Institutional Electricity Load Forecasting using Classical and Intelligent Forecasting Techniques

Surendra Singh Patel[1][2], Amit Singh[1], Navjot Kumar[1][2], PC Panchariya[1][2], SA Akbar[1][2]
[1] Cyber Physical System Area, CSIR-CEERI, Pilani, Rajasthan, India
[2] Academy of Scientific and Innovative Research (AcSIR)
Email : patelsurendra333@gmail.com

Abstract: **An accurate forecasting of Institutional Electricity load can proved to be useful asset for efficient utilization of the infrastructure available in terms of future demand and supply. Time series method of forecasting has got very wide applications like sales forecasting, yield prediction and Supply Chain Monitoring (SCM) system etc. This paper presents a classical time series models available for predicting the future demand. The classical models used to predict the future load and demand assumes the linear relationship between input and output but in the real world this doesn't seems to be practical. The intelligent and self-learning models like neural network has the lead to approximate any kind of non-linear function and can fit into these situations. Classification and prediction capabilities of Neural Network has also shown a great potential in forecasting. A neural network based time series forecasting model is also developed for electricity load forecasting. Behavioral pattern and trend of the experimental data are being studied and analyzed for accurate forecasting of electricity load.**

Keywords: Time Series Forecasting, Load forecasting, ARIMA model, Neural Network, Levenberg Marquardt Algorithm, Bayesian Regularization etc.

I. INTRODUCTION

An accurate prediction of future demand of electricity for an institution is essential for the operation
and also planning of utility. Time series is a useful technique which can provide the future insight of demand and supply for efficient utilization of available resources. A Time series is a series of sequentially recorded observations of any physical quantity or financial variable obtained over a constant and equal interval of time. For example in engineering field the data obtained from different sensors like temperature, pressure, humidity etc. constitute a discrete-time data series by sampling the continuous signal coming from the sensors [1]. Time series can play different roles in large data analysis to understand the trend that the data is following or can make predictions with the help of behavioral analysis of the data.

A rigorous attempt is required to make better interpretation or understanding of dataset so that an accurate prediction can be achieved. This type of investigation is known as time series analysis. Under this study we try to find the various characteristics of time series e.g. seasonality, stationarity, linearity etc.

Why Time Series required?

There are others structural models available for predicting the future values instead Time series analysis is becoming popular due to its inherent advantages of developing forecasting model with the help of past values of data itself. Since there are various parameters which are not observable to make predictions in the case of structural model [2] e.g. Stock price data is available on daily basis, whereas inflation & unemployment data is not. Some factors are not observable like unemployment data is difficult to find in developing countries like in India.

Time series analysis provide a great insight about the dataset which has an important role especially when it comes to forecasting and it enfolded many different forecasting models. However, which model is best suited for a particular application is required to determine based on the characteristics of dataset. Hence to identify the best forecasting model is one of the biggest challenge of food or beverage industry due to short perishability of its products which ultimately affect their loss or profit [3].

Miller and Shanklin (1988) [4] noted that forecasting becomes very critical when it comes to forecast the self-life of food products because of the perishable nature of these food items. Due to short perishability of these products, the items are prepared immediately based on the demand. Hence an inaccurate forecast may lead to overproduction or under-production and eventually the action may result to wasted food and increased food cost.

Sang-Rok Yoo et.al.[5] estimated future marine traffic volume using a popular time series model. In this work the author developed the ARIMA forecasting model for the prediction of Marine traffic volume at different ports of South Korea and compared this model with Simple Seasonal Model and Winters Multiplicative Model. They have concluded that there are chances when different model may prove to be efficient for different application.

The advancement in the field of science empower the human being to understand the various climatic pattern and consequently the predictability of a variety of events. A wide variety of forecasting methods are available to management (see, for example, Makridakis and Wheelwright, 1989). To incorporate complex behavioral pattern associated with forecasting, there are various approaches like neural network etc. are investigated in literatures [6][7][8].

