ISSN 0976 – 1411

Available online at www.internationaleJournals.com

International eJournals

INTERNATIONAL eJOURNAL OF MATHEMATICS AND ENGINEERING

International eJournal of Mathematics and Engineering 142 (2011) 1293 - 1300

www.InternationaleJournals.com

MODELING MONTHLY ELECTRICITY LOAD IN ANDHRA PRADESH

R. Ramakrishna
VJIT, Hyderabad, A.P., India. Email: ramakrishnaraavi9292@yahoo.com

M.Krishna Reddy
Department of Statistics, Osmania University, Hyderabad, A.P., India.
Email: reddymk54@gmail.com

Abstract

The purpose of this paper is to investigate the use of Box – Jenkins methodology for prediction of monthly peak load of electricity in Andhra Pradesh. A seasonal autoregressive integrated moving average model is developed for the forecasting the future peak load of electricity for planning. The forecasting performance of the model is evaluated using root mean square error, mean absolute error and mean absolute percentage error.

Keywords: ARIMA, ACF, PACF and Seasonal models.

1. Introduction

An effort is made in this paper to develop a seasonal autoregressive integrated moving average model for monthly peak load of electricity in Andhra Pradesh in giga watts (GW) . The data is monthly peak load of electricity from 1st April, 2005 to 31st March , 2010 consisting of 54 monthly observations for modeling and 6 monthly observations for forecasting collected from APTRANSCO, Government of Andhra Pradesh, Hyderabad, India.

2. Box-Jenkins Methodology

Let $\{Z_t\}$ be a time series. Then $\{Z_t\}$ is stationary if $E(Z_t)=\mu \; and \, V(Z_t)=\sigma_z^2$ for all t. Otherwise it is non – stationary. Let $Z_1, Z_2, \cdots Z_N$ be an observed sample. If trend line is parallel to x-axis and variability is uniform for all values of t in the sample time series graph, then the time series is stationary. Alternatively, if the ACF of sample dies out for higher lag is an indication for stationarity. The Box – Jenkins Methodology is valid for only stationary time series data. If the data is non – stationary, we convert it into stationary by stabilizing variance using logarithmic transformation and stabilizing mean using successive differencing. The Auto Regressive Integrated Moving Average model for the time series is denoted by ARIMA(p, d, q) and is defined by $\phi(B)\nabla^d Z_t = \theta(B)a_t$, where $\phi(B)=1-\phi_1 B-\phi_2 B^2 -\cdots-\phi_p B^p$ is polynomial in

B of order p and is known as Auto Regressive (AR) operator, $\theta(B)=1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q$ is a polynomial in B of order q and is known as Moving Average (MA) operator, $\nabla = 1-B$, B is the Backward shift operator $B^k Z_t = Z_{t-k}$ and d is the number of differences required to achieve stationarity. AR(p), MA(q) and ARMA(p, q) may be obtained as particular case of it with parameter values (p, 0, 0), (0, 0, q) and (p, 0, q) respectively.

The broad class of seasonal ARIMA models is expressed as

$$\phi(B)\Phi(B^s)w_t = \theta(B)\Theta(B^s)a_t \quad (2.1)$$

$$\phi(B)=1-\phi_1 B-\phi_2 B^2-\ldots-\phi_p B^p \quad (2.2)$$

$$\Phi(B^s)=1-\Phi_1 B^s-\Phi_2 B^{2s}-\ldots-\Phi_P B^{Ps} \quad (2.3)$$

$$\theta(B)=1-\theta_1 B-\theta_2 B^2-\ldots-\theta_q B^q \quad (2.4)$$

$$\Theta(B^s)=1-\Theta_1 B^s-\Theta_2 B^{2s}-\ldots-\Theta_Q B^{Qs} \quad (2.5)$$

and

$$w_t = \nabla_s^D \nabla^d Z_t \quad (2.6)$$

Note that equation (2.6) indicates that seasonal and consecutive differencing may be required to induce stationarity. Alternatively, we can summarize (2.1) as ARIMA (p, d, q) X $(P, D, Q)_s$, where ***s*** represents the periodicity of the seasonality.

