International Journal of Mathematical Archive-2(8), 2011, Page: 1341-1351
Available online through www.ijma.info ISSN 2229 – 5046

FORECASTING DAILY ELECTRICITY LOAD USING NEURAL NETWORKS

R. Ramakrishna*

Vidya Jyothi Institute of Technology, CB Post, Aziznagar, Hyderabad, India- 500075

****E-mail:** ramakrishnaraavi9292@yahoo.com*

Naveen Kumar Boiroju and M. Krishna Reddy

Department of Statistics, Osmania University, Hyderabad, India- 500 007

***E-mail:** nanibyrozu@gmail.com, reddymk54@gmail.com*

(Received on: 21-07-11; Accepted on: 03-08-11)

ABSTRACT

In this paper, a comparative study is carried out to investigate the forecasting capability of feed-forward neural networks model and Box-Jenkins methods, which are among those forecasting models most successfully applied in practice. This study investigates application of neural networks models and the results of which will be compared with those obtained by Box-Jenkins method.

Keywords: *Box-Jenkins methodology, Electricity load, Neural networks.*

---

1. INTRODUCTION

Accurate models for electric power load forecasting are essential to the operation and planning of a utility company. Load forecasting helps an electric utility to make important decisions including decisions on purchasing and generating electric power, load switching, and infrastructure development. Load forecasts are extremely important for energy suppliers, financial institutions, and other participants in electric energy generation, transmission, distribution, and markets. The total amount of electricity load [in GW] consumed in an electrical power system must be balanced with an equal amount of generated power. There is no efficient way of storing large amounts of electrical energy. To maintain this power balance between production and consumption the power input to the power system must be controlled. The electric power production may be planned by using various methods to forecast future power needs. To plan the production in the power generation plants, it is therefore very important to have accurate forecasts of the power consumption. There are various methods to produce such forecasts. A method may be said to be good if it at most times is able to forecast the power load with good precision.

The data used in this paper is collected from Andhra Pradesh Transmission Company (APTRANSCO), Hyderabad, India. In, practice, we require forecasts of electricity load for one day or one month or one year in advance, which is helpful in planning of the production of the electricity. The data set contains daily electricity load in Andhra Pradesh from April 01, 2005 to March 31, 2010 consisting of 1826 observations, in which 1796 daily observations used for estimation purposes (in-sample) and the remaining 30 daily observations left for forecast evaluation (out- of-sample).

Table 1: Average Daily Electricity Load (GW)

Average Daily Electricity Load (in GW) in Andhra Pradesh						
Sun	Mon	Tue	Wed	Thu	Fri	Sat
164.51	167.33	168.71	169.51	169.55	169.71	168.20

Various techniques have been developed during the past years for load forecasting, most of which are based on time series analysis. Statistical models are firstly adopted for the load forecasting problem, which include linear regression models, stochastic process and Box–Jenkins methods. Basically, most of the statistical methods are based on linear analysis. However, the load series are usually nonlinear functions of the exogenous variables. Therefore, to incorporate the nonlinearity, artificial neural networks (ANNs) have received much more attention recently. Neural networks have been shown to have the ability not only to learn the load series but also to model an unspecified nonlinear relationship.

***Corresponding author:** *R. Ramakrishna**, ***E-mail:** *ramakrishnaraavi9292@yahoo.com*

In recent years, feed-forward artificial neural networks multilayer perceptron based methods have received considerable attention for load forecasting. The ANNs as supervised models have been used to deal with the nonlinearity and non-stationarity in electricity load prediction and have produced good and satisfactory results.

2. REVIEW OF BOX-JENKINS METHODOLOGY

The Box-Jenkins method is one of the most widely used time series forecasting methods in practice. It is also one of the most popular models in traditional time series forecasting and is often used as a benchmark model for comparison with any other forecasting method [1].

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 stationary. 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.

SARIMA (p, d, q)X(P, D, Q) Model:

In practice many Time series contain seasonal periodic component, which repeats every 'S' observation. For example electrical power demand, which forms Time series, does contain a periodic component, which repeats every 'S' observations.

Box –Jenkins have generalized the ARIMA model to deal with seasonality and define a general multiplicative SARIMA model as

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

where $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$, $\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \ldots - \Phi_P B^{Ps}$,

$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$, $\Theta(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \ldots - \Theta_Q B^{Qs}$

$\nabla_s^D = (1 - B^s)^D$, $\nabla^d = (1 - B)^d$, $w_t = \nabla_s^D \nabla^d Z_t$

and ϕ, θ, Φ and Θ are polynomials of order p, P, q, Q respectively and a_t denotes purely random process. The variables w_t are formed from the original series $\{Z_t\}$ not only by simple differencing to remove trend but also by seasonal differencing ∇^s to remove seasonality. When fitting a seasonal model to data the first task is to assess the values of d and S which reduce the series to stationary and remove most of the seasonality. Then the values of p, P, q, Q need to be assessed by looking at the ACF and PACF of differenced series.