Thomas Kolarik et.al. [7] developed ARIMA model Artificial Neural Network model to forecast univariate time series. The authors showed that these intelligent techniques are more promising as compared to conventional techniques. Bogdan Oalcea and S. Cristian [8] compared the different neural network training algorithms to predict the exchange rates EUR/RON and USD/RON. The results of the work suggested the method to improve the forecasting accuracy by removing the correlation between the inputs and make them statistically independent. Jihoon Moon et.al. [9] developed two forecasting models based on artificial neural network and support vector regression. The authors also highlighted the importance of reliable electrical supply for smoothly functioning of any institute.

Various factors affect the electricity load e.g. temperature, climatic conditions and past usage pattern etc.[10]. Palit et.al. developed a neuro fuzzy forecasting model to take care of these uncertainties for making the accurate prediction of short term and medium term load consumption.

The organization of paper is as follows: we present the traditional forecasting model namely ARIMA in section 2. Section 3 briefly covers the training algorithms used in this paper for training of NN model. In result and discussion part we present results obtained and provide some comparative remarks between training algorithms. All the plots have been generated through simulation in MATLAB.

II. TRADITIONAL TIME SERIES MODEL

There are various methods of forecasting available in literatures and Time series analysis is one of them which uses historical data as the basis of estimating future outcomes. In time series method there are different models as listed below:

- Autoregressive Model (AR)
- Moving Average Model (MA)
- Autoregressive Moving Average Model (ARMA)
- Autoregressive Integrated Moving Average Model (ARIMA)
- Exponential Smoothing Model

Experimental Data Set

The experimental load data used in this research is the electricity consuption at CSIR-Central Electronics Engineering Research Institute (CSIR-CEERI). The time series plot for monthly consuption of electricity is given in Figure 1 for observation of the data. It can be observed from the plot that there is an ovious seasonal effect exist during the month of June-July of every year in which the consumption gets increases. It is also observed that the mean is not constant over the complete period so with these observations it can be concluded that the electricity load consumption data is not stationary.

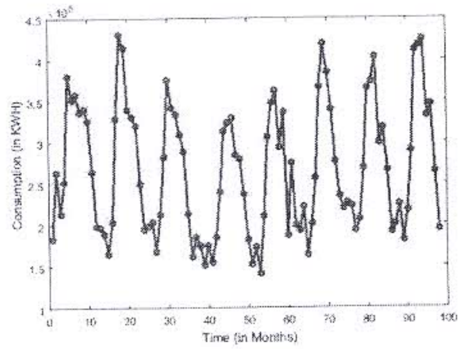

Figure 1: Monthly Electricity Consumption Time Series Plot

For making the time series stationary, there are two methods suggested in the literature namely differencing and transforming. In differencing method, in order to make series stationary we have to take difference between the data points. Once we take the difference, plot the series and see if there is any improvement in ACF and PACF curve. If not we can try a second or even third order differencing. But while differencing one should keep in mind that more we difference, the more complicated analysis we are going to come across.

If we cannot make time series stationary by differencing we can try out transforming the variables. Log transformation is probably the most commonly used transformation. But it is suggested that one should use transformation only in case when the differencing doesn't work.

In this work also we have tried with differencing only to make our time series stationary. But after differencing one time it is observed that there was no significant improvement in the seasonal effect so we have tried second order differencing and it makes the satisfactory improvement to make the series stationary. Figure 2 and Figure 3 depicts time series plot for first order and second order differencing respectively.

