Use of Artificial Neural Networks for Short-Term Electricity Load Forecasting of Kenya National Grid Power System

Christopher A. Moturi
School of Computing and Informatics
University of Nairobi
Nairobi, Kenya

Francis K. Kioko
School of Computing and Informatics
University of Nairobi
Nairobi, Kenya

ABSTRACT

This paper developed a supervised Artificial Neural Network-based model for Short-Term Electricity Load Forecasting, and evaluated the performance of the model by applying the actual load data of the Kenya National Grid power system to predict the load of one day in advance. Raw data was collected, cleaned and loaded onto the model. The model was trained under the WEKA environment and predicted the total load for Kenya National Grid power system. The test results showed that the hour-by-hour approach is more suitable and efficient for a day-ahead load forecasting. Forecast results demonstrated that the model performed remarkably well with increased number of iterations. The result suggests that incremental training approach of a neural network model should be implemented for online testing application to acquire a universal final view on its applicability.

General Terms

Load Forecasting.

Keywords

Electricity Load Forecasting, Short Term Load Forecasting, Artificial Neural Networks, ANN.

1. INTRODUCTION

Load forecasting is the prediction of future loads of a power system. Accuracy in electricity load forecasting is important in power system planning and operation and will help market participants minimize operational cost and develop a more reliable energy supply mechanism (Quaiyum et al 2011). The main problem of the planning is the demand knowledge in the future. Basic operating functions such as hydrothermal unit commitment, economic dispatch, fuel scheduling and unit maintenance can be performed efficiently with an accurate forecast. Load forecasting is also important for contract evaluations and evaluations of various financial products on energy pricing offered by a deregulated market.

Power system engineers use different methods to optimally plan, monitor, and operate different aspects of today's sophisticated power systems. Some of these methods are economic dispatch, unit commitment, state estimation, automatic generation control, security analysis, optimal power flow, and load forecast. Load forecasting can be categorized long term, medium term, or short term. Long term load forecast, one to ten years ahead, is applied in expansion planning, inter-tie tariff setting, and long-term capital investment return. Medium term load forecast, covers period of few weeks, is mainly used for scheduling fuel supply. Short term load forecast results, ahead hourly and daily, is needed in

unit commitment, maintenance, and economic dispatch problems (Feinberg and Genethliou, 2005).

There is a wide variety of both statistical and artificial intelligence techniques that have been developed for short-term load forecasting (Feinberg and Genethliou, 2005). Similar-day approach searches historical data for a day with similar characteristics as the forecast day. Regression methods model the relationship of load demand and other factors such as weather, day type, and customer category. Time series explores correlations, trends, and seasonal variations within the internal structure of the data. They include ARMA (autoregressive moving average), ARIMA (autoregressive integrated moving average), ARMAX (autoregressive moving average with exogenous variables), and ARIMAX (autoregressive integrated moving average with exogenous variables). Neural networks are used for non-linear curve fitting. Expert systems make use of rules, often heuristic in nature, to do accurate forecasting. They incorporate rules and procedures used by human experts in the field of interest into software that is then able to automatically make forecasts without human assistance. Fuzzy logic use Boolean logic while support vector machines perform a nonlinear mapping of the data into a high dimensional space (Feinberg and Genethliou, 2005). The accuracy of load forecasting depends largely on the load forecasting techniques used.

We sought to develop a supervised Artificial Neural Network-based model for short term load forecasting and evaluate the performance of the model by applying the actual load data of the national grid power system to predict the load of one day in advance. Assessments on the bases of the average forecast error was done and model performance evaluated. The model was trained by using additive regression algorithm to adjust biases and connection weights between layers.

The Kenya National Grid was used as test case. In 2005 the Kenya energy power sector was deregulated but the players in the power industry lacked the necessary accurate data to make informed investment decisions. The national power transmission and distribution company had repository of historical data which did not tell much about the future load trends. The motivation behind this research is the need for accurate short-term load forecasting for the power industry in Kenya in order to assist planners make strategic decisions on unit commitments, purchase of energy, optimal reserve capacity, system security, and reliability.

