International Journal of Electrical, Electronic and Communication Sciences
ISSN: 2517-9438
Vol:2, No:10, 2008

# Clustering based formulation for Short Term Load Forecasting

Ajay Shekhar Pandey, D. Singh, and S. K. Sinha

***Abstract*****—**A clustering based technique has been developed and implemented for Short Term Load Forecasting, in this article. Formulation has been done using Mean Absolute Percentage Error (MAPE) as an objective function. Data Matrix and cluster size are optimization variables. Model designed, uses two temperature variables. This is compared with six input Radial Basis Function Neural Network (RBFNN) and Fuzzy Inference Neural Network (FINN) for the data of the same system, for same time period. The fuzzy inference system has the network structure and the training procedure of a neural network which initially creates a rule base from existing historical load data. It is observed that the proposed clustering based model is giving better forecasting accuracy as compared to the other two methods. Test results also indicate that the RBFNN can forecast future loads with accuracy comparable to that of proposed method, where as the training time required in the case of FINN is much less.

***Keywords*****—**Load forecasting, clustering, fuzzy inference.

## I. Introduction

FOR the purpose of optimal planning and operation of large scale system, modern system theory and optimization techniques are being with the expectation of considerable cost savings. In attaining this, the knowledge of future power system, load is one of the main requirements. The accuracy of the forecasted load influences decision making in unit commitment, hydrothermal coordination, fuel allocation and off line network analysis. Since the electrical load is a function of weather variables and human social activities, the variety of short term load is very complex, sometimes it changes evenly, sometimes linearly and sometimes randomly. Therefore forecasting process has become even more complex. Owing to the importance of load forecasting, a wide variety of statistical approaches have been proposed in the last two decades such as Multiple Linear Regression, Time Series etc. These statistical approaches are not able to cope up with the variable operating conditions such as abrupt change in environmental or weather variables. These techniques have some drawbacks such as inaccurate prediction, numerical instability, and difficulty in modeling process, requirement of large historical data base and demand of high human expertise. The impacts of globalization and deregulation demands improved quality at competitive prices that is the reason why development of advanced tools and methods for planning, analysis, operation and control are needed. Artificial Neural Network (ANN) has been successfully applied for Short Term Load Forecasting (STLF). The ability of ANN to outperform conventional statistical methods especially during rapidly changing weather conditions has made these methods, an attractive alternative. High accuracy of the load forecasting for power systems improves the security of the power system and reduces the generation costs. In addition the accurate load forecasts are key data which are necessary for the electric power price forecast on the electric power markets. Researchers have concentrated on STLF using ANN, just mainly because of its capability of approximating any non-linear function and model determination through the learning process.

A wide variety of conventional models for STLF have also been reported in the literature. They are based on various statistical methods such as time series [1], exponential smoothing [2] and regression based [3]. In recent years use of intelligent techniques have increased noticeably. ANN and fuzzy systems are two powerful tools that can be used in prediction and modeling. Load forecasting techniques such as ANN [4]-[10], Expert systems [11], fuzzy logic, fuzzy expert system and fuzzy inference [12]-[17] have been developed, showing more accurate and acceptable results as compared to conventional methods. Conventional ANN model have several drawbacks, such as long training time and slow convergent speed. The RBF model is a very simple and yet intrinsically powerfully network, which is widely used in many fields because of its extensive learning and highly computing speed [8]-[10]. A neuro-fuzzy approach has been applied successively in a price sensitive environment [13]. Fuzzy Set Theory was specifically designed to mathematically represent uncertainty and vagueness and provide formalized tools for dealing with the imprecision intrinsic to many problems. A fuzzy expert system for STLF is developed in [15]. It uses fuzzy set theory to model imprecision in the load temperature model and temperature forecasts as well as operator's heuristic rules. Fuzzy set theory provides a general way to deal with uncertainty, and express the subjective knowledge about a process in the form of linguistic IF-THEN rules.

Vagueness involved in training and the selection of network parameters such as number of neurons, learning rate, and

Ajay Shekhar Pandey is with the Department of Electrical Engineering, Kamla Nehru Institute of Technology, Sultanpur, UP, India (corresponding author, phone: +91-9415091697; fax: +91-5362-240454; e-mail: shekhar.ajay04@rediffmail.com).

D. Singh is with the Department of Electrical Engineering, Institute of Technology, Banaras Hindu University, Varanasi, UP, India (e-mail: dev_singh_vns@yahoo.com).

S. K. Sinha is with the Department of Electrical Engineering, Kamla Nehru Institute of Technology, Sultanpur, UP, India (e-mail: sinhask98@engineer.com).

momentum parameters are the major obstacles in designing a neural network forecaster. In this paper, a clustering based formulation of short term load forecasting is proposed. This does not require training and uses less number of independent variables as compared to neural network forecasters. Forecaster development is formulated as an optimization problem. The continuous monitoring and trial with different parameter settings are not required, as the method does not involve any training. According to an objective function, the data partitioning approach is used to classify the data. The data are made to self-organize in to clusters by finding similarities among the variables. This simply requires computations of averaging and Euclidean distance. The proposed model corresponds to the situation where the inputs are user defined. Load is modeled as two components i.e. the base load and the temperature sensitive part. Only the temperature sensitive part of the load is modeled using the temperature variables only.

## II. Clustering Based Formulation and Implementation

For cluster analysis, the handling of interval and ratio scaled variable depends only on equal differences in different values that have same significance, and thus that the arithmetic mean and Euclidean distances are applicable. Clustering techniques involve a process of measurement, either of the magnitude of their distances between two objects or of the magnitude of the similarity to each other.

The Euclidean distance is a measure of distance between two vectors in an n-dimensional space and is given by

$$\|X - Y\| = \sum_{i=1}^{r} (X_i - Y_i)^2 \quad (1)$$

where X and Y are the n-dimensional feature vectors

The accuracy of the forecast depends upon the closeness of the match. Closest match can be found for the future load patterns from the available information in the form of historical data. The feature variables for the objects are determined through feature selection procedure. Data base is a collection of objects and their feature vectors are made from the historical data. Cluster analysis comprises interdependent process for classifying objects described by set of values of several variables.

Consider a data matrix,

$$X = \{x_{ik}\} \text{ (i=1,….,m,k=1,……,l)} \quad (2)$$

representing m objects. The m rows thus give for each object the values of l variables denoting the characteristics of these objects. A mapping d: U×U→R is called a distance function if for arbitrary x, y€U satisfies the following:

$$d(x, y) \geq d_o \quad (3)$$

$$d(x, x) = d_o \quad (4)$$

$$d(x, y) = d(y, x); \quad (5)$$

a metric distance function in addition holds

$$d(x, y) = d_o, \quad \text{then} \quad x = y, \quad (6)$$

$$d(x, z) \leq d(x, y) + d(y, z), \quad \text{then} \quad z \in U, \quad (7)$$

The distance function used in this work is Euclidean distance. If the objects in (1) be serially numbered from 1 to m, then a cluster C is defined as a non empty set of indices $C \subset \{1,2,....,m\}$. It contains $m_p$ objects $x_i$ with $i \in C$. Centroid of a cluster is a reference object in respect to which the cluster is formed to minimize a certain objective function.

For minimizing the objective function among the data itself, clustering techniques is used for partitioning the data matrix. Here the data has been partitioned in such a way using clustering that the Mean Absolute Percentage Error (MAPE) is minimized as an objective function. Problem of clustering can be addressed as the optimal partitioning of the data matrix such that the MAPE is minimum for the given data matrix. It can be formulated as

$$\text{Min. MAPE } (g, m_p) = \frac{1}{N}\sum_{i=1}^{N}(F_i - T_i) \quad \text{for MAPE,}$$

$$\text{Min. } Z = C(m_p) = \sum_{x_j \in C}\|x_{jk} - x_{ik}\| \quad (k = 1,2,....,l) \quad (8)$$

for a cluster C with centroid $x_{ik}$,

where $F_i = \frac{1}{m_p}\sum_{j \in C} O_j$

g= number of objects to be introduced between two nearest objects,

$m_p =$ Cluster size

Partitioning is performed for the historical data and the value g and $m_p$ is found for the cross-validation data. The future loads are then predicted for the whole year assuming that the model is optimal for future data. The advantage associated with this formulation is that, the actual accuracy index MAPE is minimized instead of other functions such as RMSE (Root Mean Square Error) or SSE (Sum Squared Error). The other advantage is that only two unknowns, g and $m_p$ are to be determined.

The data, from an electric utility, Canada consists of integrated hourly system load and relevant temperature. The hourly temperature of two places has been taken. To forecast the load of a week ahead, data of three weeks prior to the week, to be forecasted is used as historical data. A full week ahead of historical data is used for cross-validation set. The model designed for forecasting the load is based on the temperature sensitivity of the load. If the relationship between the load and temperature can be mapped, then the load at a particular hour can be given by adding the variation to the base load. So, in designing this model the temperature of the two places has been taken as variables and the load variation from the base load is taken as objects.

Out of many different combinations of the load time series, temperature and the values of g and $m_p$ for minimum error on the cross-validation data is tabulated in the Table I for week days. Features selected for the model are

$L(i-1), L(i-2)$ and $T1(i)$ and $T2(i)$ for week days. The optimization variables are $m_p = 4$ and g=1 for week days.

TABLE I
FEATURE AND PARAMETER SELECTION FOR WEEK DAYS

| S.No. | Feature | Min. Error | g (0-3) | $m_p$ (1-4) |
|---|---|---|---|---|
| 1. | $L_{k-1}$ | 1.5213 | 2 | 3 |
| 2. | $L_{k-1}, L_{k-2}$ | 1.2089 | 2 | 1 |
| 3. | $L_{k-1}, L_{k-2}, L_{k-3}$ | 1.3089 | 2 | 3 |
| 4. | $L_{k-1}, L_{k-2}, L_{k-3}, L_{k-4}$ | 1.5049 | 1 | 3 |
| 5. | $L_{k-1}, L_{k-2}, L_{k-4}$ | 1.4792 | 3 | 3 |
| 6. | $L_{k-1}, L_{k-2}, T_k^1$ | 1.2821 | 1 | 4 |
| 7. | $L_{k-1}, L_{k-2}, T_k^1, T_k^2$ | 1.0323 | 1 | 4 |
| 8. | $L_{k-1}, L_{k-2}, T_k^1, T_k^2, T_{k-1}^1, T_{k-1}^2$ | 1.1667 | 1 | 3 |

## III. Methods Used for Comparison

### *A. Radial Basis Function Neural Network*

RBF Network consists of two layers, a hidden layer with nonlinear neurons and an output layer with linear neurons. In RBF neural network, three parameters are needed to study: the center and the variance of the basis function and the weight connecting hidden layer to the output layer. A RBF neural network embodies both the features of an unsupervised learning based classification and a supervised learning layer. The hidden unit consists of a function called the radial basis function, which is similar to the Gaussian Density function whose output is given by

$$o_i = \exp\left(-\left(\sum_{j=1}^{r} \frac{(x_{jp} - W_{ij})^2}{\sigma}\right)\right)$$

where

$W_{ij}$ = Center of the $i^{th}$ RBF unit for input variable j

$\sigma$ = Spread of the RBF unit

$x$ = $j^{th}$ variable of the input pattern

Number of neuron determines the number of cluster centers that are stored in the network. Neurons with large spread will generalize more. Therefore, spread is an important parameter and depends on the nature of input pattern space. The values of the different parameters of the RBF networks are determined during training. These parameters are spread, cluster centers, and weights and biases of the linear layer. The number of neurons for the network and spread is determined through experimentation with a large number of combinations of spread and number of neuron. The week days structure uses 6 inputs. The model uses 5 radial basis neurons and a single neuron for weekdays.

### *B. Fuzzy Inference Neural Network (FINN)*

The inference mechanism is the process which numerically evaluates the information embedded in the fuzzy rule base to get the final result. The fuzzy rule base consists of "IF-THEN" type rules. Fuzzy inference refers to a fuzzy IF-THEN structure. A fuzzy inference engine can process mixed data. Input data received from the external world is analyzed for its validity before it is propagated into a fuzzy inference engine. The capability of processing mixed data is based on the membership function concept by which all the input data are eventually transformed into the same unit before the inference computations. To deal with the linguistic values such as high, low, and medium, architecture of ANN that can handle fuzzy input vectors is propounded. RBFNN has been integrated with fuzzy inference to form a FINN for Short Term Load Forecasting. In FINN the RBFNN plays an important role to classify input data into some clusters while the fuzzy inference engine handles the extraction of rules. The fact that the initial parameters of the FINN are not randomly chosen as in neural networks but are assigned reasonable values with physical meaning gives the training of an FNN a drastic speed advantage over neural networks.

*Input Variable Selection and Data Processing*

Parameters with effect on hourly load can be categorized into day type, historical load data and weather information. Temperature is the most effective weather information on hourly load. The input load is sorted into 5 categories and labeled as low (L), low medium (LM), medium (M), medium high (MH) and high (H). The input temperature is also sorted into 5 categories same as above. The historical data is used to design data which are further fuzzified using IF-THEN rule. Design data consists of hourly data, integrated load data and temperature of two places. The data model involves the range of data low (L), low medium (LM), medium (M), medium high (MH) and high (H), five linguistic variables for each crisp data type. These five linguistic values are defined as L (3800-4280MW), LM (4280.001-4760MW), M (4760.001-5240MW), MH (5240.001-5720MW) and H (5720.001-6200MW) using IF-THEN rule. These data are normalized and fuzzified in form of FAM table using the five linguistic variables using IF-THEN rule.

Fuzzy logic systems do not have learning and self-adaptability. Therefore, physical characteristic of the whole system must be known, and then, a series of rules must be defined, which are expressed in the if-then form. The FINN is trained using last four weeks hourly load data and then they are used to forecast the load for the next 168 hours i.e. one week.

## IV. Simulation Results

Table II, III & IV presents the forecast error in Mean Absolute Percentage Error for a week each in winter, spring and summer seasons respectively. This reflects the behaviour of the network during seasonal changes. The index used for testing the performance of forecasters is the MAPE. The designed network is used to forecast the week ahead forecast

TABLE II
COMPARATIVE PERFORMANCE (WINTER, JAN. 25-31)

| Day | Clustering | RBFNN | FINN |
|---|---|---|---|
| Monday | 1.1194 | 1.0776 | 2.5711 |
| Tuesday | 1.3289 | 1.0727 | 1.5041 |
| Wednesday | 1.6555 | 1.1105 | 2.0527 |
| Thursday | 0.8902 | 0.7494 | 2.6438 |
| Friday | 1.3887 | 1.1171 | 1.9225 |
| Saturday | 1.0929 | 1.6459 | 2.3185 |
| Sunday | 0.8067 | 1.5838 | 2.8998 |
| Average | 1.1832 | 1.1939 | 2.2732 |

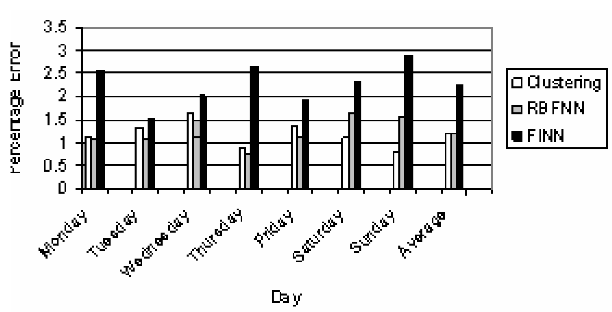



Fig. 1 Comparative chart showing accuracy of different approaches for winter

TABLE III
COMPARATIVE PERFORMANCE (SPRING, MAY 17-23)

| Day | Clustering | RBFNN | FINN |
|---|---|---|---|
| Monday | 1.2167 | 1.0856 | 1.9990 |
| Tuesday | 0.6759 | 0.7082 | 1.8797 |
| Wednesday | 0.5802 | 0.9606 | 1.9750 |
| Thursday | 2.3343 | 2.2876 | 2.0208 |
| Friday | 1.4699 | 1.1114 | 1.8356 |
| Saturday | 1.1042 | 0.7726 | 2.3826 |
| Sunday | 1.3413 | 1.7412 | 2.6110 |
| Average | 1.2461 | 1.2310 | 2.1005 |

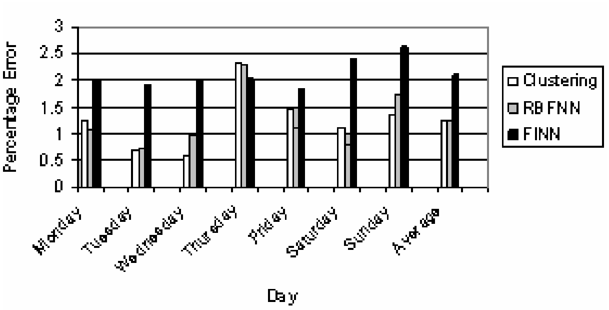



Fig. 2 Comparative chart showing accuracy of different approaches for spring

TABLE IV
COMPARATIVE PERFORMANCE (SUMMER, JULY 19-25)

| Day | Clustering | RBFNN | FINN |
|---|---|---|---|
| Monday | 0.9938 | 1.2466 | 2.2050 |
| Tuesday | 0.9387 | 2.2017 | 1.9221 |
| Wednesday | 1.2389 | 0.8057 | 1.9505 |
| Thursday | 0.6789 | 1.2365 | 1.5206 |
| Friday | 0.8277 | 0.9062 | 1.5079 |
| Saturday | 0.8950 | 1.0312 | 1.9915 |
| Sunday | 1.2590 | 1.1475 | 1.5122 |
| Average | 0.9788 | 1.2246 | 1.8014 |

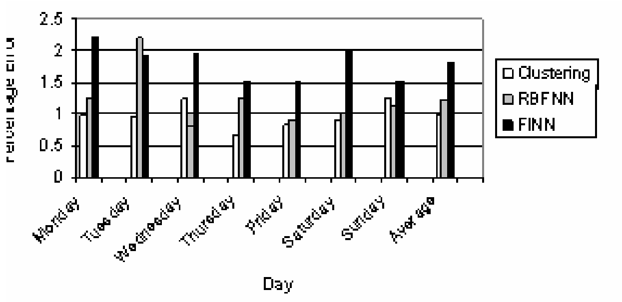



Fig. 3 Comparative chart showing accuracy of different approaches for summer

on an hourly basis. Forecasting has been done on the one year load data of Trans Alta Electric Utility for Alberta, Canada. Load varies from 3900 MW to 6200 MW.

It is observed (Fig. 1, 2&3) that the proposed clustering based formulation for STLF works well in all the seasons irrespective of the variation in temperature. However RBFNN is equally good for all the seasons. For having a comparative study, the proposed clustering method is compared with other two methods, RBFNN (Radial Basis Function Neural Networks) and FINN (Fuzzy Inference Neural Network). Comparison has been done for the same set of data and for the same period of time.

In winters there is wide variation in temperatures and therefore in the loads also. It is observed that the forecaster captures the load shape quite accurately and the forecasting errors on most of the week days are low with slightly higher error on weekend days. For winter season the accuracy of the proposed clustering based model and RBFNN is alike but better than FINN.

The hourly load profile for the summers is quite smooth. That is the reason; all the three methods of forecasting in this season have performed well. With clustering based formulation, MAPE is the minimum.

Load profile during spring season is smoother than that in winter. In spring RBFNN is giving slightly better results than the clustering based formulation for the same period. At the same time, training time required for FINN is very less as compared to RBFNN.

## V. Conclusion

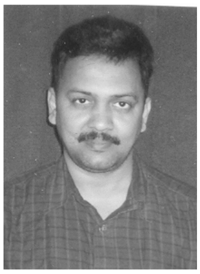
A clustering based formulation has been developed using MAPE as an objective function. Load forecasting method, proposed above is feasible and effective. Comparative study shows that the proposed clustering based approach for short term load forecasting is better and giving more accurate results than the other two methods, FINN and RBFNN for the same period of time and same set of data. The error depends on many factors such as homogeneity in data, network parameters, choice of model and the type of solution. This approach required fewer amount of data and also does not requires training. Forecasting result shows that RBFNN is equally good for week ahead forecasting and it can also forecast future loads with accuracy comparable to that of proposed method, where as the training time required in the case of FINN is much less. The proposed method is computationally simpler than a neural network and also aids in selecting the best set of independent variables due to its speed and also giving better accuracy than the other two methods.

## Acknowledgment

The authors would like to thank TransAlta, Alberta, Canada for providing the load data used in the studies.

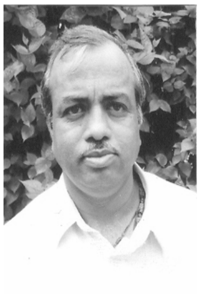
## References

[1] M.T. Hagen and S.M.Behr, "The time series approach to Short Term Load Forecasting", *IEEE Trans. On Power Systems*., PRWS-2(3), pp. 785-791, 1990.
[2] W.R.Christiaanse, "Short Term Load Forecasting using general exponential smoothing", *IEEE Trans. On Power Appar. Syst*. PAS-90 pp.900-910, 1971.
[3] A.D.Papalexopoulos, T.Hasterberg, "A Regression based Approach to Short Term System Load Forecast", *IEEE Trans. On Power Systems*. Vol.5, No.4, , pp 1535-1544, Nov. 1990.
[4] I. Mogram and S. Rahman , "Analysis and evaluation of five short term load forecast techniques", *IEEE Trans. On Power Systems.* Vol.4, No.4, pp 1484-1491, 1989.
[5] T. S. Dillon, S. Sestito, and S. Leung, "Short term load forecasting using an adaptive neural network," *Elect. Power Energy Syst.*, vol. 13, Aug. 1991.
[6] D.C.Park M.A.,El-Sharkawi, R.J.Marks, L.E.Atlas and M.J.Damborg, " Electric Load Forecasting using an Artificial Neural Networks", *IEEE Trans. on Power Systems,* vol.6,No.2, , pp. 442-449, May 1991.
[7] T.M.Peng N.F.Hubele and G.Karady, " Advancement in the application of Neural Networks for short term load forecasting", *IEEE Trans. on Power Systems,* vol.7,No.1, pp. 250-257, Feb.1992.
[8] D.K.Ranaweera,N.F.Hubele and A.D.Papalexopoulos, " Application of Radial Basis Function Neural Network Model for Short Term Load Forecasting", *IEE Proc. Gener. Trans. Distrib.*, vol. 142,No.1, Jan.1995.
[9] S.P. Singh and O.P. Malik, "Single ANN architecture for short term load forecasting for all seasons", *Int. Jour of Engineering Intelligent Systems,* vol. 3, no. 4 249-254,Dec.1995.
[10] D. Singh and S.P. Singh, "Self selecting neural network for short-term load forecasting", *Jour. Of Electric Power Component and Systems,* vol. 29, pp. 117-130,2001
[11] K.L.Ho, Y.Y.Hsu, C.F.Chen, T.E.Lee, C.C.Liang, T.S.Lai and K.K.Chen , "Short Term Load Forecasting of Taiwan Power System using a Knowledge Based Expert System", *IEEE Trans.on Power Systems,* vol.5, pp. 1214-1221, 1990.
[12] A. G. Bakirtzis, J. B. Theocharis, S. J. Kiartzis, and K. J. Satsios, "Short-term load forecasting using fuzzy neural networks," *IEEE Trans. Power Syst.*, vol. 10, pp. 1518–1524, Aug. 1995.
[13] A. Khotanzad, E. Zhou and H.Elragal, "A Neuro-Fuzzy approach to Short-term load forecasting in a price sensitive environment," *IEEE Trans. Power Syst.*, vol. 17 no. 4, pp. 1273–1282, Nov. 2002.
[14] Hiroyuki Mori and Hidenori Kobayashi," Optimal fuzzy inference for short term load forecasting ", I*EEE Trans. on Power Systems,* vol.11, No.2, pp. 390-396. Feb. 1996.
[15] K.H. Kim, J.K. Park, K.J. Hwang, and S.H. Kim, "Implementation of Hybrid Short-term Load Forecasting System Using Artificial Neural Networks and Fuzzy Expert Systems," *IEEE Trans. on Power Systems*, vol. 10, no. 3, pp. 1534–1539, Aug. 1995.
[16] Ranaweera D.K., Hubele N.F. and Karady G.G., "Fuzzy logic for short-term load forecasting", *Electrical Power and Energy Systems*," Vol. 18, No. 4, pp. 215-222, 1996.
[17] Kwang-Ho Kim, Hyoung-Sun Youn, Yong-Cheol Kang, "Short-tem Load Forecasting for Special Days in anomalous Load Conditions Using Neural Network and Fuzzy Inference Method", *IEEE Trans. On Power Systems*, Vol. 15, pp. 559-569, 2000.



**Ajay Shekhar Pandey** received his B.Tech. and M.Tech. in Electrical Engineering from K.N.I.T. Sultanpur, India in 1987 and 2001 respectively. He is working as an Asstt. Professor with the Department of Electrical Engineering, Kamla Nehru Institute of Technology, Sultanpur, U.P, India and also pursuing his Ph.D. degree. His areas of interest are Neural Network and its applications in Power System.

**D. Singh** received his B.Tech. in Electrical Engineering from R.E.C. Surat, India in 1993, the M.E. in Electrical Engineering from MNREC Allahabad, India in 1999 and the Ph.D. from Institute of Technology, BHU, Varanasi, India in 2001, Currently he is working as Reader with the Department of Electrical Engineering, Institute of Technology,BHU, Varanasi, India. His field of interest includes state estimation and AI applications in Power System.



**S. K. Sinha** received his B.Sc.Engg in Electrical from R.I.T.Jamshedpur, Bihar, India in 1984, the M.Tech. degree in Electrical Engineering from Institute of Technology,BHU, Varanasi, India in 1987, and the Ph.D. from University of Roorkee, Roorkee, India in 1997. Currently he is working as an Asstt.Professor with the Department of Electrical Engineering, Kamla Nehru Institute of Technology, Sultanpur, U.P, India. His field of interest includes Fuzzy Control, Robotics and AI applications.