Biblid: 1821-4487 (2013) 17; 2; p 97-99
UDK:621.315:621.3.017

Original Scientific Paper
Originalni naučni rad

COMPARATIVE ANALYSIS OF TEMPERATURE AND LOAD TIME SERIES INFLUENCE ON SHORT-TERM LOAD FORECASTING

UPOREDNA ANALIZA UTICAJA VREMENSKIH SERIJA TEMPERATURE I OPTEREĆENJA NA PROCES PREDVIĐANJA POTROŠNJE ELEKTRIČNE ENERGIJE

Slobodan ILIĆ, Srđan VUKMIROVIĆ, Aleksandar SELAKOV
University of Novi Sad, Faculty of Technical Science, 21000 Novi Sad, Trg Dositeja Obardovića 6
email: slobodan.ilic@uns.ac.rs

ABSTRACT

This paper presents several Multiple Linear Regressionmodels for short-term load forecasting, which are intended for use as benchmark tests. A comparative analysis based on models' performance is presented, which resulted from helping a US power utility to develop a commercial product.The main goal of the analysis carried out is to determine which variables influence the electrical load the most.Models which take into account only weather variables are developed on one side, and the ones containing only load shape information on the other. Both models are of linear regression type, while the input variables are selected based on correlation analysis. The influence of the different variables is than compared, and the model that uses both weather and load shape information is developed. The proposed benchmark models are used as a tool in developing more sophisticated methods for electrical load prediction, such as artificial neural netoworks, or support vector machines. The analysis of proposed models could provide useful results in dealing with a real-life forecasting algorithm and re-evaluation ofmethods already presented.
Key words: *short-term load forecasting, load data series, power industry, temperature influence, multiple linear regression.*

REZIME

U ovom radu je predstavljeno nekoliko modela višestruke linearne regresije, koji se koriste za navedenu problematiku, pri čemu je njihova osnovna namena mogućnost komparativne analize. Predstavljeno je upoređivanje različitih modela, koji su nastali prilikom razvoja aplikacije za predviđanje potrošnje el. energije, namenjene za korišćenje u elektroenergetskoj kompaniji iz SAD. Glavni cilj sprovedene analize je da se utvrdi koje promenljive imaju najveći uticaj na ponašanje krive potrošnje el. energije. Stoga su razvijeni posebni modeli koji uzimaju u obzir izolovani uticaj temperature i opterećenja. Na osnovu korelacione analize formiran je model koji u prvom slučaju sačinjavaju isključivo temperaturne promenljive a u drugom isključivo promenljive koje opisuju vremensku seriju el. opterećenja. Finalni model podrazumeva kombinaciju uticaja temperature i opterećenja i daje najbolje rezultate kada se aplicira na podatke koji su ovde analizirani. Analiza predstavljenih modela može biti korisna u radu sa realnim algoritmima predviđanja, pri čemu se neke od predloženih metoda mogu decidno testirati radi utvrđivanja praktične primenljivosti.
Ključne reči: *kratkoročno predviđanje potrošnje električne energije, vremenske serije potošnje, elektroenergetska industrija, uticaj temperature na potrošnju, višestruka linearna regresija.*

INTRODUCTION

Short-term load forecasting (STLF) presents an important functionality for an electric power system. It is defined as predicting the load shape for a given set of consumers in some future period. The usual horizon of forecasting is next 24 hours, or the next day, observed from the moment it is generated. Forecasts are sometimes generated for up to seven days ahead, depending on the specific needs of the system operator. Many important decisions are based on STLF, such as: unit commitment, generator scheduling, maintenance plans and economic dispatch. The task of STLF is difficult to achieve due to nonlinearity of the load series, its dependence on many different factors (environmental, social and economic), and their random-like behavior. The trend of applying new methodologies in order to tackle the challenges of power systems regarding STLF resulted in the development of numerous forecasting models. The models can roughly be categorized into groups: the classical methods *(Christiaanse, 1971; Papalexopoulos and Hesterberg, 1989; Meslier, 1978; Irisarri, et al., 1982; Shahidehpour, et al., 2002; Amjady, 2001)*, the artificial intelligence methods *(Hippert, et al., 2001; Srinivasan, et al., 1995; Bakirtzis, et al., 1995)* and more recently hybrid methods *(Song, et al., 2006; Fan, and Chen, 2006; Ilić, et al., 2012; Amjady and Keynia, 2009; Peng, et al., 1990; Amjady, et al., 2010)*. However, a large number of models proposed in the literature is hardly reproducible in real-life situations, and lacks the proper concreteness in order to be applied to a real system. This is mainly due to the fact that STLF models have mainly been developed to suit the specific needs of the load process they are being applied to. Benchmarking has been a known issue in STLF, although this category has not deserved as much attention as is deserves. Because of this tendency, the field of STLF in general lacks the well-established methodology to produce benchmarking models for comparative assessment *(Tao, et al., 2011)*. This paper analyses a naïve model that has been proposed in literature *(Tao, et al., 2011)*, for the purposes of benchmarking. It is applied to a set of load series data representing a US electric facility. The differences in behaviour between the original data set proposed in *(Tao, et al., 2011)*, the one proposed in this paper are analysed, and modifications of the model are proposed to better suit the specific needs in this case.

