Sudan University of Science and Technology

College of Engineering

Electrical Engineering

Short-Term Load Forecasting Using Artificial Neural Network Technique

التنبؤ بالأحمال قصيرة المدى باستخدام تقنية الشبكات العصبية الاصطناعية

A project Submitted In Partial Fulfillment for Requirements of the Degree of B.Sc. (Honor) In Electrical Engineering

Prepared By:

1. **AL HUSSAIN MUHAMMAD AL HASSAN AL HUSSAIN**
2. **DAFALLA ELRASHEED ELGAILY DAFALLA**
3. **OSAAMA ABAAS MUHAMMAD ABAAS**
4. **SHAIKH IDRIS JAMAL ALDEEN AL SHAIKH IDRIS**

Supervised By:

Dr. Mohammad Osman Hassan

October 2017

الأية

قال تعالي:

(قُلْ إِنَّمَا أَنَا بَشَرٌ مِثْلُكُمْ يُوحَىٰ إِلَيَّ أَنَّمَا إِلَٰهُكُمْ إِلَٰهٌ وَاحِدٌ ۖ فَمَنْ كَانَ يَرْجُو لِقَاءَ رَبِّهِ فَلْيَعْمَلْ عَمَلًا صَالِحًا وَلَا يُشْرِكْ بِعِبَادَةِ رَبِّهِ أَحَدًا)

[سورة الكهف 110]

DEDIECATION

To the fountain of patience, and optimism, and hope Mother.

To the big heart Father.

To those who have demonstrated to me what is most beautiful in life? Brothers and Sisters.

To the taste of the most beautiful moments with Friends.

To those who taught me how to find them, and taught me as well how not loss them.

ACKNOWLEDGMENT

We would like to express the great appreciation to our Supervisor Dr.Mohammed Osman Hassan for the support and endless supervision throughout the whole period of final year project, his willingness to give his time, and his advices throughout the duration of this project, so generously has been very much appreciated, my gratitude is also extended to all SUST lecturers, without them We wouldn’t be able to acquire such knowledge.

ABSTRACT

This project focused on short-term load forecasting [STLF] in power system operations. Load forecasting is future demand prediction, which assumes an essential part of power system management. Short term load forecasting [STLF] provides load prediction helps in generation scheduling, maintenance, and unit commitment decisions. Therefore, [STLF] plays significant role in power system planning, and the performance of the economic system. This project deal with most power ful Artificial Intelligent [AI] which is Artificial Neural Network [ANN], ANN model designed and compared with one of the statistical methods, which is time series model. MATLAB SIMULINK software is used to accomplish ANN model. This model used Multilayer Feed Forward ANN using MatlabR2016b NN-Tool is trained and examined using data of period from (1/7/2014 to 31/7/2014). At the end, both methods shows that the STLF using artificial neural network [ANN] more accurate than the statistical technique.

المستخلص

يُعنى هذا المشروع بعملية توقع الأحمال الكهربية للمدى القصير في منظومة القدرة الكهربية. و عملية توقع الأحمال هي عبارة عن التنبؤ بالطلب في المستقبل و تعتبر دراسة توقع الأحمال الكهربية واحدة من أهم العمليات التي تتم في منظومة القدرة الكهربية .
حيث تساعد عملية التنبوء بالأحمال للمدى القصير في التزويد بالمعلومات الكافية لجدولة الأحمال , و إجراء عمليات الصيانة كما تلعب دوراً مهماً في التخطيط لمنظومة القدرة الكهربية و تحسين الأداء الإقتصادي للمنظومة , تمت هذه الدراسة بإستخدام الشبكات العصبية الإصطناعية و التي تعتبر واحدة من أفضل تقنيات الذكاء الإصطناعية , تم تصميم نموذج الشبكات العصبية و مقارنته مع إحدى الطرق الإحصائية و هي طريقة السلاسل الزمنية , و تمت محاكاة نموذج الشبكات العصبية بإستخدام لغة البرمجة (ماتلاب) , و ذلك على البيانات في الفترة من (2014/7/1 - 2014/7/31) , و تم عرض النتائج المتحصل عليها بواسطة كل من نموذج الشبكات العصبية الإصطناعية و نموذج السلاسل الزمنية الإحصائي و كانت النتائج المتحصل عليها بواسطة نموذج الشبكات العصبية أكثر دقة من تلك المتحصل عليها بواسطة نموذج السلاسل الزمنية الإحصائية .

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

LF	Load Forecasting
KBES	Knowledge Base Expert System
AI	Artificial Intelligent
ANN	Artificial Neural Network
FLS	Fuzzy Logic System
STLF	Short Term Load Forecasting
MAPE	Mean Absolute Percentage Error
PE	Percentage Error
MTLF	Medium Term Load Forecasting
LTLF	Long Term Load Forecasting
ARMA	Autoregressive Moving Average
ARIMA	Autoregressive Integrated Moving Average
MAM / MA	Moving Average Model
ESM	Exponential Smoothing Model
SEDC	Sudanese Electricity Distribution Company
ME	Mean Error
MAE	Mean Absolute Error
MPE	Mean Percentage Error

CHAPTER ONE
INTRODUCTION

1.1. Overview

In general power system planning work on creating adequate and reliable power supply to reach for the produced load demand, in both near and further future. This is achieving at minimum possible cost keeping the quality supply satisfactory [1]. For distribution system planning the main aim is to achieve a good, and continuity service, this done by determining sizes, locations, and timing of future changes such as additions of system components (substations, lines, feeders, etc.) [2]. So for adequate planning we need to know the future load demand should be produced, to achieve this, the electrical utilities used a technique known as load forecasting [LF]. Load forecasting is a intelligent estimating of the past data and present demand manner to determine the future ones [1]. So planning the distribution system need for prediction of the future load demand in term of system components. The past studies of this term deal with the traditional (mathematical) models, often this models have low accuracy in the results, therefor the researcher spent much effort to create new models with high accuracy which focus in control theory, begin from numerical analysis to low accuracy methods of Artificial Intelligent [AI] until reach nowadays to one of the most powerful [AI] techniques used in studies of load forecasting are artificial neural network [ANN], fuzzy logic system [FL] and expert system [3].

The quality, and accuracy of system planning depends on the quality and accuracy of the data, and load forecast. Which inaccurate forecast lead to many problems such as in case an over prediction of future load will cause premature investment, and unnecessary hugeness of capital cost, whereas under prediction of future load will lead to shortage of equipment this mean increase in maintenance cost, therefore accurate results are very important, and to investigate the optimum situation it is better plan the system for load little higher than the estimated load [1]. So the accurate load forecasting propose efficient solutions of improve the distribution system service.

This project take the topic of short term load forecast by using the one of the modern [AI] techniques is artificial neural network [ANN], because it is simplicity, and high accurate compare to the mathematical models. the study applied on a part of the distribution network in Khartoum city, which is knowing Dar-El Salam substation which supplying activating area with consider the effect of the weather temperature.

1.2. Problem Statement

Nowadays load forecasting becomes one of the most issues in electrical field, it have a important part in power system operation, and planning specially in the operation side and maintenance. But it considered one of the most difficult topic in electrical Field. Firstly, the complexity of the load series, and the peck load during the Seasons. Secondly, the load forecasted uses the historical data of the load it-self, and other elements need to be considered [3].

The overview illustrate that the forecasting must be has high accuracy, to avoid either two problems (the produced electricity not enough for the actual demand, and premature investment without interest). Sudan consider one of the outgrowth countries which can’t handling these two problems, therefor the requesting of the accurate models of load forecasting increase, which help utilities to make a right decisions.

The most common methods to forecast the load can be divided into two methods, which are traditional (Statistical) methods, and artificial intelligent methods. The traditional methods depend on the statistical techniques, which have some weakness points and disadvantages such as taking a lot of time of modeling compare to other techniques, mathematical part very complex, and low accuracy. And the idea behind some of these techniques is considering the load has a linier nature, but the load usually nonlinear nature.

On the other hand , the artificial approach have been proved as efficient for short term load forecasting [STLF] that not depend on the human experience [3] .

Therefor this project deal with one of the artificial approach is [ANN] to passing the difficult, and weakness points which occur when the traditional methods were used.

1.3. Objectives

The objectives of this project are :

- Suggest the ANN model to predict the peak load for month by suggesting the input variables which related with the load , and then specify the construction of the model for the network (the number of layer and number of neurons in hidden layer) .
- Train, test the ANN model and show it's results .
- Apply the time series model and show it's results .
- Compare between the two forecasting models by using the coefficient of MAPE .

1.4. Methodology

The aim of this project is obtaining on the model which have high accuracy, fast, simply, and easy to use, this done by suggesting a construction of load forecasting by using ANN, based on specified input variables, and specified requirements of building construction such as number of layer, and number of neurons in each layer. This leads to obtain the good predicting values, which help to take a right decisions. The terms were considering in ANN model are, historical data, and temperature degree. The load curve will use to measure the accuracy between the actual, and forecasted load. the analysis, and simulation would be performed by Matlab simulation software for [ANN] and Excel software for the tradition method .

1.5. Project Layout

Chapter One : represent the background, problem statement, objectives and methodology.

Chapter Two : literature review include on load in forecast, classification of load, type of forecast based on time and forecast methods.

Chapter Three : case study consist of an information about historical data, information about Dar-El Salam substation, illustration for the steps of the time series model and ANN model, and briefly defining about the accuracy of the forecast.

Chapter Four : will be mainly explains about the results and discussion. Based on the results obtained analysis.

Chapter Five : Conclusion and Recommendation.

CHAPTER TWO
LITERATURE REVIEW

2.1. Introduction

This chapter will explain more about the background theories, methods and techniques are related to the project.

2.2. Power and Energy

In forecast technique load was expressed as energy or power (demand). Energy is the amount of work done over a specific time period, so it is measured in megawatt hours (MWH) in other words, how many megawatts (MW) of electricity are used (or produced) over how many hours.

Megawatt hours are standard unit in which power is bought and sold at the whole sale level in contracts between power generators and utility companies.

On the other hand, power is capacity of energy which being used. Megawatt is a unit for measuring power that is equivalent to the energy produced by 10 automobile.

2.3. Classification of Loads

The most common classification of electrical loads follows the billing categories used by the utility companies [4]. Total load demand of an area depend upon its population and the living standards of people. In general types of loads can be divided into the following categories [5].

