBM is added to the stack, the new DBN has a better lower bound on the log probability of the data than the previous DBN [45].

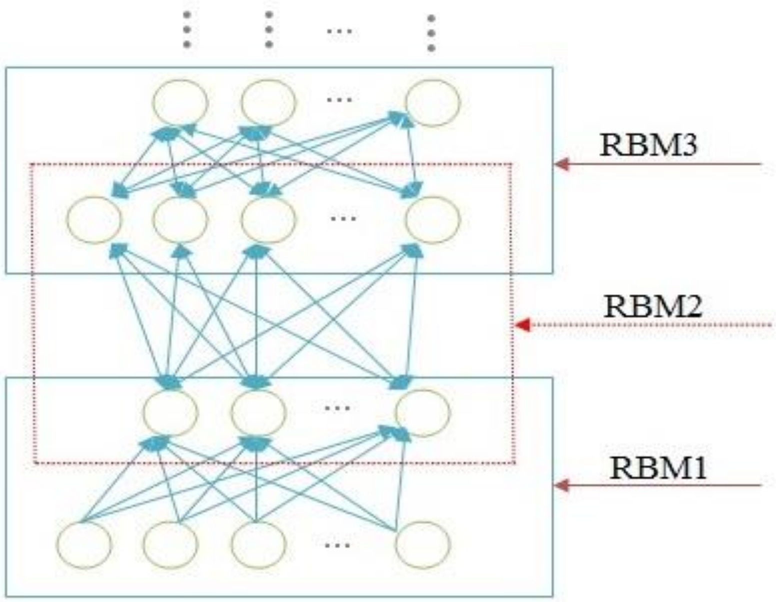


Fig 4. 3 Deep Belief Network (DBN)

Pseudo code of DBN

Training Dataset (X_t,Y_t), t ={1, 2, 3, …, n},

n = length of the input vector

% Initialization

Initialize M_Epoch, LR, E_Error, Visible and Hidden biases, Random Weight Matrix

% Start Positive Phase

%From Visible to Hidden one

E(V,H1) = Sum(W_{ij} * V_i);

P(H1=1/V) = Sigm(HB + E(V,H1));

%End V to H1

%From H1 to H2

E(H1,H2) = Sum(W^T_{ij} * V_i);

P(H2=1/H1) = Sigm(HB + E(H1,H2));

%End H1 to H2

…

While Error is not converged

…

% Start Negative Phase (From output returns to visible units by applying the same process with positive phase)

H1 ~ P(H1/H2)

V1 ~ P(V/H1)

…

%Update Weights and Biases

UpdtW_{ij} = W_{ij} + LR (+veP(V/H) - -veP(H/V)) ;

UpdtBV = BV + LR (V1 – V2)

UpdtBH1 = BH1+ LR (H1 – H2)

…

End while

In order to implement the DBN based weather forecasting, we uses two stacked restricted boltzmann machine. The algorithm most often used to train RBMs, that is to optimize the weight vector, is the CD algorithm as proposed by Hinton. The algorithm performs Gibbs sampling and is used inside a gradient descent procedure to compute weight update.

4.4 Support Vector Machine

SVMs are typically used for learning classification, regression, or ranking functions, for which they are called classifying SVM [46]. However, in this thesis concerned with regression. SVM Regression (SVR) is a method to estimate a function that maps from an input object to a real number based on training data. Similarly to the classifying SVM, SVR has the same properties of the margin maximization and kernel trick for nonlinear mapping.

Pseudo code of SVM for regression (SVR)

Training Dataset (X_t, Y_t), t ={1, 2, 3, …, n},

n = length of the input vector

Mapping into a non-linear:

$Y(X) = W^T * F(X) + B;$

% regularized Error function

Minimize the Euclidian norm

Min $1/2||W||^2$

Apply e-Sensitive error function

$L_e(Y) = 0$ if $|f(x) – Y| < e$ or $L_e(Y) = |f(x) – Y| - e$

If $L_e(Y) < e$

Break;

else

% adding slack Variables

$Y_n – e – sl <= Y_t <= Y_n + e + sl$

% Error Function for SVR

minimize: $C*Sum(a^+ + a^-) + ½ ||W||^2$

% Apply Lagrange multiplier and dual formulation

L: $½||W||^2 + C * Sum (a^+ + a^-) - Sum\alpha(e + a^+ – Y_i + <W,X_i + B) – Sum\alpha^*(e + a^- – Y_i - <W,X_i - B) – Sum(\beta a^+ + \beta^* a^-)$

% Prediction

$Y(X) = Sum((\beta - \beta^*) K(X,Xn) + B)$

% Apply Karush-Kuhn-Tucker(KKT)

determine B:

$B = Y_t - e - W^T * F(X)$

While the Error is not converged

Update B and W

Return to **prediction** step

End

End

In implementing of the SVM regression with e-insensitive loss function, Firstly, the input dataset is mapped into future space. Then in order to evaluating the kernel function, dot products are computed with the training dataset under the mapped values. Finally, the dot products are added up using the weights. And we will use the MATLAB optimization package for programming.

4.5 Overall flow chart of the existing forecasting system

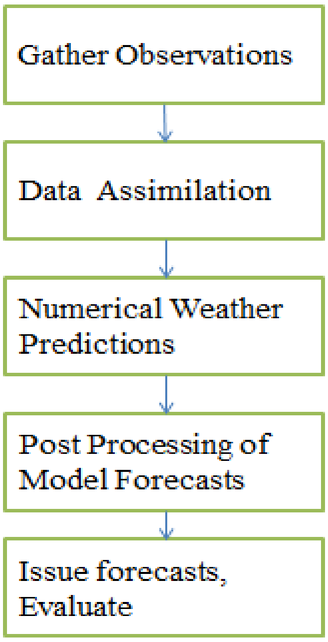


Fig 4. 4 Flow chart of numerical weather forecasting

The figure 4.4 shows the overall flowchart of the existing system and we summarize it as follows:

Step 1: collecting of the weather variables.

Step 2: Automated algorithms and manual intervention to detect, correct, and remove errors in observed data.

Step 3: interpolating the data numerically

Step 4: prediction of the weather events by applying numerical extrapolation

Step 5: Evaluating the forecasting events

Chapter Five

5. Results and Discussion

5.1. Introduction

In this section we have provided several empirical results of the predicted weather with respect to Ethiopian context based on the three experiment ranges namely short range, medium range, and long range using DBN and SVM algorithms in contrast to the existing forecasting system called numerical method. And also, in the experiments mainly concentrated on forecasting of maximum and minimum values of the selected weather variables with handling and missing the data that we have mentioned in chapter three. The three experiments are:

- ✓ Experiment one: It focuses on short range forecasting of the weather events using the DBN, SVM, and NWP algorithms based on the collected hourly data and it compares the performance of these algorithms based on time consumption and root mean square error.
- ✓ Experiment two: It focuses on medium range forecasting of the selected weather events based on the DBN, SVM, and NWP algorithms using a daily data-set and it evaluates and compares their performance by applying RMSE and the time they have taken during training the data.
- ✓ Experiment three: It focuses in the evaluating of the long range forecasting result of the selected weather events based on the DBN, SVM, and NWP algorithms using the average of one third of the daily data-set.

RMSE = $\mathrm{sqrt}((\sum_{i=1}^{n}(Yi - yi)\hat{}2)/n)$ … ..11

%RMSE = $\mathrm{sqrt}((\sum_{i=1}^{n}(Yi - yi)\hat{}2)/n) * 100\%$ … … … … … … … … … … … … … … … … ..12

Where, Yi = desired value, yi = actual value, and n = size of the data.

As in [47] describes, numerical weather prediction refers to forecasts that are obtained by using complex mathematical calculations carried out with high-speed computers; and its accuracy primarily depends on two factors. First, the more data that is available to a computer, the more accurate its result. Second, the faster the speed of the computer, the more calculations it can perform, and the more accurate its report will be.

The paper in [48] concludes, without the invention and subsequent improvement of computers, numerical weather prediction would still be in its infancy. Therefore, it is so difficult to achieve good performance of weather forecasting using numerical with the recent computers. In addition to that, as we have seen in the literature review section, a lot of researches ensure numerical weather prediction has low performance in contrast to machine learning based forecasting methods. Moreover, the drawbacks of numerical weather prediction are:

- ✓ It is parameterized
- ✓ It is difficult with convective process
- ✓ It has systematic errors.

The Statistical methods are usually associated with linear data [15]. However, the weather events are nonlinear by their nature. Therefore, it is tedious to forecast accurate future conditions using statistical methods. Besides, forecasting of an atmospheric condition using statistical methods has a number of drawbacks. For example, it requires thousands of equation to develop the forecasting, change in model requires developing a new equation, and others [49].

Therefore, SVM regression is selected instead of numerical and statistical methods and makes a comparison with the proposed algorithm called DBN. However, in order to ensure the aforementioned problems, we implemented numerical method for regression called polynomial regression and comparison of the result with the selected machine learning algorithms is done.

For implementing of the proposed and existing system, the four and half years of hourly and daily recorded dataset of NOAA is used. Besides, we uses seventy percent of the collected data for training purpose and thirty percent of the data-set for testing and validating purpose. Then, after the testing is done a graph is plotted between the actual output and the predicted output. Then, a comparison among the proposed and existing algorithms can be made.

All in all, eight towns and seven weather variables are selected as a sample to forecast atmospheric condition with respect to Ethiopian context, and more than 1000 empirical results are achieved from the experiment. However, as it is difficult to put and discuss the whole result we have achieved, we have provided the temperature and precipitation weather variables only to discuss and evaluate their result as a sample here.

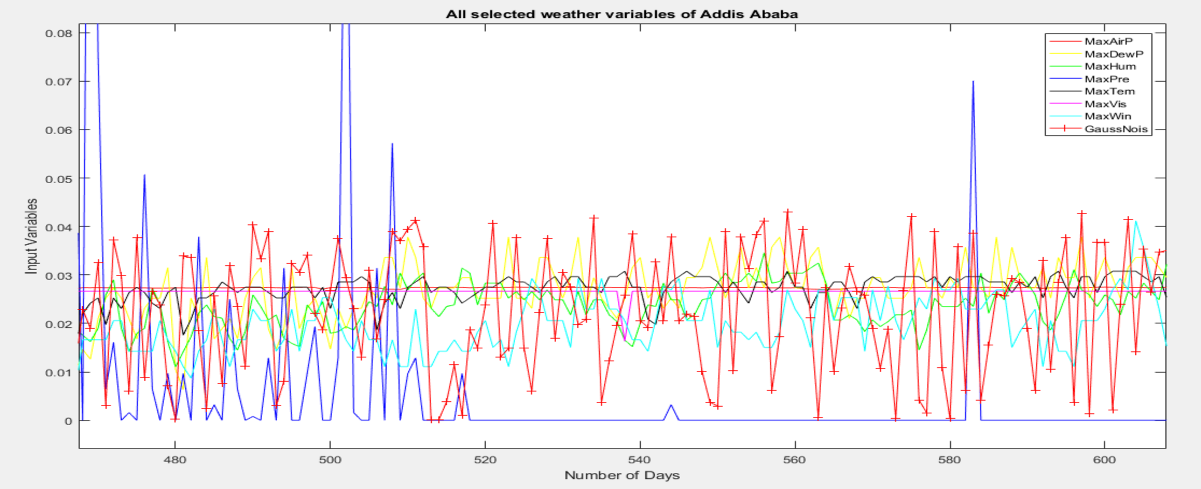


Fig 5. 1 Graph of the maximum value of the selected weather variables after normalization

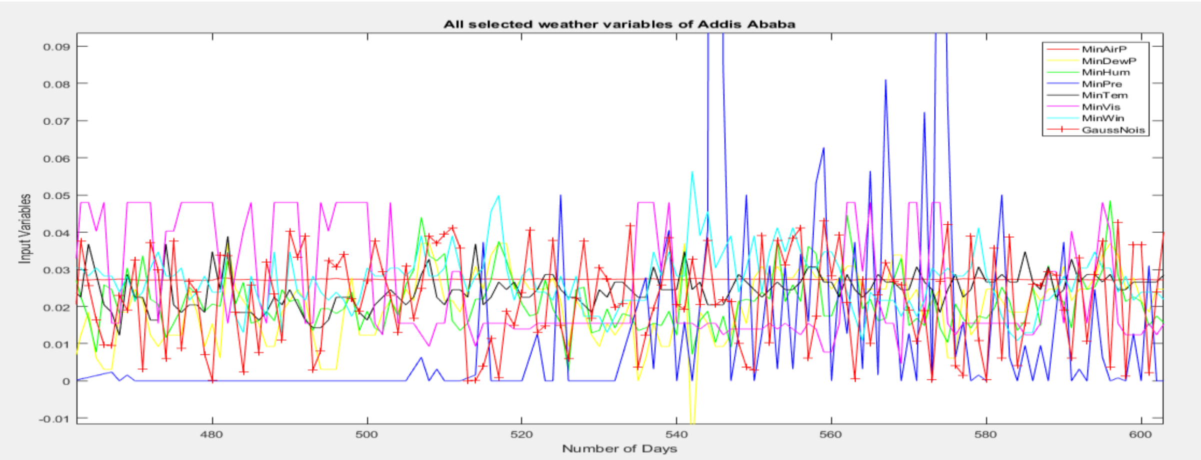


Fig 5. 2 Graph of the minimum value of the selected weather variables after being normalized

The figure 5.1 and figure 5.2 depicts a plot containing the maximum and minimum collected data such as temperature, dew point, humidity, wind, visibility, precipitation, and air pressure including the Gaussian noise over a normalized scale.

5.2. Result

In all the three forecasting algorithms namely DBN, SVM, and Numerical forecasting models, we will use the same training dataset. And also, for all the following tests provided in this section, we have utilized MATLAB 2016a on Intel(R) Core(TM) i5-3340M CPU @ 2.70GHz (4 CPUs), ~2.7GHz and 4GB RAM.

During applying SVM for regression algorithm for all tests, the training parameters we used are:

- ✓ E-insensitive tube as loss function with the value of e = 0.005,

- ✓ Additional capacity control C = 100,
- ✓ Kernel type = polynomial kernel function, and
- ✓ Quad optimization toolbox of MATLAB.