Box – Jenkins methodology consists of the following four steps:

(i) Model identification: We note that the following general qualitative properties provide hints about the structure of the ARIMA.

1. Non stationary: The sample autocorrelation function (ACF), $\hat{\rho}_k$ decays very slowly and sample partial autocorrelation function (PACF), φ_{kk}, has large positive or negative value of lag 1. This can possibly be reduced to a stationary series by differencing, that $\nabla Z_t, \nabla^2 Z_t \ \ldots.. \nabla^d Z_t$; usually $d \leq 2$
2. Seasonal non stationary: The sample autocorrelation function (ACF), is zero except at lags s, 2s, 3s, 4s, …and decay very slowly. This can be made stationary by seasonal differencing. That is, usually D=1.

3. Autoregressive behavior: The sample partial autocorrelation (PACF), φ_{kk}, is non-zero for lags k=1,2,3,…,p, and is zero thereafter. Fit an autoregressive of order p.
4. Seasonal autoregressive: The sample partial autocorrelation function (PACF), φ_{kk}, is zero except at lags m = s, 2s,3s,…Ps and is zero for m>Ps. Fit seasonal autoregressive of order P.
5. Moving averages behavior: The sample autocorrelation function (ACF), $\hat{\rho}_k$ is non zero for lags k=1, 2, 3… q and is zero thereafter. Fit moving averages of order q.
6. Seasonal moving average behavior: The sample autocorrelation function (ACF) is zero except at lags M = s, 2s, 3s….Qs and is zero lags M>Qs. Fit season moving averages of order Q.
7. 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.

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

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

(iv) Forecasting: The future values are forecasted using minimum mean squared error forecasting method.

If Z_t is the actual load for period t and $\hat{Z}_t$ is the forecast, then the error is defined as $e_t = Z_t - \hat{Z}_t$. The following measures may be considered

Mean Absolute Error (MAE) = $\frac{1}{n}\sum_{t=1}^{n}|e_t|$.

Root Mean Square Error (RMSE) = $\sqrt{\left(\frac{1}{n}\sum_{t=1}^{n}e_t^2\right)}$.

Mean Absolute Percent Error (MAPE) = $\frac{1}{n}\sum_{t=1}^{n}\frac{|e_t|}{Z_t}X100$.

The index of agreement(d_w) was proposed by Willmott (1981) to overcome the insensitivity of r^2 to differences in the observed and predicted means and variances .The index of agreement represents the ratio of the mean square error and the potential error (Willmot,1982) and is defined as

$$d_w = 1 - \frac{\sum_{t=1}^{N}\left(Z_t - \hat{Z}_t\right)^2}{\sum_{t=1}^{N}\left(\left|\hat{Z}_t - \bar{Z}\right| + \left|Z_t - \bar{Z}\right|\right)^2}$$

The potential error in the denominator represents the largest value that the squared difference of each pair can attain with the mean square error in the numerator. Index of agreement d_w is also very sensitive to peak load and insensitive for low load. The range of d_w is similar to that of r^2 and lies between 0 (no agreement) and 1 (perfect agreement).

The above measures are used in the following ways:

(1) The comparison of the accuracy of two different techniques.
(2) The measurement of a techniques usefulness or reliability.
(3) The search for an optimal technique.

3. FEEDFORWARD NEURAL NETWORKS

The recent upsurge in research activities into artificial neural networks (ANNs) has proven that neural networks have powerful pattern classification and prediction capabilities. One of the major application areas of ANNs is forecasting. There is an increasing interest in forecasting using ANNs in recent years. Forecasting has a long history and the importance of this old subject is reflected by the diversity of its applications in different disciplines ranging from business to engineering. The ability to accurately predict the future is fundamental to many decision processes in planning, scheduling, purchasing, strategy formulation, policy making, and supply chain operations. As such, forecasting is an area where a lot of efforts have been invested in the past. Yet, it is still an important and active field of human activity at the present time and will continue to be in the future. Forecasting has been dominated by linear methods for many decades. ANNs provide a promising alternative tool for forecasters. The inherently nonlinear structure of neural networks is particularly useful for capturing the complex underlying relationship in many real world problems. Neural networks are perhaps more versatile methods for forecasting applications in that not only can they find nonlinear structures in a problem, they can also model linear processes.

