Alexandria Engineering Journal (2011) 50, 137–144

FACULTY OF ENGINEERING ALEXANDRIA UNIVERSITY

Alexandria University

Alexandria Engineering Journal

www.elsevier.com/locate/aej
www.sciencedirect.com

ORIGINAL ARTICLE

A methodology for Electric Power Load Forecasting

Eisa Almeshaiei [a], Hassan Soltan [b,*]

[a] *Production Engineering Department, College of Technological Studies, Public Authority for Applied Education and Training, Sheiwck, Kuwait*
[b] *Production Engineering and Mechanical Design Department, Faculty of Engineering, Mansoura University, Mansoura 35526, Egypt*

Received 21 January 2010; accepted 26 January 2011
Available online 27 July 2011

KEYWORDS
Electric Power Load Forecasting;
Time series;
Pattern segmentation/decomposition

Abstract Electricity demand forecasting is a central and integral process for planning periodical operations and facility expansion in the electricity sector. Demand pattern is almost very complex due to the deregulation of energy markets. Therefore, finding an appropriate forecasting model for a specific electricity network is not an easy task. Although many forecasting methods were developed, none can be generalized for all demand patterns. Therefore, this paper presents a pragmatic methodology that can be used as a guide to construct Electric Power Load Forecasting models. This methodology is mainly based on decomposition and segmentation of the load time series. Several statistical analyses are involved to study the load features and forecasting precision such as moving average and probability plots of load noise. Real daily load data from Kuwaiti electric network are used as a case study. Some results are reported to guide forecasting future needs of this network.

1. Introduction

Electricity as a product has very different characteristics compared to a material product. For instance, electricity energy cannot be stored as it should be generated as soon as it is demanded. Any commercial electric power company has several strategic objectives. One of these objectives is to provide end users (market demands) with safe and stable electricity. Therefore, Electric Power Load Forecasting (EPLF) is a vital process in the planning of electricity industry and the operation of electric power systems. Accurate forecasts lead to substantial savings in operating and maintenance costs, increased reliability of power supply and delivery system, and correct decisions for future development. Electricity demand is assessed by accumulating the consumption periodically; it

* Corresponding author.
E-mail addresses: eamboabood@yahoo.com (E. Almeshaiei), hasoltan@hotmail.com (H. Soltan).

Peer review under responsibility of Faculty of Engineering, Alexandria University.
doi:10.1016/j.aej.2011.01.015

ELSEVIER Production and hosting by Elsevier

is almost considered for hourly, daily, weekly, monthly, and yearly periods.

The EPLF is classified in terms of the planning horizon's duration: up to 1 day/week ahead for *short-term*, 1 day/week to 1 year ahead for *medium-term*, and more than 1 year ahead for *long-term* [2,19]. Short-term forecasts are used to schedule the generation and transmission of electricity. Medium-term forecasts are used to schedule the fuel purchases. Long-term forecasts are used to develop the power supply and delivery system (generation units, transmission system, and distribution system). Also, refer to Weron [29] and Pedregal and Trapero [19].

The electricity demand pattern is necessarily affected by several factors including time, social, economical, and environmental factors by which the pattern will form various complex variations [1,10]. Social (such as behavior) and environmental factors are big sources of randomness (noise) found on the load pattern. Diversity and complexity in demand pattern have been leading to developing complicated EPLF methods. The literature is enriched with EPLF methods having many attempts to find the best estimation of load forecasting. The major methods include time series such as exponential smoothing, ARMA, Box–Jenkins ARIMA, regression, and transfer function (dynamic regression); expert systems; neural networks; Fuzzy logic; and support vector machine [4,20,21,27,29,3,10,18,25]. Recently, ant colony optimization is applied to EPLF [15]. Furthermore, there are many review papers on EPLF methods such as Alfares and Nazeeruddin [2], Singh and Singh [22], Metaxiotis et al. [13], Taylor et al. [24], and Hahn et al. [9].

The ARIMA models and their versions have achieved a considerable success for EPLF (see [7,29], and [16]. In general, ARIMA models can be used when the time series is stationary without missing data [8]. They can be further hybridized with artificial intelligence techniques [26].

However, the complexity of demand pattern depends on its base period; it changes from fairly smooth curve (annually based) to most noisy and cyclic complex curve (hourly based) since the effect of environmental factors increases. Combined forecasting was also introduced based on a certain linear combination of various results from different forecasting methods [5,30,31,28]. Hybrid methods were also developed for EPLF such as Niu et al. [14,15]. Also, see Weron [29] for an extensive review of EPLF and its applications.

