Statistical modelling of electricity consumption in Sri Lanka

LS Madushani[1*],KDN Kumari[1]

[1] Department of Mathematics and Philosophy of Engineering, Open university of Sri Lanka, Sri Lanka
[#] samudramadushani@gmail.com

Abstract*—Consumption of electricity is growing each year in Sri Lanka due to industrialization, urbanization, modernization. Forecasting electricity consumption is highly important when deciding on investment and construction. Therefore, this study is focused to develop suitable statistical models for forecasting monthly electricity consumption. The monthly electricity consumption data (in Gigawatt hours) of Sri Lanka were obtained from data library of central bank of Sri Lanka from January of 2010 to June of 2018.Two different approaches namely, by ARIMA approach and Holt-Winters exponential smoothing approach were used to develop two statistical models for forecasting electricity consumption. Model fitting was done by the data from January of 2010 to February of 2018 whereas electricity consumption data from March of 2018 to June of 2018 were used to validate the models. From the results,* $ARIMA\ (0,1,2)\ (0,1,1)_{[12]}$ *and Holt-Winters exponential smoothing model at* $\alpha = 0.41, \beta = 0$ *and* $\gamma = 0.77$ *were identified as optimal models. Since the observed MAPE values of both models are less than 10%, the modelling approaches described herewith can be applied to forecast electricity consumption in Sri Lanka.*

Keywords*—* **ARIMA, Holt-Winters exponential smoothing, electricity consumption, MAPE value**

I. INTRODUCTION

Demand for electricity was growing at a rate of 6% per year with the growth of population. Average daily demand was approximately 40GWh and maximum reported on 26th April 2017, was 44.97GWh. Electricity demand consists of 34% of domestic consumers, 29% from industries and 21% from general purpose consumers with the balance 16% coming from religious organizations, government institutions, hotels and street lighting (Ministry of power and renewable energy, 2018).

During last eight months of 2017, electricity generation in Sri Lanka was shared by 40% of coal, 19% of thermal IPP (Oil), 18% of thermal CEB, 16% of hydro and 7% of other renewables (Ministry of power and renewable energy, 2018). Hence, it was clear that Sri Lankan electricity generation highly depends on coal which is a non-renewable energy source import from other countries.From February to April of 2019, the supply of electricity has disrupted due to a technical failure of Norochcholaipower plant. However, Sri Lanka is not utilized non-renewable energy sources to produce electricity demand.

To identify energy sources for production of electricity, it is essential to forecast the electricity consumption. Several studies have been done regarding electricity consumption of Sri Lanka and other countries.

Ruwanthi and Wickremasinghe (1999) considered multivariate electricity demand model according to influential factors and the study was found that electricity demand on the considered period of Sri Lanka depends on the GDP at constant (1980) factor prices and price of substitutes.

Moreover, Pathberiya and Dias(2013) was developed three models which were the classical model, Winter's exponential smoothing model and the stochastic model for forecasting electricity sales in Colombo city and results indicated that stochastic model was more accurate than other two models. Furthermore,Rathnayake and Senevirathne (2014), Mean absolute deviation (MAD), mean absolute percentage error (MAPE) and mean square error (MSE) were used to compare the prediction accuracy of $ARMA\ (1,$ and $GreyModel\ (1,$ models. They have observed that from the year 1998 to 2005, $Grey\ Model\ (1,$ was more suitable than $ARMA(1$ for forecasting annual electricity production and consumption. Another study was proposed that $ARIMA(3,1,$ and $ARIMA(1,1,$ are suitable and more appropriate for predicting the future demands of electricity in Sri Lanka based on yearly data from 1970 to 2014 (Fernando,Gunawardana and Perera et al, 2018). However that study was not given a comparison regarding actual and forecast values to decide the accuracy of the fitted model.

Yasmeen F, Sharif M (2014) considered linear and non-linear models such as ARIMA, Seasonal ARIMA and ARCH/GRACH model for forecasting electricity consumption in Pakistan. According to the results of MAPE values, it was identified that the $ARIMA(3,1,$ model was the most suitable model for forecasting electricity consumption. In addition, another study was donefor forecasting electricity consumption in China which was shown that $ARIMA\ (1,1,$ model performed high prediction ability(Miao, 2015).