ANNs are data-driven non-parametric methods that do not require many restrictive assumptions on the underlying process from which data are generated. As such, they are less susceptible to the model misspecification problem than parametric methods. This “learn from data or experience” feature of ANNs is highly desirable in various forecasting situations where data are usually easy to collect, but the underlying data-generating mechanism is not known or pre-specifiable. Neural networks have been mathematically shown to have the universal functional approximating capability in that they can accurately approximate many types of complex functional relationships. This is an important and powerful characteristic, as any forecasting model aims to accurately capture the functional relationship between the variable to be predicted and other relevant factors or variables. The combination of the above-mentioned characteristics makes ANNs a very general and flexible modeling tool for forecasting.

Before a neural network can be used for forecasting, it must be trained. Neural network training refers to the estimation of connection weights. Although the estimation process is very similar to that in linear regression where we minimize the sum of squared errors (SSE), the ANN training process is more difficult and complicated due to the nature of nonlinear optimization involved. There are many training algorithms developed in the literature and the most influential one is the backpropagation algorithm by Werbos (1974) and Rumelhart et al. (1986). The basic idea of backpropagation training is to use a gradient-descent approach to adjust and determine weights such that an overall error function such as SSE can be minimized.

ANNs have achieved remarkable successes in the field of business forecasting. It is, however, important to note that they may not be a panacea for every forecasting task under all circumstances. Forecasting competitions suggest that no single method, including neural networks, is universally the best for all types of problems in every situation. Thus, it may be beneficial to combine several different models in improving forecasting performance. Indeed, efforts to find better ways to use ANNs for forecasting should never cease.

4. SARIMA Model

In this Section, we discuss the modeling of daily electricity load in Andhra Pradesh using Box-Jenkins methodology. The data is daily peak load of electricity from 1st April, 2005 to 31st March, 2010 consisting of 1826 observations in which 1796 daily observations used for modeling and 30 daily observations are used for forecasting. As we have earlier stated that development of SARIMA model for any variable involves mainly four steps: Identification, Estimation, Diagnostic checking and Forecasting.

Model Identification:

Time plot of the daily electricity load (figure 1) reveals that the data is seasonal and non stationary.

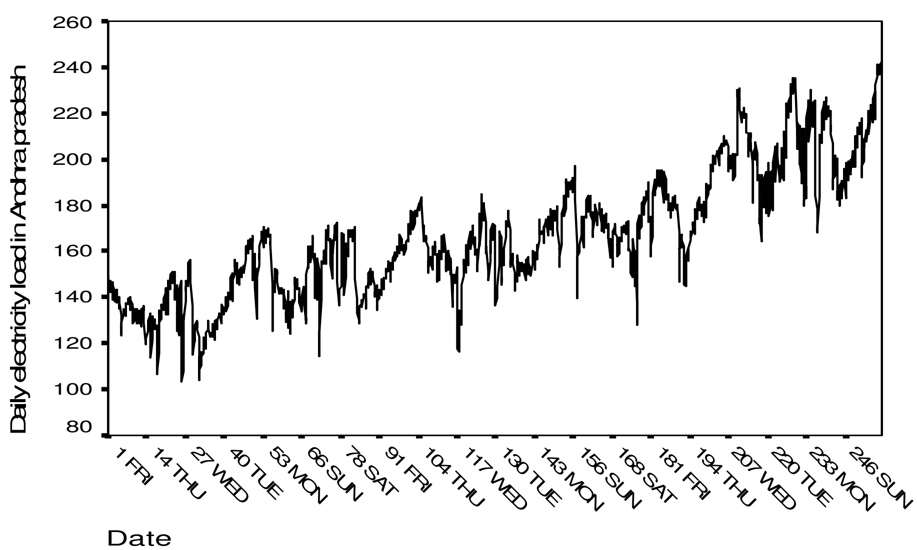

Figure: 1 Time plot of daily electricity load in Andhra Pradesh

The sample ACF for the daily electricity load in Andhra Pradesh is given below.