In this era the electric power consumption is growing fast and may be more randomly because of the increasing effect of environmental and human behavior. Therefore, the electricity demand pattern becomes more complex and unrecognized. For instance, the people all over the world are using increased number and variety of electric appliances most of them are environmentally related, that increases the cyclic variation and noise on the demand pattern. Though there are many forecasting methods, no single one can be generalized to perform enough for all cases, especially when many factors are considered. Thus, to get a proper forecast, it is not just adopting a famous method. In other words, an ideal method for a case may perform poorly for another one. Therefore, the research must be directed to specially assigned methods. In other words, each electric power plant in any country needs to follow its own EPLF method. (That is the direction of this paper.) For that purpose, the general methods can be also adopted but with efficient and effective modifications that suit the case; otherwise the results will be misleading. For instance, Ohtsuka et al. [16] have considered an ARMA process as a time series structure, for their special model, because it is suitable for the Japanese load data as they stated.

The aim of this paper is to demonstrate a pragmatic EPLF methodology for analyzing the electric load pattern and predicting the future load demand for short, medium, and/or long terms. This methodology can integrate different forecasting models. The rest of the paper is organized as follows. Section 2 is a preliminary for the proposed methodology which is presented in Section 3. Section 4 applies the methodology to a typical power load pattern. Concluding remarks are contained in Section 5.

2. Preliminaries

Time series forecasting methods and their modifications (e.g. combinations and hybridizations) are the most popular methods for forecasting in general. De Gooijer and Hyndman [7] have introduced a magnificent general review for time series forecasting. Electricity load pattern is principally a time series. Time series analysis is normally the main method used in the field of EPLF. As any time series, the load pattern can be modeled as a function of different factors [17,29,19]. Such function should comprise several components; each component can meet one or more factors (time, social, economical, or environmental factor). It has been suggested in the literature that time series of electricity load may contain a trend component [3]. This could take the form of either a deterministic or a stochastic trend. Such functional relationships can assume different forms such as additive, multiplicative, or mixed forms. An additive function and a multiplicative function for electricity load were presented by Chen et al. [6] as

$$L(t) = L_n(t) + L_w(t) + L_s(t) + L_r(t) \tag{1}$$

and

$$L(t) = L_n(t)F_w(t)F_s(t)F_r(t) \tag{2}$$

where $L(t)$ is the total load at time t; $L_n(t)$ is the normal or trend component, which is a set of standardized load shapes; $L_w(t)$ is the weather sensitive component; $L_s(t)$ is a special event component that create a substantial deviation from the usual load pattern; and $L_r(t)$ is a completely random term (the noise). The F's are positive correction factors justified in the same way.Thus, to find a proper forecast of future electricity load, the behavior of the given pattern must be studied very well. For instance, when the pattern is a multi-component time series, the impact of each component on the time series must be identified especially the noise component. It is obvious that each component itself is a time series. This necessitates decomposition of the time series in some way to identify and extract the partial loads (components) that superimposed to the main load time series and to each other.

The classical decomposition of time series is often based on four components: trend, cyclic, seasonal, and random components. The seasonal component is viewed as a periodic component with a fixed cycling period corresponding to the individual seasons. In practice, it is convenient to combine the trend and the cyclical components into a *trend-cycle* component to reduce to three components time series. Decomposition analysis is often used for *detrending* and *deseasonalizing* the time series. Review Palit and Popovic [17]. The same was frequently followed to analyze the electricity load [29]. Since the load time series may contain many more components, the decomposition task becomes very hard. However, for complex load pattern, the decomposition process may fail to find proper forecasts.

The global models for EPLF often fail to forecast near accurate loads due to the nature of time series of electricity load. Strictly, the load pattern is very complicated non-stationary time series. Therefore, the principle of partitioning the load time series into enough homogenous regions was introduced in the literature aiming at improving the accuracy of load forecasting [23,11,12]. After partitioning, each region can be analyzed separately and the relationship between those regions should be identified. In other words, a different model can be assigned to each region with monitoring the behavior of model parameters relative to other models.

3. The proposed EPLF methodology

This section presents the proposed EPLF methodology for forecasting future load without the restriction to the term length (short, medium, or long). This methodology is mainly based on the principles of time series *segmentation and decomposition*. Some additional statistical analysis will be followed to aid the decision making based on the adopted forecasts such as *probability plots*. The methodology can be presented through the flowchart shown in Fig. 1. The main phases of the methodology can be summarized as follows.