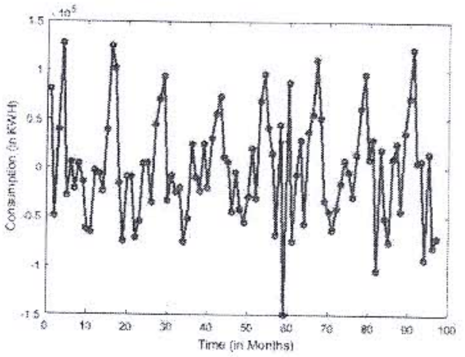

Figure 2: Time series plot after 1st order differencing

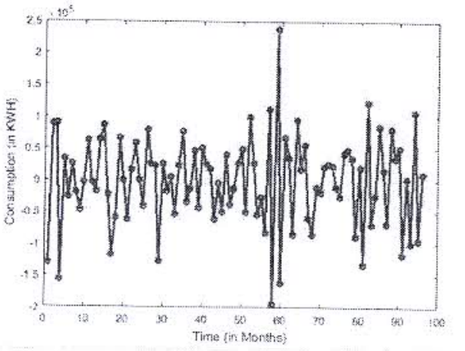

Figure 3: Time series plot after 2nd order differencing

Preprocessing of Experimental Dataset

Since the electrical load data has been recorded on monthly basis and there is very large difference observed between peak values of monthly data in summer (especially in the month of June) and in winters (especially in the month of February) in terms of magnitude and not suitable for neural network kind of techniques since the training data, validation data and testing data for neural network forecasting model are based on available electrical load only. Hence, the electrical load consumption data have been first normalized using maximum normalization method within the scale of 0 and 1.

The ARIMA Model

The ARIMA model is the variant of Box-Jenkin's ARMA model and is used when the given time series is not stationary. The nonstationary time series can be transformed into stationary series by means of differencing between two successive lags of the series and hence a new time series can be obtained. Sometimes it may happen that the non-stationarity cannot be removed by performing one time difference hence in that case we have to perform differencing again on the newly generated series and this can go on upto certain limit. As if we keep on differencing the series to get stationarity, the complexity gets increased eventually. Hence there is a trade-off between complexity and differencing. So another method has also been proposed in literatures which is known as Log transformation. But it is also suggested that log transformation has to be performed only when the given time series couldn't get stationary by differencing. In the acronym ARIMA, the letter I stands for integrated which means the number of times the series gets differentiated to make it stationary.
The commonly used mathematical expression to define the structure of ARIMA models is ARIMA(p, q, d),

$$y_t = \mu + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \ldots.. + \alpha_p y_{t-p} + \varphi_1 u_{t-1} + \varphi_2 u_{t-2} + \ldots.. \varphi_q u_{t-q} + u_t \ldots..(1)$$

Where ***'p'*** is the number of autoregressive parameters, ***'q'*** is the number of moving-average parameters, and ***'d'*** is the number of differencing passes. Here in this forecasting model the value of ***'d'*** comes out to be 2 since we have differenced twice to make our time series data stationary. The parameters $\alpha_1, \alpha_2, \ldots\ldots \alpha_p$ and $\varphi_1, \varphi_2, \ldots\ldots, \varphi_q$ are the estimating parameters. Least Square Method (LSM) or Maximum Likelihood Estimation (MLE) are most widely used methods to estimate these parameters. In this paper LSM method has been tried to estimate the model parameter.
The graphs of the sample ACF and PACF is plotted for the corrected data to identify the order of ARIMA model is shown in figure (4) and (5) respectively.
The ACF represents pure MA(q) model while PACF represents pure AR(p) model. Either ACF or PACF plots do not vanishes after certain lags hence one cannot determine the order of time series model and hence the forecasting model cannot be developed without this knowledge.

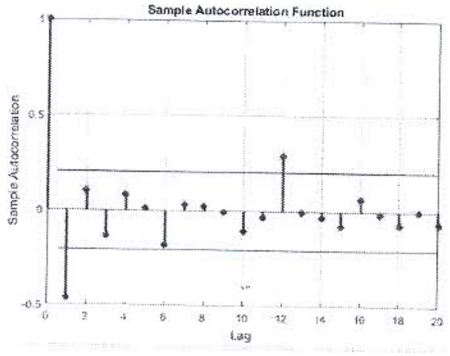

Figure 4: Sample ACF plot for Corrected Series

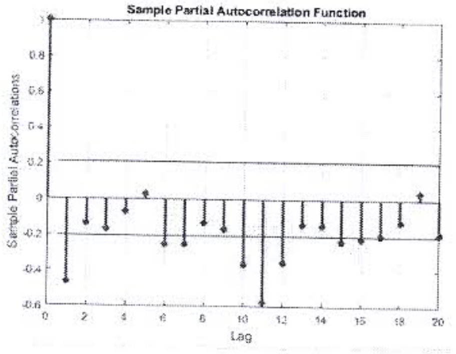

Figure 5: Sample PACF plot for Corrected Series

Since the model order cannot be identified hence it is very difficult to develop classical forecasting model for predicting the electricity load consumption. Although there are numerous techniques available in the literature which can deal such situations e.g. Neural Network, Fuzzy Logic etc. based time series model for forecasting the electricity load. So the other approach apart from the classical one is being proposed in this paper is Neural Network based time series model. Neural Network time series tool provided by MATLAB has been used for training, validation and testing of the model.

