International Journal of Science, Environment and Technology, Vol. 6, No 6, 2017, 3240 – 3251

ISSN 2278-3687 (O)
2277-663X (P)

FORECASTING AND COMPARING THE CONSUMPTION OF ELECTRICITY IN MAURITIUS WITH ARTIFICIAL NEURAL NETWORK AND AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODEL

Jayrani Cheeneebash*[1], Ashvin Gopaul and Davish Seetanah

[1]Department of Mathematics, Faculty of Science, University of Mauritius, Reduit 80837, Mauritius

E-mail: jayrani@uom.ac.mu (**Corresponding Author*)

Abstract: Electricity plays a vital role in the economy of a country and it is one of the most important and influential energy sources that exist nowadays. The demand in electricity has experienced a dramatic increase in Mauritius for the past decades due to the rapid changes in the domestic economy. Forecasting electricity is not an easy task since there is an unstable trend in the energy consumption where demand varies differently in summer and winter. In this work, we study a linear method namely the Autoregressive Integrated Moving Average and a nonlinear approach, in particular the Artificial Neural Network as forecasting tools. We examine the demand in electricity on a daily basis in Mauritius for the period of Oct 2007 to Sept 2014 where the first 70% dataset is used for training while the remaining 30% data points have been used for testing and validation. Concerning the linear method, after studying the ACF, PACF and the Ljung-Box test we fit and compare different values of p and q in a limited range of 0 to 3 in the seasonal ARIMA model(p,d,q)(P,D,Q). The best results were recorded when seasonal ARIMA (0,1,1) with seasonality 365 model was used. However, when it was compared to the ANN method, the latter proved to be more accurate in terms of least mean square error.

Keywords: Energy, Environment, Climate Change, Sustainable development.

1. INTRODUCTION

The demand of the world for electricity is expected to increase by more than half of its actual consumption in 2040 from 524 quadrillion British thermal units (Btu) to 820 quadrillion Btu [10] and that of the Republic of Mauritius is certainly no exception. Mauritius is an island nation situated off the southeast coast of Africa in the southwest Indian Ocean, about 870 km east of Madagascar. It has a population of about 1.3 million with an annual growth of 0.5% and GDP per capita of 8654 USD. It has total area of 2040km^2 and its capital is Port-Louis. The Central Electricity Board (CEB) is the sole supplier of electricity in the country and it is forecasted that the demand of the country will top around 1800Kwh by 2022 [11]. The increase in the population size and in the number of number of tourists arrival is also putting more pressure on the consumption of electricity and any failure on the supply side may result

*Received Sep 29, 2017 * Published Dec 2, 2017 * www.ijset.net*

in black outs. In fact, prolonged power shortage have many negative impacts on society such as ceasing of modern communication and banking services, economic breakdown, closure of gas stations and transportation systems, interruption in food distribution, closure of hospitals and finally social chaos. Thus it is very important to develop accurate forecasting models for electricity demand so that proper planning can be done for the electricity supply from the different power stations in the country.

In this paper, we consider the demand for electricity as a time series and study two methods for forecasting future demands. We consider a popular linear method known as the Autoregressive Integrated Moving Average (ARIMA) and another nonlinear method which is the Artificial Neural Network (ANN). The ARIMA Models was developed by Box and Jenkins in 1976 and brings a specific perspective when it comes to time series forecasting. The model can forecast a certain time series observation with high accuracy [5]. The linear model exists in several forms, namely, ARMA, ARIMA, Seasonal ARIMA models and so on. Each one of these methods consists of their own function and it brings a different approach to the model according to the observation. The models are suitable only for an appropriate trend between the past and the future. In other words, it is suitable to forecast short term basis in the future.

Nonlinear models have contributed a lot in the field of forecasting for the past years. A survey on the different applications of ANN has been carried out by Zhang et. al. [9]. ANN have also been used for weather forecasting [4] and stock price prediction [8]. This technique has also been used for forecasting the consumption of electricity as mentioned in [3,7].

The information processing of the artificial neural network works practically the same way as the human neural system. This model has a high capacity to extract information from complicated data set and identify trends that are highly complex. Artificial neural network is widely used in forecasting areas as it yields an interesting and appealing alternative tool for researchers.

2. FORECASTING METHODS

A time series process equivalent to a stochastic process is a series of data that is measured sequentially over time. The key purpose of these series is to formulate mathematical representations whose aim is to aid us to describe the fluctuation of the data over a time *t*. These erratic data are recorded as

$$\{a_t : t = 1,2,3,\ldots\} \quad (1)$$

2.1 **ARIMA**

The autoregressive moving average (ARMA) is basically used to understand a time series and predict its future values.

A time series given by Eq.(1) is ARMA (*p*, *q*) if it is stationary and the components AR and MA are given by

(AR): $$a_t = c + \sum_{i=1}^{p} \varphi_i a_{t-i} + w_t \ , \qquad (2)$$

where *c is a constant,* φ_i are the AR parameter and w_t is the white noise and

(MA): $$a_t = \mu + w_t + \sum_{j=1}^{q} \theta_j^* w_{t-j} \, , \qquad (3)$$

where μ is the expectation, w_t is the white noise, θ_i^* are the parameters and $w_t \sim W_N(0, \sigma^2)$.

We then have ARMA as

(ARMA): $$a_t = c + w_t + \sum_{i=1}^{p} \varphi_i a_{t-i} + \sum_{j=1}^{q} \theta_i^* w_{t-j} \ , \qquad (4)$$

where $\varphi_p \neq 0$, $\theta_q^* \neq 0$ and $\sigma^2 > 0$.

ARIMA [1] can be defined as being a composition of autoregressive model, differencing process and moving average process. It is used in cases where differencing is required to make a time series stationary. We consider the parameters *p* (autoregressive parameter), *q* (moving average parameter) and *d* (differencing). The general form of ARIMA is *ARIMA (p, d, q).*

To fit a time series to the ARIMA structure, Box and Jenkins recommended that data must first be identified in terms of the model to be used and based on the results of the stationary test. Then in the estimation phase, the model is specified and the parameters are determined. Finally, the forecasting step is done and it can then be tested for accuracy.

For seasonal series we use the seasonal autoregressive integrated moving average (SARIMA). The general form is denoted by *SARIMA(p,d,q)*(P,D,Q)*, where, *p*, *d*, and *q* are the AR, differencing and MA order of non-seasonal differencing and *P*, *D*, and *Q* are the order of seasonal component [2]

2.2 ARTIFICIAL NEURAL NETWORKS

The artificial neural network (ANN) is substantially described as a computational model inspired from the elementary neural system of the animal's brain. Mc Culloch and Pitts [6] developed the concept of a structure consisting of a set of layers of interconnecting neurons which are coupled together. In this dynamic method, each link has its own weight. These weights can be described as data carrier for all the processing that is done at each node. The ANN structure comprises of 3 distinct layers as shown in Fig.1 whereby each layer is designed to carry out a specific task. These layers are the input, hidden and output layer. The input layer, as its name indicates is meant for users to input their data in the network. The responsibility of the hidden layer is to execute all calculations and processing needed whereas the output layer only displays the final results.

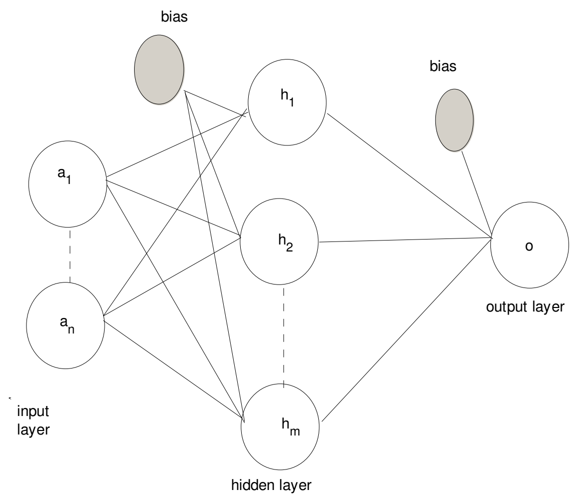

Fig 1 Structure a feed forward artificial neural network

From Fig.1, it can be clearly pointed out that it has *n* input nodes, *m* processing or hidden nodes which have as target to produce only one output in the output layer.

