Neural Networks for Short Term Wind Speed Prediction

K. Sreelakshmi, P. Ramakanthkumar

***Abstract*—**Predicting short term wind speed is essential in order to prevent systems in-action from the effects of strong winds. It also helps in using wind energy as an alternative source of energy, mainly for Electrical power generation. Wind speed prediction has applications in Military and civilian fields for air traffic control, rocket launch, ship navigation etc. The wind speed in near future depends on the values of other meteorological variables, such as atmospheric pressure, moisture content, humidity, rainfall etc. The values of these parameters are obtained from a nearest weather station and are used to train various forms of neural networks. The trained model of neural networks is validated using a similar set of data. The model is then used to predict the wind speed, using the same meteorological information. This paper reports an Artificial Neural Network model for short term wind speed prediction, which uses back propagation algorithm.

***Keywords*—**Short term wind speed prediction, Neural networks, Back propagation.

I. INTRODUCTION

THE energy is a vital input for the social and economic development of any nation. With increasing agricultural and industrial activities in the country, the demand for energy is also increasing. Formulation of an energy model will help in the proper allocation of widely available renewable energy sources as solar, wind, bioenergy and hydropower in meeting future energy needs. A study of the energy models helps energy planning, research and policy making.

II. LITERATURE REVIEW

Wind speed prediction is necessary as wind is an intermittent source of energy. It is an important alternate source of renewable energy to the fast depleting fossil fuels [1].

There are different types of models available for wind speed prediction. They are classified as Statistical, Intelligent systems, Time series, Fuzzy logic, neural networks. Models constructed based on meteorological, topological data and wind turbine technical information using numerical methods, are suited for long term predictions since they have difficulty in fast acquisition of data and complicated computations. They lack adaptive update of estimated parameters to strengthen the model. They are based on non-statistical approaches. They depend on the experience of a meteorologist. They are carried out by interpolating measured data from a large number of sources, spread over a large area and long terms of time, instead of fixing the geographical location and time periods.

Time series models are based on historical wind data and statistical methods. The simplest of these is the ARMA which is also called Persistence model and is usually used as the bench mark [2]. This analysis results in the description of the process through a number of equations involving large amount of information, which is a major drawback of time series models.

Fuzzy models are also used to estimate wind speed. This paper reports a fuzzy model trained using genetic algorithm based – learning scheme, applied to an electrical power production plant. They are found to be more efficient than the conventional ARMA models [3]. Regression techniques are found to be less efficient compared to Artificial Neural Network model (ANN) models [4].

Kalman filter models are found to be 10% better than the ARMA. These models are found to be superior to ARMA [5]. ANNs are best suited for non-linear systems and do not require mathematical models and adapts automatically to changes in the inputs to minimize mean square errors. They have the capability to deal with large data sets. It reports that advanced ANN techniques are to be applied to improve the accuracy of the predictions [6]. The current works use Back propagation algorithm [7], Radial basis functions [8], Wavelet techniques and Support vector machines for short term forecasting [9]. SVM models are found to take less computational times compared to ANN models.

III.NEED FOR FORECASTING WIND SPEED

Wind speed prediction from past observations has applications in many diverse fields such as Target tracking, Missile guidance, Satellite launch, Electrical power demand forecasting, etc.

The most important factor influencing wind power generation is the local wind speed. The immediate requirement is for the development of improved short range forecasting methods which improve power transmission scheduling and resource allocation and hence the reliability of the power grid. Bus load forecasting is required for planning and operation of power system and power distribution.

K. Sreelakshmi is with Dept of Telecommunication Engg., RV College of Engineering Bangalore – 560059 Karnataka, India. (Phone: 91-80-28601700-241,.e-mail: sansa68@yahoo.co.in)

P. Ramakanth Kumar is with Dept of Computer Science, RV College of Engineering Bangalore – 560059, Karnataka, India. (Phone: 91-80-28601700-254, e-mail: pramakanth_2000@yahoo.com)

Forecasting enables an adaptation between demand and generation [10].

IV. FEED FORWARD NEURAL NETWORK WITH BACKPROPOGATION

A *neural network* is a computational structure which resembles a biological neuron[12]. It can be defined as *a* "*massively parallel distributed processor made up of simple processing units, which has a natural propensity for storing experimental knowledge and making it available for use*." It resembles the human brain in two respects:

a) Knowledge is acquired by the network from its environment through a learning process.
b) Interneuron connection strengths, also known as synaptic weights, are used to store the acquired knowledge.

