JKAU: Eng. Sci., **vol. 11 No. 2,** pp. 69-79 (1419 A.H. / 1999 A.D.)

Time Series Approach to Hourly Electric Load Forecasting for the Western Province of Saudi Arabia

Farhat Ali Burney, Mohammad Sadiq Al-Jifry and Abu Bakr Esshack
Industrial Engineering Department,
King Abdulaziz University, Jeddah, Saudi Arabia

Abstract. The hourly electric load in the Western Province of Saudi Arabia has four distinct patterns, namely that of summer season, spring season, fall season, and winter season. This paper demonstrates fitting of an auto regressive integrated moving average model to the summer season data using the Box and Jenkins time series approach. The model is tested for forecasting by using a set of test data. The resulting forecasts, which are checked by different criteria, are found to have a high degree of accuracy. The whole exercise has been carried out by using computer programs, which were partly developed.

1. Introduction

Electric load forecasting occupies a central role in the planning and operation of electric power systems[1]. Over a period of time, different attempts have been made to achieve this goal by using a variety of quantitative approaches of which perhaps the most outstanding has been the time series analysis approach[2-7]. This paper reports such an attempt for forecasting the hourly electric load for the Western province of Saudi Arabia. An effort was first made to apply the Dynamic Data System approach[8]. When that didn't succeed, the Box and Jenkins approach was employed[9]. The resulting models were of auto-regressive integrated moving average type. The forecasts developed by these models have a high degree of accuracy as confirmed by a set of different criteria[10-11]. The hourly electric load for the Western Province of Saudi Arabia has four distinct subsets, namely, summer season, spring season, fall season, and winter season. A separate model was developed for each season. This paper reports, for the sake of illustration, the results pertaining to the summer season only[12].

2. The Data

The hourly electric load was analyzed for the Hegra year 1410 as shown in Table 1 and Figures 1 and 2. Furthermore, a typical week's data are plotted in Figures 3-6 for each of the summer, fall, winter, and spring seasons.

TABLE 1. Basic statistical analysis of hourly electric load in MW (Year 1410 H) on a month-to-month basis.

Month	Maximum	Minimum	Mean	STD
01/1410 H	2811	1879	2431	176
02/1410 H	2897	1782	2504	203
03/1410 H	2734	1284	2131	299
04/1410 H	2267	1108	1728	223
05/1410 H	1918	766	1264	246
06/1410 H	1644	721	1125	176
07/1410 H	1670	712	1079	182
08/1410 H	1653	578	1193	194
09/1410 H	2525	905	1721	352
10/1410 H	2641	1136	2073	303
11/1410 H	3165	1881	2575	252
12/1410 H	3132	1840	2606	263