2.3.1. Domestic Load

The domestic load is defined as the total energy consumed by electrical appliances in the household work. It depends on the living standard, weather and type of residence. The domestic load mainly consists of lights, fans, mixer, heaters, ovens, small pumping motors etc. The domestic loads consume very little power and also independent from frequency this load largely consists lighting, cooling or heating [6].

The residential loads have the most seasonal variations. it is generally seen that the load factor is poor, it may be taken as 10 to 12% [1].

2.3.2. Commercial Load

Commercial load mainly consist of lighting of shops, offices, and many other electrical appliances used in commercial establishments such as market places, restaurants etc [5]. This class of load is spread over more hours of the day compared to residential loads. Commercial loads are not as large as industrial loads. Commercial loads are also characterized by seasonal variations, and the variations are primarily due to the extensive use of air-conditioners and space heaters [1].

2.3.3. Industrial Load

Industrial load consist of small-scale industries, medium-scale industries, large-scale industries, heavy industries and cottage industries [5]. For example small scale industries require load up to 25 KW, medium-scale industries between 25 and 100 KW and heavy industries require load above 500 KW. Industrial load are considered base load that contains little weather dependent variation. The industrial load are composite load. The composite load is function of frequency and voltage and its form a major part of the system load [1].

2.3.4. Agriculture Load

This type of load is mainly motor pump-sets load for irrigation purposes. The load factor of this load is very small e.g. 0.15 - 0.20 [6]. It is a general practice to energize tube–well feeders for 12 hours a day during night. In rural sector in our country the percentage of agricultural load varies between 60% to 80% [1].

2.4. Types of Load Forecast

The chapter one had been defining the load forecasting as a simple process for quantum defining future load. In one fact the long range planner take in his planning 20-30 years forecasts, to investigate the adequate generation, and transmission plans for the decades ahead. And in the other fact the system operator deal with the load minutes or hours ahead, to achieve the reliable operating of system. from the both two facts the daily forecast is important to schedule the generating units for optimum economy. Weekly or monthly load forecast is very important for maintenance schedule [1]. depending upon period of interest, generally the load forecast was classified based on time horizon :

- Short term load forecasting .
- Medium term load forecasting .
- Long term load forecasting .

The classification based on the information used in forecast process, and if the forecasting period (horizon) increase more information must be applied to forecast process .

2.4.1. Short Term Load Forecasting

The period of short term load forecasting [STLF] ranging from a hour to week sometimes up to one year’s less than three months, usually need for historical data and weather information . The [STLF] is important in schedule the generation unit, reliable, and secure operation of power system, economic dispatch, and real time control [7].

2.4.2. Medium Term Load Forecasting

The period of medium term load forecasting [MTLF] ranging from one month to five years and sometimes ten or more years, usually require the historical data, weather, and economy information.

The [MTLF] used by utility to buy enough fuel, maintenance schedule, and find the capacity of distribution, transmission, and generation system [3, 7].

2.4.3. Long Term Load Forecasting

The period of long term load forecasting [LTLF] ranging from five years to twenty years or more. Usually need for weather, economy, demographic, and sometimes land information.

The [LTLF] used by engineer planning to plan the future expansion in the system, also provide an information for company can help on equipment buying [3, 7].

2.5. Methods of Load Forecasting

The methods of load forecast classified to two types are, statistical methods, and artificial intelligent methods.

This project consider some of the common methods used in short term load forecasting, these methods have been updating over the time.

2.6. Statistical Methods

These methods called the quantitative methods which use a statistical tools. In these methods had two types of models:

1. Explanatory models: which depend on the relationship between the variable to be forecasted and one or more of independent variables.

2. Time Series models: the forecasted variable depend on the past values and no relation between these variables.

The next methods are an examples of the above two models:

2.6.1. Regression Methods

The idea behind this methods is the forecast values consider a function of a certain number of factors that's affect it is output [8]. Therefore there are two types of variables which are dependent variables and independent variables.

Statistical regression technique can be further classified into the simple linear regression model, and the multiple regression models.

In the simple linear regression model, the forecast, and predictor variables are assumed to be related linearly with each other. Was shown in equation (2.1)

$$Y = \beta_o + \beta_1 X + \varepsilon_1 \quad (2.1)$$

Where:

Y : is a dependent variable .

X : is an independent variable .

$\varepsilon 1$: the deviation from the underlying straight line model (error) .

βo , $\beta 1$: are parameters determine the intercept and the slope of the line .

The multiple variable regression model is similar to simple linear model yet it involves multiple variables. The general form of multiple variables regression is given in equation (2.2) below [9].

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots \beta_k X_k + e_i \quad (2.2)$$

Where :

Y : is a dependent variable .

X_1 , X_2 , X_k : is a independent variables .

β_0 , β_1 , β_k : are parameters gives the effect of each predictor after taking account of the effect of all other predictors in the model .

e_i : represents the error involved during forecasting .

2.6.2. Time Series Methods

In previous talk about the time series model, its forecast is just based on the past values, and no functional relationship is established. It is Unnecessary to know the relation between variables. The method treated as black box [8].

This technique is based on the concept that data of the load usually has an internal structure, which contain of an, seasonal variation, trend, and random. In analysis the time series the above factors are important [7].

1. **Seasonal Variation [S]**

   It is defined as estimated, and repetitive pattern around the trend line in time period (it may be days, weeks, months and years). by this factor the forecast can simulation these variations to isolate, and identify seasonal variation, and random fluctuation on the time series [7].

2. **Trend Factor [T]**

This factor is observed over the time period, and any changes related in the time are noted, and calculated, and the trend of these changes is established. There a many type of trend, the series may be increasing or decreasing at various rates [7].

3. **Irregular (Random) Variation [I]**

The variation are accidental, these variations are wholly unpredictable and generated by abnormal stations for example: flood, war, strike, and sudden changes in demand etc. [7].

The time series methods explore, and detect the internal structure, and utilized the structure in electric load forecasting. There are many models of time series method such as :

1. **ARMA and ARIMA Models**

[ARMA] or Autoregressive Moving Average models are at most times utilized for stationary procedures, and are one of the popular linear models in time series forecasting during the past three to four decades, this model applied on stationary process [9].

[ARIMA]or Autoregressive Integrated Moving Average model is an addition to [ARMA], but it also involves non-stationary processes this were popularized by George Box, and Gwilym Jenkins in 1970 [9].

2. **Moving-Average Model [MAM]**

The idea behind this is that the prediction value affected by the recent past data. We calculate the mean by taking only N latest data.

$$At = (xt + xt - 1 + xt - 2 + \cdots + xt - n + xt - n + 1) / n \qquad (2.3)$$

where x_t represents the observation made in period [t], At denotes the moving average calculated after making the observation in period(t).

3. The Exponential Smoothing Model [ESM] and Holt's Method

The Exponential Smoothing Model [ESM] weights recent observations more than older ones, and be written in following mathematical form :

$$\text{forecast } Ft = St = \alpha xt + (1 - \alpha)St - 1 \qquad (2.4)$$

where α is the smoothing constant having values between zero, and one whereas St is the smoothed value of the observations based on our “ best guess ” as to the value of the mean.

2.7. Artificial Intelligent Methods

Artificial intelligent approaches which are include the following techniques :

- Knowledge-based expert system [KBES] model .
- Artificial neural networks [ANNs] model .
- Fuzzy inference system [FLS] model .

2.7.1. Expert system

Expert systems are a type of knowledge based system designed to embody expertise in a particular specialized domain. It use began in the 1960’s for such applications as geological prospecting, and computer design. Expert systems work best when a human expert is available to work with software developers for a considerable amount of time in imparting the expert's knowledge to the expert system software. Also, an expert's knowledge must be appropriate for codification into software rules. An expert system may codify up to hundreds or thousands of production rules.

2.7.2. Fuzzy Logic system

Fuzzy logic [FLS] is the technique of thinking that resemble human thinking. Fuzzy logic is a logic having many values, approximate reasoning, and have a vague boundary. The variables in fuzzy logic system may have any value in between 0 and 1 and hence this type of logic system is able to address the values of the variables (called linguistic variables) those lie between completely truths, and completely false. Each linguistic variable is described by a membership function which has a certain degree of membership at a particular instance.

- **Main parts of fuzzy system**

1. Fuzzification : It transforms the system inputs, which are crisp numbers, into fuzzy sets.

2. Knowledge Base : It stores IF-THEN rules provided by experts.

3. Inference Engine : It simulates the human reasoning process by making fuzzy inference on the inputs and IF-THEN rules

4. Defuzzification : It transforms the fuzzy set obtained by the inference engine into a crisp value.

2.7.3 Artificial Neural Network [ANN]

Artificial Neural Network [ANN] is An extremely simplified model of the human brain. Human brain is extremely complex, parallel information processing system, and nonlinear. Also the computing process in the human brain is different from the conventional digital computer. It has the ability to manage its structural constituent, known as neurons, so it can perform computations and calculation faster than the digital compute .

The human brain can routinely performs recognition tasks like the face recognizing in approximately 100-200 ms, where same task by a conventional computer it might take days to be done [3].

- **The Main Advantages of ANN**

1. Adaptive learning : An ability to learn how to do tasks based on the data given for training or initial experience .
2. Self-Organization : An ANN can create its own organization or representation of the information it receives during learning time [10].

- **The Application of ANN**

1. Classification : Pattern recognition, feature extraction, and image matching .
2. Noise Reduction : Recognize patterns in the inputs, and produce noiseless outputs.
3. Prediction : Extrapolation based on historical data [10] .

The human cortex have approximately 10 billion neurons, and each biological neuron is linked to some thousands of other neurons. Speed of biological neurons is measured in milliseconds.

The idea of artificial neural networks is based on the working principles of the human brain. The connection made to resemble the neurons and dendrites in the human brain by using silicon, and wires. The human brain combined hundreds billion of nerve cells named neurons. These neurons are connected to cells by axons.