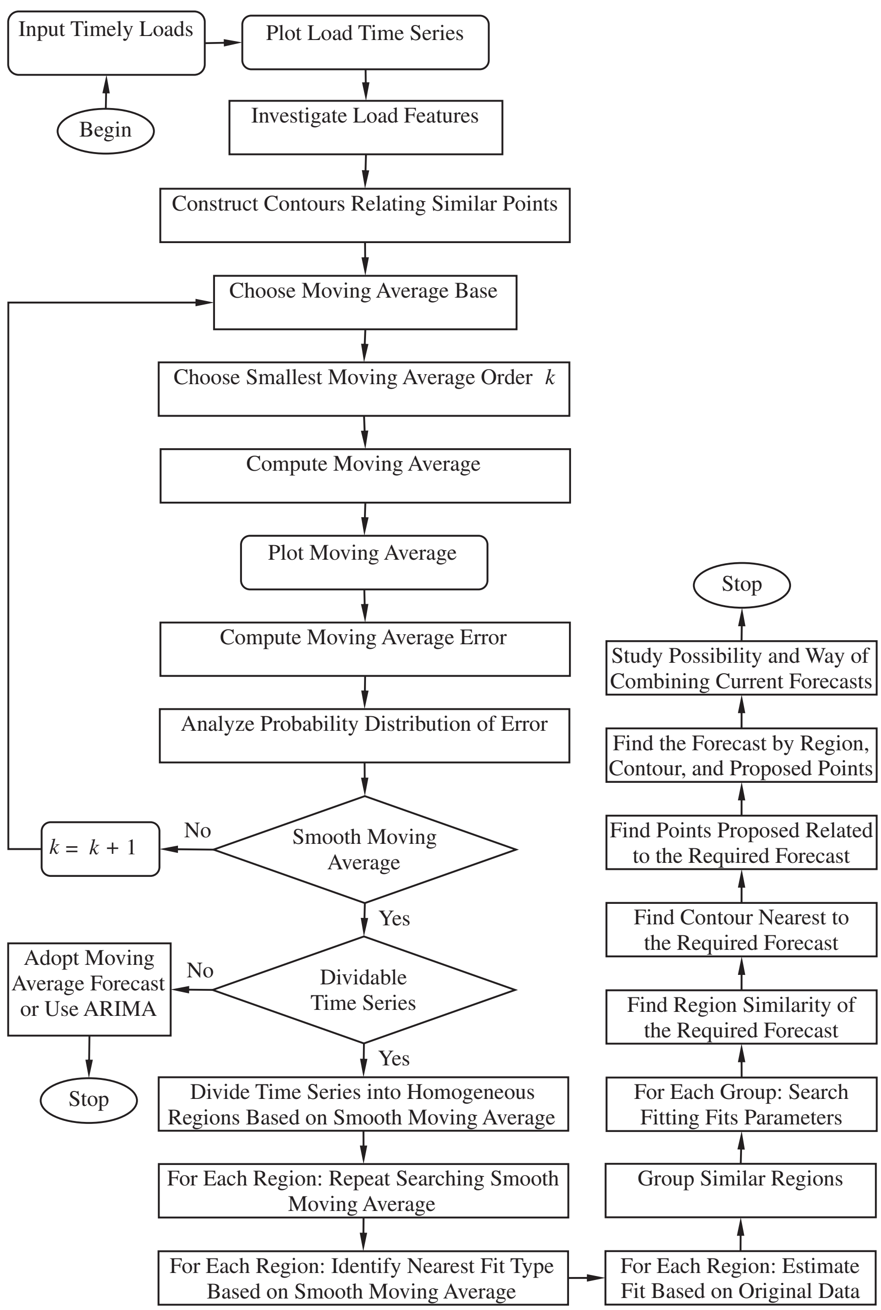

Figure 1 Flowchart for the proposed EPLF methodology.