Box – Jenkins methodology consists of the following four steps

Model identification: The appropriate time series model can be identified using ACF and PACF as given below:

i) The model is AR (p), if the sample autocorrelations dies out for higher lags and p sample PACs are significant. Thus the order of the autoregressive model is number of significant PACs.

ii) The model is MA (q), if the sample PACs dies out for higher lags and q sample autocorrelations are significant. Thus the order of moving average model is number of significant autocorrelations.

iii) The model is ARMA (p,q), if both sample autocorrelations and partial autocorrelations dies out for higher lags where q is the number of significant autocorrelations and p is the number of significant partial autocorrelations.

Estimation of parameters: Maximum Likelihood method is used for estimation of parameters with their significance.

Diagnostic checking: We test for the adequacy of the model identified in step1 using Box-Ljung Statistic. If the model is inadequate then repeat the steps 1 to 3 until an adequate model is observed.

Forecasting: The future values are forecasted using minimum mean squared error forecasting method. [1][2][3][4][5][6][7]

3. Model Building

As we have earlier stated that development of ARIMA model for any variable involves mainly three steps: Identification, estimation and diagnostic checking.

Model Identification:

ARIMA Model is estimated only after transforming the variable under forecasting into a stationary series. Time plot of the data (figure 1) reveals that the data is non stationary.

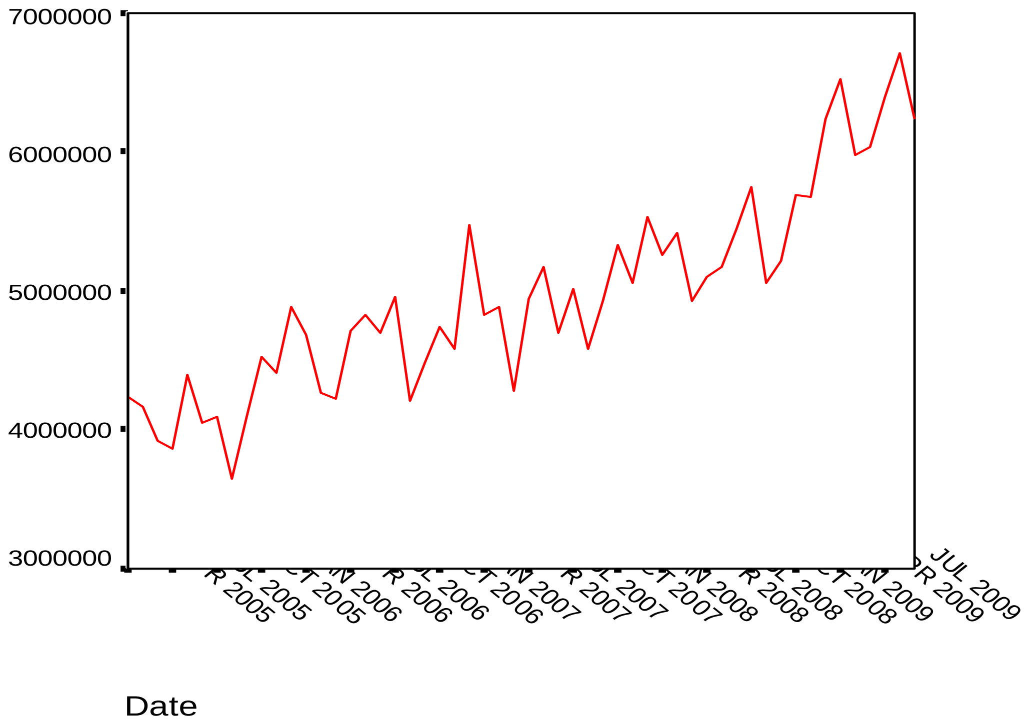


Figure 1: Time plot of monthly peak load of electricity.

Non stationarity in variance is corrected through natural log transformation and non stationarity in mean is corrected through appropriate differencing of the data. In this case, difference of order one (i.e. d=1, D=1) is sufficient to achieve stationary in mean with periodicity 12.