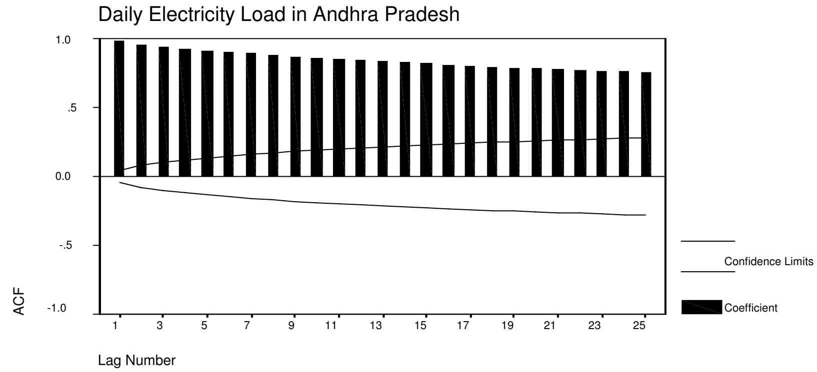

Figure: 2 Sample autocorrelation function for the daily electricity load in Andhra Pradesh

From the above time plot and ACF plot, one can observe that the given electricity load is seasonal and a seasonal autoregressive integrated moving average (SARIMA) model can fit the given data well. First we apply the seasonal difference to the given electricity load and observed the following sample ACF and PACF plots of the daily electricity load in Andhra Pradesh.

Non stationarity in variance is corrected through natural logarithm transformation and non stationarity in mean is corrected through appropriate differencing of the data. In this case, non seasonal difference of order 1 (i.e. d=1) and seasonal difference of order 1 (i.e. D=1) is sufficient to achieve stationary in mean and variance. The newly constructed variable $W_t = \nabla^1 \nabla_7^1 \tilde{Z}_t$ can now be examined for stationary.

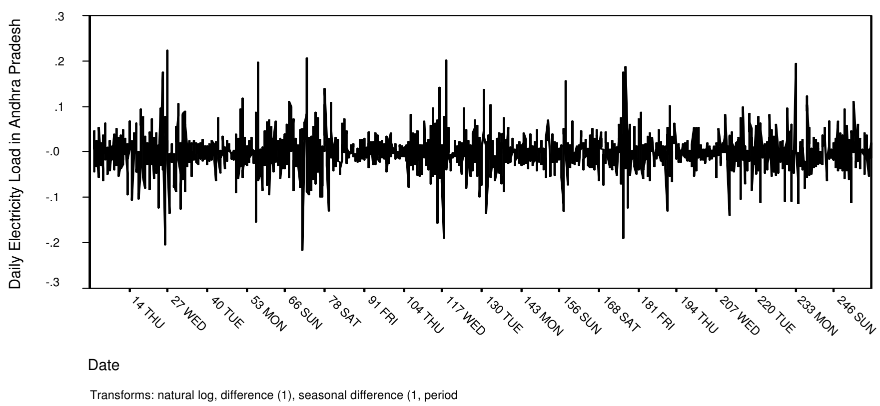

Figure: 3 Time plot of log transformed daily electricity load with non seasonal difference 1 and seasonal difference 1.

The graph (figure 3) of W_t is stationary in mean and variance. The next step is to identify the values of p, q, P and Q. Autocorrelations and partial autocorrelations for 25 lags of W_t are computed for the identification of the parameters of SARIMA model.

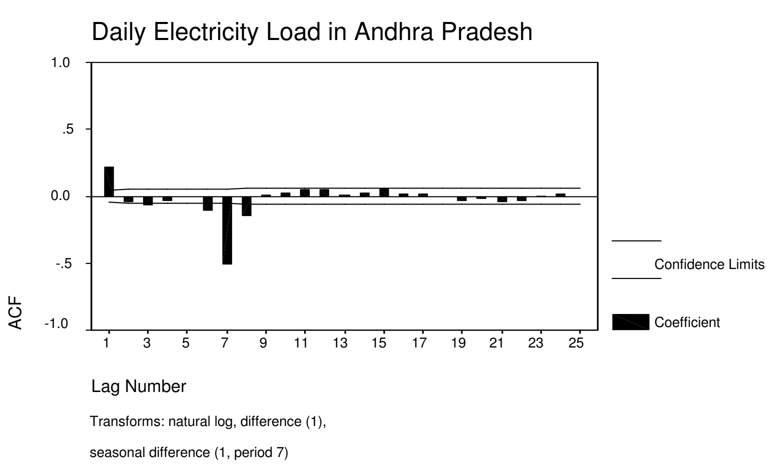

Figure: 4 Sample autocorrelation function with non-seasonal difference 1 and seasonal difference 1 of period 7.