MATERIAL AND METHOD

General MLR models for STLF

The general linear regression model with normal error terms can be defined as:

$$L_i = \alpha_o + \alpha_1 X_{i1} + \alpha_2 X_{i2} + ... + \alpha_{p-1} X_{i,p-1} + e_i \tag{1}$$

where$\alpha_0 \dots \alpha_{p-1}$ are model parameters, $X_{i,1} \dots X_{i,p-1}$ are the known constants, e_i is the independent, normally distributed random variable $N(0, \sigma^2)$. Then the response function is:

$$E[Y] = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_{p-1} X_{p-1} \quad (2)$$

where$X_1 \dots X_{p-1}$represent$p-1$ predictor variables. The definition (1) thus implies that the observations L_i are independent normal variables, with mean $E[L_i]$ as given by (2) and constant variance σ^2.

Quantitative and qualitative prediction variables

Prediction variables that constitute the prediction function are often quantitative. For example, if we model the consumption in some area as a linear function of the numberconsumers in that area, the load time series will show the increasing pattern, if the number of consumers increases. The consumers count in this case can be regarded as a quantitative variable.

The definition of (1) does not strictly imply the use of variables that are quantitative. The use of qualitative variables, sometimes-called class, or dummy variables is also possible. These represent such information as the type of the day, which is weekday or weekend, and can be included in the model. Indicators with values 0 or 1 are used to identify the classes of a quantitative variable. For example, if the load (L) prediction is carried out based on the information about the type of the day for which the prediction is being done (i.e. whether it falls on weekend or on a workday), we define a qualitative prediction variable X_i in the following way:

$$\begin{cases} X_1 = 1, \text{ if the day is a weekday} \\ X_1 = 0, \text{ if the day is a weekend} \end{cases} \quad (3)$$

The predicting function is then defined as:

$$E[L] = \alpha_0 + \alpha_1 X_1 \quad (4)$$

In case of a weekday, where $X_i = 1$, the equation (4) becomes:

$$E[L] = \alpha_0 + \alpha_1 \quad (5)$$

In case of a weekend, where $X_i = 0$, the equation (4) becomes:

$$E[L] = \alpha_0 + \alpha_1 \quad (6)$$

Polynomial regression

Polynomial regression models contain polynomials of the predictor variables, which make the response function curvilinear. For example, if one of the variables used to predict the load (L) is temperature (T_i), and the order of the polynomial is three (*Meslier, 1978*), the prediction model can be written as follows:

$$L_i = \alpha_0 + \alpha_1 T_1 + \alpha_2 T_1^2 + \alpha_3 T_1^3 + e_i \quad (7)$$

Although the equation (7) has nonlinear members, it is in fact the special case of (1), because the prediction variables T_1, T_1^2 and T_1^3 are independent, which we can present as X_{i1}, X_{i2}, and X_{i3}, while the model proposed in (7) becomes:

$$L_i = \alpha_0 + \alpha_1 X_{i1} + \alpha_2 X_{i2} + \alpha_3 X_{i3} \quad (8)$$

Dataset used for modeling load demand process

The models that are proposed in this paper require two sets of data for their construction. In first case, we design the model to take into account only the temperature variables, while in the second one, the sole information that makes up the model is the one about the past behavior of the load process at system level. Most electric utilities have access to these two sets of data; some even have their own meteorological facilities, while others have to use weather services.