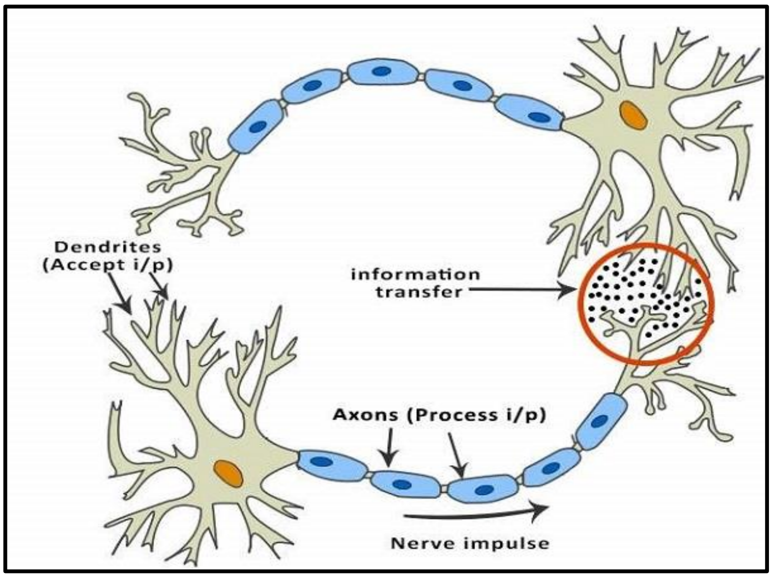

Figure (2.1): Biological neuron

Biological Neuron starts at the sensory organs which used to inputs. These inputs will change to electric impulses to transfer through the axons or the neural network as in ANN [3].

♦ Types of Neural Network

The neural network classified based on the layer organization, neurons are organized in different form of layers. These types can be outlined as follow:

1. Single-layer Feed Forward Networks

   In single-layer network, consists of one input layer, and one output layer. Since there is no mathematical operation in the input layer it is not count as layer [3]. figure (2.2) below shows single-layer feed forward network .

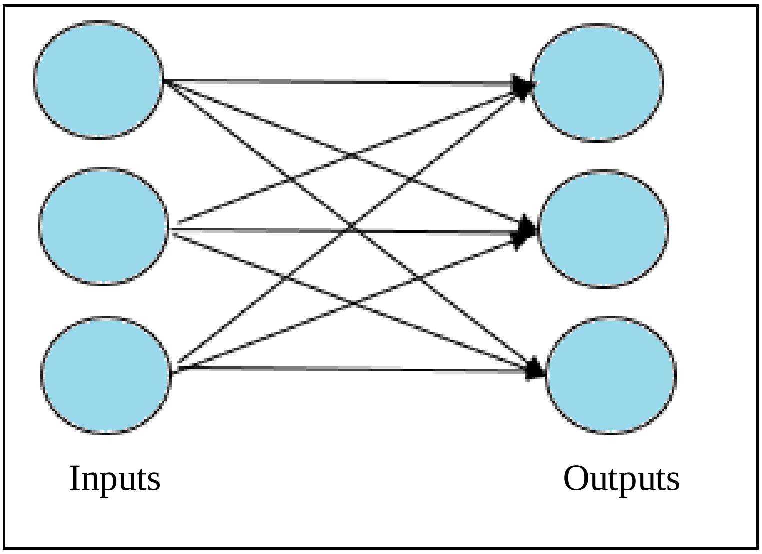

Figure (2.2) : Single layer feed forward

2. Recurrent Networks

consists of input layer with ability to Self-feedback form the output layer, hidden layer will handle all the mathematical operation, output layer to gain the results from it, and at least one feedback loop or more [3]. Figure shows multilayer feed forward networks figure (2.3) Recurrent Network.

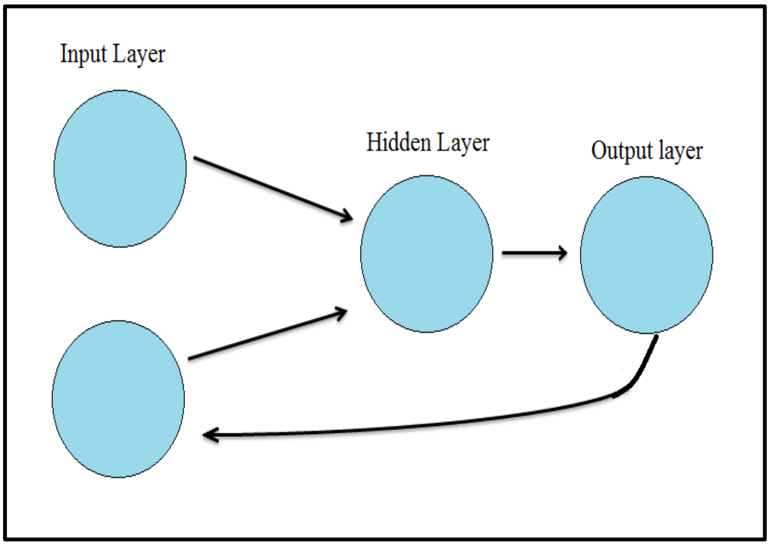

Figure (2.3) : Recurrent network

3. Multilayer Feed Forward Networks

consists of three layers input layer and hidden layer where all mathematical operation and output layer to show the results of the process [3]. Figure shows multilayer feed forward networks.

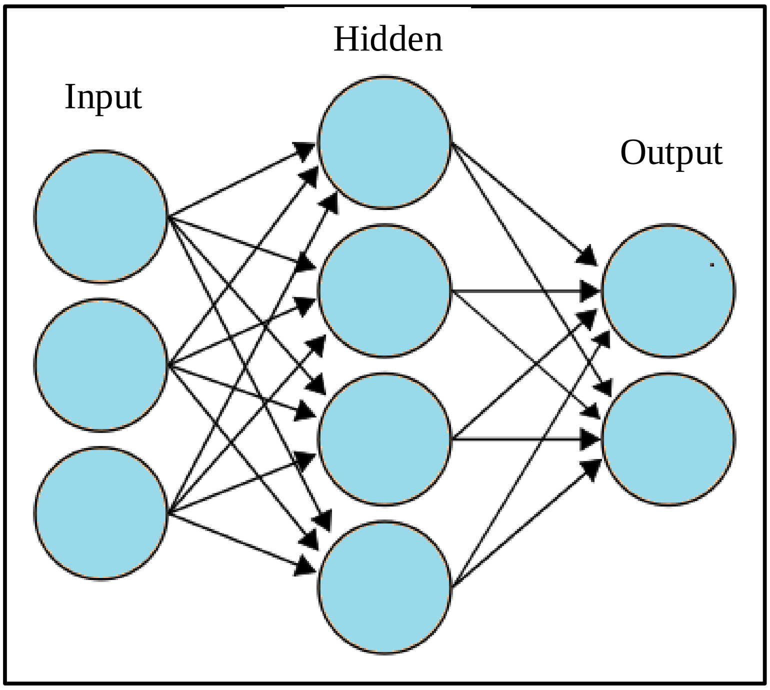

Figure (2.4) : Multilayer feed forward networks

♦ Learning Process

By learning rule a procedure was meant for modifying the weights, and biases of a network. The purpose of learning rule is to train the network to perform some task. They fall into three broad categories :

1. Supervised Learning

The learning rule is provided with a set of training data of proper network behavior. As the inputs are applied to the network, the network outputs are compared to the targets. The learning rule is then used to adjust the weights, and biases of the network in order to move the network outputs closer to the targets [11].

2. Reinforcement Learning

It is similar to supervised learning, except that, instead of being provided with the correct output for each network input, the algorithm is only given a grade. The grade is a measure of the network performance over some sequence of inputs [11].

3. Unsupervised Learning

The weights, and biases are modified in response to network inputs only. There are no target outputs available. Most of these algorithms perform some kind of clustering operation. They learn to categorize the input patterns into a finite number of classes [11].

CHAPTER THREE
CASE STUDY

3.1. Introduction

This chapter discussed the methodology of the project which consists of several parts. First part discussed the process, and the procedures of the project in general while, the second part discussed, and explained the models that been used to forecasts the load in details.

The data have been taking from a part of distribution networks of The Sudanese Electricity Distribution Company. Ltd [SEDC], this part is knowing Dar-El Salam distribution substation, the data collected in time period from 1/7/2014 to 31/7/2014, and that by taking peak value of load demand for each day of this period, also collected the maximum, and the average of the temperature degrees for Khartoum state in this period from https://www.wunderground.com/history/airport/HSSS.html .

3.2. Dar-El Salam Substation

Dar-El Salam distribution substation located in Omdurman city, it feeds area consist of domestic, and commercial loads, these loads are representing in ELMAHKIM, NEW BLOAKS, 22DAR-EL SALAM, SOUK LIYBIA, SUDATEL, and MOSTISHFA-UMBDA.

3.2.1 The Main Components of The Substation

1. Input bus bar have rated of 33KV.

2. Feeder network had six feeders shown on table 3.1.

3. Two transformers have rated of 22KVA .

4. Output bus bar have rated of 11KV .

Table 3.1 : Feeder line currents

	Line Name	Line Current
DAR EL SALAM	EL MHAKEM	230
	NEW BLOCKs	358
	22DARELSALAM	292
	SOUK LIBYA	366
	SUDATEL	228
	MOSTISHFA UMBDA	248

The figure (3.1) show the main components of the substation in the single line diagram.