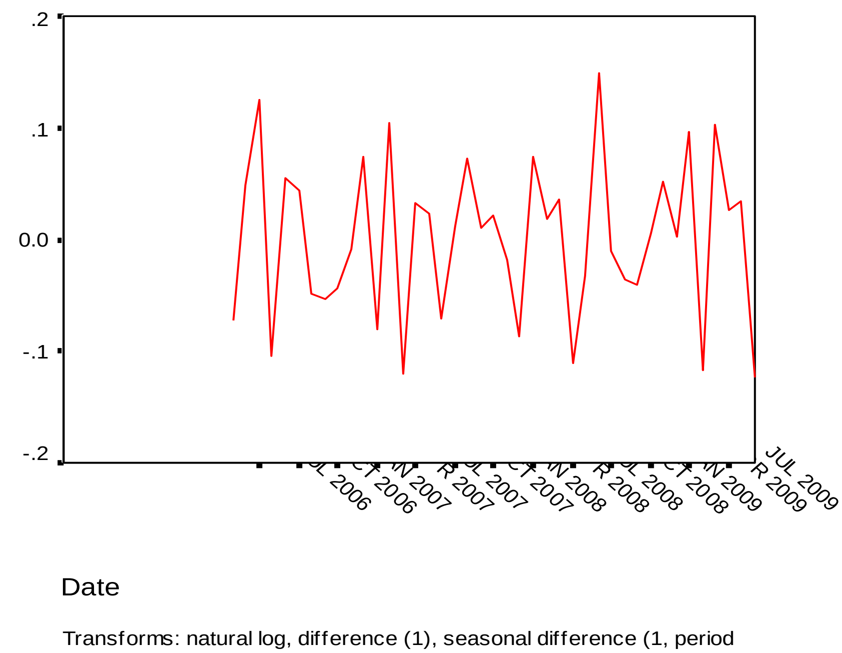
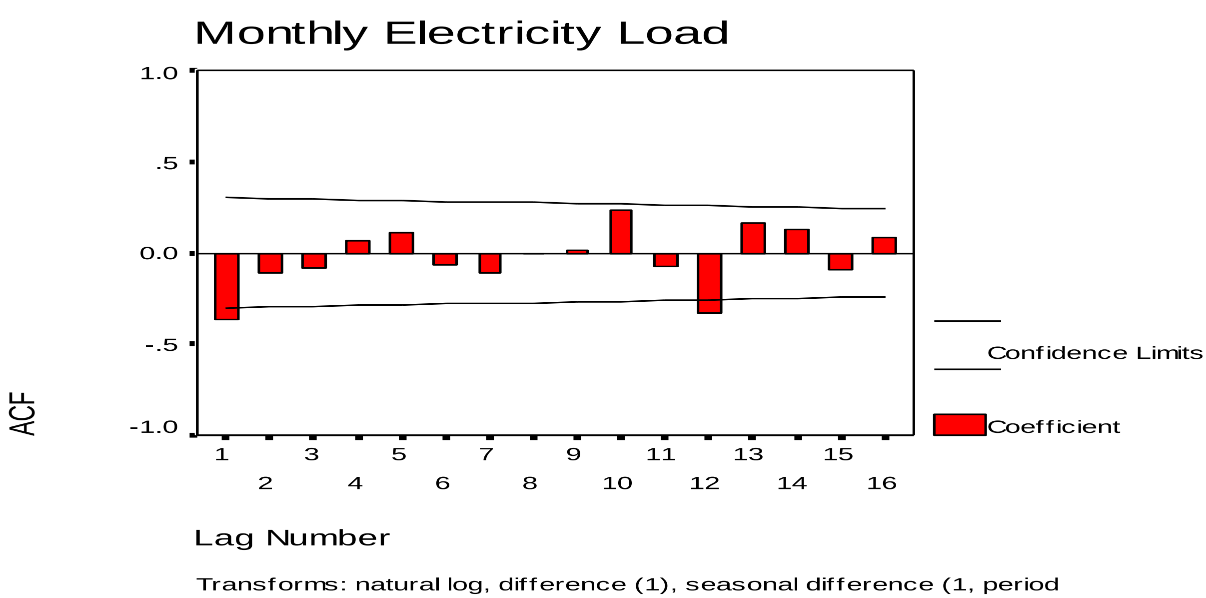


Figure 2: Time plot of the monthly peak load of electricity with difference one with weekly periodicity.