There are many different types of neural networks, from relatively simple to very complex, just as there are many theories on how biological neurons work.

A feed-forward neural network has *layers* of processing elements, which make independent computations on data that it receives and passes the results to another layer and finally, a subgroup of one or more processing elements determine the output from the network. Each processing element makes its computation based upon a weighted sum of its inputs. The first layer is always the *input layer* and the last layer is always the *output layer*. The layers placed between the first and the last layers are the *hidden layers*. The processing elements are seen as units that are similar to the neurons in a human brain, and hence, they are referred to as *cells* or *artificial neurons*. A neuron is an information processing unit that is fundamental to the operation of a neural network. The figure shows the block diagram of a neuron.

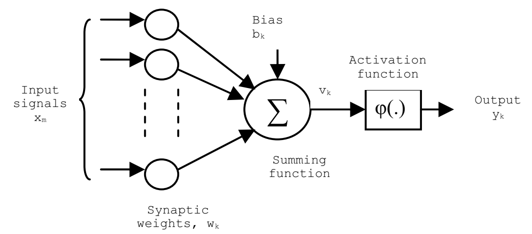

Fig. 1 Basic Structure of a Neuron

There are three basic elements in the neuron model namely; the synapses, the adder and the activation function. Synapses between neurons are referred to as *connections*, which are represented by edges of a directed graph in which the nodes are the artificial neurons. The adder is used for summing of the inputs, represented by the *sigma*. A *threshold function* or *the activation function* is sometimes used to qualify the output of a neuron in the output layer.

This neuronal model includes an externally applied bias, denoted by b_k, which has the effect of changing the net input value of the activation function. In mathematical terms, we may describe a neuron k by the following pair of equations:

$$u_k = w_{kj} x_j \tag{1}$$

$$y_k = \phi(u_k + b_k) \tag{2}$$

Here, $x_1, x_2 \ldots x_m$ are the input signals.
$w_{k1}, w_{k2} \ldots w_{km}$ are the synaptic weights
u_k is the linear combiner output due to the input signals
b_k is the bias
φ () is the activation function
y_k is the output signal of the neuron
v_k is the induced local field or activation potential

The output of any neuron is the result of thresholding, if any, of its internal activation, which, in turn, is the weighted sum of the neuron's inputs. The Activation function is denoted by φ (v), defines the output of a neuron in terms of the induced local field v. Here, sigmoid function is used as the activation function. It is defined by the equation,

$$\varphi(v) = 1/(1+\exp(-av)) \tag{3}$$

where 'a' is the slope parameter of the sigmoid function. The figure 2 shows the structure of the curve traced by a sigmoid function.

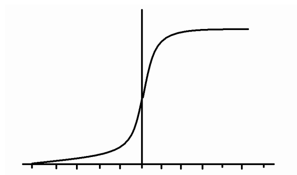

Fig. 2 Sigmoid function

V. SYSTEM MODEL

Here, raw data from six text files are read, filtered and processed to obtain normalized data. Patterns are generated and statistical analysis is performed for good correlation among the input data values. A large part of the data is fed into the training network and the remaining part into the testing network. Finally, the wind speed is predicted which is the output of the Neural Network.
Short term wind speed prediction involves the following steps:

a) Data Acquisition & Pre-processing
b) Data Conversion & Normalization
c) Statistical Analysis
d) Design of Neural Network
e) Training
f) Testing

Data Acquisition & Pre-processing

The weather report and the related parameter values are collected at a Weather Station at periodic time intervals, say every ten minutes. Our work involves the utilization of six different parameters values which are acquired from the weather station report. A historical data of 10 years is considered for the experimentation. The various parameters that are considered as input to the model are shown in the following table.