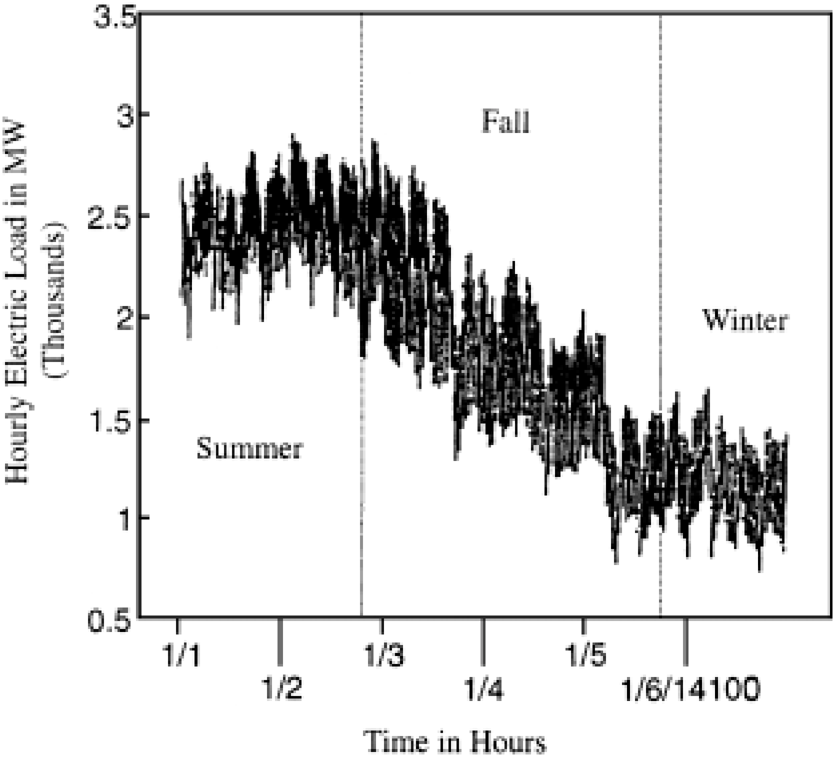

FIG. 1. Hourly electric load in first half of 1410 H year.

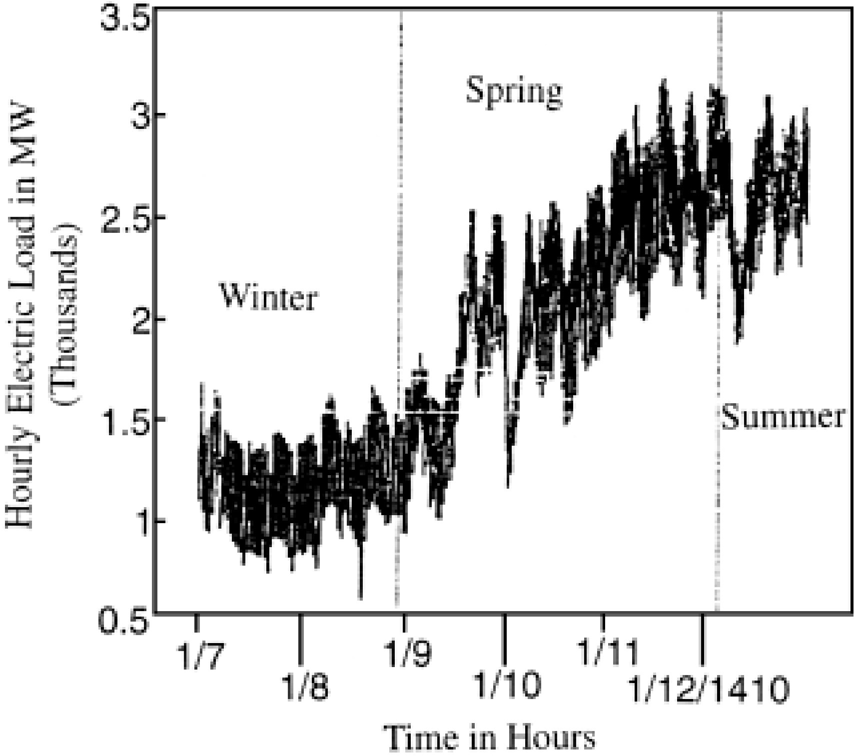

FIG. 2. Hourly electric load in second half of 1410 H year.

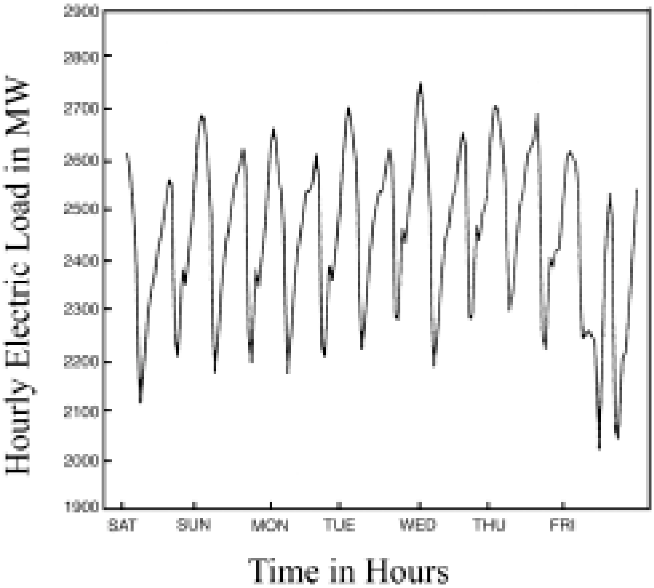

FIG. 3. A typical week in summer.

