Forecasting electricity demand in Egypt: A comparison study

Mohamed A. Ismail
Cairo University, Cairo, Egypt – mismail@feps.edu.eg

Alyaa R. Zahran
Cairo University, Cairo, Egypt – azahran@feps.edu.eg

Eman M. Abd El-Metaal*
Cairo University, Cairo, Egypt – eman_mahmoud@feps.edu.eg

Abstract

Electricity is the power of any nation. Forecasting electricity demand is critical concerning the future technical improvements. An accurate hour forecast is a vital process to balance electricity produced and electricity consumed at any time in the day and to minimize the cost. Different methods have been used all over the world for forecasting electricity demand. However, a notable feature of the electricity demand time series is the presence of both within-day and within-week seasonal patterns. A double seasonal autoregressive integrated moving average (ARIMA) model and Artificial Neural Networks (ANNs) are investigated in this paper in capturing the double seasonal patterns of the Egyptian electricity demand series. The results show the accuracy of the forecasts produced by these methods. Double seasonal ARIMA model outperforms ANNs in short forecast times ahead. While for longer time horizons, double seasonal ARIMA is outperformed by ANNs.

Keywords: Double seasonality; ARIMA models; Artificial neural networks; Post-sample forecasts.

1. Introduction

Electricity is essential that people cannot live without in our daily life. Without it, life will be so much difficult and slow. Electricity utilities all over the world have given a remarkable interest for forecasting electricity demand to satisfy the growing demand. Providing accurate hourly electricity demand forecasts up to a day ahead is critical in electricity industry planning. It balances electricity produced and electricity consumed at any time through the day, increases the reliability of electric power system, minimizes costs and provides correct decisions for development [Garcia, M.P. & Kirschen, D.S. (2006)].

Different methods have been investigated in forecasting electricity demand, including multiple linear regression [Hyde, O. & Hodnett, P.F. (1997); Al-Hamadi, H. M. & Soliman, S. A. (2005); and Aslan et al.(2011)], ARIMA model [Othman et al. (2009)], and exponential smoothing method [Ismail et al. (2015)].

Electricity demand series may contain more than one seasonal pattern. A within-day seasonal pattern is notable from the similarity of the hourly electricity demand from one day to the next and a within-week seasonal pattern is also notable from the similarity of the daily demand exists week after week. Thus, using a forecasting method that taking into account both seasonal patterns (daily and weekly) is mandatory.

Norizan et al. (2011) compared double seasonal ARIMA model which captured within-day and within-week seasonal patterns with single seasonal ARIMA model based on a half hourly electricity demand Malaysian data. They concluded that double seasonal ARIMA model outperformed the single ARIMA model. Ismail et al. (2015) also adopted double seasonal ARIMA model to forecast hourly electricity demand in Egypt. Double seasonal ARIMA model provided accurate forecasts for different time horizons.

Neural networks and ARIMA models are often compared in terms of forecasting capacity. In this paper, ANNs and double seasonal ARIMA models are going to be investigated in their forecasting capacity of capturing the double seasonal patterns in the Egyptian electricity demand series. The paper is structured as follows: Section 2 describes the Egyptian electricity demand series, Section 3 describes the forecasting evaluation measurement, Section 4 describes the forecasting methods that are used, Section 5 discusses the results, while the conclusion closes the paper.

2. Hourly Egyptian Electricity Demand Series

A data set consists of hourly Egyptian electricity demand series measured in Megawatt (MW) for a one year starting on Saturday 7 January 2012 and ending on Friday 28 December 2012 is used. Figure (1) shows the Egyptian electricity demand subseries covering the period from Friday 1 June to Thursday 28 June. In the figure, from hour 1 till hour 24 represents the first day in this sub time series, while from hour 24 till hour 48 represents the second day and so on. A within-day seasonal pattern and a within-week seasonal pattern are apparent. A within-day seasonal pattern is apparent in this sub set from the similarity of the demand from one day to the next. A within-week seasonal pattern is also apparent from comparing the demand on a certain day in different weeks. It is clear that the weekdays show similar patterns of demand, while the weekend days, have the lowest peak of electricity demand, have a different electricity demand.