During forecasting using DBN also, for all experimental tests we use three consecutive RBM networks with one input layer, one output layer which has one neuron, and two hidden layers which has ten and six neurons respectively. As well, we used learning rate (Lr = 0.78), number of epochs (Ep = 1000), randomly generated weight, and initially zeros of biases for all layers.

Experiment 1 – Short range

The figure 5.3 shows the short range temperature prediction graph using SVM; the dotted blue graph indicates the upper and lower graph of the e-insensitive tube called lose function of the SVM for regression. The red and green graph that we have depicted in the graph is the real data and predicted data respectively. Besides, the y-axis is normalized hourly temperature values and the x-axis is the number of days which are selected randomly as sample.

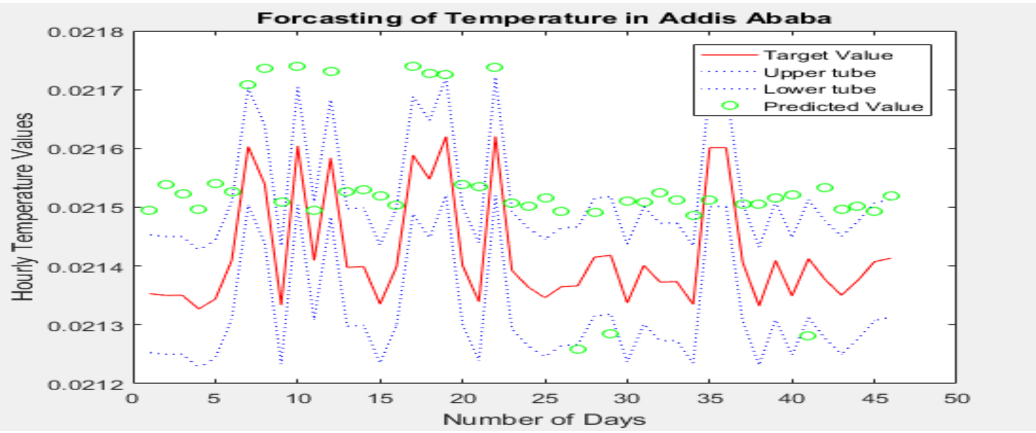


Fig 5. 3 Short range forecasting of temperature using SVM

As in the paper [50] describes, the main goal of SVM for regression is to find the estimated value with in the e-insensitive tube. However, it will not accept any deviation larger than the tube. Therefore, all the results are done based on this principle.

In the result of the short range temperature forecasting, shown in figure 5.3, the number of support vectors which are the predicted values out of the e-insensitive tube are 260 (i.e. 13.63% of the number of the trained data). Besides, based on the percentage of RMSE, the short range temperature forecasting is 79.6% accurate. Therefore, the result indicates the predicted values are approximately closely traced the actual values.

The figure 5.4 indicates the short range precipitation prediction using SVM. The dotted blue graph shows the upper and lower e-insensitive tube, the red and green graph indicates target data and predicted values respectively. Besides, the y-axis is normalized hourly precipitation values and the x-axis is the number of days which are selected randomly as sample.

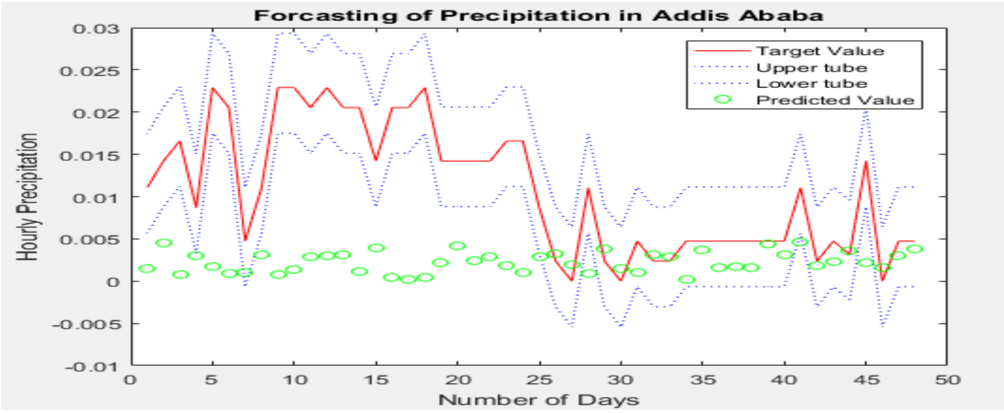


Fig 5. 4 Short range forecasting of Precipitation using SVM

As the result graph (figure 5.4) indicates, we have 685 number of support vectors (i.e. 58.75% of the number of the trained data); and also it is 87.3% accurate. The result has good accuracy; however, above 50% of the predicted values are support vectors. Therefore, forecasting of precipitation is not as good as forecasting temperature with SVM.

The figure 5.5 (a) and (b) shows the graph of short range forecasted temperature and precipitation respectively using DBN. The red and green graph indicates the actual and predicted value respectively. Besides, the y-axis is normalized hourly temperature and precipitation values (5.5 a and b respectively) and the x-axis is the number of days which are selected randomly as sample.

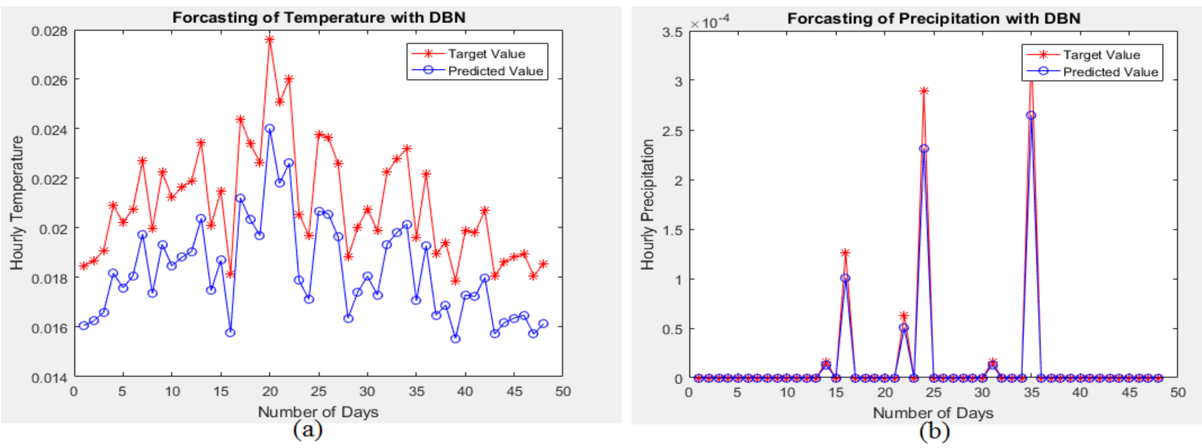


Fig 5. 5 (a) Short range of temperature forecasting with DBN (b) Short range of precipitation forecasting with DBN

As the graph 5.5 (a) and (b) indicates the trained value closely traces to the actual value. In addition to that, based on the percentage of root mean square error, they have 88.6% and 87.47% accuracy respectively. Therefore, forecasting of temperature and precipitation with large dataset using DBN gives us good fit.

The figure 5.6 and 5.7 shows the graph of short range forecasted temperature and precipitation of the forecasted precipitation respectively using numerical with polynomial regression. Besides, the y-axis is normalized hourly temperature and precipitation values and the x-axis is the number of days which are selected randomly as sample.

The result graph indicates that, using the proposed weather variables only, numerical forecasting is not appropriate to predict the future atmospheric condition. For example, forecasting of temperature and precipitation numerically has 52.5% and 78% accuracy respectively.

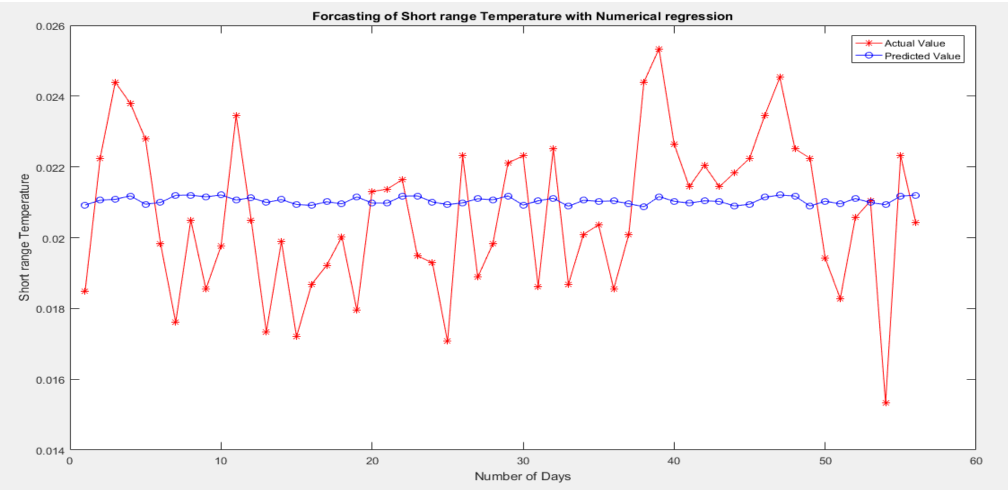


Fig 5. 6 (a) Short range temperature forecasting with numerical method

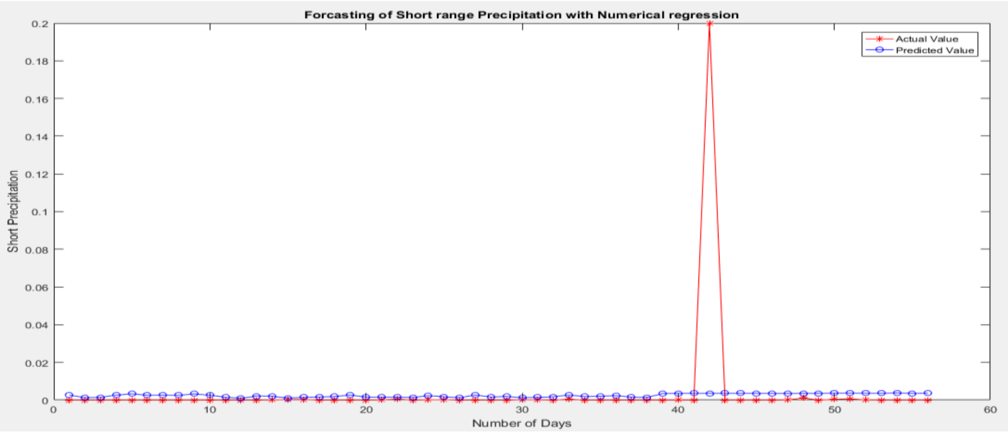
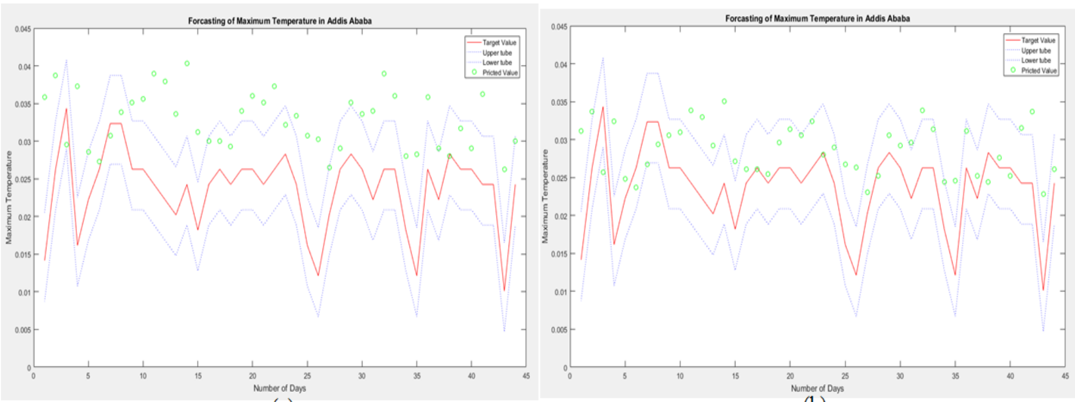


Fig 5. 7 Short range precipitation forecasting with numerical method

Experiment 2 – Medium range

The figure 5.8 (a) and (b) shows the medium range forecasting of maximum temperature before and after handling the missing values respectively using SVM algorithm. And also, the red, green, and blue indicates the target, predicted values, and upper and lower e-insensitive tube respectively. Besides, the y-axis is normalized daily temperature values and the x-axis is the number of days which are selected randomly as sample.

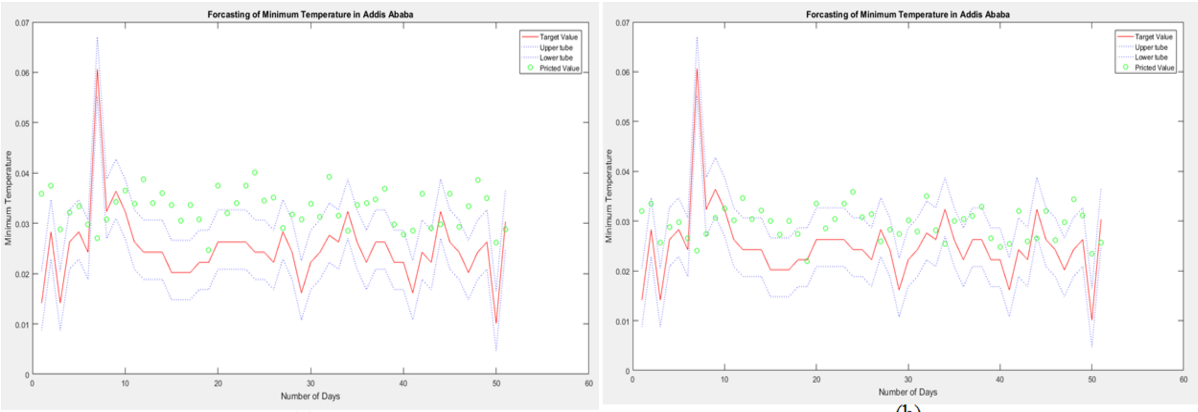
As we will see in the next figures (5.8, 5.9, and 5.10), before handling the missing values, more than 87% of the trained values are become a support vector. However, in order to reduce the number of support vectors, we enlarged the values of epsilon (e) from 0.001 to 0.015. Besides, based on the percentage of the RMSE, we have achieved greater than 79.6% accuracy in average. Furthermore, the performance and the train time consumption is summarized in table 9.



