APPLICATION OF FUZZY LOGIC TO LOAD FORECASTING

A DISSERTATION
Submitted in partial fulfillment of the
requirements for the award of the degree
of
MASTER OF TECHNOLOGY
in
ELECTRICAL ENGINEERING
(With Specialization in Power System Engineering)

By
PRADIP SHANKARRAO TALODHIKAR

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE-247 667 (INDIA)
JUNE, 2004

I.D. No.-M.T...391../.EE-NP/Pa 2004

CANDIDATE'S DECLARATION

This is to certify that the report which is being presented in this dissertation titled "**Application of Fuzzy Logic to Load Forecasting**" in partial fulfillment of the requirements for the award of the degree of **Master of Technology** in **Electrical Engineering**, with specialization in Power System Engineering, submitted in the **Department of Electrical Engineering, Indian Institute of Technology, Roorkee** is an authentic record of my own work under the supervision of **Dr. E. Fernandiz**, Assistant Professor, Electrical Engineering Department, Indian Institute of Technology, Roorkee and **Dr. N. P. Padhy**, Assistant Professor, Electrical Engineering Department, Indian Institute of Technology, Roorkee.

I have not submitted the matter embodied in this report for the award of any other degree or diploma.

Date: 30 June 2004
Place: Roorkee

Talodhikar
(PRADIP SHANKARRAO TALODHIKAR)

CERTIFICATE

This is to certify that the above statement made by the student is correct to the best of my knowledge.

(Dr. E. Fernandiz)
Assistant Professor,
Deptt. of Electrical Engineering
Indian Institute of Technology,
Roorkee

(Dr. N. P. Padhy)
Assistant Professor,
Deptt. of Electrical Engineering
IndianInstituteTechnology,
Roorkee

ACKNOWLEDGMENT

A worship is never complete without prayer, so is a dissertation work never complete without its due acknowledgment. I take this opportunity to express my deep sense of gratitude to both of my guides **Dr. N. P. Padhy, Assistant Professor,** Department of Electrical Engineering Indian Institute of Technology, Roorkee and **Dr. E. Fernandiz,** Assistant Professor, Department of Electrical Engineering, Indian Institute of Technology ,Roorkee for their continual guidance, constant encouragement, discussion and unceasing enthusiasm. I consider myself privileged to have worked under their guidance.

My heartfelt gratitude and indebtedness goes to **Dr. J. D. Sharma,** professor of Electrical Engineering Department, **Er. Bharat Gupta,** Asst. Professor, O.C, PSS Lab for providing the necessary lab facilities. My sincere thanks to all **faculty** members of Power System Engineering for their constant encouraging and caring words, constructive criticism and suggestions, have contributed directly or indirectly in a significant way towards completion of this work. I am highly indebted to them for their unbounded interest, untiring efforts and great devotion to work.

My heartfelt appreciation also goes to all those who helped me directly and indirectly to make this dissertation work a success.

Last but not the least I am highly indebted to my parents, family members and all my friends whose sincere prayers, best wishes, moral support and encouragement have a constant source of assurance, guidance, strength and inspiration to me.

Date: 30 JUNE 2004

(Pradip Shankarrao Talodhikar)

ABSTRACT

The present work is done for 30 days load data taken for Nov.2003 for IITR campus for load forecasting and is tested for Fuzzy Time Series methodology, which is an integration of conventional and non-conventional approaches.

To test the developed MATLAB program, load data for 30 days i.e. from 1st Nov. 2003 to 30th Nov. 2003 has been collected from the different substations located inside the campus of Indian Institute of Technology, Roorkee. These required data are taken from the substations located in Indian Institute of Technology, Roorkee. In campus, there are five substations, which are taking the input from 33KV main receiving substation under Uttaranchal Power Corporation supplying to the 11kv substation. These substations are located in Cautley Bhawan, Welding Research Laboratory, High Voltage laboratory, Wind tunnel substation and the old substation near Physics Deptt. For the development of fuzzy time series load forecasting, the load data is collected from these five substations and the daily record of weather variables, average temperature and relative humidity has also been obtained from Hydrology department of I.I.T. Roorkee for the same period. The data is compiled in terms of various pre-decided variables for load forecasting using regression-based analysis. The proposed method is well suited for online load forecasting as the part of the program is deal with the historical data worked out for offline and only the adaptive part of it done on the online basis. This method can be constructed in simpler procedure than the conventional approaches.

CONTENTS

Title	Page No.
CANDIDATE'S DECLARATION	I
ACKNOWLEDGEMENT	II
ABSTRACT	III
CONTENTS	IV
LIST OF FIGURES	VI
LIST OF SYMBOLS	VII
LIST OF TABLES	VIII
1. INTRODUCTION	1
1.1 General	1
1.2 Literature Review	2
1.3 An Outline of Dissertation	6
2. LOAD FORECASTING	7
2.1 Definition of Load Forecasting	7
2.2 Importance of Load Forecasting	8
2.3 Load Forecasting principle	8
2.4 Classification of Load Forecasting	10
2.5 Methods for Load Forecasting	10
2.6 Introduction to Time series	11
3. FUZZY LOGIC: AN OVERVIEW	15
3.1 Introduction	15
3.2 Fuzzy Rule	16
3.3 Membership function	17
3.4 Fuzzy Arithmetic operation	20
3.5 Defuzzification process	21
4. INTRODUCTION TO FUZZY TIME SERIES	26
4.1 Definitions of Fuzzy Time Series	26
4.2 Types of Fuzzy Time Series	26

5. FUZZY TIME SERIES TO LOAD FORECASTING 29
5.1 Introduction 29
5.2 Algorithmic Procedure 30
5.3 Flowcharts 33

6. METHODOLOGY ADOPTED 36
6.1 Basics of the method 36
6.2 Data Requirements 36
6.3 Load Forecasting using Conventional method 37
6.4 Load Forecasting using Fuzzy Time Series 42

7. RESULTS AND CONCLUSION 47

REFERENCES IX

APPENDIX A XI

APPENDIX B XIX

LIST OF FIGURES

Fig No.	Title	Page
2.1	Algorithm for general Load Forecasting	9
2.2	Block diagram for Auto-regressive Integrated moving Average model	14
3.1	Schematic diagram of Fuzzy Logic technology	17
3.2	Membership functions	18
3.3	Center-of-Area method	23
3.4	Center-of-maximum method	24
3.5	Mean-of-maxima method	25
7.1	Graph for load forecasting by using fuzzy time series approach	49
7.2	Graph for forecasting error in %ge by Fuzzy Time Series approach	50
7.3	Comparison between Conventional method, realistic load data and Fuzzy Time series approach	51
7.4	Graph for predicted load and input load using conventional method	52
7.5	Graph for variance for Load Forecasting using Conventional method	53
7.6	Graph for error in Load Forecasting by Conventional method	54

LIST OF SYMBOLS

Symbols		Description
A	:	*Set of containing data (or objects)*
a_t	:	*Shock or white noise*
. (B)	:	*Generalized +autoregressive operator or stationary operator*
d_2	:	*Variance of white noise process*
w_t	:	*Weights used*
p	:	*Autoregressive operator*
X	:	*Set of data or objects*
x	:	*Individual value of data set X*
$\mu_A(x)$	:	*Membership function that converts the set X and A*

LIST OF TABLES

Table No.	Title	Page No.
6.1	Actual load and fuzzy sets	44
6.2	Fuzzy logic relationship with relation operators	45
7.1	Forecasted value of load (KW) by fuzzy time series approach and conventional method	55
7.2	Comparison of forecasting error between fuzzy time series approach and conventional method	57
A-1	Load Data (KW) observed at different five substations located in IITR campus for 30 days from 1st Nov. 2003 TO 30th Nov. 2003.	V
A-2	Load current (Amp.) observed at different substations located in IITR campus for 30 days from 1st Nov. 2003 TO 30th Nov. 2003.	VI
A-3	Daily average temperature recorded at Hydrology department, IITR, from 1st Nov. 2003 TO 30th Nov. 2003.	VII
A-4	Daily relative Humidity recorded at Hydrology department, IITR, from 1st Nov. 2003 TO 30th Nov. 2003.	VIII
B	Table for Output membership function and predicted value	IX

CHAPTER: 1

INTRODUCTION

1.1 GENERAL

An accurate load forecasting is of great importance for power system operation. It is the basis of power system planning such as economic planning, unit commitment, hydrothermal coordination and security analysis. An approach based upon fuzzy time series is developed for an accurate load forecasting. Fuzzy time series makes a load correction inference from historical information and the past forecast load error. Adding the inferred load error to the preliminary load forecasting obtains a final forecast load.

Electric load forecasting is aimed at predicting system load over an interval of one day or one week. It plays an important role in on-line scheduling and security function of energy management system. There are numerous methods proposed for load forecasting which includes the conventional techniques such as Moving Average model, Auto regression (AR), Auto regression integrated moving average (ARIMA) model, Exponential Smoothing method, Linear Regression and multiple regression analysis. These are all time series models. Load forecasting methods can be categorized as parametric based methods and artificial intelligence based methods. [5]. The later methods use neural networks as load models, where the former methods formulate load models as mathematical function exploiting the relationship that relates the load to dominant factors affecting this function. If these conventional and non-conventional approaches are integrated, then the load forecasting error may be minimized. The accuracy of load prediction depends mainly on the model used.

So among the traditional one, fuzzy time series is a widely discussed methodology for load forecasting. However, in traditional method, there are still large errors for forecast results. To take these forecast errors into account, a fuzzy time series approach can be employed. Fuzzy system converts exact information into symbolic information through use of linguistic sets by a process called *fuzzification*. A fuzzy variable should be weather conditions such as temperature, humidity, wind speed, cloud cover, etc. Fuzzification captures the vagueness that is so prevalent in many human decision making processes. Fuzzy sets are specified in linguistic terms such as 'low', 'normal', and 'high' requiring a little domain specification of knowledge to define. These load data having a crisp output which requires a method called *De-fuzzification*, to extract a crisp value that best represents the fuzzy output. Defuzzification may be generally carried out by centroid rule, which is the most popular method used to aggregate input information and fuzzy rules into preference values for each alternative for final decision making. With such crisp inputs and outputs, a fuzzy system implements a non-linear mapping from the input space to the output space. This mapping is accomplished by a number of *if-then* rules.

1.2 LITERATURE REVIEW

Several papers have proposed for the use of Fuzzy Logic for short term load forecasting. The present application of fuzzy method for load forecasting is in the experimental stage. For the demonstration of the method, a fuzzy time series method is adopted which includes conventional and non-conventional method. Forecasting methodology to be followed depends upon the period of interest for the load forecasting.

Paper [1] suggests two methods, fuzzy time series and genetic algorithm and compares them for the prediction of engineering manpower. The models of fuzzy time series and genetic algorithm are tested and based on the error analysis; model with minimum average error value is selected and used for the assessment of engineering manpower requirement. It is the key paper for the proposed work of dissertation. The central idea is taken from this paper for the prediction of load for 30 days short term load forecasting.

The smooth working of industry depends on the availability of proper engineering manpower. If proper qualified and experienced technical personnel are not available, the industry cannot run in the most efficient way. Here, an effort is made to assess the engineering manpower requirement (personnel belonging to Mechanical Engineering) in certain industry group in the state of West Bengal in India for the next five years. The models of fuzzy time series and genetic algorithm are tested and based on the error analysis; model with minimum average error value is selected and used for assessment of engineering manpower requirement. Certain statistical functions, i.e. least square technique based on linear, exponential equations and the tables of orthogonal polynomials are applied on the estimated data value calculated earlier for the prediction of futuristic engineering manpower. The particular statistical model is chosen based on the average error of estimated data generated using statistical models with the actual data over span of years. The said particular statistical model based on the estimated data using models of fuzzy time series or genetic algorithm can be used for the generation of futuristic forecasted engineering manpower.

Some of the work has completed on the fuzzy time series for electric power consumption [2]. The author presents a data analysis and estimation procedure of electrical power consumption under uncertain conditions. Traditional methods are based on statistical and probabilistic approaches but it may not be quite suitable to apply purely mathematical models to the data generated by human activities such as the power consumption. It introduces a new approach based on possibility theory and fuzzy auto regression and applies it to the analysis of time series data of electric power consumption. The proposed fuzzy model assumes that the time series data of power consumption reflect human decision making activities rather than the probabilistic process assumed in the conventional models.

Intaek Kim and Sung-Rock Lee [3] give the fuzzy time series prediction on consecutive values for non-stationary data. It presents a time series prediction method using a fuzzy rule-based system. A serious problem of conventional methods is that they cannot properly handle non-stationary data whose long-term mean is floating. To cope this, a fuzzy time series method is used for consecutive values. Some of the new fuzzy learning method for prediction also utilizes the difference of consecutive terms in a time series. For that the authors show the simulation results the proposed method, which has two important capacities: representation of non-stationary signal and reduction of errors in predicting a time series. There is the future research work is also going on the proposed algorithm to more genuine non-stationary time series. The logic for fuzzy time series prediction is taken from this literature for fuzzy time series load forecasting.

The application of fuzzy logic technology for load forecasting is in the experimental stage at present. The demonstration of this technology is used a fuzzy expert systems that forecasts the daily peak load, is given in paper [4]. In this paper, author Dr. Keith Holbert has designed the fuzzy logic system to forecast the peak and through load. But, there is also some disadvantages in the proposed system. In selecting the fuzzification and defuzzification process, the problem for selecting the membership functions are come into picture. Such problems may be overcome by combining the neural network with the fuzzy logic. The neural network optimizes the rule base and involves the training of the network to the historical data to determine the rules that contribute to a better decision. One another technique is genetic algorithm and fuzzy time series method, which includes an integration of time series and fuzzy logic approach. This is a series of data used in time slots with the fuzzy logic, gives the better result even in the non-stationary data.

