International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-10, August 2019

Power Load Forecasting using Back Propagation Algorithm

J. VeerendraKumar, K.Sujatha, B. Chandrashaker Reddy, V. Karthikeyan

Abstract*:. The day today operation and scheduling activities of a power generation unit requires the forecast of power demand to meet the needs of the consumers. Load forecasting is categorized into short, medium and long term forecasts. The short term prediction refers to hourly load forecast for the period lasting from an hour to more than a few days. The medium term forecasts is prediction of power load ranging from one to several months ahead. Finally, the long term forecast indicates prediction of power load ranging from one to many years in near future. Excellence of short term hourly power load forecast has a considerable effect on cost effective functioning of power plants because numerous assessments are dependent on these forecasts. These assessments comprise cost effective scheduling of power plants, scheduling of coal acquisition, power plant security consideration, and planning for energy related business. The significance of precise load forecasts will reduce the wastage of power in the future because of remarkable changes taking place in the structure of power generation industries due to deregulation and heavy competition. This situation creates thrust on the power generation sectors to function at maximum possible efficiency, which leads to precise forecasting of the power load.

***Keywords:* Power load forecasting, back propagation algorithm, Artificial Neural Network**

I. INTRODUCTION

1.1. LOAD FORECASTING

The planning of power system is done to obtain realistic ballpark figures regarding the power consumption in future and this technique is called as power load forecasting. Forecasting is a statement of what will happen if certain conditions or trends continue and assumes that the causes of the events are under human control. There is another term called prediction which appears to be similar to forecasting with respect to definition. But prediction is a guess of something which will happen regardless of whether people want it or not and is beyond human control [1-4].

1.2. CONVENTIONAL METHODS OF LOAD FORECASTING

The different techniques are

- Regression analysis
- Time series analysis
- Correlation technique
- Delphi technique
- Extrapolation
- End-use method
- Probabilistic extrapolation correlation

1.3. WORK FLOW ORGANIZATION

The work flow is structured such that, section 1, describes the introduction to load forecasting, section 2 comprises of fundamentals of ANN with literature survey discussed in section 3. Section 4 discusses the method adopted to implement power load forecasting, data collection and data selection. Section 5 explains the result for the forecast data and finally section 6 concludes the work carried out.

II. ARTIFICIAL NEURAL NETWORK(ANN)

The conceptual model of ANN depicts a real nervous system in human body encloses neuron, interacting with one another via the axon connections. This kind of model possesses similarity to axons and dendrites of the central nervous system (CNS) of the human body. This self-organizing and adaptive nature of ANN has forced this model to offer a new parallely processed, more robust and user-friendly paradigm as compared with the other traditional approaches. ANN is a collection processing elements, assembled in a structured way to depict the characteristics of the real biological neural network. The mapping ability of ANN helps to solve the non linear real time problems, very different from the conventional approach. They mimic the characteristics of the human brain on a computer [4-6] using a structured algorithm. ANN is made up of artificial processing elements called 'neurons', which are in fact the scaled version of the real neurons of CNS. It is believed, some important features of the human brain are replicated by building a network that consists many artificial neurons. ANN architecture includes extraordinarily parallel connected adaptive processing elements interconnected in a well defined fashion [7-10].

These networks have processing elements (PEs) which are interpreted using differential equations [15, 16]. The PEs are interrelated and connected layer by layer and also intra-connected. Each PE computes the inner product of connection strength with the outputs from the PEs in the preceding layer. The inner product is modified using a non-linear function which facilitates the convergence.

Manuscript published on 30 August 2019.
*Correspondence Author(s)
J. VeerendraKumar, Research Scholar, Dept. of EEE, Dr. MGR Educational and Research Institute, Chennai, T.N, India.
Dr. K.Sujatha, Professor, Dept. of EEE, Dr. MGR Educational and Research Institute, Chennai, T.N, India.
B. Chandrashaker Reddy, Assistant Professor, Department of ECE, SoE, NNRG, Hyderabad, Telangana, India.
Dr. V. Karthikeyan, Professor, Dept. of EEE, Dr. MGR Educational and Research Institute, Chennai, T.N, India.