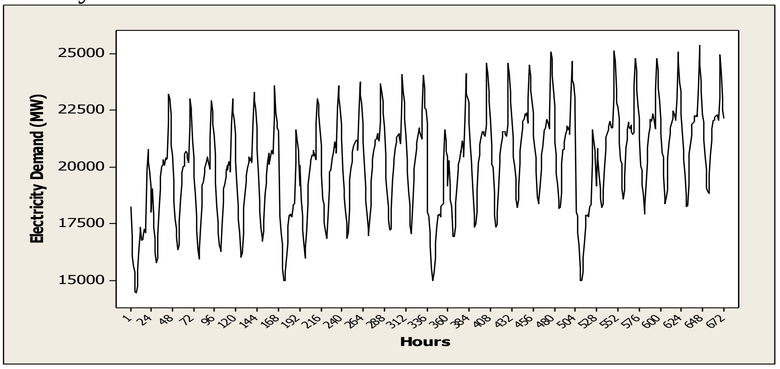

Figure (1): Egyptian electricity demand from Friday 1 June to Thursday 28 June

3. Forecasting Evaluation Measurement

For the purpose of evaluating the forecasting accuracy of the forecasting methods, the mean absolute percentage error (MAPE) is often considered [Taylor, J. W. (2003); Ismail et al. (2015)]. This statistic depends on the idea of calculating the difference between the observed and forecasted values. Low values of MAPE are preferred. MAPE measure is defined as follows:

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| * 100 \qquad (1)$$

where y_i and $\hat{y}_i$ are the actual and the forecasted values, respectively, while n is the number of forecasted values.

4. Forecasting Methods

4.1. Double Seasonal ARIMA model

ARIMA models are a combination of past series values and past errors, denoted as ARIMA (p, d, q), where p is the number of autoregressive terms, d is the order of the difference and q is the number of moving average terms. ARIMA (p, d, q) can be written as [Box, G.E. & Jenkins, G.M (1976)]:

$$\emptyset_p(B)\nabla^d y_t = \theta_q(B)\varepsilon_t \qquad (2)$$

where y_t is the electricity demand in period t ; $\nabla^d$is the nonseasonal differencing operator ; B is the backward shift operator; $\varepsilon_t$is a white noise process with mean zero and variance σ^2; $\emptyset_p(B)$ and $\theta_q(B)$ are polynomial functions of order p and q respectively ; $\emptyset_p(B)$ represents the autoregressive part and $\theta_q(B)$ represents the moving average part.

Box, G.E. & Jenkins, G.M (1976) have introduced also a multiplicative seasonal ARIMA model. They have added seasonal autoregressive terms, moving average terms and seasonal differencing to ARIMA model. Seasonal ARIMA model is denoted as ARIMA (p, d, q) $(P, D, Q)_s$ where P is the number of seasonal autoregressive terms, D is the order of seasonal differencing, Q is the number of seasonal moving average terms and s is the seasonal period in the season. Seasonal ARIMA can be expressed as

$$\emptyset_p(B)\Phi_P(B^s)\,\nabla^d\,\nabla_s^D\,y_t = \theta_q(B)\,\Theta_Q(B^s)\,\varepsilon_t \qquad (3)$$

Where ∇_s^D is the seasonal differencing operator; $\Phi_P(B^s)$, is a polynomial function of order P, represents the seasonal autoregressive terms ; and $\Theta_Q(B^s)$, is a polynomial function of order Q, represents the seasonal moving average terms.

Seasonal ARIMA model can be extended for double seasonal ARIMA model [Box et al. (1994)]. Taylor, J. W. (2003) has expressed extended seasonal ARIMA model that captures two seasonality patterns (within-day and within-week seasonal patterns). The multiplicative double seasonal ARIMA model, which is denoted as ARIMA (p, d, q) $(P_1, D_1, Q_1)_{s_1}$ $(P_2, D_2, Q_2)_{s_2}$, can be written as

$$\emptyset_p(B)\,\Phi_{P_1}(B^{s_1})\,\Omega_{P_2}(B^{s_2})\,\nabla^d\,\nabla_{s_1}^{D_1}\,\nabla_{s_2}^{D_2}\,y_t = \theta_q(B)\,\Theta_{Q_1}(B^{s_1})\,\Psi_{Q_2}(B^{s_2})\,\varepsilon_t \quad (4)$$

where $\nabla_{s_1}^{D_1}$ is the daily seasonal differencing operator; $\nabla_{s_2}^{D_2}$ is the weekly seasonal differencing operator; s_1 and s_2 are the two seasonal periods; $\Phi_{P_1}(B^{s_1})$ and $\Omega_{P_2}(B^{s_2})$ are autoregressive polynomials of orders P_1 and P_2 respectively; and $\Theta_{Q_1}(B^{s_1})$ and $\Psi_{Q_2}(B^{s_2})$ are moving average polynomials of orders Q_1 and Q_2 respectively.

Different double seasonal ARIMA models are applied to the Egyptian electricity demand series. All the data is used to estimate parameters except for the last 4 weeks that are put aside to evaluate post-sample accuracy of forecasts. We set s_1=24 to model the within-day seasonal pattern, and s_2=168 to model the within-week pattern. Lag polynomials up to order three are considered for the seasonal autoregressive polynomials and seasonal moving average polynomials. Different double seasonal ARIMA models have been estimated by maximum likelihood method using SAS software. The Schwartz Bayesian Criterion (SBC) for the different models are calculated and compared. By choosing the model corresponding to the minimum value of SBC, one is attempting to select the model corresponding to the highest Bayesian posterior probability. Double seasonal ARIMA (3,0,1) $(1, 1, 1)_{24}$ $(2, 1, 3)_{168}$ model is selected with the lowest SBC.