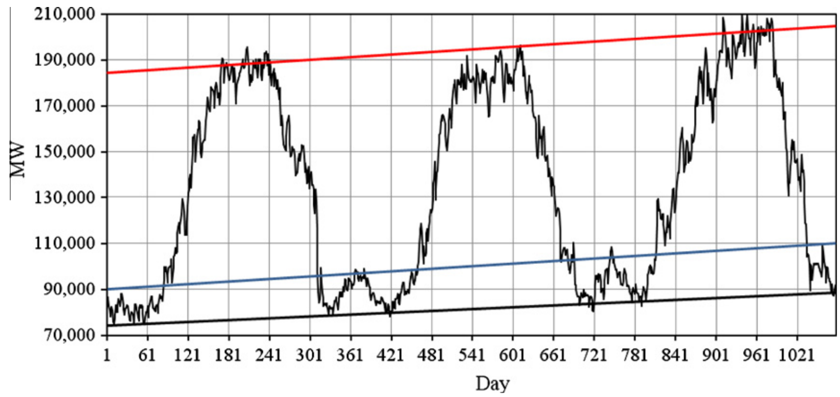

Figure 2 Daily load in Kuwaiti network during 2006–2008.

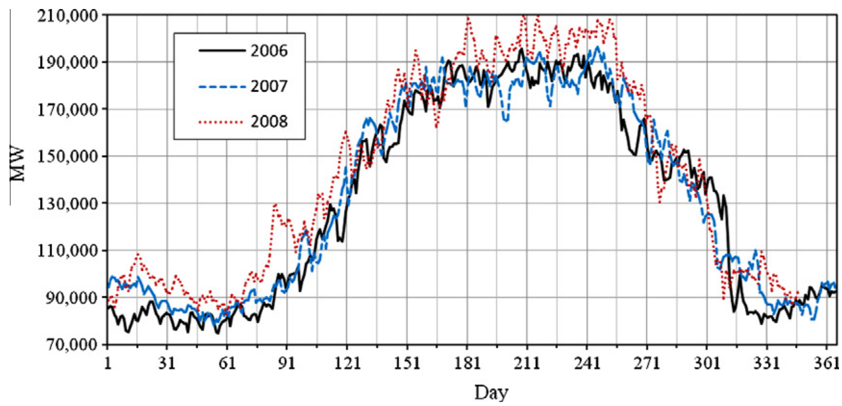

Figure 3 Daily meshed plot for load during 2006–2008.

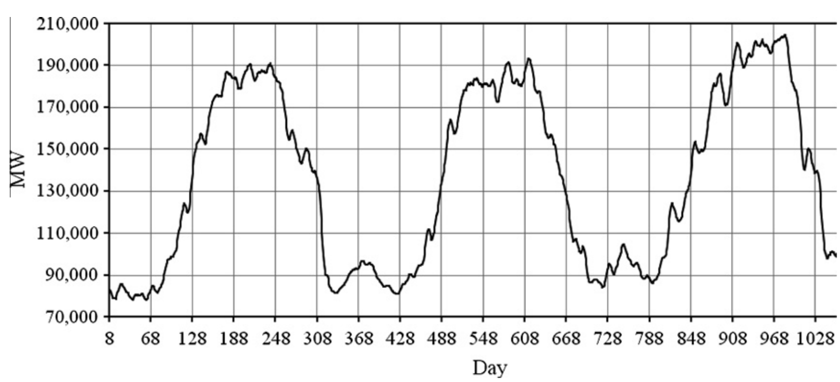

Figure 4 The 7-day moving average for load during 2006–2008.

3.1. Primal visual and descriptive statistical analysis

This gives a primal conclusion about the behavior of the features of load time series.

3.2. Contour construction

This locates similar points along the time series around a single trend line, if that is available.

3.3. Load pattern decomposition

This is based on only two components; the noise superimposed on the main load component. Moving average, exponential smoothing, or ARMA/ARIMA models can be used for that purpose. Here, moving average is proposed. The main component is expected to be smooth. Therefore, moving average order is experimented. The smoothed moving average leaves a noise component with smallest mean value

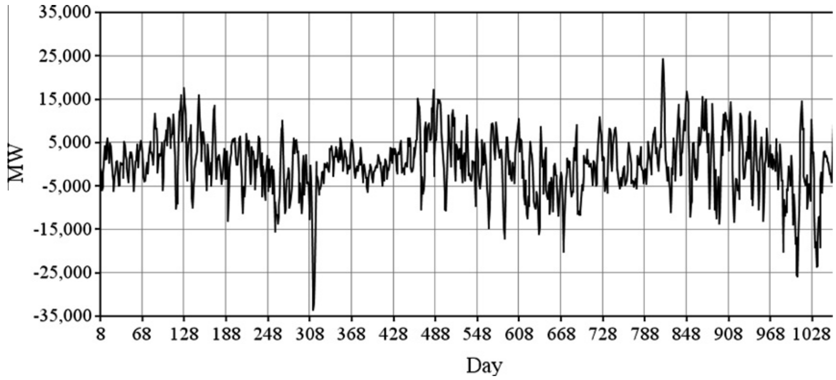

Figure 5 Noise based on 7-day moving average for load during 2006–2008.