III. FEED-FORWARD NEURAL NETWORK

Neural network approximate the mapping from input to output and have the valuable capability for time series to take care noise effect accompanying with the input data and mapping function and possesses self-learning and prediction capability in the case if there is any missing information. One of the elegant ability of neural network techniques is to approximate arbitrary non-linear functions comprised of past values of series itself. The network function in neural network is the nonlinear function of the linear combination of input hence it can approximate any linear or nonlinear relationship between input and output. This ability of neural network has great importance in time series analysis and forecasting applications. Furthermore the neural network have the potential to build relationship between more number of inputs and outputs and hence can support multivariate time series forecasting. Due to these intelligent abilities, Feed-Forward neural networks are commonly used in time series forecasting.

Unlike the linear model for regression and classification, neural network model is simple nonlinear function of input variables and output variables controlled by a vector 'w' known as weight parameter. The parametric nonlinear basis function of Feed-Forward neural network has the form so that each basis function is itself a nonlinear function of a linear combination of the inputs [11]. Hence the basic neural network model can be described by a series of linear transformation as follows:

$$a_j = \sum_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)} \quad \ldots..(2)$$

Where the parameters $w_{ji}^{(1)}$ and $w_{j0}^{(1)}$ are the weights and biases of the model respectively and a_j are known as input activations. Now each of these activations can be transformed using a differentiable, nonlinear activation $h(.)$ as

$$z_j = h(a_j) \quad \ldots..(3)$$

These quantities (eq. 3) are corresponds to the hidden layers of the network. The nonlinear function $h(.)$ can have various forms but generally sigmoidal functions such as logistic sigmoid or $'tanh'$ functions has been used. Like the input activation, output unit activation function can also be described by considering the hidden layer units as:

$$a_k = \sum_{j=1}^{M} w_{kj}^{(2)} z_j + w_{k0}^{(2)} \quad \ldots...(4)$$

Finally considering the output activation function to be logistic sigmoid, the overall network function is given as:

$$y_k(x, w) = \sigma\left(\sum_{j=1}^{M} w_{kj}^{(2)} z_j h\left(\sum_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)}\right) + w_{k0}^{(2)}\right) \quad \ldots..(5)$$

This type of function can be represented as two layer network diagram of feed-forward neural network as shown in figure 6:

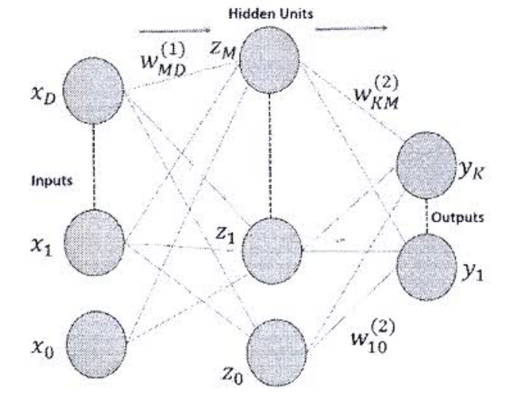

Figure 6: Feed-forward two layer Neural Network Diagram [11]

The electricity consumption data of 59 months is being used for training and validation and 39 months data is used for testing of the algorithm. Three training algorithm namely Levenberg Marquardt, Bayesian Regularization and Complex Conjugate Gradient methods have been given in the tool. In this paper we have tried with first two training algorithms separately and the results are quite satisfactory with both of the techniques.