4.2. Artificial Neural Network

ANNs are highly interconnected processing elements designed in a way to model how the human brain performs to do a particular task. ANNs have been developed as a generalization of mathematical models of human cognition or neural biological. ANN is based on the assumptions that:

- Information processing occurs at many elements (called neurons).
- Signals are passed between neurons over connection links
- Each connection link has an associated weight
- Each neuron applies an activation function to its input to determine its output signals.

A feedforward network refers to the direction of signals flow from the input units to the output units. Input signals are passed through the neural network once to the output neurons. Feedforward networks are the most widely used for forecasting economic time series data [Kaastra, I. & Boyd, M. (1996); Zhang et al. (1998)].

Feedforward networks are designed based on three stages: i) the feedforward of the input training pattern, ii) the backpropagation of the associated error, and iii) the adjustment of the weights. The artificial neurons are organized in layers and send their signals "forward", and then the errors are propagated backwards. The network receives inputs by neurons in the input layer, and the output of the network is given by the neurons on an output layer. The backpropagation algorithm uses supervised training, which means that we provide the algorithm with the actual values and want to calculate the error (difference between actual and expected results). The idea of the backpropagation is to reduce this error. The training begins with random weights, and the goal is to adjust them so that the error will be minimized.

ANNs are investigated in forecasting the Egyptian electricity demand series. Different number of single hidden layer feedforward NNs are obtained using R package. Different electricity demand lagged values are used as input variables such that different seasonal lags are considered. For comparison purposes, we use three seasonal daily and weekly lags to match the maximum seasonal order used in the previous double seasonal ARIMA models. Daily seasonal lags that are 24, 48, and 72 and weekly seasonal lags that are 168, 336, and 504 are considered.

The following lags are considered and the MAPE of the forecasts produced by these ANNs up to one week horizon, two weeks horizon, three weeks horizon and a month horizon are calculated as follows:

- 1 to 72, 168,336, 504 with MAPE (3.96, 3.83, 3.69 and 4.29)
- 1 to 72, 168, 336 with MAPE (3.26, 3.41, 4.14 and 4.61)
- 1 to 72, 168 with MAPE (3.43, 4.76, 5.93 and 7.13)
- 1 to 48, 168, 336, 504 with MAPE (3.55, 3.44, 3.28 and 3.79)
- 1 to 48, 168, 336 with MAPE (3.08, 3.20, 3.87 and 4.28)
- 1 to 48, 168 with MAPE (3.20, 4.42, 5.52 and 6.65)
- 1 to 24, 168, 336, 504 with MAPE (2.33, 2.23, 2.25 and 2.70)
- 1 to 24, 168, 336 with MAPE (2.21, 2.30, 2.59 and 2.90)
- 1 to 24, 168 with MAPE (2.39, 3.09, 3.77 and 4.60)

The lagged values (1 to 24, 168,336, 504) and (1 to 24, 168, 336) with the lowest MAPE for different time horizons are preferred.

5. Comparative Results

In this section, double seasonal ARIMA $(3,0,1)\ (1\ ,\ 1,1)_{24}\ (2\ ,1\ ,3)_{168}$, ANN with lags 1 to 24, 168, 336, 504 and ANN with lags 1 to 24, 168, 336 are compared based on the MAPE of out-sample forecasts to determine the preferred recommended methods for forecasting the Egyptian electricity demand series. Table (1) shows the MAPE of out-sample forecasts produced by the pervious methods for one week horizon, two weeks horizon, three weeks horizon and a month horizon, while Figure (2) shows how further apart the forecasts produced by these methods from the actual values of the electricity series than others.

Table (1): The MAPE of out-sample forecasts

	Out-sample 1 week forecast	Out-sample 2 week forecast	Out-sample 3 week forecast	Out-sample 1 month forecast
DSARIMA (3,0,1) $(1,1,1)_{24}$ $(2,1,3)_{168}$	1.32	1.79	2.58	3.73
ANN with lags 1 to 24, 168, 336, 504	2.33	2.23	2.25	2.70
ANN with lags 1 to 24, 168, 336	2.21	2.30	2.59	2.90