(a) (b)

Fig 5. 8 (a) Maximum temperature with missing data (b) Maximum temperature after handling the missing data

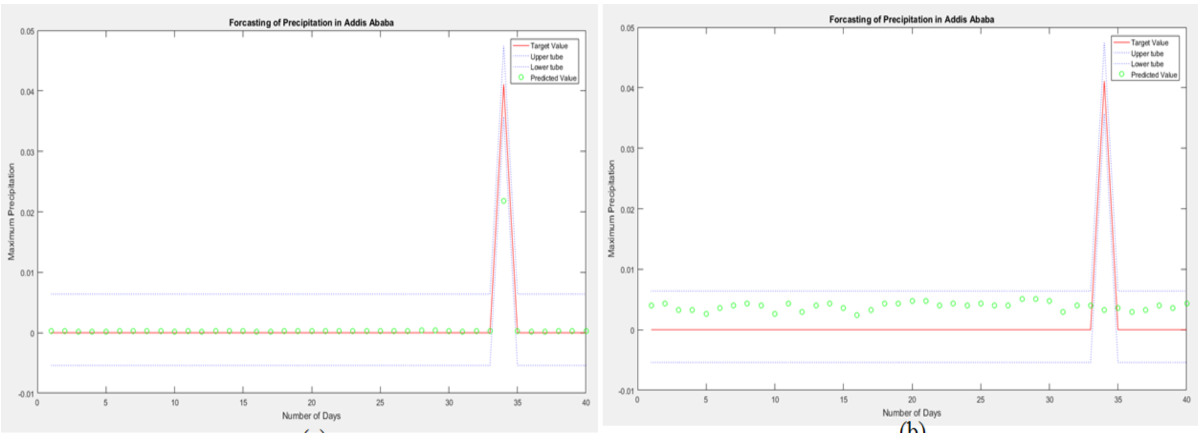
The figure 5.9 (a) and (b) shows the medium range minimum temperature prediction graph before and after handling the missing data using SVM; the dotted blue graph indicates the upper and lower graph of the e-insensitive tube. The red and green graph that we have plotted in the below graph are the actual and predicted value respectively.



(a) (b)

Fig 5. 9 (a) Minimum temperature with missing data (b) Minimum temperature after handling the missing data

The figure 5.10 (a) and (b) shows the prediction of maximum precipitation before and after handling the missing data using SVM. And also, the red, green, and blue indicates the actual, predicted values, and upper and lower e-insensitive tube respectively. Besides, the y-axis is normalized daily precipitation values and the x-axis is the number of days which are selected randomly as sample.



(a) (b)

Fig 5. 10 (a) Maximum Precipitation with missing data (b) Maximum Precipitation after handling the missing data

The figure 5.11 (a) and (b) shows the prediction of maximum precipitation before and after handling the missing data using SVM. Besides, the red, green, and blue indicates the actual, predicted values, and upper and lower e-insensitive tube respectively. Besides, the y-axis is normalized daily precipitation values and the x-axis is the number of days which are selected randomly as sample.

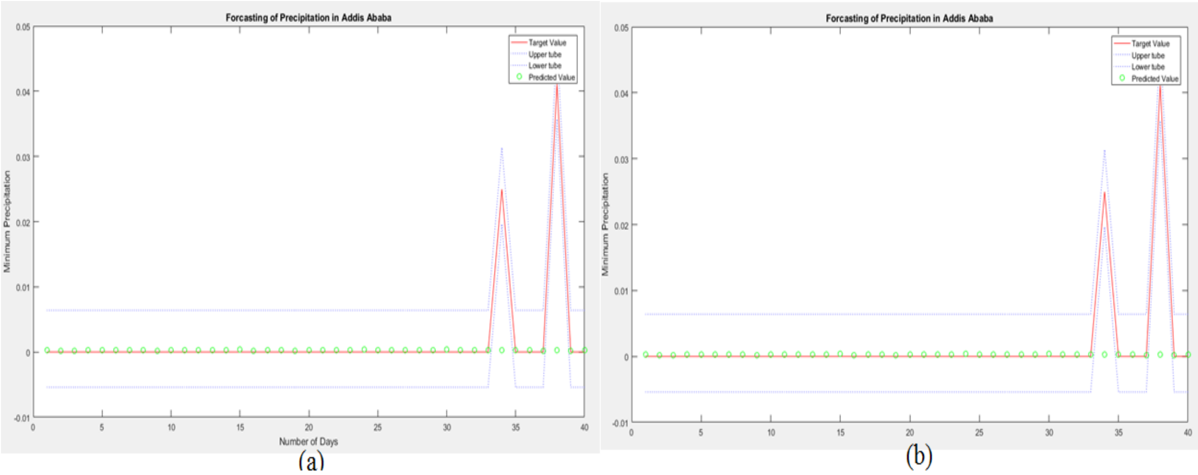


Fig 5. 11 (a) Minimum Precipitation with missing data (b) Minimum Precipitation after handling the missing data

The figure 5.12 (a) and (b) shows the forecasting of minimum temperature after and before handling the missing values with DBN respectively. The actual data and predicted values are plotted using red and blue color respectively.

As we will see in figure 5.12, 5.13, and 5.14, the predicted data closely traced to the actual value. But, in contrast to before and after handling the missing values, forecasting after handling the missing contiguous values are more accurate than forecasting without handling the missing values. For example, the DBN forecasting method has greater than 83.76% and 87.82% in average accuracy before and after handling the missing values.

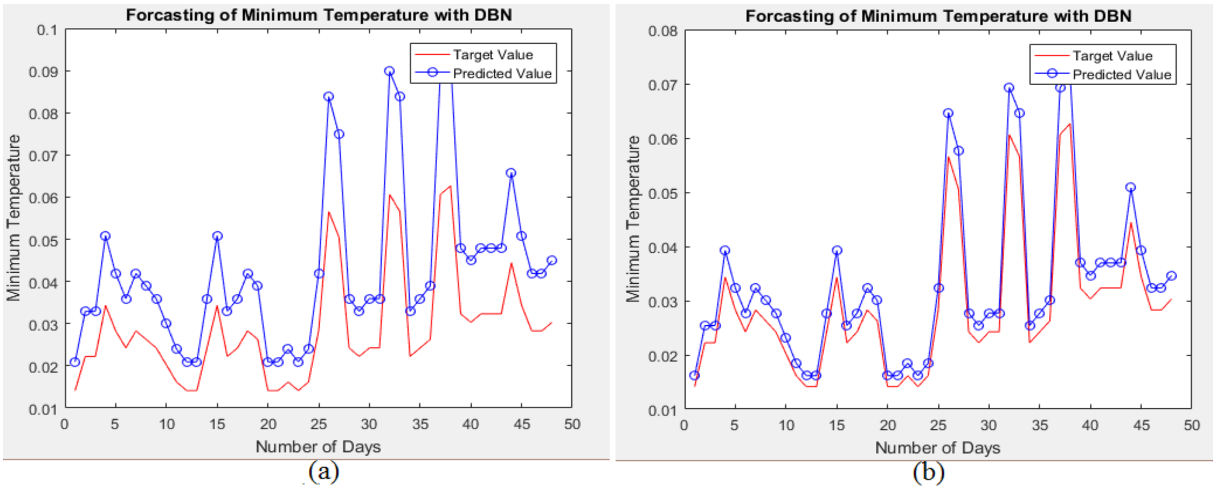


Fig 5. 12 (a) Minimum temperature after handling the missing data (b) Minimum temperature before handling the missing data

The figure 5.13 (a) and (b) shows, prediction of maximum temperature after and before handling the missing data using DBN respectively. The red and green colors indicate the actual data and predicted values respectively. Additionally, figure 5.24 (a) and (b) shows the root mean square error graph of figure 5.23 (a) and (b) respectively.

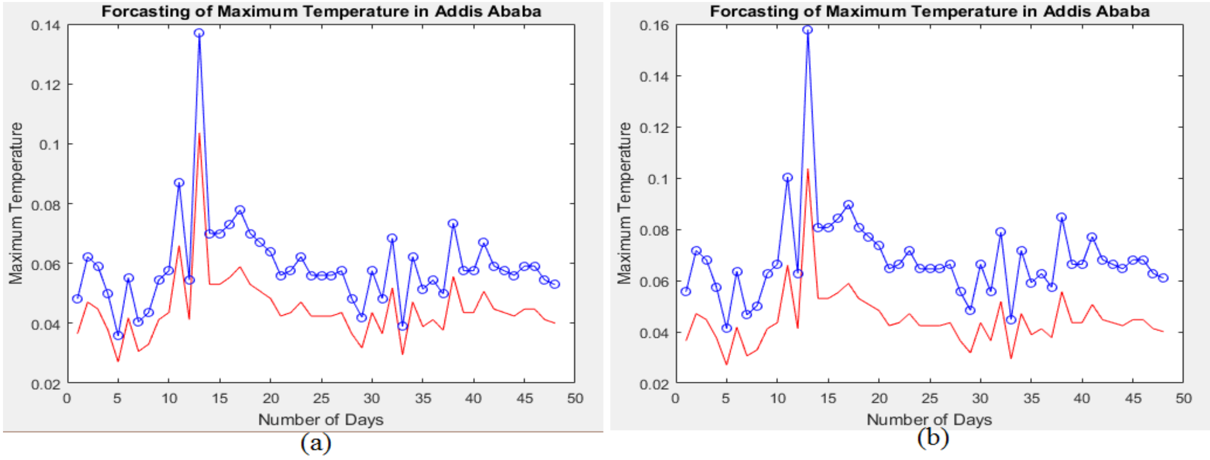


Fig 5. 13 (a) Maximum temperature after handling the missing data (b) Maximum temperature before handling the missing data

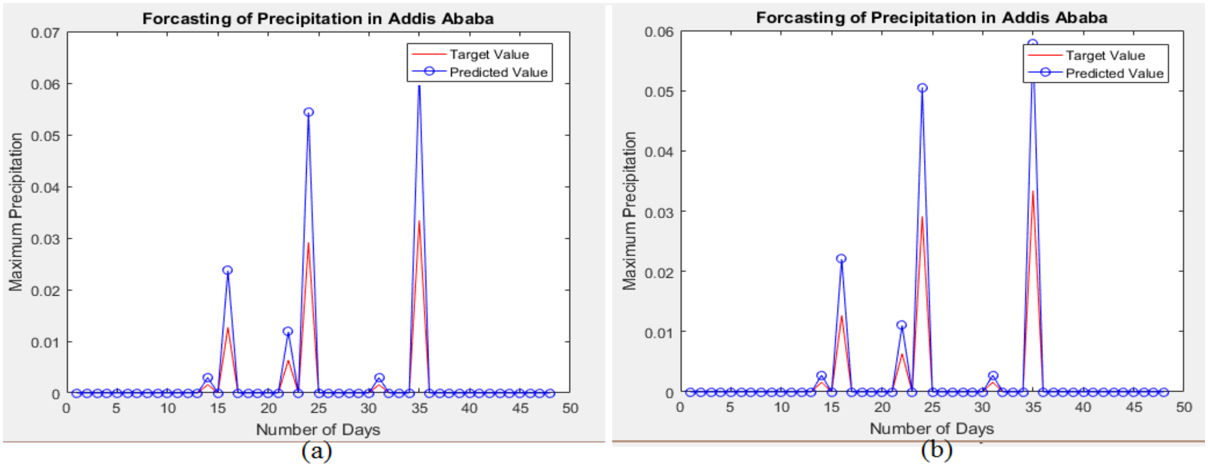


Fig 5. 14 (a) Maximum precipitation before handling the missing data (b) Minimum precipitation before handling the missing data

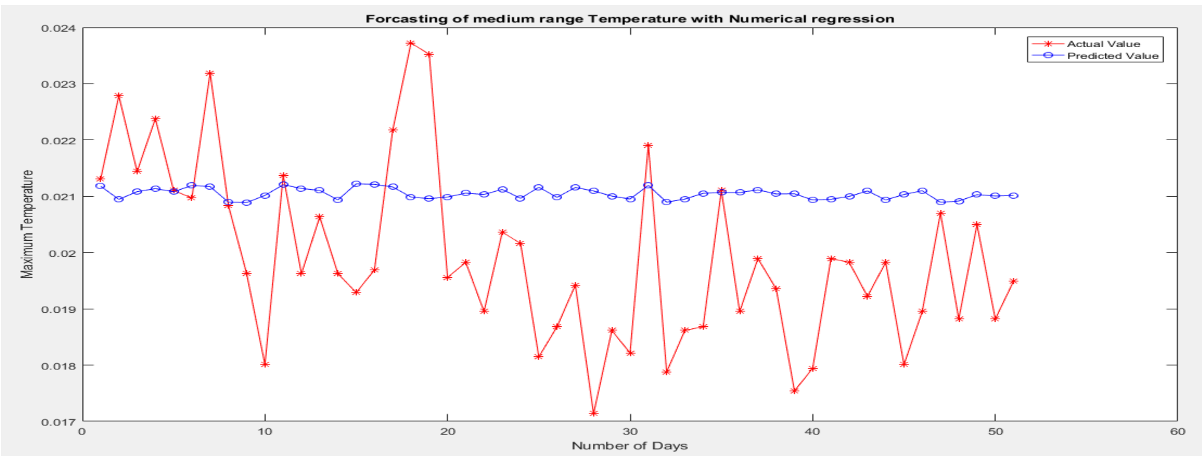


Fig 5. 15 Forecasting of maximum temperature using numerical with polynomial regression after handling the missing data

