International Journal of Business and Management; Vol. 11, No. 5; 2016
ISSN 1833-3850 E-ISSN 1833-8119
Published by Canadian Center of Science and Education

Energy Consumption Forecasting Using Seasonal ARIMA with Artificial Neural Networks Models

Abdoulaye Camara[1], Wang Feixing[1] & Liu Xiuqin[1]

[1] Department of Information and Computing Science, School of Mathematics and Physics, University of Science and Technology Beijing, Beijing, China

Correspondence: Abdoulaye Camara, Department of Information and Computing Science, School of Mathematics and Physics, University of Science and Technology Beijing, Beijing, China. E-mail: ablozi@yahoo.fr

Received: February 26, 2016 Accepted: March 7, 2016 Online Published: April 18, 2016

doi:10.5539/ijbm.v11n5p231 URL: http://dx.doi.org/10.5539/ijbm.v11n5p231

Abstract

In many areas such as financial, energy, economics, the historical data are non-stationary and contain trend and seasonal variations. The goal is to forecast the energy consumption in U.S. using two approaches, namely the statistical approach (SARIMA) and Neural Networks approach (ANN), and compare them in order to find the best model for forecasting. The energy area has an important role in the development of countries, thus, consumption planning of energy must be made accurately, despite they are governed by other factors such that population, gross domestic product (GDP), weather vagaries, storage capacity etc. This paper examines the forecasting performance for the residential energy consumption data of United States between SARIMA and ANN methodologies. The multi-layer perceptron (MLP) architecture is used in the artificial neural networks methodology. According to the obtained results, we conclude that the neural network model has slight superiority over SARIMA model and those models are not directional.

Keywords: forecasting, seasonality, time series, artificial neural networks, feed forward neural network, arima, multilayer perceptron (mlp) architecture, energy consumption

1. Introduction

Many researches were done on energy predictions with different solution techniques in recent years. There exist some approaches such as statistical approach, neural networks approach. Statistical approaches include many techniques such as autoregressive integrated and moving average (ARIMA), linear regression e.g. (Nowicka-Zagrajek & Weron, 2002; Mucuk & Uysal, 2009; Kandananond, 2011). The Autoregressive Integrated Moving Average models (ARIMA) models were developed by George Box and Gwilym Jenkins in the early 1970s. It is known as the Box-Jenkins methodology is generally used in time series analysis and forecasting. It is widely recognized as the most appropriate forecasting technique in many areas and is used extensively for time series e.g. (Mucuk & Uysal, 2009; Kandananond, 2011; Junttila, 2001; Adebiy, Adewumi, & Ayo, 2014). The use of ARIMA models for time series forecasting do not necessitate the use of the independent variables in the models construction. They rely on past information in the series itself to make forecasts. In the ARIMA models, the autocorrelation functions of the data are indispensables in order to carry out the use of the models (Mucuk & Uysal, 2009; Kandananond, 2011; Shumway & Stoffer, 2006). The time series forecasting using ARIMA methodology is different from most methods, it does not assume knowledge of any underlying model or relationships and any particular pattern of the series data to be predict. Hence, ARIMA models are more robust and efficient than some complex models.

An interactive approach was used to identify a possible ARIMA model from a general class of models. Diagnostic tests of the model are then carried out. If the model is accepted, then the model is investigated to forecast the future values of the data. The ARIMA model appears as a stochastic differential equation that is frequently utilized to model stochastic processes (Mucuk & Uysal, 2009; Kandananond, 2011). It can be used to represent the stationary process or non-stationary and also used to represent the seasonal process, namely SARIMA (Mucuk & Uysal, 2009; Kandananond, 2011; Junttila, 2001; Md Maarof, Zuhaimy, & Fadzli, 2014; Suhartono, & Guritno, 2005). In fact a lot of works are done using ARIMA models in the forecasting procedure, mainly in the energy area e.g. (Nowicka-Zagrajek & Weron, 2002; Mucuk & Uysal, 2009; Kandananond, 2011;

Junttila, 2001; Suganthi & Samuel, 2012; Juan, Graff, & Rodriguez, 2012).economics/financial area e.g. (Singh & Mishra, 2015; Junttila, 2001; Adebiy, Adewumi, & Ayo, 2014).

Artificial neural networks (ANNs) is a technique of machine learning approaches, they are more used in some areas such as economic, business, finance, energy, hydrologic as forecasting models (Kandananond, 2011; Singh & Mishra, 2015; Suganthi, & Samuel, 2012; Richard, Anthony, & Wanjoya Anthony, 2014; Oludolapo, Jimoh, Kholopane, 2012; Bodyanskiy & Popov, 2006; Adebiy, Adewumi, & Ayo, 2014; Jain, & Kumar, 2007). Thus, they appear as the most accurate and widely used approach for forecasting (Suganthi & Samuel, 2012; Zhang, Patuwo Eddy, & HuMichael, 1998; Richard, Anthony, & Wanjoya, 2014; Oludolapo, Jimoh, Kholopane, 2012). The study the relationship between input variables and output variables is the basis of the development of ANN models and they are also used for seasonal tie series (Singh & Mishra, 2015; Suhartono, Subanar, & Guritno, 2005; Claveria, Monte, & Torra, 2014; Benkachcha, Benhra, & El Hassani, 2015). As the statistical approach, many researchers are interested to use the ANN to forecast phenomena in the different areas e.g (Kandananond, 2011; Suganthi & Samuel Anand, 2012; Zhang, Patuwo Eddy, & HuMichael, 1998; Richard, Anthony, & Wanjoya Anthony, 2014; Bodyanskiy & Popov, 2006; Jain & Kumar, 2007; Panigrahi, Karali, & Behera, 2013; Claveria, Monte, & Torra, 2014). The input variables are predictor with the ability to make general observations from the results learnt from original data, and allow us to make correct inference the latent part of the population (Adebiy, Adewumi, & Ayo, 2014). The ability to model the relationship not linear without any knowledge about its nature is one of the strengths of the neural network approach; it can approximate a function to the desired level of accuracy. This is in contrary to the many techniques for forecasting models, such as ARIMA, linear regression, which suppose that the data are generated from linear processes and as a result might be inadequate for some real data (Singh & Mishra, 2015). ANNs can give the similar results obtained by the various traditional techniques, such as (seasonal) ARIMA model. The neural network approach can also be used to model and forecast the multivariate time series (Wutsqa, Subanar2, Guritno, Soejoeti, Claveria, Monte, & Torra, 2014). In that case, the causality and cointegration studies are necessary.