The graph of the monthly peak load of electricity data is stationary is observed from the figure 2. The next step is to identify the value of p, q, P and Q, for this, autocorrelations and partial autocorrelation of various orders of the daily peak load of electricity data is computed.



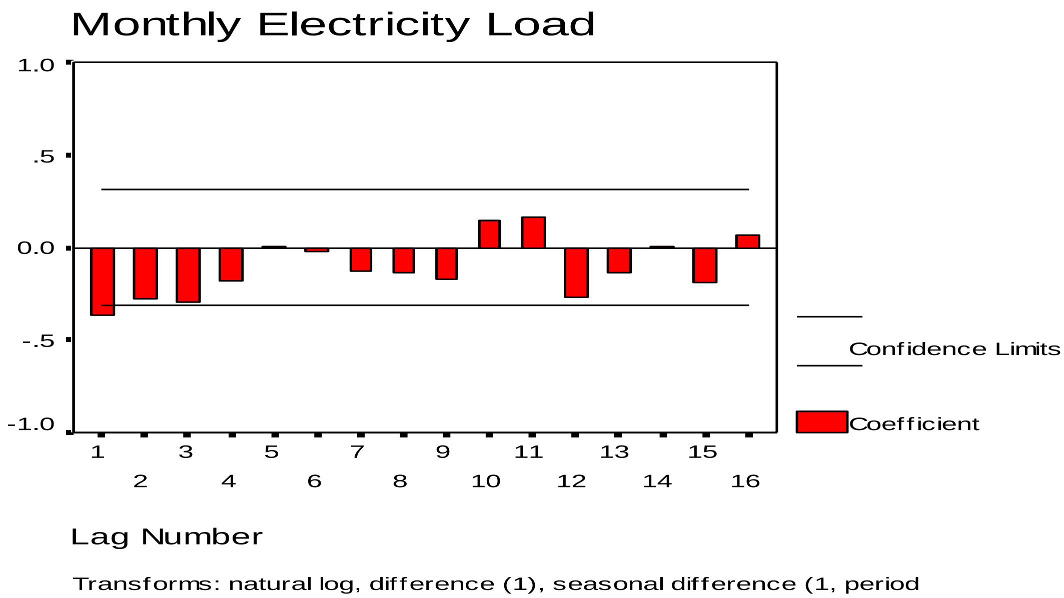


Figure 3: Sample autocorrelations and partial autocorrelations of the monthly peak load of electricity data with seasonal difference one and periodicity 12.

The order of p, P can be at most 1, the order of q can be at most 2 and Q can be at most 1 is observed from the above ACF and PACF plots. After entertaining the tentative SARIMA(p,d,q)(P,D,Q) models, we chose the model which has minimum Akaike's Information Criterion (AIC) and Schwartz Bayesian Criterion (SBC) values.

ARIMA $(p,d,q)X(P,D,Q)_s$	AIC	SBC	Standard Error
$(0,1,1)X(1,1,0)_{12}$	-118.734	-115.307	0.053

Model Estimation:

