2012

Finding the Best ARIMA Model to Forecast Daily Peak Electricity Demand

Mohamad As'ad
University of Wollongong, uow@asad.edu.au

Abstract
Time series models of peak daily electricity demand (June 2010-May 2011) are constructed using half hourly demand data from New South Wales, Australia. We are interested in predicting the peak electricity demand for the first seven days of June 2011 starting from 31 the May 2011. How much of the past data should be used for constructing an appropriate model which is able to provide a better forecast for the peak demand? Four appropriate ARIMA (autoregressive integrated moving average) models based past three, six, nine and twelve months of data are considered. Using RMSE (root mean square error) and MAPE (mean absolute percentage error) to measure forecast accuracy, it is shown that the ARIMA model build based on past three months data is the best model in term of forecasting two to seven days ahead and ARIMA model based on past six months data is the best model to forecast one day ahead.

Keywords
Electricity demand, ARIMA, RMSE, MAPE

Publication Details
As'ad, Mohamad, Finding the Best ARIMA Model to Forecast Daily Peak Electricity Demand, Proceedings of the Fifth Annual ASEARC Conference - Looking to the future - Programme and Proceedings, 2 - 3 February 2012, University of Wollongong.

Finding the Best ARIMA Model to Forecast Daily Peak Electricity Demand

Mohamad As'ad

Centre of Statistics and Survey Methodology, School of Mathematics and Applied statistics, University of Wollongong, NSW, 2522, AUSTRALIA
maa406@uowmail.edu.au

Abstract

Time series models of peak daily electricity demand (June 2010-May 2011) are constructed using half hourly demand data from New South Wales, Australia. We are interested in predicting the peak electricity demand for the first seven days of June 2011 starting from 31 the May 2011. How much of the past data should be used for constructing an appropriate model which is able to provide a better forecast for the peak demand? Four appropriate ARIMA (autoregressive integrated moving average) models based past three, six, nine and twelve months of data are considered. Using RMSE (root mean square error) and MAPE (mean absolute percentage error) to measure forecast accuracy, it is shown that the ARIMA model build based on past three months data is the best model in term of forecasting two to seven days ahead and ARIMA model based on past six months data is the best model to forecast one day ahead.

Keywords: Electricity demand, ARIMA, RMSE, MAPE.

1. Introduction

Electricity demand is defined by AEMO (Australia Electricity Market Operator) as the amount of electrical power generated and distributed from power plants [2]. Volatility of spot price in electricity markets is a reflection of supply and demand. Demand forecasting is focused on predicting hourly, daily, weekly and annually of the system demand and peak demand [6]. Furthermore, forecasting of daily peak electricity demand is very important for decision making processes in the electricity sector [4]. A number of methods (such as double seasonal ARIMA model, Exponential Smoothing for Double Seasonal, Artificial Neural Network model(ANN), Regression with Principal Component Analysis) are employed to forecast electricity demand in England and Wales[5]. A neuro-fuzzy model to forecast electricity demand in Victoria is used, and it is compared with ANN and ARIMA [1]. ANFIS (adaptive neural fuzzy inference system) is used to forecast monthly electricity demand and compared with autoregressive model (AR) and autoregressive moving average model (ARIMA). Results show, that ANFIS is better than others models[3]. ARIMA model is used to forecast annual electricity demand in Thailand and it is shown than it performs better than ANN and multiple linear regressions [8].

In this paper, we forecast daily peak electricity demand for the first seven days of June 2011 in NSW, based on data up dated to the end of May 2011, using four ARIMA models and then the best model based on forecasting performance using root mean square error (RMSE) and mean absolute percentage error (MAPE) is selected.

2. Data

Half hourly data for electricity demand and price for every state in Australia from December 1998 are available at www.aemo.com.au. Peak electricity demand for any day is the maximum of 48 half hourly demand during that day. Descriptive statistics for NSW of daily peak demand from June 1, 2010 to May 31, 2011 are shown in Table 1 below:

Table 1: Descriptive statistics of peak electricity demand

N	Minimum	Maximum	Mean	Std.Dev
365	7692.00	14579.86	10354.01	1191.32

Figure 1 shows the plot of time at which half hour the daily peak demand occurs for each day.

