Electrical Power and Energy Systems 65 (2015) 94–101

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

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

Forecasting methods for balancing energy market in Poland

Tomasz Popławski, Grzegorz Dudek *, Jacek Łyp

Czestochowa University of Technology, Institute of Power Engineering, Al. Armii Krajowej 17, 42-200 Częstochowa, Poland

ARTICLE INFO

Article history:
Received 5 September 2013
Received in revised form 10 September 2014
Accepted 22 September 2014

Keywords:
Electricity demand forecasting
Balancing energy market
Coordination plans

ABSTRACT

In the paper the scope of research carried out for the Transmission System Operator of the Polish Power System PSE-Operator S.A. in order to make up forecasting tools supporting creation of coordination plans is described. The Transmission System Operator is obliged legally to make up such plans for traffic and maintenance of the transmission grid. The article describes in detail forecasting models examined for different time horizons, for which the coordination plans are made up. These models were designed for preparing the daily, monthly and annual coordination plans by the PSE-Operator S.A. and they are currently in the implementation phase. The model based on fuzzy estimators supporting daily coordination plans, standard load curve model supporting monthly coordination plans and hybrid model supporting annual coordination plans are presented. The models were verified using real data examples.

Introduction

Forecasting of future loads of the power system is a crucial operation for the operator of transmission system (OTS), who is responsible for making up instant energy balance and adjustment of energy supplies and deliveries to technical requirements and actual energy demands.

Forecasting may be carried out for different time horizons – within annual, monthly, weekly and daily plans. Prediction of a day-to-day load variation remains a particularly challenging task, therefore a number of the so-called daily coordination plans have been developed. These include technical–commercial daily balance (TCDB), initial daily coordination plan (IDCP), daily coordination plan (DCP) and current daily coordination plan (CDCP). The classification was discussed in more detail in [32].

In the process of power system operation the fundamental role is played by the DCP and CDCP. The IDCP and the TCDB include aggregated general-type data. Within the framework of daily planning the present OTS's knowledge and the control possibilities of supply units, the forecasted energy demand, the quantity of determined production, international exchange and system limitations are taken into account. The CDCP is created during the commercial 24 h period for the demands of traffic maintenance. Within this framework the hourly power demand is divided into 15-min long periods.

For every planning stage the forecast tools supporting the planning processes, with specific features in dependence on the planning horizon, are needed. Many of the forecasting methods described in [6] could be adopted for this purpose.

A number of short-term load forecasting (STLF) models have been designed in recent years. Conventional STLF models use smoothing techniques (e.g. [4,35], regression methods (e.g., [28,9], and statistical analysis. Regression methods are usually applied to model the relationship between load consumption and other factors (e.g., weather, day type, and customer class) [13]. ARIMA and related models, where the load is modeled by the autoregressive moving average difference equation, are very popular [16,22]. These models are based on the assumption that the data exhibit specific features, such as autocorrelation, trend, and seasonal variation. Conventional STLF methods have a strong theoretical basis and are still competitive with newer methods [36].

In recent years, artificial intelligence methods (AI) have been widely applied to STLF [25]. AI methods of forecasting have shown the capability to perform better when dealing with non-linearities and other difficulties in modeling of the time series. They do not require any complex mathematical formulations or quantitative correlation between inputs and outputs. The AI methods most often used in STLF can be divided as follows: neural networks (e.g. multilayer perceptron, radial basis function network, Kohonen network, recurrent networks) (e.g. [29,20,5]), fuzzy and neuro-fuzzy systems (e.g., [24,38,34,21,10]) and expert systems (e.g., [33,18]).

New STLF methods are still being created. Some of them are based on machine learning and pattern recognition techniques, for example regression trees, cluster analysis methods (e.g., [14]),

* Corresponding author. Tel.: +48 34 3250 896; fax: +48 34 3250 803.
E-mail addresses: poptom@el.pcz.czest.pl (T. Popławski), dudek@el.pcz.czest.pl (G. Dudek), jackrat@el.pcz.czest.pl (J. Łyp).

http://dx.doi.org/10.1016/j.ijepes.2014.09.029
0142-0615/

support vector machines (e.g., [26,2]), wavelets [1], fractal geometry, point function method [23], canonical distribution of the random vector method [30], and the artificial immune system [7].

The models for medium-term and long-term load forecasting developed in recent years are based mainly on the AI and machine learning methods – neural networks (e.g., [15]), fuzzy and neuro-fuzzy approaches (e.g. [3]), support vector regression (e.g., [19]), and swarm intelligence (e.g., [37]).

In the subsequent part of the paper the proposed forecasting methods supporting the daily, monthly and annual plans are presented in detail.

Forecasting model supporting daily coordination plans

The basic input data for setting up the construction procedures of daily coordination plans are the load forecasts. The daily forecasts of the Polish power system quarter-hour demand are prepared according to the following schedule:

- for DCP the forecasts for the day $t+1$ are prepared once a day, at the midday of the day t,
- for IDCP the forecasts for the day $t+2$ are prepared once a day, at the midday of the day t,
- for TCDB the forecasts for the following 7 days $t+3$, $t+4,\ldots$, $t+9$ are prepared once a day, at the midday of the day t,
- for CDCP the forecasts for the day $t+1$ are prepared on demand after the midday of the day t and during the day $t+1$.

The subject of study in the present paper is the fuzzy estimator of the regression function, which has recently been developed in our team. In the course of subsequent research it has turned out, that it outperforms other STLF models, such as the adopted predictor of standard load curve and multilayer perceptron [31].

Similarity-based STLF models

The proposed forecasting model belongs to a class of similarity-based methods (SB) of STLF [8]. These are nonparametric regression methods, where the regression function is estimated from data using mutual similarities between the data points.