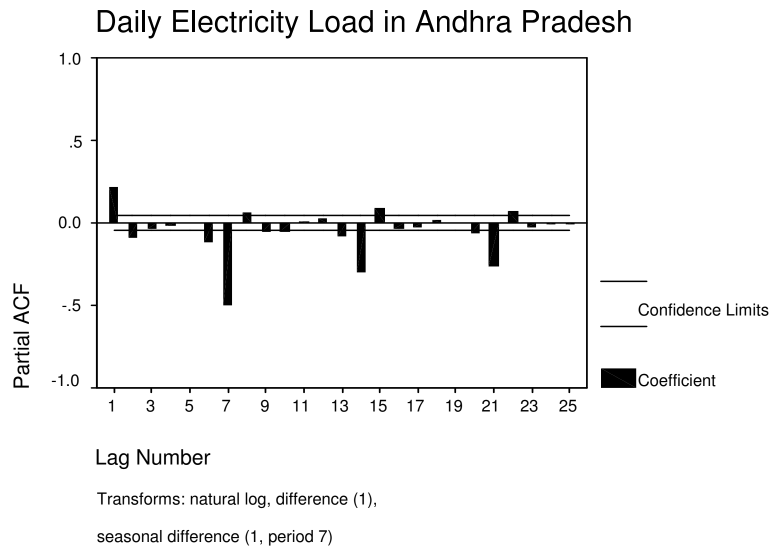

Figure: 5 Sample partial autocorrelation function with non-seasonal difference 1 and seasonal difference 1 of period 7.

From the above sample ACF and PACF, it is observed that the order of p is at most 2, d is at most 1 and q is at most 2, and the order of P is at most 3 and, D is at most 1 and Q is at most 2.We entertained the following tentative SARIMA models and chosen that the model, which has minimum BIC value. We considered the residual analysis of each model by computing MAPE, RMSE, MAE, Box-Ljung Q-Statistic and its significant probability for 25 lags are used to identifying a suitable model for the given time series on daily electricity load in Andhra Pradesh

Table: 2 Tentative adequate SARIMA Models for forecasting daily electricity load

SARIMA(p ,d ,q)X(P,D,Q)$_7$ Model	BIC	MAPE	RMSE	MAE
SARIMA(2,1,1)x(0,1,1)$_7$	2.966	1.895	4.369	3.075
SARIMA(2,1,1)x(0,1,2)$_7$	2.973	1.895	4.376	3.075

So the most suitable model is SARIMA (2, 1, 1) X (0, 1, 1)$_7$ as this model has the lowest BIC and RMSE values.
Model Estimation:

Model parameters (without constant term in the model) are estimated using PASW18 for selected model. Results of estimation of parameters are given below.

Table: 3 Model Parameters of the SARIMA (2, 1, 1) X (0, 1, 1)$_7$ Model

Parameters	B	S.E.(B)	T-Ratio	Prob.
AR1	1.094	0.031	35.844	0.000
AR2	-0.277	0.023	-11.976	0.000
MA1	0.904	0.023	38.714	0.000
SMA1	0.999	0.088	11.377	0.000

So the fitted model for the daily electricity load in Andhra Pradesh is

$$(1-1.094B+0.277B^2)\nabla^1\nabla_7{}^1\tilde{Z}_t = (1-0.904B)(1-0.999B^7)a_t\,.$$

Diagnostic Checking:

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

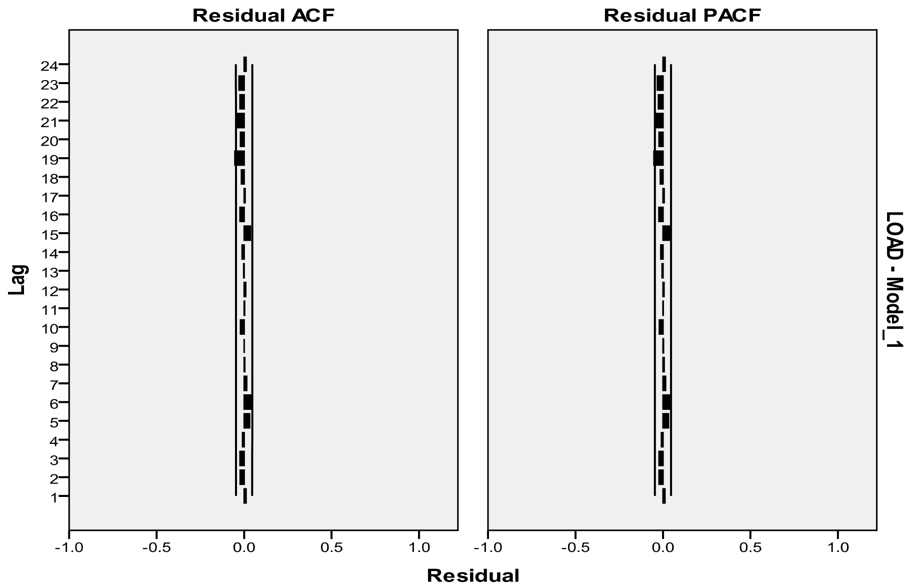

Figure: 6 Residual autocorrelations and partial autocorrelations up to the 25 lags.