In this paper, ANN and SARIMA models are used to forecast the energy consumption in U.S. and the comparative study is carried out in order to find the best forecasting model and the forecasting performance parameters, such as Mean Square Error (MSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Forecast Error (FE) were used to clarify and/ or confirm the contradictory opinions presented in literature about superiority of each model over one another.

In this paper, the introduction was found in the Section 1, the literature review in the section 2, the methodology used in the Section 3, while the experimental results were found in Section 4. Section 5 proposed the comparison results obtained from SARIMA and ANN models the, while useful conclusion is provided in Section 6.

2. Literature Review

In the forecasting literature, there are many papers using different approaches such as: time series, econometric, ANN, hybrid models etc from different areas of engineering, economics, science and technology (Nowicka-Zagrajek & Weron, 2002; Kandananond, 2011; Md Maarof, Zuhaimy, & Fadzli, 2014; Suganthi & Samuel Anand, 2012; Jain, & Kumar, 2007; Panigrahi, Karali, & Behera, 2013). It appears that the forecasting accuracy depends on the chosen models and the types of time series. The recursive and rolling regressions forms of ARIMA were used by Junttila (2001) to investigate structural breaks impact on the forecasts of Finnish inflation. It is found that the result obtained from this method produces more forecasting accuracy. Evolutionary Neural Network to forecast time series is proposed by Sibarama Panigrahi et al.; they evaluate the effectiveness of three different methods of artificial neural network (ANN) models to forecast time series. Jain A. et al. (2007) presented a hybrid neural network approach, which composites of the conventional and ANN techniques to forecast hydrologic time series. It is found that the results can be applied for some areas other than hydrology. Richard K., et al. used ANN for modeling the revenue returns of mobile payment service in Kenya. They conclude that the ANN forecast accurately the revenue returns if it is correctly trained. Yevgeniy Bodyanskiy and Sergiy Popov forecasted the quasi-periodic financial time series using neural network approach. It is showed that the ANN models are better than conventional approaches for financial time series forecasting (Bodyanskiy & Popov, 2006).

Several works in the forecasting area compared the different methods, such as ARIMA an ANN e.g. (Kandananond, 2011; Singh & Mishra, 2015; Suhartono, Subanar, & Guritno, 2005; Juan, Graff & Rodriguez, 2012); SARIMA and Genetic Algorithm (Md Maarof, Zuhaimy, & Fadzli, 2014), and the hybrid methods of ANN and ARIMA were also proposed in some papers e.g. (Jain, & Kumar, 2007; Zhang Peter, 2003). Evolutive design of ARMA and ANN models were presented by Juan J. Flores et al., 2012 [10] for time series forecasting, the result indicated that the ARIMA models forecast less than ANN models. Peter Zhang (2003) compared ARIMA, artificial neural networks (ANNs) and the combined model (hybrid approach). The results indicated that the combined model can be an effective way to improve forecasting accuracy achieved by either of the models used separately. Adebiy A. A. et al. (2014) also presented the comparison between ANN and ARIMA models to predict the stock price, the obtained results showed the slight superiority of ANN model on ARIMA model, but the difference between the actual and predicted values from the two models is not significant.

In the energy area some studies have been done on demand and consumption of energy forecasting around the world and many methods were investigated in order to find the best model with the most forecasting accuracy. Among these studies, Kandananond (2011) forecasted electricity demand in Thailand for the period 1986-2010 using time series and causal models, namely, ARIMA, MLR and ANN. The obtained results indicated that the ANN approach outperformed the ARIMA and MLR approach, but the last two approaches can be preferable to the ANN one because of the principle of parsimony. Mucuk and Uysal (2009) used the Box-Jenkins methodology to predict the primary energy demand in Turkey. The ARIMA model was built on the historical data from 1970 to 2006 and project the future values of energy demand for the period 2007-2015. The results showed that the energy demand has the same trend in training set as well as in the prediction period. Suganthi L. and Samuel Anand (2012) proposed the demand forecasting using Energy models. Consequently, they examined different forecasting models mainly traditional methods including time series, regression, econometric, ARIMA; soft computing techniques etc. In that paper, the result indicated that ARIMA and neural networks models are linked to energy demand. Nowicka-Zagrajek and Weron (2002) dealt modeling and forecasting electricity loads in California by using ARMA models with hyperbolic noise. It is found that the results obtained in the proposed methods are better than the one obtained by the official forecasts of the California System Operator. The comparative study between two types of networks, namely MLP and RBF is propose by Oludolapo et al. (2012) to forecast South Africa's energy consumption. According to them the MLP model is less accurate than RBF model.

In the neural networks approach different types of neural were used, which the most widely is feed-forward multilayer perceptron e.g. (Kandananond, 2011; Singh & Mishra, 2015; Suhartono, Subanar, & Guritno, 2005; Zhang, Patuwo Eddy, & HuMichael, 1998; Juan, Graff, & Rodriguez, 2012; Oludolapo, Jimoh, & Kholopane, 2012; Wutsqa, Subanar, Guritno, Soejoeti, Claveria, Monte, & Torra, 2014; Benkachcha, Benhra, & El Hassani, 2015), Radial basis function (RBF) (Oludolapo., Jimoh, & Kholopane, 2012; Claveria, Monte, & Torra, 2014), Elman neural network (Claveria, Monte, & Torra, 2014). For the comparative study, the ANN models generally forecast best than ARIMA models (Kandananond, 2011; Singh, & Mishra, 2015; Suhartono, Subanar, & Guritno , 2005; Juan, Graff, & Rodriguez, 2012; Adebiy, Adewumi, & Ayo, 2014), but the ARIMA model is parsimony than ANN (Kandananond, 2011; Suhartono, Subanar, & Guritno, 2005). The hybrid approach of ARIMA and ANN predict better than each model ARIMA and ANN used alone (Suhartono, Subanar, & Guritno, 2005; Jain, & Kumar, 2007; Zhang Peter, 2003).