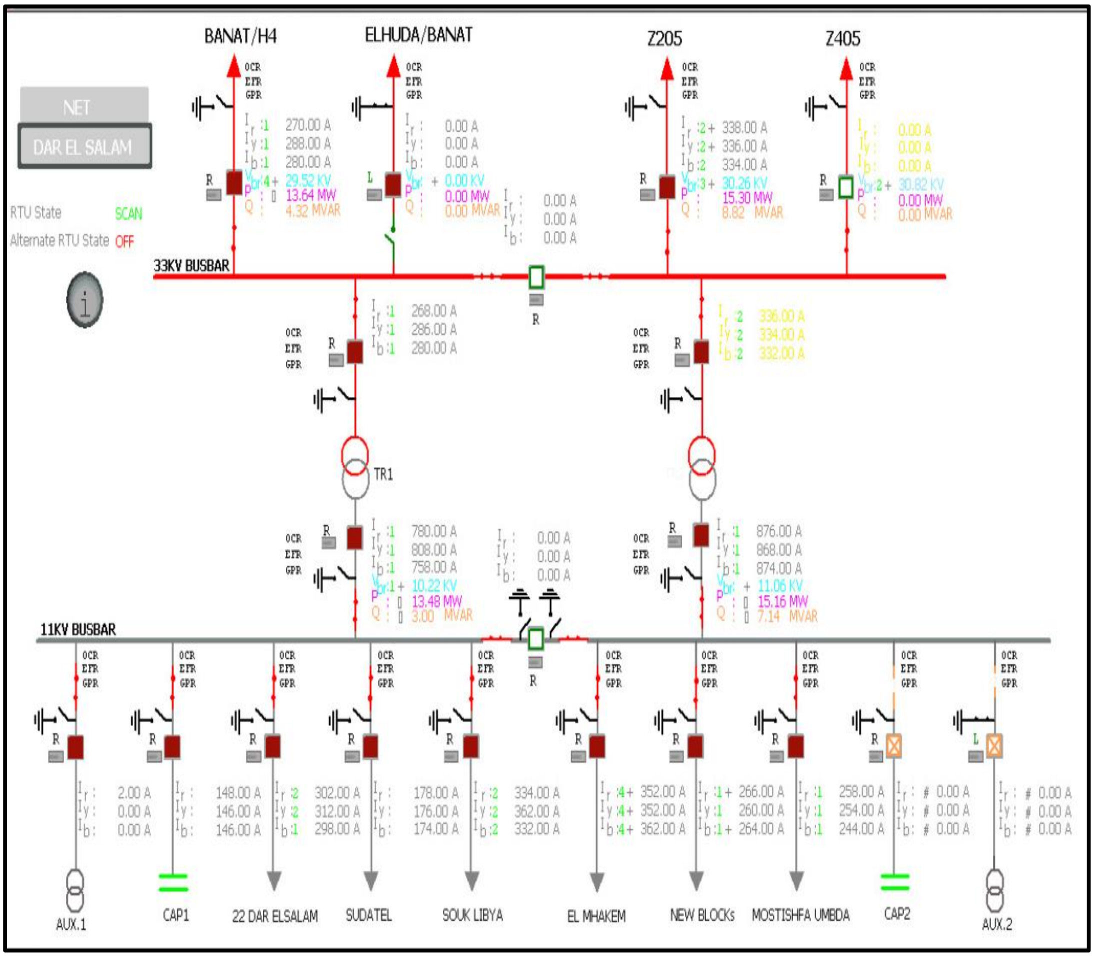

Figure (3.1) : Dar-El Salam single line diagram

3.3. Time Series Predicted Model

The load in time series model express by the equation (3.1)

$$Y = T * S * I \quad (3.1)$$

Where :

Y = the load .

T = trend factor .

S = seasonal variation .

I = random variation [8].

To make a prediction by this model, the following steps will be applied [7]:

1. Step One

Specify the input variables, these variables were :

1- The peak load of the day from (1/7/2014 to 31/7/2014) .

2- The time variable (Days of Month) .

2. Step Two

Trend analysis, this step down by considering the time series has linear trend, therefor the regression analysis will be applied to predict the trend factor and use the trends into forecasting the future values of the variables unvalued. The model have a non-linear relationship between the independent variable (time t), and dependent variables (peak load Y) was

$$Ft = a * e^{b*t} \quad (3.2)$$

Where :

F_t = forecasted value of the trend analysis in period (t) .

a = intercept of the trend line in y-axis .

b = slop of the trend line .

t = time period .

to make the equation (3.2) in liner form the simple regression will be employed this done by taking the natural logarithm to yield

$$\ln Ft = \ln a + b * t \ln e$$

But because ln e =1

$$\ln Ft = \ln a + b * t \qquad (3.3)$$

The equation (3.3) converted to (3.4)

$$M = Z + b * t \qquad (3.4)$$

The value of the two constants can be calculated by using the two relationships

$$b = \frac{n \sum t * M - \sum tM}{n \sum t^2 - (\sum t)^2} \qquad (3.3)$$

$$Z = Mavg - b * tavg \qquad (3.4)$$

Where :

Ft = the actual value of the trend analysis in period (t) .

n = number of period .

M = ln F_t .

Z = ln a .

M_{avg} = average value $= \frac{\sum M}{n}$

t_{avg} = average value of time variable = $\frac{\sum t}{n}$

the values of the constant (a) and (b) are :

a = 900.49

b = 0.0027

3. Step Three

Use moving average to make a smoothing for the time series to reduce the effect of the fluctuation in data by taking MA(6), MA(4), and MA(2) to reach for suitable data .

4. Step Four

Extract the S*I component from the data. and this done by

$$S * I = \frac{Ft}{MA(2)} \qquad (3.5)$$

Where :

F_t = the peak load .

MA(2) = the last moving average value .

5. Step Five

Extract the S component (the seasonal index) .

6. Step Six

Make the deseason value, this value without the season and random components , this done by

$$\text{deseason} = Ft * S \qquad (3.6)$$

And from the deseason value the error was calculated

$$\text{error} = y - \text{deseason} \qquad (3.7)$$

where :

y = the peak load value

Then calculate Mean error .

$$\text{Mean error} = \frac{\sum \text{error}}{\text{n}} \qquad (3.8)$$

Where :

n = number of period .

7. **Step Seven**

Now the forecasted value calculated by

$$\text{Yt} = \text{deseason} + \text{Mean error} \qquad (3.9)$$

Where :

Y_t = the forecasted value

3.4. ANN Predicted Model

A broad spectrum of factors affect the system's load level such as trend effects, cyclic-time effects, weather effects, and random effects like human activities. This project was developed a system that predicted one month ahead. The inputs were taken the past month peak load, and the day of the month. Dar-El Salam was chosen for Omdurman region, and used the daily temperature as input parameters. The city chosen was the major city in Omdurman governorate, and as such gave sufficient representation to the change in weather parameter across the state. The inputs were fed into Artificial Neural Network [ANN], and after sufficient training were used to predict the peak load for the next month.

3.4.1. The Selected Inputs

The inputs are using in the ANN model indicated below :

1. Daily peak load of the past month .
2. Day of the week (7 days).
3. Work day / holiday day.

4. Max/ Average daily temperature of the past month (July month)
5. Max/Average predict daily temperature of the forecasted month (August month) .
6. Daily average load of days of the past month.
7. Previous hour load from the peak load.
8. Date of the day.

And the output obtained was the predicted peak load for the next month. The input matrix was prepared by using Excel program before used them in Matlab, the input matrix consist of load matrix (31 x 3), which had the peak load, the average load of the day values, and the previous hour load form the peak load. The temperature matrix (31 x4) had the average, and max values for both the predicted, and the past data, the day of the week matrix (31 x1), and the values of this matrix from 1 to 7, type of the day matrix (is a working day ?) (31 x 1), and the values of this matrix from 0 for the working day to 1 for the holiday, and the serial date matrix (31 x 1). The serial date is a Microsoft Excel® numerical representation of the calendar date as a total number of days starting with January 1, 1900 as serial date 1. Because the MATLAB® serial date system starts on January 0, 0000 which is different from the starting date in Excel® which is January 1, 1900. For example the value of, July1, 2014 is equivalent to serial date 41821.

The block diagram of ANN model is shown in figure (3.2)

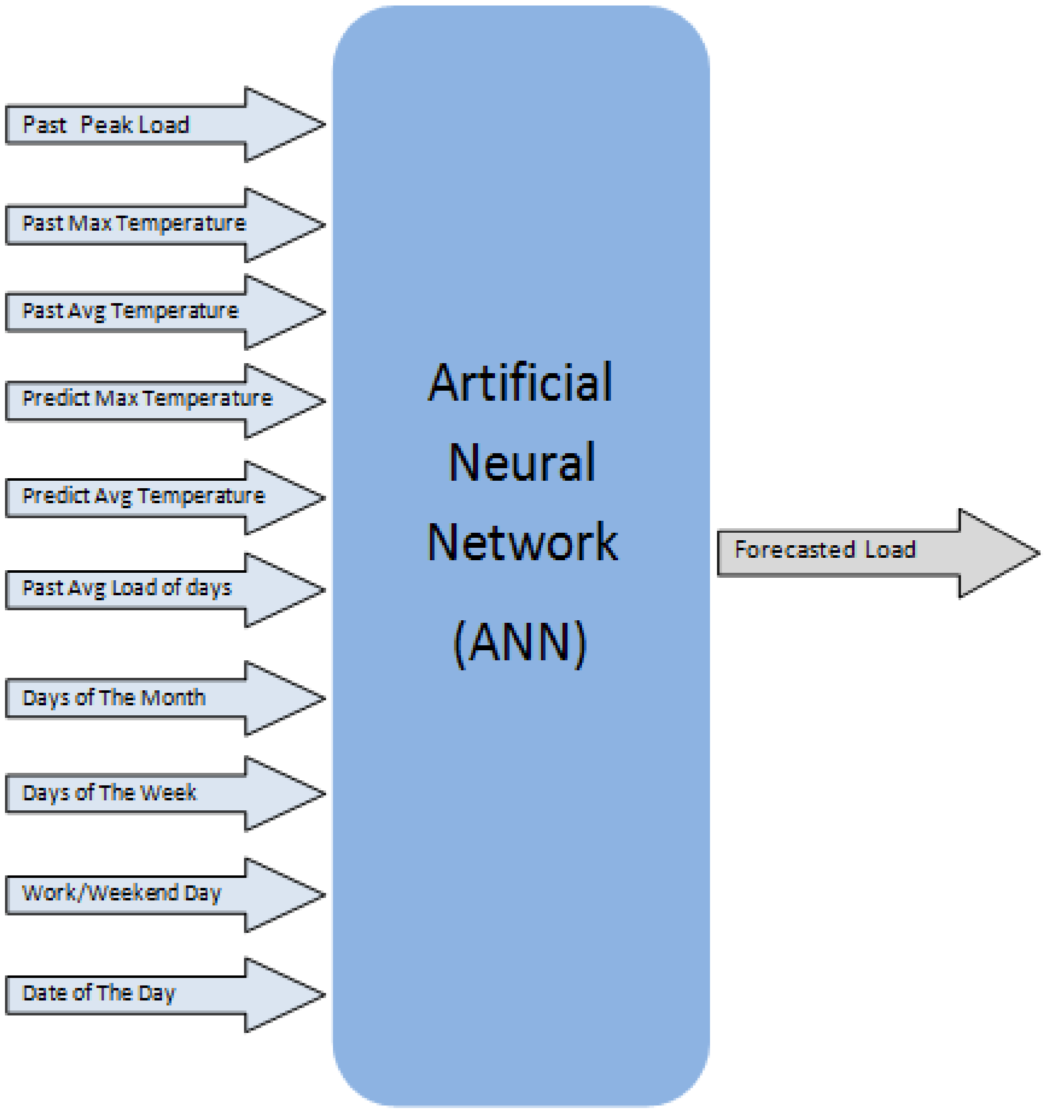

Figure (3.2) : ANN model block diagram

3.4.2. ANN Model Construction

For Neural Network model a Multi-Layer Perception [MLP] network was used with a single hidden layer. The number of neurons in the hidden layer was varied between 9 and 16 before finally being set at 14 neurons . The activation function used in the hidden layer neuron was "Tan-sigmoid ".