Let $A = (a_1, a_2, \ldots, a_n)^T$ and denote the input vector and the weight to be

$$W = \begin{bmatrix} w_{1,1} & \cdots & w_{1,n} \\ \vdots & \ddots & \vdots \\ w_{n,1} & \cdots & w_{n,n} \end{bmatrix} \tag{5}$$

We denote the total input by net. The latter corresponds to the sum of each input multiplied by their weight.

$$net = \sum_{i=1}^{n} a_i w_{i,j} \tag{6}$$

The neural network consists of a bias node having value *1* which is also included in the total input. Its aim is to shift the function to the left.

$$net = a_0 + \sum_{i=1}^{n} a_i w_{i,j} \tag{7}$$

Now, let the hidden layer vector to be denoted by $H = (h_1, h_2, \ldots, h_n)^T$. Each entry is obtained by using the activation function *G*, as follows:

$$h_j = G\left(b + \sum_{i=1}^{n} a_i w_{i,j}\right) \tag{8}$$

where *j=1,2,3,…,m*.

The calculated data in the hidden layer becomes the input of another function so as to get the output given as,

$$O = \sigma\left(a_0 + \sum_{j=1}^{m} z_j h_j\right), \tag{9}$$

where a_0 is the bias neuron and z_j is the weight between node *j* and the output node.

Finally, replacing Eq. (8) into (9), we get

$$O = \sigma\left(a_0 + \sum_{j=1}^{m} z_j G\left(b + \sum_{i=1}^{n} a_i w_{i,j}\right)\right). \tag{10}$$

Another characteristic of the feed forward neural network is that each network may consist of a different activation function.

3. METHODOLOGY

In this section, we examine the demand in electricity, calculated in Kilowatts per hour, on a daily basis in Mauritius for the period of Oct 2007 to Sept 2014 which is represented in Fig. 2. Since we have a dataset consisting of very large values, we use the following normalization formula

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}}.$$

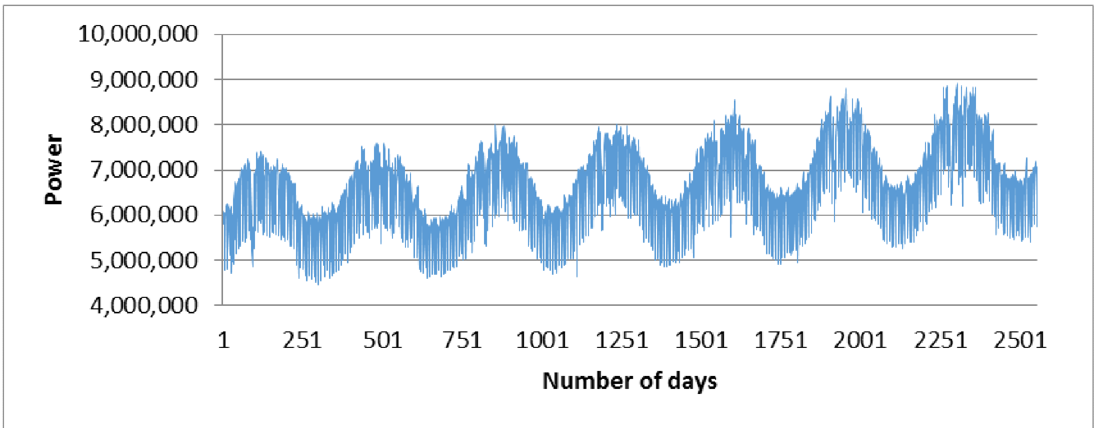

Fig 2 Electricity demand in Mauritius

3.1 FITTING THE ANN MODEL

We use the Neural Network Time Series Tool (*ntstool*) in Matlab to solve a nonlinear time series problem with a dynamic neural network. It takes the past values of a time series data set to predict future values thus enabling us to work and to solve the nonlinear time series problems.