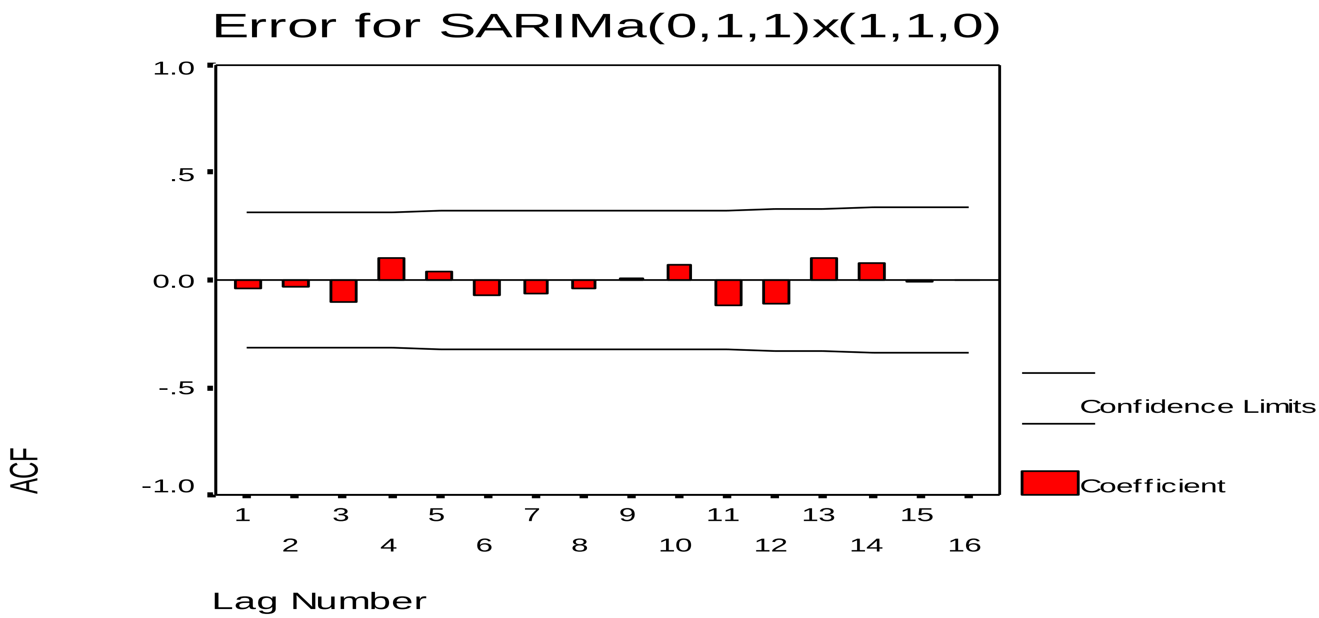
Model parameters (without the constant term in the model) are estimated using SPSS. Results of estimation of parameters are given below.

Parameter	Coefficient	SE	T-RATIO	Probability	Remarks
MA1	0.639	0.122	5.253	0.000	Significant
SAR1	-0.514	0.149	-3.443	0.001	Significant

Therefore the forecasting model is $(1 + 0.514\, B^{12})\nabla^1\, \nabla_{12}^1\, Z_t = (1 - 0.639B)a_t$

Diagnostic Checking :

This is done through examining the autocorrelations and partial autocorrelations of the residuals of various orders.