IV. TRAINING ALGORITHMS FOR NEURAL NETWORK

Levenberg Marquardt (LM) Algorithm

The LM algorithm is one of the popular optimization algorithm used in neural network to train the wide variety of models. The algorithm was first proposed by Kenneth Levenberg [12] in 1944 and later rediscovered by Donald Marquardt [13] in 1963. The algorithm provides a numerical solution for the problem called Nonlinear Least Squares Minimization. The primary application of this algorithm is to estimate the parameter value such that sum of squared error should minimized. The algorithm has the tendency to provide fast and stable convergence.

The Levenberg-Marquardt algorithm is the blend of gradient descent and Gauss-Newton iteration because the update function used in this algorithm comes from these two methods. For initialization of algorithm the gradient descent is attractive because it gives optimized starting point which leads to the algorithm convergence near solutions but at the cost of slow convergence. But this slow convergence problem can be improved significantly if we have some information about the second derivatives of objective function. At this level the role of Newton's method comes into play. The non-linear function to be minimized using this algorithm can be of the form as given below:

$$f(x) = \frac{1}{2}\sum_{j=1}^{m} e_j^2(x) \dots\dots (6)$$

Where $x = (x_1, x_2, \dots\dots, x_n)$ represents the input vector and each e_j is a function $\square^n$ to $\square$ and referred to as a residuals and it is assumed that $m \geq n$.

Bayesian Regularization

Bayesian regularization is the probabilistic approach used in the neural network learning process. It has got one of the great advantage over other training algorithm is that it eliminates the need of a complicated cross validation process and provides the constructive helps in the case where the dataset is comparatively small. This regularization also remove the problem of overfitting in neural network model. In this regularization, the weights of the model are considered to be random variables and hence the algorithm implicitly provides the optimal number of weights or parameters that should be utilized in model. The cost function used in this method can be given as:

$$Y(w) = \beta E_d(w) + \alpha E_w(w) \dots\dots\dots. (7)$$

Where,
$E_d(w)$ is the sum of squared errors i.e. difference between actual and predicted network output and $E_w(w)$ is the sum of squared weights.

V. RESULTS AND DISCUSSION

In this work around 98 months of electricity consumption data has been used for training, validation and testing of Neural Network based time series model. Both the training algorithms were separately tested for forecasting the electricity load. Once the training is completed and the adjustable parameters are set then the same parameters has been used for forecasting. The number of hidden layer and the number of input lags has been chosen heuristically by minimizing the Mean squared error. Regression 'R' value is also being considered one of the parameter for convergence of the algorithm. The results obtained individually using both training algorithm are shown in the figure 7 and figure 8 respectively:

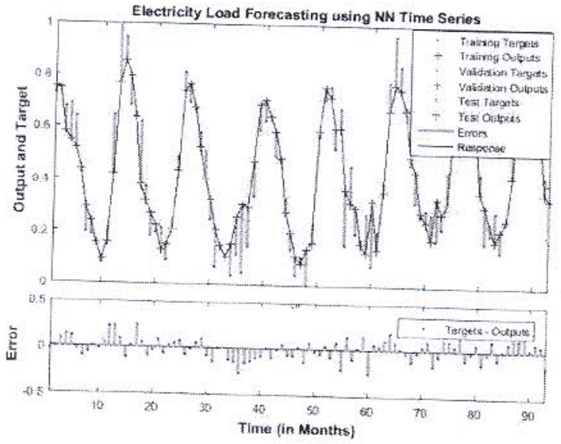

Figure 7: Electricity Load Forecasting output and error plot using NN Time Series with LM Training Algorithm

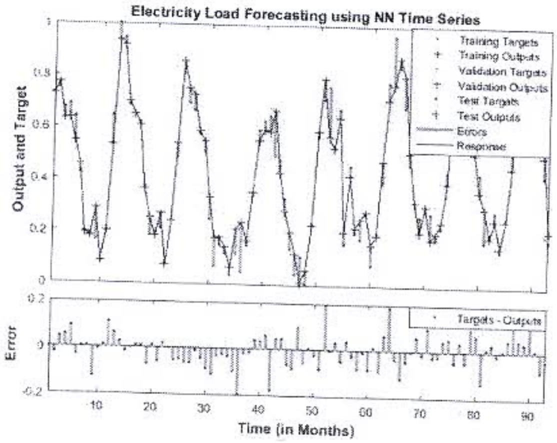

Figure 8: Electricity Load Forecasting output and error plot using NN Time Series with Bayesian Regularization Training Algorithm

