Forecast Daily Night Peak Electric Power Demand in Sri Lankan Power System

A Ananthasingam, A Atputharajah (*Senior member IEEE)*
University of Jaffna, Sri Lanka

***Abstract*— Electricity has become vital for day today life. Electricity consumption has enormously increased during the recent past. However it cannot be conveniently stored. Therefore, at every instant of time there should be a sufficient amount of electricity generation to meet the consumer demand. Depending on the demand, power producers adjust their production level. Therefore forecasting electricity demand would be very helpful to power system operators to maintain the power quality and reliability of the power system. The highest daily electric power demand occurs in night peak in Sri Lankan power system. Therefore this paper describes forecasting the Daily Night Peak Electric Power Demand (DNPEPD) of Sri Lankan power system by using past daily data in time series analysis. Final model can be used to forecast the DNPEPD a week ahead. It is shown that the developed approach can produce more accurate results for the forecast for short term. The developed model is more beneficial for planning of a power generation pattern.**

Keywords*— *quality; reliability; time series; peak demand; short term; and power generation pattern.

I. Introduction

Typical daily electric power consumption in Sri Lanka, as shown in Fig. 1, shows a temporal variation which consist of three major segments, off peak, day peak and night peak. While the maximum electric power demand of 2164.2 MW occurs in night peak and minimum of 558.5 MW occurs in off peak in 2013 [1] [2].

Power generation should meet consumer demand. Therefore as per the consumer demand, power system needs its power generation level to be changed instantly and it should meet the level of consumer demand without any shortage as well as any wastage. Therefore, producing power to match the consumer demand is a really challenging operation as electricity cannot be stored conveniently. If there is a way to forecast the future demand in all three peak times, it would be very helpful to power system operators to decide the power generation pattern, which leads the system to improve the power quality as well as the reliability. It can provide supportive decisions for power system operators whether to rely on their own generators or to seek another independent power producer to satisfy the demand as the Ceylon Electricity Board rely on some independent power producers too[1] [2].

Time series analysis, neural network and wavelets are available to predict electric power demand, however, extensive analysis revealed that time series techniques are more effective than wavelet transform and neural network techniques for short term forecasting[3][4]. Time series analysis with Box Jenkin's Auto Regressive Integrated Moving Average (ARIMA) model methodology is the appropriate way to perform the forecasting. Taylor et al applied seasonal ARIMA to capture daily and weekly seasonality within demand data [5] [6].

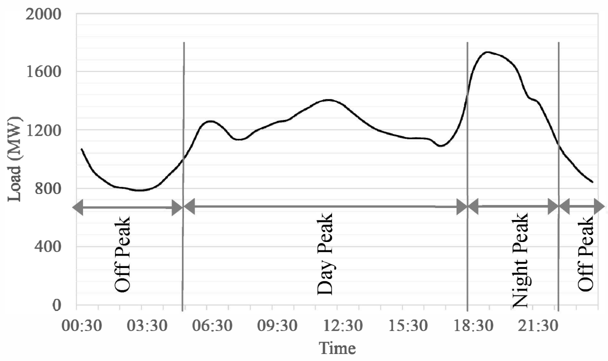

Fig. 1. Typical daily electric power consumption in Sri Lanka

A noteworthy model for a time series analysis is Auto Regressive Moving Average (ARMA) model. ARIMA is based on ARMA Model. The difference is that ARIMA Model converts non-stationary data to a stationary data set before working on it. ARIMA model is widely used to forecast linear time series data [7]. The ARIMA models are often referred to as Box-Jenkins models as ARIMA approach was first popularized by Box and Jenkins. The univariate version of this methodology is a self- projecting time series forecasting method [8]. The underlying technique is to find an appropriate formula so that the residuals are as small as possible and exhibit no pattern.

Normally, the ARIMA model is represented as ARIMA(p, d, q) where p is the number of autoregressive terms, d is the number of non-seasonal differences and q is the number of moving average terms, If there is a seasonal effect in the data set the model will be ARIMA(p, d, q),(P, D, Q)n, where p, d and q are the same as the above and P is the number of seasonal auto regressive terms, D is the number of seasonal differences, Q is the number of seasonal moving average terms and n is the order of seasonal differencing [8]. General mathematical representation of seasonal ARIMA model is given in appendix.