RESULTS AND DISCUSSION

Analysis of temperature influences

Temperature influence was tested independently from any other variable. Temperature variables that constitute the prediction model were chosen based on their correlation to the load series. Time scope that was included for the possible choice of temperature variables was three days. This choice was based upon trial and error, and on common sense, having in mind the heat capacity of the environment. The following choice of temperature variables was made in order to form purely temperature model:

- $T_{d-1}(h)$
- $T_{d-2}(h)$
- $T_{d-1}^2(h)$
- $T_{d-2}^2(h)$

from which we can construct a model:

$$L_d(h) = \alpha_1 T_{d-1}(h) + \alpha_2 T_{d-2}(h) + \alpha_3 T_{d-1}^2(h) + \alpha_4 T_{d-2}^2(h) \quad (9)$$

The proposed model was tested with recorded temperature measurements, for the period from January 1, 2009 to March 7, 2009. The training period was from January 1, 2009 to February 28, 2009, while the testing period was from February 28, 2009 to March 7, 2009. The model yielded results which are shown in Figure 1.

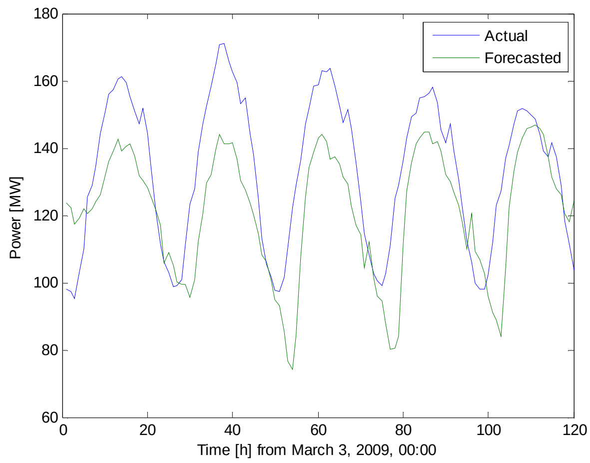

Fig. 1. Prediction based on purely temperature model

Analysis of load curve shape influences

The influence of the past behavior of load time series on prediction accuracy was tested isolated from other variables. Load variables that would enter the model were chosen based on the correlation analysis to the prediction signal. At the time the forecasting is executed, only the previously recorded data are available. Thus the following variables were chosen to construct the prediction model:

- $L_{d-1}(h)$
- $L_{d-5}(h)$
- $L_{d-7}(h)$

giving the following model:

$$L_d(h) = \alpha_1 L_{d-1}(h) + \alpha_2 L_{d-5}(h) + \alpha_3 L_{d-7} \quad (10)$$

The proposed model was tested with recorded load measurements, for the period from January 1, 2009 to March 7, 2009. The training period was from January 1, 2009 to February 28, 2009, while the testing period was from February 28, 2009 to

March 7, 2009. The model yielded results which are shown in Figure 2.

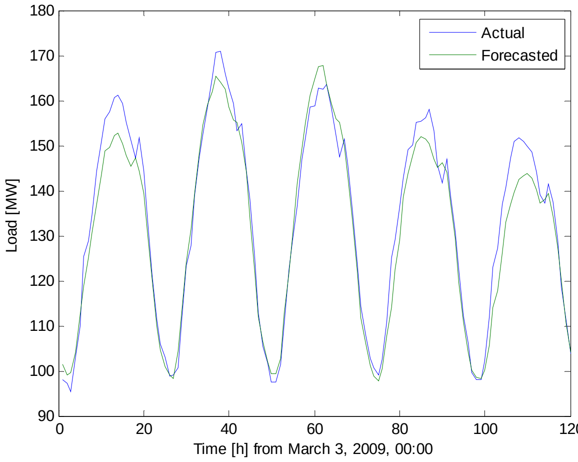

Fig. 2. Prediction based on purely load model

Combined model

The combined model is formed by integrating two previous models, that used isolated influence of temperature and load variables. The model can be defined as:

$$L_d(h) = \alpha_1 L_{d-1}(h) + \alpha_2 L_{d-5}(h) + \alpha_3 L_{d-7} + \alpha_4 T_{d-1}(h) + \cdots + \alpha_5 T_{d-2}(h) + \alpha_6 T_{d-1}^2(h) + \alpha_7 T_{d-2}^2(h) \quad (11)$$

with the total of seven parameters. The final results are shown in Figure 3, the training and testing periods being the same as in previous two cases.

