Electrical Demand Load Forecasting by ARIMA Regression and Artificial Neural Networks

Thays Abreu*, Carlos R Santos Junior, Anna P. Lotufo
Electrical Engineering Department
University EstadualPaulista, Unesp
Ilha Solteira, Brazil
* Email: thays7abreu [AT] gmail.com

Mara L. M. Lopes
Mathematics Department
University Estadual Paulista, Unesp
Ilha Solteira, Brazil

Abstract— **The use of hybrid techniques has been applied to solve several problems, including convergence and precision. This paper proposes a hybrid methodology using SARIMA regression models and Multilayer Perceptron Neural Network trained by Levenberg-Marquardt algorithm applied to short term load forecasting. This forecasting is a notably important task for the planning and operation of electrical power systems to anticipate when and how much generation and transmission must be available to provide the required electrical load without interruption. The results are presented for a Brazilian electrical demand energy time series.**

Keywords- SARIMA Models by Box & Jenkins, term load forecast, artificial neural network.

I. Introduction

Electrical power systems have been growing in size and complexity and it is necessary to provide alternatives to minimise the system operation and generation costs. Electrical load forecasting is a fundamental task in operation and control centres, therefore, this task must be precise to enable the system to operate with security and reliability to guarantee the energy is provide safely, economically and continuously. Some techniques that are emphasied in the literature for load forecasting [1] include: simple or multiple linear regression, exponential smoothing, state estimation, Kalman filter, and the Auto Regressive Integrated Moving Average (ARIMA) of Box & Jenkins [2]. These methods require the previous load modeling for further application. The parameters that aid the load modeling can include: meteorological conditions such as cloudy days, wind velocity, temperature or others conditions such as non-typical days (holidays, strikes, etc.) [3].

Presently, the use of intelligent techniques such, as artificial neural networks (ANN) [4] is an efficient alternative procedure for effective load forecasting. The great advantage of the ANN is the ability to handle with complex problems by training, the ability to identify and incorporate characteristics of the time series without needing a theoretical formulation and previous load modelling such as those required by statistical methods.

An interesting proposal is to conjugate statistical techniques with intelligent systems resulting in a hybrid model. There are several published using ARIMA of Box&Jenkins and artificial neural networks (Multilayer Perceptron with Back propagation algorithm) [5-9]. These works are applied to time series of several real problems.

Koutroumanidis [5] presentes a description of alternative energy sources to produce electricity. ARIMA and ANN are used, where ARIMA is responsible for generating the load parameters and ANN are used to predict the future prices of selling wood in Greece. The use of the hybrid model has obtained good results, allowing the industries to proceed in the rational planning of producing and selling wood. In the works of Khashei and Bijari [6], [7], a hybrid model with ARIMA and ANN is used to obtain a precise prediction applied to annually register the sunspot, the number of linces that are captured per year in the Mackenzie river in Canada and the dolllar/pound rate tax. The ARIMA models are used in the first phase to generate the necessary data from the historical series, and in the second phase, the ANN is used to model the data generated by ARIMA models and to predict the time series.

The works of Valenzuela [8] and Khashei [9] besides using ARIMA and ANN, also combine Fuzzy Rules and Genetic Algorithms to improve the results.

Considering the good results obtained in the literature using hybrid models, the objective of this work is to presente a hybrid methodology using ARIMA and ANN (multilayer perceptron with Levenberg-Marquardt training) to predict short term electrical loads.

At the first step of the proposed model, the ARIMA is used to model the electrical load series, providing the simulated electrical load series that will be used as input to the neural network (second step) that executes the load prediction.

The proposed methodology is tested with the electrical load historical data from 30, 60 and 90 days to predict the electrical load of the next day. To validate the proposed methodology, the results were compared with other neural network architectures [10] and with the results obtained by the classical models (ARIMA and perceptron muiltilayer neural network). In this way, it is possible to verify that the use of the hybrid model is viable and provides the best results compared to the models that are used separately.

II. Time Series Prediction Models

There are several different approachs to execute time series forecasting, for for example the traditional statistical modesl such as: exponential smoothing and ARIMA of Box&Jenkins in which the future prediction values are limited by being described by a linear function of previous observations. Another class of prediction models is nonlinear, for example the artificial neural networks. The hybrid model proposed in this paper use two techniques, i.e., the linear model (ARIMA) and the nonlinear (ANN).