As the results indicate, none of these autocorrelations is significantly different from zero at 5% level. This proves that the model is an appropriate model.
Portmanteau Test:

For this purpose, the various autocorrelations of residuals for 25 lags are computed and the same along with their significance which is tested by Box-Ljung Q- test statistic. Let the hypothesis on the model is
H_o: The selected model is adequate.
H_1: The selected model is inadequate.

Table: 4 Portmanteau Test

Ljung-Box Q-Test		
Statistics	DF	Sig.
12.943	14	0.531

Since the probability corresponding to Box-Ljung Q-statistic is greater than 0.05, therefore, we accept H_o and we may conclude that the selected seasonal autoregressive integrated moving average model is an adequate model for the given time series on daily electricity load in Andhra Pradesh.
Forecasting:
]
One can forecast the future daily electricity load in Andhra Pradesh by the equation (fitted model) by minimum mean square error method.

Forecasts of Daily Electricity Load in Andhra Pradesh:

We have forecasted the daily electricity load in Andhra Pradesh (in GW) for the out-of-sample set (2nd March, 2010 to 31st March, 2010) and the forecasts using the selected SARIMA(2,1,1)X$(0,1,1)_7$ model is tabulated in Section 6.

5. FFNN Forecasting Model

In this Section, we develop a Feed forward neural networks (FFNN) model for forecasting of daily electricity load (GW) in Andhra Pradesh State. PASW 18 software is used to build a feed forward neural network for the forecasting of electricity load in Andhra Pradesh State.

Data: The data of daily electricity load in Giga Watts (GW) is collected from APTRANSCO, Hyderabad. This data contains total 1826 observations from Friday, April 01, 2005 to Wednesday, March 31, 2010.

Rescaling:

Scale-dependent variables and covariates are rescaled to improve network training. In the present study, we use adjusted normalized method to rescale the variables. The adjusted normalized values fall between -1 and +1. The given data is partitioned into three samples namely training, testing and hold out samples. The training sample comprises the data records used to train the neural networks; the testing sample is an independent set of data records used to track errors during training in order to prevent over training. The hold out sample is another independent set of data records used to assess the final neural network; the error for the hold out sample gives an honest estimate of the predictive ability of the model because the hold out cases are not used to build the model.

We have considered the following partitions of the data for searching of an optimal FFNN model.

Table: 5 Partitions of the time series data

Partition	I	II	III	IV	V	VI
Training (%)	95	90	85	80	75	70
Testing (%)	3	8	13	18	23	28
Hold-out (%)	2	2	2	2	2	2
Total (%)	100	100	100	100	100	100

Structure of the Network:

The model is a three layer feed forward neural network and it consists of an input layer, a hidden layer and an output layer. Total number of input neurons needed in this model is two, each representing the values of lag_1 (previous day load in the same week) and lag_7 (same day in the previous week).

In this model only one output unit is needed and it indicates the forecasts of daily electricity load. There is no easy way to determine the optimum number of hidden units without training and testing. The best approach to find the optimal number of hidden units is trial and error. In practice, we can use either the forward selection or backward selection to determine the hidden layer units. We apply forward selection method, in which we select a small number of hidden neurons then record the network performance by computing the RMSE, MAE and MAPE. Next increase the hidden neurons one by one, train and test until the error is acceptably small or no significant improvement is noted. The following results are obtained for the each partition set.

Table: 6 Results of forward selection method

Partition set	Number of neurons in the layer			Error measures		
	Input	Hidden	Output	MAE	RMSE	MAPE
I	2	1	1	0.62	0.795	0.397
	2	2	1	0.752	1.041	0.474
	2	3	1	0.401	0.546	0.251
	2	4	1	0.515	0.753	0.325
II	2	1	1	0.625	1.025	0.437
	2	2	1	0.435	0.57	0.266
	2	3	1	0.757	0.994	0.494
	2	4	1	0.496	0.659	0.31
III	2	1	1	0.462	0.758	0.316
	2	2	1	0.472	0.642	0.293
	2	3	1	0.359	0.55	0.24
	2	4	1	0.374	0.56	0.245
IV	2	1	1	0.44	0.756	0.305
	2	2	1	0.418	0.583	0.275
	2	3	1	0.471	0.728	0.323

	2	4	1	0.512	0.675	0.325
V	2	1	1	0.505	0.724	0.325
	2	2	1	0.383	0.55	0.248
	2	3	1	0.413	0.687	0.268
	2	4	1	0.431	0.625	0.277
VI	2	1	1	0.527	0.827	0.356
	2	2	1	0.309	0.416	0.195
	2	3	1	0.389	0.549	0.247
	2	4	1	0.357	0.516	0.229

From the above table, we have obtained the optimum network is 2-2-1 and optimum number of hidden neurons are two in the hidden layer and the best partition is given below for which the error measures are minimum.