2. LITERATURE REVIEW

2.1 Load Forecasting

According to (Feinberg and Genethliou, 2005) the important factors to consider for short-term load foresting include time factors (time of year, day of week, and hour of day), weather data, and possible customer categories. For medium- and long-term forecasts, the factors to consider include the historical load and weather data, the number of customers in different categories, the number and types of appliances connected and their characteristics such age, the economic and demographic data and their forecasts.

Electricity load forecasting is a challenging problem whose modeling may depend to a great extent on the geographical region where the forecasting is needed. Basically there are three main ways in which load forecasting can be done: *Expert Judgments*: The method relies heavily on the experience and knowledge in similar forecast environments and its success depends on historically maintained databases on load forecasts and their accuracy; The main advantage is that it can adapt to exceptional circumstances like unforeseeable random events; While the main disadvantage is that it can only be as good as the person doing the forecast in terms of expertise and objectivity (Mishra, 2008a). *Linear Models:* Models of the unknown parameters are estimated from the data using linear functions (Mishra, 2008a); and can be can be implemented using linear regression or time series. *Nonlinear Models:* Nonlinear regression fits a curve through the data by finding the values of those parameters that generate the curve that comes closest to the data; Examples of nonlinear methods include Artificial Neural Networks, Nonlinear Regression, Fuzzy Logic and Bayesian Network.

2.2 Load Forecasting Methods

(Alfares and Nazeeruddin, 2002) have reviewed various electricity load forecasting techniques including multiple regression, exponential smoothing, iterative reweighted least-squares, adaptive load forecasting, time series, genetic algorithms, fuzzy logic, expert systems, and neural networks. They observed a trend particularly towards neural network based load forecasting. Islam (2011) observes that modern load forecasting techniques, such as expert systems, ANN, and fuzzy logic are particularly attractive as they have the ability to handle the nonlinear relationships between load and the factors affecting it directly from historical data. ANN are used for nonlinear short term load forecasting owing to their powerful nonlinear mapping capabilities (Mishra and Patra 2008b). (Quaiyum et al 2011) studied short term hourly load forecasting using different types of Artificial Neural Networks and showed that the Particle Swarm Optimized Elman Recurrent Neural Network (PSOERNN) is efficient. Other methods have been used to forecast short-term electricity forecasts such statistical methodology (Fan and Hyndman, 2012).

2.3 Artificial Neural Networks

An Artificial Neural Network (ANN) is a system based on the operation of biological neural networks. The model is founded upon the functionality of a biological neuron. Software implementation of a neural network can be made with their advantages and disadvantages. The advantages include: A neural network can perform tasks that a linear program cannot; when an element of the neural network fails, it can continue without any problem by their parallel nature; a neural network learns and does not need to be reprogrammed; can be implemented in any application; and can be implemented without any problem. The disadvantages include: neural network needs training to operate; architecture of a neural network is different from the architecture of microprocessors therefore needs to be emulated; requires high processing time for large neural networks. ANN is an adaptive, most often nonlinear system that learns to perform a function from data. ANN is built with a systematic step-by-step procedure to optimize a performance criterion or to follow some implicit internal constraint, which is commonly referred to as the learning rule. The input/output training data are fundamental in neural network technology, because they convey the necessary information to "discover" the optimal operating point. The nonlinear nature of the neural network processing elements provides the system with lots of flexibility to achieve practically any desired input/output map.

The key factors that make ANN highly suitable for short-term load forecasting include: it does not require any pre-assumed functional relationship between electric load and other variables such as weather conditions; provides a nonlinear mapping between weather variables and previous load patterns; electric load without the need for predetermined model; usually fault tolerant and robust.