A. ARIMA of Box&Jenkins models

First, confirm that you have the correct template for your paper size. This template has been tailored for output on the A4 paper size. If you are using US letter-sized paper, please close this file and download the file for "MSW_USltr_format".

The ARIMA of Box&Jenkins models are based on statistical concepts that use the correlation between observations at different times. The identification of the model structure is difficult, but there is some software [11] that executes this step automatically, avoiding the most difficult step of the analysis.

The ARIMA models assume that the future value is a linear function of past observations and random errors. The process that generates the time series is according equation (1) [12]:

$$\emptyset(B)\Delta^d Z_t = \theta(B)a_t \tag{1}$$

of order *(p,d,q)* written as ARIMA(p, d, q), where p and q are orders of $\emptyset(B)$ and $\theta(B)$ respectivelly.

where:

Z_t : current values at time *t*;

a_t : random errors at time *t*;

$\emptyset(B)$: auto regressive operator of order p;

$= 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$

$\theta(B)$: moving average operator of order q;

$= 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$

Δ^d : series differentiation;

$= (1 - B)^d$

d : indicates number of times the series differs until it becomes stationary.

If the series is already stationary, the term d does not exist and the model is termed ARMA (p, q), according to equation (2) [12].

$$\emptyset(B)Z_t = \theta(B)a_t \tag{2}$$

Box&Jenkins methodology [2] is most frequently used to execute time series prediction, consisting in adjusting ARIMA (p,d,q) to a data set. The construction of ARIMA models includes three steps: identification, estimation and verification [2].

The seasonal ARIMA models are an extention of ARIMA models, known as the Seasonal Auto Regressive Integrated Movin Average (SARIMA), represented by SARIMA (p, d, q)(P ,D, Q)m. They are according equation (3) [12].

$$\phi(B)\Phi(B^m)\Delta^d\Delta_m^D Z_t = \theta(B)\Theta(B^m)a_t \tag{3}$$

where:

$\Phi(B^m)$: seasonal auto regressive operator of order P;

$= 1 - \Phi_1 B^m - 1 - \Phi_2 B^{2m} - \cdots - 1 - \Phi_P B^{Pm}$

$\Theta(B^m)$: seasonal moving average operator of order Q;

$= 1 - \Theta_1 B^m - 1 - \Theta_2 B^{m2} - \cdots - 1 - \Theta_Q B^{mQ}$

Δ_m^D : seasonal series differentiation;

$= (1 - B^m)^D$

D : number of seasonal diferences.

The identification, estimation and verification procedures of the seasonal models are identical to those without seasonality, except at the identification phase, where it is necessary to differentiate the series in relation to Δ and Δ_m to produce a stationary state, obtaining the values of d and D. The estimation and verification phases of SARIMA (p, d, q) (P,D,Q)m are similar to the ARIMA (p, d, q) [13].

B. Artificial Neural Networks

Artificial Neural Networks have some advantages in predicting time series due to model series with nonlinear characteristics.

The advantages of using ANN are due to the capacity of generalisation in non-stationary environments whose characteristics can change with time. The generalization of the neural networks occurs to produce adequate outputs to inputs that were not presented during the training, allowing complex and difficult problems to be solved [4].

The multilayer perceptron neural network (MLP) is an important architecture, and it is also one of the most widely used architectures that are applicable to different application problems. MLPs contain at least one hidden layer between the input and the output layers. MLP is a feed-forward network with multiple layers and supervised training [4].

One of the most important characteristics of this neural network is the ability to learn and improve the performance, thus, the neural network must be trained adjusting the weights by a training algorithm. In this study the Levenberg-Marquardt algorithm is used [14].

The weight adaptation by the Levenberg-Marquard algorithm is given by equation (4) [15]:

$$W(k+1) = W(k) - [J^T(W)\, J(W) + \mu_k I]^{-1}\, J^T(W)\, e(W) \tag{4}$$

where:

W	:	synaptic weight vector;
J	:	Jacobian matrix;
I	:	identity matrix;
μ_k	:	Levenberg-Marquardt constant;

$$e(W) = \sum_{i=1}^{n}(y_i - y_{ei})$$

y_i	:	output provided by the neural network;
y_{ei}	:	real output.

