Sri Lankan Journal of Applied Statistics. Vol.9, 2008, p121-131

Day and Night, Electricity Load Forecasting for Peak Values in Sri Lanka

T. M. J. A. COORAY AND T. S.G. PEIRIS

Department of Mathematics
Faculty of Engineering
University of Moratuwa, Sri Lanka

ABSTRACT

Short term load forecasting (STLF) of electricity and power demand plays a very important role in the economic and secure operation of power systems. Accurate load forecast will lead to appropriate scheduling and planning with much lower costs in the operation of power systems. Various classes of load forecast¬ing models of power systems of different nature have been suggested as reported in the literature. However, at present there is no such technique used in Sri Lanka regarding the daily day and night load forecasting. This paper develops a state space based on structural time (SSST) series model to forecast day and night peak values of electricity demand in Sri Lanka for week-days and week-ends. The model was developed using hourly electricity consumption in Sri Lanka for May 2007. The model adequacy was tested using various statistical indicators and errors were found white noise. The efficiency of the models was further tested for an independent set of data. In order to derive the coefficients of the model parameter a computer program was developed using Matlab. It is shown that the developed approach can produce more accurate results for the STLF than the conventional techniques. The developed model is more beneficial for dispatching centers of a power system.

Key words: ARIMA, electricity consumption, Short term load forecasting, State space models, Structural time series

INTRODUCTION

Electricity supply planning requires efficient management of existing power systems and optimization of the decisions concerning additional capacity. Demand prediction is an important aspect in the development of any model for electricity supply planning. The form of the demand depends on the type of planning and the accuracy that is required, hence it can be represented as an annual demand (GW), a peak demand (MnW), or load duration curves like daily, weekly or annual. Short Term Load Forecasting (STLF) is required for the control and scheduling of power systems. The focus varies from minutes to several hours ahead. The predictions are required as inputs to scheduling algorithms for the generation and transmission of electricity.

The load forecasts help in determining which devices to operate in a given period, so as to minimize costs and secure demand even when local failures may occur in the system. Short term (24 hours ahead) prediction of future load demand is important for the economic and secure operation of power systems. Fundamental operational functions such as unit commitment, hydro-thermal coordination, interchange evaluation; scheduled maintenance and security assessment require a reliable STLF. Owing to the importance of the STLF, research in this area over the past 40 years has resulted in the development of numerous forecasting methods. These methods are mainly classified into two categories: classical approaches and stochastic approaches.

Classical STLF approaches are based on various statistical methods for processing the numerical information. These approaches forecast current value of a variable by using a mathematical combination of the previous values of those variables and previous or current values of other variables. However, such models are not suitable for short-term load forecasting (Ansley and Kohn, 1985). But stochastic approaches are not being devéloped for load forecasting. For example, operators of dispatching centers for scheduled maintenance or adequacy assessment require to know daily peak load in advance for their planning. Thus in this paper, forecasting of daily peak load is considered in stochastic STLF environment.

METHODOLOGY

Stochastic time series method appears to be the most popular approach that is applied to short term load forecasting electric power industry. The theory of stochastic time series time series is discussed in many time series text books (Anderson, 1971; Akaike, 1974; Wei, 1994) and this approach have been used in many load forecasting publications. In this methodology the load series, X(t), is modeled as the output from the linear filter that has a random series input, e_t usually called a white noise as shown in Figure 1.

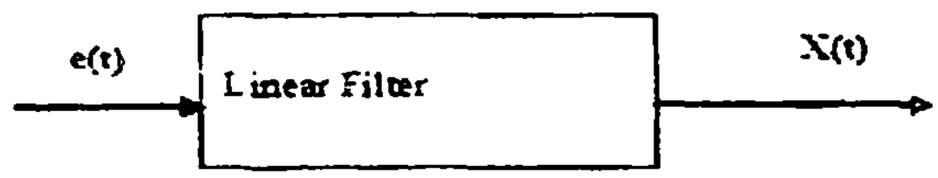

Figure 1: Load Time Series Modelling

This random input has zero mean and unknown fixed variance σ^2. Depending on the characteristic of the linear filter, different versions of Autoregressive Integrated Moving-Average (ARIMA) can be modeled.