The most important analytical tools with ARIMA modelling are Auto Correlation Function (ACF) and the Partial Auto Correlation Function (PACF), which measure the statistical relationships between observations in a single data series [9]. The ACF gives big advantage of measuring the amount of linear dependence between observations in a time series that are separated by a lag k. The PACF plot is used to decide how many auto regressive terms are necessary to expose one or more of the time lags where high correlations appear, seasonality of the series, trend either in the mean level or in the variance of the series [10] [11].

The prime objective of this paper is to statistically forecast the Daily Night Peak Electric Power Demand (DNPEPD) of Sri Lankan power system a week ahead.

978-1-4799-1876-8/15/$31.00 ©2015 IEEE

Outcome of this research can be used in different endeavors in power sectors such as planning, system expansion and operational and maintenance schedule. Short term forecasting like a week ahead and a month ahead are considered necessary in unit commitment analysis and coordination of different type of plants (hydro/thermal/wind). The main contribution is to decide the reliable and cost effective generation pattern at the planning stage.

The reminder of this paper is organized as follow. Section II describes adopted methodology and Section III provides results and detail analysis. The next section summarizes the conclusion drawn from this study, description on future works. Required mathematical details are given in appendix.

II. METHODOLOGY

The flowchart of the adopted methodology is depicted in Fig. 2. The methodology has been developed from the Box Jenkin's methodology for ARIMA modelling [7] [8]. To start the process, time series data is plotted in suitable time scale. Then, variance of the plotted series is checked whether stable or not. If it is not stable, suitable data transformation is applied on the data to stabilize the variance. After getting the variance stabilized, ACF and PACF of plotted data are analysed and mean is checked whether stable or not. If not, suitable regular or seasonal transformation is applied to get the mean stable.

After obtaining the stabilized data, model selection and testing, where more than one tentative models are selected. By choosing one of the selected tentative model, parameter checking is carried out to check whether they are significant and uncorrelated or not, if it is not, goes to next model, if it is yes, residual checking is carried out to check whether the residuals are uncorrelated or not. If it is not, goes to next model, if it is yes this is the successful model to forecast.

A. Data

DNPEPD data for a period of 365 days, during the year 2013 was collected from Ceylon Electricity Board. To determine the model, data from January 2013 to November 2013 were used as training data set and efficiency of the model was tested with an independent set of remaining data of December 2013. Descriptive statistics of DNPEPD of Sri Lanka in 2013 are shown in Table I.

1. Data Stabilization

In order to make the variance stable, the training data set was plotted (see Fig. 3). It showed that the time series plot was not stable in variance and a transformation was applied. Among many traditional transformations available, for improving normality, the Box-Cox transformation (1) provides a family of power transformations that incorporates and extends the traditional options to achieve the optimal normalizing transformation for each variable [12].

$$x_t' = \begin{cases} (x_t^{\lambda} - 1)/\lambda & \text{where } \lambda \neq 0 \\ \log_e(x_t) & \text{where } \lambda = 0 \end{cases} \quad (1)$$

