Electric Power Systems Research 130 (2016) 139–147

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Pattern-based local linear regression models for short-term load forecasting

Grzegorz Dudek*

Department of Electrical Engineering, Czestochowa University of Technology, Al. Armii Krajowej 17, 42-200 Czestochowa, Poland

ARTICLE INFO

Article history:
Received 2 March 2015
Received in revised form 2 September 2015
Accepted 4 September 2015

Keywords:
Linear regression
Partial least-squares regression
Patterns of seasonal cycles
Short-term load forecasting
Time series

ABSTRACT

In this paper univariate models for short-term load forecasting based on linear regression and patterns of daily cycles of load time series are proposed. The patterns used as input and output variables simplify the forecasting problem by filtering out the trend and seasonal variations of periods longer than the daily one. The nonstationarity in mean and variance is also eliminated. The simplified relationship between variables (patterns) is modeled locally in the neighborhood of the current input using linear regression. The load forecast is constructed from the forecasted output pattern and the current values of variables describing the load time series. The proposed stepwise and lasso regressions reduce the number of predictors to a few. In the principal components regression and partial least-squares regression only one predictor is used. This allows us to visualize the data and regression function. The performances of the proposed methods were compared with that of other models based on ARIMA, exponential smoothing, neural networks and Nadaraya–Watson estimator. Application examples confirm valuable properties of the proposed approaches and their high accuracy.

1. Introduction

Short-term load forecasting (STLF) is necessary for economic generation of power and system security. It refers to forecasts of system load from hours to several days ahead. The accurate load forecasts lead to lower operating cost which contributes to savings in electric utilities. The STLF accuracy is also important for the deregulated electricity markets. The amount of energy which the utility has to buy or sell in the real time market at unfavorable prices depends on the forecast error. Thus STLF is a very important problem for electric utilities, regional transmission organizations, energy suppliers and financial institutions. This is reflected in the literature by many forecasting methods that have been applied, including conventional methods and new computational intelligence and machine learning methods. A large research activity in the field of STLF is related with the problem complexity: the load time series is nonstationary in mean and variance, expresses trend, multiple seasonal variations (daily, weekly and annual) and random noise. In addition, load is affected by many external factors such as weather, time, demography, economy, electricity prices, geographical conditions, consumer types and their habits.

Among the conventional STLF methods the most commonly employed are: the Holt–Winters exponential smoothing (ES) and the autoregressive integrated moving average (ARIMA) models [1]. In ES the time series is decomposed into a trend component (expressed by level and growth terms) and seasonal components which can be combined additively or multiplicatively. ES allows us to model nonlinear and heteroscedastic time series but the exogenous variables cannot be introduced into the model. Other important disadvantages of ES are overparameterization and a large number of starting values to estimate. In [2] to reduce the dimension of the model new ES formulation called parsimonious seasonal ES was proposed. But there are still dozens or hundreds of terms to initialize and update in the model. The recently developed exponentially weighted methods in application to STLF are presented in [3].

ARIMA processes are a very rich class of possible models and allows us to model multiple seasonal cycles. The stochastic nature of load is often modeled with seasonal ARIMA models in practice. A disadvantage of ARIMA models is their linear nature. The order selection process of ARIMA is usually considered subjective and difficult to apply, which is a main obstacle in using these models. To simplify the forecasting problem the time series is often decomposed. The components: trend, seasonal components and irregular component, showing less complexity than the original series, are modeled independently (e.g. [4]). Another time series

* Corresponding author. Tel.: +48 343250896; fax: +48 343250803.
E-mail address: dudek@el.pcz.czest.pl

http://dx.doi.org/10.1016/j.epsr.2015.09.001
0378-7796/

decomposition method using lifting scheme (the second generation wavelet transform) was described in [5].

The most popular computational intelligence methods applied in STLF are neural networks. They have many attractive features such as: universal approximation property, learning capabilities, massive parallelism, robustness in the presence of noise, and fault tolerance. The drawbacks of neural network include: disruptive and unstable training, difficulty in matching the network structure to the problem complexity, weak capacity of extrapolation and many parameters to estimate (hundreds of weights). Some examples of using neural networks in STLF are: [6], where the complexity of the network is controlled by the Bayesian approach, [7], where a new hybrid forecasting method composed of wavelet transform, multilayer perceptron and evolutionary algorithm is proposed, [8], where a generic framework combining similar day selection, wavelet decomposition, and multilayer perceptron is presented, and [9], where the neural network generates the prediction intervals.

Another branch of computational intelligence, fuzzy logic, allows us to enter information by facts and rules formulated verbally by experts and describing the behavior of complex systems by using linguistic expressions. With the help of fuzzy rules the imprecise, incomplete and ambiguous information can be introduced into the STLF models. When it is difficult to gain knowledge directly from the experts, to generate a set of if-then rules the neuro-fuzzy approach is applied, which learns from examples. But the neuro-fuzzy system structure is complex and the number of parameters is usually large (it depends on the problem dimensionality and complexity), so the learning is difficult and does not guarantee convergence to the global minimum. Examples of STLF models based on fuzzy logic are: [10], where the neuro-fuzzy system is used to adjust the results of load forecasting obtained by radial basis function neural network, [11], where two neuro-fuzzy networks are proposed: a wavelet fuzzy neural network using the fuzzified wavelet features as the inputs, and fuzzy neural network employing the Choquet integral as the outputs, [12], where an integrated approach which combines a self-organizing fuzzy neural network learning method with a bilevel optimization method is described, and [13], where the forecasting model combines fuzzy logic, wavelet transform and neural network. Another useful computational intelligence tools for STLF are: support vector machines (SVM) [14,15], ensembles of models [16,17] and artificial immune systems [18] (description of more STLF models you can find on the website http://gdudek.el.pcz.pl/publications).