Autoregressive Integrated Moving-Average (ARIMA) Models

The time series models using ARMA platform are assumed to be stationary process (Box and Jenkins, 1976). This means that the mean of the series of any of this processes and the covariance among its observations do not change with time. If the process is not stationary, transformation of the series to a stationary process has to be performed first. This can be achieved, for the time series that are non stationary in the mean, by differencing the process. By introducing the ∇ operator, the differenced series of order d differenced time series is written as $\nabla^d x_t = (1-B)^d x_t$.

The differenced stationary time series can be model as an AR, MA, or an ARMA to yield an ARIMA time series processes. The final model of the series that needs to be differenced d (in general d = 0, 1, or 2) times and has a orders p and q respectively for the AR and MA components, ARIMA(p,d,q) and SARIMA(p,d,q)(P,D,Q)$_s$ models, can be written as:

$$\phi(B)\nabla^d x_t = \theta_q(B)e_t \tag{1}$$

$$\Phi(B^S)\phi(B)\nabla^d x_t = \Theta(B^S)\theta_q(B)e_t \tag{2}$$

where $B^k x_t = x_{t-k}$, and the autoregressive and moving average polynomial operators are defined respectively as

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 .... - \phi_p B^p \tag{3}$$

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3 .... - \theta_q B^q \tag{4}$$

$$\Phi_P(B^s) = 1 - \delta_1 B^s - \delta_2 B^{2s} - \delta_3 B^{3s} .... - \delta_p B^{Ps} \tag{5}$$

$$\Theta_Q(B^s) = 1 - \beta_1 B^s - \beta_2 B^{2s} - \beta_3 B^{3s} .... - \beta_q B^{Qs} \tag{6}$$

If we assume that the zeros in (3) to (6) are outside the unit circle. To illustrate the model identification we consider the general SARIMA(p,d,q)(p,D,Q)$_s$ model. In the identification stage preliminary estimates for *p, d,* and *q* are obtained using various techniques such as simultaneous inspection of auto and partial correlations functions

(Box and Jenkins,1976), extended sample autocorrelations (Tsay and Tiao, 1984), and order determination criteria such as in (Lutkrpohlth, 1985). Based on these preliminary values the actual autoregressive (AR) and /or moving average (MA) parameters are estimated using linear procedures proposed by Koreisha and Pakkila (1993). Subsequently residuals from such models undergo various diagnostic checks to determine the adequacy of the model. If necessary modifications are made to the model structure to ensure that the residuals behave like white noise.

Implementation of the proposed approach for STLF

To implement the proposed approach, we statistically studied the load demand of the Sri Lanka network. From this statistical study, the following results have been obtained: at first, it is noted that in Sri Lanka, the shape of the load curve on all weekdays, and weekends are almost the same. The typical load curves for weekdays and weekend have been shown in the Figures 2 and 3 respectively.