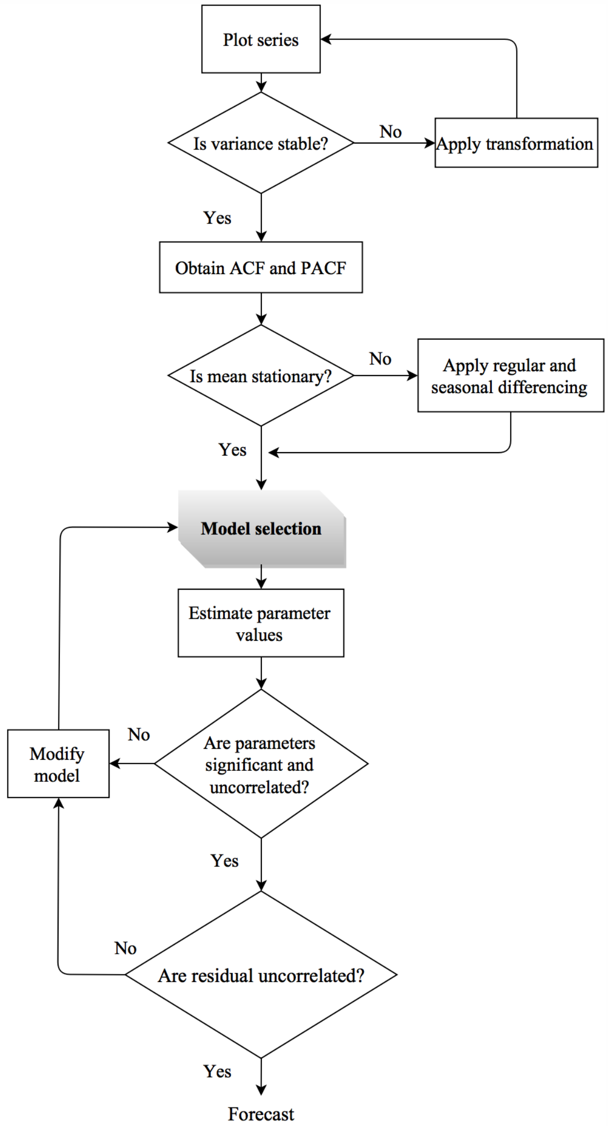

Fig. 2. Methodology

Where, x_t denotes original data and x_t' denotes transformed data and λ refers to Box Cox transformation coefficient.

The value for λ was chosen from a trial and error method, where it was given from -2 to 2 in 0.5 increment and for each instance variance was checked [10]. The stable variance was found at λ = 0.5 and data was transformed, which is square root transformation.

ACF and PACF of Box Cox transformed data were plotted (see Fig. 4.a and Fig. 4.b) to investigate whether the mean of

TABLE I: DESCRIPTIVE STATISTICS OF DNPEPD OF SRI LANKA IN 2013

Number of observations (N)	365
Maximum	2164.20 MW
Minimum	1383.00 MW
Mean	1830.05 MW
Standard Deviation	134.17 MW

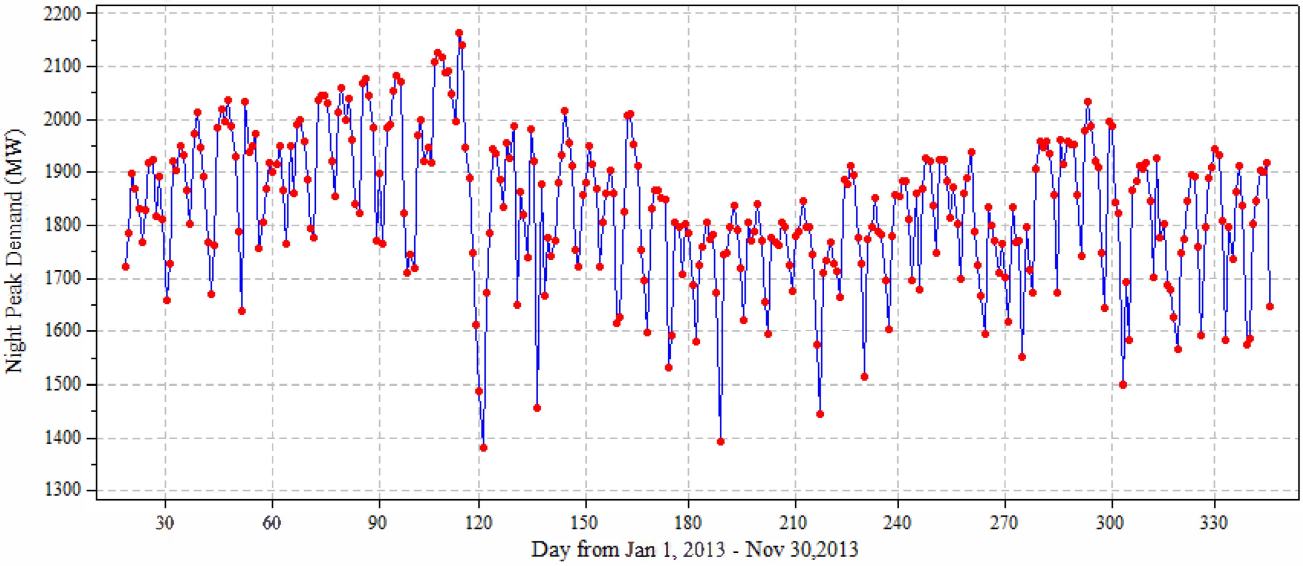

Fig. 3. Time series plot of DNPEPD from Jan 2013 - Nov 2013

the transformed data is stationary and whether it has any seasonal effect, where the blue lines indicate the correlation values of Box Cox transformed data at each lag and the red lines indicate the 5% significance limits (95% confidence limits) of correlation values of transformed data. Most of the ACF and PACF lags were beyond the confidence limit, it showed that the transformed DNPEPD data was not stationary in mean. Further, ACF of lag 1, 7, 14, 21 and so on, went high, it indicated that there was a seasonal effect in weekly. Therefore first order non seasonal difference (d) was taken to get the mean stationary and seasonal difference (D) was taken to remove the seasonal trend through subtracting the current observation from the previous 7^{th} observation. Finally stabilized data was ready for model selection.