It is observed that LM algorithm is fast as compared to Bayesian Regularization algorithms and also showed more accurate forecast results in terms of mean squared error and R value. But if number of hidden layers and number of lags get increased, a significant improvement has been observed in the case of Bayesian regularization training in terms of MSE and R value. It can also be stated that these intelligent methods of forecasting has great advantage over classical techniques since in classical methods the model parameters estimation is of great importance and require significant effort to correctly estimate them for model accuracy.

ACKNOWLEDGEMENT

The authors thank Director; CSIR-CEERI; for the guidance and availing the lab facilities. The authors would also like to thanks CSIR for financial support. The authors also thank the team members of SAG lab; CSIR-CEERI and Mr. Punit Khatri, BITS, Pilani for their help and support. Authors are also thankful to Electrical section of CSIR-CEERI for providing the electricity data to perform the experimental simulations.

REFERENCES

[1] Ajoy K. Palit and D. Popovic, "Computational Intelligence in Time Series Forecasting: Thoery and Engineering Applications", Advances in industrial control, Springer.

[2] Analytics University, YouTube channel.

[3] N. de P. Barbosa, E.da S.Christo, and K. A. Costa, "Demand Forecasting for Production Planning in a Food Company", ARPN Journal of Engineering and Applied Sciences, Vol. 10, No. 16, September 2015 ISSN 1819-6608.

[4] Miller, J. L., & Shanklin, C. W. (1988), "Forecasting menu-item demand in foodservice operations", Journal of the American Dietet ic Association, 88 (4), 443449.

[5] Sang-Rok Yoo, Jung Sik Jeong, Jae-Yong Jong, and Gyei-Kark Park, "Forecast of Marine Traffic Volume using Time Series Model", Proceedings of 2013 International Conference on Fuzzy Theory and Its Application National Taiwan University of Science and Technology, Taipei, Taiwan, Dec. 6-8, 2013.

[6] Makridakis, S. and S.C. Wheelwright, "Forecasting methods for management", 5th ed., New York: John Wiley & Sons.

[7] Thomas Kolarik and Gottfried Rudorfer , "Time series forecasting using neural network", CiteSeerX scientific digital library.

[8] Bogdan Oancea, Ştefan Cristian Ciucu, "Time Series Forecasting Using Neural Networks", Challenges of the Knowledge Society. IT in Social Sciences.

[9] Jihoon Moon, Jinwoong Park, Eenjun Hwang, Sanghoon Jun, "Forecasting power consumption for higher educational institutions based on machine learning", Springer Science+Business Media New York 2017.

[10] Ajoy K Palit, Walter Anheier, Dobrivoje Popovic, "Electrical Load Forecasting Using a Neural-Fuzzy Approach", Studies in Computational Intelligence,Volume 250, Springer.

[11] Christopher M. Bishop, "Patter Recognition and Machine Learning", Information science and statistics, Springer.

[12] K. Levenberg, A method for the solution of certain problems in least squares, Quarterly of Applied Mathematics, 5, 164–168, 1944.

[13] D. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, SIAM Journal on Applied Mathematics, 11(2), 431–441, June 1963.

[14] Eric Lynn Taylor, "Short-term Electrical Load Forecasting for an Institutional/Industrial Power System Using an Artificial Neural Network", Masters Theses, University of Tennessee, Knoxville.

[15] Jenkins, G. M., Practical Experiences with Modelling and Forecasting Time Series, Lancaster: G J P Publications, 1979.

[16] Paul Newbold, "ARIMA Model Building and the Time Series Analysis Approach to Forecasting", Journal of forecasting, Vol.2, 1983.

[17] Box, G. E. P. and Jenkins, G. M., Time Series Analysis, Forecasting and Control, San Francisco: Holden Day, 1970.

[18] Cleveland, W. S., 'The inverse autocorrelations of a time series and their applications', *Technometrics*, 14, (1972), 277-298.