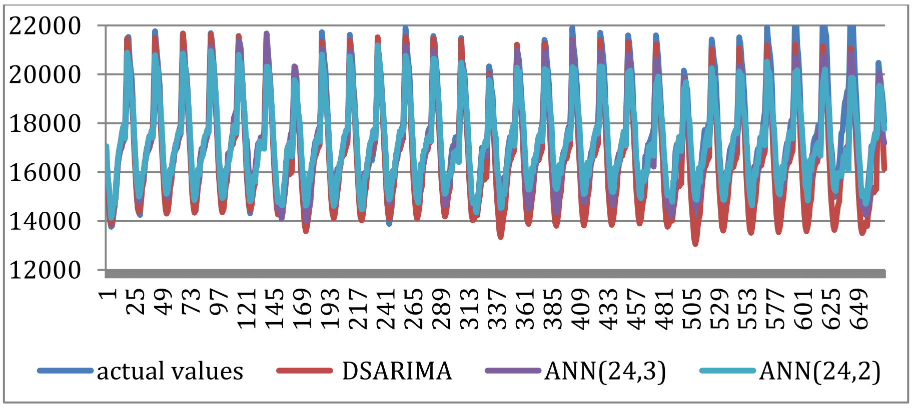

Figure (2): Time plot of the actual values versus the forecasts

From the previous results, it is shown that the forecasts produced by double seasonal ARIMA (3,0,1) $(1,1,1)_{24}$ $(2,1,3)_{168}$, ANN with lags 1 to 24, 168, 336, 504 and ANN with lags 1 to 24, 168, 336 are so close to the actual values. However, one could conclude from Table (1) and Figure (2) that double seasonal ARIMA (3,0,1) $(1,1,1)_{24}$ $(2,1,3)_{168}$ outperforms ANN models in short lead times up to two weeks ahead. Despite that for longer time horizons, double seasonal ARIMA (3,0,1) $(1,1,1)_{24}$ $(2,1,3)_{168}$ model is outperformed by the ANNs.

6. Conclusions

In this paper, we compared the performance of double seasonal ARIMA model and ANNs in forecasting the Egyptian electricity demand. A notable feature of electricity demand series is the presence of both within-day and within-week seasonal patterns. A double seasonal ARIMA model and ANNs are used to capture these seasonal patterns. Forecasts produced by these methods are accurate. Double seasonal ARIMA models and ANNs are competitive to each others for different time horizons. Despite that, double seasonal ARIMA model is better for short lead time horizons, while ANN is better for longer horizons.

References

1- Aslan, Y., Yavasca, S., & Yasar, C. (2011). Long term Electric Peak load forecasting of Kutahya using different approaches. International Journal on Technical and Physical Problems of Engineering, 3, 87-91.
2- Al-Hamadi, H. M., & Soliman, S. A. (2005). Long-term/mid-term electric load forecasting based on short-term correlation and annual growth. Electric Power System Research, 74, 353–361.
3- Box, G. E., Jenkins, G. M., & Reinsel, G. C. (1994). Time Series Analysis: Forecasting and Control. 3rd ed., Englewood Cliffs, New Jersey: Prentice Hall.
4- Box, G.E., & Jenkins, G.M. (1976). Time Series Analysis: Forecasting and Control. Revised ed., Holden-day, San Francisco.
5- Garcia, M.P., & Kirschen, D.S. (2006). Forecasting system imbalance volumes in competitive electricity markets. IEEE Transactions on Power Systems, 21, 240-248.
6- Hyde, O., & Hodnett, P.F. (1997). An Adaptable Automated Procedure for Short-Term Electricity Load Forecasting. IEEE Transactions on Power Systems, 12, 84-93.
7- Isamil, M.A., Zahran, A.R., & Abd El-Metaal, E.M. (2015). Forecasting hourly electricity demand in Egypt. Proceedings of the 60th ISI World Statistics Congress, Rio de Janeiro, Brazil, 805-809.
8- Isamil, M.A., Zahran, A.R., & Abd El-Metaal, E.M. (2015). Forecasting hourly electricity demand in Egypt using double seasonal autoregressive integrated moving average model. Proceedings of the First International Conference on Big Data, Small Data, Linked Data and Open Data, Bercelona, Spain, 42-45.
9- Kaastra, I., & Boyd, M. (1996). Designing a Neural Network for Forecasting Financial and Economic Time Series. Neurocomputing, 10, 215-236.
10- Norizan, M., Ahmad, H.M., Suhartono, & Ismail, Z. (2011). Improving Short Term Load Forecasting Using Double Seasonal ARIMA Model. World Applied Sciences Journal, 15, 223-231.
11- Othman, M.M., Abd Rahman, A.K., & Musirin, I. (2009). The Application of Box-Jenkins Models in Short-Term Load Forecasting. International Journal power Energy and Artificial Intelligence, 1, 104-110.
12- Taylor, J. W. (2003). Short-Term Electricity Demand Forecasting using Double Seasonal Exponential Smoothing. Journal of the Operational Research Society, 54, 799– 805.
13- Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with Artificial Neural Networks: The state of the Art. International Journal of Forecasting, 14, 35– 62.