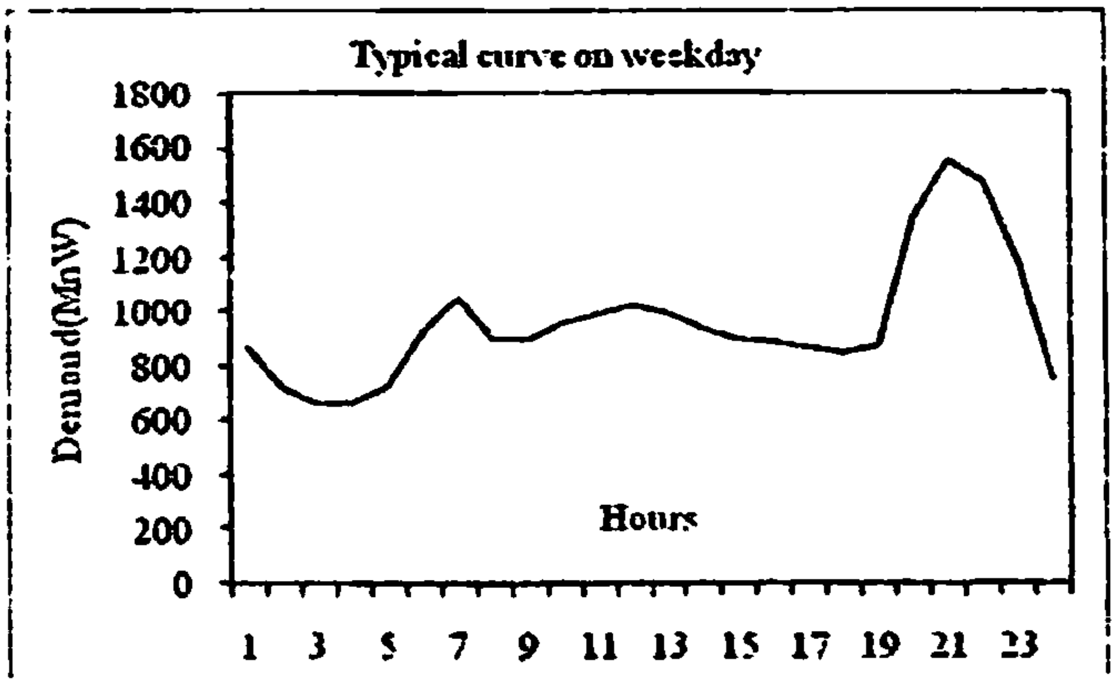

Figure 2: Typical curve of hourly electricity demand in Sri Lanka on a weekday

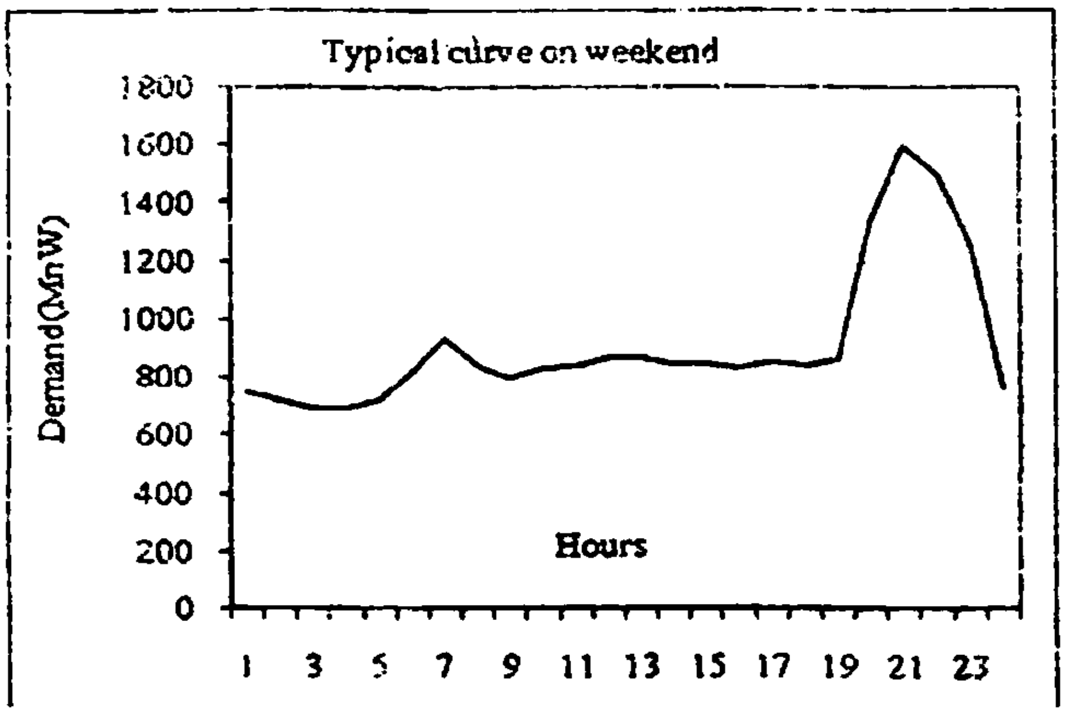

Figure 3: Typical curve of electricity demand in Sri Lanka on a weekend

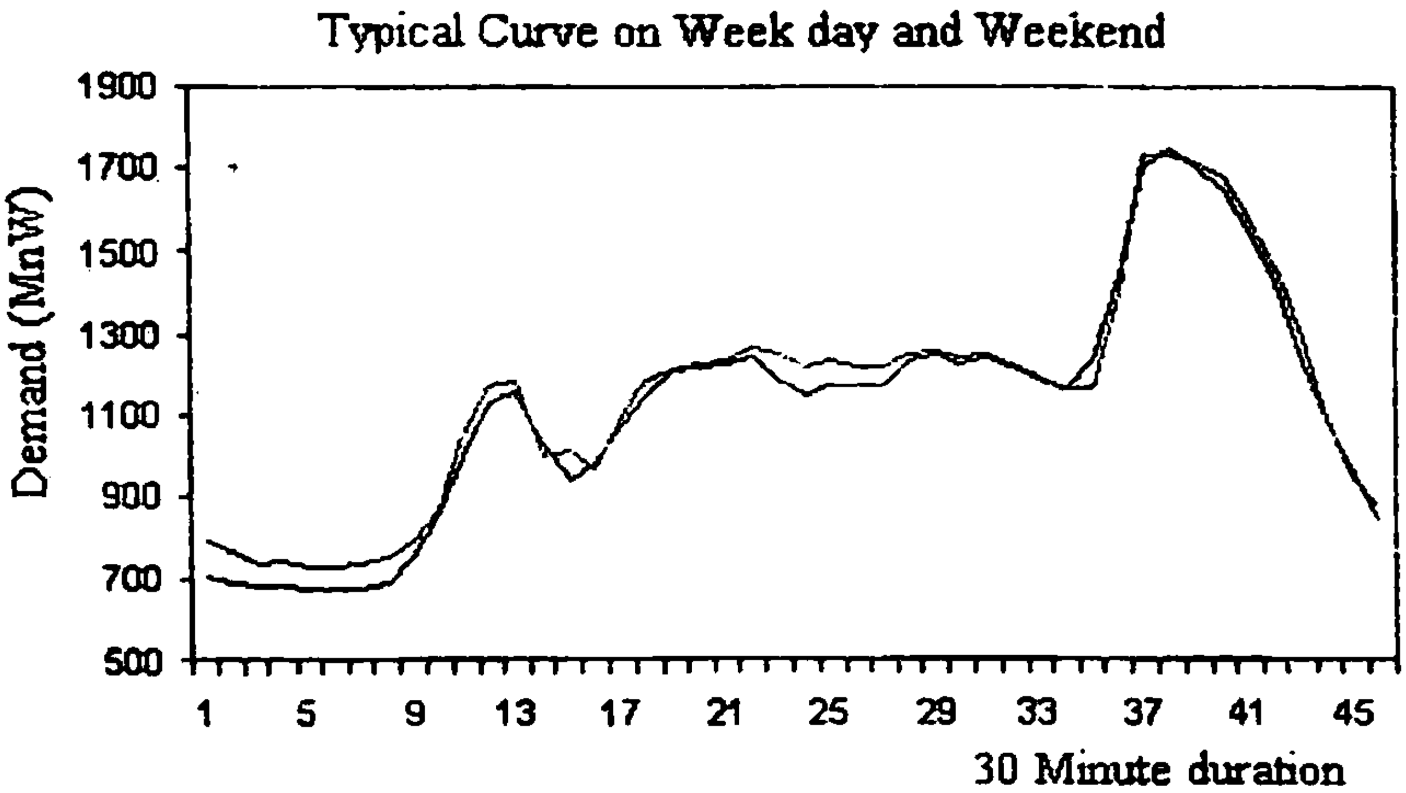

Figure 4: Typical curve of electricity demand in Sri Lanka on weekday and weekend

Figure 4 indicates that there is no significant change of the electricity demand between weekdays or weekends for ½ hour intervals. It can be seen that maximum hourly load is approximately 1700(MnW) and minimum load approximately is 600(MnW) respectively between 9.00 p.m to 10.00 p.m and 1.00 p.m to 2.00 p.m. in each day.

