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.

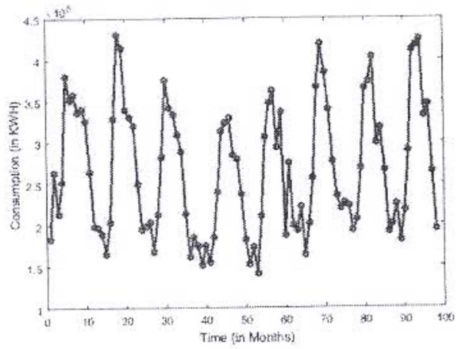

Figure 1: Monthly Electricity Consumption Time Series Plot

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.

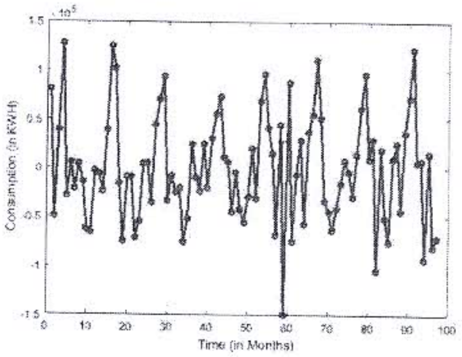

Figure 2: Time series plot after 1st order differencing

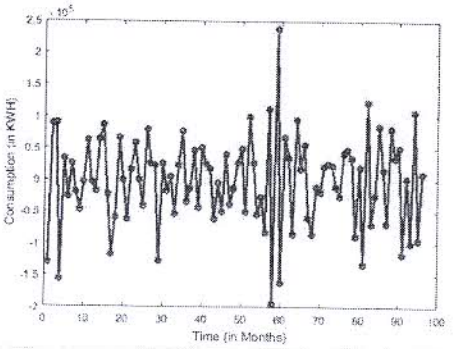

Figure 3: Time series plot after 2nd order differencing

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.