Figure (3.3) shows the NN-TOOL graphical user interface consists of three layers, input layer with ten input, hidden layer with fourteen neurons, and one output layer to obtained the forecasted load .

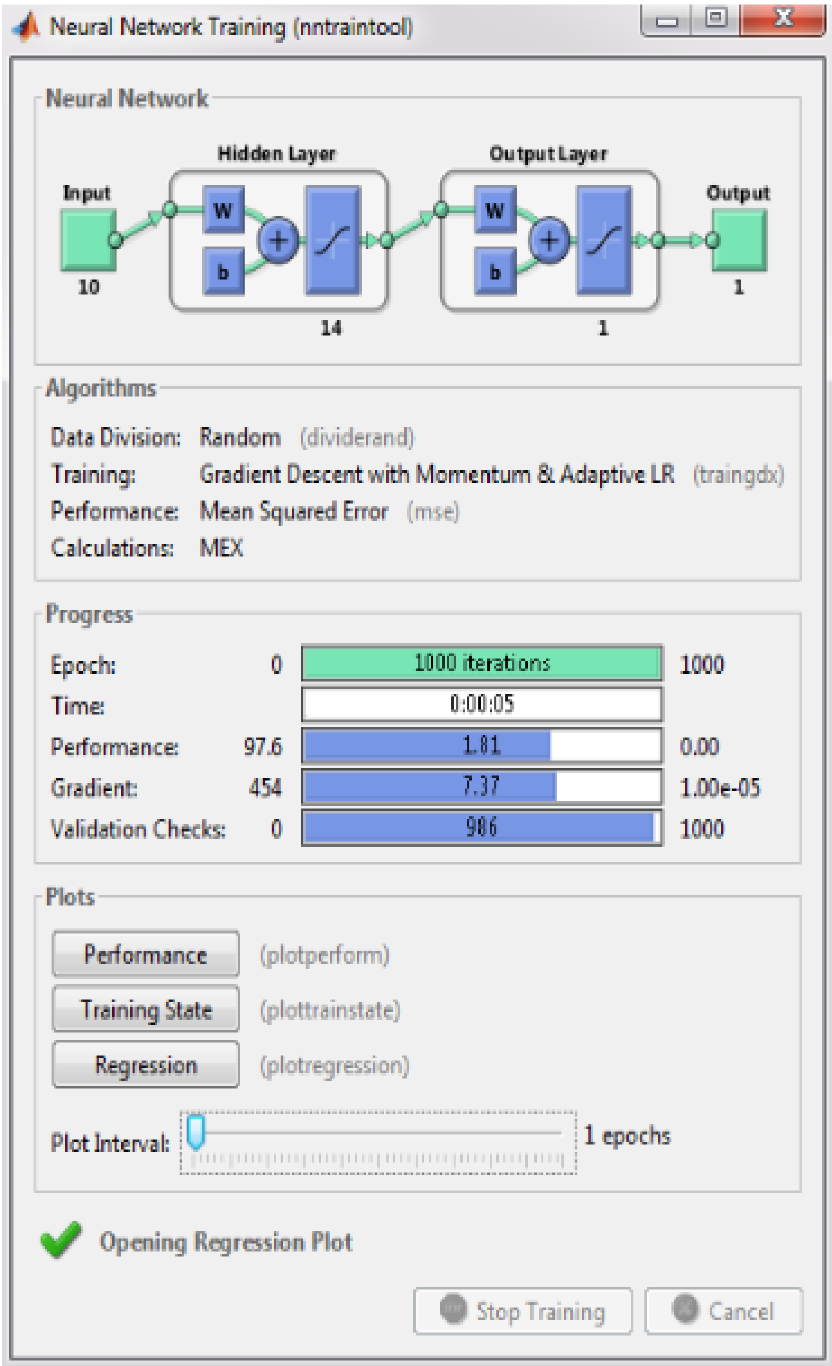

Figure (3.3) : NN-Tool

The flowchart of the project is shown below in figure (3.4)

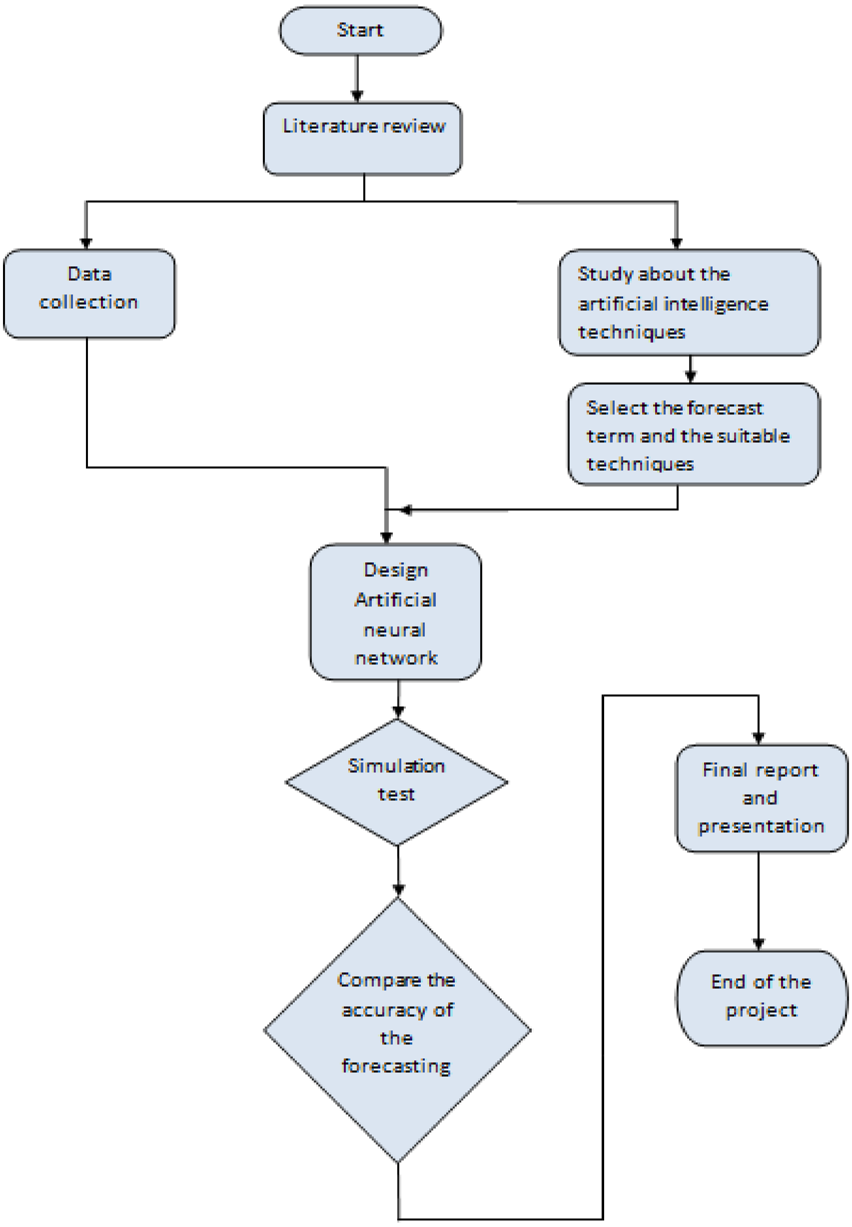

Figure (3.4) : Flowchart of the project

3.5. The Forecast Accuracy

Accuracy is overriding criterion for selecting the particular forecast model [8]. The forecast model use to estimate data already known to predict the accuracy of the model if (Y_t) is the observed value for time (t), and (F_t) is the forecast value for the period (t) , the error defined by

$$e = Yt - Ft \quad (3.9)$$

Some standard statistical measures used to measure the accuracy of the method are

$$\text{Mean error [ME]} = \frac{1}{n} * \sum_{t=1}^{n} et \quad (3.10)$$

$$\text{Mean absolute error [MAE]} = \frac{1}{n} * \sum_{t=1}^{n} |et| \quad (3.11)$$

$$\text{Percentage error [PE]} = \frac{(Yt - Ft)}{Yt} * 100\% \quad (3.12)$$

$$\text{Mean percentage error [MPE]} = \frac{1}{n} * \sum_{t=1}^{n} PE \quad (3.13)$$

$$\text{Mean absolute percentage error [MAPE]} = \frac{1}{n} * \sum_{t=1}^{n} |PE| \quad (3.14)$$

The above measurements of error were used to quantify the accuracy of the forecast model. The MAPE measurement is the best one of these measurements, it is giving fairly good idea about the accuracy of the model, since a percentage error will be taken [8].

This project consider MAPE, and PE as the accuracy measurements for the two forecasted models .

CHAPTER FOUR
RESULTS AND DISCUSSION

4.1. Introduction

This project focused on artificial intelligence model. One statistical approach built to compare the accuracy between the two models using mean absolute percentage error [MAPE]. All the operations, and procedures, to obtain the results of these models are attained via simulation by using Matlab software, and Excel software**.**

4.2. The Results of Time Series Model

In this model the matrix input was (31 x 2), which consist the peak load, and the time factor (day of month) samples, the results shown in table 4.1 which represent the daily peak load for August 2014, and have the predicted values, and the PE column.