The function of the artificial neuron is revealed in Figure 1. Non-linearity is said to exist for the vectors which are binary, bipolar and hard limited in nature. For analog vectors, thrashing functions like unipolar sigmoid functions (0 to 1), hyperbolic tan, Gaussian, logarithmic and exponential functions are used [1].

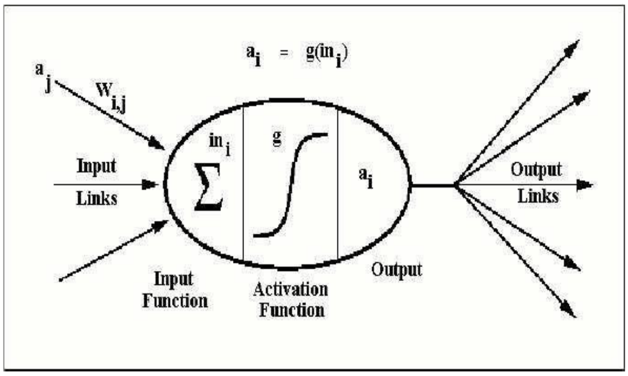

Figure 1. Function of an Artificial Neuron

A discrete network has two states for the node i.e., '0 or 1', and '–1 or 1' which is the same case with analog network in which the output is continuous. In a synchronous network, the discrete outputs of the PEs are updated for each and every instant of time and for an asynchronous network the state of only one PE is updated. If there is no connection existing from output layer to input layer then the network is a feed forward network if a closed loop exists then it is called as feedback neural network. The ANN is said to be static, if the outputs depend on the current inputs and does not depend on the past values. This kind requires no memory. On the other hand for a dynamic or recurrent neural network the output depends upon past inputs or outputs. If the interlink among the PEs vary with time, then it is an adaptive network else it is a non-adaptive network. The synaptic weight modification is done by supervised or unsupervised learning methods. For supervised learning both inputs and outputs are required; whereas the unsupervised learning the outputs are not needed and only inputs are required.

The supervised learning rules include Perceptron learning law, adaptive linear element (ADALINE), Widrow Hoff LMS rule, multi layer perceptron (MLP), hidde Markov models and temporal dynamic models. Self organizing feature Maps (SOM), Kohonen learning rule, Adaptive Resonance Theory (ART) and the PCA are categorized under unsupervised learning rule [11, 12].

III. REVIEW OF LITERATURE

In early 1940s the artificial neural networks was first reported by Mc Culloh and Pits. Whenever the excitatory inputs exceed the threshold, a neuron is fired. As long as the inhibitory inputs are present this process continues. To compute any logical function such kind of networks can be used. It is concluded by Rosenblatt that the Mc-Culloh pitts model was un-biological. To overcome the deficits, a new model called perceptron model was introduced which was able to learn and generalize. On further investigations, several mathematical models including competitive learning or self-organization, and forced learning were also used [1-3].

Adaptive Linear Neuron (ADALINE) was incorporated in supervised learning method Widrow and Hoff, introduced the ADALINE. It is a single neuron, with gradient steepest descend method for error reduction using the supervised learning rule. The linearly separable patterns are discriminated using ADALINE which is a linear neuron. The concept of multi layer ADALINEs or multi layer perceptron (MLP) model was developed for patterns which are non linearly separable. The architecture for MLP is illustrated in Figure 2. Werbos, in his Ph.D. dissertation described the training of the MLP network as back-propagation algorithm (BPA). But it was not popularized. Rumelhart and his group conducted studies on BPA published them. Lippmann briefed this concept as back-propagation algorithm which became popular for training the multi layer network [13, 14].