An economic secure operation of electric power systems requires good prediction of electric load for twenty-four hrs. [5] gives the application of fuzzy system to electric short term load forecasting in two model formulation problem. These two models are used for modeling the electric load. In the first model, the load power depends only on hour on question and the second model is a function of the temperature as well as the hour in question. Here the problem is formulated as a linear fuzzy regression where the linear programming is used for estimation of load parameters. This regression based technique uses in the convention method of load forecasting and compares with the fuzzy time series model. Here the triangular membership function is used which is symmetrical around the middle. The first model is a harmonic model that accounts only the time question and does not account for weather factors. The coefficients of the model are assumed as fuzzy while the load power is assumed either crisp or fuzzy. In second model, the hybrid combination is used that accounts for the hour in question. The model can implement for summer of for winter weekdays or weekend also. The second part of paper [6] represents about the computational results for application of fuzzy systems to load forecasting. The computational results are obtained using fuzzy models; the variations in the load power can be accounted for since the estimated load power is within the upper as well as lower limits. The actual load never violets these limits.

Mo-yuen Chow and Hahn Tram [7] explain the general methodology to use fuzzy logic to fuse the available information for spatial load forecasting. The proposed scheme can provide distribution planners other alternatives to aggregate their information for the spatial load forecasting. It describes the fuzzy logic formulation in spatial load forecasting land usage inference and how the fuzzy sets and fuzzification process is used for load forecasting. Fuzzy rules for fuzzification are given with the membership functions. The general defuzzification method used is centroid rule approach. The approach has been discussed in terms of ease of heuristic implementation and feature resolution. The proposed work for fuzzy time series load forecasting uses the fuzzy rule and defuzzification idea is taken from this paper.

Two days workshop was held under the auspices of the Khwarzimic Science Society at University of Lahore [8] explains the fuzzy logic control using the matlab software package. The central idea about the fuzzy logic technology is taken from the same paper, which gives the knowledge of fuzzy logic in matlab programming. It gives the detailed information about the membership functions, if and then fuzzy rule and defuzzification.

The further discussion about the time series methodology to load forecasting is given in the paper [9]. It is shown than Box and Jenkins time series models (ARIMA, Periodic ARIMA and Transfer function) [11], in particular manner, which is well suited to the further application. But there is a drawback of inability to accurately represent nonlinear relationship between load and temperature, which is overcome by comparing several Box and Jenkins models with a forecasting procedure currently used.

The paper [10] presents a regression based adaptive weather sensitive load forecasting algorithm, which is developed in electric power utility of Serbia. It consists of two steps, one is daily load model and other is hourly load model. The Euclidean distance between the forecast day and days from the database is used as criteria of defining the identification period in both models. The weather sensitive model is taken for the further work in conventional method.

Nima Amjdy [12] represents a new time series modeling for short term load forecasting, which can accurately forecast the hourly loads of weekdays as well as of weekends and public holidays. These methods provide the more accurate results than the conventional techniques, such as ANN or Box –Jenkins models [11]. The work has shown the exact forecasting for the daily peak load of a power system. This incorporates the time series modeling of the ARIMA with the knowledge of experienced human operators.

The proposed fuzzy auto-regressive (AR) model and its applications described in paper [13]. The identification and estimation of the model proposed is optimized by the linear programming problem under some condition. Hence, the performance of the proposed model is tested by some random data. Here, the improvement of fuzzy AR model is done. This method is applied to load forecasting process. It describes the behavior of fuzzy time series as compared with the stochastic model. The performance of its model is discussed using a few applications, such as living expend of worker's household and price indexes of meats. The concept of fuzzy time series is taken for the proposed work for load forecasting. The details about the papers presented in journals regarding the load forecasting are given in the bibliography [14]. The representative cross-section of publications related to load forecasting is listed in this paper.

1.3 AN OUTLINE OF DISSERTATION

The dissertation work is organized in the seven chapters. Chapter 2 describes the definition of electric load forecasting, its application to the power system planning, hydrothermal scheduling coordination. Its importance to load forecasting with their principle used for electric load prediction. The load forecasting is classified in this chapter. The conventional methods and non-conventional methods used for load forecasting are given here. The time series is introduced in the same chapter with its definition and time series model such as Autoregressive (AR) model and Autoregressive integrated moving average (ARIMA) model.

Chapter 3 gives the detailed information about the fuzzy logic concept whose features that makes it superior to classical theory, is also discussed here. It also gives the fuzzy if and then rule with their membership function types. It includes the fuzzy arithmetic and basic operations to fuzzy logic concept. The final membership function is defuzzified by using defuzzification process, such as center of area, center of maximum and maxima of minima methods. The most superior and common used method for defuzzification is center of area method.

Chapter 4 describes the time series approach. It includes fuzzy time series invariant and variant time series methods.

Chapter 5 describes the fuzzy time series to load forecasting and the seven stepwise algorithmic procedure described for fuzzy time series with problem evaluation with the flowcharts shown.

Chapter 6 includes the total program evaluation for the 30days load data using fuzzy time series approach to load forecasting. It gives the basics of the methods used and its data requirements taken from the different substations located in IITR campus. The load forecasting using conventional method as multiple regression analysis is also discussed here. Then it compares with the fuzzy time series approach.

Chapter 7 gives the result and conclusion about the present work of dissertation.

CHAPTER: 2

LOAD FORECASTING

2.1 DEFINITION OF LOAD FORECASTING

Load is a general term meaning either demand or energy, where demand is the time rate of change of energy. The term forecast refers to the projected load requirement using systematic process of defining future loads in sufficient quantitative details so that important power system expansion decisions can be made. For effective power system expansion, estimates of both power and energy are crucial. An accurate forecast depends on the judgment of the forecaster and it is impossible to rely strictly on analytical procedures to obtain an accurate forecast. Also good judgment alone cannot be over emphasized in forecasting future requirements and analytical tools must be used.

In some cases, a total forecast is obtained by combining forecast for various classes of customers such as residential, commercial, industrial, agriculture and other. For the better prediction of load, it is imperative for the forecaster to be well versed with the characteristics of varieties of loads i.e. growth rate with time and effect of weather fluctuations. In many cases, seasonal variations in residential component are responsible for the seasonal variations in system peak. The extent of residential influence will depend upon the percentage of total system load that is residential. Increase in per capital consumption due to wide spread use of weather sensitive devices, need to include weather effects in forecasting future requirements become imperative.

Forecasting methodology to be followed depends upon period of interest for load forecasting. Short and medium term forecasting can be done using conventional procedures i.e. by extrapolation of certain factors characterizing hourly modulation of load. For long term forecasting, consumption is broken down into homogeneous sectors. The preceding extrapolation techniques applied to different consumption classes make it possible to generate individual load curves, thus giving a global level by summation.

2.2 IMPORTANCE OF LOAD FORECASTING

Load forecasting has always been an integral part of power system planning and operation. However, it did not receive as much attention in the past as it deserves because the fuel supplies, especially hydrocarbons were cheap and abundant and utilities could find funds for erecting enough gas or oil generation plants at relatively short lead time. In the last few years, conditions have considerably changed and the past practices have to be suitably modified. Thus, load forecasting will assume greater importance in the years to come.

Long term demand forecast is used to determine the capacity of generation, transmission and distribution system additions and energy forecasts, which determine the type of facilities required i.e. whether peaking generating unit or base load unit should be installed as there is substantial difference in their cost. Whereas the importance of short term load forecasting is realized in service reliability and efficient operating performance of power supply system, particularly in the case of thermal power plant which demands accuracy in the forecast of system load. This is necessary in order that generating capacity adequate to supply the system load, maintain the system security and supply the necessary spinning reserve. The scheduling of spinning reserve as coverage for the loss of largest unit or the loss of that transmission capacity which renders unavailable the greatest amount of generating capacity, is an accepted the principle of reliable system operation.

2.3 LOAD FORECASTING PRINCIPLE

Most of the planners agree that forecasting is affected by pricing the policy and economic prosperity, economists point out that the willingness and the ability of consumers to pay a given price at any given time in the future are not necessarily the same. The willingness and ability to pay depend on future economic wealth, future income distribution and future prices. The wealth of each person, village, town, country and indeed the world are all bound up together in a little understood and highly complex way. Therefore, load forecasting requires a study at all levels of disaggregation, starting very briefly with the world, the national macro-economic and overall energy sector and finally national and local power sectors. To make the proper load forecasts, it must be studied the likely future worldwide and electricity prices. In developing countries, the likely future connection cost for new consumers is also important, together with prices of fuels, which can substitute for electricity and vice versa.

For that, the general load forecasting algorithm [10] gives the forecasting techniques for prediction of load on daily basis and also hourly basis, which is shown in Fig. 2.1.

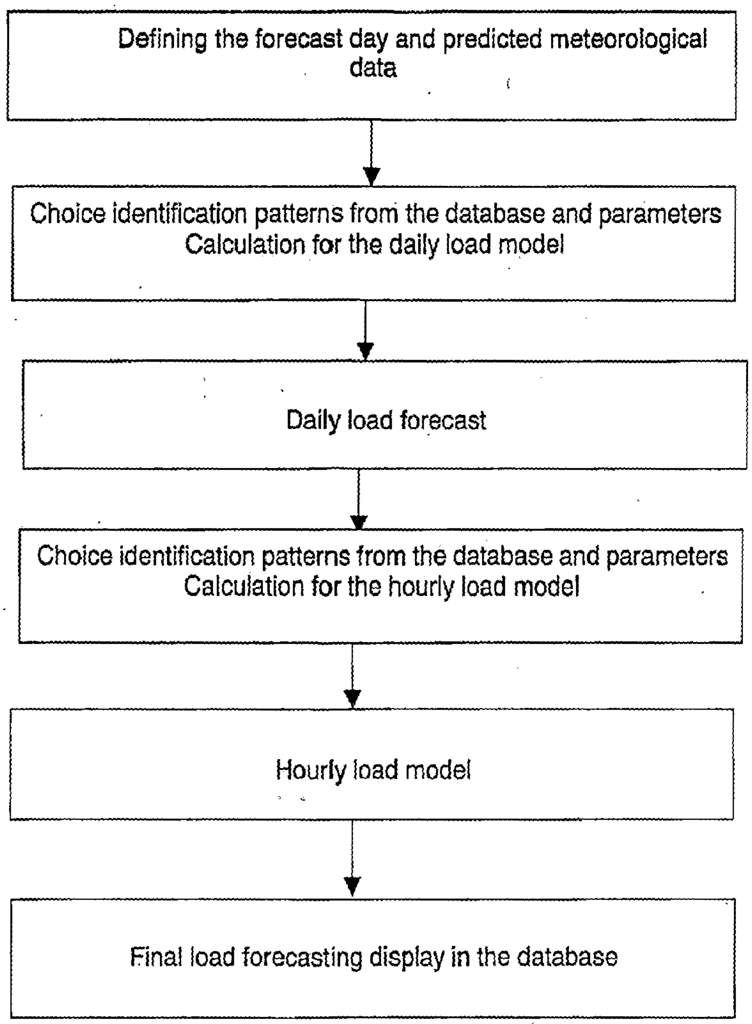


Fig. 2.1 Algorithm for general Load Forecasting

The main presumption used in the algorithm development is that the total daily load forecasting is much easier task than the hourly load forecasting. That is why two different models, one for the total daily load forecast and the other for hourly load forecast, are established. So, the first step in the proposed general approach is to forecast the total daily load and the second step is the daily load shape definition. Hourly, loads are expressed in per unit as parts of the total daily loads. All parameters of the both models are automatically adapted in the identification procedure.

2.4 CLASSIFICATION OF LOAD FORECASTING

Since there is no formula by which the optimum amount of electric power for any country or geographic area can be determined though few countries are using an empirical formula developed by them. According to their past experience and records and few are using fixed rates of power requirements. Forecasting the future demands for electricity requires assumptions with regards to the basic policies governing the electric power industry. This is especially true because the demand for electricity depends heavily upon the degree of which it is available and the price at which it is offered for sale.

Depending upon the time period of interest, a specific load forecasting procedure may be classified as:

a. Short term load forecasting
b. Long term load forecasting

2.5 METHODS TO LOAD FORECASTING

The electric load forecasting uses the following methods:

A. Conventional Approaches

The conventional approaches [14] for electric load forecasting are stated as shown:

1. Time Series Model
   a. Moving Average Model
   b. Exponential Smoothing
   c. Linear Regression
   d. Auto Regression (AR)
   e. Auto Regressive Integrated Moving Average (ARIMA)
2. Extrapolation Model
   a. Trend fitting Model
   b. Probabilistic Model

3. Econometric Model
   a. Probit Model
   b. Multiple Regression Model

4. Other Miscellaneous Model
   a. Delphi Technique (Survey Model)
   b. Kalman Filter Modeling
   c. State Estimation Model

B. Non Conventional Approaches

1. Artificial Neural Network (ANN)
2. Expert System
3. Fuzzy Logic Concept
4. Genetic Algorithm

2.6 INTRODUCTION TO TIME SERIES

2.6.1 Definition

A *time series* is defined as a sequential set of data measured over time. The sequence is the most crucial idea in time series analysis because the information on the source is embedded in the sequence. Therefore, it is able to analysis a time series by assuming that a source of a time series is governed by a deterministic dynamic system.

Time series prediction involves forecasting the future by understanding the past. It has been widely studied by many researchers. Most of works on prediction have been conducted from the viewpoint of stochastic models such as *Moving average (MA), Integrated Moving Average (IMA), and Auto Regressive Integrated Moving Average (ARIMA)* [11].

2.6.2 Time Series Model

The most fundamental time series models [9] are the autoregressive model and the moving average model. In the autoregressive model (AR), the current value of the process is expressed as a linear combination of previous values of the process and a random shock, a_t.

The different time series models are discussed as follows:

a. Auto Regressive (AR) Model

A stochastic model, which can be extremely useful in the representation of certain practically occurring series, is the so called 'autoregressive model'. In this model, the current value of the process is expressed as a finite, linear aggregate of previous values of the process and a shock 'a_t'. [11].

Let us denote the values of a process at equally spaced times *t, t-1, t-2, by* y_t, y_{t-1}, y_{t-2}...

Also let y_t, y_{t-1}, y_{t-2},.... be the deviations from . ;

For example,

$$\overline{y_t} = y_t - \alpha \tag{2.1}$$

Then

$$\overline{y_t} = \phi_1 \overline{y}_{t-1} + \phi_2 \overline{y}_{t-2} + \dots\dots\dots\dots + \phi_p \overline{y}_{t-p} + a_t \tag{2.2}$$

is called 'an autoregressive (AR) process [9] order p'. The reason for this name is that a linear model.

$$\overline{Y} = \phi_1 \overline{x_1} + \phi_2 \overline{x_2} + \dots\dots\dots\dots + \phi_p \overline{x_p} + a_t \tag{2.3}$$

Relating a 'dependent' variable y to a set of 'independent' variables x_1, x_2, ... x_p plus an error term 'e' is often referred to as 'regression model' and *y* is said to be 'regressed' on x_1, x_2, ... x_p. In the variable y is regressed on previous values of itself; hence the model is autoregressive.