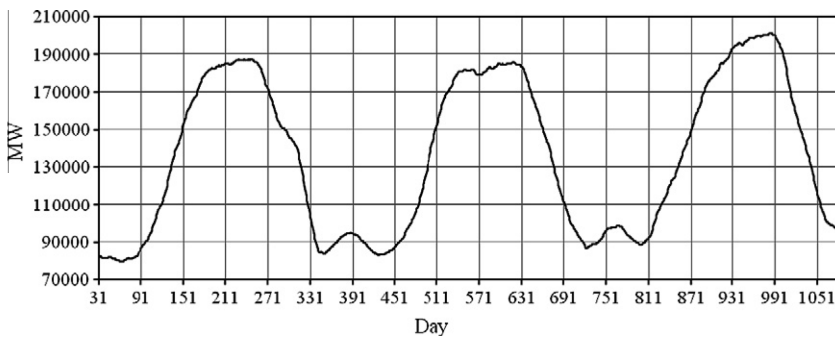

Figure 6 The 30-day moving average for load during 2006–2008.

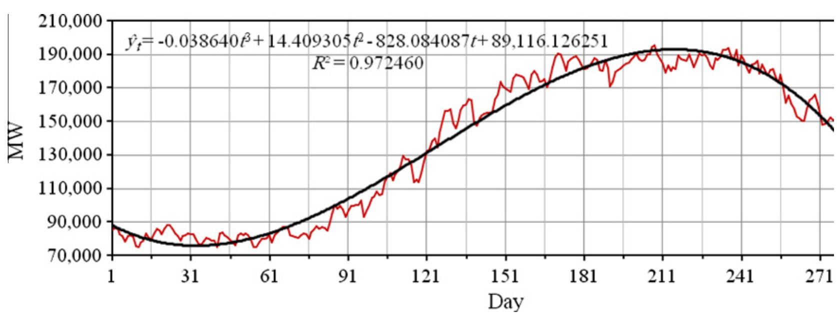

Figure 7 Trend of the first region of year 2006.

(near to zero) and standard deviation greater than zero. The decomposition process is carried out for original time series and each homogeneous region. It is proposed here that the moving average base can be a suitable measure of central tendency. The order of moving average should be chosen carefully.

3.4. Load pattern segmentation

This is based on the decomposition of original time series into two components to identify segments (regions) of time series. Each region is identified by a set of points demonstrates some homogeneity (approximate a trend). This can be done using *analysis of variance* or other methods such as *self organizing maps* [11,12]. Furthermore, this process can be done visually or by using simpler statistics. Notice that, in rare cases, the original time series itself may be found homogeneous.

3.5. Future load forecasting

This can be achieved aided by one or more of three bases – region similarity, contour, and proposed related points – and their combined forecast.

Curve fitting is conducted to each region. The parameters of resulting trends are studied to find if there is a correlation between the parameters of some regions. The regions that have correlation between their parameters are considered similar and assigned into a group. Hence, we can forecast the growth of these parameters. The effect of each region, contour, and proposed related point on the required forecast can be

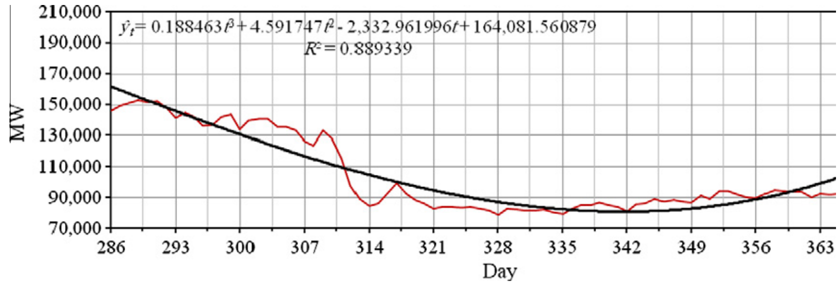

Figure 8 Trend of the second region of year 2006.