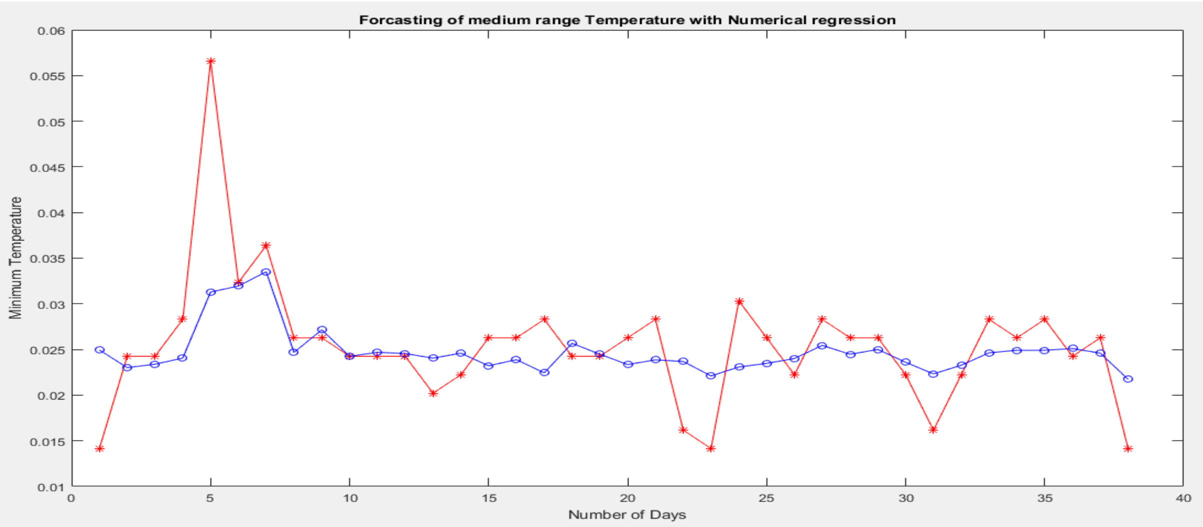


Fig 5. 16 forecasting of minimum temperature using numerical with polynomial regression after handling the missing data

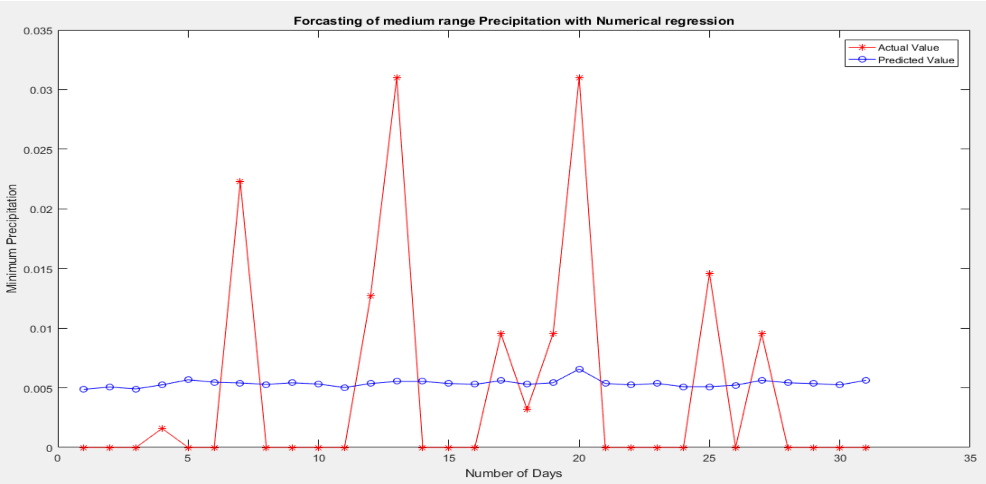


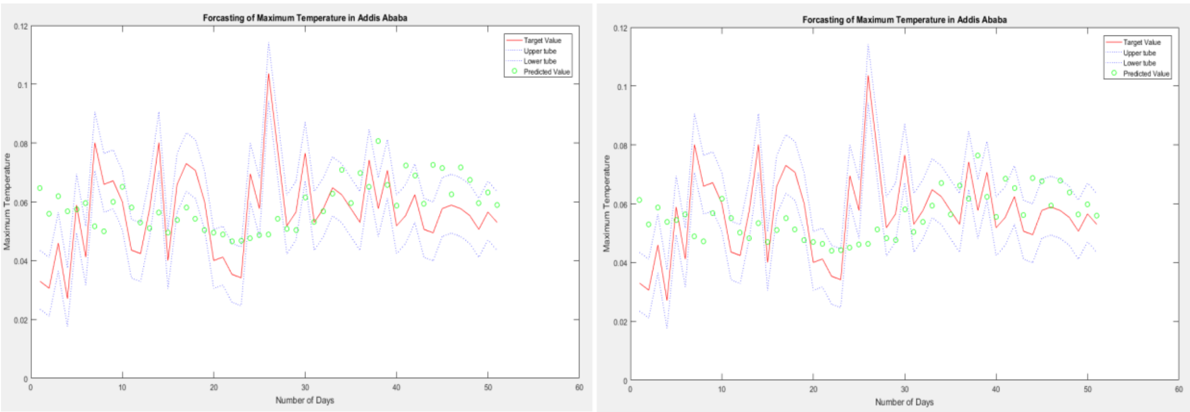
Fig 5. 17 Forecasting of precipitation using numerical with polynomial regression after handling the missing data

In [47] paper describes, numerical based weather forecasting without powerful computer is not give us good fit. Therefore, as we have seen from figure 5.27, 5.29, and 5.31, the estimated values are not closely traced to the target value.

Experiment 3 – Long range

The figure 5.18 (a) and (b) shows the long range prediction of maximum temperature before and after handling the missing data using SVM. The red, green, and blue indicates the actual data, predicted values, and upper and lower e-insensitive tube respectively. Besides, the y-axis is normalized one third average of daily temperature values and the x-axis is the number of days which are selected randomly as sample.

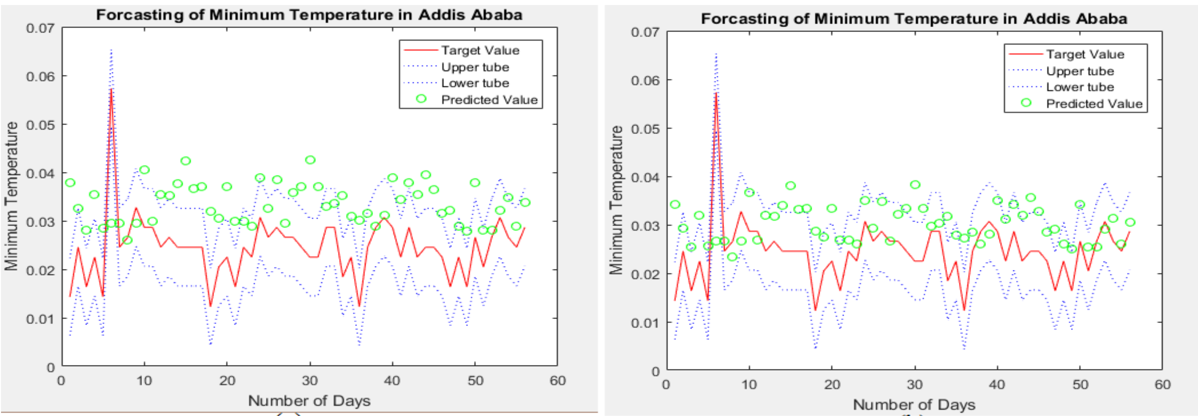
As we will see in the figures 5.18, 5.19, 5.20, and 5.21, almost 75% of the predicted values are support vectors. However, they have greater than 79.34% accuracy. This result indicates us even most of the predicted values are out of the e-tube, but they are well trained.



(a) (b)

Fig 5. 18 (a) Maximum temperature with missing data (b) Maximum temperature after handling the missing data

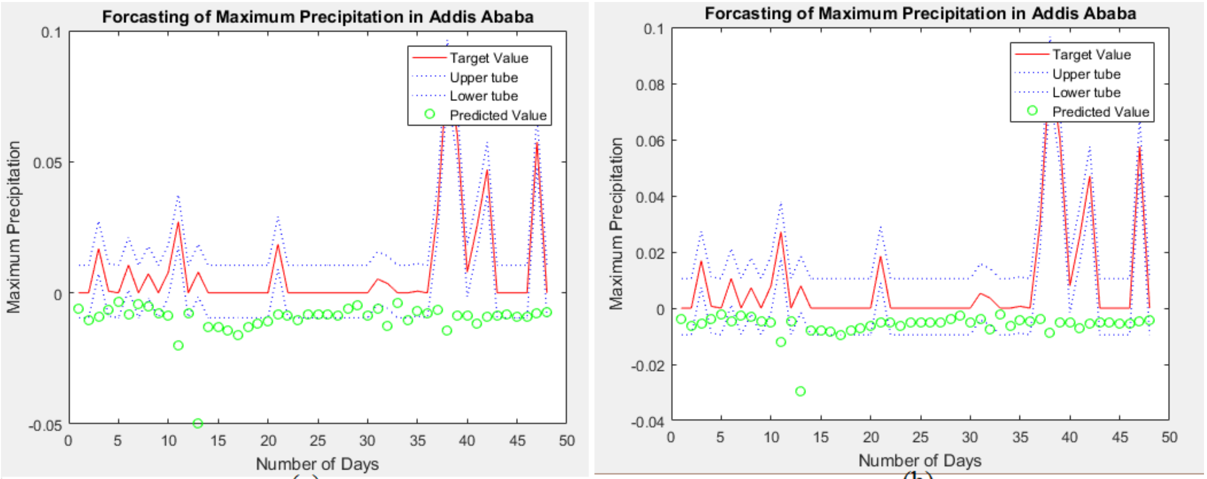
The figure 5.19 (a) and (b) shows the long range prediction of minimum temperature before and after handling the missing data using SVM. Besides, the red, green, and blue indicates the actual data, predicted values, and upper and lower e-insensitive tube respectively.



(a) (b)

Fig 5. 19 (a) Minimum temperature with missing data (b) Minimum temperature after handling the missing data

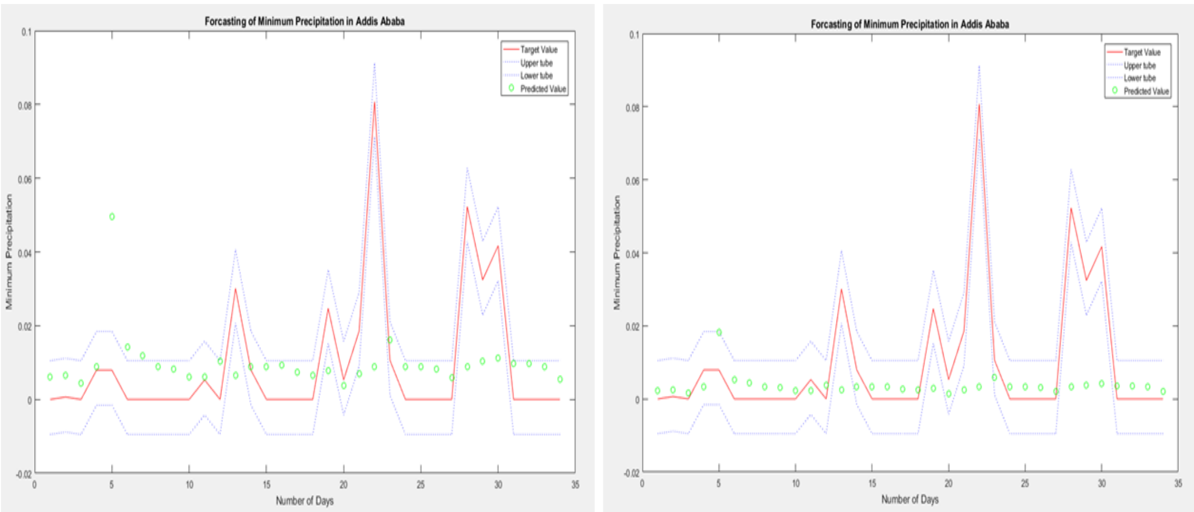
The figure 5.20 (a) and (b) shows the long range prediction of maximum precipitation before and after handling the missing data using SVM. Besides, the red, green, and blue indicates the actual data, predicted values, and upper and lower e-insensitive tube respectively. Besides, the y-axis is normalized one third average of daily precipitation values and the x-axis is the number of days which are selected randomly as sample.



(a) (b)

Fig 5. 20 (a) Maximum Precipitation with missing data (b) Maximum Precipitation after handling the missing data

The figure 5.21 (a) and (b) shows the long range prediction of minimum precipitation before and after handling the missing data using SVM. Besides, the red, green, and blue indicates the actual data, predicted values, and upper and lower e-insensitive tube respectively.



(a) (b)

Fig 5. 21 (a) Minimum precipitation with missing data (b) Minimum precipitation after handling the missing data

The figure 5.22 shows the long range maximum temperature prediction and error graph of the forecasting using DBN. The red, green, and light blue colors indicate the graph of actual, predicted, and error respectively. Besides, the y-axis is normalized one third average of daily temperature values and the x-axis is the number of days which are selected randomly as sample.

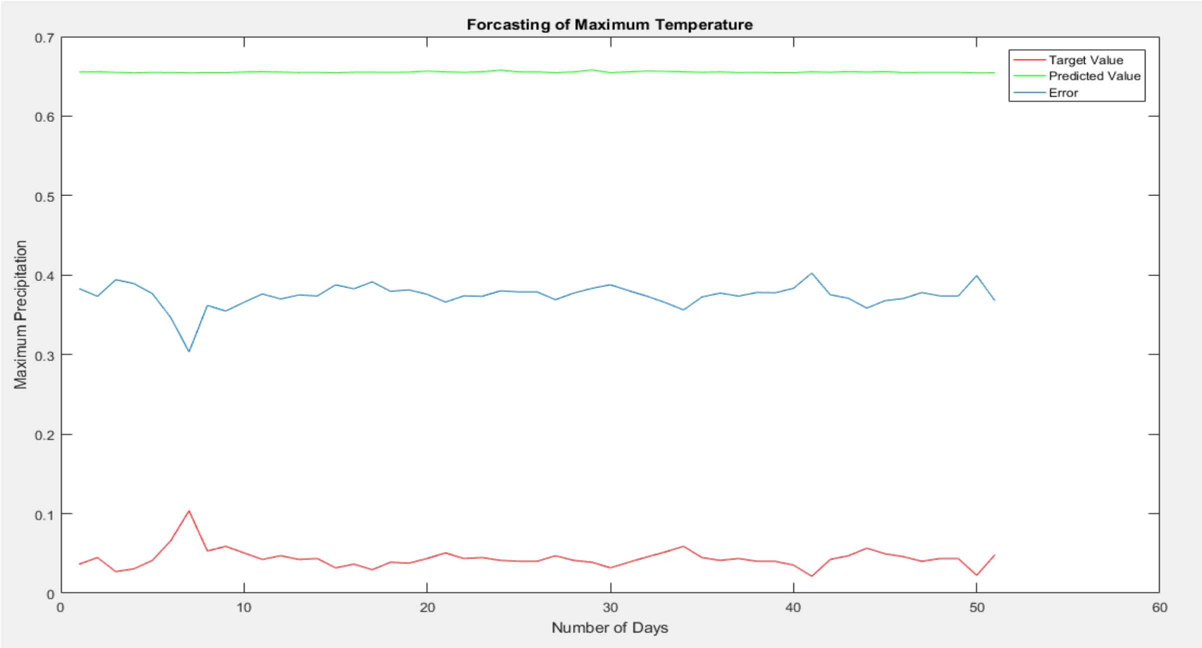


Fig 5. 22 Long range forecasting of maximum temperature using DBN

From the result figure 5.22 observed, unfortunately the forecasting system seems not to learn very well and we have achieved 34.8% accuracy. This is because due to the small dimensionality input of the historical dataset.