Table 4.1 : Time series model

Date	Day	Actual Peak load in KW	Time Series forecast in KW	Time Series PE
01/08/2014	Fri	806	793.1474	1.5946
02/08/2014	Sat	931	879.7795	5.5017
03/08/2014	Sun	872	900.7046	-3.2918
04/08/2014	Mon	935	923.5320	1.2265
05/08/2014	Tue	991	986.4215	0.4620
06/08/2014	Wed	993	980.9954	1.2089
07/08/2014	Thu	980	975.7774	0.4309
08/08/2014	Fri	847	764.4452	9.7467
09/08/2014	Sat	809	832.4490	-2.8985
10/08/2014	Sun	914	930.8851	-1.8474

11/08/2014	Mon	905	917.1870	-1.3466
12/08/2014	Tue	949	972.6642	-2.4936
13/08/2014	Wed	1038	1013.4451	2.3656
14/08/2014	Thu	977	982.9493	-0.6089
15/08/2014	Fri	863	778.9729	9.7366
16/08/2014	Sat	871	869.6934	0.1500
17/08/2014	Sun	975	970.9661	0.4137
18/08/2014	Mon	951	949.0807	0.2018
19/08/2014	Tue	1048	1034.4899	1.2891
20/08/2014	Wed	936	968.5797	-3.4807
21/08/2014	Thu	1002	1005.0581	-0.3052
22/08/2014	Fri	977	828.8844	15.1602
23/08/2014	Sat	899	891.5082	0.8333
24/08/2014	Sun	999	992.2574	0.6749
25/08/2014	Mon	975	970.1222	0.5003
26/08/2014	Tue	1034	1036.0445	-0.1977
27/08/2014	Wed	1019	1021.3951	-0.2350
28/08/2014	Thu	1009	1017.8423	-0.8763
29/08/2014	Fri	844	790.2072	6.3736
30/08/2014	Sat	922	911.2004	1.1713
31/08/2014	Sun	970	986.5639	-1.7076

Form the above table the values of the time series model was indicating, The results show the variation between the actual, and forecasted values. Because this model of time series models deal with two variables were the peak load, and time factor, which neglected many factors such as the weather factor, type of the day (working day / holiday), and day of week, therefor in some points the value of forecast not satisfied, and this indicted by PE column, the negative sign (-) in PE produced in equation (3.13), which indicated on the high value of forecasted peak load than actual value. The model show a behavior of the load depending on its technique, The model couldn't simulate the behavior of the actual load in right form, because there are a missing factors not considered in this model which had effect on load behavior, this shown in figure (4.1) which illustrate the results in table 4.1, and show the comparison between the actual values, and the predicted values.

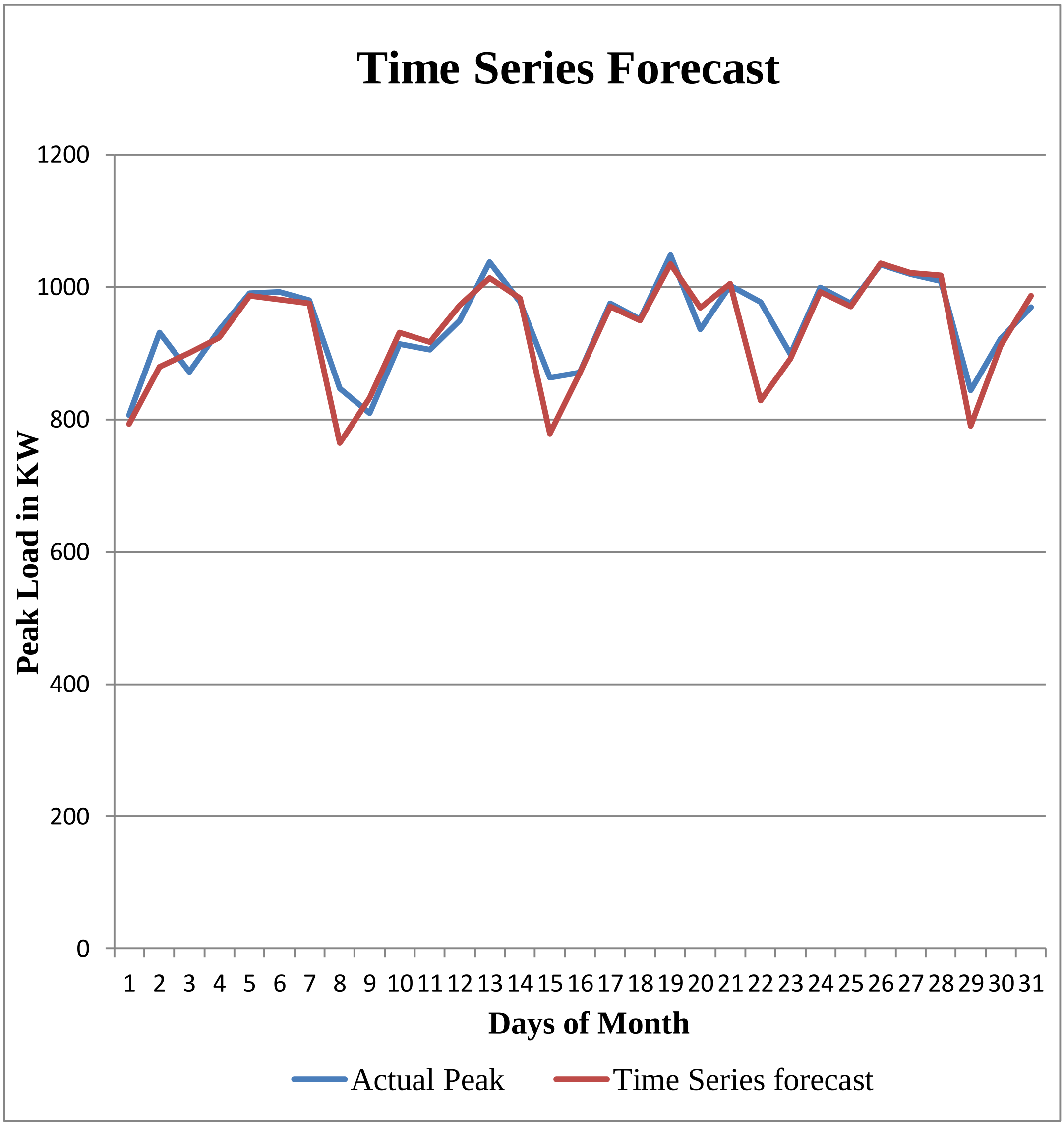

Figure (4.1) : Load curve of time series model

The figure (4.1) represent the previous discussion about the time series model, the high deviation in some points of the load curve lead to mismatch between the two curves. To obtain the accuracy measurement of this model The PE values were used to get on the MAPE value. The PE range of the time series model from (-3.4807% to 15.1602%). The MAPE value calculated by using equation (3.14), and the value was 2.5268%. In general the value of the predicted model not effected by the historical values (peak load), and time factor only, but there are many factors must be considering, to obtain on reliable, and accurate results.

4.3. The Results of Artificial Neurons Network [ANN] Model

In this model the input matrix was (31 x 10), which consisted on 31 samples for the 10 elements, which specified in the construction of the ANN, and matrix (31 x 1) represented the actual load. 21 samples used as training data, 5 samples used as validation, and 5 samples used as testing. The regression R values measure the correlation between the outputs, and targets. All the regression results are very near to one which mean that the data has close relationship, this shown on figure (4.2).

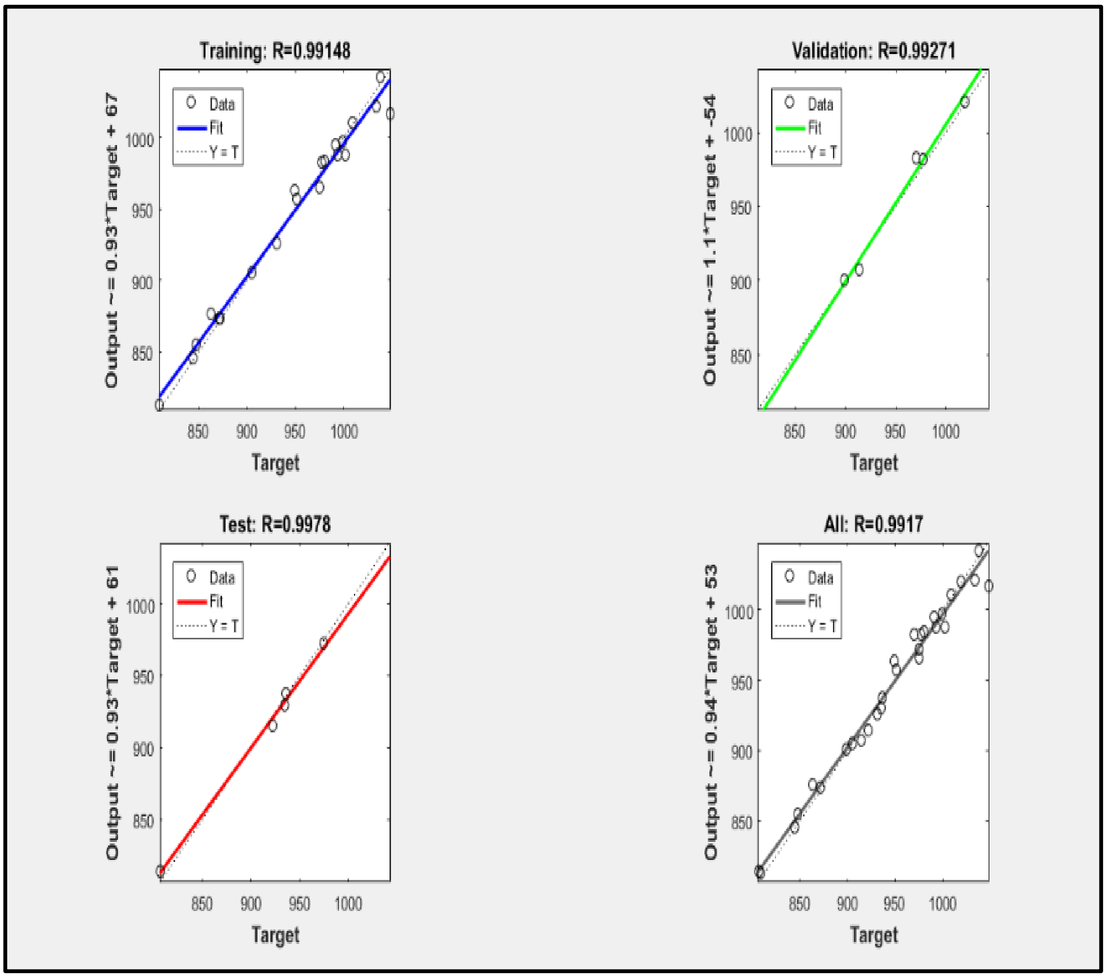

Figure (4.2) : The regression interface window.

The results shown in table 4.2 which represent the day peak load for August 2014, and have the predicted values, and the PE column.