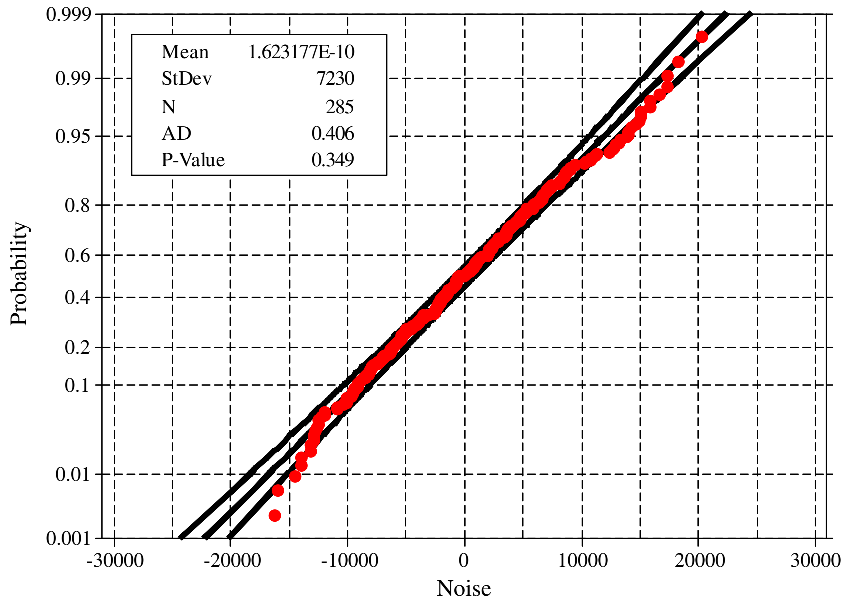

Figure 9 Probability plot for noise of first region of year 2006.

weighted. Thus, the process of estimating a forecast can be aided with a weighting decision support system. Other similar systems such as ranking systems can also be used.

The performance of any forecasting methodology can be evaluated using several proposed statistical metrics (see [29] and [17]. Most common and reasonable metric is the *mean absolute percentage error* that is

$$\text{MAPE} + \frac{1}{n}\sum_{i=1}^{n}\left(\frac{|y(t) - \hat{y}(t)|}{y(t)}\right) \tag{3}$$

where $y(t)$ and $\hat{y}(t)$ are the estimated load and actual load at time t; and n is the number of data used. Less MAPE means more precise estimates. However, the performance metric suitability depends on the features of the load time series.

4. Analysis of electric load case

The proposed methodology is basically a practical guide to design specific EPLF models for forecasting future electricity consumption. As cited before, an accurate forecast can not be obtained blindly through global methods. Thus, this methodology is case dependent. For demonstration, the case of daily loads data of Kuwaiti electric network during 2006–2008 is analyzed. These data are obtained from the actual records in MWh. Original time series of the load is depicted in Fig. 2. Excel 2007 and MINITAB 14 will be used as aids for analysis. It is obvious that the load shows repetitive cyclic features for nearly equal lengths. Three main types of *contours* can be drawn. The upper, middle, and lower contours are approximately increasing straight lines. These contours can be used to forecast some special day's loads. Also, other contours can be drawn.

Fig. 3 illustrates the meshed plot for the load during the three years. This gives sight about the growth of daily loads, which increases in most of days. This plot can be used for load weighting. The simple moving average is used for data filtering. It is found that the 7-day moving average (Fig. 4) is the nearest smooth one with mean error about 30.55 MW and MAPE of 0.0384. This error represents the noise (Fig. 5) based on moving average. Test of hypothesis proved that the current noise can not be said normally distributed with having the apparent odd values. This moving average seems sufficient to discover the regional homogeneity of the time series. Further smoothing will not yield the actual noise because of starting to distort the main load. For instance, 30-day moving average (Fig. 6) gives more information about homogeneity with mean