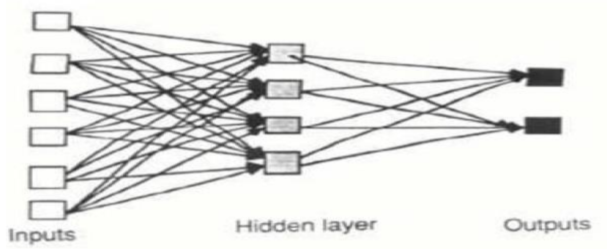

Figure 2. Multilayer perceptron model

A large volume of work has been done, with respect to the choice of number of hidden layers and number of hidden layer nodes during the training process where the input patterns are presented, with initial random weights and weight updation at different stages by optimizing the error function. The training procedure in ANN is unique and is application oriented; most of the real time applications are solved with only one hidden layer containing nodes using supervised learning rule. Many studies revealed that only one hidden layer holds good for various training strategies and this was revealed by Sietsma and Dow. But anyway, only few applications, like the one claimed by Chester has two hidden layers. Optimal number of nodes is to be used in the hidden layer. Excessive number of hidden layer nodes will result in the oscillation of the mean squared error (MSE) without any convergence; or occasionally the network may converge to one of the local minima. Similarly, very less number of nodes in the hidden layer will just help to learn the patterns, but generalization is not possible. As a result, some technique is needed to identify the exact number of nodes in the hidden layer. Heuristic and meta heuristic algorithms can be used to find the hidden layer nodes as proposed by Mario, Wang and Robert. Hirose, yamashita and Hijiya have adopted an algorithmic approach, to estimated the minimized MSE. Gaussian function in the hidden nodes simplifies the work of finalizing the value for number of nodes in hidden layer and also

Retrieval Number: J10310881019/19©BEIESP
DOI: 10.35940/ijitee.J1031.0881019
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-10, August 2019

Weymaere and Martens suggested sigmoid function to be used in the output nodes. Fujita has analyzed hidden unit function [3, 4, 5].

In this method mathematical models are not needed and it learns from the input output patterns contributing to robust and adaptive computing capabilities. By using SOM the adaptive learning method are enhanced and is immune to noise [1, 2].

IV. IMPLEMENTATION OF ANN FOR POWER LOAD FORECASTING & RESULTS

In the normal environment, the Tamilnadu Electricity Board knows the power demand. Due to rapid modernization forecast the load with the existing electrical and electronic items and population of the city it is very difficult to find the amount of power to be generated. To overcome this problem, an artificial neural network approach has been adopted. The ANN approach does not need mathematical model for modeling the process with process parameters and state variables. ANN offers mapping of inputs and outputs which is application oriented. The process parameters and state variables serve as inputs and outputs for the ANN respectively. A feature represents a process parameter or state variable. A pattern is defined as the combination of input data and output data and a number of such patterns together will form a collection of data.

The application of power load forecasting is realized using ANN as shown in the flowchart in Figure 3. The training of ANN for power load forecasting is depicted in Figure 3. The weight updation algorithm is used to train the ANN during which the network learns the patterns. A desired performance index is used to stop the training process in ANN. Finally, the weights obtained are stored used further for testing process. When ANN is used for applications like power load forecasting, the data is altered using the updated weights obtained by training the ANN. Now and then, the outputs of the network are verified for correctness. If desired values of outputs for power load forecast data is obtained, then it is accepted to be trained.

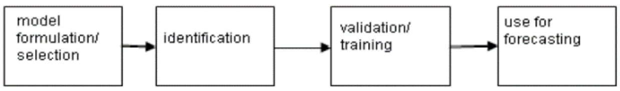

Figure 3. Block diagram for implementation of ANN for power load forecasting

4.1. COLLECTION OF DATA

The power consumption data is obtained from the Tamilnadu electricity board. The data includes data, month, day, peak hours (time1 & time2), year, production of power (megawatts), and power consumption (million units). The collected data is for the period January 1998 and November 1998. The data is given in Figure 4.

4.2. PATTERN NORMALIZATION

Normalization is done to convert all the patterns having values between ‘zero’ to ‘one’ to reduce the complexity in computation. Equation 1 represents the formula for normalization.

$$Y_i = Y_i / Y_{max} \quad (1)$$

Where Y_i is the value of that feature, and Y_{max} is the maximum value of the feature.