Table 4.2 : ANN model

Date	Day	Actual Peak load in KW	ANN Forecast in KW	ANN PE
01/08/2014	Fri	806	813.6253	-0.9461
02/08/2014	Sat	931	926.2706	0.5080
03/08/2014	Sun	872	873.6487	-0.1891
04/08/2014	Mon	935	929.8417	0.5517
05/08/2014	Tue	991	995.2427	-0.4281
06/08/2014	Wed	993	987.6251	0.5413
07/08/2014	Thu	980	984.1287	-0.4213
08/08/2014	Fri	847	854.7425	-0.9141
09/08/2014	Sat	809	812.4134	-0.4219
10/08/2014	Sun	914	906.9483	0.7715
11/08/2014	Mon	905	904.9021	0.0108
12/08/2014	Tue	949	963.4630	-1.5240
13/08/2014	Wed	1038	1042.7410	-0.4567
14/08/2014	Thu	977	982.5741	-0.5705
15/08/2014	Fri	863	876.0672	-1.5142
16/08/2014	Sat	871	873.5602	-0.2939
17/08/2014	Sun	975	965.2887	0.9960
18/08/2014	Mon	951	957.0091	-0.6319
19/08/2014	Tue	1048	1033.0028	1.4310
20/08/2014	Wed	936	937.5233	-0.1627
21/08/2014	Thu	1002	988.0625	1.3910
22/08/2014	Fri	977	982.0289	-0.5147
23/08/2014	Sat	899	900.7708	-0.1970
24/08/2014	Sun	999	997.3685	0.1633
25/08/2014	Mon	975	972.2748	0.2795
26/08/2014	Tue	1034	1021.5555	1.2035
27/08/2014	Wed	1019	1020.5475	-0.1519
28/08/2014	Thu	1009	1010.4428	-0.1430
29/08/2014	Fri	844	845.7636	-0.2090
30/08/2014	Sat	922	914.8157	0.7792
31/08/2014	Sun	970	982.2074	-1.2585

From the above table the ANN model forecasted results was indicating. The results give a high percentage of match between the ANN values, and the actual ones, and a small deviation in some points was occurred. This cause a small value of PE, which give a good indicating about the high value of accuracy. The pervious problems in time series model was solving in this model, the strong relation between the suggestion inputs, and the output occur. The model give a high degree of simulation of load behavior, the both curves were plotted in figure (4.3) which illustrate the results in table 4.1, and show the comparison between the actual values, and the predicted values.

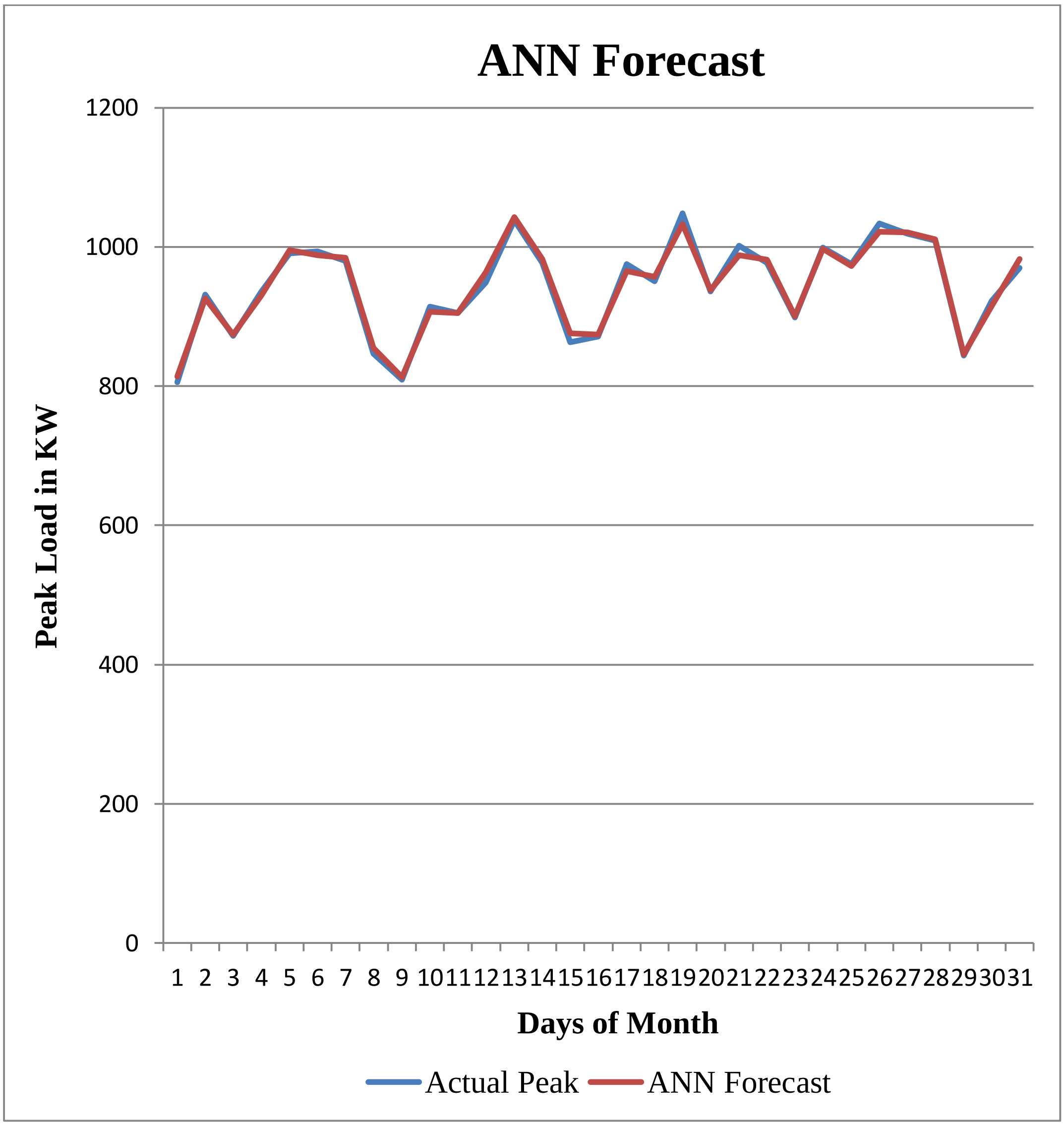

Figure (4.3) : The load curve of ANN model

The figure (4.3) agree with the previous discussion about the ANN model, the identification occur in the most points in load curve, there a small deviation in few points of curve, and the high accuracy occur in the ANN curve, the forecast values very close to the real values. To specify the accuracy of the ANN model use the PE values to obtain on the MAPE value. The PE in range from (-0.9461% to 2.9577%). The value of MAPE calculated by using the equation (3.14), this value was 0.6315% . The value of MAPE of ANN model express on the high value of accuracy.

Table 4.3 shows the forecasting comparison between the time series model. and the artificial neural network model. Based on the results of both models the artificial neural network model has very high accuracy.

Table 4.3 : Daily peak load forecast

Date	Day	Actual Peak load in KW	Time Series forecast in KW	ANN Forecast in KW
01/08/2014	Fri	806	954.9739	813.6253
02/08/2014	Sat	931	905.3658	926.2706
03/08/2014	Sun	872	907.8136	873.6487
04/08/2014	Mon	935	910.2680	929.8417
05/08/2014	Tue	991	912.7290	995.2427
06/08/2014	Wed	993	915.1967	987.6251
07/08/2014	Thu	980	917.6711	984.1287
08/08/2014	Fri	847	920.1522	854.7425
09/08/2014	Sat	809	922.6399	812.4134
10/08/2014	Sun	914	925.1344	906.9483
11/08/2014	Mon	905	927.6357	904.9021
12/08/2014	Tue	949	930.1437	963.463
13/08/2014	Wed	1038	932.6585	1042.741
14/08/2014	Thu	977	935.1800	982.5741
15/08/2014	Fri	863	937.7084	876.0672
16/08/2014	Sat	871	940.2437	873.5602
17/08/2014	Sun	975	942.7858	965.2887
18/08/2014	Mon	951	945.3347	957.0091
19/08/2014	Tue	1048	947.8906	1017.0028
20/08/2014	Wed	936	950.4533	937.5233
21/08/2014	Thu	1002	953.0230	988.0625
22/08/2014	Fri	977	955.5997	982.0289
23/08/2014	Sat	899	958.1833	900.7708
24/08/2014	Sun	999	960.7739	997.3685
25/08/2014	Mon	975	963.3715	972.2748
26/08/2014	Tue	1034	965.9761	1021.5555
27/08/2014	Wed	1019	968.5877	1020.5475
28/08/2014	Thu	1009	971.2064	1010.4428
29/08/2014	Fri	844	973.8322	845.7636
30/08/2014	Sat	922	976.4651	914.8157
31/08/2014	Sun	970	979.1052	982.2074

The figure (4.3) shows the forecasting comparison between the time series model, and the artificial neural network model from the above table.