PROPOSED MODEL

The state-space model with fixed coefficients

The fixed-coefficients SS formulation supported by MATLAB:

$$x_{t+1} = \phi x_t + \Gamma u_t + E w_t \tag{7}$$

$$z_t = H x_t + D u_t + C v_t \tag{8}$$

where x_t is a (nx1) vector of state variables, u_t is a (rx1) vector of exogenous variables, z_t is a (mx1) vector of observable variables, w_t and v_t are white noise processes such that:

$$E(w_t) = 0, \quad E(v_t) = 0 \tag{9}$$

$$E\left[\begin{pmatrix} w_{t_1} \\ v_{t_2} \end{pmatrix}\begin{pmatrix} w_{t_1}^T v_{T_2}^t \end{pmatrix}\right] = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix}\delta_{t_1 t_2} \tag{10}$$

being Q and R positive semi-definite matrices. $\phi, \Gamma, E, H, D, C, Q, R$ *and* S are matrices constant over time. The additive structural decomposition of a time series is defined by:

$$z_t = t_t + c_t + s_t + e_t \tag{11}$$

where: t_t is the trend component, representing the long-term behavior of the series, c_t is the transitory component, or cycle, describing short-term fluctuations, s_t is the seasonal component, associated to persistent variability patterns repeated along a season, and e_t is an irregular component. These classes of models are supported by Matlab through the use of the formulation (7) to (12). For more details about these formulation is given by Hannan and Rissanen (1982).

Vector Auto Regressive Moving Averages (VARMA) Models

The VARMA model is defined by:

$$FR(B)FS(B^s)y_t = G(B)u_t + AR(B)AS(B^s)z_t$$

where y_t, u_t, z_t are defined by

$$FR(B) = I + FR_1B^2 + FR_2B^3 + ... + FR(B^p)$$

$$FS(B^s) = I + FS_1B^S + FS_2B^{2S} + ... + FS_p(B^{p_s})$$

$$G(B) = G_0 + G_1B^2 + G_2B^3 + ... + G_p(B^p) \qquad (12)$$

$$AR(B) = I + AR_1B^2 + AR_2B^3 + ... + AR_q(B^q)$$

$$AS(B) = I + AS_1B \; + AS_2B^{2S} + ... + AS_Q(B^{QS})$$

These representation are Important and frequent particular cases of this formulation are the univariate ARMA models and the VARMA models. Standard MATLAB software toolbox uses internally the (State Space) SS representation (5) to (6) for most computations. However, its basic representation is the "THD format" which uses to translate any State Space representation to defined new toolbox.

State Space and Structural Time Series (SSST) Models

The function *ss2thd* obtains the representation of the SS model (5)-(6) in THD format. Its syntax is: $[t_1, d_1, l_1]$ = ss2thd(ϕ, Γ, E, H, D, C, Q, S, R), where the input arguments correspond to the parameter matrices in the standard representation (5)-(10).

Analysis of VARMA Model

The following section provides details about the implementation of composite models in MATALB. The tentative time series model as the form of

$$(1-\alpha_1B-\alpha_2B^2)(1-\varphi_1B^{24})x_t = (1-\beta_1B-\beta_1^2B^2)(1-\delta B^{24})e_t$$