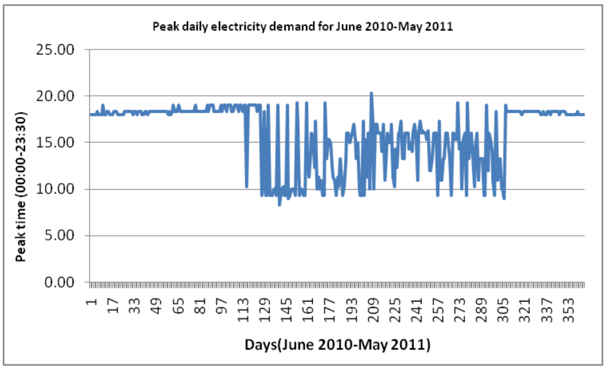

Figure 1: Time of peak demand

Figure 1 shows that the peak demand occurs during half hour interval starting at 6:00 PM, 6:30 PM, or 7:00 PM for the first 115 observations (June, 1 2010 to November,23 2010) and for the last 59 observations (April 4, 2011 to May 31, 2011). From November 24, 2010 to April 3, 2011 the time at which peak demand occurred times at 23 difference times. Table 2 shows the frequency of the time at which peak demand occurred.

Table 2: Table frequency of electricity demand

	1 June 2010 to 23 Nov 2010		4 April 2011 to 31 May 2011		24 Nov 2010 to 3 April 2011	
	Time	Freq	Time	Freq	Time	Freq
1	18:00	25	18:00	20	16:00	30
2	18:30	69	18:30	38	09:30	27
3	19:00	21	19:00	1	10:00	17
4	-	-	-	-	13:30	16
5	-	-	-	-	others	101

Daily peak electricity demand occurred at 18:30, on 110 days.

3. Data Analysis

Four set of data are constructed from the data of daily peak electricity demand for June 2010 to May 2011. The first data set is data for past 3 months (March 2011-May 2011); the second one is data for past 6 months (December 2010-May2011), the third part is data for past 9 months (September 2010-May 2011) and the last one is data for 12 months (June 2010-May 2011). Descriptive statistics for 4 time periods are reported in Table 3.

Table 3: Mean and Standard deviation

Time	Mean	Std.Dev
3 months	10157.30	894.134
6 months	10183.46	1183.438
9 months	9977.67	1073.241
12 months	10354.01	1191.321

To forecast peak demand for the first week in June 2011, we use ARIMA seasonal model. Let $BY_t = Y_{t-1}$, the seasonal ARIMA model is

$$\Phi_P(B^s)\phi_p(B)(1-B)^d(1-B^s)^D Y_t = \theta_q(B)\Theta_Q(B^s)a_t$$

Where,

$$\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_P B^{Ps},$$
$$\Theta_Q(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \cdots - \Theta_Q B^{Qs},$$
$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p,$$
$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q,$$

Y_t : data series, t: time, a_t: random error (a_t is mean zero and variance $\widehat{\sigma_a}$), B: backward shift operator, Φ: coefficient seasonal autoregressive, Θ: coefficient seasonal moving average, ϕ: coefficient nonseasonal autoregressive, θ: coefficient nonseasonal moving average, P: order seasonal autoregressive, p: order nonseasonal autoregressive, d: order of differencing nonseasonal, D: order of differencing seasonal, s: long seasonal period, q: order nonseasonal moving average, and Q: order seasonal moving average [7].

To build an ARIMA model, we use steps as below:

Step 1. Identification.
Using autocorrelation function (ACF) and partial autocorrelation function (PACF), degree of differencing and appropriate autoregressive and moving average terms are determined.

Step 2. Parameter Estimation.
Parameters of ARIMA model in step 1 are estimated using SAS. Non-significant terms are deleted and the final model has all terms significant.

Step 3. Diagnostic Checking.
Q-statistics for residuals of ARIMA model in step 2 is calculated[9]. If the Q statistic is significant then the model is not adequate, go to step 1 and if the Q-statistics is not significant then the fitted ARIMA model is appropriate.

Step 4. Forecasting.
Use ARIMA model in step 3 to obtain forecasts for first seven days of June 2011.

Let $\{Y_t\}$ denote the observed time series of peak demand, t = 1, 2, ...,n (in this term n= 365)
The first step to build an ARIMA model of peak demand is to difference Y_t by lag 1 and then by lag 7.
Let $(1-B)(1-B^7)Y_t = Z_t$.
For the four time periods data under consideration, following ARIMA models were used for forecasting peak half hourly demand for first seven days of June 2011.