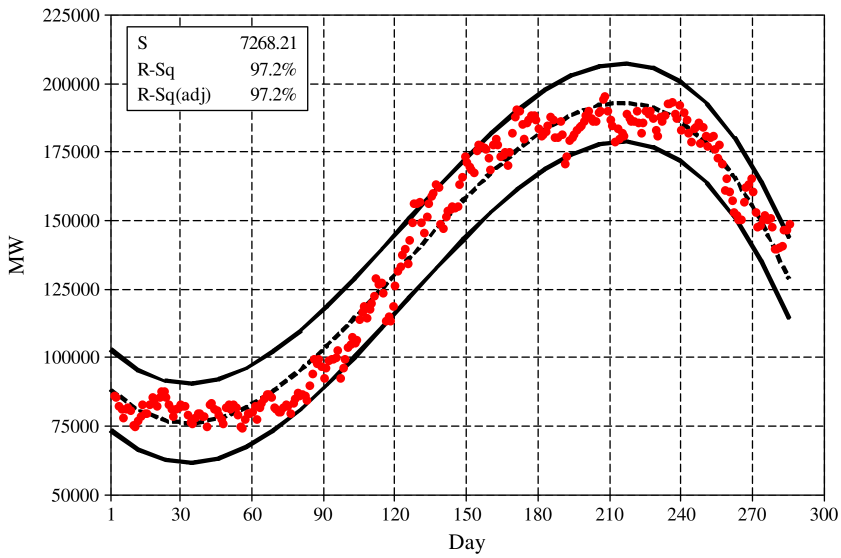

Figure 10 Fitted polynomial plot with confidence 95% for first region of year 2006.

error of 174.47 MW, while this error can't be considered as noise. As shown later, the segmentation process will yield normally distributed noise with mean about zero.

Referring to Figs. 4 and 6, the load time series of Kuwaiti network for the given period can be divided into six regions, each year into two regions. The first region of each year is around 285 days (about 9.5 months) horizon. All regions can be fitted to third order polynomial as shown for instance in Figs. 7 and 8 of year 2006. The first regions (first group) have a specific similarity and the second regions (second group) have another similarity. In other words, for each group, each polynomial parameter demonstrates a trend. Probability plots, as that shown in Fig. 9, are constructed to test the hypothesis of the noise probability distribution based on the estimated trend polynomials. The noise along the time series can be assumed normally distributed with mean about zero. Therefore, a forecast estimate can be obtained with a specific confidence level, as shown in Fig. 10. Fortunately, in this case study, each region can be used to forecast a daily future load by applying slight modification to the trend parameters.

5. Concluding remarks

Electricity demand forecasting represents the main task in the planning of electricity production because it determines the required resources to operate the electricity plants such as daily consumption of fuels. Furthermore, it is the corner stone of planning for electric plants and networks. The literature reports that the electric load pattern is very complex. It is therefore necessary to develop new methods for EPLF to reduce the uncertainty of the predictions. This paper has reported that every electric network and plant needs to its special forecasting method because each country is indifferent in the factors that affect the electricity demand. In the developing countries the electric demand grows with dynamic and high growth rate.

The methodology of this paper is developed as a guide for constructing specialized EPLF models. It proposes three bases – *region similarity*, *contour*, and *proposed related points* – and their *combined forecasts*. Forecasting analysis is conducted to the daily load time series of Kuwaiti electric network during three years. The analysis proved that the proposed methodology accommodates this case study. The segmentation process results in homogeneous regions for which polynomial trends have been identified. Furthermore, it is found that the superimposed noise can be considered normally distributed, which enable constructing confidence intervals for future forecasts.

References

[1] M.R.G. Al-Shakarchi, M.M. Ghulaim, Short-term load forecasting for baghdad electricity region, Electric Machines and Power Systems 28 (2000) 355–371.

[2] H.K. Alfares, M. Nazeeruddin, Electric load forecasting: literature survey and classification of methods, International Journal of Systems Science 33 (2002) 23–34.

[3] L.F. Amaral, R.C. Souza, M. Stevenson, A smooth transition periodic autoregressive (STPAR) model for short-term load forecasting, International Journal of Forecasting 24 (2008) 603–615.

[4] C. Álvarez, S. Añ, Stochastic load modeling for electric energy distribution applications, Top 2 (1994) 151–166.

[5] G.J. Chen, K.K. Li, T.S. Chung, H.B. Sun, G.Q. Tang, Application of an innovative combined forecasting method in power system load forecasting, Electric Power Systems Research 59 (2001) 131–137.

[6] H. Chen, C.A. Cañizares, A. Singh. ANN-based short-term load forecasting in electricity markets. in: Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference 2, 2001, p. 411–415.

[7] J.G. De Gooijer, R.J. Hyndman, 25 Years of time series forecasting, International Journal of Forecasting 22 (2006) 443–473.

[8] V.Ş. Ediger, S. Akar, ARIMA forecasting of primary energy demand by fuel in Turkey, Energy Policy 35 (2007) 1701–1708.

[9] H. Hahn, S. Meyer-Nieberg, S. Pickl, Electric load forecasting methods: tools for decision making, European Journal of Operational Research 199 (2009) 902–907.