INPUT TO THE ANN   OUTPUT FROM THE ANN

Day	Hour	Load 1	Load 2	Load 3	Load 4
1	10.3	3747.4	3953.8	4090.4	3972.8
1	11	3754	3847	4127	3927.8
1	11.3	3553.1	3862.1	3975.7	3903.2
1	12	3523.9	3899.7	3972.7	3903.8
1	12.3	3462.7	3709.4	3901.2	3842.4
1	13	3480.1	3784.2	3958.7	3842.5
1	13.3	3660.1	3918.3	3889	3857.1
1	14	3576.5	3906.3	3904.2	3895.1
1	14.3	3443.3	3924.9	3979.2	3830.9
1	15	3445.1	3931.7	3904.4	3802.5
1	15.3	3558.3	3853	3855.6	3785.8
1	16	3557.9	3873.4	3893	3795.8
1	16.3	3568.4	3785.8	3885.4	3941.2
1	17	3608.8	3852.4	3965.6	3956.4
1	17.3	3673.5	3986.5	4131.2	3911.8
1	18	3660.1	4009.3	4069.8	3985
1	18.3	3811.7	4106.1	3940.8	4046.6
1	19	4081.5	4291.3	4146.8	4228.2
1	19.3	4201.2	4252.4	4208.1	4303.4
1	20	4132.6	4273.2	4219.9	4211.4
1	20.3	3858.4	4100.5	4159	4193.3
1	21	3769.6	3999.5	4164.8	4108.5
1	21.3	3705.8	3692.1	3798.5	3954.3
1	22	3606.6	3573.1	3732.5	3919.5
1	22.3	3346.6	3521.4	3566	3647.2
1	23	3349.6	3460.7	3544	3634

Figure 4. Sample data collected from TNEB

4.2.1 Training based on pattern selection

Training and Testing the ANN is done after categorizing the patterns for the intended purpose. The classification performance during testing is done with the remaining patterns and the correctness of the forecasted data is verified. The entire population should be represented by the patterns selected for training the network. The procedure is as follows:

- Among the given patterns, the first pattern in the month of January 1998 is taken as the starting point. For example, Thursday will be the first pattern as it is evident from Figure 4.
- Then a pattern belonging to next week is chosen and it will be Friday. This will be the second pattern. Hence, only one pattern will be chosen in a week and the successive pattern chosen will be the day following the pattern of the day of the previous week. This process is applied for the entire data set at the starting point.

4.3. ANN TRAINING STRATEGIES

Learning takes place with the patterns, using different learning rules where weights are modified. This technique is called as supervised learning and unsupervised learning methods. The problem of power load forecasting comes under supervised learning where the inputs and outputs are considered during the training of ANN. Currently, forecasting of power demand in the city of Chennai is a tedious work. The network for power load forecasting uses supervised learning rule. The pattern consisting of inputs are presented to ANN and the actual output obtained from the output layer is evaluated as the target value of the corresponding pattern.

The deviation between the actual and the target values is called as MSE with respect to the pattern subjected to training. This error value is transmitted in the reverse direction, such that the connection strengths between different layers are optimally modified. Thus the MSE, which is the objective function, is minimized for the pattern under study. This procedure is put together till the last training pattern is presented and the ANN has learnt the entire training set after many of iterations so that the MSE, which was large initially is reduced now, by presenting all the data needed to train the network. Anyway, the attainment of global minimum by the ANN is not assured and instead, may land on any one of the local minima. If MSE value increases, divergence is taking place instead of convergence. At times, the ANN is said to oscillate between convergence and divergence.

At this juncture, the end of the training is denoted and can be stopped with the help of MSE or with the help of some performance criterion for classification. Once the desired value of MSE is obtained, training the ANN is to be halted. The challenge is that, the exact value of MSE may not be known and it needs to be inferred after the training has taken placed. Over fitting may occur if the ANN is over trained to achieve very low MSE. Hence to avoid the network getting over trained, an optimal value of MSE is considered. Generality is lost due to over fitting which may lead to exact classification of patterns which are already classified during training process and does not classify the test patterns which are unknown to the network.