2. Model Selection

After the data stabilization, ACF and PACF were checked again and some tentative models were taken into consideration, which are shown in Table II.

Table III illustrates the parameter estimates, coefficients and corresponding p values of all the selected tentative models, where all the p values are nearly zero (less than 0.05) can be considered as appropriate. Model B contained a p value of 0.407 and model A and model C contained maximum 0.115 and 0.092 respectively, where model A and model C were comparatively better than model B.

Further, Table IV illustrates the Ljung-Box Chi-Square statistics, which tests the overall randomness of the model, where, if all the p values are significant (greater than 0.05) the model can be considered as appropriate. Model A and model C satisfied the condition as all the p values of Chi-Square statistics are significant and model B failed in this test. Therefore at this stage model B has been dropped out from this analysis.

Thereafter, the Akaike information criterion (AIC) (2) was calculated for the selected models, which is a measure of the relative quality of statistical models for a given set of data and the smallest value of AIC will be considered as successful model. Model C contained smallest value of AIC which was

TABLE II: SELECTED TENTATIVE MODELS

Model	Description
A	ARIMA(2, 1, 3)(1, 1, 1)$_7$
B	ARIMA(3, 1, 2)(1, 1, 1)$_7$
C	ARIMA(3, 1, 3)(1, 1, 1)$_7$

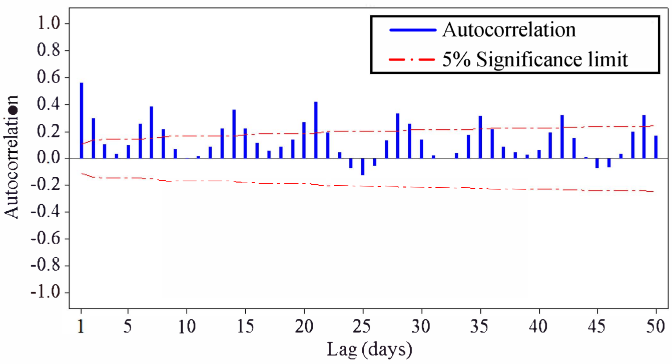

Fig. 4.a. Autocorrelation plot for Box Cox transformed DNPEPD with 5% significance limits for the autocorrelations

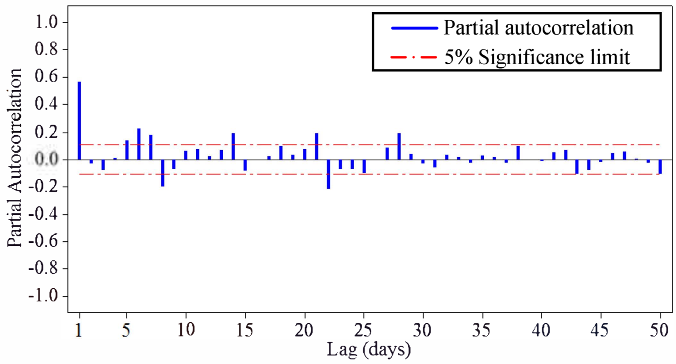

Fig. 4.b. Partial Autocorrelation plot for Box Cox transformed DNPEPD with 5% significance limits for the partial autocorrelations

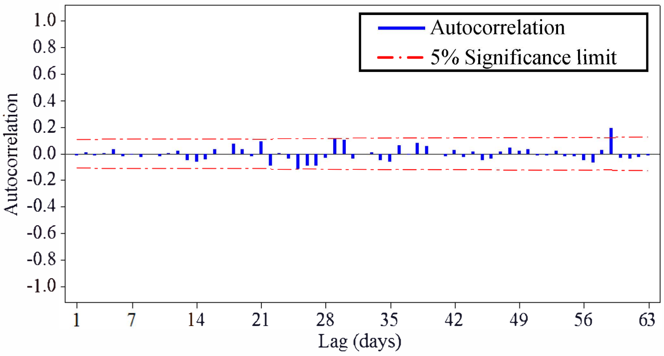

Fig. 5.a ACF of residuals of Box Cox transformed DNPEPD, ARIMA (3, 1, 3)(1, 1, 1)$_7$ model with 5% significance limits for the autocorrelations