[10] H.R.S. Keyno, F. Ghaderi, A. Azade, J. Razmi, Forecasting electricity consumption by clustering data in order to decrease the periodic variable's effects and by simplifying the pattern, Energy Conversion and Management 50 (2009) 829–836.

[11] M. Martín-Merino, J. Román, A new SOM algorithm for electricity load forecasting, in: I. King et al. (Eds.), ICONIP 2006, Part I, LNCS 4232, Springer-Verlag, 2006, pp. 995–1003.

[12] M. Martín-Merino, J. Román, Electricity load forecasting using self organizing maps, in: S. Kollias et al. (Eds.), ICANN 2006, Part II, LNCS 4132, Springer-Verlag, 2006, pp. 709–716.

[13] K. Metaxiotis, A. Kagiannas, D. Askounis, J. Psarras, Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher, Energy Conversion and Management 44 (2003) 1525–1534.

[14] D.-X. Niu, Q. Wang, J.-C. Li, Short term load forecasting model based on support vector machine, in: D.S. Yeung et al. (Eds.), ICMLC 2005, LNAI 3930, Springer-Verlag, 2006, pp. 880–888.

[15] D. Niu, Y. Wang, D.D. Wu, Power load forecasting using support vector machine and ant colony optimization, Expert Systems with Applications 37 (2010) 2531–2539.

[16] Y. Ohtsuka, T. Oga, K. Kakamu, Forecasting electricity demand in Japan: a bayesian spatial autoregressive ARMA approach, Computational Statistics and Data Analysis 54 (2010) 2721–2735.

[17] A.K. Palit, D. Popovic, Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications, Springer-Verlag London Ltd., 2005.

[18] S.Sp. Pappas, L. Ekonomou, P. Karampelas, D.C. Karamousantas, S.K. Katsikas, G.E. Chatzarakis, P.D. Skafidas, Electricity demand load forecasting of the hellenic power system using an ARMA model, Electric Power Systems Research 80 (2010) 256–264.

[19] D.J. Pedregal, J.R. Trapero, Mid-term hourly electricity forecasting based on a multi-rate approach, Energy Conversion and Management 51 (2010) 105–111.

[20] I.S. Qamber, E.A. Al-Gallaf, Electric load forecasting model for the state of Bahrain network, Electric Power Components and Systems 29 (2001) 259–276.

[21] J.V. Ringwood, D. Bofelli, F.T. Murray, Forecasting electricity demand on short, medium and long time scales using neural networks, Journal of Intelligent and Robotic Systems 31 (2001) 129–147.

[22] D. Singh, S.P. Singh, Self organization and learning methods in short term electric load forecasting: a review, Electric Power Components and Systems 30 (2002) 1075–1089.

[23] H.A. Soltan, On forecasting with blind data, Mansoura Engineering Journal 25 (2000) M28–M34.

[24] J.W. Taylor, L.M. de 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.

[25] J.W. Taylor, Triple seasonal methods for short-term electricity demand forecasting, European Journal of Operational Research 204 (2010) 139–152.

[26] O. Valenzuela, I. Rojas, F. Rojas, H. Pomares, L.J. Herrera, A. Guillen, L. Marquez, M. Pasadas, Hybridization of intelligent techniques and ARIMA models for time series prediction, Fuzzy Sets and Systems 159 (2008) 821–845.

[27] H. Wang, B.-S. Li, X.-Y. Han, D.-L. Wang, H. Jin, Study of neural networks for electric power load forecasting, in: J. Wang et al. (Eds.), ISNN 2006, LNCS 3972, Springer-Verlag, 2006, pp. 1277–1283.

[28] J. Wang, W. Zhu, W. Zhang, D. Sun, A trend fixed on firstly and seasonal adjustment model combined with ε-SVR for short-term forecasting of electricity demand, Energy Policy 37 (2009) 4901–4909.

[29] R. Weron, Modeling and forecasting electricity loads and prices: a statistical approach, John Wiley & Sons Ltd., England, 2006.

[30] Y. Yao, Z. Lian, S. Liu, Z. Hou, Hourly cooling load prediction by a combined forecasting model based on analytic hierarchy process, International Journal of Thermal Sciences 43 (2004) 1107–1118.

[31] P. Zhou, B.W. Ang, K.L. Poh, A trigonometric grey prediction approach to forecasting electricity demand, Energy 31 (2006) 2839–2847.