The parameter μ_k is a training stabiliser factor that is based, on the Newton method, althougt it is faster because it, avoid long steps that can lead to a convergence error. This algorithm is considered the fastest, and it is executed with few iterations when using an adequate quantity of data. However, when this number is high this training is not adequate due to the computational effort generated by inversing matrices [14, 15].

III. Prepare Your Paper Before Styling

The proposed methodology of this article is a modification of the work presented by Abreu et. al [16]. In [16], as shown in Figure 1, a hybrid model whose objective is to obtain the prediction of electric load in the short term through the best SARIMA model has been developed .

In this article the proposed hybrid model aims at reducing the number of steps used in [16], and consequently, obtaining a good prediction of electric charges with a shorter processing time. It is important to emphasize that the difference of this work is associated with the direct process of simulated series obtained by R software [16] that is used directly in the prediction of electric load through an artificial neural network, a multi-layer perceptron, with Levenberg-Marquardt algorithm.

The methodology developed in this work consists of realising short term load forecasting using the ARIMA of the Box&Jenkins model and the Multilayer perceptron neural network trained by the Levenberg-Marquardt algorithm. The load time series contain hourly data and is considered with seasonality presenting correlations in different times, i.e., correlation between loads at night must be high and are likely greater than the correlation between loads during the morning of the same day.

Before initiating the first step, the real data are normalised to avoid saturation of the neural network. The normalisation is given as follows:

$$CN = (C \times 1{,}2)/\, Vmax(C)$$

where:

CN	=	normalised load;
C	=	real load;
$Vmax\ (C)$	=	maximum value of the real load.

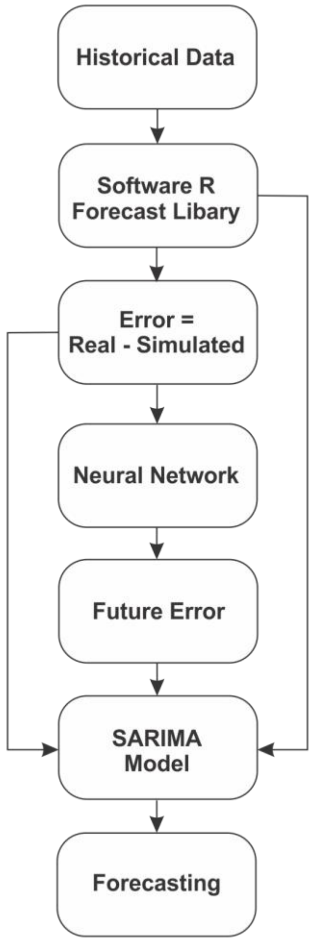

Fig. 1. Modelo Híbrido desenvolvido por Abreu [16]

After being normalised, the load data are applied to R software by forecast library [17, 18] to determine the simulated and predicted series of the best SARIMA model. The simulated series is used as the input of the neural network. Figure 2 below shows the structure.

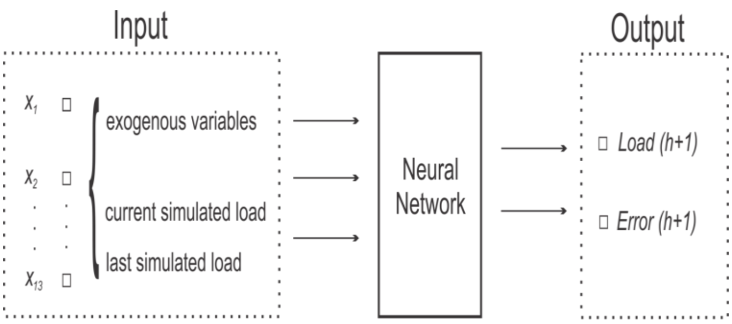

Fig. 2. Neural Network structure

The neural network specified in Figure 2 is composed of three layers, where the input layer has 13 neurons, the hidden layer has 27 neurons and the output layers contain 2 neurons.

The set containing the input and output vector is defined respectively by equations (5) and (6):

$$X_h = [\, te \;\; L(h-3) \;\; L(h-2) \;\; L(h-1) \;\; L(h)]^T, \quad X \in R^b \tag{5}$$

$$Y = [\, C(h+1) \quad E(h+1)], \;\; Y \in R^2 \tag{6}$$

where:

b	=	dimension of vector X;
$L(h-v)$	=	simulated load v hours previous of current hour h;
$C(h+1)$	=	load value to the subsequent hour to current hour h;
$E(h+1)$	=	error of the subsequent hour to current hour h;

At the input layer, vector te representes the nine inputs of the neural network composed by binary data corresponding to the exogenous variables: holidays, day of the week and hour of the day. The last four input $L(h-3)$ $L(h-2)$ $L(h-1)$ $L(h)$, correspond to the simulated loads obtained by SARIMA at time $h-3$, $h-2$, $h-1$ and h. The output of the neural network is the predicted load and error 24 hours in advance.