This study focuses on forecasting electricity consumption by Auto-Regressive Moving Average (ARIMA) and exponential smoothing approach. The monthly electricity data were obtained from data library of central bank of Sri Lanka. Here, data are available from January of 2010 to June of 2018. Therefore, the model fitting was done by the data from January of 2010 to February of 2018 and remaining data were used for model validation. The analysis was carried out using R software which is an open source software.

II. METHODOLOGY

The following two approaches are used for forecasting electricity consumption in Sri Lanka.

A. *ARIMA approach*

ARIMA is the combination of Auto Regressive and Moving Average models (Pankratz, 2009). Since ARIMA approach can be applied only for stationary data, firstly, time series plots were used to identify the behavior of the data which indicates whether a trend or seasonality available with data. If the data set contains at least one of the trend or seasonality components, then data were said to be non-stationary.

If so, log transformation and differencing operation were applied to obtain stationarity of data. Furthermore, Augmented Dickey-Fuller (ADF) test was then applied to confirm the stationarity of data. In this test, if the series does not have a unit root, then data can be taken as stationary. Time series models were identified by examining the behavior of auto correlation function (ACF) and partial auto correlation function (PACF). The optimal model was obtained by the lowest Akaike Information Criteria (AIC) value (Robert& David 2017).

ARIMA models can be classified as non-seasonal ARIMA and seasonal ARIMA (SARIMA). ARIMA model consists of three parts which are auto regressive parameter $($, number of differences passes $($ and moving average parameter $($. The general form of seasonal model $SARIMA(p,d,q)(P,D,Q$ is given by,

$$\phi_p(B^s)\varphi(B)\Delta_s{}^D\Delta^d X_t = \Theta_Q(B^s)\theta(B)W_t \qquad (1)$$

where $\{W$= Gaussian white noise process

=period of the time series

$\varphi($= ordinary auto regressive component

$\theta($= ordinary moving average component

$\phi_p(B$=seasonal auto regressive component

$\Theta_Q(B$=seasonal moving average component

= Back Shift Operator

=ordinary component

Δ=seasonal difference component

Diagnostic checks are performed on residuals of the fitted models for checking the assumptions that residuals are randomly and normally distributed. Ljung-Box Q test was used for that. Then accuracy of forecast error of fitted model was measured by Mean Absolute Percentage Error (MAPE).

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{e_t}{y_t}\right| \qquad (2)$$

where y_t actual value

$e_t = y_t - f_t$

f_t Fitted value

n Number of observations

Lewis (1982) was developed a criteria to measure the accuracy of the model as follows (Table 1).

Table 1. Deciding the most suitable model according to MAPE values

MAPE value range	Evaluation
MAPE 10%	High accuracy forecasting
10% < MAPE 20%	Good forecasting
20% < MAPE 50%	Reasonable forecasting
MAPE 50%	Inaccurate forecasting

B. *Exponential smoothing approach*

Exponential smoothing is a time series forecasting method. Holt-Winters method is one of exponential smoothing method use for time series forecasting. This method uses a prediction that is a weighted linear sum of past observations. Furthermore, this method is used to model three aspects of the time series which are typical average (, trend (over time and seasonality ((Holt, 1957).

III. Results and Discussion

A. Preliminary analysis

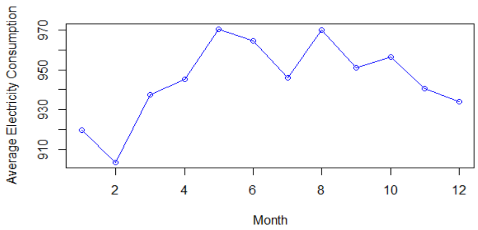

Figure 1. Average monthly consumption variations

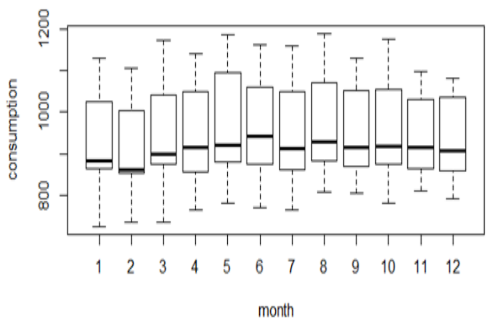

Figure 2. Boxplots for monthly electricity consumption

According to above figure 1 and 2, it was clear that average electricity consumption from April to November (excluding July) was higher than rest months. The reason for this may be the fact that using higher amount of electricity for air conditioners and fans during these warm periods. Electricity consumption of March was in wide range and electricity consumption of November and December were in small range when compares with other months.