where $\nabla_{24} = (1-B^{24})$

```
load EE.dat;
y=transdif(EE,1,1,1,24);
[theta,din,lab]=arma2thd([0 0],[0],[0],[0 0],0,24);
theta=e4preest(theta,din,y);
```

```
prtmod(theta,din,lab);
[theta,it,lval,g,h]=e4min('lffast',theta,'',din,y);
[std,corrm,varm,Im]=imod(theta,din,y);
prtest(theta,din,lab,y,it,lval,g,h,std,corrm,(toc/60));
[e,vt,wt,ve]=residual(theta,din,y);
titD='residuals from series electricity demand';
descser(e,titD);
plotsers(e,0,titD);
uidents(e,20,titD);
```

```
************************** Model ***************************
VARMA model (innovations model)
1 endogenous v., 0 exogenous v.
Seasonality: 24
Parameters :
FR1(1,1)      -0.9639
FR2(1,1)       0.4483
FS1(1,1)       0.6590
AR1(1,1)      -0.0580
AS1(1,1)       0.1729
AS2(1,1)      -0.1880
V(1,1)      2444.6248
```

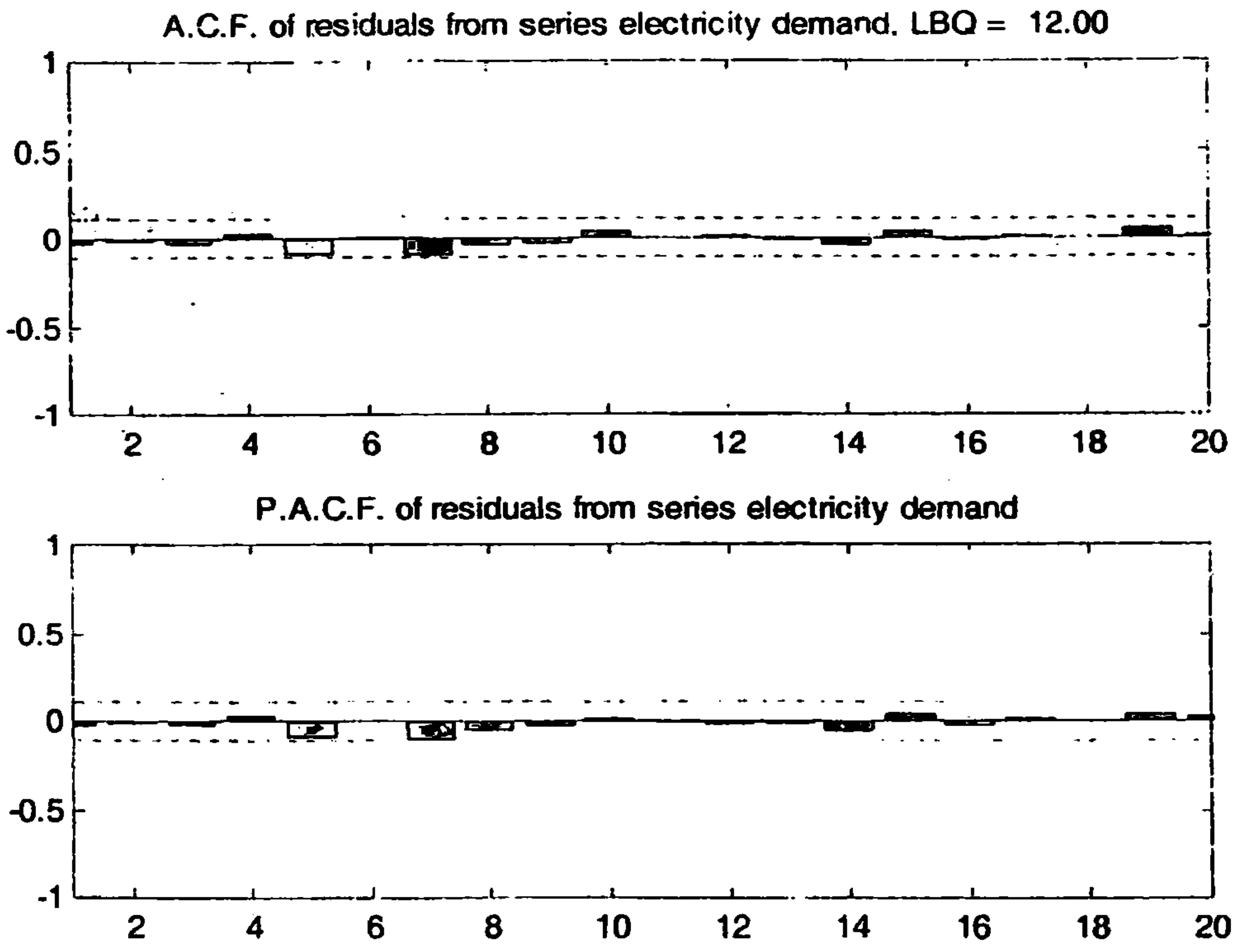

Figure 5: Sample ACF and PACF of the residual from VARMA Model

Based on estimated VARMA model sample autocorrelation function (ACF) and sample partial auto correlations (PACF) calculated (Fig. 5), it indicates that all higher order lag terms of ACF's are not insignificant at 95% level of significance. Further, it was found, that the both AIC and BIC criterion, these two values are minimized at that level of significance. The mean absolute percentage error (MAPE) of the selected model was only 4% for 6-ahead forecast horizon.

Analysis of ARIMA model

To fit an appropriate ARIMA model for the daily day and night peak load forecast, the above discussed three models are used. Those methods are State Space modeling and Box-Jenkins approach for ARIMA Modeling. By using these methods for all the data sets, best fit model was decided. To verify the model, forecasted the values for last three periods using the fitted model. It is also calculated the MAPE for those forecasted values which was 15%. For the fitted model under ARIMA methodology, MINITAB out put of the final estimates of parameters is shown in Table 1.

Table 1: Estimated parameters of the selected model

Type	Coefficients	Std Deviation	t-value	Probability
AR 1	0.5227	0.0796	6.56	0.000
AR 2	-0.3876	0.0792	-4.90	0.000