4.4. BACK-PROPAGATION ALGORITHM (BPA)

Newton's Steepest descent rule is used in the learning to achieve the global minimum. The flowchart illustrated in Figure 5 depicts the BPA. Training the ANN, involves the finalization of architecture size, MSE, the number of hidden layers and number of hidden layer nodes. At the start random weights are assigned for the connections between the nodes. A pattern from the training set is presented in the input layer of the network and the error at the output layer is calculated. The error is propagated backwards towards the input layer and the weights are updated. This procedure is repeated for all the training patterns. At the end of each iteration, test patterns are presented to ANN, and the classification performance of ANN is evaluated. Further training of ANN is continued till the desired classification performance is reached. The architecture for BPA is shown in Figure 6. The algorithm for BPA is as follows:

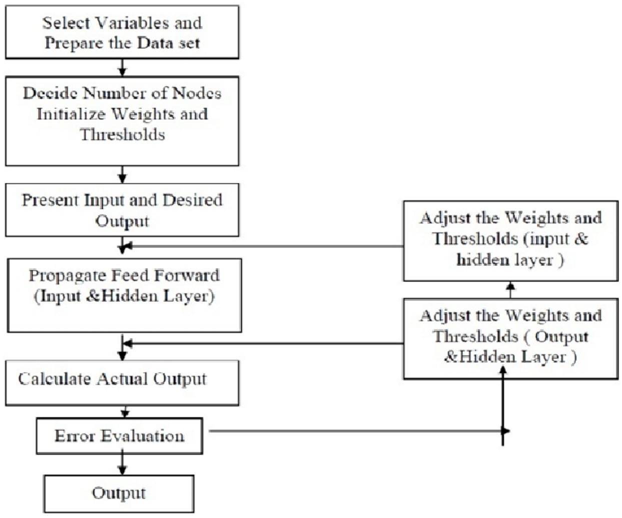

Figure 5. Flowchart for BPA

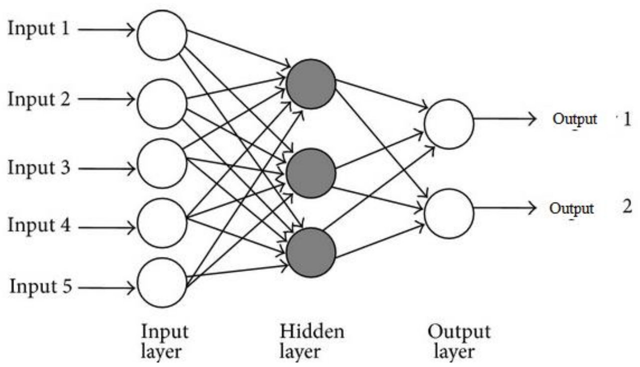

Figure 6. Architecture for BPA

4.5. ALGORITHM FOR BPA

FORWARD PROPAGATION

- Initialize the connection strength values and threshold.
- Present the input-output pattern to the ANN.
- Calculate the output of each output node in the consecutive layers.

$$OUT_{(\text{output of a node})} + 1/(1+\exp(-x))$$

- The error of a pattern is calculated

$$Er(p) = (1/2) \sum(\text{desired}(p) - \text{actual}(p))^2$$

REVERSE PROPAGATION

- Calculate in the output layer the error for the corresponding nodes.

$$\mathcal{E}_{(\text{output layer})} = \text{actual } (1-\text{actual})(\text{desired} - \text{actual})$$

- Modify the connection strength in between the output and hidden layer.

$$W(n+1) = W(n) + \eta\, \mathcal{E}_{(\text{output layer})}\, O_{h(\text{hidden layer})}$$

- Calculate the error for the nodes in the hidden layer

$$\mathcal{E}_{(\text{hidden layer})} = \text{actual } (1-\text{actual}) \sum \mathcal{E}_{(\text{output layer})}$$

W_{ho}(updated weights for hidden and output layer)

Retrieval Number: J10310881019/19©BEIESP
DOI: 10.35940/ijitee.J1031.0881019
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

- Modify the connection strength in between the hidden and input layer.

$$W(n+1) = W(n) + \eta\, \mathcal{E}_{(\text{hidden layer})}\, O_{i(\text{input layer})}$$

The above steps complete one weight updation

- Present the next pattern and repeat the above steps to modify the weights for second time.
- One epoch is over if all the training patterns are processed.
- Calculate the errors for all the training patterns and display MSE.

$$E\ (MSE) = \sum E(p)$$

V. RESULTS AND DISCUSSION

5.1 SUMMARY OF PRESENT WORK

Power load forecasting by ANN has been considered in spite of the existing conventional technique. The main reason to use ANN for power load forecasting is its model free nature. Power consumption data was collected for all the 24 hours. The number of patterns collected is 336. The patterns include the features like day, hour and power consumption.