B. ARIMA approach

Electricity consumption data consist 102 total monthly observations ranging from January of 2010 to June of 2018.

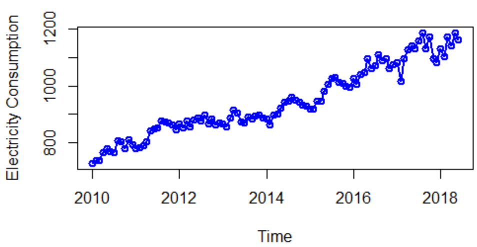

Figure 3. Monthly electricity consumption (January 2010-June 2018)

Since this plot shows both trend and seasonal component, it is clear that data are non-stationary. The Augmented Dickey-Fuller Test $(p\ value = 0.1553 > 0.0$ also confirms the non-stationality. Then data was made stationary by taking first order difference $(d =$ to remove trend component and lag transformation was used to create a constant seasonal effect. Then again Augmented Dickey-Fuller Test was applied to modified data. In that case $p\ value = 0.$ is less than the significant level of 5. Therefore, it is apparent that differenced and transformed data were stationary at 5% level of significance.

In the next stage, ACF and PACF were drawn for the modified data. Appropriate time series models were identified by ACF and PACF graphs. Then the best fitted ARIMA model was selected according to the lowest AIC value. Table 2 represents the parameters of the best ARIMA model with AIC value.

Table 2. The parameters of ARIMA (0, 1, 2) (0, 1, 1) [12]

	MA1	MA2	SMA1
Coefficients	-0.3805	-0.2003	-0.5158
Standard Error	0.1133	0.1153	0.1212
sigma^2 estimated as 402.6, log likelihood=-376.01, AIC=760.5			

Above result was confirmed by the "auto.arima" function in R statistical software (Hyndman &Khandakar, 2008).Furthermore, model assumptions are checked by considering residuals. The $p-values$ for the Ljung-Box Q test all are above 0.05, indicating "non-significance" or residuals are uncorrelated and the normality is checked by the following graph.

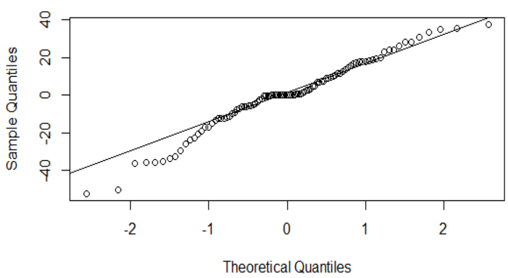

Figure 4. Normality Plot

Since the values lie well on the line, residuals are normally distributed.

C. Exponential smoothing approach

Inthe analysis of part B, it was identified that the electricity consumption in Sri Lanka contains long term trend and seasonal variations. Hence, Holt Winters method was used to develop an appropriate exponential smoothing model. The most appropriate Holt-Winters method for this study was observed as α=0.41, β=0 and γ=0.77 with the lowest MAPE value.

D. Comparison of models and forecasts for 4 months

Using both $ARIMA(0,1,2)(0,1,1)_{[}$ and Holt-Winters exponential smoothing model at $\alpha = 0.41, \beta =$ and $\gamma = 0.$, electricity consumptions were forecasted for the period from March of 2018 to June of 2018. Since the MAPE values are almost similar in both approaches, any of these models can be applied for forecasting purposes. Results were shown in the Table 3.

Table 3. Comparison of actual and forecasted values

		Forecasting	
Month	**Actual Value**	**ARIMA**	**Exponential smoothing**
March	1172	1152.821	1162.779
April	1141	1167.951	1176.447
May	1186	1193.498	1195.597
June	1163	1183.761	1182.878
MAPE value		1.604	1.603

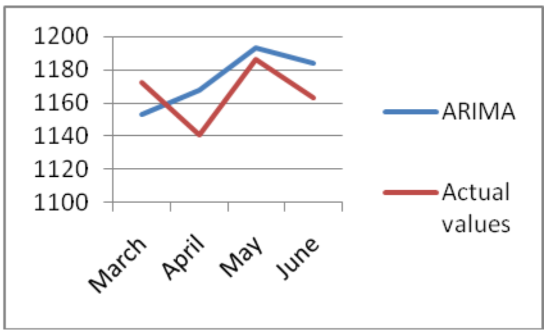

Figure 5. Comparison of fitted ARIMA model and actual time series

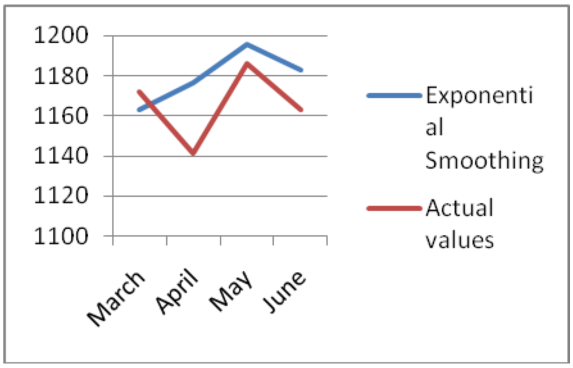

Figure 6. Comparison of exponential smoothing series and actual time series

IV CONCLUSION

This study was aimed to forecast monthly electricity consumption in Sri Lanka.

According to the monthly consumption analysis, average electricity consumption has shown a considerable increase between April and November except July due to extreme weather conditions.

From the plot of electricity consumption data, it was identified that electricity consumption in Sri Lanka contains a long-term trend and seasonal variations.Thus, ARIMA and Holt-Winters exponential smoothing approaches were used for modeling and forecasting electricity consumption. Data from January of 2010 to February of 2018 were used to build the$ARIMA(0,1,2)(0,1,1)_{[12]}$by converting the data into stationary data and Holt-Winters exponential smoothing model at $\alpha = 0.41, \beta = 0$and$\gamma = 0.77$. The data from March of 2018 to June of 2018 were used to forecast the electricity consumption.

According to the forecasting results, one can see that both models are accurate for forecasting electricity consumption in Sri Lanka.

Actually, electricity consumption forecasting can be used for the country's development progress by strategically planning the consumption and production of electricity.Coal was the main source of energy generation methodology of Sri Lanka in 2017. However, it is not possible to supply the growing demand along with the coal source in the near future. Therefore, the special attention is required to the renewable sources such as hydro, wind, solar and biomass sources etc.

REFERENCES

Fernando JLLS, Gunawardana JRNA and Perera KAIT, et al, "Time series modeling approach for forecasting electricity demand in Sri Lanka", 2017 General Sir John KotelawalaDefence University conference, Sri Lanka, 2018.

Holt CC (1957) Forecasting trend and seasonal by exponentially weighted average, *Carnegie Institute of Technology*.

Hyndman RJ &KhandakarY (2008), Automatic Time Series Forecasting: The forecast package for R, *Journal of Statistical Software*, vol.27, no. 3.

Lewis C (1982)*International and business Forecasting Methods*, Butterworths, London.

Miao J," The energy consumption forecasting in China based on ARIMA model", International Conference on Materials Engineering and Information Technology Applications (MEITA), 2015.

Ministry of power and renewable energy. *Performance 2017 and Programs for 2018*,2018.

Pankratz A (2009*) Forecasting with univariate Box-Jenkins models: Concepts and cases*, Vol. 224, John Wiley & Sons.

Pathberiya HA & Dias P (2013) Modeling electricity sales in Sri Lanka and Colombo city using different statistical methods, *Journal of National Science Foundation of Sri Lanka*, 41(1), 41-51.

Rathnayaka RMKT &Senevirathna DMKN (2014) GM (1, 1) Analysis and Forecasting for Efficient Energy Production and consumption, *International Journal of Business, Economics and Management works*,1, 6-11.

Robert HS & David SS (2017) *Time series analysis and its applications: with R examples,* 4thedition, New York, USA.

Ruwanthi KDR &Wickramasinghe (1999) Modelling sector-wise demand for electricity in Sri Lanka using a multivariate regression approach, *Journal of National Science Foundation of Sri Lanka*, 27(1), 55-64.

Yasmeen F & Sharif M (2014) Forecasting electricity consumption for Pakistan, *International Journal of Emerging Technology and Advanced Engineering*,4.