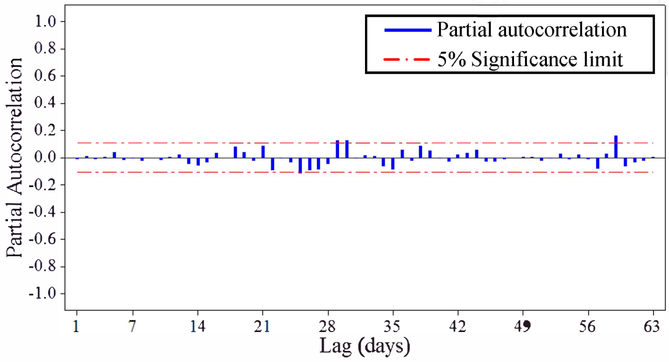

Fig. 5.b. PACF of residuals of Box Cox transformed DNPEPD ARIMA (3, 1, 3) (1, 1, 1)$_7$ model with 5% significance limits for the partial autocorrelations

982.71 and model A contained 983.00. At this stage, it could be clearly seen that, the model C was the successful model among all the selected tentative models.

$$AIC = N \log (1+2\pi) + N \log (\sigma^2) + 2\,(p + q + P + Q) \quad (2)$$

where, N denotes number of observation and σ denotes residual variance.

Then the residuals from the model C was examined for adequacy. Residuals of an adequate model should be uncorrelated as well as white noise, where the residual should follow normal distribution with zero mean and constant variance [7].

Correlation of residuals were identified by testing the ACF and PACF plots (see Fig. 5.a and Fig. 5.b), where the blue lines indicate the correlation values of residuals at each lag and the red lines indicate the 5% significance limits (95% confidence limits) of correlation values of residuals, it showed most of the autocorrelation and partial autocorrelations of the residuals at different lags were within the 95 % confidence limits and only one lag in each ACF and PACF were beyond the limit. It can be ignored as random error, which means there was an unusual happening in DNPEPD in that particular day. Therefore it can be decided that the residual of model C was uncorrelated.

The graphical measure of residuals (see Fig. 6) of model C were checked for adequacy of the model, where the first measure of Fig. 6 was the normal probability plot of the residuals (see Fig.6.a) and the second measure for adequacy of model was the histogram of the residuals (see Fig.6.b), which test the normality of the residuals. The normal probability plot showed that the most of the residuals were approximately on the straight line and histogram showed that the distribution and spread of the residuals was bell-shaped, hence good normality of the residuals.

The third measure was the plot of residuals against fitted values (see fig. 6.c), where the mean and variance of residuals can be checked. Here the residuals scattered randomly around zero, which indicated that the residuals had a mean of zero. The vertical width of the scatter did not appear to increase or decrease across the fitted values, this suggests that the variance of the residual terms was constant. The final measure was the plot of residuals against fitted order of the data (see Fig. 6.d). In this plot the data did not follow any symmetric pattern with the run order value. Almost all of the residuals were within acceptable limits which indicated the adequacy of the recommended model.

TABLE III: PARAMETER ESTIMATES OF SELECTED ARIMA MODELS

Type	**Model A**		**Model B**		**Model C**	
	Coefficient	**P Value**	**Coefficient**	**P Value**	**Coefficient**	**P Value**
AR1	0.8673	0.000	-1.3155	0.000	0.8022	0.000
AR2	-0.3506	0.000	-1.2948	0.000	-0.7404	0.001
AR3	-	-	-0.2906	0.000	0.2422	0.037
SAR7	-0.0981	0.115	-0.0509	0.407	-0.1192	0.092
MA1	1.3343	0.000	-1.0042	0.000	1.2672	0.000
MA2	-0.6853	0.000	-0.9829	0.000	-1.0342	0.000
MA3	0.2913	0.000	-	-	0.6848	0.000
SMA7	0.9436	0.000	0.9559	0.000	0.9584	0.000

TABLE IV: LJUNG-BOX Q TEST OF THE RESIDUALS OF SELECTED ARIMA MODELS

	Model A				**Model B**				**Model C**			
Lag	12	24	36	48	12	24	36	48	12	24	36	48
Chi Square	3.7	15.7	40.6	47.7	22.9	37.4	72.5	79.9	1.2	13.3	36.4	43.5
Degree of Freedom	5	17	29	41	5	17	49	21	4	16	28	40
P Value	0.600	0.545	0.074	0.219	0.000	0.003	0.000	0.000	0.872	0.654	0.133	0.325