The next step is to input the target and the desired output time series data, and then the *ntstool* splits the target values into three types of target time steps. They are known as training, validation and testing. The known output target values are assigned to a certain percentage of the target time step and each of them does a specific process. In the training step 70% of the target time steps are introduced to the network when training process is effectuated and the network is altered according to its error. The validation and training process gives us the choice to modify the percentage of the time steps from 5% to 35% with a lag of 5% in between but by giving an equal amount of target time steps to each one of them is proven to be the most effective solution. Validation uses the target time steps to measure the network generalization, and to stop the training process as soon as the generalization stops improving. The network performances are measured independently during and after the training and do not affect the training process.

Now for the network architecture, a hidden layer is used in the feedforward neural network. The number of neurons in the hidden layer can be adjusted in a small range. The number of neurons used varies from 1 to 10 as well as for the time delay. The activation function used in the hidden layer is the tansigmoid function and a linear function is used for the output layer. The network will train in a loop process as shown above. When training the open loop (single step) prediction is more effective than the closed loop (multi-step) prediction because it permits us to provide a good feedback inputs to the network and good feedback outputs as well.

3.2 FITTING THE SARIMA MODEL

From a visual inspection of the data for electricity demand in Fig.2, we find that there exists a strong seasonal pattern. Therefore we proceed in testing for seasonality as described below. The seasonal and non-seasonal part of an ARIMA model constitute of an identical structure and is denoted as ARIMA $(p,d,q)\times(P,D,Q)$. The differencing is the difference from one season to the previous one and the equation is given by:

$$Y = Y_t - Y_{t-s}, \tag{11}$$

where Y is the difference time series for a period t and s is the seasonality. The seasonal differencing time series will help to keep the seasonal pattern for a long period.

After studying the autocorrelation function test (ACF), the partial autocorrelation test (PACF) and the Ljung-Box test to diagnose the most appropriate model for the forecasting process of the Seasonal ARIMA model, we fit and compare different values of p and q in a limited range of 0 to 3.

The SARIMA model is composed of three stages, which is the identification stage, the estimation and diagnostic checking stage and the forecasting stage. In the identification stage, we identify clearly the feedback of the time series and establish the appropriate SARIMA model. The time series data is examined and we compute the differencing process and use the ACF and the PACF plot to find the corresponding values of the polynomials. This process requires one or more computations in order to obtain the best result. To find out if the differencing process is needed or not, the stationary test is carried out. The estimation and the diagnostic checking stage, we consider the SARIMA model that is going to fit to the variable obtained in the identification stage, as well as to choose the appropriate parameters to the model. For the forecasting stage, we use the parameters obtained by the previous stages to forecast the future values of the observations as well as to simulate the forecasted confidence interval from the SARIMA model.

4. RESULTS AND DISCUSSION

The artificial neural network with different parameters was applied to the data set as explained in the above section. The daily demand of electricity in Mauritius for the period of Oct 2007 to June 2014 is trained to forecast the demand for the following three months that is July, August and September 2014. Different runs were done with the attempt of obtaining the least mean squared error (MSE). The latter is calculated as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n} \hat{Y}_i - Y_i \,,$$

where $\hat{Y}_i$ is the predicted value of Y_i. The optimal solution was obtained by assigning the number of delays to be 2 and the simulation was then performed by varying the number of hidden neurons as shown in Table 1. The best result is obtained when putting the number of hidden neurons equal to 2 with values for MSE and regression as 0.0000482027 and 0.96611 respectively. Having found the best model, we have applied it on the data from July 2014 to Sept 2014 and have compared the forecasted values with the actual values. Fig. 6 shows that the forecasted and actual values have a strong correlation.

Now we shall implement the SARIMA method on the same dataset described earlier so as to compare its accuracy with that of ANN. First we apply first and second differencing on the data.

We observe a random fluctuation in both the first and second differencing, and they seem to be stationary as shown in Fig. 3 and Fig. 4. It can be seen that the daily demand of electricity in the first differencing gives a better approach to the mean than in the second differencing.