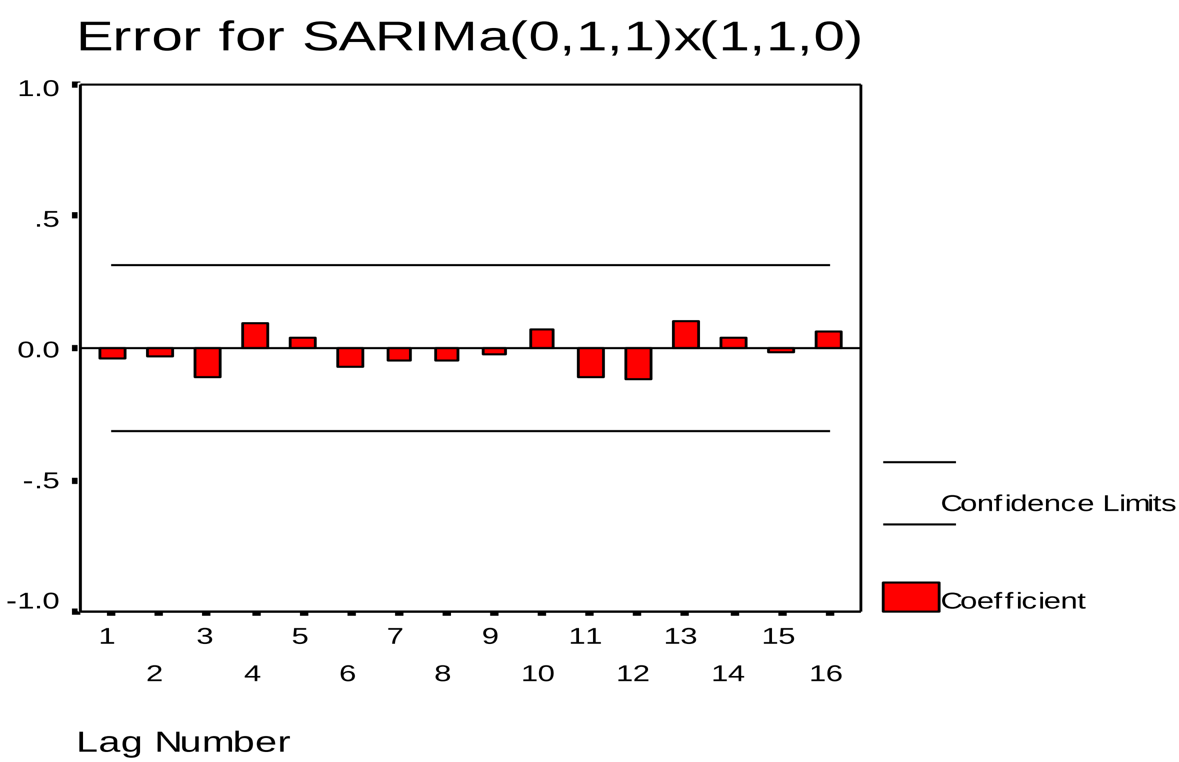


Figure 4: ACF and PACF of errors of the selected model.

As the results indicate, none of the correlations is significantly different from zero at 5% level. This proves that the model is an appropriate model. Adequacy of the model is tested using the Portmanteau test. For this purpose the various correlations of errors for 16 lags were computed and the same along with their significance which is tested by Box-Ljung test for testing the hypothesis Ho: the selected model is adequate against H1: the selected model is inadequate. The test statistic value and its significant probabilities are given below.

Lag	Box-Ljung Statistic	Sign.Prob.
5	1.219	0.943
10	2.059	0.996
15	4.644	0.995

Since the probability corresponding to Box-Ljung statistic is greater than 0.05, therefore we accept Ho and we may conclude that the selected model is an adequate model.

Forecasting

We can forecast the future monthly peak load of electricity using the model
$(1 + 0.514\, B^{12})\nabla^1\ \nabla_{12}^1\, Z_t = (1 - 0.639B)a_t$

4. Conclusions

We considered the residual analysis to measure the performance of the model using root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). These performance measures calculated for the sample data as well as computed for the future values are given in the following table.

Data	N	MAE	RMSE	MAPE
Modelled	54	229.609	284.358	4.417
Forecasts	06	294.468	308.393	4.443

It suggests that the model performance is good for a short term forecasts. Like any other method, this ARIMA technique also does not guarantee perfect forecasts nevertheless it can be successfully used for forecasting long time series data and it should be updated from time to time with incorporation of current data.

5. References

1. **Box G.E.P, Jenkins G.M, (1994),** “Time Series Analysis Forecasting and Control”, 3rd ed., Pearson Education.
2. **Brockwell, P.J. and Davis R. A. (1996),** “An Introduction to Time Series and forecasting”, Springer-verlag.
3. **Chatfield, C. (1991),** “The analysis of time series: an introduction”, 5th ed., London: Chapman and Hall.
4. **Choi, B. (1992),** “ARMA model identification”, Berlin: Springer Verleg.
5. **Krishna Reddy. M , Naveen Kumar .B (2008)**, “Forecasting foreign exchange rates using time delay neural networks” Proceedings of IV International Conference on Data Mining 2008, Las Vegas, Nevada, USA.

6. **Makridakis, S.,Wheelwright, S.C. and Hyndman, R.J. (2003),** “Forecasting methods and applications”, John Wiley and Sons.
7. **Vandaele, Walter. (1983),** “Applied time series and Box-Jenkins Models”, Academic press Inc. San Diego.