5.2. TRAINING STRATEGIES USED

The network was trained with supervised algorithm (back propagation algorithm). The algorithm uses inputs and target outputs. From the Figure 7, it is clear that only 5 inputs and one output have been used for training the ANN. The inputs are day, time and previous three-week power consumption.

To have an understanding of the forecasting power of ANN, the total number of training patterns taken into consideration 42. The left over patterns are treated as patterns for testing phase. The network was trained with five number of PEs in the input layer and a single PE in the output layer. The hidden nodes were altered from three to ten nodes. Figure 8 shows the variation of nodes in the hidden layer for η =1.0 and α = 0.8. It can be observed that the error change is more when the configuration of the network is 5-5-l.

The Figure 8 shows the influence of η on convergence for given α =0.8. It can be observed from the figure that when η reduces from the value 1.0 to 0.5, it takes more iterations for the network to reach the given MSE=0.01.

The Figure 8 shows the influence of α for given value of η = l.0. It can be observed from the figure that when α reduces from 0.8, it takes more iterations for the network to reach the given MSE=0.01.

The Figure 9 gives a comparison for actual power consumption and the forecast power consumption for 50 test data. During testing only the inputs were present to the ANN. The propagated input with trained weights and output from output layer is obtained. This value is multiplied a constant 4800 to get the original value. It can be noted that during training the loads were divided by 4800 and hence during testing, the output of the ANN should be multiplied by 4800. From this figure, if 20 % variation is allowed then the forecast power consumption is correct. Hence a result of 93% correct forecasting of power consumption is achieved.