This paper, thus confirms or not the superiority of the ANN model on the seasonal ARIMA model relative to the energy consumption prediction in residential sector of U.S.

3. Methodology

In this study, it is aimed to forecast the residential energy consumption in U.S. using the Box-Jenkins methodology and Artificial Neural Network approach and compared their results in order to know the best model for predicting energy consumption in U.S. The methodologies used in this study are presented below. The residential energy consumption data in U.S. was used in this study and the data are available in U.S. Energy Information Administration (EIA) (http://www.eia.gov/). The function arima and the function feedforwardnet (Haykin, 2009; Mark Hudson Beale, Martin Hagan, & Howard Demuth, 2015) available in Matlab were used for seasonal ARIMA and ANNs models, respectively. The statistical parameters were used for the forecasting performance of training set fit and testing set forecast. They are different performance parameters using in this paper, namely, the mean absolute error (MAE), the mean squared error (MSE), the forecast error (FE), and the mean absolute percentage error (MAPE).

3.1 Input Data

The quarterly energy consumption of the United States from January 1973 to June 2015 is used in this paper. The original data are monthly, but we converted them to quarterly data in this study. The number of observations is 170 they are measured in Billion of British thermal unit (Btu). We take logarithm of the data in order to stabilize the variance and the seasonal differenced is taken for stationarization.

The description of the data is available in the table 1 and the plot of energy data is given in the Figure 1.

Table 1. The description of the data

Parameters	Training set	Test set
Total	154	16
Mean	8.3907	8.5543
Median	8.3905	8.5244
Standard Deviation	0.2206	0.1689
Variance	0.0486	0.0285
Maximum	8.8275	8.8827
Minimum	7.9796	8.3388

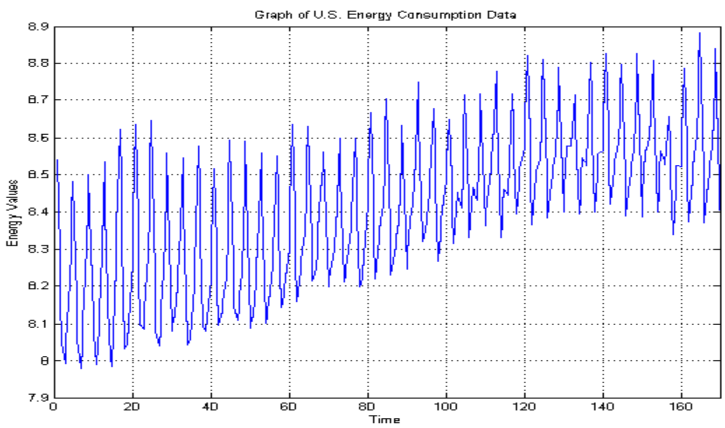

Figure 1. Graph of the energy consumption data

We observe that the energy data shown in the Figure 1 present a clear ascending trend as well as a seasonal pattern. In other hand there are ascending trend and a seasonal pattern for the energy data. Since it is clearly that they are nonstationary and seasonal because in this study all data used are quarterly data. The autocorrelations functions confirm that the data is non-stationary and has the seasonal pattern.

3.2 Autoregressive Integrated Moving Average (ARIMA) Models

In the early 1970s, George Box and Gwilym Jenkins have developed Autoregressive Integrated and Moving Average (ARIMA) for forecasting. Box and Jenkins (1970) have put together exhaustively the relevant information required to understand and study ARIMA models and later Box et al. worked in the same direction. Since then, a number of univariate ARIMA models have been published in the time series and forecasting literature (Nowicka-Zagrajek, & Weron, 2002; Mucuk & Uysal, 2009; Kandananond, 2011; Junttila, 2001; Suganthi & Samuel Anand, 2012; Shumway Robert & Stoffer David, 2006). In the ARIMA model, the estimated value of a variable is a linear combination of the past values and the past errors according to (Singh & Mishra, 2015). Normally the forecasting procedure will be performed after modeling the time series. The Box-Jenkins methodology is a popular methodology for modeling and forecasting the time series, and particular ARIMA model, it consists of the following steps: model identification, parameters estimation and diagnostics tests. Autoregressive terms (AR) and Moving average terms (MA) constitute the two different parts of the equation in the ARIMA approach. The non-seasonal ARIMA model is generally is defined as follow (Kandananond, 2011):

$$\nabla^d Y_t = \mu + \phi_1 \nabla^d Y_{t-1} + \phi_2 \nabla^d Y_{t-2} + \cdots + \phi_p \nabla^d Y_{t-p} + U_t + \theta_1 U_{t-1} + \theta_2 U_{t-2} + \cdots + \theta_q Y_{t-q} \quad (1)$$

Where the difference operator is ∇, ϕ and θ are the coefficients of the autoregressive component and moving average respectively; Y_t and U_t are the actual values and white noise at time t, respectively.

The numbers p, d and q determine the order of ARIMA model, where the order of the autoregressive term is indicated by p, while d denotes the degree of differencing (I) involved and q for the order of the moving average part μ is the constant term. In the lag operator form the ARIMA model can be writing as follow:

$$\phi(L)(1-L)^d Y_t = c + \theta(L) U_t \quad (2)$$

Where $\{U_t\}$ is a white noise with mean zero and variance σ^2, L is the backshift operator i.e. $LY_t = Y_{t-1}$; $\phi(z)$ and $\theta(z)$ are polynomials of orders p and q, respectively, $c = \mu(1-\phi_1-\phi_2-\cdots-\phi_p)$.

Similarly, a seasonal ARIMA model was introduced by Box and Jenkins (1970), and it is the most common model used in many applications economics, energy, financial, industries etc e.g. (Singh, & Mishra, 2015; Md Maarof, Zuhaimy, & Fadzli, 2014; Suhartono, Subanar, & Guritno, 2005; Shumway Robert, & Stoffer David, 2006). The seasonal ARIMA model is flexible linear models for time series that can be used to model many different types of real data including the seasonal terms or not. It is mainly based on the standard Box-Jenkins method. The seasonal autoregressive terms (SAR) and the seasonal moving average terms (SMA) are included in this method.