If AR operator, *p* is defined by

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots\dots\dots\dots - \phi_p B^p \tag{2.4}$$

Then the AR model may be written economically as,

$$\phi(B)\overline{y}_t = a_t \tag{2.5}$$

The model contains *p+2* unknown parameters *a*, . $_1$, . $_2$, *p*, d_2, which in practice have to be estimated from the data. The additional parameter 'd_2 'is the variance of the white noise process at. It is not difficult to see that the AR model is a special case of the linear filter model. AR processes can be stationary or nonstationary. For the process to be stationary, the ß's must be so chosen that the weights will form a convergent series.

b. Auto Regressive Integrated Moving Average (ARIMA) Model

Many series actually encountered in industry or business exhibits nonstationary behavior and in particular do not vary about a fixed mean. Such series may nevertheless exhibit homogeneous behavior of a kind. In particular, although the general level about which fluctuations are occurring may be different at different times, the broad behavior of the series, when differences in level are allowed for, may be similar. This behavior may be represented by generalized autoregressive operator. *(B)*, in which one or more of the zeroes of the polynomial. *(B)* is unity.

Thus the operator can be written as,

$$\phi(B) = \beta(B)(1-B)^d \tag{2.6}$$

where, . *(B)* is a stationary operator.

Thus, a general model which can be represented as the homogeneous non-stationary behavior, is in the form of

$$\phi(B)y_t = \beta(B)(1-B)^d y_t = \theta(B)a_t \tag{2.7}$$

That is

$$\phi(B)w_t = \theta(B)a_t \tag{2.8}$$

where,

$$w_t = \Delta^d y_t \tag{2.9}$$

Therefore, homogeneous nonstationary behavior can be represented by a model, which calls for the d'th difference of the process to be stationary. In practice, d is usually 0, 1, or at most 2.

The process defined above, provides a powerful model for describing the stationary and nonstationary time series and is called an "Autoregressive Integrated Moving Average" (ARIMA) process of order (p.d.q). The general ARIMA model [11] can be generated from white noise 'a_t' by means of three filtering operations as Moving average filter, stationary auto regressive filter and non-stationary summation filter which is indicated by the block diagram of Fig. 2.2.

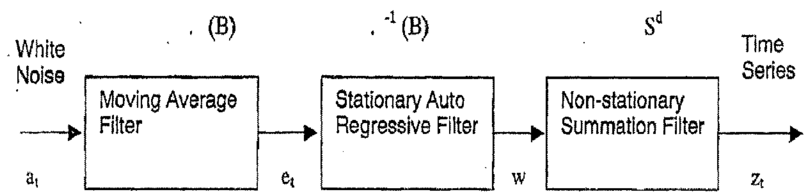


Fig. 2.2 Block diagram for Auto-Regressive Integrated Moving Average model

CHAPTER: 3

FUZZY LOGIC: AN OVERVIEW

3.1 INTRODUCTION

Among the various paradigmatic changes in science and mathematics in this century, one such change concerns the concept of uncertainty. Uncertainty is considered essential to science. It is not only an unavoidable plague, but it has in fact a great utility.

It is generally agreed that an important point in the evolution of the modern concept of uncertainty, was the publication of a paper by Lotfi A. Zadeh (1965). In his paper, Zadeh introduced a theory whose objects – Fuzzy sets – are sets with boundaries that are not precise. The membership in a fuzzy set is not a matter of affirmation but rather a matter of a degree.

The significance of Zadeh's paper was that it challenged not only probability theory as the sole agent for uncertainty but the very foundations upon which probability theory is based: two valued logic. The capability of fuzzy sets to express gradual transition from membership to non-membership and vice-versa has a broad utility. It provides us not only with a meaningful and powerful representation of measurement of uncertainty, but also with a meaningful representation of vague concepts expressed in natural language. For example instead of describing the whether today in terms of the exact percentage of cloud cover, we can just say that it is sunny. While the latter description is vague and less specific, it is often more useful. A research on the theory of fuzzy sets has been growing steadily since the inception of the theory in the mid - 1960. The body of concept and results pertaining to the theory is now quite impressive.

The four features that make the fuzzy logic superior to classical theory are:

1. The fuzzy logic allows us to express irreducible observation and measurement uncertainties in their various manifestations and make these uncertainties intrinsic to empirical data. Such data, which are based on graded distinction among states of relevant variables are usually called fuzzy data when fuzzy data, are processed, their intrinsic uncertainties are processed as well and the result obtained are more meaningful than their counterpart obtained by processing the usual crisp data.

2. The fuzzy logic offers far greater resources for managing complexity and controlling computational cost.
3. The fuzzy logic has considerably greater expressive power; consequently it can effectively deal with a broader class of problem. In particular, it has the capability to capture and deal with meanings of sentences expressed in natural language.
4. The fuzzy logic has greater capability to capture human common sense reasoning, decision making and other aspects of human cognition.

3.2 FUZZY RULE

Fuzzy sets and fuzzy operators are the subjects and verbs of fuzzy logic. Conditional statements, *if-then* rules are the things that make fuzzy logic useful.

A single fuzzy if-then rule assumes the form:

If x is A then y is B

Where, *A* and *B* are linguistic values defined by fuzzy sets on the ranges *X* and *Y*, respectively. The if-part of the rule "*x is A*" is called *the antecedent or premise* [8], while the then-part of the rule "*y is B*" is called *the consequent or conclusion* [8]. Antecedent is an interpretation that returns a single number between 0 and 1, whereas the consequent is an assignment that assigns the entire fuzzy set *B* to the output variable *y*.

The fuzzy system is a popular computing framework based on the concepts of '*fuzzy set theory*', '*fuzzy if then rules*' and '*fuzzy reasoning*' [4]. The structure of fuzzy inference consists of three conceptual components, namely:

- Rule Base containing a selection of fuzzy rules.
- Database defining the membership functions. These are used in the fuzzy rules.
- Reasoning mechanism that performs the inference procedure upon the rules and given facts and derives a reasonable output or conclusion.

Sometimes it is necessary to have crisp output. This requires a method called *De-fuzzification*, to extract a crisp value that best represents the fuzzy output. With such crisp inputs and outputs, a fuzzy expert system implements a non-linear mapping from the input space to the output space. This mapping is accomplished by a number of *if-*

then rules, each of which describes a local behavior of the mapping, which is described in the following overall process (Fig.3.1).

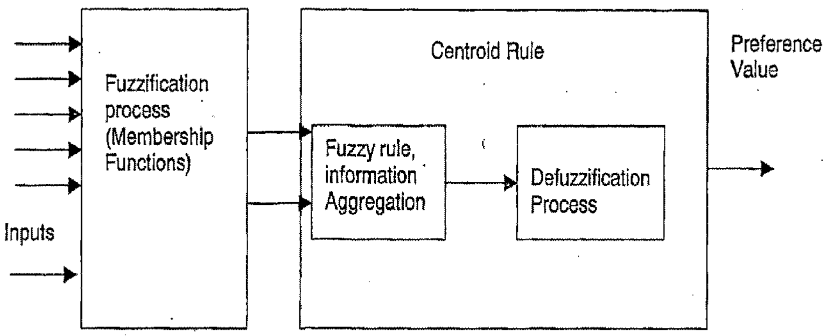


Fig. 3.1 Schematic Diagram of Fuzzy Logic Technology

3.3 MEMBERSHIP FUNCTIONS

A membership function is a curve that defines how each point in the input space is mapped to a membership value of degree of membership between 0 and 1. The input space is sometimes referred to as the *universe of discourse*. The output axis is a number known as the membership value between 0 and 1. The curve is known as a membership function and is often given in the designation of μ.

To illustrate this let us consider:

X: a set of data or objects. (Example, Forecast temperature values)

A: another set containing data (or objects)

x: an individual value of the data set *X*

$\mu_A(x)$ is the membership function that connects the set *X* and *A*. The membership function $\mu_A(x)$,

- Determines the degree that x belongs to *A*
- Its value varies between 0 and 1
- The high value of $\mu_A(x)$ means that it is very likely that *x* is in *A*

The membership function is selected by trial and error. There are four basic membership functions namely:

- Triangular.
- Trapezoidal.
- Gaussian.
- Generalized bell.

The above functions are shown in Fig. 3.2.

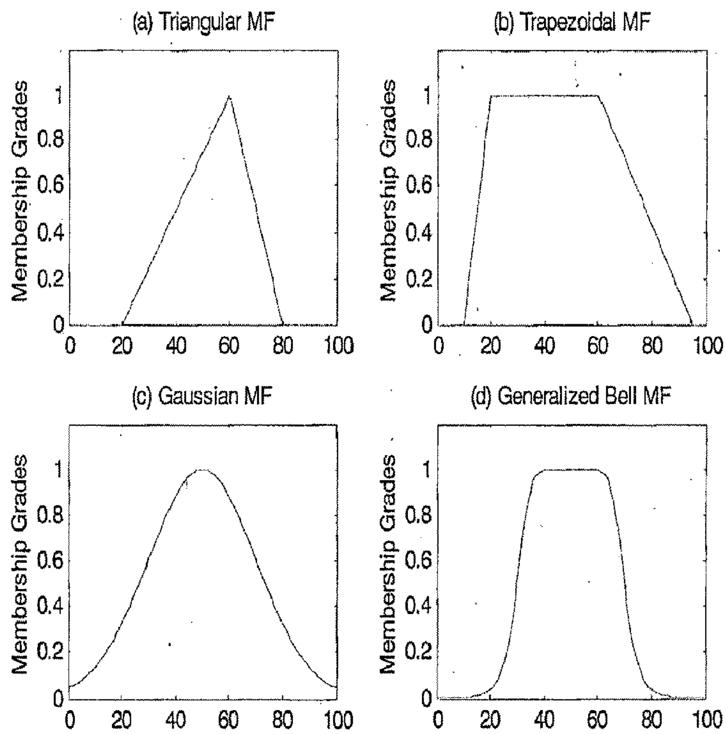


Fig. 3.2 Membership Functions

The triangular function *"triangle (x, a, b, c)"* is defined as:

$$A = \begin{cases} 0, x \le a \\ (x-a)/(b-a), x \in (a,b) \\ (c-x)/c-b), x \in (b,c) \\ 0, x \ge c \end{cases}$$

It has three parameters '*a*' (minimum), '*b*' (middle) and '*c*' (maximum) that determines the shape of the triangle as shown in Fig. 3.2(a).

A trapezoidal membership function is specified by four parameters given by:

A = trapezoid (x, a, b, c, d)

The function is described as:

$$A = \begin{cases} 0, x \le a \\ (x-a)/(b-a), x \in (a,b) \\ 1, x \in (b,c) \\ (d-x)/d-c), x \in (c,d) \end{cases}$$

The above function determines the shape of the trapezoidal membership function as shown in Fig. 3.2 (b).

Similar definitions for Gaussian and generalized bell [4] can be given. However, triangular and trapezoidal functions are simple and most frequently used. The membership functions are not restricted to these four. One can have their own tailor-made functions. The functions above were mere one dimensional in nature. In principle, one can even have multi-dimensional membership functions. Coming back to our sets *A* and *X*, the fuzzy set *A* in *X* as a set of ordered pairs are defined as,

$$A = \{(x, \mu_a(x)) \forall x \in X\}$$

A set, *A* of points or objects in some relevant universe, *X* is defined as those elements of *X* that satisfy the membership property defined for *A*. in traditional or crisp set theory each element of *X* either is or is not an element of *A*. Elements in fuzzy set can have a continuum of degrees of membership ranging from complete membership to complete non-membership.

The membership function *μ(X)* gives the degree of membership for each element x. *X*. *μ(X)* is defined on [0 1], where 1 represents elements that are completely in A and 0 represents elements that are completely not in A and values between 0 and 1 represents partial inclusion in A.

Formally, A is represented as the ordered pair $(x, \mu(X))$.

$$A = \{|x, \mu(X)) |_{x\,.\,X,} \text{ and } 0\ .\ \mu(X)\ .\ 1\}$$

The use of a numerical scale for the degree of membership provides a convenient way to represent gradations in the degree of membership. Precise degrees of membership generally do not exist; instead it tends to reflect a sometimes subjective ordering of the elements in the universe.

Fuzzy numbers are numerical approximations such as "about 5". Formally, a fuzzy number is defined as a closed interval on R (the real line), N (the integers) or any totally ordered set whose membership function is normal and convex and reaches its maximum value (1.0) at the number. The width of membership function shows the range of possible values.

3.4 FUZZY ARITHMETIC OPERATIONS

Basic Operations of Fuzzy Set

1. A fuzzy set A is contained in fuzzy set B, $A.\ B$ if $\mu_A(X)\ .\ \mu_B(X)$ for all $x\ .\ X$.
2. The basic set operations for two fuzzy sets A and B:

Intersection:

$A\ .\ \ B = min\ [\mu_A\ (X), \mu_B\ (X)]$

Union:

$A\ .\ \ B = max\ [\mu_A\ (X), \mu_B\ (X)]$

Complement:

$\bar{A} = 1 - \mu_A\ (X)$

These operations satisfy the *associativity* and *distributivity* properties of ordinary sets.

3. A .*–cut* is the set of elements in a fuzzy set that have a degree of membership greater than . :

$A_. = \{X\ |_{x\,.\,X,}\ and\ \mu(X) > .\ \}$

4. Arithmetic operations on fuzzy numbers

Fuzzy numbers are represented by,

A = [a1, a2]

B = [b1, b2]

C = [c1, c2]

Fuzzy arithmetic operations as defined above are equivalent to the corresponding interval arithmetic operations for each . –cut. For the basic arithmetic operations, we have:

$$A_a + B_a = [a1_a, a2_a] + [b1_a, b2_a]$$
$$= [a1_a + a2_a, b1_a + b2_a]$$
$$A_a - B_a = [a1_a, a2_a] - [b1_a, b2_a]$$
$$= [a1_a - a2_a, b1_a - b2_a]$$
$$A_a X B_a = [a1_a, a2_a] X [b1_a, b2_a]$$
$$= [a1_a . a2_a, b1_a . b2_a]$$
$$A_a / B_a = [a1_a, a2_a] / [b1_a, b2_a]$$
$$= [a1_a / b2_a, a2_a / b1_a]$$

$$\frac{1}{B_A} + \frac{1}{[b1_a, b2_a]} = [\frac{1}{b2_a}, \frac{1}{b1_a}]$$

$A. \times K = [a_1', a_2'] \times K = [K a_1', K a_2']$

Operations (4) and (5) are undefined if the interval contains 0 as resulting interval goes to infinity.

It is easy to show from the proportionality properties of triangles that additions and subtraction of triangular fuzzy numbers and a multiplications, division and inversion of fuzzy number generally do not give a triangular result. Fig. 3.1 illustrates these properties of fuzzy numbers. Even though the result of multiplication of fuzzy numbers is not triangular, it can be approximated by a triangular fuzzy number (1). This approximation simplifies the fuzzy arithmetic.

3.5 DEFUZZIFICATION PROCESS

The process that converts a fuzzy set or fuzzy number into a crisp value or number is known as '*defuzzification*'.

Defuzzification is such inverse transformation, which maps the output from the fuzzy domain back into the crisp domain. The following defuzzification methods are of practical importance:

- Center-of-Area (C-o-A)
- Center-of-Maximum (C-o-M)
- Mean-of-Maximum (M-o-M)

3.5.1 Center-of-Area Method

The *Center-of-Area* method is often referred to as the *Center-of-Gravity* method because it computes the *centroid* of the composite area representing the output fuzzy term.

The centroid rule is one of the most popular defuzzification techniques and its approach can compromise and /or resolve the conflicts or inconsistencies of the different preferences of decision making. Therefore the decision is not only based on some specific points of the membership curves but on the entire membership functions under consideration [7]. Centroid method is the most prevalent and physically appealing of all the defuzzification methods.

Fig. 3.3 shows the membership functions of a linguistic output variable *MotorPower* where the areas of *ZE* and *PM* are combined by the *union* operator and thus their contour becomes the composite fuzzy output for *MotorPower*. *C-o-A* defuzzification method computes the centroid of this area.

$$u^* = \frac{\sum_{i=1}^{N} u_i \mu_{OUT}(u_i)}{\sum_{i=1}^{N} \mu_{OUT}(u_i)}$$

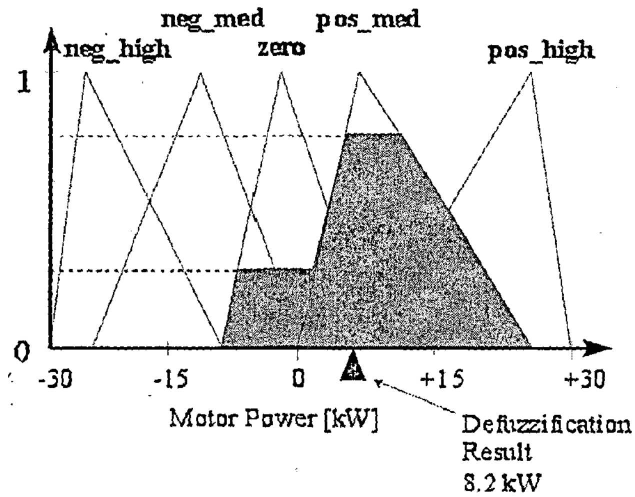


Fig. 3.3 Center-of-Area Method

3.5.2 Center-of-Maximum Method

The *Center-of-Maximum* requires only the peaks of the membership functions. Defuzzified value is determined by finding the fulcrum where the weights are balanced. This method is also called by *Height-Method.*

The crisp output is computed as a weighted mean of the term membership maxima, weighted by the inference results.

Equations are very similar, except that for C-o-A it is used the areas of each membership functions. For C-o-M, it is used only their maxima. Naturally, the results are slightly different.

$$u^* = \frac{\sum_{i=1}^{N} u_i \cdot \sum_{k=1}^{n} \mu_{O,k}(u_i)}{\sum_{i=1}^{N} \sum_{k=1}^{n} \mu_{O,k}(u_i)}$$

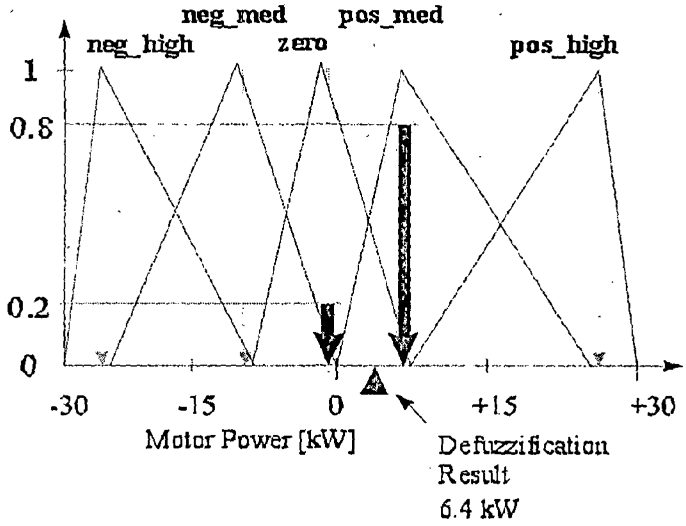


Fig. 3.4 Center-of-maximum method

3.5.3 Mean-of-Maxima Method:

The *Mean-of-Maxima* is used when the maxima of the membership functions are not unique, one can then take the mean of all maxima. *Max (ml)* is the inferred fuzzy output term with the highest degree of truth and *M* is the integer number of such peaks.

$$u^* = \sum_{m=1}^{M} \frac{u_m}{M}$$

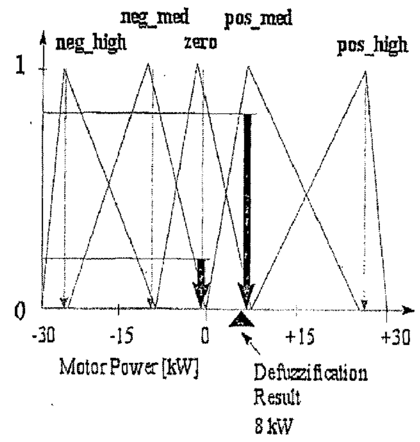


Fig. 3.5 Mean-of-Maxima method

CHAPTER: 4

INTRODUCTION TO FUZZY TIME SERIES

4.1 Fuzzy Time Series

Fuzzy time series approach is the integration of the conventional time series approach and non-conventional fuzzy logic concept. Some of the definitions of fuzzy time series are discussed as follows.

Definition 1:

It is assumed that Y (t), (t =......, 0, 1, 2, ...), a subset of R' as the universe of discourse on which fuzzy sets fi (t), (i = 1, 2...), are defined and F (t) as a collection of f_1 (t), f_2 (t)... then F (t) is termed as fuzzy time series defined on,

Y (t), (t =......, 0 1, 2...).

Definition 2:

If it is assumed that F (t) is caused by F (t -1) only, i.e. F (t -1) → F (t), then this relation can be expressed as F (t) = F (t -1). R (t, t -1), where R (t, t -1) is the fuzzy relationship between F (t -1) and F (t). F (t) = F (t -1). R (t, t -1) is called the first order model of F (t).

Definition 3:

Suppose R (t, t -1) is a first order model of F (t) and if for any t, R (t, t -1) is independent of t, i.e. For any t, R (t, t -1, t -2), then F (t) is the time variant fuzzy time series.

4.2 Types of Fuzzy Time Series

The fuzzy time series have the following two types:

1. Variant Fuzzy Time Series

Variant fuzzy time series is the time dependent approach, which definitions are already given above.

2. Invariant Fuzzy Time Series

If it is assumed that F (t) is a fuzzy time series and if for any t, F (t) = F (t -1) and F (t) has only finite elements, then F (t) is a time invariant fuzzy time series.

If F (t) is a fuzzy time series, F (t) = F (t -1) for any t and F (t) has only finite elements f_i (t) (i=1, 2, n), then

R (t, t -1) = ..., f_{i1} (t-1) f_{i0} (t) U f_{i2} (t-2) fj1 (t-1) U... ... f_{im} (t-m) f_{jm-1} (t-m+1)

Where, m>0 and all pairs of fuzzy sets are different.

Fuzzy Time Series

Let U be the universe of discourse, where

U = {u_1,u_2,...,u_n}.A fuzzy set Ai of U is denoted by

$$A_i = fA_i(u_1)/u_1 + fA_i(u_2)/u_2 + ... + fA_i(u_n)/u_n$$

where,

fAi is the membership function of fuzzy set Ai ,

fAi :U->[0,1].

u_k is the element of fuzzy set Ai ,and

fAi(u_k) is the degree of belongingness of u_k to Ai :fAi(u_k) .[0,1]

where 1 <= k <= n.

The fuzzy time series is also given by,

Y (t) (t =...,0,1,2,...), is a subset of R. Let Y (t) be the universe of discourse denoted by fuzzy set fi (t).

If F (t) consists of fi (t) (i =1, 2,...)

F (t) is denoted as a fuzzy time series on Y (t) (t =..., 0, 1, 2,...).

If there exists a fuzzy relationship (t-1, t), such that F (t) = F (t-1) * R (t-1, t),

where * is an operator, then F (t) is said to be caused by F (t .1).The relationship between F (t) and (t-1) can be denoted by F (t-1) -> F (t).

Suppose F (t-1) = Ai and F (t) = Aj, a fuzzy logical relationship is denoted as Ai-> Aj.

Then fuzzy logical relationships can be further grouped together into fuzzy logical relationship groups according to the same left-hand sides of the fuzzy logical relationships.

For example, there are fuzzy logical relationships with the same left-hand sides (Ai):

Ai -> Aj1

Ai -> Aj2

....

These fuzzy logical relationships can be grouped into a fuzzy logical relationship group as follows:

Ai -> Aj1, Aj2...

Suppose F (t) is caused by F (t-1) only, and F (t) = F (t-1) × R (t-1, t). For any t, if R (t-1, t) is independent of t,

Then F (t) is named a time-invariant fuzzy time series; otherwise it will be a time-variant fuzzy time series.

CHAPTER: 5

FUZZY TIME SERIES TO LOAD FORECASTING

5.1 INTRODUCTION

One of the breakthroughs in the area of load forecasting began with the development of computational intelligence paradigms, including fuzzy logic in the 1980s. This method is well known not only as universal approximations but also for their capability of learning. A system with unknown dynamics is approximated or represented by a number of fuzzy rules. Extracting fuzzy rules from data has been investigated. This is true as long as the dynamics can be described in terms of fuzzy rules. Because of the intelligence of describing nonlinear systems using fuzzy rule-based systems has been successful in nonlinear areas [3].

Many applications however show that a time series is not necessarily stationary. This example includes indices related to economy, consumption, production and many other fields. A non stationary system, in strict sense, cannot be represented by fuzzy rules, because a fixed number of rules can describe a time invariant system, which obviously rules out the non-stationarity [3].

In the present work of dissertation, a load forecasting using conventional method integrated with the unconventional method known as Fuzzy Time Series method is presented. The data required for this dissertation work is collected from the Hydrology Deptt. and different substations in IIT Roorkee campus. The traditional methods are based on the statistical and probabilistic approaches but it may not be quite suitable to apply purely mathematical modes to the data generated. Fuzzy time series is the new approaches based on the possibility theory and fuzzy auto-regression and apply it to the analysis of time-series data of the electric load forecasting. The proposed fuzzy time series mode can be constructed in simpler procedure than the conventional approaches.

5.2 ALGORITHMIC PROCEDURE

Step 1:

The historical load data for 30 days is taken randomly selected during November 2003. The universe of discourse 'U' is defined within the available data based on which the fuzzy sets will be defined [1]. If the minimum engagement (D_{min}) is 289 and the maximum engagement (D_{max}) is 766.36, then the universe U is defined as [D_{min}-D_1, D_{max}+ D_2] where D_1 and D_2 are two positive numbers. For D_1=89 and D_2=33.14. Hence the universe of discourse 'U' for the method is the interval of U [400, 1100].

Step 2:

The universe of discourse 'U' is partitioned into five equal lengths. The intervals are chosen as,

U_1 = [400, 540], U_2 = [540, 680], U_3 = [680, 720], U_4 = [720, 960], U_5 = [960, 1100].

Step 3:

Fuzzy sets are defined on the universe. First some linguistic values are determined. If A_1= (many), A_2= (many, many), A_3= (very many), A_4= (too many), A_5= (too many many) be the possible values, next all fuzzy sets are labeled with possible linguistic values. If U_1, U_2, ...,U_5 be chosen as the elements of each fuzzy set, then each U_k (k=1,2,.......,5) belonging to Ai determines the membership of U_1, U_2,.........,U_5 to each A_i(i=1,2,.....,5). If U_k belongs to A_i, the membership will be 1; if U_k does not belong to Ai at all, the membership will be zero; otherwise a number from (0, 1) is chosen as the degree to which U_k belongs to A_i.

Step 4:

The available data are justified based on the Gaussian function, which is furnished in Table 1 as shown in Appendix: A.

Step 5:

Then the fuzzy sets and fuzzy conditional statements have developed such logical relationships as "if and then statements" like if the engagement of year I is A_k, then that of year i+1 is A_j and so on.

Using this, 'if and then statements', all relationships are developed as follows:

$A_2 \to A_1$, $A_1 \to A_3$, $A_3 \to A_4$, $A_4 \to A_5$, $A_5 \to A_3$, $A_3 \to A_5$, $A_5 \to A_1$, $A_1 \to A_3$, $A_3 \to A_4$, $A_4 \to A_3$, $A_3 \to A_4$, $A_4 \to A_3$, $A_3 \to A_1$, $A_1 \to A_3$, $A_3 \to A_4$, $A_4 \to A_3$, $A_3 \to A_2$, $A_2 \to A_1$, $A_1 \to A_3$, $A_3 \to A_4$, $A_4 \to A_2$, $A_2 \to A_4$, $A_4 \to A_5$, $A_5 \to A_2$, $A_2 \to A_1$, $A_1 \to A_1$, $A_3 \to A_3$, $A_4 \to A_4$, $A_3 \to A_3$

The repeated relationship is being counted only once.

If an operator 'X' of two vectors is defined and if C and B are row vectors of dimension m and

$D = (d_{ij}) = C^T \times B$, then the element d_{ij} of matrix D at row i and column j is defined as;

$$d_{ij} = \min (C_i, B_j) \ (i, j = 1, m) \tag{5.1}$$

where

C_i and B_j are the i^{th}, and the j^{th} element of C and B respectively.

If it is assumed that,

$$\begin{aligned}
&R_1 = A_2^T \times A_1,\ R_2 = A_1^T \times A_3,\ R_3 = A_3^T \times A_4,\ R_4 = A_4^T \times A_5,\\
&R_5 = A_5^T \times A_3,\ R_6 = A_3^T \times A_5,\ R_7 = A_5^T \times A_1,\ R_8 = A_1^T \times A_3,\\
&R_9 = A_3^T \times A_4,\ R_{10} = A_4^T \times A_3,\ R_{11} = A_3^T \times A_4,\ R_{12} = A_4^T \times A_3,\\
&R_{13} = A_3^T \times A_1,\ R_{14} = A_1^T \times A_3,\ R_{15} = A_3^T \times A_4,\ R_{16} = A_4^T \times A_3,\\
&R_{17} = A_3^T \times A_2,\ R_{18} = A_2^T \times A_1,\ R_{19} = A_1^T \times A_3,\ R_{20} = A_3^T \times A_4,\\
&R_{21} = A_4^T \times A_2,\ R_{22} = A_2^T \times A_4,\ R_{23} = A_4^T \times A_5,\ R_{24} = A_5^T \times A_2,\\
&R_{25} = A_1^T \times A_1,\ R_{26} = A_1^T \times A_1,\ R_{27} = A_3^T \times A_3,\ R_{28} = A_4^T \times A_4,\\
&R_{29} = A_3^T \times A_3
\end{aligned} \tag{5.2}$$

From this above equation, the following relation is obtained,

$$R\ (t, t\text{-}1) = R = U.\ R_i \tag{5.3}$$

where, R is a 5 x 5 matrix,
U is a union operator

Using this, relation R for load forecasting is calculated.
Then the forecasting model is defined as,

$$A_i = A_{i-1} . R \tag{5.4}$$

where A_{i-1} is the load of year 'i -1' and

A_i is the load forecasted of year 'i 'in terms of fuzzy sets and

'.' is the max- min operator.

Step 6:

The forecasted output is interpreted. The calculated results using step (5) are actually all fuzzy sets. If is necessary to translate the fuzzy output into a regular number which is an equivalent scalar. This step is called *"defuzzification."*

The following principles are used to interpret the forecasting results:

1. If the membership of an output has only one maximum, the midpoint of the interval corresponding to the maximum is selected as the forecasted value.
2. If the membership of an output has two or more consecutive maximum, the midpoint of the corresponding conjunct intervals is selected as the forecasted value.
3. The fuzzy output is standardized and the midpoint of each interval is used to calculate the centroid of the fuzzy set as the forecasted value.

The predicted values for the load forecasting are furnished in Table 2, given in appendix.

Step 7:

Finally, the forecasted error and the average forecasting error are calculated by using the following formulae,

$$\text{Forecasting error} = \frac{|(\text{Forecasted value - Actual value})|}{(\text{Actual value})} \times 100$$

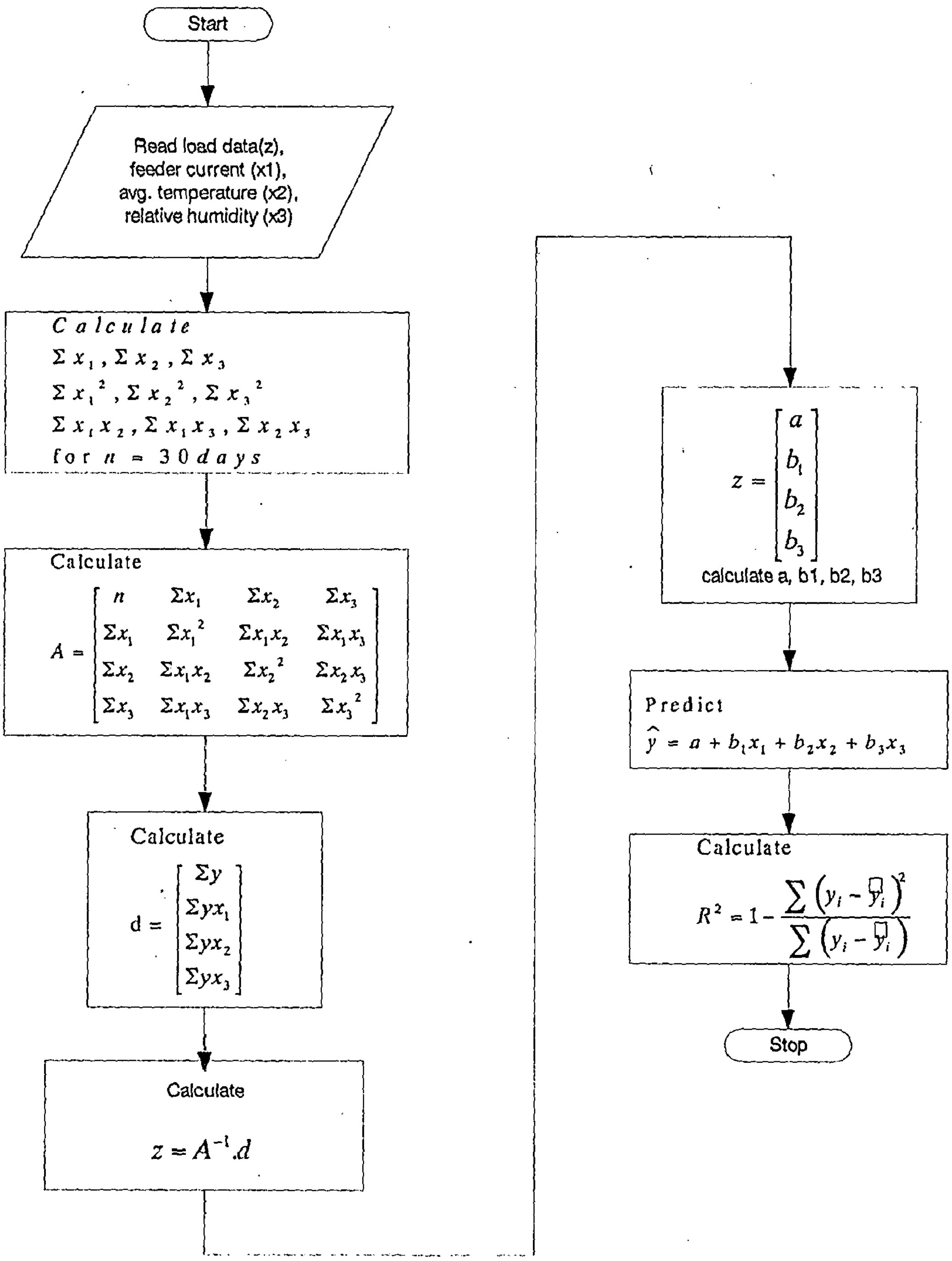
$$\text{Average forecasting error} = \frac{\text{Sum of forecasting errors}}{\text{Total number of errors}} \tag{5.5}$$

It is observed that the forecasting error changes and average forecasting error is also calculated by using above formulas. The change of estimated values is shown in Fig. and the forecasting error is also furnished in Table 3 given in the appendix. B.

5.3 FLOWCHARTS

The flowchart for the development of program for load forecasting techniques has been developed in following two sections. Flowchart given in subsection stands for the initializing and identifying the parameters of load model using the conventional method as multiple regression analysis. The values of parameters calculated in these sections are passed in the non-conventional techniques used as fuzzy time series for the next flowchart. And these two methods are finally compared. First flowchart gives the details about the conventional method used for load forecasting using the one month load data, current values and the weather variables such as average temperature and relative humidity to check the prediction of next 30 month load data.

2. Flowchart for Load Forecasting by Multiple Regression method



CHAPTER: 6

METHODOLOGY ADOPTED

6.1 BASICS OF THE METHOD

The present work is done for 30 days load data taken for Nov.2003 for IITR campus for load forecasting and is tested for Fuzzy Time Series methodology which is the integration of conventional and non-conventional method by using MATLAB programming.

The second model for load forecasting is conventional method used contains the adaptive load model which uses the multiple regression analysis which takes the variables as load current, weather variables such as average temperature and relative humidity. These data are collected from the Hydrology Deptt. and different substations in IIT Roorkee campus.

These two models are tested for the one month data collection for IITR campus and compared with the realistic load data. This feature enables the model to update its own parameters according to its past performance. So, the load forecasting model is a suitable for online usage. Simplified flowchart of load forecasting procedure using these techniques is already shown. The program evaluation for the load forecasting techniques is explained under the following sections. For that the data required for load forecasting techniques are as follows.

6.2 DATA REQUIREMENTS

To test the developed MATLAB program, load data for 30 days i.e. from 1^{st} Nov. 2003 to 30^{th} Nov. 2003 has been collected from the substation located inside the campus of Indian Institute of Technology, Roorkee. The required data are taken from 11KV feeder supplying power to Indian Institute of Technology, Roorkee. In campus, there are five substations, which are taking the input from 33KV main receiving substation under Uttaranchal Power Corporation supplying to the 11kv substation. The five substations are located near, Cautley Bhawan, Welding Research Laboratory, High Voltage laboratory, Wind tunnel substation and the old substation near Physics Deptt. The load data for three months (1^{st} Oct.--31^{st}Dec.2003) is collected from these five substations. The daily record of weather variables, average temperature and relative humidity has also been obtained from Hydrology department of I.I.T. Roorkee for the same period. The data is compiled in terms of various pre-decided variables for load forecasting using regression-based analysis.

6.3 LOAD FORECASTING USING CONVENTIONAL METHOD

In the present work, computer program in MATLAB for the probabilistic forecasting of daily power system loads with lead times of 1 to 30 days is developed. The forecasts produced are based on both the historical daily load data and the information from the latest weather forecast. The conventional method used for load forecasting is multiple regression analysis.

Multiple Regression model is the econometric model, involves the multiple regression between the load demand and various other variables that can be influenced on the load demand. The metrological information is reduced to a number of specific factors designated as follows:

i. Effective temperature, T
ii. Cooling power of the wind, W
iii. Illumination index, L
iv. Rate of precipitation, P

The general model can be expressed as:

$$y_i = \alpha + \beta_1 x_{1i} + \beta_1 x_{1i} + \dots\dots\dots\dots + \beta_1 x_{1i} + u_i \tag{6.1}$$

where, i=1, 2,……, n

The residuals u_i is due to the measurement errors in y and the errors in the specification of the relationship between y and the x. Under the assumption pf least squares gives the estimators of . , . $_1$, . $_2$, . $_k$. It has minimum variance among the class of linear unbiased estimators.

Let's assume for three explanatory variables, x_1, x_2 and x_3.

Hence it becomes,

$$y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i \tag{6.2}$$

where,

i = 1, 2,……., n

According to the least squares method, the estimators

$$Q=\sum(y_i-\hat{\alpha}-\hat{\beta}_1x_{1i}-\hat{\beta}_2x_{2i}-\hat{\beta}_3x_{3i})^2 \qquad (6.3)$$

$$S_{33}=\sum x_{3i}^2-n\bar{x}_3^2 \qquad (6.4)$$

Differentiate Q with respect to α, β_1, β_2, β_3

$$\frac{\partial Q}{\partial\hat{\beta}_1}=0$$
$$\Rightarrow\sum 2(y_i-\hat{\alpha}-\hat{\beta}_1x_{1i}-\hat{\beta}_2x_{2i}-\hat{\beta}_3x_{3i})(-x_{1i})=0$$

$$\frac{\partial Q}{\partial\hat{\beta}_2}=0$$
$$\Rightarrow\sum 2(y_i-\hat{\alpha}-\hat{\beta}_1x_{1i}-\hat{\beta}_2x_{2i}-\hat{\beta}_3x_{3i})(-x_{2i})=0$$

$$\frac{\partial Q}{\partial\hat{\beta}_3}=0$$
$$\Rightarrow\sum 2(y_i-\hat{\alpha}-\hat{\beta}_1x_{1i}-\hat{\beta}_2x_{2i}-\hat{\beta}_3x_{3i})(-1)=0$$

$$\frac{\partial Q}{\partial\hat{\alpha}}=0$$
$$\Rightarrow\sum 2(y_i-\hat{\alpha}-\hat{\beta}_1x_{1i}-\hat{\beta}_2x_{2i}-\hat{\beta}_3x_{3i})(-1)=0 \qquad (6.5)$$

These four equations are called the normal equations. It can also be simplified as:

$$\sum y_i=n\hat{\alpha}+\hat{\beta}_1\Sigma x_{1i}+\hat{\beta}_2\Sigma x_{2i}+\hat{\beta}_3\Sigma x_{3i}$$
$$\bar{y}=\hat{\alpha}+\hat{\beta}_1\bar{x}_1+\hat{\beta}_2\bar{x}_2+\hat{\beta}_3\bar{x}_3 \qquad (6.6)$$

where,

$$\bar{y}=\frac{1}{n}\sum y_i,\bar{x}_1=\frac{1}{n}\sum x_{1i},\bar{x}_2=\frac{1}{n}\sum x_{2i},\bar{x}_3=\frac{1}{n}\sum x_{3i} \qquad (6.7)$$

The equation (6.4) can also be written as,

$$S_{23}=\sum x_{2i}x_{3i}-n\bar{x}_2\bar{x}_3 \qquad (6.8)$$

Substituting the value of α from eq. (6.5), we get

$$\sum x_{1i}y_i = n\bar{x}_1(\bar{y} - \hat{\beta}_1\bar{x}_1 - \hat{\beta}_2\bar{x}_2 - \hat{\beta}_3\bar{x}_3) + \hat{\beta}_1\sum x_{1i}^2 + \hat{\beta}_2\sum x_{1i}x_{2i} + \hat{\beta}_3\sum x_{1i}x_{3i} \qquad (6.9)$$

Now let us define the following:

$$\begin{aligned} S_{11} &= \sum x_{1i}^2 - n\bar{x}_1^2 \\ S_{12} &= \sum x_{1i}x_{2i} - n\bar{x}_1\,\bar{x}_2 \\ S_{13} &= \sum x_{1i}x_{3i} - n\bar{x}_1\,\bar{x}_3 \end{aligned} \qquad (6.10)$$

Then Eq. (6.10) can be written as

$$\begin{aligned} S_{1y} &= \hat{\beta}_1 S_{11} + \hat{\beta}_2 S_{12} + \hat{\beta}_3 S_{13} \\ S_{2y} &= \hat{\beta}_1 S_{12} + \hat{\beta}_2 S_{22} + \hat{\beta}_3 S_{23} \\ S_{3y} &= \hat{\beta}_1 S_{13} + \hat{\beta}_2 S_{23} + \hat{\beta}_3 S_{33} \end{aligned} \qquad (6.11)$$

$$\begin{aligned} S_{1y} &= \sum x_{1i}y_i - n\bar{x}_1\,\bar{y} \\ S_{2y} &= \sum x_{2i}y_i - n\bar{x}_2\,\bar{y} \\ S_{3y} &= \sum x_{3i}y_i - n\bar{x}_3\,\bar{y} \\ S_{yy} &= \sum y_i^2 - n\bar{y}^2 \end{aligned} \qquad (6.12)$$

By similar simplifications Eq. (6.11) and (6.12) can be written as

Now, we solve Eqs. (6.10) and (6.11) for β_1, β_2, and β_3. Then, we substitute these values in (6.12) and get an estimate of α.