The load time series are characterized by annual, weekly, and daily cycles due to the changes in industrial activities and climatic conditions. In Fig. 1 the load time series of the Polish power system is shown.

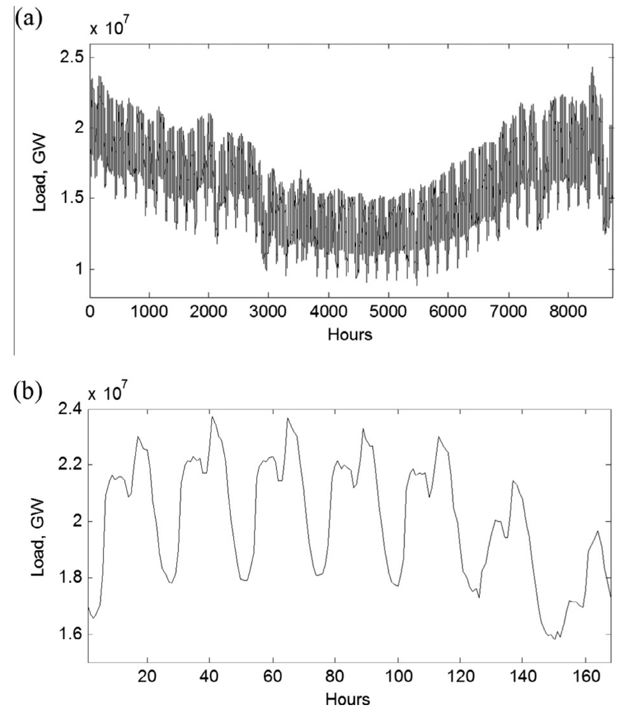

Fig. 1. The load time series of the Polish power system in the yearly (a) and weekly (b) intervals.

The SB methods use the analogies between time series sequences with periodicities. The course of a time series can be deduced from the behavior of this time series in similar conditions in the past or from the behavior of other time series with similar changes in time. At first of the SB forecasting procedure, the time series is divided into sequences, which usually contain one period (in the considered STLF problem, the period is equal to 96 quarters). To eliminate weekly and annual variations, the sequence elements are preprocessed to extract their patterns. The pattern is a vector with components that are functions of real time series elements, that is, quarter-hourly loads in this case. The input and output (forecast) patterns are defined: $\mathbf{x} = [x_1\ x_2\ \ldots\ x_{96}]$ and $\mathbf{y} = [y_1\ y_2\ \ldots\ y_{96}]$, respectively. The patterns are paired ($\mathbf{x}_t$, $\mathbf{y}_t$), where $\mathbf{y}_t$ is a pattern of the time series sequence succeeding the sequence represented by $\mathbf{x}_t$, and the interval between these sequences (forecast horizon τ) is constant. The SB methods are based on the following assumption: If the pattern $\mathbf{x}_a$ in a period preceding the forecast moment is similar to the pattern $\mathbf{x}_b$ from the history, then the forecast pattern $\mathbf{y}_a$ is similar to the forecast pattern $\mathbf{y}_b$. Patterns $\mathbf{x}_a$, $\mathbf{x}_b$, and $\mathbf{y}_b$ are determined from the history of the process. Pairs $\mathbf{x}_a$–$\mathbf{x}_b$ and $\mathbf{y}_a$–$\mathbf{y}_b$ are defined in the same way and are shifted in time by the same number of series elements (usually this is a multiple of the daily period).

The similarity measures are based on the distance measures (most often Euclidean or Manhattan), correlation measures, or a function similarity measure.

The way the $\mathbf{x}$ and $\mathbf{y}$ patterns are defined depends on the nature of the time series (seasonal variations and trends) and the forecast horizon. Functions transforming series elements into patterns should be defined so that patterns could carry most information about the process, and the model quality becomes maximal. Moreover, functions transforming forecast sequences into patterns $\mathbf{y}$ should ensure the possibility of calculating the real forecast of the time series elements.

Taking into account the schedule of the coordination plan preparation, the forecast patterns $\mathbf{y}$ encode real loads (L) in the following quarters of the forecast day $t+\tau$: $\mathbf{L}_{t+\tau} = [L_{t+\tau,1}\ L_{t+\tau,2}\ \ldots\ L_{t+\tau,96}]$, and the input patterns $\mathbf{x}$ map the quarter-hourly loads preceding the forecast day – $\mathbf{L}_{t'} = [L_{t',1} L_{t',2} \ldots L_{t',96}]$, where t is the day number in which the forecasting procedure is carried out, τ is a forecast horizon equal to 3, 4, …, 9 for TCDB, 2 for IDCP, and 1 for DCP, t' denotes 24-h period from the midday of the day t-1 to the midday of the day t including 96 quarters. $\mathbf{L}_{t'}$ contains the most current information about system loads available at the moment of forecast preparation. Vectors $\mathbf{y}$ are encoded using actual process parameters (determined from the period t'), which allows us to take into consideration the current variability of the process and ensures the possibility of decoding.

On the basis of earlier experiments the following pattern definitions $\mathbf{x}_t = [x_{t,1}\ x_{t,2}\ \ldots\ x_{t,96}]$ and $\mathbf{y}_t = [y_{t,1}\ y_{t,2}\ \ldots\ y_{t,96}]$ are adopted:

$$x_{t,i} = \frac{L_{t',i} - \bar{L}_{t'}}{\sqrt{\sum_{l=1}^{96}\left(L_{t',l} - \bar{L}_{t'}\right)^2}}, \tag{1}$$

$$y_{t,i} = \frac{L_{t+\tau,i} - \bar{L}_{t'}}{\sqrt{\sum_{l=1}^{96}\left(L_{t',l} - \bar{L}_{t'}\right)^2}}, \tag{2}$$