The multiplicative seasonal ARIMA model is represented by $ARIMA(p,d,q)\times(P,D,Q)$, where P denotes number of seasonal autoregressive (SAR) terms, the seasonal differences is denoted by D, the number of seasonal moving average (SMA) terms is Q and k denotes the seasonal period (Singh & Mishra, 2015; Md Maarof, Zuhaimy, & Fadzli, 2014; Suhartono, Subanar, & Guritno, 2005; Shumway Robert, & Stoffer David, 2006). As mentioned above, the Box and Jenkins methodology has three steps in time series analysis, namely, model identification, parameters estimation and diagnostic checking as mentioned in many papers e.g. (Nowicka-Zagrajek, & Weron, 2002; Mucuk & Uysal, 2009; Kandananond, 2011; Singh & Mishra, 2015; Md Maarof, Zuhaimy, & Fadzli, 2014; Shumway Robert & Stoffer David, 2006).

In the lag operator polynomial form the seasonal ARIMA model (multiplicative) can be writing as follow:

$$\phi_p(L)\Phi_P(L^k)(1-L)^d(1-L^k)^D Y_t = \theta_q(L)\Theta_Q(L^k)U_t \tag{3}$$

In the identification stage, a time plot of the data was constructed, and examines the stationarity of the time series. If the series is not stationary, then it was made stationary by data differentiating e.g. (Mucuk & Uysal, 2009; Kandananond, 2011; Singh & Mishra, 2015; Md Maarof, Zuhaimy & Fadzli, 2014; Shumway Robert & Stoffer David, 2006). In this process, the Correlogram function is used to identify the potential models e.g. (Mucuk & Uysal, 2009; Kandananond, 2011; Shumway Robert & Stoffer David, 2006). In order to build the best seasonal ARIMA model for residential energy consumption data, the non-seasonal and seasonal components of autoregressive and moving average parameters have to be computed for an appropriate model. The information criteria, such as Akaike Information Criterion and Bayesian Information Criterion are used to determine the best model and a Ljung-Box test is used to investigate if the residuals of the selected model are white noise. One can use the correlogram to prove that the residuals are white noise or not. For the potential models, the parameters of the models are estimated using either least squares or maximum likelihood e.g. (Mucuk & Uysal, 2009; Kandananond, 2011; Shumway Robert & Stoffer David, 2006). For our project the maximum likelihood estimator was investigated to estimate the parameters. In those models, the best model is chosen using suitable criteria. Diagnostic tests of the model are then carried out; those tests are based on the residuals e.g. (Nowicka-Zagrajek, & Weron, 2002; Mucuk & Uysal, 2009; Kandananond, 2011; Singh & Mishra, 2015; Md Maarof, Zuhaimy, & Fadzli, 2014; Suhartono, Subanar, & Guritno, 2005). If the model is accepted, then the model is used to forecast the future values.

3.3 Artificial Neural Networks (ANN) Approaches

Artificial neural networks (ANN) are a kind of Artificial Intelligence technique that mimics the behavior of the human brain (Haykin, 2009; Mark Hudson Beale, Martin Hagan, & Howard Demuth, 2015). In the ANN approach, the neurons are fully or partially connected. A neuron is a processing unit in a neural network. The neurons are connected to one another and weights assigned for the connections. Each neuron has three components, namely, inputs, an activation function and outputs. A weight is affected to each input. The output is calculated using the weighted inputs and a bias value. An activation function generates the calculated output, and the final output is generated. It is natural that the forecasting performance of artificial neural networks is influenced by those elements mentioned above, so they should be considered carefully e.g. (Kandananond, 2011; Singh & Mishra, 2015; Suhartono, Subanar, & Guritno, 2005; Oludolapo, Jimoh, & Kholopane, 2012; Wutsqa, Subanar, Guritno, Soejoeti, Claveria, Monte, & Torra, 2014).

Mcculloch and Pitts (1943) for the first time proposed the idea of the artificial neural network but because of the computing problems, they were not in much use until the back propagation algorithm was discovered by Rumelhart and others in 1986, see (Singh & Mishra, 2015). The neural network has the capacity to model the non-linear relationship without a priori assumptions of the nature of the relationship. Thus, this is its greatest advantage comparatively traditional statistical models. The ANN model performs a non-linear functional mapping from the past observations to the future value. In this study the multilayer perceptron model is employed e.g. (Kandananond, 2011; Oludolapo, Jimoh, & Kholopane, 2012; Haykin, 2009; Mark Hudson Beale, Martin, & Howard, 2015; Benkachcha, Benhra, & El Hassani, 2015), and the training is performed with one kind of the back-propagation algorithm. The multilayer perceptron neural networks (MLP) consist of multiple layers of computational units interconnected in a feedforward way e.g. (Kandananond, 2011; Suganthi & Samuel Anand, 2012; Zhang Peter, 2003; Claveria, Monte & Torra, 2014; Benkachcha, Benhra, & El Hassani, 2015). MLP are supervised neural networks that use as a building block a simple perceptron. The parallel perceptrons constitute the used topology, with connections between layers that include optimal connections. The capacity to approximate a given function by MLP network is determined by the number of neurons in the hidden layer (Kandananond, 2011; Singh & Mishra, 2015; Suhartono, Subanar, & Guritno, 2005; Haykin, 2009; Wutsqa, Subanar, Guritno, Soejoeti, Claveria, Monte, & Torra, 2014). An arbitrary function can be elegantly approximate by the feedforward networks with sufficiently many hidden units and properly adjusted parameters. The ANN model is defined as follows (Suhartono, Subanar & Guritno, 2005; Adebiy, Adewumi, & Ayo, 2014; Zhang Peter, 2003):

$$Y_t = \beta_0 + \sum_{j=1}^{q} \beta_j g\left(\sum_{i=1}^{p} \omega_{ij} X_{t-i} + \omega_{0j} \right) + \varepsilon_t \tag{4}$$

Where Y_t is the output vector of the MLP at time t ; g is the transfer function of the neurons in the hidden layer; p is the number of input nodes; X_{t-i} is the input value at time $t-1$; q is the number of neurons in the hidden layer; ω_{ij} are the weights of neuron j connecting the input with the hidden layer; β_j are the weights connecting the output of the neuron j at the hidden layer with the output neuron and ε_t is the error terms in time period t.