The figures 5.23 and 5.24 shows the forecasting of the long range of the maximum and minimum temperatures using numerical method respectively. Besides, the y-axis is normalized one third average of daily temperature values and the x-axis is the number of days which are selected randomly as sample.

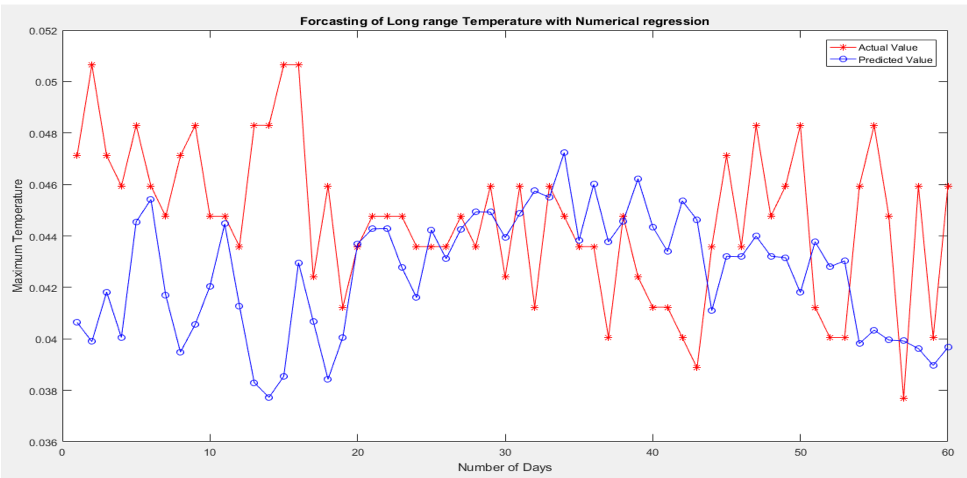


Fig 5. 23 forecasting of maximum temperature using numerical with polynomial regression after handling the missing data

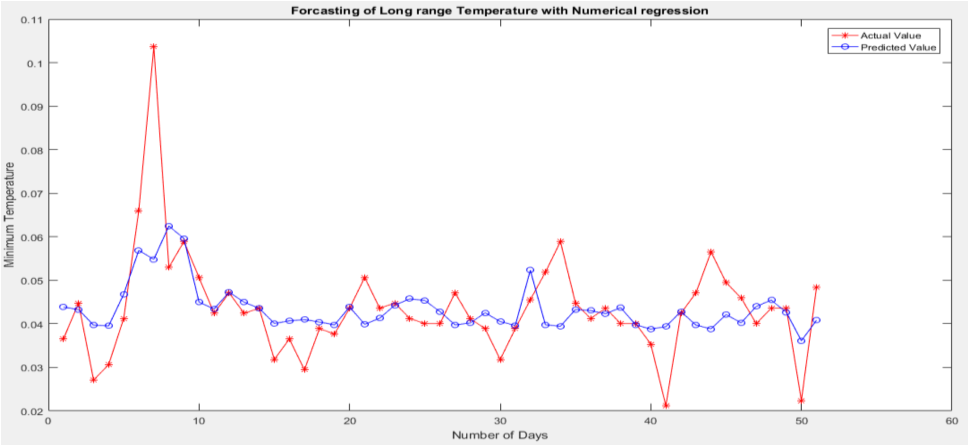


Fig 5. 24 forecasting of minimum temperature using numerical with polynomial regression after handling the missing data

As we have clearly seen from the figure 5.42 and 5.44, the predicted data is not closely traced to the actual data and we have achieved 58% accuracy.

5.3. Discussion

The correctness and accuracy of the proposed model is checked using Root Mean Square Error (RMSE) and the time consumption. The RMSE is measured by square rooting the average of the squares of errors and the time consumption measures the time taken to execute the result, that is the difference between starting time and final time. Besides, the accuracy of forecasting is calculated by the product of RMSE and 100%. Finally, we conclude, if the result has lesser RMSE and time consumption values, as more accurate the result are.

As we have seen in the result section, in this research work, the proposed algorithms comprehensively simulated using maximum and minimum values of temperature and precipitation before and after handling the missing values with respect to Ethiopian context based on the three forecasting ranges such as short range, medium range, and long range. Moreover, the network is trained with four and half years hourly and daily historical data recorded from the eight Ethiopian towns that listed in chapter three.

In **experiment one** (i.e. short range forecasting), we uses similar hourly historical data-set to train SVM and DBN networks. Since the change of temperature and precipitation in short term is very stable, the stations are not recorded their maximum and minimum values. Therefore, in this experiment, we did not provide the result based on their maximum and minimum values. The objective of this experiment is to forecast the short term (i.e. hourly) temperature and precipitation using the proposed algorithms in contrast to the existing system. From figure 5.3

and figure 5.5 (a), we observed that both SVM regression and DBN for regression approximately similar. Furthermore, figure 5.4 and 5.5 (b) shows that DBN has better accuracy than SVM regression. However, as we have observed the figure 5.6 and 5.7, numerical method forecasting less accurate than both SVM and DBN.

Empirically, forecasting of temperature and precipitation using SVM regression are 79.6% and 87.3% accurate respectively. Whereas, when we implement them using DBN, they are 88.6% and 88.3% accurate respectively. Therefore, as in [20] [21] [25] ensures the accuracy of DBN in forecasting of a big data and as we clearly observed from the figures as well as the experimental results, the DBN gives a better performance than SVM regression as well as numerical based regression.

In **experiment two** (i.e. Medium range forecasting), we have use the same daily recorded data-set to implement the weather prediction using DBN, SVM networks, and numerical regression method; based on their maximum and minimum values. Moreover, we predict the atmospheric condition based on before and after handling the missing values except during forecasting numerically. As we observe the empirical results from figures 5.8, 5.9, 5.10, 5.11, 5.12, 5.13, 5.14, 5.15, 5.16, and 5.17 as well as from table 9, DBN has better performance to forecast the future value of maximum and minimum temperature and precipitation before and after handling the missing values in contrast to SVM and numerical regressions.

For instance, the accuracy of predicting maximum temperature before and after handling the missing data is 80.4% and 81.8% using SVM regression and 87.7% and 92.2% using DBN respectively achieved. And also, the accuracy of forecasting minimum temperature before and after handling the missing data is 80.8% and 83.3% using SVM regression and 90.13% and 90.72% using DBN respectively. Furthermore, DBN is better for forecasting of the precipitation values.

In **experiment three** (i.e. Long range forecasting), we uses the average value of one third of the daily recorded data-set; and implemented using the same dataset for all DBN, SVM, and numerical regression algorithms. Then, the total average of one third of the four and half years data recorded is around 546 dataset. According to [23], DBN requires high dimensional input dataset for train the network and to get better performance. Due to this reason, from figure 5.18 (a), 5.22, and 5.23 clearly observed, the SVM and numerical regressions are completely outperforms DBN based forecasting methods. For example, the accuracy of forecasting of maximum temperature before handling the missing data is 86.02% and 32.88% using SVM regression and DBN respectively. Therefore, as we have seen from the empirical result, in low dimension input dataset DBN has very low performance relatively to SVM as well as numerical regressions.

The table 9 gives us the result of the three experiments of Maximum temperature before handling the missing data (MaxTM), Maximum temperature after handling the missing data (MaxTH), Minimum temperature before handling the missing data (MinTM), Minimum temperature after handling the missing data (MinTH), Maximum precipitation before handling the missing data (MaxPM), Maximum precipitation after handling the missing data (MaxPH), Minimum precipitation before handling the missing data (MinPM), and Minimum precipitation after handling the missing data (MinPH) based on RMSE and time consumption.

Table 8 Summary of the time required to train the system with the RMSE values in each of the experiments

Forecasting Range			**SVM**	**DBN**
Short range	Temp	RMSE	0.204	0.114
		TimeCo	147.6 seconds	127.3 seconds
	Prec	RMSE	0.127	0.117
		TimeCo	346.9 seconds	115.9 seconds
Medium Range	MaxTM	RMSE	0.196	0.123
		TimeCo	40.6 seconds	32.0 seconds
	MaxTH	RMSE	0.167	0.078
		TimeCo	42.1 seconds	37.12 seconds
	MinTM	RMSE	0.192	0.0987
		TimeCo	146.3 seconds	72 seconds
	MinTH	RMSE	0.182	0.0928
		TimeCo	171.3 seconds	49.7 seconds
	MaxPM	RMSE	0.1259	0.159
		TimeCo	172.2 seconds	23.0 seconds
	MaxPH	RMSE	0.1253	0.073
		TimeCo	218.0 seconds	21.92 seconds
	MinPM	RMSE	0.0267	0.0089
		TimeCo	149.0 seconds	19.2 seconds
	MinPH	RMSE	0.0266	0.0082
		TimeCo	146.7 seconds	17.6 seconds
Long Range	MaxTM	RMSE	0.139771	0.6712
		TimeCo	4.7 seconds	3.0 seconds
	MaxTH	RMSE	0.183137	0.6515
		TimeCo	4.5 seconds	2.75 seconds
	MinTM	RMSE	0.159	0.6891
		TimeCo	6.9 seconds	3.2 seconds
	MinTH	RMSE	0.119	0.6874
		TimeCo	6.9 seconds	2.55 seconds

	MaxPM	RMSE	0.1444	0.7152
		TimeCo	7.0 seconds	2.37 seconds
	MaxPH	RMSE	0.1433	0.4317
		TimeCo	7.3 seconds	2.50 seconds
	MinPM	RMSE	0.2460	0.6815
		TimeCo	5.0 seconds	3.15 seconds
	MinPH	RMSE	0.2457	0.6358
		TimeCo	4.2 seconds	1.77 seconds

All in all, if the input dataset has high dimension, DBN based forecasting provides high performance in contrast to SVM and numerical regressions; and in small dimension input dataset, forecasting with respect to SVM and numerical regression outperforms the DBN forecasting methods.

5.3. Summary

In this chapter several empirical result is provided based on the forecasting ranges using DBN in contrast to SVM and NWP. And also, as it is difficult to put and discuss the whole achieved result and since the economy of Ethiopia is highly depends on agriculture which factor affects by temperature and rainfall distribution, the temperature and precipitation weather events are selected and delivered as a sample. Besides, the whole experimental result performance are evaluated based on the percentage of RMSE and the time taken during training the network.

Chapter Six

6. Conclusion and Future Works

6.1. Conclusion

Forecasting of the accurate atmospheric condition is becoming an essential necessity for the individuals, agriculturalist, and other socio-economics in the world. Therefore, this research proposed to investigate a new forecasting model with respect to Ethiopian context based on the three forecasting range such as short range, medium range, and long range. The proposed system is implemented using DBN, SVM, and numerical regression algorithms by classifying the daily and hourly collected historical data based on the Ethiopian seasons and rainfall regimes.

The collected time series data from different station are incomplete and inconsistent. Therefore, to resolve this issue, we apply linear and cubic spline interpolation technique. Cubic spline interpolation is most effective on handling a small amount of missing data-sets (i.e. four contiguous missing values) in contrast to linear interpolation. However, linear interpolation is most effective on handling large amount of missing data. Additionally, we apply Spearman's correlation for future extraction of the selected weather variables.

This research work compares the performance of SVM for regression and DBN based forecasting of the weather events with the existing forecasting method; that is numerical regression based on RMSE and time consumption. As described in [46], SVM for regression does not care about the forecasting of errors as long as they are less than ε-tube (i.e. the upper and lower bound of the target value). However, DBN focuses on how the target value traces the actual value. In experiment one, which is short range forecasting, both SVM regression and DBN networks have approximately similar and have higher performance than the numerical based forecasting. For instance, the forecasting accuracy of temperature is 88.6%, 78.2%, and 52.5% using DBN, SVM, and Numerical algorithms respectively. But, forecasting using DBN is closely traced the actual data. Similarly, in experiment two, which is medium range forecasting, also have approximately similar results with experiment one was achieved. But, in the last experiment, which is long range forecasting, SVM and numerical regressions are completely outperforms DBN based prediction. As we have seen from the result, DBN based forecasting of a big data is more efficient in contrast to the SVM as well as numerical regressions. However, for forecasting of small dimension input data, SVM and numerical regressions has better performance than DBN.

Moreover, this thesis implements the forecasting of the weather events before and after handling the missing values and their performance are quite similar in DBN and a little bit difference in SVM regression. However, relatively, forecasting after handling the missing data gives us better performance.

Generally, if we apply big data of weather events as input, the DBN has a better performance than SVM regression and forecasting with respect to SVM regression outperforms the DBN forecasting methods, if we use small dimension dataset as input values. Besides, the proposed system is useful for forecasting of high dimension weather events input data with respect to Ethiopian context.