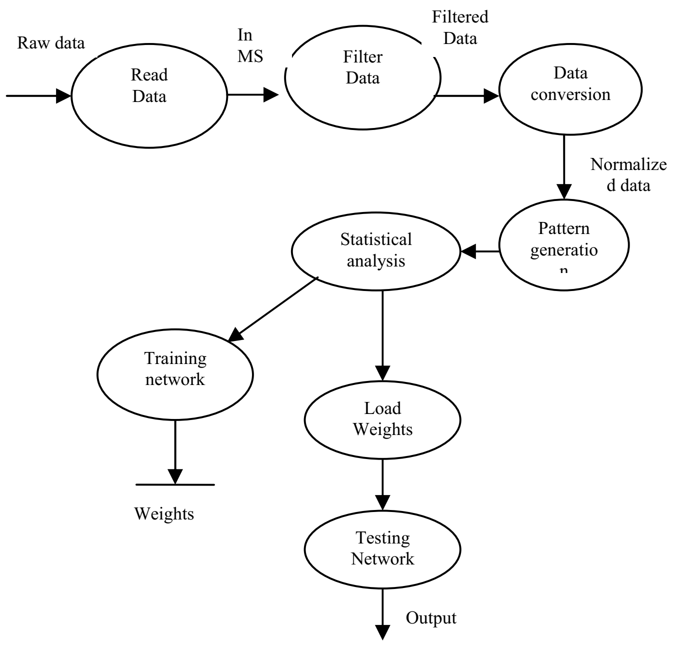

Fig. 3 System architecture

TABLE I LIST OF PARAMETERS FOR THE NETWORK

PARAMETERS	UNITS
Mean temperature	Deg.C
Humidity	%RH
Wind gust	m/s
Wind direction	Deg.M
Barometric pressure	Mb
Wind speed	m/s

ba.txt - Notepad

DATE	TIME	VALUE
2004-02-08	15:50	992.4
2004-02-08	16:00	992.3
2004-02-08	16:10	992.3
2004-02-08	16:20	992.2
2004-02-08	16:30	992.2
2004-02-08	16:40	992.1
2004-02-08	16:50	991.9
2004-02-08	17:00	992
2004-02-08	17:10	992
2004-02-08	17:20	992
2004-02-08	17:30	992
2004-02-08	17:40	992
2004-02-08	17:50	992
2004-02-08	18:00	992.1
2004-02-08	18:10	992.1

Fig. 4 Sample input file

Data Conversion & Normalization

The data collected after preprocessing then goes through the stage of data conversion and Normalization. The maximum value in each of the parameters is found out and each value in that column is divided by the maximum value so that the value lies between zero and one.

OP.TXT - Notepad

DATE	TIME	MTEMP	HUM	WGUST	WDIR	BARO	WSPEED
2004-02-08	15:50	21.4500	53.5040	6.3000	238.1007	992.4000	4.7000
2004-02-08	16:00	21.1800	53.7088	5.7000	223.0474	992.3000	3.9000
2004-02-08	16:10	20.9100	53.1968	6.0500	168.1029	992.3000	5.0000
2004-02-08	16:20	20.6300	53.8112	6.6500	169.7964	992.2000	5.2000
2004-02-08	16:30	20.3900	52.0704	8.0000	160.5763	992.2000	6.0000
2004-02-08	16:40	20.2400	53.3504	8.0500	176.1940	992.1000	6.1000
2004-02-08	16:50	20.0100	54.6304	9.4500	182.9680	991.9000	5.9000
2004-02-08	17:00	19.7200	56.2688	6.6500	251.6487	992.0000	5.3000
2004-02-08	17:10	19.3900	57.6512	7.2000	272.9114	992.0000	4.4000
2004-02-08	17:20	19.0500	59.6992	5.6500	239.9824	992.0000	4.1000
2004-02-08	17:30	18.6800	60.9792	5.8000	236.0309	992.0000	4.6000
2004-02-08	17:40	18.2800	62.2080	6.5500	205.5480	992.0000	4.8000
2004-02-08	17:50	17.8900	63.1808	5.5500	221.7302	992.0000	4.8000
2004-02-08	18:00	17.5100	63.6416	6.0500	217.2143	992.1000	4.5000
2004-02-08	18:10	17.1200	64.0000	6.6000	198.3976	992.1000	4.7000
2004-02-08	18:20	16.7500	64.4096	6.3500	169.6082	992.1000	4.9000
2004-02-08	18:30	16.3900	64.6656	6.9000	190.6828	992.0000	5.1000
2004-02-08	18:40	16.0700	65.2800	6.1500	194.8225	992.1000	4.6000
2004-02-08	18:50	15.7900	66.7648	5.1000	160.3881	992.1000	3.4000
2004-02-08	19:00	15.5000	67.5328	4.5500	183.7207	992.3000	2.9000
2004-02-08	19:10	15.2100	66.2528	3.7500	192.1882	992.3000	3.0000
2004-02-08	19:20	14.9200	64.5632	4.0500	187.8603	992.3000	2.9000
2004-02-08	19:30	14.7200	64.0512	3.7000	163.5869	992.4000	2.5000