SAR 24	-0.9598	0.0477	-20.13	0.000
Constant	3.600	3.881	0.93	0.355

Thus final model can be written as

$$(1-\alpha_1 B-\alpha_2 B^2)(1-\varphi_1 B^{24})x_t = e_t \text{ where}$$
$$x_t = 3.6+0.55x_{t-1}-0.38x_{t-2}-0.95x_{t-24} = e_t$$

CONCLUSIONS

To predict the daily load forecasting in Sri Lanka, the developed state space VARMA methodology using MATLAB algorithm is recommended. As the model was developed and validated using observed data for only one month, it is suggested to test the accuracy of this procedure using a longer time-series. Moreover, this study illustrated how customer concerns can be incorporated into the planning stage of power sector development. The models should be designed to complement existing approaches to environmental assessment that are now a routine put of project development.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transaction in Automatic Control,* 19: 716-723.

Anderson, T. (1971). *The Statistical Analysis of Time Series.* Wiley, New York.

Ansley, C.F. and R. Kohn (1985). Estimation, filtering and smoothing in state space models with incompletely specified initial conditions. *Annals of Statistics,* 13: 1286-1316.

Box, O.E.P. and G.M. Jenkins (1976). *Time Series Analysis Forecasting and Control.* Holden-Day, SanFrancisco.

Hannan, E.J. and J. Rissanen (1982). Recursive estimation of mixed autoregressive-moving average order. *Biometrika,* 69: &1-94.

Lutkeohl, H. (1985). Comparison of criteria for estimating the order of a vector autoregressive process. *Journal of Time Series Analysis,* 6: 35-52.

Koreisha, S. and G. Yoshimoto (1991). A comparison among identification procedures for autoregressive-moving average models. *Intern. Statist. Rev.* 59: 37-57.

Koreisha, S. and T. Pukkila (1993). Determining the order of a vector autoregression when the Number of Component Series is Large. *Journal of Time Series Analysis,* 14: 47 69.

Tsay. R.S. and G.C. Tiao (1984). Consistent estimates of autoregressive parameters and extended sample, *J. of Time Series Anal* 5:15-21.

Wei, W.W.S. (1994). *Time series Analysis: Univariate and Multivariate Methods.* Addison-Wisly Publishing Company.