6.2. Future Works

Since numerical weather forecasting method need high performance computer to predict accurate weather events and difficult forecasting of accurate atmospheric condition with the recent computers, for the next:

- ✓ We should focus on applying different machine learning algorithms to forecast accurate weather condition.
- ✓ With considering the effect of interpolation during missing data handling, to forecast an accurate atmospheric condition, it is also better to add different types of weather events namely radiation, cloud distribution, wind direction and speed.

Furthermore, in order to achieve accurate weather events, we recommended designing and implementing of hybrid system called combination of DBN and SVM algorithms regression for the Ethiopian context can give us much accurate results.

Reference

[1] Slovenian environment agency. *Weather and Climate* [online]. Available: http://www.arso.gov.si/en/Weather, [Accessed on: Nov 27, 2017]

[2] National Meteorological Services Agency. *Brief History* [online]. Available: http://www.ethiomet.gov.et/about/history, [Accessed on: Nov 27, 2017]

[3] CAeM Working Group on Advanced Techniques Applied to Aeronautical Meteorology, *methods of interpreting numerical weather prediction output for aeronautical meteorology*, Second edition. Geneva Switzerland: World Meteorological Organization (WMO–No. 770), 1999

[4] Ethiopian National Meteorological Services Agency. (2001E.c, Oct. 06). *Brief History* [online]. Available: http://www.ethiomet.gov.et/about/history, [Accessed on: Feb 29, 2017]

[5] Seifu Admassu and Dr. Ing. Abdulkarim H. Seid, "Rainfall Variation and its Effect on Crop Production in Ethiopia," M.S. thesis, Dept. Civ. Eng., Addis Ababa Univ., Ethiopia, MSc, Aug 2004.

[6] *Initial National Communication of Ethiopia to the United Nations Framework Convention on Climate Change (UNFCCC),* 1st ed., FDRE Ministry of Water Resources NMSA., Addis Ababa, Eth., Jun 2001. pp. 2-30.

[7] Demel Teketay, "Deforestation, wood famine and environmental degradation in highland ecosystems of Ethiopia: urgent need for actions". *International Conference on African Development Archives*. Paper 3., 2001.

[8] Dawit Abebe, "Future climate of Ethiopia from PRECIS Regional Climate Model Experimental Design," M.S. thesis, Dept. Civ. Eng. Addis Ababa Univ., Ethiopia, MSc, Nov 2010.

[9] Aklilu Mekasha, Alan J. Duncan (PhD), Kindie Tesfaye (PhD), Lisanework Nigatu (PhD), and Chilot Yirga (PhD), "Climate and Land Use /Land Cover Change and Their Interconnections with Livestock Feed Resource Management in Ethiopia," Ph.D. dissertation, Dept. Agro./physio., Haramaya Univ., Ethiopia, MA, Feb 2014.

[10] Environments of Africa. *Climates Of Africa - Forming Of The Sahara Desert* [online]. Available: https://courseware.e-education.psu.edu/courses/earth105new/content/lesson07/03.html, [Accessed on: Feb 28, 2017]

[11] W. Taylor James and Roberto Buizza, “Neural Network Load Forecasting with Weather Ensemble Predictions”, IEEE Trans. on Power Systems, 2002, Vol. 17, pp. 626-632.

[12] Malik Pooja, Prof. Singh Saranjeet, Arora Binni, et al., "An Effective Weather Forecasting Using Neural Network," *International Journal of Emerging Engineering Research and Technology,* Volume 2, Issue 2, May 2014, PP 209-212

[13] Abhishek Kumar, Kumar Abhay, Ranjan Rajeev, Kumar Sarthak, et al., " A Rainfall Prediction Model using Artificial Neural Network," *IEEE Control and System Graduate Research Colloquium (ICSGRC),* 2012,

[14] Kumar Neeraj , Kumar Jha Govind, et al., "A Time Series ANN Approach for Weather Forecasting," *International Journal of Control Theory and Computer Modeling (IJCTCM),* Vol.3, No.1, January 2013

[15] Narvekar Meera, Fargose Priyanca, et al., “Daily Weather Forecasting using Artificial Neural Network,” *International Journal of Computer Applications,* Volume 121 – No.22, July 2015

[16] Hung N. Q., Babel M. S., Weesakul S., Tripathi N. K., “An artificial neural network model for rainfall forecasting in Bangkok, Thailand,” School of Eng. and Tech., Asian Institute of Technology, Thailand, 7 August 2009

[17] Maqsood Imran, Riaz Khan Muhammad, Abraham Ajith,“ An ensemble of neural networks for weather forecasting” Dep. of Comp.Scie., Univ. of Victoria, 20 May 2004

[18] Ranjan Nayak Deepak, Mahapatra Amitav, Mishra Pranati, et al., “A Survey on Rainfall Prediction using Artificial Neural Network,” *International Journal of Computer Applications,* Volume 72– No.16, June 2013

[19] Chen Kai, Jun Liu, ShanxinGuo, Jinsong Chen, Ping Liu, Jing Qian, Huijuan Chen, Bo Sun, et al., “Short-Term Precipitation Occurrence Prediction for Strong Convective Weather Using Fy2-G Satellite Data: A Case Study of Shenzhen, South China,” *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* Volume XLI-B6,2016 XXIII ISPRS Congress, Jul 2016.

[20] Grover, Aditya, Ashish Kapoor, and Eric Horvitz. "A deep hybrid model for weather forecasting." *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2015.

[21] N.K. Liu James, Hu Yanxing, Jia You Jane, Wai Chan Pak, “Deep Neural Network Based Feature Representation for Weather Forecasting” Dep. of Computing, Hong Kong Polytechnic Univ.,nd

[22] Gao Bing, Saad Malik, YudiSantoso, Zhanxue Zhu,“Using Data Mining Technique to Predict Seasonal Climate Change” Dep. of Comp.Scie., Univ. of Victoria, May 12, 2016

[23] Palmer T.N., “Predicting uncertainty in forecasts of weather and climate”, ECMWF technical Memorandum No. 294**,** Nov 1999,

[24] Yadav Rohit Kumar, Ravi Khatri, et al., “A Weather Forecasting Model using the Data Mining Technique, ”*International Journal of Computer Applications (0975 – 8887),*Volume 139 – No.14, April 2016.

[25] Wang H.Z., Wang G.B., Li G.Q., Peng J.C., Liu Y.T., “Deep belief network based deterministic and probabilistic wind speed forecasting approach,” College of Mechatronics and Control Engineering, Shenzhen University, 2016

[26] Taillardat Maxime, Foug`eres Anne-Laure, Naveau Philippe, Mestre Olivier., “Forest-based methods and ensemble model output statistics for rainfall ensemble forecasting,” Laboratoire des Sciences du Climat et de l’Environnement, Univ Lyon, 11 novembre 1918 [27] B. Iversen Emil, M. Morales Juan, K. Møller Jan, Madsen Henrik., “Short-term probabilistic forecasting of wind speed using stochastic differential equations,” Technical University of Denmark, 2015

[28] Kuhn Max and Johnson Kjell, Applied Predictive Modeling, 3rd ed. New York, 2013.

[29] National Centers for Environmental Information. *Climate Data Online* [online]. Available: https://www7.ncdc.noaa.gov/CDO/cdo

[30] Nakagawa Shinichi, “Missing data: mechanisms, methods, and messages,” *Ecological Statistics: Contemporary Theory and Application*. First Edition. Edited by Gordon A. Fox, Simoneta Negrete-Yankelevich, and Vinicio J. Sosa. Eds. Oxford University Press, 2015.

[31] Soley-Bori Marina, “Dealing with missing data: Key assumptions and methods for applied analysis,” Dept. Health policy and Manag., Boston Univ., **USA**, Tech. Rep. No. 4, May 6, 2013.

[32] Sheung Chi Fung David, " Methods for the Estimation of Missing Values in Time Series ". M.S. thesis, Dept. Mathematics and Planning. Edith Cowan Univ., Perth, Western Australia, MSc, 2006.

[33] Schagat Trista, Paguio Aileen, Kopish Kevin, Corporation Promega, “Normalizing Genetic Reporter Assays: Approaches and Considerations for Increasing Consistency and Statistical Significance,” Cell Notes Issue, 17 2007.

[34] Github. *CS231n Conventional Neural Networks for Visual Recognition* [online]. Available: http://cs231n.github.io/neural-networks-2/, [Accessed on: Dec 21, 2017]

[35] Sola J., Sevilla J., et al., “Importance of Input Data Normalization for the Application of Neural Networks to Complex Industrial Problems,” *IEEE transactions on nuclear science*, Vol. 44, No. 3, JUNE 1997.

[36] Kim Daehyon, “Normalization methods for input and output vectors in backpropagation neural networks,” International Journal of Computer Mathematics, 71:2, 161-171, 20 Mar 2007.

[37] Abdi Herv´e, “Normalizing Data,” *Program in Cognition and Neurosciences*, Texas Univ., Dallas, 2010.

[38] I. Davies Philip, J. Higham Nicholas, et al., “Numerically Stable Generation of Correlation Matrices and Their Factors,” Vol. 40, No. 4, pp. 640–651, 2000

[39] Robin Beaumont, “An Introduction to statistics Correlation,” [Online], Sep 2012. Available: http://www.robin-beaumont.co.uk/virtualclassroom/stats/course1.html

[40] Shong Chok Nian, “Pearson’s Versus Spearman’s and Kendall’s Correlation Coefficients for Continuous Data,” M.S. thesis, Dept. Public Health, Pittsburgh Univ., MSc, 2010.

[41] *Deep Learning Tutorial*, 1st ed., LISA lab., Montreal Univ., Sep 01, 2015, pp. 91–117.

[42] *Neural Networks and Deep Learning,* Available: www.cs.wisc.edu/~dpage/cs760/, [Accessed on: Apr 19, 2017]

[43] E. Hinton Geoffrey and Osindero Simon, "A Fast Learning Algorithm for Deep Belief Nets," Comput. Sci., Univ. Toronto, Canada, nd.

[44] Liu Feng, Liu Bingquan, Sun Chengjie, Liu Ming, Wang Xiaolong, " Deep Belief Network-Based Approaches for Link Prediction in Signed Social Networks," School of Computer Science and Technology, Harbin Institute of Technology,, No. 92, 10 April 2015

[45] Batres-Estrada Gilberto. *Deep Learning for Multivariate Financial Time Series*. Stockholm, June 4, 2015.

[46] Drucker, H., Burges, C.J., Kaufman, L., Smola, A.J. and Vapnik, V., 1997. Support vector regression machines. In *Advances in neural information processing systems* (pp. 155-161).

[47] Encyclopedia. *Weather Forecasting* [online]. Available: https://www.encyclopedia.com/earth-and-environment/atmosphere-and-weather/weather-and-climate-terms-and-concepts/weather-0#3437800646 [Accessed on: Apr 23, 2017]

[48] Stefanie Eminger. *The History of Weather Forecasting*[online] Available: http://www-history.mcs.st-andrews.ac.uk/HistTopics/Weather_forecasts.html, [Accessed on: Apr 23, 2017]

[49] Meteorology 311. (Fall 2015). www.meteor.iastate.edu/classes/mt311/powerpoint/mteor311_NWP_2015.pdf, [Accessed on: May 29, 2017]

[50] Smola A. and Schölkopf B. *A Tutorial on Support Vector Regression*. NeuroCOLT Technical Report NC-TR-98-030, Royal Holloway College, University of London, UK, 1998

Appendix A.