The computational procedure for the load forecasting formulation would be:

1. First obtain all the means of variables : x_1, x_2, x_3, z

   where, x_1 is the load current in ampere

   x_2 is the average temperature in degree cel.

   X_3 is the relative humidity in percentage

   Z is the load demand in KW.

2. Then obtain all the sums of squares and sums of products:

$$\sum x_{1i}^2, \sum x_{2i}^2, \sum x_{3i}^2, etc$$

$$\sum x_1, \sum x_2, \sum x_3, \sum z,$$

3. Then obtain S_{11}, S_{12},..........., etc and solve eqs. (6.10), (6.11) and (6.12) to get

$\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3.$

4. Substituting these values in eqn. (6.6).we get

$$\hat{\alpha} = \bar{z} - \hat{\beta}_1 \bar{x}_1 - \hat{\beta}_2 \bar{x}_2 - \hat{\beta}_3 \bar{x}_3$$

The solution of eqs. (6.10), (6.11) and (6.12) has to be done by successive elimination. This can be done systematically by a process known as, "Doolittle method".

$$\bar{S}_{zz} - \hat{\beta}_1 S_{xz}$$

In simple regression, the residual sum of squares was defined as

$$S_{22} = \sum x_{2i}^2 - n\bar{x}_2^2 \tag{6.13}$$

and r^2_{xy} is defined as,

$$\hat{\beta}_1 S_{xz} / S_{zz} \tag{6.14}$$

The analogous expressions in multiple regression are,

$$R^2_{z.x_1x_2x_3} = \frac{\hat{\beta}_1 S_{1z} - \hat{\beta}_2 S_{2z} - \hat{\beta}_3 S_{3z}}{S_{zz}} \tag{6.15}$$

The notation $y.x_1x_2x_3$ is used in multiple correlation. The variable before dot is the explained variable and the variables after the dot are the explanatory variables.

The formula derived above can be written as matrix form as given below;

and

$$\bar{S}_{xx}\hat{\beta} = S_{xz} \tag{6:16}$$

The equations can also be written as;

$$RSS = S_{zz} - \hat{\beta}_1 S_{1z} - \hat{\beta}_2 S_{2z} - \hat{\beta}_3 S_{3z} \tag{6.17}$$

Assume that there are no linear dependencies among the x's.

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{bmatrix}$$

Then S_{xx} is a singular matrix and its inverse exists.

Pre multiplying both side of eq. (6.17) by S_{xx}^{-1}, we get

$$\sigma^2 C = \sigma^2 S_{xx}^{-1} \qquad (6.18)$$

This equation is analogous to the equation obtained in simple regression. Also if we denote as,

$$\hat{\beta} = S_{xx}^{-1} S_{xz} \qquad (6.19)$$

the equation (6.18), (6.19) and (6.20) can be written as:

$$S_{xx} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix}$$

$$S_{xz} = \begin{bmatrix} S_{1z} \\ S_{2z} \\ S_{3z} \end{bmatrix}$$

This formula for solving the regression analysis is very analogous to that of simple regression.

Lastly, if $\sigma^2 C$ is the matrix of variances and covariance of $\beta_1, \beta_2, \beta_3$.

Then,

$$\hat{\alpha} = \bar{z} - \hat{\beta}'\bar{x}$$

$$RSS = S_{zz} - \hat{\beta}' S_{xz}$$

$$R_{z.x_1x_2}{}^2 = \frac{\hat{\beta}' S_{xz}}{S_{zz}}$$

The variance for simple regression is also the same as for variance. In deriving the functional relationships between the variation of demand and the specific metrological factors to which it is sensitive, the basic assumption is again made that the weather sensitive component of the demand may be expressed as sum of functions of the respective metrological factors.

6.4 LOAD FORECASTING USING FUZZY TIME SERIES

The historical load data for 30 days is selected during November 2003. The universe of discourse 'U' is defined within the available data based on which the fuzzy sets is defined. The detail of calculations of universe of discourse for the same load is given in section 5. Hence the universe of discourse 'U' for the method is the interval of U [400, 1100].

This universe of discourse 'U' is partitioned into five equal lengths. The intervals for universe of discourse are so chosen in the decided zone as like 140 in between these intervals,

U_1 = [400, 540], U_2 = [540, 680], U_3 = [680, 820], U_4 = [820, 960], U_5 = [960, 1100].

Fuzzy sets are defined on the basis of this universe. For this, some linguistic values are determined as A_1= (many), A_2= (many, many), A_3= (very many), A_4= (too many), A_5= (too many many) and all fuzzy sets are labeled with possible linguistic values. If U_1, U_2, ...,U_5 be chosen as the elements of each fuzzy set, then each U_k (k =1, 2,.......,5) belonging to Ai determines the membership of U_1, U_2,........,U_5 to each A_i(i =1, 2,.....,5). If U_k belongs to A_i, the membership will be 1; if U_k does not belong to Ai at all, the membership will be zero; otherwise a number from (0, 1) is chosen as the degree to which U_k belongs to A_i.

The available data are justified based on the Gaussian function, which is furnished in Table 1 as shown in Appendix: A. Then the fuzzy sets and fuzzy conditional statements have developed such logical relationships as *"if and then statements"* like if the engagement of year I is A_k, then that of year i+1 is A_j and so on.

Using this, 'if and then statements', all relationships are developed as given in (5.2) for fuzzy sets and the repeated relationship is being counted only once.

If an operator 'X' of two vectors is defined and if C and B are row vectors of dimension m and $D = (d_{ij}) = C^T \times B$, where C_i and B_j are the i^{th}, and the j^{th} element of C and B respectively. The final relation for fuzzy sets is calculated from the equation (5.3)

The load forecasting model can be defined as in (5.4) by using max-min operator function.

Hence, the forecasted output is interpreted. The calculated results obtained are actually all fuzzy sets. For that, it is necessary to convert the fuzzy output into a regular number which is an equivalent scalar, known as *"defuzzification."* There are different methodologies for defuzzification process such as center-of-area, center of sum and maximum-of-minima methods, etc. Here, the most convenient and general used method for defuzzification is center of area or centroid of center of gravity method. But in the present work, the defuzzification process for converting the fuzzy output into the desired one is done by some principle used from the centriod method, discussed in chapter 5.

Then, the predicted values for the load forecasting are furnished in Table 2, given in appendix A. and the forecasted error and the average forecasting error are calculated by using (5.5)

PROBLEM FORMULATION

Universe of Discourse:

U = [400, 1100]

Universe of discourse is partitioned into equal lengths.

U1 = [400, 540]

U2 = [540, 680]

U3 = [680. 820]

U4 = [820, 960]

U5 = [960, 1100]

Some linguistic values for fuzzy logics are defined as,

A1 = (many)

A_2= (many, many)

A_3= (very many)

A_4= (too many)

A_5= (too many many)

Fuzzy logical relationship is detailed in the table 6.1.

There will be 29 fuzzy logical relationships, which is calculated by the fuzzy sets from given table. The relational operators on the basis of fuzzy *if and then* rule are also given in table 6.2:

Nov. 2003

Table 6.1 Actual Load and Fuzzy Set

Days	Load (KW) y	A_1	A_2	A_3	A_4	A_5	Fuzzy set
1.	599.66	0.8	1.0	0.6	0.0	0.0	$A_2 = b_1$
2.	485.90	1.0	0.8	0.0	0.0	0.0	$A_1 = b_2$
3.	759.15	0.0	0.6	1.0	0.8	0.0	$A_3 = b_3$
4.	851.60	0.0	0.0	0.8	1.0	0.6	$A_4 = b_4$
5.	985.90	0.0	0.0	0.0	0.8	1.0	$A_5 = b_5$
6.	815.40	0.0	0.6	1.0	0.8	0.0	$A_3 = b_6$
7.	1021.9	0.0	0.0	0.0	0.8	1.0	$A_5 = b_7$
8.	460.80	1.0	0.8	0.0	0.0	0.0	$A_1 = b_8$
9	521.00	1.0	0.8	0.0	0.0	0.0	$A_1 = b_9$
10.	748.27	0.0	0.8	1.0	0.6	0.0	$A_3 = b_{10}$
11	806.50	0.0	0.6	1.0	0.8	0.0	$A_3 = b_{11}$
12	958.70	0.0	0.0	0.6	1.0	0.8	$A_4 = b_{12}$
13	825.34	0.0	0.0	0.8	1.0	0.6	$A_4 = b_{13}$
14	735.50	0.0	0.8	1.0	0.6	0.0	$A_3 = b_{14}$
15	756.55	0.0	0.6	1.0	0.8	0.0	$A_3 = b_{15}$
16	878.00	0.0	0.0	0.8	1.0	0.6	$A_4 = b_{16}$
17	722.50	0.0	0.8	1.0	0.6	0.0	$A_3 = b_{17}$
18	471.30	1.0	0.8	0.0	0.0	0.0	$A_1 = b_{18}$
19	812.35	0.0	0.6	1.0	0.8	0.0	$A_3 = b_{19}$
20	870.30	0.0	0.0	0.8	1.0	0.6	$A_4 = b_{20}$
21	713.05	0.0	0.8	1.0	0.6	0.0	$A_3 = b_{21}$
22	585.90	0.8	1.0	0.6	0.0	0.0	$A_2 = b_{22}$
23	435.70	1.0	0.8	0.0	0.0	0.0	$A_1 = b_{23}$
24	738.95	0.0	0.8	1.0	0.6	0.0	$A_3 = b_{24}$
25	828.75	0.0	0.0	0.8	1.0	0.6	$A_4 = b_{25}$
26	645.80	0.6	1.0	0.8	0.0	0.0	$A_2 = b_{26}$
27	893.44	0.0	0.0	0.6	1.0	0.8	$A_4 = b_{27}$
28	962.70	0.0	0.0	0.0	0.8	1.0	$A_5 = b_{28}$
29	620.85	0.6	1.0	0.8	0.0	0.0	$A_2 = b_{29}$
30	534.00	1.0	0.8	0.0	0.0	0.0	$A_1 = b_{30}$

Table 6.2 Fuzzy Logic Relationship with Relational Operators

Fuzzy logic relationship	Relational operators	Fuzzy logic relationship
$A_2 \rightarrow A_1$	R_1	$A_2^T \times A_1$
$A_1 \rightarrow A_3$	R_2	$A_1^T \times A_3$
$A_3 \rightarrow A_4$	R_3	$A_3^T \times A_4$
$A_4 \rightarrow A_5$	R_4	$A_4^T \times A_5$
$A_5 \rightarrow A_3$	R_5	$A_5^T \times A_3$
$A_3 \rightarrow A_5$	R_6	$A_3^T \times A_5$
$A_5 \rightarrow A_1$	R_7	$A_5^T \times A_1$
$A_1 \rightarrow A_3$	R_8	$A_1^T \times A_3$
$A_3 \rightarrow A_4$	R_9	$A_3^T \times A_4$
$A_4 \rightarrow A_3$	R_{10}	$A_4^T \times A_3$
$A_3 \rightarrow A_4$	R_{11}	$A_3^T \times A_4$
$A_4 \rightarrow A_3$	R_{12}	$A_4^T \times A_3$
$A_3 \rightarrow A_1$	R_{13}	$A_3^T \times A_1$
$A_1 \rightarrow A_3$	R_{14}	$A_1^T \times A_3$
$A_3 \rightarrow A_4$	R_{15}	$A_3^T \times A_4$
$A_4 \rightarrow A_3$	R_{16}	$A_4^T \times A_3$
$A_3 \rightarrow A_2$	R_{17}	$A_3^T \times A_2$
$A_2 \rightarrow A_1$	R_{18}	$A_2^T \times A_1$
$A_1 \rightarrow A_3$	R_{19}	$A_1^T \times A_3$
$A_3 \rightarrow A_4$	R_{20}	$A_3^T \times A_4$
$A_4 \rightarrow A_2$	R_{21}	$A_4^T \times A_2$
$A_2 \rightarrow A_4$	R_{22}	$A_2^T \times A_4$
$A_4 \rightarrow A_5$	R_{23}	$A_4^T \times A_5$
$A_5 \rightarrow A_2$	R_{24}	$A_5^T \times A_2$
$A_2 \rightarrow A_1$	R_{25}	$A_2^T \times A_1$
$A_1 \rightarrow A_1$	R_{26}	$A_1^T \times A_1$
$A_3 \rightarrow A_3$	R_{27}	$A_3^T \times A_3$
$A_4 \rightarrow A_4$	R_{28}	$A_4^T \times A_4$
$A_3 \rightarrow A_3$	R_{29}	$A_3^T \times A_3$

Using equation (5.3),

The following relation is obtained.

$$R=\begin{bmatrix} 1.0 & 0.8 & 1.0 & 0.8 & 0.48 \\ 1.0 & 0.8 & 0.8 & 1.0 & 0.8 \\ 1.0 & 1.0 & 1.0 & 1.0 & 1.0 \\ 0.8 & 1.0 & 1.0 & 1.0 & 1.0 \\ 1.0 & 1.0 & 1.0 & 0.64 & 0.8 \end{bmatrix}$$

Using this R, relation operator, the load forecasting model is defined as,

$A_i = A_{i-1} . R$

where, A_{i-1} is the load of year '*i -1*' and

A_i is the load forecasted of year '*i* 'in terms of fuzzy sets and

'.' is the max- min operator

With the help of defuzzification process, generally centroid method is used, the fuzzy output for load forecasting is translated into a regular number. For this some of the forecasting result carried out by defuzzification principles as discussed later.