The transfer functions may be a linear or a non-linear function. There are different kind of the transfer functions, such as Logistic, Hyperbolic tangent, Gussian, and Sine e.g. (Kandananond, 2011; Singh & Mishra, 2015; Zhang, Patuwo Eddy, & HuMichael, 1998; Oludolapo, Jimoh, & Kholopane, 2012; Haykin, 2009; Mark Hudson Beale, Martin T. Hagan, & Howard, 2015; Benkachcha, Benhra, & El Hassani, 2015). MATLAB provides these transfer functions. The Hyperbolic Tangent Sigmoid (tansig) is used in this study for the hidden layers transfer function defined as:

$$g(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{5}$$

In this section the mainly idea is to convert the time series prediction task into a function approximation task using MLP architecture. The task is to predict for example, the quantity $S(t)$ of energy consumption at time t given as much as some previous values. The time lags $S(t-1)$; $S(t-2)$; $S(t-n)$ are used to construct the inputs and targets for training and testing sets in order to predict $S(t)$. The inputs and targets from the time series with the time lags $S(t-1)$; $S(t-4)$ and $S(t-5)$ are given the table 2.

The creation of the ANN model in Matlab includes some steps. In this case we use the MLP model. It is created by the following steps: creating, training, and testing.

1. Creating a network topology, this involves the selection of the number of input, the number of hidden layers nodes, the number of output and the activation function.

2. Training the network, this include select the optimization algorithm, input the training and target data, specify the learning parameters for training, select the performance function (MSE), training network.

3. Testing the network, in this part the trained network capacity is then evaluated by testing the model using the testing data set.

Table 2. The data set for MLP architecture

Inputs			Targets
S(1)	S(4)	S(5)	S(6)
S(2)	S(3)	S(4)	S(7)
S(3)	S(4)	S(5)	S(8)
⋮	⋮	⋮	⋮
S(N-5)	S(N-4)	S(N-1)	S(N)

In this study, the feedforwardnet (Haykin, 2009; Mark Hudson Beale, Martin, & Howard, 2015) is used to create the neural network; the training function for this work is trainbfg algorithm, the activation function tansig (relation 5) is used for hidden neuron and the activation function purelin ($purelin(x) = x$) is used for output; the ANN model was built and evaluation as in the statistical approach and the learning procedure was repeated several times for each model in order to increase the chances of getting the global minimum e.g. (Adebiy, Adewumi, & Ayo, 2014; Haykin, 2009; Mark Hudson Beale, Martin, & Howard, 2015).

In order to determine the best model, different architectures were used in this experiment. In fact the training and testing data sets were formed carefully. The network was trained and the network structure is evaluated for each training session using mean squared error (MSE). The network structure that returns the smallest MSE will be considered the best forecasting model.

4. Experimental Results and Discussion

4.1 Result of Seasonal ARIMA Model

Data is transformed by the logarithm function; the trend and seasonality are removed using seasonal differentiated of the data and the seasonal ARIMA is estimated. The first 154 data were used for data modeling and parameter estimation and the 16 last data were used to test the model and performed the forecasting evaluation as well as the comparative study. As the data are quarterly, so we have the seasonality 4. It is important to observe that the seasonal component seems to be of irregular size. To remove the trend and the seasonal variation, the seasonal differencing is applied. Because the data are non-stationary with seasonal variation, so we will first take a seasonal difference. The seasonally differenced data are shown in figure 2.The seasonal differenced data appear to be stationary, according to the figure 1; there is no remaining trend to observe. This is confirmed by unit root tests. The augmented Dickey-Fuller test rejects the null hypothesis (H_0) of a unit root with p-values smaller than 0.001 for the data.

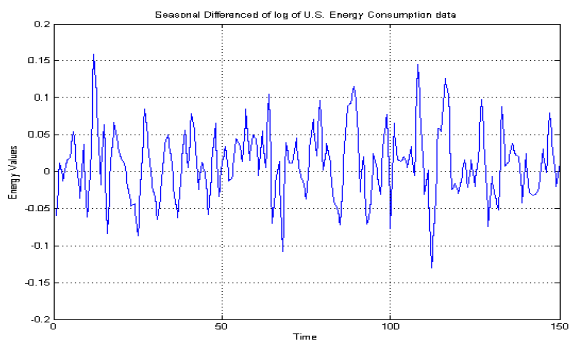

Figure 2. Graph of seasonality differenced of logarithm data

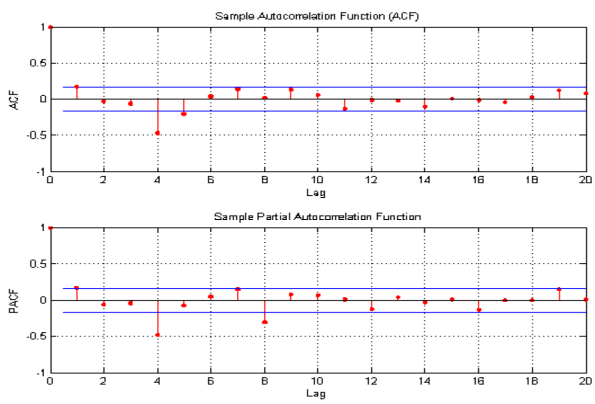

Figure 3. Seasonal differenced of the data

This suggests that the series are not integrated and hence that the seasonally differenced series is stationary. But there remains a seasonal pattern in the data. The seasonal component is not a regular sinusoidal pattern, nor is it a repetition of a fixed form, but rather it is of irregular amplitude.

Examination of the correlogram of differenced data (Figure 3) suggests at the seasonal lags a model with a seasonal moving average of order Q = 1, and a seasonal autoregressive of order P = 2. In the non-seasonal lags, there is one significant spike in the PACF suggesting a possible AR (1) term and also one significant spike in the ACF suggesting possible MA (1). Consequently, this initial analysis suggests that a possible model for these data is $ARIMA(1,0,1)\times(2,1,1)_4$.