According to the sample autocorrelation function plot of the difference series above, we observe that there are significant autocorrelation at a gap of 7 lags and at smaller lags as well. Both the sample autocorrelation and the sample partial autocorrelation function are considered to be effective to figure out the autocorrelation at single lags. There exists a more quantitative method to effectuate the test for autocorrelation at numerous lags; conduct the Ljung-Box Q-test. When conducting the Ljung-Box test, we find out that the null hypothesis is rejected which brings us to the conclusion that the significant autocorrelation is found in the series.

Table 1 Accuracy of ANN method for delay = 2.

Hidden Neurons	MSE	Regression
1	0.000225784	0.9422
2	0.0000482027	0.96611
3	0.000170123	0.963804
4	8.49975E-05	0.965716
5	0.000262697	0.913444
6	0.000205353	0.911585
7	0.000135448	0.953528
8	0.000235209	0.926541
9	0.000119818	0.924443
10	9.60862E-05	0.95313

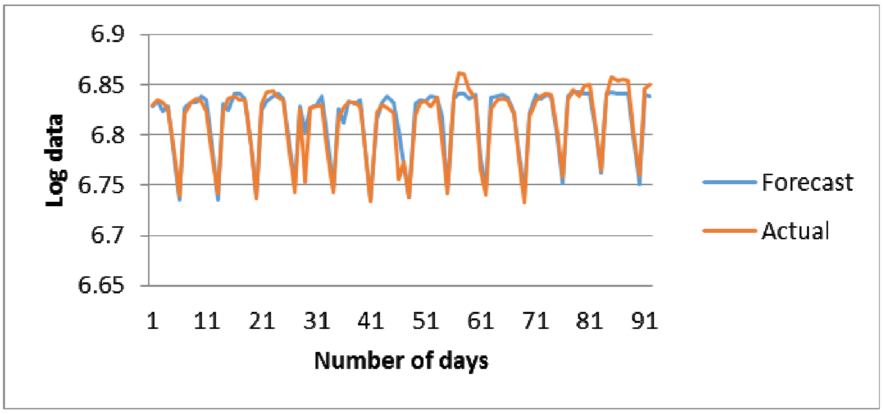

Fig.6 ANN forecasted values and the actual values

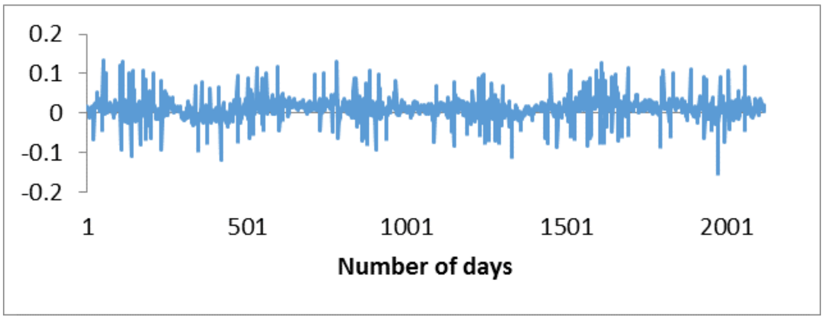

Fig. 7 First differencing of the demand of electricity

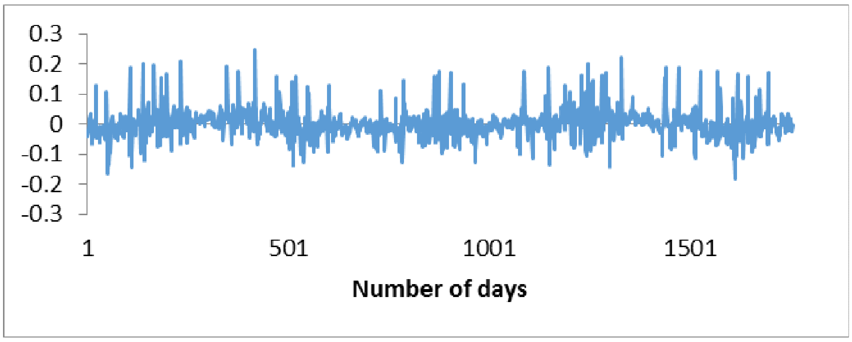

Fig. 8 Second Differencing of the demand of electricity

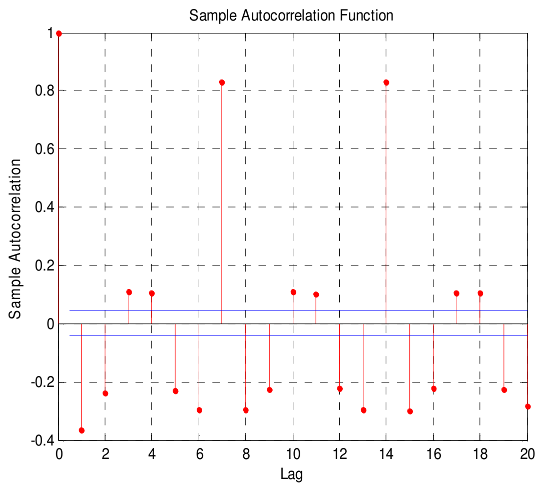

Fig. 9 Sample Autocorrelation of the daily demand of electricity

We next test different values for the parameters on the seasonal ARIMA model in order to determine the best combinations. Table 2 shows the different models and we find out that there are 3 optimal solution, (0,1,1), (1,1,1) and (3,1,1) with yearly seasonality s=365 for Seasonal ARIMA model.