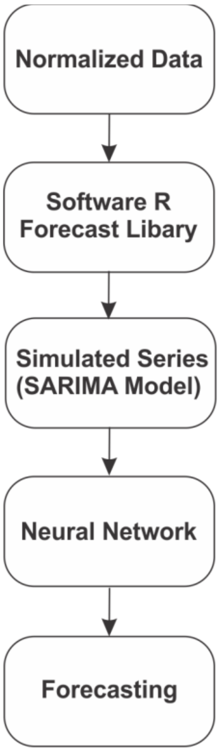

Fig. 3. Flowchart of the proposed model

The load prediction is executed by adding the outputs of the neural network according to equation (7). The proposed methodology is shown in Figure 4.

$$Prediction = \; L(h+1) + E(h+1) \tag{7}$$

The training and test phases of the neural network were executed using the MATLAB software with the Neural Network toolbox [19].

IV. Application and Results

The proposed model aims to predict short term load forecasting 24 hours in advance.

To test the efficiency of the hybrid model, three different applications were executed according to the quantity of training days. The predited day is the 24 hours of the day following the last day of the historical training.

TABLE I. PERIODS USED IN THE THREE APPLICATIONS

Application	Period	Number of vectors
1	31 days	744
2	61 days	1464
3	84 days	2208

To evaluate the precision of the results, the MAPE (mean absolute perceptual error) and the maximum prediction error are calculated. The real load values are compared with the estimated values calculated by the hybrid model. These errors are calculated according to equations (8) and (9), respectively [20].

$$MAPE = \frac{1}{NT}\sum_{h=1}^{NT}\{|C(h) - \underline{C}(h)|/C(h)\} \times 100\% \tag{8}$$

$$Max\ error(\%) = máx\{|C(h) - \underline{C}(h)|/C(h)\} \times 100\% \tag{9}$$

where:

$C(h)$: real load referred to hour h;

$\underline{C}(h)$: estimated load by hybrid model refered to hour h;

NT : total quantity of hours.

The obtained results with the hybrid model are compared with the following architectures: multilayer perceptron with descent gradient and momentum, multilayer perceptron with Levenberg-Marquardt training [10], and also with SARIMA models obtained for each application.

Tables 2, 3 and 4 show the results obtained analysing the MAPE and maximum error performance for applications 1, 2 and 3, respectively.

Figures 4, 5 and 6 present the load prediction curves 24 hours in advance.

TABLE II. MAPE AND MAXIMUM ERROR FOR APPLICATION 1

Models	MAPE (%)	Maximum Error (%)
Hybrid Model	1,353	1,726
SARIMA	9,096	18,623
Perceptron trained by descent gradient with momentum	1,998	4,845
Perceptron with Levenberg-Marquardt training	1,178	3,560

TABLE III. MAPE AND MAXIMUM ERROR FOR APPLICATION 2

Models	MAPE (%)	Maximum Error (%)
Hybrid Model	1,187	1,286
SARIMA	7,774	16,955
Perceptron trained by descent gradient with momentum	2,093	5,022
Perceptron with Levenberg-Marquardt training	1,490	4,267

TABLE IV. MAPE AND MAXIMUM ERROR FOR APPLICATION 3

Models	MAPE (%)	Maximum Error (%)
Hybrid Model	1,093	0,913
SARIMA	8,747	18,630
Perceptron trained by descent gradiente with momentum	2,313	7,088
Perceptron with Levenberg-Marquardt training	1,698	4,687

Analysis of Tables II, III and IV shows that the hybrid model is superior to the other neural networks and the SARIMA model for the three applications proposed. Considering the economical generation point of view, results with less than 5% precision present a significant reduction in the generation costs [21]. Thus, decreasing this precision is economical and the proposed model obtains results below 1%.

As the quantity of vectors increase, better results are obtained,which is a fundamental property of the ANN, which assures that, as more data are used for training better results are obtained at the diagnosis (test phase).