We experimented with different parameters of seasonal and non-seasonal terms of the autoregressive and the moving average in order to find an appropriate model with the best performance. To confirm the goodness of suggested models, we use the values of information criteria in the neighborhood of this model. The model $ARIMA(1,0,1)\times(0,1,1)_4$ is considered as the best model. In fact, the selected model for forecasting can be expressed as follows:

$$(1-L^4)(1-\phi_1 L)Y_t = (1+\theta_1 L)(1+\Theta_1 L^4)U_t \tag{6}$$

The parameters estimated values are given in the following Table 3. According to the estimated values; we conclude that the estimated seasonal ARMA model is stable and invertible.

Table 3. Parameter estimates for $\mathbf{ARIMA}\ (\mathbf{1},\mathbf{0},1)\times(\mathbf{0},\mathbf{1},1)_4$ model

Parameter	Estimate	Std. Errors	t-test
AR(1)	-0.820047	0.0458554	-17.8833
MA(1)	0.97315	0.0317527	30.6477
SMA(1)	-0.341767	0.0849414	-4.02356
Variance	0.00201405	0.00021095	9.54752

The standard errors are very small relative to the estimated parameter values and the t-tests are greater than 2; this indicates that they are significant. This suggests that leaving one or more parameters out would result in an inadequate model.

The correlogram of residuals from the selected model is shown in Figure 4. All the spikes are now within the significance limits in both the ACF and PACF, and so the residuals appear to be a white noise. A Ljung-Box test also indicates that the residuals are white noise. Once the seasonal ARIMA model passes the required checks, thus, it can be used for forecasting.

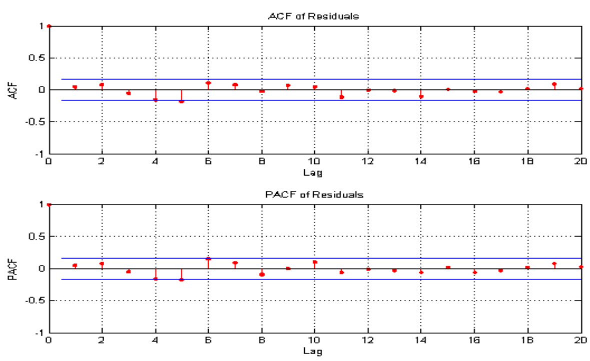

Figure 4. Correlogram of Residuals of the fitted $ARIMA(1,0,1)\times(0,1,1)_4$ model

The actual value and predicted values of residential energy consumption are given in table 4, while the graph of predicted values against actual values of energy data is found in Figure 5. The forecast error is determined by:

$$Forecast\ Error\ (EF) = \frac{(Actual\ Values - Predicted\ Values)}{Actual\ Values} \tag{7}$$

Table 4. Sample of the results of $ARIMA(1,0,1)\times(0,1,1)_4$ model with forecast error

Period	Actual values	Forecast values	Forecast error
2011-03	8.5631	8.5369	0.00305964
2011-04	8.5261	8.5897	-0.00745945
2012-01	8.6559	8.8082	-0.01759494
2012-02	8.3398	8.3981	-0.00699058
2012-03	8.5228	8.5363	-0.00158399
2012-04	8.5197	8.5901	-0.0082632
2013-01	8.7861	8.8079	-0.00248119
2013-02	8.3734	8.3983	-0.0029737
2013-03	8.4910	8.5361	-0.00531151
2013-04	8.6073	8.5903	0.00197507
2014-01	8.8827	8.8077	0.00844338
2014-02	8.3706	8.3985	-0.00333309
2014-03	8.4777	8.5360	-0.00687687
2014-04	8.5752	8.5904	-0.00177255
2015-01	8.8390	8.8077	0.00354112
2015-02	8.3388	8.3985	-0.0071593

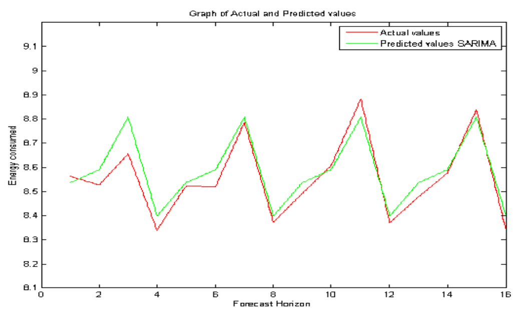

Figure 5. Graph of actual values and predicted values

4.2 Results of ANN Model

Model estimation of the type ANN is applied using the same data set as in model building of SARIMA.

The lags 1, 4 and 5 are included in the model. Different architectures were used for this experiment using the ANN approach. The network structure that returns the smallest MSE will be considered the best forecasting model. The architecture $MLP(3,10,15,1)$ is considered as the best model, i.e. it has the lower MSE. The MSE value in the training set is 0.0023, while its value is 0.0039 in the testing set. However for this architecture the training set has the best forecasting performance. We observe that the relationship between the data and the numbers of nodes in hidden layers as well as the number of layers are important in the neural network approach for time series forecasting.

In the neural network approach, a major challenge is the internal structure of the model that makes it is difficult to analyze and understand the steps by which the output is reached. Therefore, it is difficult to explain exactly why such a model has a good generalization error than other models. To get a good generalization error for a neural network model, suggest that the model has to have reached a global minimal during training and also have optimal number of weights. The Table 5 indicates the predicted values with the forecast error. The forecast error of the selected architecture is also small, which indicated good news for the forecast model. The graph of the actual and predicted values is given in the Figure 6.