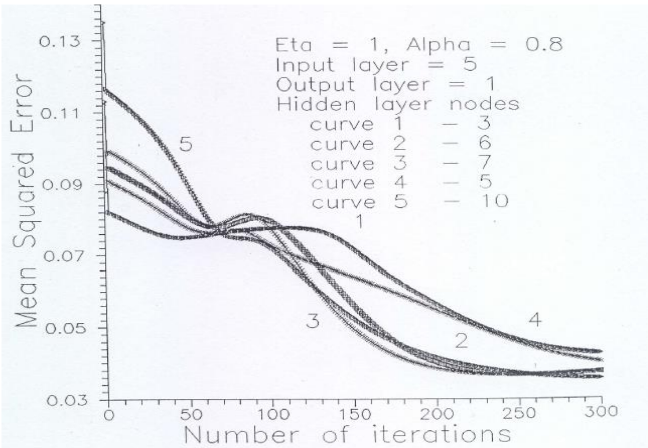

Figure 7. Influence of hidden nodes on error convergence

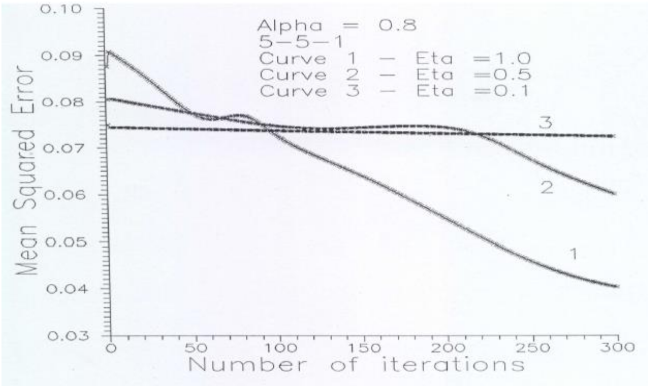

Figure 8. Influence of η values on error convergence

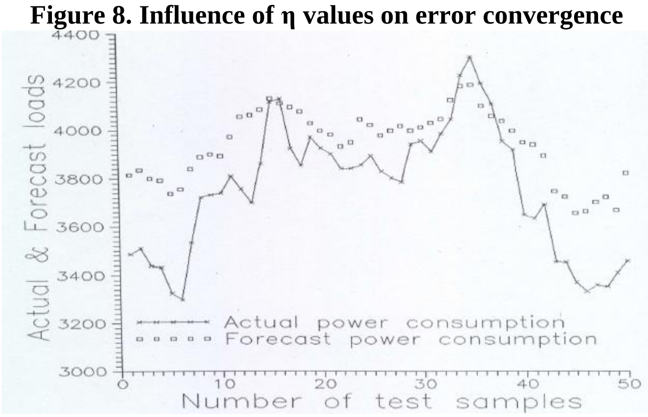

Figure 9. Comparison of actual power consumed and Forecast power consumption

VI. CONCLUSION

The Power load forecasting by ANN has been considered inspite of the existing conventional techniques because of its robustness. Moreover the generalization capacity is the main reason to use ANN for power load forecasting. Power consumption data was collected for all the 24 hours. Nearly 168 patterns are collected. The patterns include the features like day, hour and power consumption. The network was trained with supervised algorithm (Back propagation algorithm). The forecasting efficiency was found to be 93% during testing of the Feed forward architecture.

REFERENCES

1. Purushothaman S and Srinivasa Y.G., " A procedure for training an artificial neural network with the application of tool wear monitoring ", International Journal of Prod. Res., pp 635-657, 1998.
2. D. C. Park, M. A. El-Sharkawi, R. J. Marks II, L. E. Atlas, M. J. Damborg. "Electric load forecasting using an artificial neural network, "IEEE Transactions on Power Systems, Vol. 6, no. 2, pp. 442-449, May 1991.
3. D. Srinivasan, S. S. Tan, C. S. Cheng, Eng Kiat Chan. "Parallel neural network-fuzzy expert system strategy for short-term load forecasting: system implementation and performance evaluation", IEEE Transactions on Power Systems, vol.14, no.3, pp.1100-1106, Aug 1999.
4. Singh A, Kalra P.K, Emmanuel, "Point of view of knowledge based systems for load forecasting", Symposium on expert systems application to power systems, 1998, pp 7~1- 7~4.
5. Taylor W. J, McSharry E.P, de Menezes M.L, "A comparison of univariate methods used for forecasting electricity demand up to a year ahead", International Journal of Forecasting,2006, Vol.22, pp 1-16.
6. Feinberg E.A, Genethliou D, "Chapter 12 Load Forecasting", Applied mathematics for power systems, pp 269-282.
7. Galkin I, "Crash introduction to artificial neural networks" found on http://ulcal.uml.edu/~iag/CS/intro-to-ANN.html
8. Galiana F.D, Gross G, "long term load forecasting", Proceedings of the IEEE, 1987, Vol. 75 No.12, pp 1558-1573.
9. Singh A, Chen H, Canizares A.C, "ANN-based short term load forecasting in electricity markets", Proceedings of the IEEE power engineering society transmission and distribution conference, 2001, pp2:411-415.
10. Komane T, "Short term peak load forecasting using neural networks", Department of Electrical Engineering, University of Cape Town, 2001.
11. EM 1110-2-1701,"Appendix B load forecasting methods", 1985, pp B-1-B-5 found On http://www.usace.army.mil/publications/eng-manuals/em1110-21701/a-b.pdf.
12. http://www.cee.hw.ac.uk/~alison/ai3notes/subsection 2 5 2 1.htm, Expert system Architecture.
13. Anderson D, McNeil G, "Artificial neural networks technology", A DACS State-of the Art Report, Data & Analysis Center for Software, 1990, pp 1-35.
14. stergiou C, Siganos D, "Neural Networks" found on http://www.doc.ic.ac.uk/~nd/surprise96/journals/vol4/cs11/report.html.
15. Khothanzad A, Hwang R.C, Abaye A, Maratukulam D, "An adaptive modular artificial neural network yearly load forecaster and its implementation at electric utilities", IEEE transactions on power systems, Vol. 10, NO. 3, 1995, pp1716-1722.
16. Netlab Toolbox, help files, © Ian T Nabney (1996-2001), www.ncrg.aston.uk/netlab/index.php

Retrieval Number: J10310881019/19©BEIESP
DOI: 10.35940/ijitee.J1031.0881019
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)