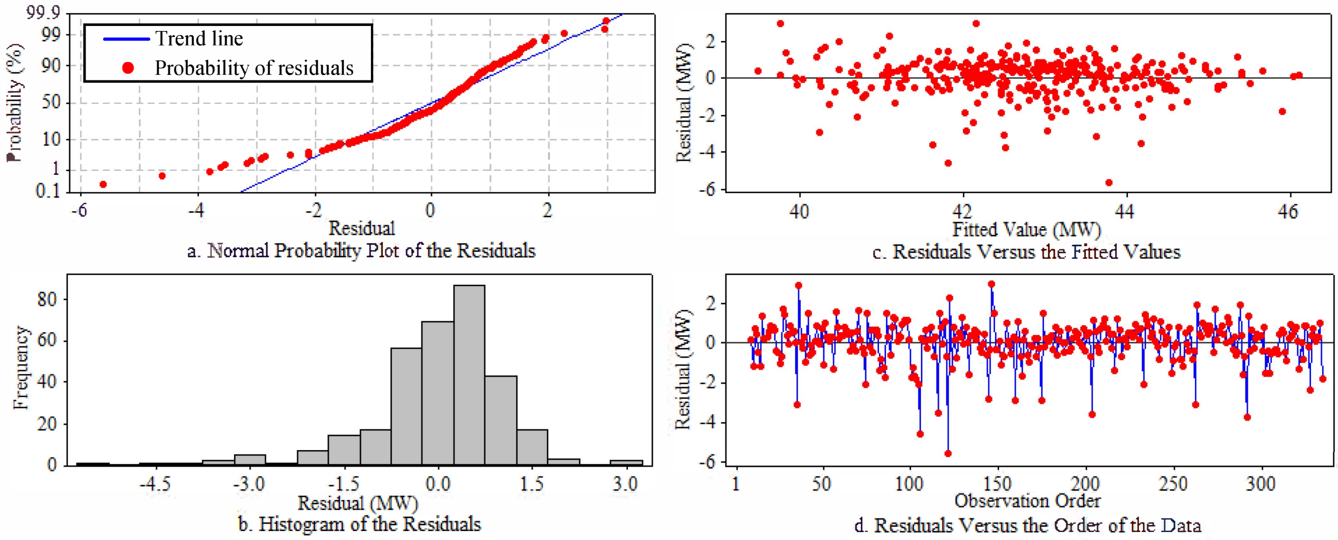

Fig. 6. Residual plots of Box Cox transformed DNPEPD, ARIMA (3, 1, 3) (1, 1, 1)$_7$ model

Finally DNPEPD of December 2013 were forecast from the successful model and results were compared with actual December 2013 data.

III. Results and Analysis

Both, the forecast (black line) and actual (red line) DNPEPD of Sri Lankan power system were plotted on the same axis (see Fig. 7) to indicate the model adequacy, performance and comparison purposes. The similarity and matching between the forecast and actual DNPEPD were appropriate.

The Mean Absolute Percent Error (MAPE) (3) of a week ahead forecast and a month ahead forecast were calculated, 1.855% and 4.195%, respectively. Both MAPE are in the acceptable limit as MAPE of 10% is considered as good [13] [14], however deviation between actual and forecast values of DNPEPD in the month ahead forecast was little higher than in the week ahead forecast, because the cumulative error was increasing with number of forecast days, which means the error in each forecast was added to next forecast [15]. Therefore a week ahead forecast seems to be very good than a month ahead forecast.

$$MAPE = \frac{100}{n}\sum\left(\frac{e_i}{x_i}\right) \quad (3)$$

Where, n denotes number of forecast observations, e_i denotes error in i^{th} forecast value and x_i denotes i^{th} actual value.

Therefore it is suggested that the model C ARIMA (3, 1, 3) (1, 1, 1)$_7$ is the most appropriate model to forecast the DNPEPD a week ahead. Mathematical representation of this model is shown below in (4),

$$(1-0.8022B+0.7404B^2-0.2422B^3)(1+0.1192B)W_t = (1+1.2672B-1.0342B^2+0.6848B^3)(1+0.9584B)Z_t \quad (4)$$

Conclusion

The DNPEPD was studied using the Box-Jenkins (ARIMA) model methodology. Here 334 data from January 2013 to November 2013 were used for analysing and modelling purposes. The performance of the resulting successful model ARIMA (3, 1, 3) (1, 1, 1)$_7$ was evaluated by using the December 2013 data through graphical comparison