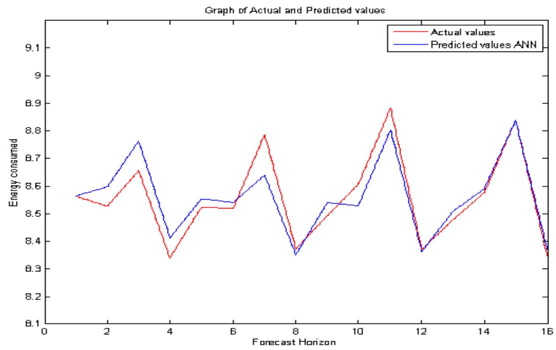

Figure 6. Graph of predicted values and actual values

Table 5. Sample results of ANN model with forecast error

Period	Actual values	Forecast values	Forecast error
2011-03	8.5631	8.5645	-0.00016349
2011-04	8.5261	8.5982	-0.00845639
2012-01	8.6559	8.7645	-0.01254636
2012-02	8.3398	8.4127	-0.00874122
2012-03	8.5228	8.5533	-0.00357864
2012-04	8.5197	8.5404	-0.00242966
2013-01	8.7861	8.6405	0.01657163
2013-02	8.3734	8.3511	0.0026632
2013-03	8.4910	8.5389	-0.00564127
2013-04	8.6073	8.5287	0.00913178
2014-01	8.8827	8.8042	0.0088374
2014-02	8.3706	8.3642	0.00076458
2014-03	8.4777	8.5089	-0.00368024
2014-04	8.5752	8.5897	-0.00169092
2015-01	8.8390	8.8376	0.00015839
2015-02	8.3388	8.3644	-0.00306999

5. Comparison of SARIMA and ANN Models

As mentioned in the introduction, the main goal of the project was to forecast the energy consumption using two approaches, statistical and neural network approaches. In this section, we deal the comparative study between the two approaches by evaluating the forecast performance of the used models. The forecast performance is evaluated using the different manners, the comparative study can be done both within the training data set in which estimation of the unknown parameters of the model was carried out and in the testing data set e.g. (Kandananond, 2011; Adebiy, Adewumi, & Ayo, 2014; Claveria, Monte, & Torra, 2014). The training data set allows us to measure the performance of the fitted model and the testing data set is to measure the performance of the forecast model e.g. (Singh & Mishra, 2015; Md Maarof, Zuhaimy, & Fadzli, 2014; Zhang Peter, 2003). Using the results given in Table 6, Table 7 and Figure 7, we observed that the ANN model and the SARIMA model have the forecasting accuracy comparatively close. Thus, we conclude that the forecasting accuracy is not quite significant. But, the performance of ANN model is better than SARIMA model in terms of forecasting accuracy from the test data using MAE and MAPE, the opposite result is happened for MSE. While the SARIMA model fits better the historical data (training data) than ANN models using all performance parameters. The SARIMA and ANN models are not directional according to the Figure 6, i.e. they are toward value forecasting. The forecasts error are very small for the used models, this can be interpreted as good news about models; this means that any model can be used to predict the residential energy consumption, but the ANN approach indicated superior performance over the SARIMA models using some performance parameters.

Statistical significance test was performed, which indicates no significant difference between the actual values and predicted values of the two models, the p-values of the test relating to ANN and SARIMA models are 0.9189 and 0.6156, respectively. Thus, this study allows us to clarify the contrary opinions reported in literature relating to the superiority of ANN model over ARIMA model in time series forecasting such as (Kandananond, 2011; Singh & Mishra, 2015; Suganthi & Samuel Anand, 2012).

According to what we have just seen above, some sources of uncertainty in forecasts models can be identified. These can arise from different reasons, such as the inherently stochastic nature of the data, the uncertainty in model specification, measurement error in the data etc.

Table 6. Sample result of ANN and seasonal ARIMA models with forecast error

Period	Actual values	Forecasting values		Forecast Error	
		SARIMA	ANN	SARIMA	ANN
2011-03	8.5631	8.5369	8.5645	0.00305964	-0.00016349
2011-04	8.5261	8.5897	8.5982	-0.00745945	-0.00845639
2012-01	8.6559	8.8082	8.7645	-0.01759494	-0.01254636
2012-02	8.3398	8.3981	8.4127	-0.00699058	-0.00874122
2012-03	8.5228	8.5363	8.5533	-0.00158399	-0.00357864
2012-04	8.5197	8.5901	8.5404	-0.0082632	-0.00242966
2013-01	8.7861	8.8079	8.6405	-0.00248119	0.01657163
2013-02	8.3734	8.3983	8.3511	-0.0029737	0.0026632
2013-03	8.4910	8.5361	8.5389	-0.00531151	-0.00564127
2013-04	8.6073	8.5903	8.5287	0.00197507	0.00913178
2014-01	8.8827	8.8077	8.8042	0.00844338	0.0088374
2014-02	8.3706	8.3985	8.3642	-0.00333309	0.00076458
2014-03	8.4777	8.5360	8.5089	-0.00687687	-0.00368024
2014-04	8.5752	8.5904	8.5897	-0.00177255	-0.00169092
2015-01	8.8390	8.8077	8.8376	0.00354112	0.00015839
2015-02	8.3388	8.3985	8.3644	-0.0071593	-0.00306999

Table 7. Forecasting performance of the models

Models	Training Data			Testing Data		
	MSE	MAE	MAPE	MSE	MAE	MAPE
SARIMA	0.0020	0.0349	0.0042	0.0034	0.0475	0.0056
ANN	0.0023	0.0368	0.0044	0.0039	0.0474	0.0055

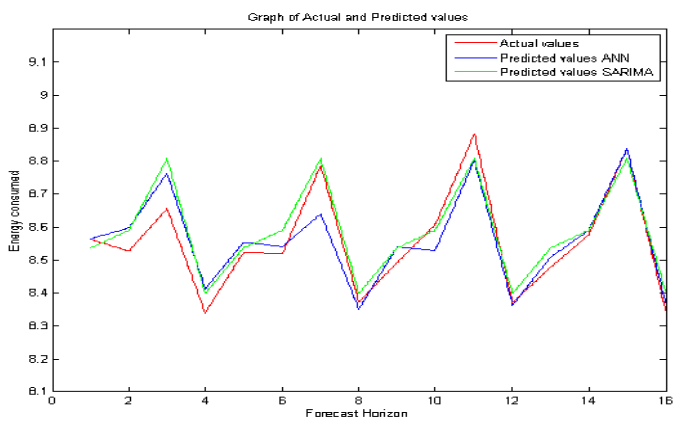

Figure 7. Graph of predicted values and actual values of ARIMA and ANN models

The above performance parameters can be determined by the following formulas:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(Y_i - F_i)^2 \tag{8}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |Y_i - F_i| \tag{9}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N} \frac{|Y_i - F_i|}{Y_i} \tag{10}$$

Where Y_i is the actual value; F_i is the predicted value; N is the number of data.