2.4 Training of Artificial Neural Networks

ANN has to be configured such that the application of a set of inputs produces the desired set of outputs. Various methods to set the strengths of the connections exist. One way is to set the weights explicitly, using a priori knowledge. Another way is to 'train' the neural network by feeding it teaching patterns and letting it change its weights according to some learning rule (Desouky and Elkateb, 2000). The learning situations can be categorized in two distinct sorts: *Supervised Learning* or Associative Learning in which the network is trained by providing it with input and matching output patterns; *Unsupervised Learning* or Self-Organization in which an output unit is trained to respond to clusters of pattern within the input and discover statistically salient features of the input population. *Reinforcement Learning* may be considered as an intermediate form of the above two types of learning. Here the learning machine does some action on the environment and gets a feedback response from the environment. The learning system grades its action good (rewarding) or bad (punishable) based on the environmental response and accordingly adjusts its parameters until an equilibrium state occurs. In this research we used a supervised ANN model using the actual load data of the national grid power system.

2.5 Review of Applications of ANN in Load Forecasting

2.5.1 Application in Nigerian

Adepoju et al 2007 used a multi-layer perceptron (MLP) networks with back propagation to adjust the weights of the neural networks. The neural network was trained using data from the Nigerian Power Company and results were obtained by testing the trained network on new data for 24 hours of a day over a one-week period. They obtained a mean error of 2.54% and were able predict hourly load. Their results confirmed that ANNs are appropriate and efficient in short-term load forecasting.

(Buhari and Adamu, 2012) developed a model using the Levenberg-Marquardt optimization technique as a back propagation algorithm for the Multilayer Feed Forward ANN model using MATLAB. The results obtained showed a high degree of accuracy in forecasting future load demands for the

daily operational planning of a power system distribution sub-stations.

2.5.2 Application in an Engineering College in India

(Sheikh and Unde, 2012) carried out short-term load forecasting for an Engineering College in Ahmednagar, India, using ANN implemented on MATLAB-10. They showed that their Multi-Layer Perceptron (MLP) network model has a good performance and reasonable prediction accuracy by obtaining a mean absolute percentage error (between the actual and predicted values) of 1.066%. Their results suggest that ANN model with the developed structure can perform good prediction with least error.

2.5.3 Application in a Telecommunication Station in Italy

Caciotta et al, 2009 used Hybrid Neural Network system for power electric load forecasting of a telecommunication station for a daily forecast of both main load and its oscillation. Using a three layers multilayer perceptron (MLP) for the mean daily forecast, and MLP and self organizing map (SOM) for the oscillation forecasting, they obtained an accuracy of 2% for a whole year the combined components.

2.5.4 Application in Metropolitan France

(Mordjaoui et al, 2010) presented a short term electric load forecasting model using an adaptive neuro-fuzzy inference system (ANFIS) applied on historical data for metropolitan France for the year 2009. Satisfactory results were obtained for the prediction of the weekly load curve and showed that the accuracy of model is not affected by rapid fluctuations in power demand, which is the main drawback of ANN models. The forecasting performance obtained showed that it is possible to build a high accuracy model with less historical data using a combination of neural network and fuzzy logic which can be used in real time.

2.5.5 Application in Iran

Hayati (2007) used three architectures of ANN Multi-Layer Perceptron (MLP), Elman Recurrent Neural Network (ERNN) and Radial Basis Function Network (RBFN) to model STLF systems for the West of Iran and the results showed that RBFN is much faster and more reliable for STLF systems.

3. RESEARCH METHODOLOGY

3.1 Basic Assumptions and Limitations

This study was carried out under the following assumptions:

Short term load was forecast using limited historical data and the ANN model had the ability to map the non-linear characteristic of the load given a number of inputs

The study was based on simulations and experimentations; no mathematical justifications was provided for the same

i) The study had considered only one input factor for forecasting i.e. the previous two days load data. However, other factors such as weather data and temperature could also be used as input parameters

ii) Planned and unplanned load shedding was assumed to have negligible effects on the grid loading

iii) 'Acts of God' like floods, lightening etc. were assumed to be absent

iv) The effect of weather changes caused minimal changes to the national grid loading.