Experiment 1

Short_SVM_temp.m short

```
clear all;
close all;
clc
format short
% Initialization
Slack_Val = 0.0005;
%Control_Val = 100;
Control_Val = input('Control_Val: ');
if Control_Val == Inf
    tol = 0.0005;
else
    tol = Control_Val * 1e-6;
end
bias = 0;
Addis_Max_Temp =
xlsread('F:/Thesis/All_In_One_Codes/SVM_Code/AddisAbaba/Temperature/MaxTemp
erature/Temp.xlsx');
Max_Temp = Addis_Max_Temp( :,1); % Target Value in degree f
Max_Visi = Addis_Max_Temp( :,2);
Max_Wind = Addis_Max_Temp( :,3);
Max_Prec = Addis_Max_Temp( :,4); % Precipitation in 0.1 Inch
Gaus_Noi = Addis_Max_Temp( :,5);
% Replacing the NaN value into zero
Max_Temp(isnan(Max_Temp))=0;
Max_Visi(isnan(Max_Visi))=0;
Max_Wind(isnan(Max_Wind))=0;
Max_Prec(isnan(Max_Prec))=0;
% Change the value of temperature into degreeC
Max_Temp = 5/9 * (Max_Temp - 32);
% Convert the precipitation value to milli meter
Max_Prec = (Max_Prec./2.54);
% Normalize the input variables a = a./sqrt(sum(a.^2))
Max_Temp = Max_Temp./sqrt(sum(Max_Temp.^2));
Max_Visi = Max_Visi./sqrt(sum(Max_Visi.^2));
Max_Wind = Max_Wind./sqrt(sum(Max_Wind.^2));
Max_Prec = Max_Prec./sqrt(sum(Max_Prec.^2));
Gaus_Noi = Gaus_Noi./sqrt(sum(Gaus_Noi.^2));
% Augmented matrix of the selected input variables
%%Training input
Input_Value =
[Max_Visi(1:2000),Max_Wind(1:2000),Max_Prec(1:2000),Gaus_Noi(1:2000)]; %
Training input variables
%Training target
Max_Temp_Train = Max_Temp(1:2000);
%%Test Inputvalues
Input_Value_Test =
[Max_Visi(2001:end),Max_Wind(2001:end),Max_Prec(2001:end),Gaus_Noi(2001:end
)]; % Test input variables
%Test target
Max_Temp_Test = Max_Temp(2001:end);
fprintf('Support vector regressing...\n')
e = 0.05;
%tolorance for support vector detection
Epsilon = tol;
```

```
%construct the kernel matrix
% Mapping of the input variables into higher dimension
n = size(Max_Temp_Train,1);
phi = zeros(n,n); % Initialize the Constracting of the kernel matrix
fprintf(' Constracting of the kernel matrix...\n')
%phi =  Input_Value * Input_Value';
for i = 1:n
    for j = 1:n
       phi(i,j) = (Input_Value(i,:) * (Input_Value(j,:))');
    end
end
% Ploting of the e-tube
Y1 = Max_Temp_Train + Epsilon;
Y2 = Max_Temp_Train - Epsilon;
plot(Max_Temp_Train,'r-')
hold on
plot(Y1,':b')
plot(Y2,':b')
% Randomize Weight then predict the new value
weight = rand(n,1);
Pre_Value = weight' * phi + bias;
Pre_Value = Pre_Value';
% Calculate the e-insensitive error
Error = abs(Pre_Value - Max_Temp_Train);
Sum_Error = sum(Error);
if Sum_Error <= Epsilon
    fprintf('Error is converged. \n')
    plot(Pre_Value,'og')
else
% Adding Slack number to the epsilon tube
    hold off
    plot(Max_Temp_Train(1:50:2000),'r-')
hold on
plot(Y1(1:50:2000) + Slack_Val,':b')
plot(Y2(1:50:2000) + Slack_Val,':b')
end
% Solve the Optimizing Problem
phii = [phi -phi; -phi phi];
        c = [(e*ones(n,1) - Max_Temp_Train); (e*ones(n,1) +
Max_Temp_Train)];
vlb = zeros(2*n,1);
        vub = Control_Val*ones(2*n,1);
        x0 = zeros(2*n,1);
               bias = 0;
           A = [ones(1,n) -ones(1,n)]; b = 0;
% Regularisation
phii = phii + 1e-10*eye(size(phii));
fprintf('Optimizing... \n')
%tol = 0.5;
%iter = 1000;
Initial_time = cputime;
options = optimset('maxiter',1000);
[alpha lambda how] = quadprog(phii,c,A,b,[],[],vlb,vub,[],options);
%[alpha how] = quadprog(phii,f,A,b,Aeq,beq,vlb,vub);
fprintf('Execution time: %4.1f seconds\n', cputime - Initial_time);
fprintf('Status: %s\n', how);
beta = alpha(1:n) - alpha(n+1:2*n); % the difference of lagrangian
multiplier
fprintf('Sum beta: %f\n',sum(beta));
% Compute the number of support vectors
svi = find(abs(beta) > Epsilon & abs(beta) < (Control_Val - Epsilon));
```

```
nsv = length(svi); % Number of support vectors
fprintf('Support Vectors: %d (%3.1f%%)\n',nsv,100*nsv/n);
%Z = (phi'*phi).^2; % polynomial/quadratic/ kernel function
Z = sum(beta'* Input_Value);
% find bias from average of support vectors with interpolation error e
%SVs with interpolation error e have alphas: 0 < alpha < C
svii = find(abs(beta) > Epsilon & abs(beta) < (Control_Val - Epsilon)) ;
if length(svii) > 0
    bias = (1/length(svii))* sum(Max_Temp_Train(svii))- Epsilon *
sign(beta(svii)) - (phi(svii,svi)* beta(svi));
else
fprintf('No support vectors with interpolation error e cannot compute
bias.\n');
bias = (max(Max_Temp_Train)+ min(Max_Temp_Train))/2;
end
%%constracting of tesfting of kernel matrix
k = size(Max_Temp_Train,1);
m = size(Max_Temp_Test,1);
Tes_Ker = zeros(k,m);
for i = 1:k
    for j = 1:m
       Tes_Ker(i,j) = (Input_Value_Test(j,:) * (Input_Value(i,:))');
    end
end
if length(bias)>= 240
Test_Max_Temp = beta'*Tes_Ker + (bias(1:240))';
else
   Test_Max_Temp = beta'*Tes_Ker + ([bias;zeros((240-length(bias)),1)])';
end
%w = beta' * phi;
if length(svii) > 0
Pre_Value = (beta' * (Input_Value*Input_Value' + 1).^2 +
([bias;zeros((2000-length(svii)),1)])');
else
   Pre_Value = (beta' * (Input_Value*Input_Value' + 1).^2 + bias);
end

%Pre_Value = (beta' *(phi*phi').^2 + ([bias;zeros((2000-
length(svii)),1)])');
Pre_Value = abs(Pre_Value');
Errorr = abs(Max_Temp_Train - Pre_Value);
MSE = sum(Errorr.^2)/n;
plot(Pre_Value((1:50:2000)),'og');
%plot(Error(1955:2000),'k-*')
xlabel('Number of Days')
ylabel('Maximum Temperature')
title('Forcasting of Maximum Temperature in Addis Ababa')
legend('Target Value','Upper tube','Lower tube','Predicted Value')
figure;
plot(Pre_Value(15:3:150),'r-')
hold on
plot(Max_Temp_Test(15:3:150)- Epsilon,':b')
plot(Max_Temp_Test(15:3:150)+ Epsilon,':b')
plot(Test_Max_Temp(15:3:150),'og')
xlabel('Number of Days')
ylabel('Hourly Temperature Values')
title('Forcasting of Temperature in Addis Ababa')
legend('Target Value','Upper tube','Lower tube','Predicted Value')
figure;
Errorr1 = abs(Test_Max_Temp(15:3:150) - (Max_Temp_Test(15:3:150))');
plot(Errorr1,'k')
```

```
xlabel('Number of Days')
ylabel('Error Value')
title('Forcasting Error of Temperature in Addis Ababa')
```

Appendix B.

Experiment 2

Max_Prec_Num.m

```
clear all;
close all;
clc
Addis_Min_Temp =
xlsread('F:/Thesis/All_In_One_Codes/SVM_Code/AddisAbaba/Precipitation/MinPr
ecipitation/Min_Pre_Hand.xlsx');
Min_Temp = Addis_Min_Temp( :,1); % Target Value
Min_Visi = Addis_Min_Temp( :,2);
Min_Wind = Addis_Min_Temp( :,3);
Min_Prec = Addis_Min_Temp( :,4); % Precipitation in 0.1 Inch
% Replacing the NaN value into zero
Min_Temp(isnan(Min_Temp))=0;
Min_Visi(isnan(Min_Visi))=0;
Min_Wind(isnan(Min_Wind))=0;
Min_Prec(isnan(Min_Prec))=0;
% Convert the precipitation value to milli meter
Min_Temp = (Min_Temp./2.57);
% Normalize the input variables
% a = a./sqrt(sum(a.^2))
Min_Temp = Min_Temp./sqrt(sum(Min_Temp.^2));
Min_Visi = Min_Visi./sqrt(sum(Min_Visi.^2));
Min_Wind = Min_Wind./sqrt(sum(Min_Wind.^2));
Min_Prec = Min_Prec./sqrt(sum(Min_Prec.^2));
%Gaus_Noi = Gaus_Noi./sqrt(sum(Gaus_Noi.^2));
%construct and solve the simultaneous equation using the vandermode matrix
vander_mode = [ones(size(Min_Temp)),Min_Visi,Min_Wind,Min_Prec];
Const_val = vander_mode\Min_Temp;
%poly1 = polyfit(range_data',Short_Temp,3);
Min_Temp_pred = Const_val(1) + Const_val(2)*Min_Visi +
Const_val(3)*Min_Wind + Const_val(4)*Min_Prec;
Error = abs(Min_Temp_pred-Min_Temp);
% Mapping of the input variables into higher dimension
n = size(Error,1);
RMSE = sqrt(sum(Error.^2)/n);
plot(Min_Temp(1:50:1509),'-r*')
hold on
%plot(range_data',polyval(poly1,range_data'),'bo-')
plot(Min_Temp_pred(1:50:1509),'-bo')
xlabel('Number of Days')
ylabel('Minimum Precipitation')
title('Forcasting of medium range Precipitation with Numerical regression')
legend('Actual Value','Predicted Value')
```

Appendix C.

Experiment 3

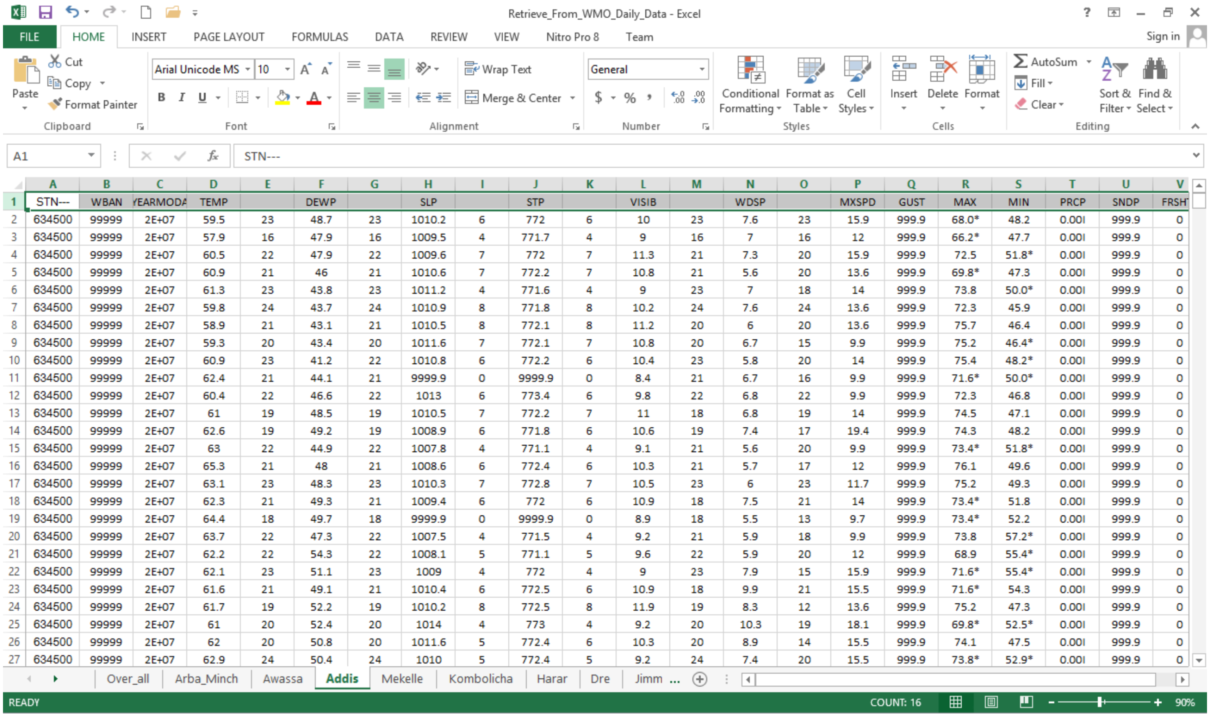
Max_Temp_DBN.m

```
clear
clc
LW = 0.1; LVB = 0.1; LHB = 0.1; WC = 0.1132;
IM = 0.5; FM = 0.9; LOB = 0.1;
Expected_Error = 0.0001;
EP = 1000;
count = 1;
% Initializing symmetric weights and biases.
Addis_Max_Temp =
xlsread('F:/Thesis/All_In_One_Codes/SVM_Code/AddisAbaba/Temperature/MinTemp
erature/Temperature_Hand.xlsx');
W = rand(4,3);
W2 = W';
VB = zeros(400,4);
HB = zeros(400,3);
OB = zeros(400,1);
POBH1 = zeros(400,1);
POBH2 = zeros(400,1);
POBO = zeros(400,1);
WN = zeros(4,3);
VBN = zeros(400,4);
HBN = zeros(400,3);
OBN = zeros(400,1);
Max_Temp = Addis_Max_Temp( :,1); % Target Value
Max_Visi = Addis_Max_Temp( :,2);
Max_Wind = Addis_Max_Temp( :,3);
% Convert the precipitation value to milli meter
Max_Prec = (Max_Prec./2.54);
% Reshape into quarter dimension
Max_Temp = reshape(Max_Temp,3,[]);
Max_Visi = reshape(Max_Visi,3,[]);
Max_Wind = reshape(Max_Wind,3,[]);
% Calculate the mean of the reshaped data
Max_Temp = mean(Max_Temp)';
Max_Visi = mean(Max_Visi)';
Max_Wind = mean(Max_Wind)';
% Normalize the input variables
Max_Temp = Max_Temp./sqrt(sum(Max_Temp.^2));
Max_Visi = Max_Visi./sqrt(sum(Max_Visi.^2));
Max_Wind = Max_Wind./sqrt(sum(Max_Wind.^2));
% Augmented matrix of the selected input variables
Input_Value = [Max_Visi,Max_Wind,Max_Prec,Gaus_Noi]; % Training input
variables
 Max_Temp_Train = Max_Temp(1:400);
Input_Value_Test =
[Max_Visi(401:end),Max_Wind(401:end),Max_Prec(401:end),Gaus_Noi(401:end)];
% Test input variables
%Test target
Max_Temp_Test = Max_Temp(401:end);
```

```
% Start positive phase
EH1 = W(1,1)*Max_Visi + W(2,1)*Max_Wind + W(3,1)*Max_Prec +
W(4,1)*Gaus_Noi;
EH2 = W(1,2)*Max_Visi + W(2,2)*Max_Wind + W(3,2)*Max_Prec +
W(4,2)*Gaus_Noi;
EH3 = W(1,3)*Max_Visi + W(2,3)*Max_Wind + W(3,3)*Max_Prec +
W(4,3)*Gaus_Noi;
POBH1 = 1./(1 + exp(-HB(:,1) - EH1));
POBH2 = 1./(1 + exp(-HB(:,2) - EH2));
% From H1 to Hidden two
EHO = W2(1,1)*POBH1 + W2(2,1)*POBH2 + W(3,1)*POBH3 ;
POBO = 1./(1 + exp(-OB - EHO));
%%%% End Positive Phase %%%%%%%%%%
Error = (Max_Temp - POBO).^2;
Sum_Error = sum(Error.^2)/length(Error);
RMSE = sqrt(Sum_Error);
% Start Negative Phase
EHON1 = W2(1,1)*POBO ;
EHON3 = W2(3,1)*POBO ;
POBO1 = 1./(1 + exp(-HB(:,1) - EHON1));
POBO2 = 1./(1 + exp(-HB(:,2) - EHON2));
%%%%%% From hidden to visible
W = W';
EHV1 = W(1,1)*POBO1 + W(2,1)*POBO2 + W(3,1)* POBO3;
EHV2 = W(1,2)*POBO1 + W(2,2)*POBO2 + W(3,2)* POBO3;
PH2V1 = 1./(1 + exp(-VB(:,1) - EHON1));
PH2V2 = 1./(1 + exp(-VB(:,2) - EHON1));
%%%%%%%%%%% End of Negative Phase
Max_Visi = Max_Visi - PH2V1;
Max_Wind = Max_Wind - PH2V2;
Max_Prec = Max_Prec - PH2V3;
Gaus_Noi = Gaus_Noi - PH2V4;
%%%%% Update weight and biases
count = 0;
%while RMSE < 0.38
for i = 1:EP
    if RMSE > Expected_Error
    count = count + 1;
    WN = W + LW.*(sum(POBO) - sum(PH2V1 + PH2V2 + PH2V3 + PH2V4));
    VBN = [(VB(:,1) + LVB * (PH2V1 - Max_Visi)),(VB(:,2) + LVB * (PH2V2 -
Max_Wind)),(VB(:,3) + LVB * (PH2V3 - Max_Prec)),(VB(:,4) + LVB * (PH2V3 -
Max_Prec))];
    HBN = [(HB(:,1) + LHB .*(POBH1 - POBO1)),(HB(:,2) + LHB .*(POBH2 -
POBO2)),(HB(:,3) + LHB .*(POBH3 - POBO3))];
    OBN = OB + LOB * POBO;
   W = WN'; VB = VBN; HB = HBN; OB = OBN;
   else
    disp('The error is converged.')
    end
if count == 1000
    plot(Max_Temp(1:10:503),'r-')
hold on
plot(POBO(1:10:503),'g-')
plot(Error(1:10:503))
xlabel('Number of Days')
ylabel('Maximum Precipitation')
title('Forcasting of Maximum Temperature')
legend('Target Value','Predicted Value','Error')
end
end
```

Appendix D.