5. Conclusion

In this study, two approaches, statistical and neural network, were deployed of identifying models for time series. Later, we applied these methodologies on forecasting of the U.S. energy consumption based on the historical data from 1973 to 2015.

The empirical results obtained about the forecasting performance of seasonal ARIMA model and ANN model to energy consumption prediction have been presented in this project. The forecasting performances of the used models were performed in this study. The performance of ANN model was compared with SARIMA model, which is frequently used for time series analysis. We observe that both SARIMA and ANN models can get good forecast in application to real data with seasonal pattern and can be effectively engaged for energy forecasting. Although the performance of ANN model is better than SARIMA model using the error measurement, statistical significance test showed that there is no significant difference between the actual values and predicted values of the two models, because the actual and forecast values of the developed forecasting models are quite close. We also find that the SARIMA and ANN models are not directional.

In the forecasting procedure, the principle of parsimony is the important aspect to find the best model. If some models are equal, complex models are left to favor of simple models. However, the SARIMA model can be preferred model to ANN, because of the simplicity of its structure.

In future work, one can do further project on time series with trend and seasonal by combining some forecasting methods, especially the hybrid of machine learning approaches with the statistical approaches (Kandananond, 2011; Suganthi & Samuel, 2012; Juan, Graff, & Rodriguez, 2012; Jain &Kumar, 2007; Zhang, 2003).This technique can be used to improve existing predictive models with recent energy data. In additional the causal method for time series can be used to predict the energy consumption, which uses the independent variables that influence the energy consumption (Kandananond, 2011; Oludolapo, Jimoh, & Kholopane, 2012; Panigrahi, Karali, & Behera, 2013; Benkachcha, Benhra, & El Hassani, 2015).

Acknowledgments

The author would like to thank Mr. Wang Feixing and Mme Liu Xiuqin for their support for their support, advice and suggestions.

References

Adebiy A. A., Adewumi, A. O., & Ayo, C. K. (2014). Comparison of ARIMA and Artificial Neural Networks Models for Stock Price Prediction. *Journal of Applied Mathematics, 7.*

Benkachcha, S., Benhra, J., & El, H. H. (2015). Seasonal Time Series Forecasting Models based on Artificial Neural Network. *International Journal of Computer Applications*, *116*(20).

Bodyanskiy, Y., & Popov, S. (2006). Neural network approach to forecasting of quasi-periodic financial time series. *European Journal of Operational Research, 175*, 1357-1366.

Claveria, O., Monte, E., & Torra, S. (2014). *A multivariate neural network approach to tourism demand forecasting*. Working Paper.

Haykin, S. (2009). *Neural networks and learning machines* (3rd ed.).

Jain, A., & Kumar, A. M. (2007). Hybrid neural network models for hydrologic time series forecasting. *Applied Soft Computing, 7*, 585-592.

Juan, J. F., Graff, M., & Rodriguez, H. (2012). Evolutive design of ARMA and ANN models for time series forecasting. *Renewable Energy, 44*, 225-23.

Junttila, J. (2001). Structural breaks, ARIMA model and Finnish inflation forecasts. *International Journal of Forecasting, 17,* 203-230.

Kandananond, K. (2011). Forecasting Electricity Demand in Thailand with an Artificial Neural Network Approach. *Energies, 4*, 1246-1257. http://dx.doi.org/10.3390/en4081246

Mark, H. B., Martin, T. H., & Howard, B. D. (2015). *Neural Network Toolbox™ User's Guide.* The MathWorks, Inc.

Md Maarof, M. Z., Zuhaimy, I., & Fadzli, M. (2014). Optimization of SARIMA Model Using Genetic Algorithm Method in Forecasting Singapore Tourist Arrivals to Malaysia. *Applied Mathematical Sciences, 8*(170), 8481-8491. http://dx.doi.org/10.12988/ams.2014.410847

Mucuk, M., & Uysal, D. (2009). Turkey's Energy Demand. *Current Research Journal of Social Sciences*, *1*(3), 123-128. Retreived from http://maxwellsci.com/print/crjss/(3)123-128.pdf

Nowicka-Zagrajek, J., & Weron, R. (2002). Modeling electricity loads in California: ARMA models with hyperbolic noise. *Signal Processing, 82*, 1903-1915. http://prac.im.pwr.edu.pl/~hugo/publ/NowickaRWeron02_SP.pdf

Oludolapo, O. A., Jimoh, A. A., & Kholopane, P. A. (2012). Comparing performance of MLP and RBF neural network models for predicting South Africa's energy consumption. *Journal of Energy in Southern Africa, 23*(3).

Panigrahi, S., Karali, Y., & Behera, H. S. (2013). Time Series Forecasting using Evolutionary Neural Network. *International Journal of Computer Applications*, *75*(10).

Richard, K., Anthony, W., & Wanjoya, A. (2014). Artificial neural network application in modeling revenue returns from mobile payment services in Kenya. *American Journal of Theoretical and Applied Statistics, 3*(3), 60-64.

Shumway, R. H., & Stoffer, D. S. (2006). *Time Series Analysis and Its Applications with R Examples* (2nd ed. pp. 84-164).

Singh, A., & Mishra, G. C. (2015). Application of Box-Jenkins Method and Artificial Neural Network Procedure for Time Series Forecasting of Prices. *Statistics in Transition New Series, 16*(1), 83-96.

Suganthi, L., & Samuel, A. A. (2012). Energy models for demand forecasting—A review. *Renewable and Sustainable Energy Reviews, 16*, 1223-1240.

Suhartono, S., & Guritno, S. (2005). A Comparative Study of Forecasting Models for Trend and Seasonal Time Series: Does Complex Model Always Yield Better Forecast Than Simple Models? *Jurnalteknikindustri, 7*(1), 22-30.

Wutsqa, D. U., Subanar, Guritno, S., & Soejoeti, Z. (n. d.). Statistical Inference for Modeling Neural Network in Multivariate Time Series.

Zhang, G., Patuwo, E. B., & HuMichael, Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting, 14*, 35-62.

Zhang, P. G. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing, 50*, 159-175.