The forecasting error and the average forecasting error are determined with the help of the error formulae as listed in equation (6.23).

The realistic load data for the prediction month for proposed method is collected from the different five substations and observed with the 1st conventional method and the 2nd fuzzy time series method, which is shown in the graphs (Appendix B). The forecasting error for both the method is calculated and compared.

CHAPTER: 7

RESULTS AND CONCLUSIONS

The developed program for load forecasting has been run on actual load data and weather variable records obtained from different five substations located in IITR campus and Hydrology department of IIT Roorkee, respectively. The following assumptions are made;

1. The daily load data for a period of 30 days from 1st Nov. 2003 to 30th Nov. 2003 are tested for fuzzy time series model for the next 30 days load prediction.
2. The accuracy of this adopted methodology can be further improved by using the hourly load forecasts instead of daily average values and considering the weekly holidays.
3. This model for fuzzy time series is compared with the conventional approach and found an average load forecasting error of -6.11%.

This method is well suited for load forecasting as the part of the program is deal with the historical data worked out for offline and only the adaptive part of it done on this basis. But, this program does not have the provision for bad data rejection and also for special holiday forecasting. The load forecasting for next month is done by running the same program on time series basis for 30 times. The comparison between the fuzzy time series approach and conventional method is given in the graph. Fig. 7.4.

After calculating the average error, it is found that fuzzy time series has an average error of -6.11% whereas conventional approach for load forecasting has an average error of -29.04%. Therefore the estimated data based on fuzzy time series will be used for the assessment of futuristic forecasted

CONCLUSIONS

To check the accuracy of the fuzzy time series approach over the conventional method, same steps are repeated about 30 times for next 30 days load forecasting. This is the new approach based on the conventional time series analysis and non-conventional fuzzy logic concept, applied to the analysis of time series data0 for electric load. This proposed method has two important capacities as representation of non-stationary data and reduction of errors in predicting a time series.

In time series approach, the inability to accurately describe nonlinear relationship between loads and weather relative data is not described. This drawback in time series conventional method is overcome by using fuzzy time series for short term load forecasting. Fuzzy time series approach is tested for the 30 days load data for better result. Thus the proposed fuzzy time series model is simpler procedure than the conventional approach.

The accuracy of this adopted methodology can be further improved by considering the hourly weather forecasts instead of daily average values and special weekly holidays. The future scope of the work is that the new development of fuzzy time series can be added with the fuzzy autoregressive model for load forecasting considering the weather sensitive variables with the special holidays which may give the better result as compared to the proposed one. In fuzzy time series, the holidays and Sunday is not considered as in that day, there may be load decreasing. This is not come into picture in fuzzy time series load forecasting. This may be further improved by considering it for the next 30 days prediction.

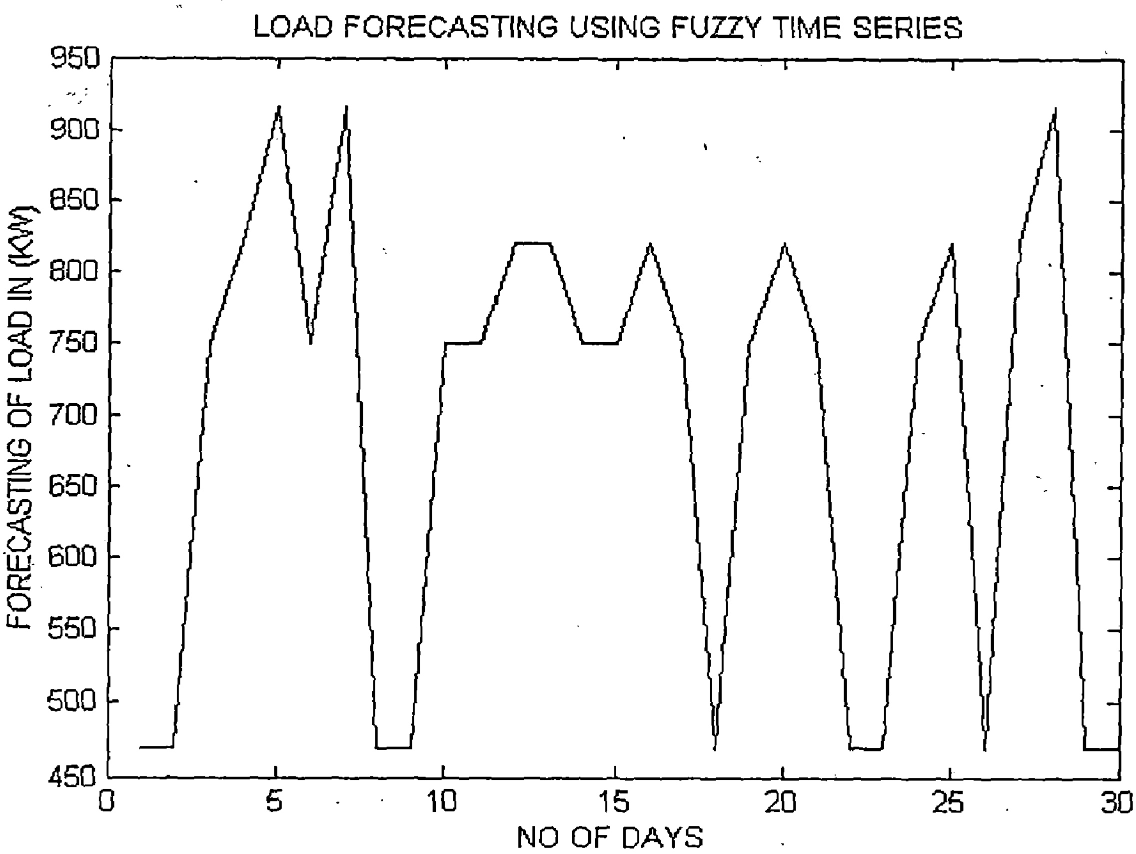


FIG. 7.1 GRAPH FOR LOAD FORECASTING BY USING FUZZY TIME SERIES APPROACH

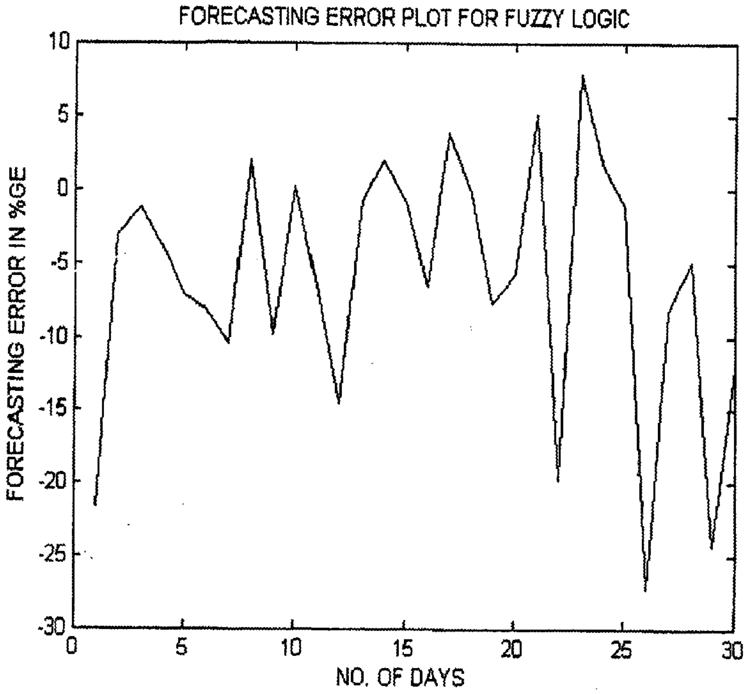


FIG. 7.2 GRAPH FOR FORECASTING ERROR IN %GE BY FUZZY TIME SERIES APPROACH

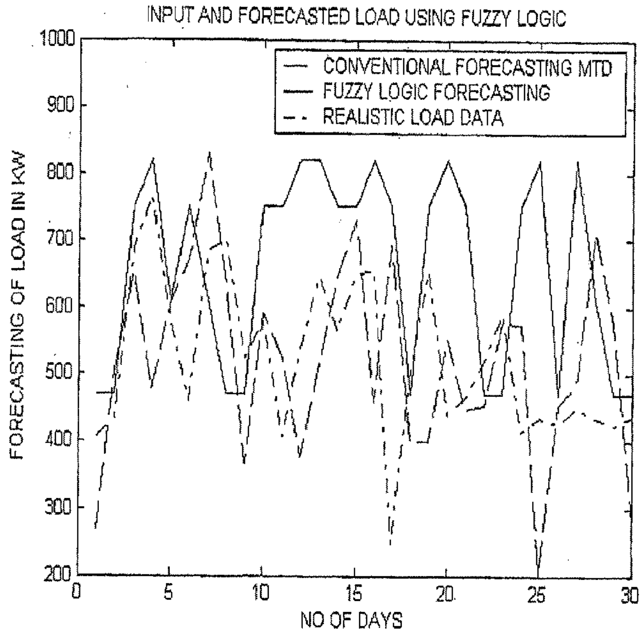


FIG. 7.3 COMPARISION BETWEEN CONVENTIONAL METHOD, REALISTIC LOAD DATA AND FUZZY TIME SERIES

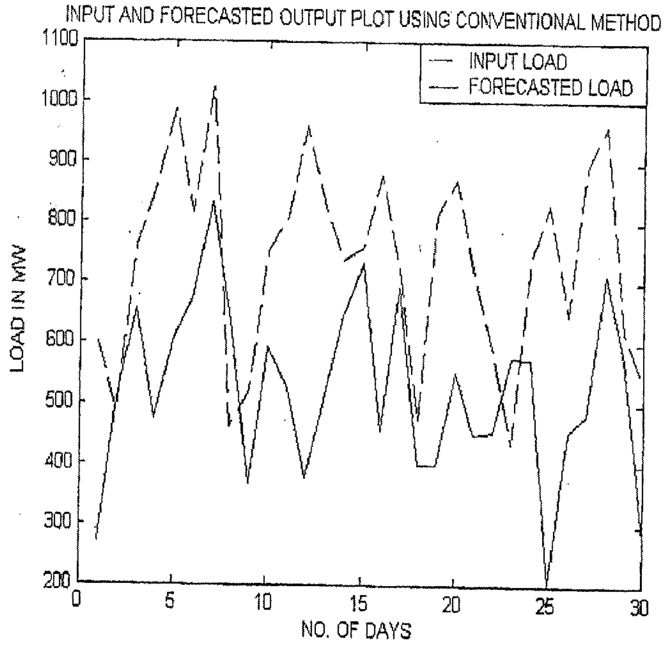


Fig. 7.4 GRAPH FOR PREDICTED LOAD AND INPUT LOAD USING CONVENTIONAL METHOD

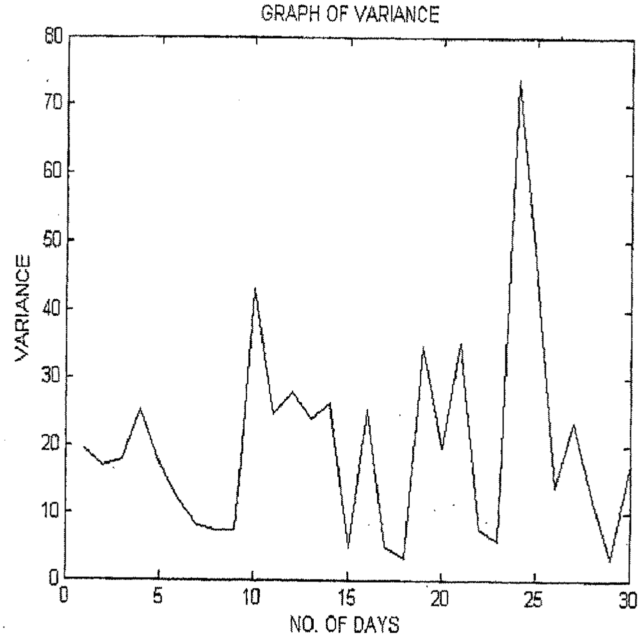


FIG. 7.5 GRAPH OF VARIANCE FOR LOAD FORECASTING USING CONVENTIONAL METHOD

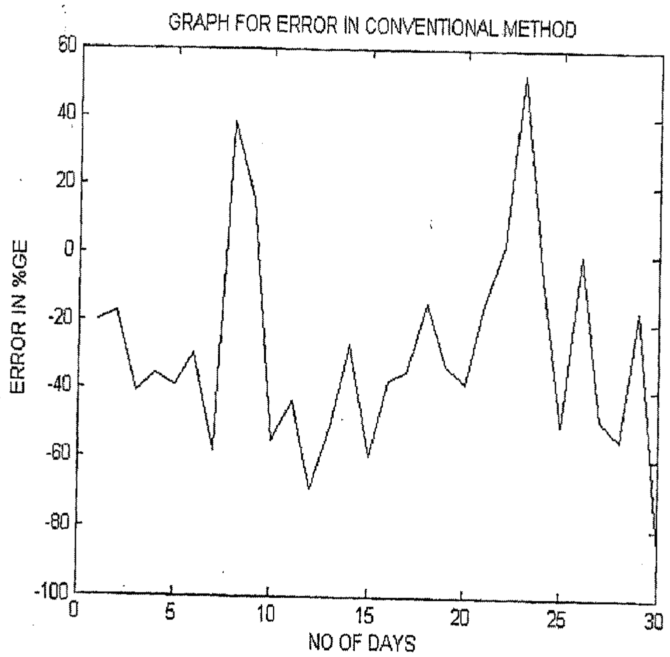


FIG. 7.6 GRAPH FOR ERROR IN LOAD FORECASTING BY CONVENTIONAL METHOD

TABLE 7.1 PREDICTED LOAD BY FUZZY TIME SERIES APPROACH

No. of Days	Actual Load in (KW)	Predicted Load in (KW) by Fuzzy Time Series Approach
1	599.66	470
2	485.90	470
3	759.15	750
4	851.60	820
5	985.90	915
6	815.40	750
7	1021.9	915
8	460.80	470
9	521.00	470
10	748.27	750
11	806.50	750
12	958.70	820
13	825.34	820
14	735.50	750
15	756.55	750
16	878.00	820
17	722.50	750
18	471.30	470
19	812.35	750
20	870.30	820
21	713.05	750
22	585.90	470
23	435.70	470
24	738.95	750
25	828.75	820
26	645.80	470