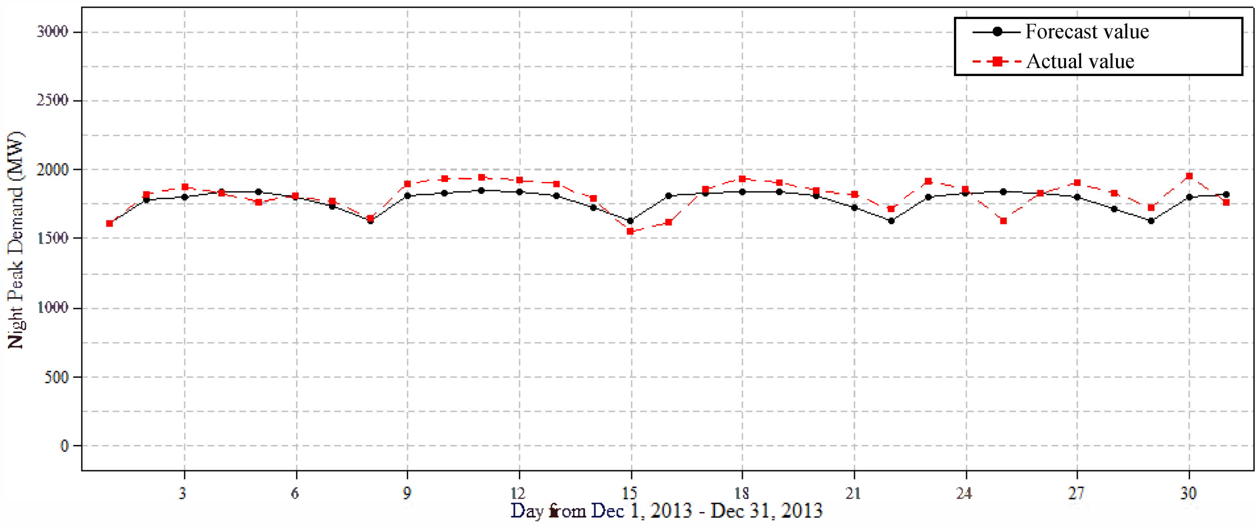

Fig. 7. Forecast and actual value of DNPEPD in December 2013

between the forecast and actual recorded data. The forecast of DNPEPD showed acceptable agreement with the actual recorded data.

Mean Absolute Percent Error (MAPE) for a month ahead DNPEPD forecast is 4.195% and week ahead is 1.855%, both are in the acceptable limit. However a week ahead forecast is more accurate than a month ahead forecast. It indicates that the selected model is the most appropriate for a week ahead DNPEPD forecast.

The study reveals that this methodology could be used as an appropriate tool to forecast the DNPEPD in Sri Lanka. The results achieved from this forecasting will be helpful to estimate the future DNPEPD a week ahead and power system also can reach the best level of power quality and reliability by having prior plan to meet the electric power demand.

FUTURE WORK

Immediate future work of this research study is to further improve the above forecasting model by considering all the proposed loads which is going to be connected to the system in near future and model the day peak and off peak as well[16].

Hydro power generation, thermal power generation, wind power generation and solar power generation are available in Sri Lankan power system, where hydro power generation is the cheapest. In case of sudden failure of a low cost generating unit, it will lead to its substitution by a higher cost generation and also there is another case, if the demand is unexpectedly very high, sometime power producers may not have any other options and they need to import expensive power from another source [1] [2].

In case of above mentioned problems statistically the random variable is the cost of producing electric power, it depends upon two uncertain quantities, such as available generators and the demand. Purpose of this future work is to propose a statistical model for relationship between electric power production cost and the above mentioned uncertain quantities. The major goal of this future work is to planning the cost effective generation pattern for Sri Lankan power system.

ACKNOWLDGMENT

The authors would like to thank Ceylon Electricity Board, for giving all the needful data and also sincere thanks goes to Mr A Laheetharan, senior lecturer in statistics, University of Jaffna for his valuable discussion and advice on this research work.

REFERENCES

[1] Ceylon Electricity Board, "Long Term Transmission Development Plan 2013 – 2022," December 2013.

[2] Ceylon Electricity Board, "Long Term Generation Expnasion Plan 2013 – 2032," December 2013.