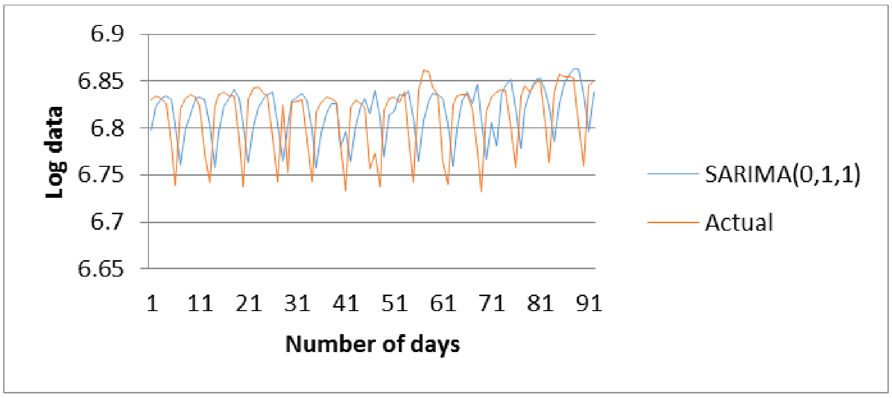

Fig. 10 Comparison of forecasted values by seasonal ARIMA with the actual values

From Table 2, we find that the mean square errors are relatively low for almost every SARIMA models. The order of polynomials for the Autoregressive (AR) model and the Moving Average (MA) model are varied from 0 to 3. In Table 6 it is clearly observed that some of the errors are alike and it is the case for the best performances as well. The Seasonal ARIMA model (0,1,1), (1,1,1) and (3,1,1) have the best performances recorded. To distinguish the best one among these three models, the actual values and the forecasted values of these three models are plotted below.

Table 2 Seasonal ARIMA Performance of the daily demand of electricity

Model	RMSE	MSE
(0,1,0)	0.0092	0.00008464
(1,1,1)	0.0082	0.00006724
(1,1,0)	0.0094	0.00008836
(0,1,1)	0.0082	0.00006724
(2,1,2)	0.0101	0.00010201
(2,1,1)	0.0083	0.00006889
(2,1,0)	0.0106	0.00011236
(0,1,2)	0.0092	0.00008464
(1,1,2)	0.0083	0.00006889
(3,1,3)	0.0088	0.00007744
(1,1,3)	0.0099	0.00009801
(2,1,3)	0.0098	0.00009604
(0,1,3)	0.0094	0.00008836
(3,1,0)	0.0093	0.00008649
(3,1,1)	0.0082	0.00006724
(3,1,2)	0.01	0.0001

Figure 11 shows that the Seasonal ARIMA (0,1,1) with seasonality 365 model gives a better approach to the actual values compared to the other two models. Although, we register very low RMSE errors for the Seasonal ARIMA models, ANN has produced a superior performance as it fits almost perfectly to the actual values as shown in Fig. 7.