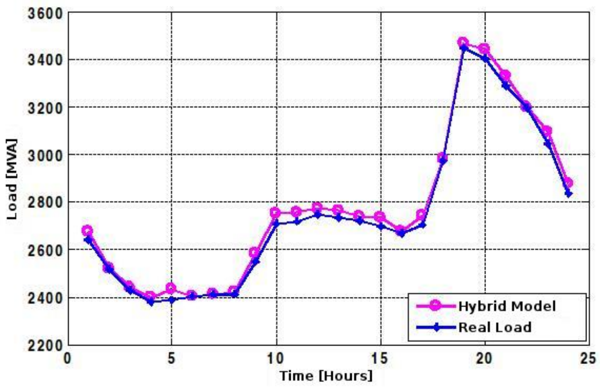

Fig. 4. Predicted and real load by hybrid model for application 1.

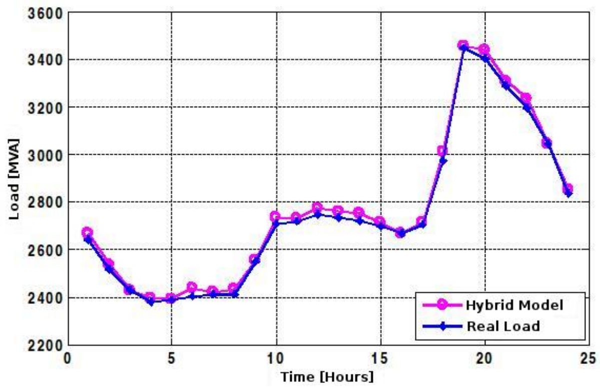

Fig. 5. Predicted and real load by hybrid model for application 2.

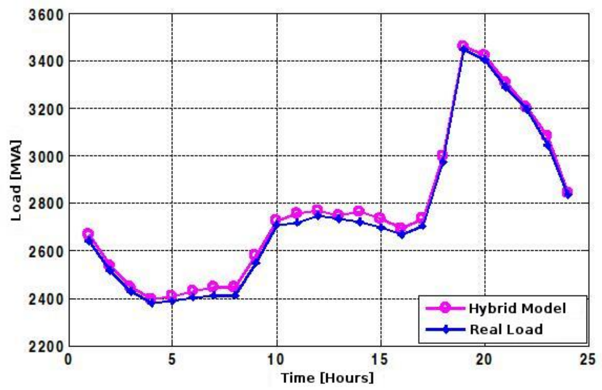

Fig. 6. Predicted and real load by hybrid model for application 2.

Analysing the results described in Tables 2 to 4, it is observed that the hybrid model has the least maximum error when compared to the three other applications, while the

MAPEs are all below 1,4%. It is emphasised that the results of the hybrid model are superior to those of the ARIMA models for the three applications, and they are superior to the Perceptron multilayer with the Levenberg-Marquardt algorithm for applications 2 and 3. This is due to the fundamental property of artificial neural networks that assures that as the quantity of data increases, the better the quality of the obtained results becomes at the diagnosis phase (test phase).

It is observed that the curves obtained by the proposed model follow the pattern of the real data for the three applications, indicating a good precision of the hybrid model.

V. CONCLUSIONS

This work developed a hybrid model using ARIMA of Box&Jenkins and artificial neural networks with the multilayer perceptron trained by Levenberg-Marquardt algorithm aiming to improve the results of short term load forecasting.

The ARIMA models are used to address the linear portion and the ANN for the nonlinear portion of the time series to execute the forecasting.

To evaluate the performance, three tests were executed containing different periods of historical data. The short term load forecasting is executed considering data from 24 hours in advance with MAPEs less than 1.34% and maximum errors less than 1.8%. It is observed that the statistical model with the intelligent system has the best results when compared to the models that are used separately.

ACKNOWLEDGMENT

The authors thank CAPES (Brazilian Research Foundation) and FAPESP (State of São Paulo Research Support Foundation) - 2010/04608-5 and 2013/03853-4 for financial support.

REFERENCES

[1] T. M. O'Donovan, Short Term Forecasting: An Introduction to the Box-Jenkins Approach, New York: John Wiley & Sons, 1983.

[2] G.E. Box, G.M. Jenkins, Times series analysis: forecasting and control. San Francisco, USA: Holden- Day, 1976.