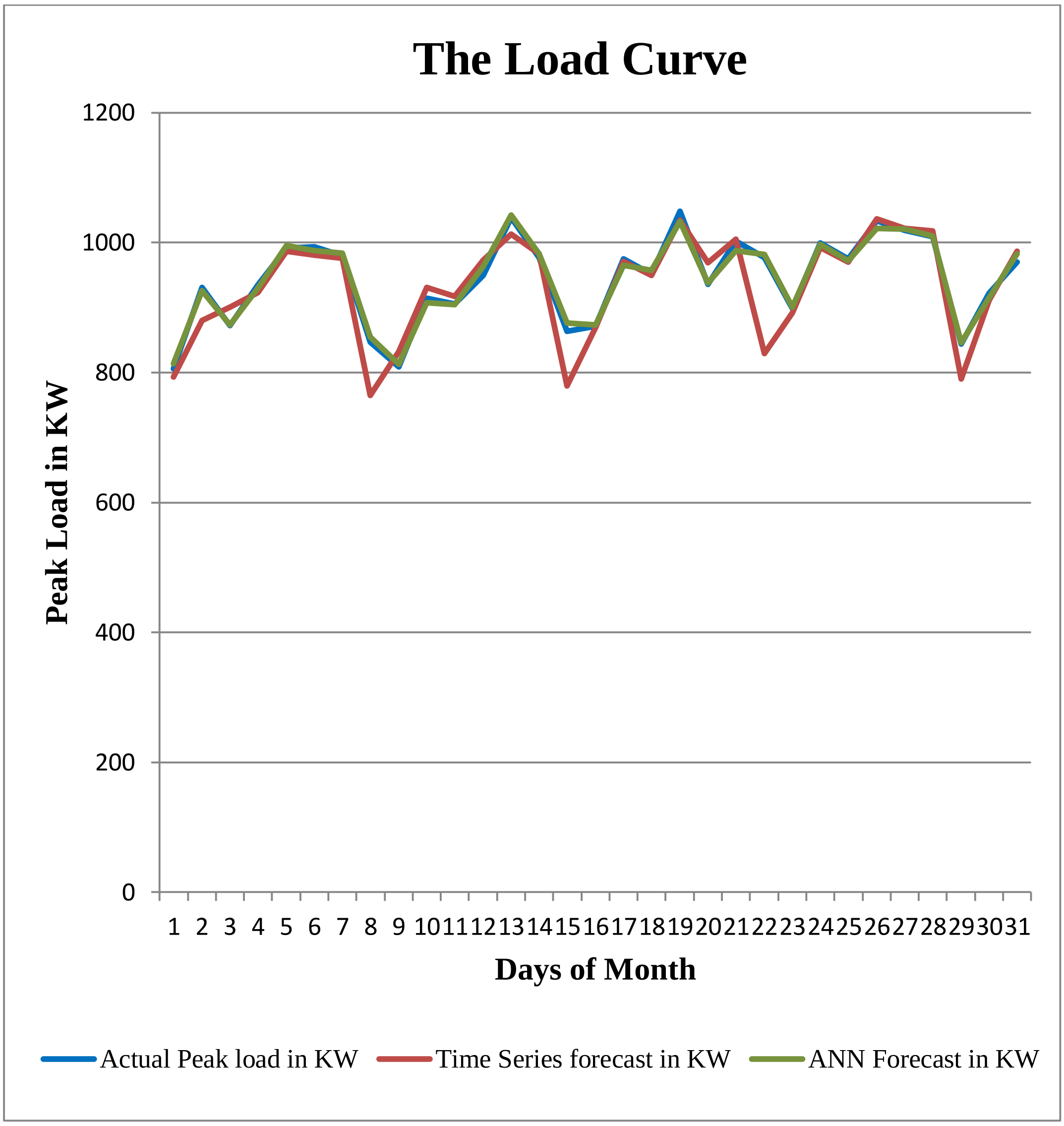

Figure (4.4) : The load curve of both models

For more clearance The figure (4.5) show the difference of PE values of the two models. this figure illustrate the accuracy of the two models, and it is clear the high accuracy occurred in ANN model.

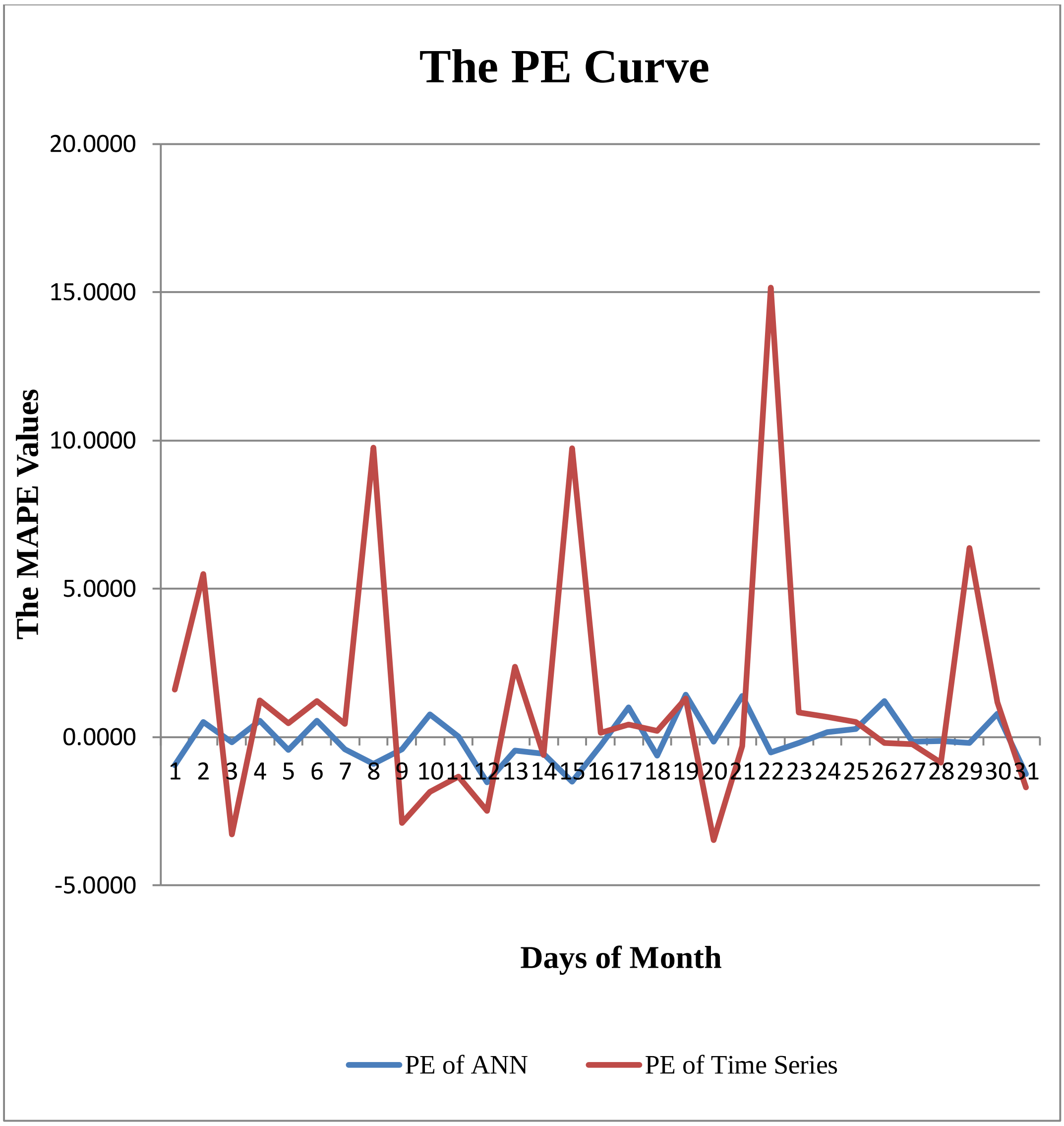

Figure (4.5) : PE curve

CHAPTER FIVE
CONCULSION AND RECOMMENDATION

5.1. Conclusion

In the management, and planning of distribution systems, load forecasting plays significant role. The load forecasting is one of the most critical task in electrical engineering, due to non-linear load's behavior. Besides, load pattern of any particular day, technically depends on several difference factors such as whether variance, and peak load data which can extremely affect the load forecasting.

The results of time series method done by using the Excel software. The range of PE from (-3.4807% to 15.1602%). The MAPE value calculated by using equation (3.14), and the value was 2.5268%, there a high deviations between the actual values, and the predicted values, this deviations occurred due to not considered important factors such as weather features, load behavior, and…etc.

The short load forecasting models designed using ANN method the implementation of the networks architecture, training, and simulation of the results were done using Matlab R2016b, and the load curves were plotting by Excel software. The artificial techniques has very high accurate results with PE range from (-0.9461% to 2.9577%). The value of MAPE was 0.6315%. The results showed that the ANN model more accurate than the time series model. To sum up, all the objectives of this project had been achieved successfully.

5.2. Recommendations

To have more accurate results it is must be takes large range of the historical data. In addition, must be consider weather’s features specially temperature variance, and load behavior of the day by day. Upgrade the model to become medium or even long term forecaster with more realistic results.

REFERENCES

[1] M.V. Deshpande, "Design And Testing of Electrical Machines", PHl Learning Pvt. Ltd, New Delhi, 2010.

[2] Paladin Design Base.com, from

file://C:/DesignBase5/actrix/help/designbase/designbase.html

[3] Musab Ibrahim Adam Yagoub,"Intelligence Systems Modeling For Energy Demand Forecasting", School of Electrical System Engineering University Malaysia Perlis [UniMAP], 2016.

[4] William H. Kersting,"Distribution System Modeling And Analysis", CRC PressLLc,NewYork,2002.

[5] D.DAS, "principle of power system", New Age International (P) Ltd, New Delhi, 2006.

[6] https://www.Circuitglode.com/electrical load.html

[7] Ehab Elsaied, Atif Abd Alwahid, Mogahid Shareef and Fatima Mohammed,"Long Term Load Forecasting", School of Electrical And Nuclear Engineering Sudan University of Science and Technology [SUST], 2016.

[8] Dr. K. Uma Rao,"Power System Operation & Control", Wiley India Pvt. Ltd., New Delhi, 2013.

[9] Mosad Alkhathami," Introduction to Electric Load Forecasting Methods", Columbia International Publishing Journal of Advanced Electrical and Computer Engineering, 2015.

[10] Vincent Cheung and Kevin Cannons, "Introduction to Neural Networks", Electrical & Computer Engineering Unvirsity of Mantitoba, 2002.

[11] Manoj Kumar, "Short-Term Load forecasting Using Artificial Neural NetWork Techniques", Electrical Engineering National Institute of Technology, Rourkela, 2012.

APPENDIX

NN-TOOL MATLAB CODE

```
% Created 10-Oct-2017 17:36:27
% This script assumes these variables are defined:

%   simplefitInputs - input data.

%   simplefitTargets - target data.

x = simplefitInputs;

t = simplefitTargets;

% Choose a Training Function

% 'trainscg' uses less memory. Suitable in low memory situations.

trainFcn = 'traingdx';

% Create a Fitting Network

hiddenLayerSize = 14;

net = fitnet(hiddenLayerSize,trainFcn);

% Setup Division of Data for Training, Validation, Testing

net.divideParam.trainRatio = 70/100;

net.divideParam.valRatio = 15/100;

net.divideParam.testRatio = 15/100;

% Train the Network

[net,tr] = train(net,x,t);

% Test the Network
```

```
y = net(x);

e = gsubtract(t,y);

performance = perform(net,t,y)

% View the Network

view(net)

% Plots

% Uncomment these lines to enable various plots.

%figure, plotperform(tr)

%figure, plottrainstate(tr)

%figure, ploterrhist(e)

%figure, plotregression(t,y)

%figure, plotfit(net,x,t)


%Generate forecaster

Newpredict = net(y)
```