[3] Conejo, A.J, J. Contreras, R. Espinola and M. A. Plaza, "Forecasting Electricity for day-ahead pool-based electric energy market," International journal of forecasting,2005, Vol 29, No 13, pp 1145-1154

[4] LO, K.L. and Y.K. WU, "Risk assessment due to local demand forecast uncertainity in the competetive supply industry," IEE proceedings – Generation, Transmission and Distribution, Vol 150, No 5, pp 573-581, 2003.

[5] Taylor, J. W. and S. Majithia, "Using combined forecasts with changing weights for electricity demand profiling," Journal of the operational research soceity, Vol 52, pp. 72-82, 2000.

[6] Taylor, J. W "Short term electricity demand forecastinf using double seasonal exponential smoothing," Journal of the operational research soceity,2000, Vol 54, No 8, pp 799-805.

[7] Spyros Makridakis, Steven C. Wheelwright, Rob J. Hyndman "Forecasting Methods and Application," Copy right 2013.

[8] S. Makridakis and M. Hison, "ARMA Models and the Box Jenkins Methodology," 95/45/TM (Revised version of 95/33/TM).

[9] John M. Gottman, Richard M. Mcfall And Jean T. Barnett "Design And Analysis Of Research Using Time Series," Psychological Bulletin, Vol. 72, No. 4, pp. 299-306, 1969

[10] Eisa Almeshaiei , Hassan Soltan, "A methodology for Electric Power Load Forecasting," Alexandria Engineering Journal, Vol 50, Issue 2, pp. 137–144, June 2011.

[11] Dhaval P. Patel , Mitul M. Patel, Devendra R. Patel, "Implementation of ARIMA model to predict Rain Attenuation for KU-band 12 Ghz Frequency,"IOSR Journal of Electronics and Communication Engineering (IOSR-JECE), e-ISSN: 2278-2834,p- ISSN: 2278-8735.Vol 9, Issue 1, Ver. III, pp. 83-87, January 2014.

[12] Jason W. Osborne, "Improving your data transformations: Applying the Box-Cox transformation," Practical Assessment, Research & Evaluation, ISSN 1531-7714, Vol 15, No 12, October, 2010.

[13] S. Vasantha Kumar and Lelitha Vanajakshi, "Short-term traffic flow prediction using seasonal ARIMA model with limited input data," Eur. Transp. Res. Rev. 7:21, 2015.

[14] Rob J Hyndman and George Athanasopoulos, "Forecasting: principles and practice"," section 2.5, OTexts , Copyright 2015.

[15] Rob J. Hyndman, Anne B. Koehler, "Another look at measures of forecast accuracy,"International Journal of Forecasting 2, pp. 679 – 688, 2006.

[16] G.V.Buddhika De Silva and Lalith A. Samaliarachchi, "Peak Electricity Demand Prediction Model for Sri Lanka Power System," The Institution of Engineers, Sri Lanka, Vol. 46, No. 04, pp. 53-60, 2013.

APPENDIX

Mathematical representation for ARIMA(p, d, q)(P, D, Q)$_n$ is given below,

$$\varphi_p(B)\phi_P(B^n)W_t = \theta_q(B)\Theta_Q(B^n)Z_t$$

Where, $W_t = \Delta^d \Delta_n^D X_t$, difference operator

$\varphi_p(B) = 1-\alpha_1 B-\alpha_2 B^2-...-\alpha_p B^p$,non-seasonal AR operator

$\phi_P(B) = 1-c_1 B-c_2 B^2-...-c_P B^P$, seasonal AR operator

$\theta_q(B) = 1+\beta_1 B+\beta_2 B^2+...+\beta_q B^q$, non-seasonal MA operator

$\Theta_Q(B) = 1+\lambda_1 B+\lambda_1 B^2+...+\lambda_Q B^Q$, seasonal MA operator

$B^n W_t = W_{t-n}$, white noise

$B^n Z_t = Z_{t-n}$, back shift operator

For an example, mathematical representation for ARIMA(3, 1, 3) (1, 1, 1)$_7$ is given below,

$$\begin{aligned} W_t &= \Delta^1\Delta_7^1 X_t \\ &= \Delta_7 X_t - \Delta_7 X_{t-7} \\ &= (X_t - X_{t-7}) - (X_{t-1} - X_{t-8}) \end{aligned}$$

$$(1-\alpha_1 B-\alpha_2 B^2-\alpha_3 B^3)(1-c_1 B)W_t = (1+\beta_1 B+\beta_2 B^2)(1+\lambda_1 B)Z_t$$