Table: 7 Optimum partition data for the FFNN(2-2-1) model

Partition	Percentage of data set	Number of observations
Training Set	70%	1278
Testing Set	28%	0511
Hold-out Set	02%	0037
Total	100%	1826

Network Structure:

A feed forward neural network consists with one input layer, one hidden layer and one output layer. An input layer consists of two neurons representing the lag_1 and lag_7 of the electricity load, hidden layer consists of two neurons and output layer consisting of one neuron representing the forecast values of the electricity load. The following figure of feed forward neural network gives clear idea about selected model for the given data.

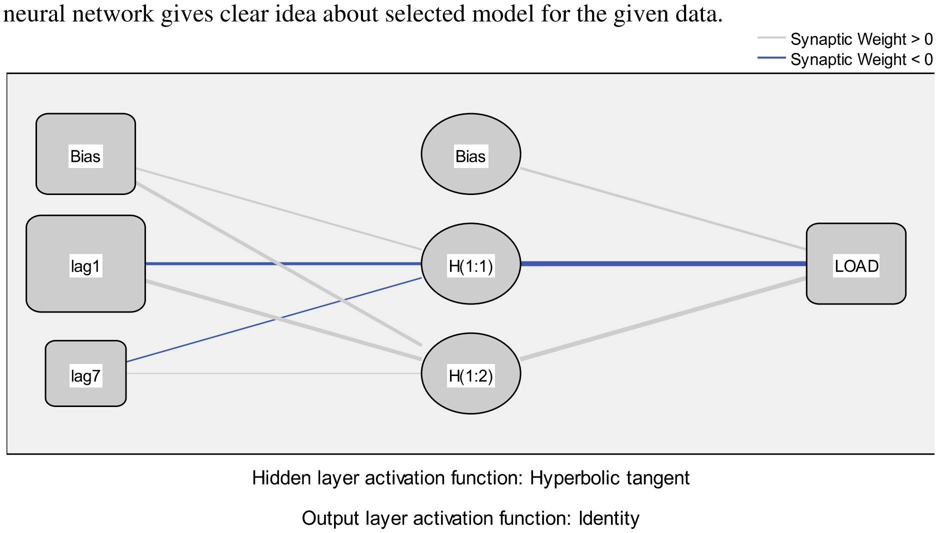

Figure: 7 Feed-forward neural network for forecasting daily electricity load.

Network Information:

The following table displays information about the neural network, including the dependent variable, number of input and output units, rescaling method, number of hidden layers and units, and activation functions.
Learning method: Supervised Learning method
Training Criteria: Online
Optimization Algorithm: Gradient descent
Initial learning rate: 0.3
Lower boundary of learning rate: 0.001
Momentum: 0.0001
Learning rate reduction, in Epochs: 10
Interval center: 0
Interval offset: ± 0.5

Synaptic Weights:

The below table displays the coefficient estimates that show the relationship between the units in a given layer to the units in the following layer. The synaptic weights are based on the training sample even if the active data set is portioned into training, testing and holdout data. Note that the number of synaptic weights can become rather large and these weights are generally not used for interpreting network results.

Table 8. Parameter Estimates

Predictor		Predicted		
		Hidden Layer 1		Output Layer
		H(1:1)	H(1:2)	Daily Load
Input Layer	(Bias)	0.620	0.788	
	lag1	-0.754	0.875	
	lag7	-0.026	0.025	
Hidden Layer 1	(Bias)			0.699
	H(1:1)			-2.969
	H(1:2)			2.129

Hidden activations:

$h_1 = \tanh(0.620 - 0.754\tilde{Z}_{t-1} - 0.026\tilde{Z}_{t-7})$, $h_2 = \tanh(0.788 + 0.875\tilde{Z}_{t-1} = 0.025\tilde{Z}_{t-7})$ where $\tilde{Z}_{t-k}$ is the rescaled variable at lag k.