Training and testing of the model was conducted offline.

3.2 Data Collection

In this study quantitative data as (summarised in Table 1) was collected as follows.

i. Measurements – the actual total load for the entire Grid with the help of SCADA (supervisory control and data acquisition) information system.

ii. Observing and recording well-defined load trend (e.g. holiday break periods, special events, anomalous days etc.).

iii. Obtaining relevant data from network operators.

Table 1: Summary of Data Collected

Data Type	Data Source	Period	Unit	Duration
Generating Plant Load	Generating Plant	30 minutes	Mw	6 months
Total Load	National Control Centre	30 minutes	Mw	6 months
Primary Substation Load	Primary substation	30 minutes	Mw	6 months

The data collected was stored in Excel worksheet and was converted to WEKA readable format, the format that is acceptable prior to presenting it the model for training or any forecasting attempt. WEKA is a tool for doing numerical computations using existing functions and user defined programs to analyse data and visualize the output. The WEKA format used was the ARFF (Attribute-Relation File Format). ARFF files have two distinct sections: the Header and Data. The Header contained the name of the relation, a list of columns and their types. The Data section contained the data declaration line and the actual instance lines. The data in ARFF file format was uploaded into WEKA and training of the same started. Sample data conversion from Excel to ARFF is shown in tables 2 and 3 below.

Table 2: Sample Raw Data-Excel Format

Time	Load (Mw)	Time	Load (Mw)	Time	Load (Mw)
0:30:00	659	8:30:00	829	16:30:00	833
1:00:00	586	9:00:00	862	17:00:00	833
1:30:00	586	9:30:00	862	17:30:00	833
2:00:00	571	10:00:00	862	18:00:00	833
2:30:00	571	10:30:00	862	18:30:00	862
3:00:00	571	11:00:00	862	19:00:00	1017
3:30:00	571	11:30:00	845	19:30:00	1032
4:00:00	571	12:00:00	840	20:00:00	1032
4:30:00	586	12:30:00	846	20:30:00	1032
5:00:00	610	13:00:00	833	21:00:00	1012
5:30:00	663	13:30:00	823	21:30:00	961
6:00:00	748	14:00:00	813	22:00:00	896
6:30:00	824	14:30:00	823	22:30:00	821
7:00:00	840	15:00:00	833	23:00:00	751
7:30:00	824	15:30:00	833	23:30:00	682
8:00:00	824	16:00:00	833	0:00:00	640

Table 3: Sample Raw Data-ARFF Format
@relation Octoload
@attribute time date HH:mm:ss
@attribute load numeric
@data
Time,Load(Mw),Time,Load(Mw), Time,Load(Mw),,,,,,
0:30:00,659,8:30:00,829,16:30:00,833,,,,,,
1:00:00,586,9:00:00,862,17:00:00,833,,,,,,
1:30:00,586,9:30:00,862,17:30:00,833,,,,,,
2:00:00,571,10:00:00,862,18:00:00,833,,,,,,
2:30:00,571,10:30:00,862,18:30:00,862,,,,,
3:00:00,571,11:00:00,862,19:00:00,1017,,,,,,
3:30:00,571,11:30:00,845,19:30:00,1032,,,,,,
4:00:00,571,12:00:00,840,20:00:00,1032,,,,,,
4:30:00,586,12:30:00,846,20:30:00,1032,,,,,
5:00:00,610,13:00:00,833,21:00:00,1012,,,,,,
5:30:00,663,13:30:00,823,21:30:00,961,,,,,,
6:00:00,748,14:00:00,813,22:00:00,896,,,,,
6:30:00,824,14:30:00,823,22:30:00,821,,,,,
7:00:00,840,15:00:00,833,23:00:00,751,,,,,
7:30:00,824, 15:30:00,833, 23:30:00,682,,,,,
8:00:00,824,16:00:00,833, 0:00:00,640,,,,,

3.3 Model Training

The model was trained by applying all the inputs to the network before the weights and biases are updated (batch training mode). The model was then trained using the incremental training mode (i.e. the weights and biases of the network are iteratively adjusted every time an input is applied to the network). The load forecasted by the WEKA was compared to the actual load data and the error was calculated using the following statistical indices: *Mean Absolute Error –* to measure how close the forecasts were to the actual load; *Correlation Coefficient – to give* the quality of a least squares fitting to the original data; *Root Mean Squared Error –* to quantify the difference between the forecast and the actual load.

3.4 Network Architectural Design for the Prediction Model

The network parameters for the model are given in the Table 4 below. The model was first trained using this arbitrary designed structure. The network comprised six hidden layer neurons and one output neuron. Both momentum factor and learning rate were held constant throughout the training at 0.2 and 0.3 respectively.

Table 4: Network Topology for the Load Model

Filter	Classifier Model	No. of Iterations	Learning Rate (Shrinkage)	Momentum	Hidden Layers	Base Classifier
None	Additive regression	10	1.0			Trees. Decision Stump
	Multilayer Percetron	20	0.3	0.2	6	

Additive Regression was the meta classifier that enhanced the performance of a regression base classifier. Each iteration fitted a model to the residuals left by the classifier on the previous iteration. Prediction was accomplished by adding the predictions of each classifier. Smoothing was accomplished through varying the shrinkage. Broadly, Additive Regression refers to any way of generating predictions by summing up contributions obtained from other models. This class of algorithms starts with an empty ensemble and incorporates new members sequentially. At each stage the model that maximizes the predictive performance of the ensemble as a whole is added, without altering those already in the ensemble. Optimizing the ensemble's performance implies that the next model should focus on those training instances

on which the ensemble performs poorly. The differences between predicted and observed values are called residuals.

4. RESULTS AND DISCUSSION
4.1 Simulation with Test Data

To evaluate generalization capability of the network, the model was simulated using the data it has not seen before. An example of the response of the network is shown in Table 5.

The model was simulated with iterations of 10, 20, 50, 75, 100, 125, and 150 with actual and predicted load curves plotted as shown in Figure 1 and Figure 2 below. Load curve or load profile is a chart showing power (watts) supplied, plotted against time of occurrence, and illustrating the varying magnitude of the load during the period covered.

Table 5: Results Obtained for 150 Iterations

Instant	Actual	Predicted	Error
1	662.033	660.836	-1.197
2	663.635	666.608	2.973
3	642.009	635.556	-6.453
4	623.551	623.964	0.413
5	615.033	614.11	-0.923
6	614.061	614.11	0.049
7	608.088	614.11	6.022
8	616.155	614.11	-2.045
9	610.728	614.11	3.382
10	611.763	614.11	2.347
11	659.499	654.929	-4.57
12	663.161	667.72	4.559
13	716.207	716.199	-0.008
14	727.377	716.199	-11.178
15	776.198	784.851	8.653
16	780.615	784.851	4.236
17	823.198	822.77	-0.428
18	855.692	855.77	0.078
19	860.308	861.668	1.36
20	888.874	874.91	-13.964
21	852.995	852.877	-0.118
22	835.863	839.805	3.942
23	833.689	839.805	6.116
24	827.313	822.984	-4.329
25	823.923	822.984	-0.939
26	817.122	822.984	5.862
27	796.451	787.292	-9.159
28	790.85	787.292	-3.558
29	778.648	780.226	1.578
30	785.153	780.226	-4.927
31	751.56	758.528	6.968
32	759.124	758.528	-0.596
33	758.434	758.528	0.094
34	753.668	758.528	4.86
35	755.553	758.528	2.975
36	769.929	769.289	-0.64
37	798.99	796.014	-2.976
38	835.901	826.134	-9.767
39	947.852	958.835	10.983
40	1016.698	1011.646	-5.052
41	1016.099	1011.646	-4.453
42	992.103	993.917	1.814
43	949.532	947.262	-2.27
44	920.581	928.519	7.938
45	852.253	845.467	-6.786
46	802.83	808.718	5.888
47	733.297	734.195	0.898
48	699.117	701.452	2.335

Correlation Coefficient = 0.9989
Mean Absolute Error = 4.0139
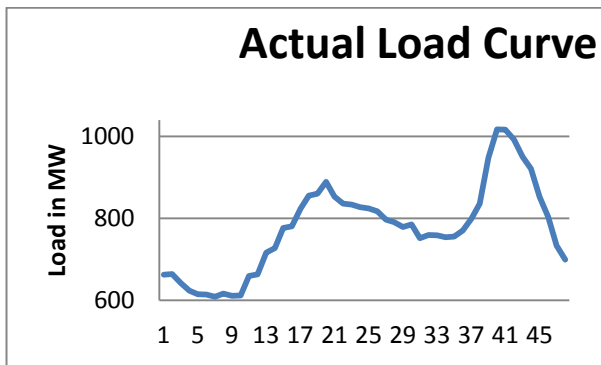Root Mean Squared Error = 5.2194
Relative Absolute Error = 4.6535
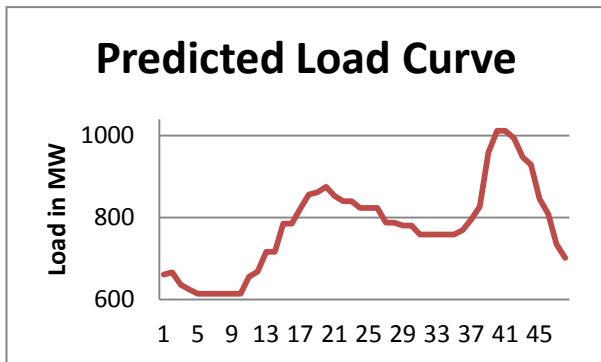
Figure 1: Actual Load Curve

Figure 2: Predicted Load Curve

The accuracy of the results depends on the number of iterations i.e. the mean absolute error was found to decrease with the increase of number of iteration. From the results obtained the mean absolute error decreased from 22.2046% for 10 iterations to 4.0139% for 150 iterations. The results obtained compare well with those obtained by other researchers (Adepoju et al 2007, Sheikh and Unde 2012, Caciotta et al 2009, Hayati 2007) as indicated in section 2.5.

4.2 Discussion
The Kenya Electricity Grid Code defines the 'normal operating frequency excursion band' as the allowable sustained operating frequency range. It gives the range to be between 45Hz (5Hz or 10% negative error allowed) to 52Hz (2Hz or 4% positive error allowed). Outside this range, the system operator may use load shedding to aid recovery of frequency to within the normal frequency tolerance band. The utility frequency or mains frequency in Kenya National Electricity Grid is 50Hz. It is the frequency at which alternating current is transmitted from a power plant to the end user. Using the sample result Table 5 we get:

Mean for the test results = 776.5Mw

Mean absolute percentage error = ±2.85%

Correlation coefficient = 0.9639 (almost perfect correlation)

The simulation results obtained using the forecasting model showed that Short Term Load Forecast can be accurately implemented using only past load values.

5. CONCLUSION
The objective of the study was to develop an accurate and reliable model for short term load forecasting (STLF) for the Kenya National Grid which may assist to economically optimize power system operations. The study showed that better results in STLF were achieved using ANN with Additive Regression as the classifier. The forecasted results

demonstrate that the model performs remarkably well with increased number of iterations. The mean absolute error was found to decrease with the increase of number of iteration.

The study proposed a new ANN methodology to forecast the short-term electricity demand. The proposed additive model allows nonlinear terms within the regression framework, which can capture the complex non-linear relationship between electricity demand and its input variables.

The training and testing of the model were conducted offline. The hourly load data used was collected from one grid. The model needs to be tested on data set from other grids, so that utility of the model can be verified for other load patterns. The results were obtained from simulations and experimentations. No mathematical justifications could be provided for the same.

6. ACKNOWLEDGMENTS

7. REFERENCES
[1] Adepoju, G.A., Ogunjuyigbe, S.O.A. and Alawode, K.O. (2007). Application of Neural Network to Load Forecasting in Nigerian Electrical Power System. *The Pacific Journal of Science and Technology,* Vol 8. No., pp 68- 72

[2] Alfares, H.K., Nazeeruddin, M., (2002). Electric Load Forecasting: Literature Survey and Classification of Methods. *International Journal of Systems Science,* Vol. 33, No. 1, pp 23-34

[3] Buhari, M., Adamu, S.S., (2012). Short-Term Load Forecasting Using Artificial Neural Network. *Proceedings of the International Multi-Conference of Engineers and Computer Scientist*, Vol 1.

[4] Caciotta, M., Giarnetti,S., and Leccese, F., (2009). Hybrid Neural Network System for Electric Load Forecasting of Telecommunication Station, *XIX IMEKO World Congress Fundamental and Applied Metrology, pp 657-661*

[5] Desouky, A.A., and Elkateb, M.M., (2000). Hybrid Adaptive Techniques for Electric-Load Forecast using ANN and ARIMA. IEE *Proceedings of Generation, Transmission and Distribution. 147(4), 213 – 217*

[6] Islam, B.U. (2011). Comparison of Conventional and Modern Load Forecasting Techniques based on Artificial Intelligence and Expert Systems. *International Journal of Computer Science Issues,* Vol. 8, Issue 5, No 3, pp 504-513

[7] Fan, S., and Hyndman, R.J., (2012). Short-Term Load Forecasting Based on a Semi-Parametric Additive Model. *IEEE Transactions on Power Systems*, 27(1), 134-141

[8] Feinberg, E.A., Genethliou, D., (2005). Load Forecasting. In *Applied Mathematics for Restructured Electric Power Systems: Optimization, Control, and Computational Intelligence* (J. H. Chow, F.F. Wu, and J.J. Momoh, eds.), Spinger, pp. 269-285

[9] Hayati, M. (2007). Short Term Load Forecasting Using Artificial Neural Networks for the West of Iran. *Journal of Applied Sciences,* 7(12), 1582 - 1588

[10] Mishra, S. and Patra, S.K.(2008a). Short Term Load Forecasting using Neural Network Trained with Genetic Algorithm & Particle Swarm Optimization. *1ˢᵗ International Conference on Emerging Trends in Engineering and Technology, ICETET.*

[11] Mishra, S. and Patra, S.K., (2008b). Short Term Load Forecasting using a Neural Network Trained by a Hybrid Artificial Immune System. *IEEE Region 10 Colloquium and the Third International Conference on Industrial and Information Systems*, Kharagpur, India.

[12] Mordjaoui, M., Boudjema, B., Bouabaz, M., Daira, R., (2010). Short-Term Electric Load Forecasting Using Neuro-Fuzzy Modeling for Nonlinear System Identification. *Proceedings of International 3ʳᵈ Conference on Nonlinear Science and Complexity (NSC'10),*No. 64, Ankara, Turkey

[13] Quaiyum, S. Khan,Y.I., Rahman, S., Barman, P., (2011). Artificial Neural Network based Short Term Load Forecasting of Power System. *International Journal of Computer Applications,* Vol. 30, No.4, pp 1-7

[14] Sheikh, S. K. and Unde, M.G. (2012). Short-Term Load Forecasting Using ANN Technique. *International Journal of Engineering Sciences & Emerging Technologies*, Vol. 1, Issue 2, pp 97-107