[3] M.L.M. Lopes, A.D.P Lotufo, C.R. Minussi, "Electric Load Forecasting Using a Fuzzy ART&ARTMAP Neural Network", Journal of Applied Soft Computing, vol. 5, n.2, p. 235-244, 2005.

[4] S. Haykin, Neural networks: A comprehensive foundation, Prentice-Hall, Upper Saddle River, New Jersey, USA, 1994.

[5] T. Koutroumanidis, K. Ioannou e G.Arabatziz, "Predicting fuelwood prices in Greece with the use of ARIMA models, artificial neural networks and a hybrid ARIMA–ANN model", Energy Police, v. 37 pp. 3627-3624, 2009.

[6] M. Khashei e M. Bijari, "An artificial neural network (p, d, q) model for time series forecasting", Expert Systems with Applications, vol. 37, pp. 479-489, 2010.

[7] M. Khashei e M. Bijari, "A novel hybridization artificial neural networks and ARIMA models for time series forecasting", Applied Soft Computing, vol. 11, pp. 2664-2675, Nov. 2011.

[8] O. Valenzuela, I Rojas, F. Rojas, H. Pomares, L. J. Herrera, A. Guillen, L. Marquez, M. Pasadas, Hybridization of intelligent techniques and ARIMA models for time series prediction, Fuzzy Sets and Systems, v.159, n.7, p. 821-845, 2008.

[9] M. Khashei, M. Bijari, G. A. R. Ardali, A. Improvement of auto-regressive integrated moving average models using fuzzy logic and artificial neural networks. Neurocomputing, v. 72, p. 956-967, 2009.

[10] K. Nose Filho, K. A. M. Araujo, J. L. Y. Maeda e A. D. P. Lotufo, "Previsão de cargas elétricas utilizando diferentes arquiteturas de redes neurais artificiais com o auxilio das toolboxes do MATLAB", In: 8th Latin American Congress Electricity Generation and Transmission, 2009, Ubatuba, SP. (in Portuguese).

[11] G. E. P. Box, G. M. Jenkins e G. Reinsel, Times Series Analysis: Forecast and Control, 3 rd ed. Englewood Cliffs: Prentice Hall, 1994.

[12] P. J. Brockweell e R. A. Davis, Times Series: Theory and Methods, 2nd ed. New York: Springer, Nov. 1990.

[13] P. A. Morettin and C. M. C. Toloi, Análise de Séries Temporais, 2nd ed., São Paulo: Edgard Blucher, 2006. (in Portuguese).

[14] L. M. Saini and M. K. Soni. "Artificial neural network based peak load forecasting using Levenberg-Marquardt and quasi-Newton methods" IEEE Proceedings on Generation Transmission, Distribution, vol.149, n.5, p. 578- 584, 2002.

[15] A. H. Barbosa, M. S. R. Freitas e F.A. Neves, "Confiabilidade estrutural utilizando o método de Monte Carlo e redes neurais" Revista Escola de Minas, vol. 58, no 3, pp. 247-255, Jul/Set 2005. (in Portuguese).

[16] T. Abreu, M. Paredes, K. A. M. Araújo, M. L. M. Lopes and A. D. P. Lotufo, " Electrical Load Forecasting using a Hybrid Model with Artificial Neural Networks" International Journal of Computer and Information Technology, vol. 04, pp. 2279-0764, March 2015.

[17] R. J. Hyndman, K. Yeasmin. "Automatic time series forecasting: the forecast package for R". Journal of Statistical Software, v. 27, 2008.

[18] J. D. Cryen, K. S. CHAN, Time series analysis: with applications in R. 2. ed. New York: Springer, 2008.

[19] H. Demuth and M. Beale, "Neural Network Toolbox for use with MATLAB". User's Guide, 1998.

[20] D. Srinivasan, S. S. Tan, C. S. Chang, e E. K. Chan, "Practical Implementation of a Hybrid Fuzzy Neural Network For One-Day-Ahead Load Forecasting," IEE Proceedings Generation, Transmission and Distribution, vol. 145, no 6, pp. 687-692, Nov. 1998.

[21] K. D. Ranaweera, G. G. Karady and G. R. Farner, "Economic impact Analysis of Load Forecasting," IEEE Transactions on Power Systems, Vol. 12, No. 3, pp. 1388-1392, Aug. 1997.