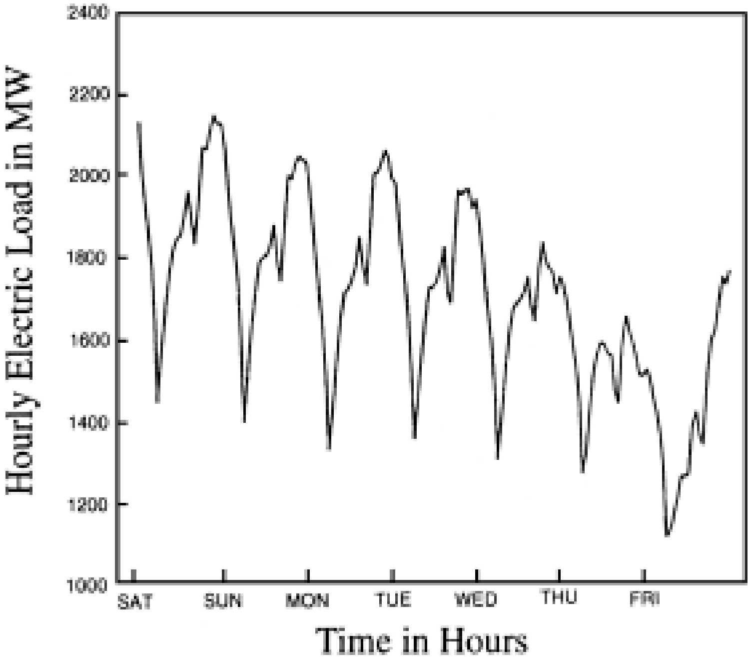

FIG. 4. A typical week in fall.

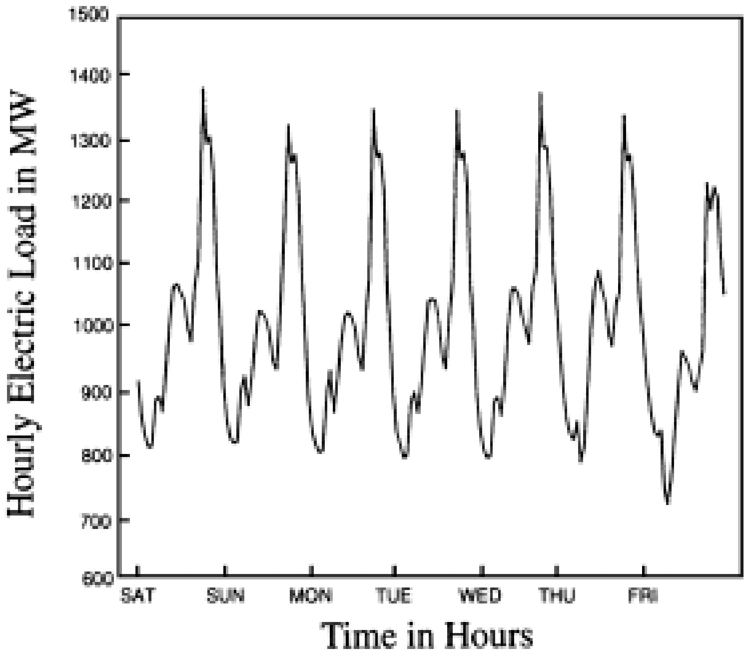

FIG. 5. A tyical week in winter.