When we compare the results in Table 1 and Table 2, we find that even though the best performance was obtained with the ANN model, it is noted that 90 % of the MSE errors acquired by the seasonal ARIMA model were significantly low compared to ANN. ANN will however perform better most of the time due to its ability to capture the nonlinear trend of the data.

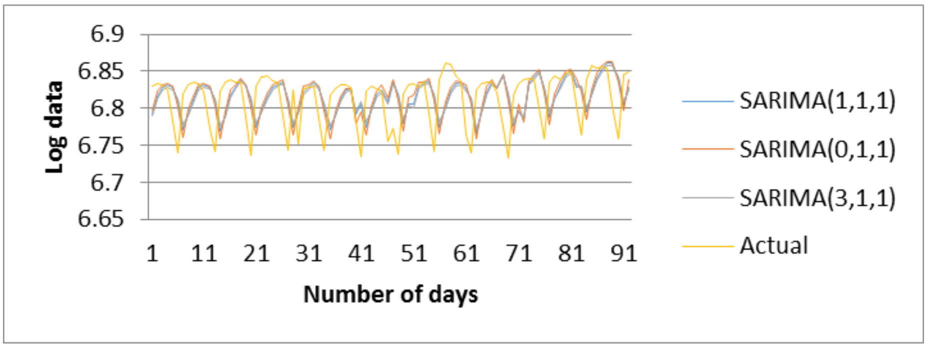

Fig. 11 Comparison of the 3 best Seasonal ARIMA Model

Conclusions

In this paper we have compared two forecasting methods based on artificial neural networks and the autoregressive integrated moving average respectively. Both methods were tested extensively so as to determine the optimal parameters which give the least mean square error. We note that for the prediction of the demand of electricity in Mauritius, it is more appropriate to use the artificial neural networks tool.

References

[1] Box, G. and Jenkins G., 1970. Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day.

[2] Chang X., Gao M., Wang Y. and Hou X., 2012 Seasonal autoregressive integrated moving, Average model for precipitation time series, Journal of Mathematics and Statistics 8 (4): 500-505. ISSN 1549-3644

[3] Kandananond K., 2011. Forecasting Electricity Demand in Thailand with an Artificial Neural Network Approach, Energies, 4, 1246-1257; doi:10.3390/en4081246.

[4] Kumar N., Jha G.K., 2013, A Time Series ANN Approach for Weather Forecasting, International Journal of Control Theory and Computer Modeling 3(1)

[5] Makridakis, S. and Hison, M., 1995 —ARMA MODELS and The Box Jenkins methodology by revised version of 95/33/TM.

[6] McCulloch W.S. and Pitts W. (1943) A logical calculus of the ideas immanent in nervous activity, Bulletin of Mathematical Biophysics, 5, 115-133.

[7] Nasr G.E., Badr E.A. and Younes M.R., 2001. Neural Networks in Forecasting Electrical Energy Consumption, FLAIRS-01 Proceedings American Association for Artificial Intelligence.

[8] Niaki, S.T.A. & Hoseinzade S., 2013. Forecasting S&P 500 index using artificial neural networks and design of experiments, J. Ind Eng Int (2013) 9: 1. doi:10.1186/2251-712X-9-1

[9] Zhang G., Patuwo B. E., Hu M.Y., 1998. Forecasting with artificial neural networks:: The state of the art, International Journal of Forecasting, 14(1), 35–62.

[10] Online1 https://www.eia.gov/todayinenergy/detail.php?id=12251

[11] Online2
http://ceb.intnet.mu/CorporateInfo/IEP2013/Chapter4_Demand%20Forecast%20for%20
Mauritius.pdf