Fig. 5 Sample output file

NORMAL.TXT - Notepad

DATE	TIME	MTEMP	HUM	WGUST	WDIR	BARO	WSPEED
2004-02-08	15:50	0.4729	0.5019	0.1356	0.6831	0.8278	0.2070
2004-02-08	16:00	0.4669	0.5038	0.1227	0.6399	0.8277	0.1718
2004-02-08	16:10	0.4610	0.4990	0.1302	0.4823	0.8277	0.2203
2004-02-08	16:20	0.4548	0.5048	0.1432	0.4871	0.8276	0.2291
2004-02-08	16:30	0.4495	0.4885	0.1722	0.4607	0.8276	0.2643
2004-02-08	16:40	0.4462	0.5005	0.1733	0.5055	0.8275	0.2687
2004-02-08	16:50	0.4411	0.5125	0.2034	0.5249	0.8273	0.2599
2004-02-08	17:00	0.4347	0.5279	0.1432	0.7220	0.8274	0.2335
2004-02-08	17:10	0.4275	0.5408	0.1550	0.7830	0.8274	0.1938
2004-02-08	17:20	0.4200	0.5600	0.1216	0.6885	0.8274	0.1806
2004-02-08	17:30	0.4118	0.5720	0.1249	0.6772	0.8274	0.2026
2004-02-08	17:40	0.4030	0.5836	0.1410	0.5897	0.8274	0.2115
2004-02-08	17:50	0.3944	0.5927	0.1195	0.6361	0.8274	0.2115
2004-02-08	18:00	0.3860	0.5970	0.1302	0.6232	0.8275	0.1982
2004-02-08	18:10	0.3774	0.6004	0.1421	0.5692	0.8275	0.2070
2004-02-08	18:20	0.3693	0.6042	0.1367	0.4866	0.8275	0.2159
2004-02-08	18:30	0.3613	0.6066	0.1485	0.5471	0.8274	0.2247
2004-02-08	18:40	0.3543	0.6124	0.1324	0.5589	0.8275	0.2026
2004-02-08	18:50	0.3481	0.6263	0.1098	0.4602	0.8275	0.1498
2004-02-08	19:00	0.3417	0.6335	0.0980	0.5271	0.8277	0.1278
2004-02-08	19:10	0.3353	0.6215	0.0807	0.5514	0.8277	0.1322
2004-02-08	19:20	0.3289	0.6057	0.0872	0.5390	0.8277	0.1278
2004-02-08	19:30	0.3245	0.6009	0.0797	0.4693	0.8278	0.1101

Fig. 6 Normalized data output file

The normalized values obtained form the pattern. This pattern is called as the *time window*. Each of these patterns is fed into the neural network in the form of input. The code for data conversion involves the following steps:

a. During the stage of data acquisition and preprocessing, the maximum value among each of the parameters is computed.
b. Normalization is carried out for all the parameters.
c. The normalized values are then written into a file using loop line count previously calculated.

Statistical Analysis

Statistical analysis is carried out to find the amount of dependency between each of the meteorological values and to get rid of the redundant values that might be present in the data set. "*Spearman rank correlation*" and an optional "*Pearson Correlation*" are applied. The purpose of correlation is to measure and interpret the strength of a linear or nonlinear relationship between two continuous variables. When conducting correlation, we use the term *association* to mean "linear association". Here, we focus on the Pearson and Spearman 'ρ' correlation coefficients. Both correlation coefficients take on values between -1 and +1, ranging from being negatively correlated (-1) to uncorrelated (0) to positively correlated (+1). The sign of the correlation coefficient (i.e., positive or negative) defines the direction of the relationship. The absolute value indicates the strength of the correlation.

Spearman's correlation

It is a technique used to test the direction and strength of the relationship between two variables. In other words, it is a device to show whether any one set of numbers has an effect on another set of numbers. It uses the statistic Rs which falls between -1 and +1.

Procedure to carry out Spearman's correlation:

a. State the null hypothesis i.e. "There is no relationship between the two sets of data."
b. Rank both sets of data from the highest to the lowest. Make sure to check for tied ranks.
c. Subtract the two sets of ranks to get the difference d.
d. Square the values of d.
e. Add the squared values of d to get $\sum d^2$.
f. Use the formula Rs = $1-(6.\sum d^2/n^3-n)$ where n is the number of ranks you have. If the Rs value between the 2 sets of data
g. If the Rs value is 0, state that null hypothesis is accepted. Otherwise, say it is rejected

The following table shows the results of correlation. It is observed that the data is symmetrical with respect to the main diagonal.

TABLE II RESULTS OF CORRELATION

Correlation	Mean temp	Humidity	Wind gust	Wind dir	Baro pres	Wind speed
Mean temp	1					
Humidity	-0.68531	1				
Wind gust	0.210195	-0.33041	1			
Wind dir	0.005796	-0.00528	-0.02107	1		
Baro pres	-0.35301	0.031238	0.030917	-0.0965	1	
Wind speed	0.193882	-0.31846	0.960779	-0.02305	0.053841	1

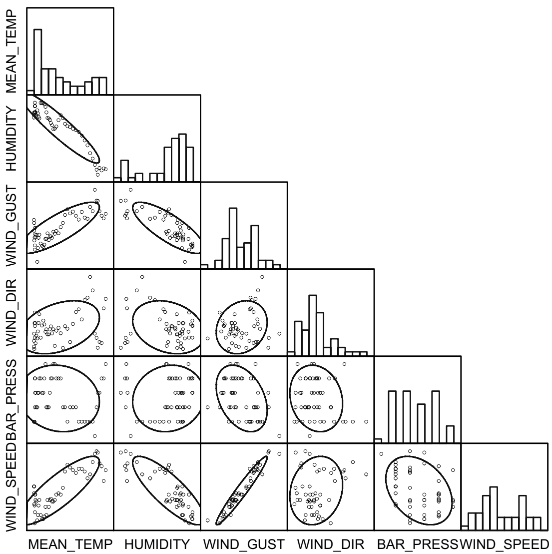

Fig. 7 Graphical representation of the correlation matrix

Design of neural network

The design of the neural network involves designing the three fields of neurons: one for input neurons, one for hidden processing elements, and one for the output neurons. The connections are for the feed forward activity. There are connections from every neuron in field 1 to every one in field 2, and in turn, from every neuron in field 2 to every neuron in field 3. Thus, there are two sets of weights, those figuring in the activations of hidden layer neurons, and those that help determine the output neuron activations. Using back propagation algorithm, in each training set, the weights are modified so as to reduce the mean squared error [MSE] between the network's prediction and the actual target value. These modifications are made in the reverse direction, from the output layer, through each hidden layer down to the first hidden layer, till the terminating condition is reached.

The steps in the algorithm are:

- Initialize the weights
- Propagate the inputs forward
- Back propagate the error
- Terminating condition

Testing

The remaining input values are utilized for testing and validation wherein the wind speed is predicted for the test input and compared with the actual values. The setup giving minimum error is obtained by varying the number of hidden layers and other parameters such as learning parameter, number of hidden layers in the network, number of epochs for test, error tolerance, number of neurons in each layer, etc.

VI RESULTS

The following are the results of the above tests:

TABLE III TEST SPECIFICATIONS FOR 5 OUTPUT LAYERS AND 50 EPOCHS

Sl No. of test case	3
Name of test	Back propagation Test 3
Error Tolerance	0.01
Learning parameter	0.05
Number of Epochs	50
Number of layers	5
Number of Neurons in each layer	5 4 3 2 1

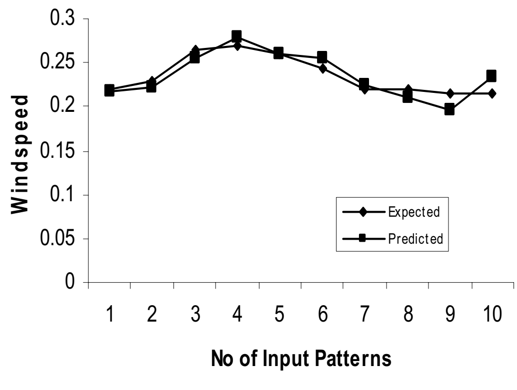

Fig. 8 Expected v/s predicted outputs for 5 layers and 50 epochs

TABLE IV TEST SPECIFICATIONS FOR 5 OUTPUT LAYERS AND 200 EPOCHS

Sl No. of test case	5
Name of test	Backpropagation Test 5
Error Tolerance	0.002
Learning parameter	0.05
Number of Epochs	200
Number of layers	5
Number of Neurons in each layer	5 4 3 2 1

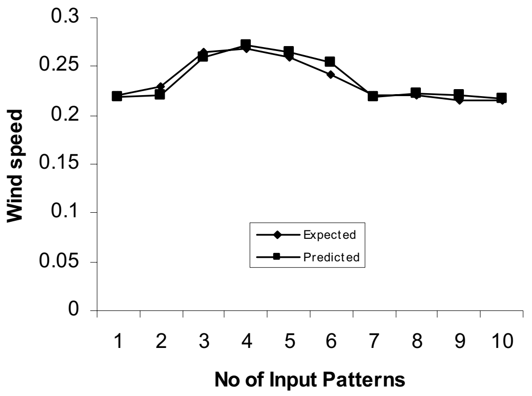

Fig. 9 Expected v/s predicted outputs for 5 layers and 200 epochs

VII. CONCLUSIONS

The literature available for wind speed modeling reveals that majority of the models is being utilized for electrical power demand forecasting. Though many short term models are presented, the accuracy of the models still need to be improved. In this model, the predicted wind speed differs from the actual value by max 5%.

REFERENCES

[1] X.Wang, G. Sideratos, N. Hatziargyriou and L. H. Tsoukalas, " Wind speed forecasting for power system operational planning" , 8th Intl. Conf. on Probabilistic Methods Applied to Power Systems, Iowa State University, Iowa, September 12-16, 2004, pp 470 - 474.

[2] Anthanasios Sfetsos and Costas Siriopoulos, "Time series forecasting of averaged data with efficient use of information", IEEE Trans on Systems, Man and Cybernatics- Part A: Systems and Humans", Digital object identifier- 10.1109/TMSCA2005.851133.

[3] Ioannis.G.Damousis, Mians C. Alexadis, John B Theocharis and Petros,S.Dokopoulos ,"A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation", IEEE Trans. on Energy Conversions, Vol. 19, No.2, PP 352 – 361, June 2004.

[4] Shuhui Li, Donald C Wunsch, Edgar O'Hair and Michael G Giesselmann, " Comparative analysis of regression and artificial neural network models for wind turbine power curve estimation", Journal of Solar Energy Engineering, Vol. 123, pp 327 – 332, November 2001

[5] Shuhui Li, " Wind power prediction using recurrent multilayer perceptron neural networks", 0-7803-7989-6/03/$17.00 ©2003 IEEE, PP- 2325 – 2330

[6] Henrique Steinhertz Hippert, CarlosEduardo Pedreira and Reinaldo Castro Souza, "Neural networks for short term load forecasting: A review and evaluation", IEEET Trans. on Power Systems, Vol 16. No.1 pp 44- 55, February 2001.

[7] T.G.Barbounis and J B Theocharis, " Locally recurrent neural networks optimal filtering algorithms: application to wind speed prediction using spatial correlation", Proceedings of The Intl. Joint Conf. on Neural Networks, Canada, July 31 – August 4, 2005, pp 2711 - 2716.

[8] G. Sideratos and N. Hatziargyriou, "Using radial basis neural networks to estimate wind power production", 1-4244-1298-6/07/$25.00 © 2007 IEEE.

[9] Enrique Romero and Daniel Toppo, "Comparing support vector machines and feed forward neural networks with similar hidden layer weights", IEEE Trans. on Neural Networks, Vol 18, no 3. May 2007, pp 959 – 963.

[10] T.G. Barbounis, J B Theocharis, Minas. C. Alexadis and Petros. S. Dokopoulos, " Long term wind speed and power forecasting using local recurrent neural network models", IEEE Trans. on Energy Conversion, Vol 21, no. 1, March 2006, pp 273 – 284.

[11] Alireza Khotanzad, Reza Afkhami Rohani and Dominic Maratukulam, "ANNSTLF – artificial neural network short-term load forecaster – generation three." IEEE Trans. on Power Systems, Vol 13, No.4, November 1998, pp 1413-1422.

K. Sreelakshmi, Assistant Professor, Dept of Telecommunication Engg., RV College of engineering Bangalore – 560059.
She obtained her BE and ME degrees from Bangalore university, with specializations in Electronics and Communication and is currently pursuing her PhD degree. Her research interests are in the field of Neural network based models and their applications, Electromagnetics and microwaves. She is a member of IETE and ISTE. She has worked in Indian Institute of Science, Bangalore as a JRF. Her work at IISc included characterization of electronic devices at cryogenic temperature, which has resulted in an international paper publication in the Journal, "Cryogenics ".

Dr. Ramakanthkumar P, Professor, Dept. of Computer science and engineering, RV College of engineering Bangalore – 560059.
He completed his B.E in Computer Science and Engineering from S.J.College of engineering, Mysore, and M.S in Software Systems from BITS, Pilani. He completed his Ph.D. from Mangalore University in the area of Pattern Recognition. He has authored more than five publications in reputed journals. His area of interest is Image Processing, Pattern Recognition and Natural Language Processing. He is a member of the Computer Society of India (CSI) and member of Indian Society for Technical Education (ISTE).