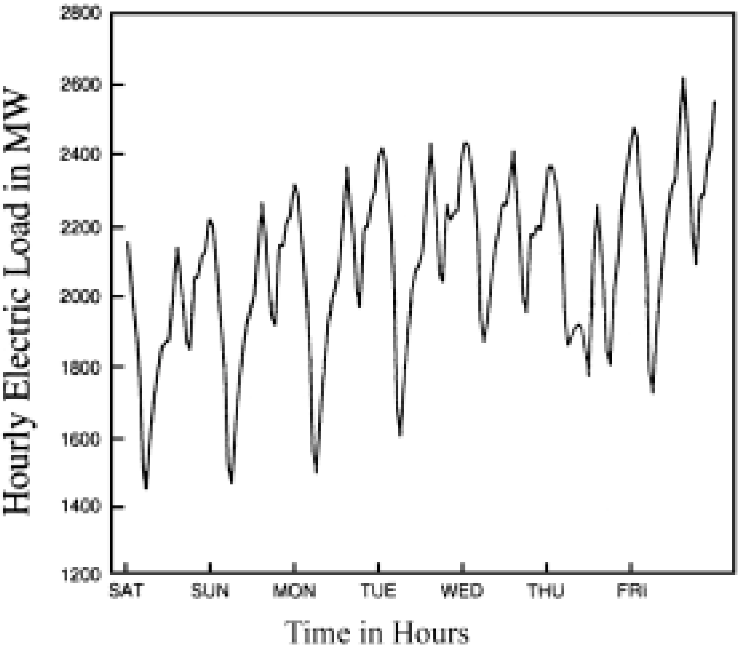

FIG. 6. A typical week in spring.

These amply demonstrate the need of modeling separately for each of the four seasons.

The summer season is divided into two parts: one part in the beginning of the year 1410 H from 1/1/1410 H to 22/2/1410 H and the other at the end of the year from 29/11/1410 H to 30/12/1410 H. The second part is abnormal as it contains the Hajj season. So the first part was chosen for this exercise. Of this, 840 data points of the 5-week period from 4/1/1410 H to 8/2/1410 H were chosen for modeling purposes and 168 data points for the one-week period from 9/2/1410 H to 15/2/1410 H were chosen for testing the forecast. Plot of these data is shown in Figure 7.

3. The Model

Let Z_t, Z_{t-1}, Z_{t-2}, ... represent observations of an equispaced time series in time t, t – 1, t – 2, ... Using B as a back shift operator such that $\beta^j Z_t = Z_{t-j}$ an autoregressive moving average model for this series can be written as:

$$\varphi(B)\,\tilde{Z}_t = \theta(B)\,a_t\ , \tag{1}$$

where $\tilde{Z}t = Zt - \mu$, the deviation of the process from its mean,

$\varphi(B) = 1 - \varphi_1 B - \varphi_2 \beta^2 - ... - \varphi_p \beta^2$ is the autoregressive operator with φ_i, i = 1, 2, ... , p as the autoregressive parameters, $\theta(B) = 1 - \theta_1 B - \theta_2 \beta^2 - ... - \theta_q \beta^q$ is the moving av-