27	893.44	820
28	962.70	915
29	620.85	470
30	534.00	470

TABLE 7.2 COMPARISION OF FORECASTING ERROR BETWEEN FUZZY TIME SERIES APPROACH AND COVENTIONAL METHOD

No. of Days	Forecasting Error in Fuzzy Time Series Approach	Forecasting Error in Conventional Method
1	-21.6223	-20.1257
2	-3.0928	-17.4435
3	-1.2053	-40.9414
4	-3.7107	-35.7803
5	-7.1914	-38.8356
6	-8.0206	-29.7225
7	-10.4609	-58.2525
8	1.9965	38.2345
9	-9.7889	15.6973
10	0.2312	-55.1135
11	-7.0056	-43.6523
12	-14.4675	-69.0674
13	-0.6470	-50.5179
14	1.9714	-27.0036
15	-0.8658	-59.4039
16	-6.6059	-37.6614
17	3.8062	-34.5791
18	-0.2758	-14.7498
19	-7.6753	-32.8242
20	-5.7796	-38.1670
21	5.1820	-14.5748
22	-19.7815	2.6039
23	7.8724	52.8743
24	1.4954	-7.0085
25	-1.0558	-50.1060
26	-27.2221	-0.3271

27	-8.2199	-48.0894
28	-4.9548	-54.6094
29	-24.2973	-16.5849
30	-12.0509	-85.3275

REFERENCES

[1]. J. P. Chaudhary, Dr. B. Sarkar, Prof. S. K. Mukherjee, "*A Comparison between the Fuzzy Time Series and Genetic Algorithm in the Prediction of Engineering Manpower*", IE (India) Journal-CP, Vol. 83, May 2002, pp. 6-11.

[2] K. Ozawa, T. Niimura, T. Nakashima, *"Fuzzy Time-Series Model of Electric Power Consumption"*, proceedings of 1999 IEEE Canadian Conference and Computer Engg., May 9-12 1999, Vol. II, pp. 1195-1198.

[3] Chao-Chih Tsai, Shun-Jyh Wu, *"A Fuzzy Time Series Prediction method based on Consecutive Values"*, IEEE Trans. Conference on Electrical and Computer Engg. 2001, Vol. II, pp. 703-707.

[4] Dr. Keith Holbert, "*Short Term Load Forecasting with Fuzzy Logic Systems*", http://www.ceaspub.eas.asu.edu/powerzone/FuzzyLogic/index.htm

[5] A. M. Al-Kandari, S. A. Soliman, M. E. El-Hawary, *"Fuzzy Systems Application to Electric Short-Term Load Forecasting: Part I Problem Formulation*", IEEE PWRS-2003, pp. 125-130.

[6] A. M. Al-Kandari, S. A. Soliman, M. E. El-Hawary, *"Fuzzy Systems Application to Electric Short-Term Load Forecasting: Part II Computational Results*", IEEE PWRS- 2003, pp. 131-137.

[7] Mo-Yuen Chow, Hahn Tram, "Application *of Fuzzy Logic Technology for Spatial Load Forecasting"*, IEEE Trans. on Power Systems, Vol. 12, No. 3, Aug. 1997, pp. 1360-1366.

[8] Naveed Ahmed, "Fuzzy Logic Control using Matlab: Part I", workshop on Khwarzimic Science Society, University of Lahore, April 2003, http://www.khwarzimic.org .

[9] Martin T. Hagan, Suzanne M. Behr, *"The Time Series Approach to Short Term Load Forecasting"*, IEEE Trans. on Power Systems, Vol. PWRS-2, No. 3, Aug 1987, pp. 785-791.

[10] Slobodan Ruzie, Aca Vuckovie, Nikola Nikolie, *"Weather Sensitive method for Short Term Load Forecasting in Electric Power Utility of Serbia"*, IEEE Trans. on Power Systems, Vol. 18, No. 4, Nov. 2003, pp. 1581-1586.

[11] Box, Jenkins, *"Time Series Analysis- Forecasting and Control"*. Holden Day Publication, San Framscosco, 1976, pp. 47-72.

[12] Nima Amjady,*"Short-Term Hourly Load Forecasting Using Time-Series Modeling with Peak Load Estimation Capability"*, IEEE Trans. on Power Systems, Vol. 16, No. 3, Aug. 2001, pp. 498-505.

[13] Kazuhiro OZAWA, Takumi WATANABE, Masayasu KANKE, *"Fuzzy Auto-Regressive Model and Its Applications"*, IEEE First Conference on Knowledge-Based Intelligent Electronic Systems, Adelaide, Australia, 21-23 May 1997, pp. 112-117.

[14] M. S. Sachdeo, R. Billinton and C. A. Paterson,*" Representative Bibliography on Load Forecasting"*, IEEE Trans. on Power Apparatus and System, Vol. PAS-96, No. 2, March/April 1977, pp. 697-700.

[15] Pradeep C. Gupta and Keigo Yamad,*"Adaptive Short Term Load Forecasting of Hourly loads Using Weather Information"*, IEEE Trans. On Power Apparatus and Systems, Vol. PAS-91, No. 5, Sept. /Oct. 1972, pp. 2085-2094.

[16] R. Campo and P. Ruiz, *"Adaptive Weather Sensitive Short Term Load Forecast"*, IEEE Trans. On Power Systems, Vol. PWRS-2, No.3 Aug. 1987, pp. 592-600.

[18] E. H. Barakat, J. M. Al – Qasem. *"Methodology for weekly load forecasting"*, IEEE Trans. On Power System, Vol. 13, No. 4, Nov-1998, pp. 1548-1555.

APPENDIX: A

TABLE: A-1

LOAD DATA (KW) OBSERVED AT DIFFERENT FIVE SUBSTATIONS LOCATED IN INDIAN INSTITUTE OF TECHNOLOGY, ROORKEE CAMPUS FOR 30 DAYS FROM 1^{ST} NOV. 2003 TO 30^{TH} NOV. 2003.

No. of days.	Cautley Bhawan Substation	Wind Tunnel Substation	Physics Deptt. Substation	High Voltage Lab. Substation	W.R.L. substation	Total Load of campus in KW
1	176.40	96.00	154.56	17.60	155.10	599.66
2	188.60	72.80	29.00	21.60	173.90	485.90
3	164.85	101.50	306.00	30.40	156.40	759.15
4	217.10	112.00	325.80	35.70	161.00	851.60
5	173.80	115.50	446.60	33.80	216.10	985.9
6	167.50	94.50	319.50	31.50	202.40	815.4
7	186.00	254.80	369.00	28.80	183.30	1021.9
8	156.40	81.00	35.00	22.80	165.60	460.8
9	167.900	89.900	66.000	27.000	170.200	521.000
10	122.550	168.000	327.320	38.400	92.0.00	748.270
11	156.200	164.500	274.400	32.000	179.400	806.500
12	136.800	290.000	307.800	21.600	202.500	958.700
13	110.400	152.000	326.340	34.200	202.400	825.340
14	151.200	77.160	316.540	28.600	162.000	735.500
15	146.050	126.000	274.400	39.900	170.200	756.550
16	226.800	180.000	260.000	45.600	165.600	878.000
17	202.500	144.000	156.600	36.100	183.300	722.250
18	131.100	84.0.00	35.0.00	28.000	193.200	471.300
19	184.950	140.400	286.000	40.000	161.000	812.350
20	186.300	139.200	320.400	45.000	179.400	870.300
21	135.450	126.000	237.200	44.200	170.200	713.050

22	217.000	105.400	45.000	39.100	179.400	585.900
23	141.900	66.000	23.000	30.000	174.800	435.700
24	141.450	59.700	311.400	38.400	188.000	738.950
25	195.750	136.800	242.200	70.000	184.000	828.750
26	197.100	92.400	102.000	61.600	192.700	645.800
27	172.800	155.800	324.000	54.600	186.240	893.440
28	202.800	144.300	437.000	75.400	193.200	962.700
29	168.750	122.400	86.000	41.600	202.100	620.850
30	151.800	112.200	34.600	46.800	189.000	534.400

TABLE: A-2

LOAD CURRENT (AMP.) OBSERVED AT DIFFERENT FIVE SUBSTATIONS LOCATED IN INDIAN INSTITUTE OF TECHNOLOGY, ROORKEE CAMPUS FOR 30 DAYS FROM 1ST NOV. 2003 TO 30TH NOV. 2003.

Date	Cautley Bhawan Substation	Wind Tunnel Substation	Physics Deptt. Substation	High Voltage Lab. Substation	W.R.L. substation	Total Load Current in AMP.
1	240	320	250	160	240	970
2	230	260	240	180	260	1170
3	210	350	450	160	280	1450
4	240	300	450	210	280	1500
5	230	350	550	260	250	1630
6	210	350	450	210	260	1520
7	240	520	500	180	280	1720
8	230	270	220	190	250	1160
9	210	290	240	180	280	1220
10	240	400	490	240	320	1665
11	230	400	400	200	290	1460
12	210	350	450	180	320	1690
13	240	500	490	180	240	1530
14	230	380	490	220	320	1610
15	210	350	400	210	280	1520
16	240	350	400	240	360	1670
17	230	400	300	190	300	1380
18	210	360	300	200	280	1290
19	240	280	400	250	320	1600
20	230	360	450	250	320	1770
21	210	480	400	260	300	1520
22	240	350	260	230	320	1400
23	230	310	240	200	330	1180
24	210	200	450	240	320	1590
25	240	350	350	280	400	1700

26	230	280	240	290	360	1430
27	210	380	450	260	360	1720
28	240	370	500	260	340	1730
29	230	340	280	260	350	1480
30	210	340	270	260	340	1440

TABLE: A-3

DAILY AVERAGE TEMPERATURE RECORDED AT HYDROLOGY DEPARTMENT, INDIAN INSTITUTE OF TECHNOLOGY, ROORKEE FROM 1ST NOV. 2003 TO 30TH NOV. 2003.

DATE	AVERAGE TEMPERATURE IN DEGREE CENT.
1	26.5
2	25.0
3	24.5
4	26.5
5	26.7
6	26.0
7	26.5
8	26.0
9	26.0
10	24.0
11	24.5
12	24.0
13	21.8
14	24.5
15	23.5
16	25.5
17	25.0
18	26.0
19	24.5
20	23.5
21	22.5
22	23.5
23	24.5
24	25.0
25	24.2

26	23.2
27	26.5
28	22.0
29	22.8
30	23.0

TABLE: A-4

DAILY RELATIVE HUMDITY RECORDED AT HYDROLOGY DEPARTMENT, INDIAN INSTITUTE OF TECHNOLOGY, ROORKEE FROM 1ST NOV. 2003 TO 30TH NOV. 2003.

DATE	RELATIVE HUMIDITY IN PERCENTAGE
1	78
2	78
3	64
4	71
5	76
6	68
7	65
8	71
9	75
10	75
11	75
12	48
13	56
14	66
15	68
16	60
17	61
18	66
19	72
20	68
21	66
22	67
23	67
24	73
25	73

26	73
27	65
28	70
29	68
30	74

APPENDIX: B

Table for output membership function and predicted value

No. of Days	Actual Load in (KW)	Output Membership Function					Predicted value
1	599.66	1.0000	0.8000	0.8000	1.0000	0.8000	470
2	485.90	1.0000	0.8000	1.0000	0.8000	0.8000	470
3	759.15	1.0000	1.0000	1.0000	1.0000	1.0000	750
4	851.60	0.8000	1.0000	1.0000	1.0000	1.0000	820
5	985.9	1.0000	1.0000	1.0000	0.8000	0.8000	915
6	815.4	1.0000	1.0000	1.0000	1.0000	1.0000	750
7	1021.9	1.0000	1.0000	1.0000	0.8000	0.8000	915
8	460.8	1.0000	0.8000	1.0000	0.8000	0.8000	470
9	521.000	1.0000	0.8000	1.0000	0.8000	0.8000	470
10	748.270	1.0000	1.0000	1.0000	1.0000	1.0000	750
11	806.500	1.0000	1.0000	1.0000	1.0000	1.0000	750
12	958.700	0.8000	1.0000	1.0000	1.0000	1.0000	820
13	825.340	0.8000	1.0000	1.0000	1.0000	1.0000	820
14	735.500	1.0000	1.0000	1.0000	1.0000	1.0000	750
15	756.550	1.0000	1.0000	1.0000	1.0000	1.0000	750
16	878.000	0.8000	1.0000	1.0000	1.0000	1.0000	820
17	722.250	1.0000	1.0000	1.0000	1.0000	1.0000	750
18	471.300	1.0000	0.8000	1.0000	0.8000	0.8000	470
19	812.350	1.0000	1.0000	1.0000	1.0000	1.0000	750
20	870.300	0.8000	1.0000	1.0000	1.0000	1.0000	820
21	713.050	1.0000	1.0000	1.0000	1.0000	1.0000	750
22	585.900	1.0000	0.8000	0.8000	1.0000	0.8000	470
23	435.700	1.0000	0.8000	1.0000	0.8000	0.8000	470
24	738.950	1.0000	1.0000	1.0000	1.0000	1.0000	750
25	828.750	0.8000	1.0000	1.0000	1.0000	1.0000	820
26	645.800	1.0000	0.8000	0.8000	1.0000	0.8000	470

27	893.440	0.8000	1.0000	1.0000	1.0000	1.0000	820
28	962.700	1.0000	1.0000	1.0000	0.8000	0.8000	915
29	620.850	1.0000	0.8000	0.8000	1.0000	0.8000	470
30	534.400	1.0000	0.8000	1.0000	0.8000	0.8000	470