Sample collected hourly data set



Retrieve_From_WMO_Daily_Data - Excel

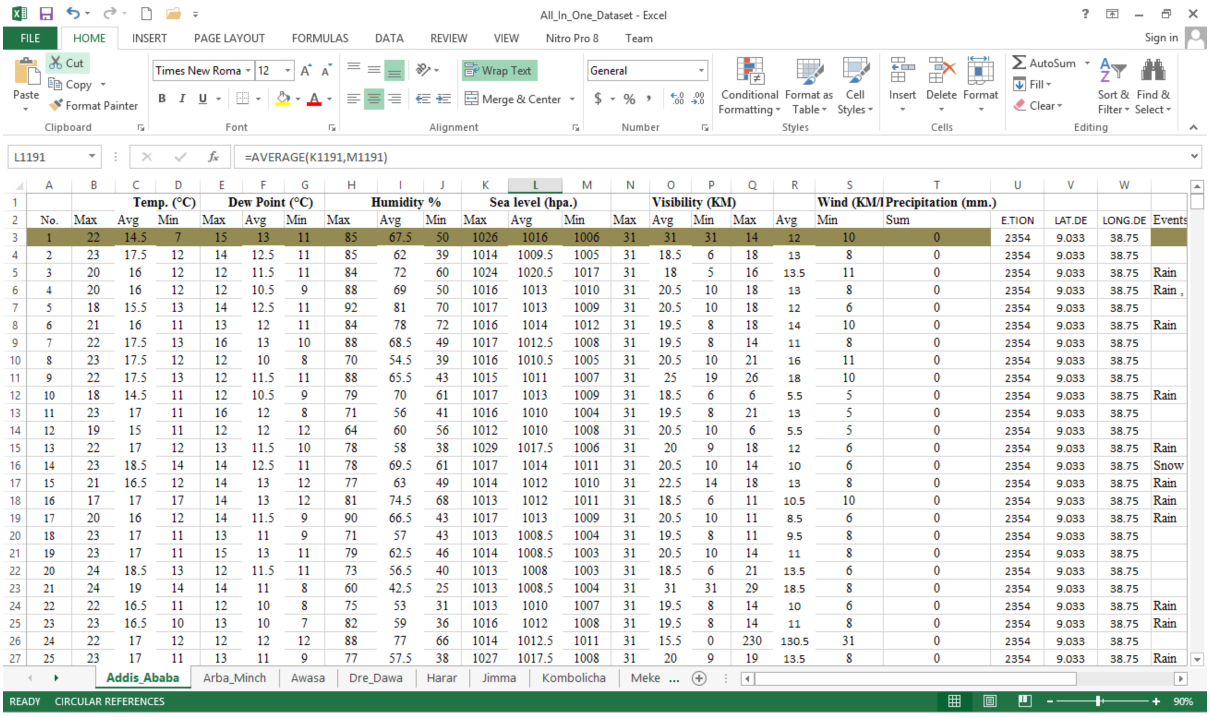
A1 | STN---

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1	STN---	WBAN	YEARMODA	TEMP		DEWP		SLP		STP		VISIB		WDSP		MXSPD	GUST	MAX	MIN	PRCP	SNDP	FRSHT
2	634500	99999	2E+07	59.5	23	48.7	23	1010.2	6	772	6	10	23	7.6	23	15.9	999.9	68.0*	48.2	0.00I	999.9	0
3	634500	99999	2E+07	57.9	16	47.9	16	1009.5	4	771.7	4	9	16	7	16	12	999.9	66.2*	47.7	0.00I	999.9	0
4	634500	99999	2E+07	60.5	22	47.9	22	1009.6	7	772	7	11.3	21	7.3	20	15.9	999.9	72.5	51.8*	0.00I	999.9	0
5	634500	99999	2E+07	60.9	21	46	21	1010.6	7	772.2	7	10.8	21	5.6	20	13.6	999.9	69.8*	47.3	0.00I	999.9	0
6	634500	99999	2E+07	61.3	23	43.8	23	1011.2	4	771.6	4	9	23	7	18	14	999.9	73.8	50.0*	0.00I	999.9	0
7	634500	99999	2E+07	59.8	24	43.7	24	1010.9	8	771.8	8	10.2	24	7.6	24	13.6	999.9	72.3	45.9	0.00I	999.9	0
8	634500	99999	2E+07	58.9	21	43.1	21	1010.5	8	772.1	8	11.2	20	6	20	13.6	999.9	75.7	46.4	0.00I	999.9	0
9	634500	99999	2E+07	59.3	20	43.4	20	1011.6	7	772.1	7	10.8	20	6.7	15	9.9	999.9	75.2	46.4*	0.00I	999.9	0
10	634500	99999	2E+07	60.9	23	41.2	22	1010.8	6	772.2	6	10.4	23	5.8	20	14	999.9	75.4	48.2*	0.00I	999.9	0
11	634500	99999	2E+07	62.4	21	44.1	21	9999.9	0	9999.9	0	8.4	21	6.7	16	9.9	999.9	71.6*	50.0*	0.00I	999.9	0
12	634500	99999	2E+07	60.4	22	46.6	22	1013	6	773.4	6	9.8	22	6.8	22	9.9	999.9	72.3	46.8	0.00I	999.9	0
13	634500	99999	2E+07	61	19	48.5	19	1010.5	7	772.2	7	11	18	6.8	19	14	999.9	74.5	47.1	0.00I	999.9	0
14	634500	99999	2E+07	62.6	19	49.2	19	1008.9	6	771.8	6	10.6	19	7.4	17	19.4	999.9	74.3	48.2	0.00I	999.9	0
15	634500	99999	2E+07	63	22	44.9	22	1007.8	4	771.1	4	9.1	21	5.6	20	9.9	999.9	73.4*	51.8*	0.00I	999.9	0
16	634500	99999	2E+07	65.3	21	48	21	1008.6	6	772.4	6	10.3	21	5.7	17	12	999.9	76.1	49.6	0.00I	999.9	0
17	634500	99999	2E+07	63.1	23	48.3	23	1010.3	7	772.8	7	10.5	23	6	23	11.7	999.9	75.2	49.3	0.00I	999.9	0
18	634500	99999	2E+07	62.3	21	49.3	21	1009.4	6	772	6	10.9	18	7.5	21	14	999.9	73.4*	51.8	0.00I	999.9	0
19	634500	99999	2E+07	64.4	18	49.7	18	9999.9	0	9999.9	0	8.9	18	5.5	13	9.7	999.9	73.4*	52.2	0.00I	999.9	0
20	634500	99999	2E+07	63.7	22	47.3	22	1007.5	4	771.5	4	9.2	21	5.9	18	9.9	999.9	73.8	57.2*	0.00I	999.9	0
21	634500	99999	2E+07	62.2	22	54.3	22	1008.1	5	771.1	5	9.6	22	5.9	20	12	999.9	68.9	55.4*	0.00I	999.9	0
22	634500	99999	2E+07	62.1	23	51.1	23	1009	4	772	4	9	23	7.9	15	15.9	999.9	71.6*	55.4*	0.00I	999.9	0
23	634500	99999	2E+07	61.6	21	49.1	21	1010.4	6	772.5	6	10.9	18	9.9	21	15.5	999.9	71.6*	54.3	0.00I	999.9	0
24	634500	99999	2E+07	61.7	19	52.2	19	1010.2	8	772.5	8	11.9	19	8.3	12	13.6	999.9	75.2	47.3	0.00I	999.9	0
25	634500	99999	2E+07	61	20	52.4	20	1014	4	773	4	9.2	20	10.3	19	18.1	999.9	69.8*	52.5*	0.00I	999.9	0
26	634500	99999	2E+07	62	20	50.8	20	1011.6	5	772.4	6	10.3	20	8.9	14	15.5	999.9	74.1	47.5	0.00I	999.9	0
27	634500	99999	2E+07	62.9	24	50.4	24	1010	5	772.4	5	9.2	24	7.4	20	15.5	999.9	73.8*	52.9*	0.00I	999.9	0

Over_all | Arba_Minch | Awassa | Addis | Mekelle | Kombolicha | Harar | Dre | Jimm ...

READY COUNT: 16 90%

Sample collected daily data set



All_In_One_Dataset - Excel

L1191 | =AVERAGE(K1191,M1191)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	
1			Temp. (°C)			Dew Point (°C)			Humidity %			Sea level (hpa.)			Visibility (KM)				Wind (KM/l	Precipitation (mm.)				
2	No.	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Sum	E.TION	LAT.DE	LONG.DE	Events
3	1	22	14.5	7	15	13	11	85	67.5	50	1026	1016	1006	31	31	31	14	12	10	0	2354	9.033	38.75	
4	2	23	17.5	12	14	12.5	11	85	62	39	1014	1009.5	1005	31	18.5	6	18	13	8	0	2354	9.033	38.75	
5	3	20	16	12	12	11.5	11	84	72	60	1024	1020.5	1017	31	18	5	16	13.5	11	0	2354	9.033	38.75	Rain
6	4	20	16	12	12	10.5	9	88	69	50	1016	1013	1010	31	20.5	10	18	13	8	0	2354	9.033	38.75	Rain ,
7	5	18	15.5	13	14	12.5	11	92	81	70	1017	1013	1009	31	20.5	10	18	12	6	0	2354	9.033	38.75	
8	6	21	16	11	13	12	11	84	78	72	1016	1014	1012	31	19.5	8	18	14	10	0	2354	9.033	38.75	Rain
9	7	22	17.5	13	16	13	10	88	68.5	49	1017	1012.5	1008	31	19.5	8	14	11	8	0	2354	9.033	38.75	
10	8	23	17.5	12	12	10	8	70	54.5	39	1016	1010.5	1005	31	20.5	10	21	16	11	0	2354	9.033	38.75	
11	9	22	17.5	13	12	11.5	11	88	65.5	43	1015	1011	1007	31	25	19	26	18	10	0	2354	9.033	38.75	
12	10	18	14.5	11	12	10.5	9	79	70	61	1017	1013	1009	31	18.5	6	6	5.5	5	0	2354	9.033	38.75	Rain
13	11	23	17	11	16	12	8	71	56	41	1016	1010	1004	31	19.5	8	21	13	5	0	2354	9.033	38.75	
14	12	19	15	11	12	12	12	64	60	56	1012	1010	1008	31	20.5	10	6	5.5	5	0	2354	9.033	38.75	
15	13	22	17	12	13	11.5	10	78	58	38	1029	1017.5	1006	31	20	9	18	12	6	0	2354	9.033	38.75	Rain
16	14	23	18.5	14	14	12.5	11	78	69.5	61	1017	1014	1011	31	20.5	10	14	10	6	0	2354	9.033	38.75	Snow
17	15	21	16.5	12	14	13	12	77	63	49	1014	1012	1010	31	22.5	14	18	13	8	0	2354	9.033	38.75	Rain
18	16	17	17	17	14	13	12	81	74.5	68	1013	1012	1011	31	18.5	6	11	10.5	10	0	2354	9.033	38.75	Rain
19	17	20	16	12	14	11.5	9	90	66.5	43	1017	1013	1009	31	20.5	10	11	8.5	6	0	2354	9.033	38.75	Rain
20	18	23	17	11	13	11	9	71	57	43	1013	1008.5	1004	31	19.5	8	11	9.5	8	0	2354	9.033	38.75	
21	19	23	17	11	15	13	11	79	62.5	46	1014	1008.5	1003	31	20.5	10	14	11	8	0	2354	9.033	38.75	
22	20	24	18.5	13	12	11.5	11	73	56.5	40	1013	1008	1003	31	18.5	6	21	13.5	6	0	2354	9.033	38.75	
23	21	24	19	14	14	11	8	60	42.5	25	1013	1008.5	1004	31	31	31	29	18.5	8	0	2354	9.033	38.75	
24	22	22	16.5	11	12	10	8	75	53	31	1013	1010	1007	31	19.5	8	14	10	6	0	2354	9.033	38.75	Rain
25	23	23	16.5	10	13	10	7	82	59	36	1016	1012	1008	31	19.5	8	14	11	8	0	2354	9.033	38.75	Rain
26	24	22	17	12	12	12	12	88	77	66	1014	1012.5	1011	31	15.5	0	230	130.5	31	0	2354	9.033	38.75	
27	25	23	17	11	13	11	9	77	57.5	38	1027	1017.5	1008	31	20	9	19	13.5	8	0	2354	9.033	38.75	Rain

Addis_Ababa | Arba_Minch | Awasa | Dre_Dawa | Harar | Jimma | Kombolicha | Meke ...

READY CIRCULAR REFERENCES 90%