It is noteworthy that many of the STLF models developed in recent years are hybrid solutions. They combine data preprocessing methods (e.g. wavelet transform) with approximation methods (such as neural and neuro-fuzzy networks or SVM) and optimization or learning methods (e.g. evolutionary and swarm algorithms).

The disadvantages of the above mentioned complex forecasting models with many parameters are: hard and time-consuming training, problems with generalization, unclear structure and uninterpretable parameters. Most often time series with multiple seasonal cycles and trend, expressing nonstationarity in mean and variance cannot be modeled directly and additional treatments such as detrending, deseasonality or decomposition are needed.

In contrast to the complex models commonly used in STLF in this work simple methods of linear regression are proposed. The number of parameters here is small and they can be estimated using simple least squares approach. The key element of the proposed methods is data preprocessing: defining patterns of seasonal cycles. This simplifies the STLF problem eliminating nonstationarity, and filtering trend and seasonal cycles longer than the daily one.

The paper is organized in a theoretical and an empirical part. In the beginning the patterns of daily cycles of load time series are defined. Then the main concepts of the linear regression models for STLF are introduced. In the last section the real load data are used to provide examples of model building and forecasting in practice. The results of the proposed methods are compared to results of other STLF methods: ARIMA, ES, multilayer perceptron and Nadaraya–Watson estimator.

2. Patterns of the times series seasonal cycles

Data preprocessing based on patterns simplifies the forecasting time series with multiple seasonal cycles. In our case the patterns of the daily cycles are introduced: the input patterns $\mathbf{x}$ and output ones $\mathbf{y}$. The input pattern is a vector $\mathbf{x} = [x_1\, x_2\, \ldots\, x_n]^T \in X = \mathbb{R}^n$, representing the vector of loads in successive timepoints of the daily period: $\mathbf{L} = [L_1\, L_2\, \ldots\, L_n]^T$, where $n = 24$ for hourly load time series, $n = 48$ for half-hourly load time series and $n = 96$ for quarter-hourly load time series. The functions mapping the time series elements L into patterns are dependent on the time series (trend, seasonal variations), the forecast period and horizon. They should maximize the model quality. In this study the input pattern $\mathbf{x}_i$, representing the ith daily period, is defined as follows:

$$x_{i,t} = \frac{L_{i,t} - \bar{L}_i}{\sqrt{\sum_{l=1}^{n} \left(L_{i,l} - \bar{L}_i\right)^2}}, \tag{1}$$

where $i = 1, 2, \ldots, N$ is the daily period number, N is the number of days in the time series, $t = 1, 2, \ldots, n$ is the time series element number in the period i, $L_{i,t}$ is the tth time series element (load) in the period i, $\bar{L}_i$ is the mean load value in the period i.

According to definition (1), first we subtract the vector $\mathbf{L}_i$ mean from its components and then we divide the resulting vector by its length. As a result we get the normalized vectors $\mathbf{x}_i$ with length 1, zero mean and the same variance. Note that the time series which is nonstationary in mean and variance is represented now by x-patterns having the same mean and variance. The trend and additional seasonal variations (weekly and annual ones in our case) are filtered. The x-patterns contain information only about the shapes of daily curves.

Whilst the x-patterns represent input variables (predictors), i.e. the loads for the day i, the y-patterns represent the output variables, i.e. the forecasted loads for the day $i + \tau$, where τ is a forecast horizon in days. The components of the n-dimensional output pattern $\mathbf{y}_i = [y_{i,1}\, y_{i,2}\, \ldots\, y_{i,n}]^T \in Y = \mathbb{R}^n$, representing the load vector $\mathbf{L}_{i+\tau}$ are defined as follows:

$$y_{i,t} = \frac{L_{i+\tau,t} - \bar{L}_i}{\sqrt{\sum_{l=1}^{n} \left(L_{i,l} - \bar{L}_i\right)^2}}, \tag{2}$$

where $i = 1, 2, \ldots, N$, $t = 1, 2, \ldots, n$.

This is the similar equation to (1) but in this case we do not use the mean load of the day $i + \tau$ $\left(\bar{L}_{i+\tau}\right)$ in the numerator and $\sqrt{\sum_{l=1}^{n} \left(L_{i+\tau,l} - \bar{L}_{i+\tau}\right)^2}$ in the denominator, because these values are not known in the moment of forecasting. We use known values of $\bar{L}_i$ and $\sqrt{\sum_{l=1}^{n} \left(L_{i,l} - \bar{L}_i\right)^2}$ instead. This is very important because when the forecast of pattern $\mathbf{y}_i$ is generated by the model we can determine the forecast of vector $\mathbf{L}_{i+\tau}$ using transformed Eq. (2):

$$\widehat{L}_{i+\tau,t} = \widehat{y}_{i,t} \sqrt{\sum_{l=1}^{n} \left(L_{i,l} - \bar{L}_i\right)^2} + \bar{L}_i, \tag{3}$$

where $\widehat{y}_{i,t}$ is the forecasted tth component of the pattern $\mathbf{y}_i$.

Note that $\bar{L}_i$ and the value of square root in (3) are known at the time of forecasting and can be used for decoding of $\widehat{y}_{i,t}$ to get $\widehat{L}_{i+\tau,t}$. Note also that $\sqrt{\sum_{l=1}^{n} \left(L_{i,l} - \bar{L}_i\right)^2}$ is the carrier of the dispersion of the current daily cycle. Using this square root in the denominator of