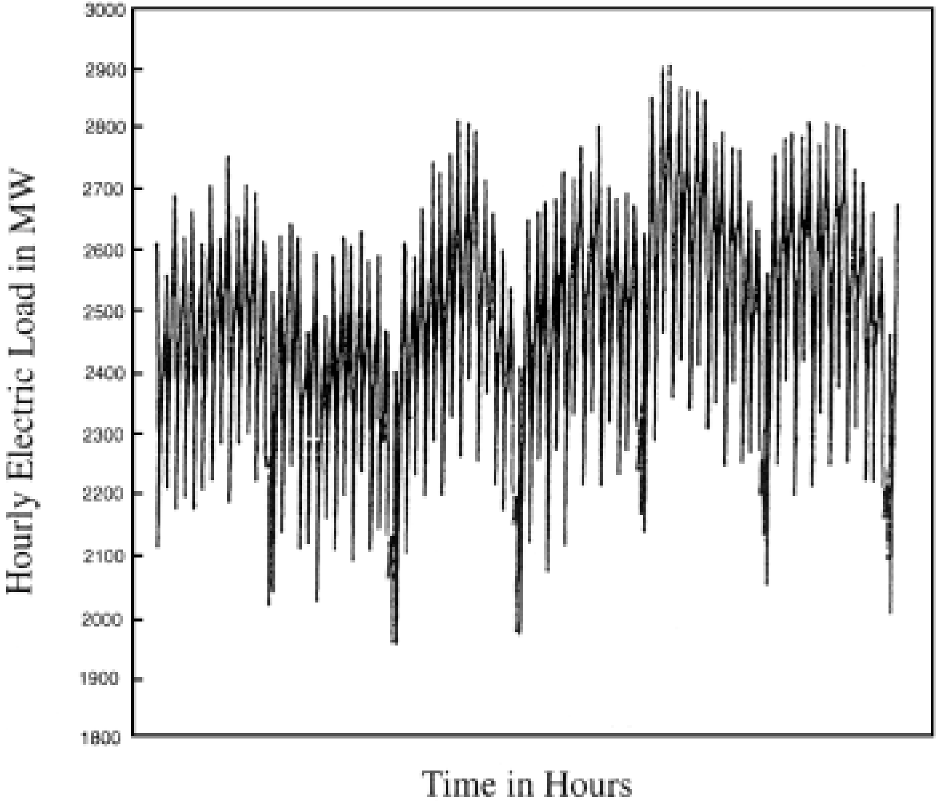

FIG. 7. Selected summer data.

erage operator with θj, j = 1,2, ... , q as the moving average parameters, and a_t = Normally distributed residuals with mean zero and variance σ_a^2.

Introducing the difference operator ∇ such that

$$\nabla \tilde{Z}_t = \tilde{Z}_t - \tilde{Z}_{t-1} ,$$

$$\nabla^2 \tilde{Z}_t = \nabla(\nabla \tilde{Z}_t)\ \nabla(\tilde{Z}_t - \tilde{Z}_{t-1}) = \tilde{Z}_t - 2\tilde{Z}_{t-1} + \tilde{Z}_{t-2} ,$$

and so on, the model now becomes as autoregressive integrated moving average model;

$$\varphi(\beta)\ \nabla^d \tilde{Z}_t = \theta(\beta)\ a_t , \qquad (2)$$

with $\nabla^d = (1 - B)^d$.

Finally, when seasonality is introduced for both the autoregressive and moving average parts the model becomes;

$$\varphi(\beta)\ \Phi(B^s)\ \nabla^d\ \nabla_s^D\ \tilde{Z}_t = \theta(\beta)\ \Theta(B^s)\ a_t , \qquad (3)$$

where $\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps}$ is the seasonal autoregressive operator , with Φ_i, i = 1,2, ... , p as the seasonal autoregressive parameters,

$\nabla_s^D = (1 - B^s)^D$ is the seasonal differencing operator such that $\nabla_s \tilde{Z}_t = \tilde{Z}_{t-s}$,

and

$\Theta(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_q B^{qs}$ is the seasonal moving average operator with Θ_j, j = 1, 2, ... , q as the seasonal moving average parameters.

4. The Data Fitted to Model

The Box and Jenkins approach to modeling consists of three stages, namely, identification, parameter estimation, and diagnostic checking[9].

The following briefly describes how these three stages were gone through interactively to arrive at a proper model for the data.

The autocorrelation plot of the data showed an oscillating pattern with a number of significant autocorrelations indicating the need for differencing. After a number of trials the hourly-weekly basis was found to be the best differencing method to achieve stationarity. Next, the partial autocorrelation plot of the differenced data was made. Since it showed two significant spikes, the following model was tentatively fitted.

$$(1 - \beta)\ (1 - \beta^{168})\ (1 - \varphi_1 B^{24} - \varphi_2 B^{48})\ \tilde{Z}_t = a_t\ . \tag{4}$$

The autocorrelation plot of the residuals shows a significant value at lag 168. Since the trial and error runs showed best fit for the moving average factor, the following model was tried;

$$(1 - \beta)\ (1 - \beta^{168})\ (1 - \varphi_1 B^{24} - \varphi_2 B^{48})\ Z_t = (1 - \theta_1 \beta^{168})\ a_t\ . \tag{5}$$

The autocorrelation plot of the residuals showed two significant values at lags 46 and 48. Including these caused the model to be unstable and inadequate. Hence nonseasonal factors were introduced and values of p - 1 and q = 1 were found to be appropriate. Thus the model turned out to be as follows:

$$\begin{aligned}&(1 - \beta)\ (1 - \beta^{168})\ (1 - \varphi_1 \beta)\ (1 - \varphi_2 B^{24})\ (1 - \varphi_3 \beta^{48})\ \tilde{Z}_t = (1 - \theta_1 \beta)\\ &(1 - \theta_2 \beta^{46})\ (1 - \theta_3 \beta^{48})\ (1 - \theta_4 \beta^{168})\ a_t\ .\end{aligned} \tag{6}$$

The model was now fitted to the data. The parameter estimates and their confidence intervals are shown in Table 2. The chi-square test of adequency results are shown in Table 3. Furthermore, the autocorrelations and partial autocorrelations of the residuals were also plotted. The results indicated:

a) no significant autocorrelation or partial autocorrelation,

b) the confidence intervals of the parameters do not include zero, and

c) the model passes the chi-square test of adequancy.

Hence the final model adopted was:

$$(1 - B)\ (1 - B^{168})\ (1 - .97127B)\ (1 - .31B^{24})\ (1 - .99489\ B^{48})\ \tilde{Z}_t$$

TABLE 2. Parameters estimates and their confidence intervals.

Parameter	Lag	Estimate	Standard error	Upper 95% conf. limit	Lower 95% conf. limit	Parameter Ok?
MA1, 1	1	0.99807	0.0028103	1.003578188	0.992561812	Yes
MA2, 1	46	0.1427	0.0399	0.220904	0.064496	Yes
MA3, 1	48	0.71397	0.05978	0.8311388	0.5968012	Yes
MA4, 1	168	0.50906	0.04105	0.589518	0.428602	Yes
AR1, 1	1	0.97127	0.01138	0.9935748	0.94895652	Yes
AR2, 1	24	0.31	0.04102	0.3903992	0.2296008	Yes
AR3, 1	48	0.99489	0.03964	1.0725844	0.9171956	Yes

Table 3. Chi-Square test on the summer model.

To Lag	Autocorrelations					
6	0.06185	– 0.0155	– 0.04104	– 0.03195	– 0.01609	– 0.044885
12	– 0.01868	– 0.00258	– 0.03152	– 0.0496	0.008	– 0.0255
18	0.00706	– 0.00081	0.0226	– 0.01855	0.04021	0.00524
24	0.06145	0.01706	– 0.01501	0.03697	0.09092	– 0.002

To Lag	Squared autocorrelations / (n – k)					
6	5.70959E-06	3.59118-07	2.52138E-06	1.53044E-06	3.88721E-07	3.58845E-06
12	5.22516E-07	1.00398E-08	1.50077E-06	3.7288E-06	9.69697-08	9.86722E-07
18	7.57502E-08	9.9863E-10	7.78598E-07	5.25347E-07	2.47224E-06	4.20484E-08
24	5.79157E-06	4.44072E-06	3.47716E-07	2.10598E-06	1.27569E-05	6.18238E-11

To Lag	DF	Calculated Chi-Square	Tabulated Chi-Square	Is Calc. < Tabu.?	Conclusion
6	0	–	–		
12	5	9.455963894	11.07	Yes	
18	11	11.21487099	19.68	Yes	
24	17	20.90049485	27.59	Yes	Model is adequate

$$= (1 - .99807B) (1 - .1427 B^{46}) (1 - .7139 B^{48}) (1 - .50906 B^{168}) a_t \ . \qquad (7)$$

5. Forecasts and their Validity

Using the test data (168 points) three different forecasts were made, namely one hour ahead, 12 hours ahead, and 24 hours ahead. To check the forecast validity, the mean absolute percentage error, MAPE, was calculated. As shown in Table 4 the values of

MAPE are .67%, 1.03% and 1.14% which indicate high reliability of the forecasts. This is further confirmed by the comparative graph of the 24 hours ahead forecasts vs actual data as shown in Figure 8. This graph shows how closely the forecast figures follow the actual values.

TABLE 4. Forecasting performance indicators of summer model.

Indicator	Description	Data set – hours ahead			
		Fit – 1	Test – 1	Test – 12	Test – 24
SSE	Sum squared error	310556	83926.8	90232	204499
MSE	Mean squared error	462.826	499.564	1132.33	1217.25
RMSE	Root mean squared error	21.5134	22.3509	33.6501	34.8891
MAE	Mean absolute error	16.2702	16.6685	25.444	28.4507
MAPE	Mean absolute percentage error	0.66%	0.67%	1.03%	1.14%

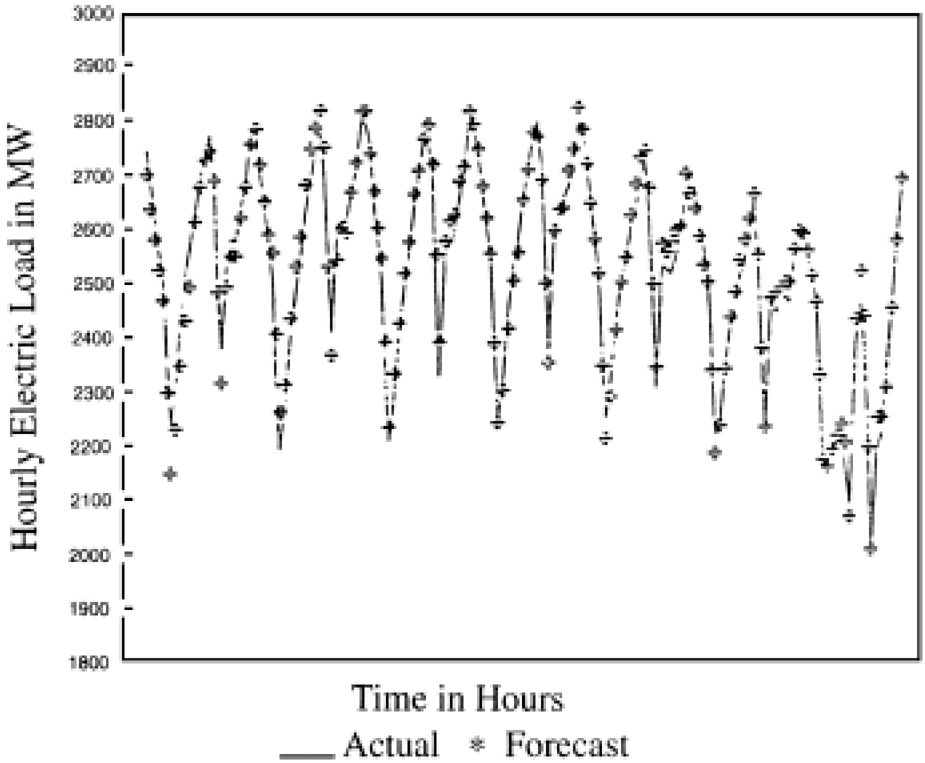

FIG. 8. Actual loads vs forecasts in summer season.

Conclusion

This paper has demonstrated the application of Box and Jenkins approach to build a time series model for a set of hourly electric load data for the Western Province of Saudi Arabia. Following the standard strategy of identification, parameter estimation and diagnostic checking, a seasonal autoregressive integrated moving average model is fitted to a set of data comprising of 840 points. Another set of 168 data points are then used for testing the one hour ahead, 12 hours ahead, and 24 hours ahead forecasts. The forecasts are found to have high accuracy. The whole exercise has been carried out by using computer programs, which were partly developed.