- ARIMA model for March 2011 to May 2011 data (3 months data).

$$(1 - 0.342\, B^2)Z_t = (1 - 0.558\, B^7)a_t$$

with $\widehat{\sigma_a}$ = 473.498

- ARIMA model for December 2010 to May 2011 data (6 months data).

$$(1 + 0.244\, B^2 + 0.176\, B^4 + 0.181\, B^{13})Z_t = (1 - 0.864\, B^7)a_t$$

with $\widehat{\sigma_a}$= 676.498

- ARIMA model September 2010 to May 2011 data (9 months data).

$$(1 + 0.270\, B^2 + 0.208\, B^4 + 0.136\, B^{13})Z_t = (1 - 0.921\, B^7)a_t$$

with $\widehat{\sigma_a}$= 568.115

- ARIMA model for June 2011 to May 2011 data (12 months data).

$$(1 + 0.241\, B^2 + 0.220\, B^4 + 0.113\, B^7 + 0.126\, B^{11} + 0.151\, B^{13}\,)Z_t = (1 - 0.921\, B^7)a_t$$

with $\widehat{\sigma_a}$= 521.410

Q-statistic for four ARIMA models above are presented in Table 4 below:

Table 4: Q-Statistic and p-value

Model	Lag 6 (P-value)	Lag 12 (P-value)	Lag 18 (P-value)	Lag 24 (P-value)	Lag 30 (P-value)
3 months	7.31 (0.120)	11.85 (0.295)	20.33 (0.206)	23.50 (0.374)	29.53 (0.386)
6 months	1.41 (0.494)	11.65 (0.167)	14.63 (0.404)	20.73 (0.413)	30.10 (0.263)
9 months	5.96 (0.051)	15.06 (0.058)	18.80 (0.173)	25.10 (0.198)	35.05 (0.110)
12 months	-	12.31 (0.055)	15.71 (0.205)	21.27 (0.266)	30.69 (0.163)

Based on the result in Table 4, each of ARIMA model can be considered an adequate model. Forecasting upto7 days ahead (1 week) using 4 models is shown in Figure 2 below:

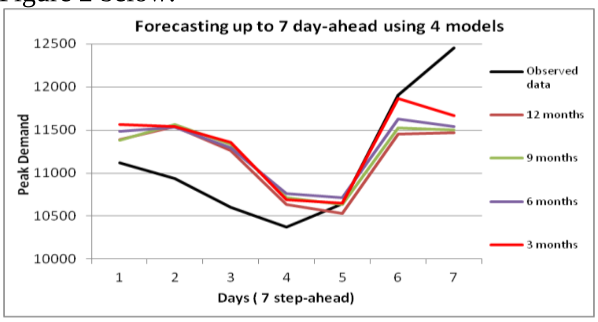

Figure 2: Observed data for 1-7 June 2011 and forecast value up to 7 days-ahead on May,31 2010 for 4 models.

Forecasting performance of the four ARIMA models given in this section are measured by using root mean square error (RMSE) and mean absolute percentage error (MAPE) criteria.

4. Forecasting

Problem of determining the best ARIMA model to forecast peak demand for the first seven days of June 2011 is considered in this section.

There are seven one day ahead forecast possible during the period May 31, 2011 to June 6, 2011.

For t=365,...,371; let $e_t(1)$ denote the forecast error for predicting one day ahead peak demand at time t, i.e, $e_t(1) = Y_{t+1} - \hat{Y}_t(1)$,

where Y_{t+1} is the observed peak demand on $t + 1$, $\hat{Y}_t(1)$ is the predicted one day peak demand.

Similarly, it is possible to make six two day ahead forecasts starting from May 31, 2011 to June 5, 2011, five three day ahead forecasts starting from May 31, 2011 to June 4, 2011..., and it is possible to make only one forecast for seven day ahead on May 31, 2011.

RMSE

For l =1, 2, ...,7; define:

$$RMSE(l) = \sqrt{\frac{1}{(8-l)} \sum_{j=0}^{(7-l)} e_{t+j}^2(l)}$$

where $e_{t+j}(l)$ = error of forecasting l days ahead demand at time t (t=365).

For the four ARIMA models, value of RMSE (l) are given in Table 5,

Table 5: RMSE for 4 models

Time	l = 1	2	3	4	5	6	7
Marc 2011-May 2011	408.363	569.826	690.048	768.902	717.198	736.883	721.635
Dec 2010-May 2011	369.678	617.429	746.333	815.573	817.695	849.605	913.991
Sept 2010-May 2011	370.735	650.977	779.785	838.162	839.88	845.417	957.796
June 2010-May 2011	357.205	630.571	772.293	851.014	872.336	892.404	989.961

Figure 3 shows value of RMSE (l) for the four models.