Neural networks model:

$$\tilde{Z}_t = I(0.699 - 2.969h_1 + 2.129h_2)$$

The selected FFNN model is used to forecast the future daily electricity load and the forecasts are presented in Section 6.

6. CONCLUSION

The forecasts obtained using two models presented in the following table.

Table: 9 Forecasts of Electricity Load (in GW) using SARIMA and FFNN models

Original Load	SARIMA Forecasts	FFNN Forecasts	Original Load	SARIMA Forecasts	FFNN Forecasts
243.47	236.24	243.65	251.33	238.55	251.63
248.19	236.63	248.15	256.02	239.03	255.86
252.17	237.04	251.86	254.12	239.32	253.94
255.23	237.44	254.75	250.64	237.21	250.83
254.72	235.49	254.2	244.21	231.93	244.79
250.16	230.37	249.87	248.52	236.12	248.83
252.57	234.63	252.32	249.34	238.08	249.56
255.59	236.64	255.25	256.4	239.19	256.23
260.09	237.8	259.46	257.58	239.68	257.25
257.99	238.32	257.66	257.64	239.97	257.2
254.86	238.63	254.79	249.79	237.85	249.76
249.41	236.55	249.59	250.82	232.55	250.86
252.92	231.29	252.93	251.4	236.75	251.42
253.43	235.47	253.5	250.75	238.71	251.04
252.29	237.43	252.59	248.3	239.83	248.79

Here the results obtained by SARIMA and Neural networks forecasting models for daily electricity load in Andhra Pradesh State, are compared to see which the best is.

Table: 10 Performance of the SARIMA and FFNN models

Forecasting Model	Error Measusres	In-Sample Set	Out-of-Sample Set
SARIMA Model	MAPE	1.89	15.93
	RMSE	4.37	15.51
	MAE	3.07	6.13
	d_w	0.96	0.29
FFNN Model	MAPE	0.19	0.1
	RMSE	0.42	0.31
	MAE	0.31	0.26
	d_w	0.99	0.99

From the above table, it is clear that neural networks model is the best to forecast the future values, because it has minimum measures of forecasting errors such as MAPE, RMSE, MAE and maximum index of agreement (d_w). Therefore we can conclude that the forecasting of daily electricity load with feed forward neural networks is more efficient than the Box-Jenkins methods.

REFERENCES:

[1] Anderson, J. A., 1995, “Introduction to Neural Networks”, Cambridge, MA: MIT Press.

[2] Bakirtzis, A. G., Petridis, V., Kiartzis, S. J., Alexiadis, M. C., and Maissis, A. H., 1996, “A Neural Network Short Term Load Forecasting Model for The Greek Power System”, IEEE Transactions on Power Systems, 11(2), 858-863.

[3] Box, G. E. P., Jenkins, G. M. And Reinsel, G. C., 1994, “Time Series Analysis Forecasting and Control”, 3rd ed., Englewood Cliffs, N. J. Prentice Hall.

[4] De Gooijer, J.G., Hyndman, J.R., 2006, “25 Years of Time Series Forecasting”, International Journal of Forecasting, 22, 443– 473.

[5] Haykin, S. S., 1999, “Neural Networks: A Comprehensive Foundation”, Upper Saddle River, N.J., Prentice Hall.

[6] Hippert, H. S., Pedreira, C. E., & Souza, R. C., 2001, “Neural networks for short-term load forecasting: A review and evaluation”, IEEE Transactions on Power Systems, 16(1), 44-55.

[7] 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, 267-273.

[8] Ljung, G. M. and Box, G. E. P., 1978, “On A Measure of Lack of Fit in Time Series Models”, Biometrika, 65 297–303.

[9] Tang, Z., Almeida, C. D. and Fishwick, P. A., 1991, “Time Series Forecasting using Neural Networks Vs. Box-Jenkins Methodology”, Simulation, Vol. 57, No. 5, 303-310.

[10] Vermaak, J. & Botha, E. C., 1998, “Recurrent Neural Networks for Short-Term Load Forecasting”, IEEE Transactions on Power Systems, 13(1), 126-132.

[11] Willmott, C.J., 1981, “On the Validation of Models”, Physical Geography, 2, 184-194.

[12] Willmott, C.J., 1982, “Some Comments on the Evaluation of Model Performance”, Bulletin of the American Meteorological Society, 63, 1309-1313.

[13] Zhang, G., Patuwo, B. E. and Hu, M. Y., 1998, “Forecasting with Artificial Neural Networks: The State of the Art”, International Journal of Forecasting, 14, 35-62.

****