References

[1] **Bunn, D.W.** and **Farmer, E.D. (editors),** *Comparative Models for Electric Load Forecasting,* John Wiley and Sons Inc., N.Y., USA (1985).

[2] **Hagan, M.** and **Klein, R.,** Off-line and Adaptive Box and Jenkins Models for Load Forecasting, *Proceedings of Lawrence Symposium on Systems and Decision Sciences,* pp. 221-225, Oct. (1977).

[3] **Vahakyla, P.,** Short Term Forecasting of Grid Load using Box and Jenkins Techniques, *Electrical Power and Energy System,* **2**(1): 29-34 (1980).

[4] **Goh, T.N., Ong, H.L.** and **Lee, Y.O.,** A Study on the Accuracy of Time Series Modeling of Daily Power Demand, *Electric Power Systems Research,* **9:** 73-77 (1985).

[5] **Hagan, M.T.** and **Behr, S.M.,** The Time Series Approach to Short Term Load Forecasting, *IEEE Transactions on Power Systems,* vol. **pwrs-2,** No. 3, pp. 785-79 (1987).

[6] **Rajrukar, K.P.** and **Nissan, J.L.,** Data Dependent System Approach to Short Term Load Forecasting, *IEEE Transactions on Systematic Manufacturing Cybernatics,* **SMC-15**(4): 532-536 (1985).

[7] **Uri, N.D.,** Forecasting Peak System Load Using a Combined Time Series and Econometric Model, *Applied Energy,* **4**(3): 219-227 (1977).

[8] **Pandit, S.M.** and **Wu, S.M.,** *Time Series and System Analysis with Applications,* John Wiley & Sons Inc., N.Y., USA (1983).

[9] **Box, G.E.** and **Jenkins, G.M.,** *Time Series Analysis Forecasting and Control,* Prentice-Hall, Englewood Cliffs, New Jersey, U.S.A. (1976).

[10] **Makridakis, S.** and **Wheelwright, S.C.,** *Forecasting: Methods and Applications,* John Wiley and Sons Inc., N.Y., USA (1978).

[11] **Thompolus, N.T.,** *Applied Forecasting Methods,* Prentice-Hall, Englewood Cliffs, New Jersey, U.S.A. (1980).

[12] **Esshack, A.B.A.R.,** *Short Term Electric Load Forecasting in Western Region of Saudi Arabia by Time Series Method,* a unpublished master's thesis, Department of Industrial Engineering, College of Engineering, King Abdulaziz University, Jeddah, Saudi Arabia (1994).

توقـع الحمـل الكهربـائي لكل ساعـة في المنطقـة الغربيـة للمملكة العربية السعودية باستخـدام المتواليـات الزمنيـة

فرحات علي برني ، محمد الصادق الجفري و أبو بكر عبد الرزاق اسحاق

قسم الهندسة الصناعية ، كلية الهندسة ، جامعة الملك عبدالعزيز

جـــدة – المملكة العربية السعودية

المستخلص . دراسة القراءات الساعية للحمل الكهربائي في المنطقة الغربية للملكة العربية السعودية تظهر أربعة أنماط متمايزة عن بعضها البعض ، وهي علي وجه التحديد تتبع المواسم الفصلية الأربعة : الصيف والربيع والخريف والشتاء . هذا البحث يعرض كيفية تطبيق نموذج توقع من النوع المتراجع ذاتيًا المتكامل ذي المتوسط المتحرك علي بيانات فصل الصيف باستخدام طريقة بوكس وجنكتر لتحليل المتواليات الزمنية ، وقد تم إختبار قدرة النموذج على التوقع باستخدام مجموعة مستقلة من البيانات . وعند تقييم التوقعات الناتجة باستخدام معايير إحصائية متعددة وجد أن هذه التوقعات على مستوى جيد من الدقة ، وقد استخدم خلال هذا البحث عدد من برامج الحاسب الآلي التي تم تطوير بعضها خصيصًا لهذا الغرض .