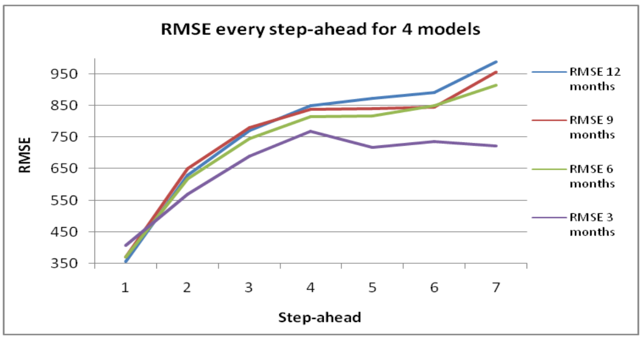

Figure 3: RMSE every step-ahead for 4 models.

It is seen from Table 5 and Figure 3 that the ARIMA model using data for December 2010-May 2011 is the best model for one day ahead forecast and ARIMA model using data for March 2011 to May 2011 is the best model for other lead times.

MAPE

For l =1, 2, ...,7; define:

$$MAPE(l) = \frac{1}{(8-l)} \sum_{j=0}^{(7-l)} \left| \frac{e_{t+j}(l)}{Y_{t+1+j}} \right| X\ 100\ \%$$

where Y_{t+1+j} = observed peak demand on $t+1+j$

day (t=365).

For the four ARIMA models, values of MAPE (l) are given in Table 6,

Table 6: MAPE for 4 models

Time	l						
	1	2	3	4	5	6	7
3 months	3.280	4.824	5.465	5.583	4.403	4.626	6.354
6 months	2.908	4.701	5.484	5.554	5.184	5.859	7.336
9 months	3.004	5.185	5.937	5.692	5.137	6.147	7.668
12 months	2.921	5.184	5.793	5.885	5.751	6.642	7.946

Figure 4 shows values of MAPE (l) for the four ARIMA models.

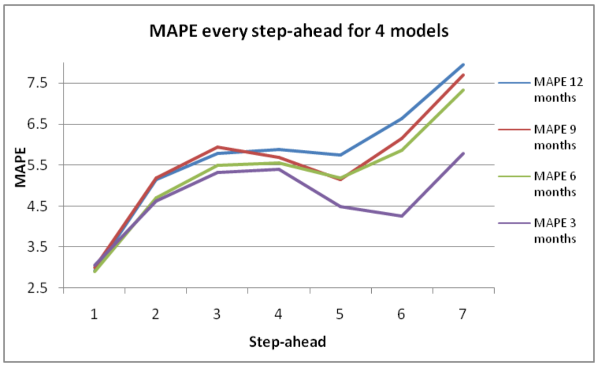

Figure 4: MAPE every step-ahead for 4 models.

It is seen from Table 6 and Figure 4, that the ARIMA model using data for December 2010-May 2011 is the best model for one day ahead forecast and ARIMA model using data for March 2011 to May 2011 is the best model for other lead times.
Therefore, same conclusion is reached by using either RMSE or by using MAPE criteria.

5. Conclusion

In this paper, the problem of finding the best model to forecast peak electricity demand for up to seven days one year data is considered. It is shown that to forecast one day ahead, six months of past data will provide better result; to forecast two or more days ahead, three months of past data may be used. It will be of interest to see if the same conclusion is valid if predictions of first seven day of another month, say December is required. Another measure of forecast accuracy is given by mean absolute deviation (MAD).

References

[1] A. Abraham, B. Nath, "A neoro-fuzzy approach for modeling electricity demand in Victoria", *Applied Soft Computing*, 1 (2001)127-138

[2] Australia Electricity Market Operator (AEMO) website,http://www.aemo.com.au/data/price_demand.html, accessed 04-11-2011.

[3] C. Ucenic, A. George, "A Neoro-fuzzy Approach to Forecast the Electricity Demand", International Conference & Environment System, (2006) 299-304.

[4] C. Sigauke, D. Chikobvu, "Prediction of daily peak electricity demand in south Africa using volatility forecasting model", *Energy Economics*, 33 (2011,) 882-888.

[5] J. W. Taylor, L. M. Menezes, P. E. McSharry, "A Comparison of Univariate Methods for Forecasting Electricity Demand Up to a Day Ahead", *International Journal of Forecasting*, 22 (2006) 1-16

[6] R. J. Hyndman, S. Fan, "Density forecasting for long-term peak electricity demand", *Power System*, 25(2)(2010). 1142-1153

[7] W. S. Wei, "*Time Series Analysis Univariate and Multivariate Methods*", Addison Wesley Publishing Company, USA, 1994.

[8] K. Kandananond, "A Forecasting Electricity Demand in Thailand with an Artificial Neural Network Approach", *Energies*, 4(2011) 1246-1247.