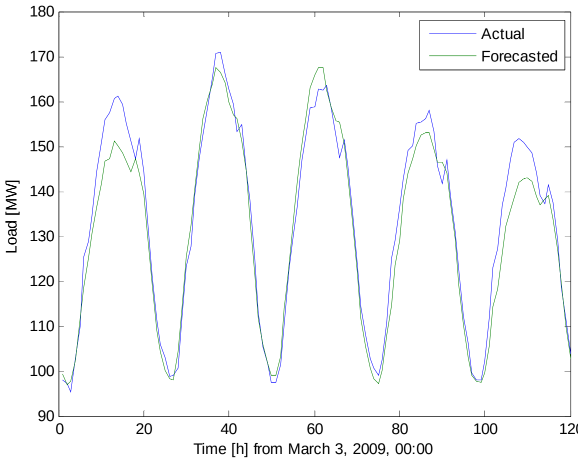

Fig. 3. Prediction based on combined load and temperature model

CONCLUSION

Producing benchmark tests in the field of STLF presents one of the challenges, mostly because the performance of different models varies greatly in dependence of the particular process being observed. One model may perform with exceptional precision on one class of electrical consumer, while it may completely fail on another one. In this paper we have investigated the benchmark test proposed in [HYPERLINK \l "Hon11" 16]. Base on the models performance, which was not satisfying for our problem class, we designed our own benchmark test, based on correlation analysis. The model was obtained by separate consideration of temperature and load variables, which were then combined into a single model. The performance of the model proposed in this paper is satisfying for the purposes of comparison. This paper can provide useful insight into benchmarking process in the field of STLF.It can thus serve as a standing point to researchers and practitioners who are designing sophisticated models, to suite particular purposes. The presented model is fully reproducible by following the documented procedure.

REFERENCES

Christiaanse, W.R. (1971). Short term load forecasting using general exponential smoothing, IEEE Trans Power ApparSyst, 90, 900-911.

Papalexopoulos, A.D., Hesterberg, T.C. (1989). A regression based approach to short term system load forecasting, in Proceedings of PICA conference, Seattle, Washington.

Meslier, F. (1978). New advances in short term load forecasting using Box and Jenkins approach, in IEEE/PES Winter Meeting, New York, 1978.

Irisarri, G.D., Widergren, S.E., Yehsakul, P.D. (1982). On-line load forecasting for energy control center application, IEEE Trans Power ApparSyst, 101, 71-78.

Shahidehpour, M., Li, Z., Yamin, H. (2002). Market operations in electric power systems. New York: John Wiley and sons Inc.

Amjady, N. (2001). Short-Term Hourly Load Forecasting Using Time Series Modeling With Peak Load Estimation Capability, IEEE Trans Power Syst, vol. 16, no. 4, pp. 798-805.

Hippert, H.S., Pedreira, C.E., Souza, R.C. (2001). Neural networks for short-term load forecasting: A review and evaluation, IEEE Trans Power Syst, 16, 44-45.

Srinivasan, D., Chang, D.S., Liew, A.C. (1995). Demand forecasting using fuzzy neural computation, with special emphasis on weekend and public holiday forecasting, IEEE Trans Power Syst, 10, 1897–1903.

Bakirtzis, A.G., Theocharis, J.B., Kiartzis, S.J., Satsois, K.J. (1995). Short term load forecasting using fuzzy neural networks, IEEE Trans Power Syst, 10, 1518–1524.

Song, K.B., Ha, S.K., Park, J.W., Kweon, D.J., Kim, K.H. (2006). Hybrid load forecasting method with analysis of temperature sensitivities. 21, 869-876.

Fan, S., Chen, L., (2006). Short-term load forecasting based on an adaptive hybrid method, 21, 392–401.

Ilić, S., Erdeljan, A., Kulić, F., Vukmirović, S. (2012). Hybrid artificial neural network system for short-term load forecasting, Thermal Science, 16, 215-224.

Amjady, N., Keynia, F. (2009). Short-term load forecasting of power systems by combination of wavelet transform and neuro-evolutionary algorithm, Energy, 34, 46-57.

Peng, T.M., Hubele, N.F., Karady, G.G., (1990). Conceptual approach to the application of neural network for short-term load forecasting, in IEEE IntSymp on Circuits and Systems, New Orleans, LA.

Amjady, N., Keynia, F., Zareipour, H. (2010). Short-term load forecast of microgrids by a new bilevel prediction strategy, IEEE Transactions on Smart Grid, 1, 286-294.

Hong, T., Wang, P., Willis, H.L. (2011). A Naïve Multiple Linear Regression Benchmark for Short Term Load Forecasting, in Power and Energy Society General Meeting, 2011 IEEE, Raleigh, NC, USA, 24-29 July 2011, 1-6.

Received: 05.03.2013. Accepted: 06.04.2013.

Beleške: