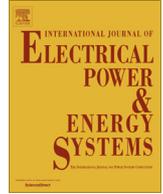




Contents lists available at ScienceDirect

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 Czestochowa, 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.

© 2014 Elsevier Ltd. All rights reserved.

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).

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 $t + 3, t + 4, \dots, t + 9$ 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 STLFL models, such as the adopted predictor of standard load curve and multilayer perceptron [31].

Similarity-based STLFL models

The proposed forecasting model belongs to a class of similarity-based methods (SB) of STLFL [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.

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 STLFL 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: $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_{96}]$ and $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_{96}]$, 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

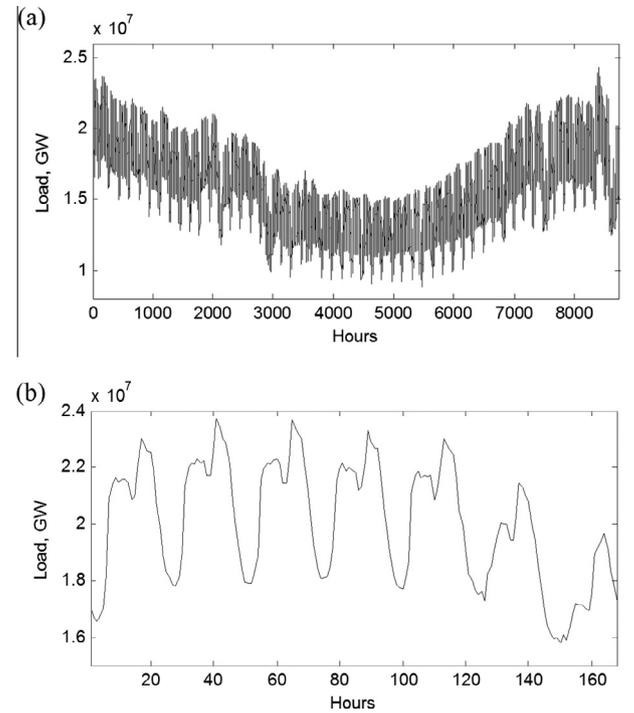


Fig. 1. The load time series of the Polish power system in the yearly (a) and weekly (b) intervals.

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 $t + \tau$: $\mathbf{L}_{t+\tau} = [L_{t+\tau,1} \ L_{t+\tau,2} \ \dots \ L_{t+\tau,96}]$, and the input patterns \mathbf{x} map the quarter-hourly loads preceding the forecast day – $\mathbf{L}_{t'} = [L_{t',1} \ L_{t',2} \ \dots \ L_{t',96}]$, 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 $\mathbf{x}_t = [x_{t,1} \ x_{t,2} \ \dots \ x_{t,96}]$ and $\mathbf{y}_t = [y_{t,1} \ y_{t,2} \ \dots \ y_{t,96}]$ are adopted:

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

$$y_{t,i} = \frac{L_{t+\tau,i} - \bar{L}_{t+\tau}}{\sqrt{\sum_{l=1}^{96} (L_{t+\tau,l} - \bar{L}_{t+\tau})^2}}, \quad (2)$$

where t is the day number or the pattern number corresponding to day t , $i = 1, 2, \dots, 96$ is the component number of \mathbf{L} , \mathbf{x} or \mathbf{y} , t' is the 24-h period from the midday of the day $t - 1$ to the midday of the day t including 96 quarters, $L_{t',i}$ is the load at quarter i of period t' , $\bar{L}_{t'}$ is the mean load in period t' .

Definition (1) expresses the normalization of the vectors $\mathbf{L}_{t'}$. After normalization, they obtain unity length, zero mean, and the same variance. Forecast pattern (2) is analogous to input patterns (1), however, it is encoded using the current loads determined from the process history, which enables us to decode the forecast vector $\mathbf{L}_{t+\tau}$ after the forecast of pattern \mathbf{y} is determined.

In the case of the forecast preparation procedure for CDCP we have two variants:

- we prepare forecast at quarter $k = 49, 50, \dots, 96$ of the day t for all quarters of the day $t + 1$ ($\tau = 1$),
- we prepare forecast at quarter $k = 1, 2, \dots, 95$ of the day $t + 1$ for the remaining quarters of this day ($\tau = 0$).

In variant (a) the period t' in (1) and (2) includes the quarters from $k + 1$ of the day $t - 1$ to k of the day t . In variant (b) the period t' includes the quarters from $k + 1$ of the day t to k of the day $t + 1$, and i in (2) denotes the quarters from $k + 1$ to 96, thus the forecast pattern \mathbf{y} in the case of (b) has the components $i = k + 1, k + 2, \dots, 96$.

The SB forecasting methods are based on the nonparametric approach to the estimation of regression function. Flexibility of nonparametric models is very useful in the preliminary analysis of a dataset and may be helpful in the construction of parametric models, which are usually more convenient to use but not necessarily more precise. The general model of nonparametric regression is of the following form:

$$y = m(x) + \varepsilon, \quad (3)$$

where y is the response variable, x is the predictor, ε is the error which is assumed to be un-biased and have a normal distribution with a zero mean and constant variance and $m(x) = E\{Y|X = x\}$ is the regression curve.

The goal of the nonparametric regression is to estimate the function $m(x)$. Most methods implicitly assume that this function is smooth and continuous. The most popular nonparametric regression models are the kernel estimators, k -nearest neighbor estimators, orthogonal series estimators, and spline smoothing [17]. In the present paper the fuzzy estimators of the regression function are applied to the STLF problem described above.

STLF model based on fuzzy estimators

The forecasting model uses a set of reference pattern pairs $(\mathbf{x}_t, \mathbf{y}_t)$ from the history of the process. For a given input pattern \mathbf{x}^* , representing the current daily curve in the period t' , the most similar patterns \mathbf{x}_t in the reference set are found, and the forecast pattern \mathbf{y}^* is formed from the patterns \mathbf{y}_t paired with them.

The model is based on the fuzzy set theory. Vectors \mathbf{x} correspond to the points in the 96-dimensional space. We assign the reference points \mathbf{x}_t , representing the same days of the week as \mathbf{x}^* (Monday to Sunday), to the neighborhood of the input point \mathbf{x}^* . (The distinction of days of the week during this procedure is caused by the diversity of the load curve shapes for the different days of the week.) Assigning the reference points to the neighborhood of \mathbf{x}^* is not sharp but fuzzy. The function of membership the patterns \mathbf{x}_t to the neighborhood of pattern \mathbf{x}^* , depending on the distance $d(\mathbf{x}^*, \mathbf{x}_t)$, is defined:

$$\mu(\mathbf{x}^*, \mathbf{x}_t) = \exp \left[- \left(\frac{d(\mathbf{x}^*, \mathbf{x}_t)}{\sigma} \right)^\alpha \right], \quad (4)$$

where σ is the width parameter and α is the shape parameter.

The membership function (4) is a Gaussian function with the center at point \mathbf{x}^* (see Fig. 2). Alternative membership functions are presented in [8].

Now the estimator $m(x)$ can be defined as follows:

$$m(\mathbf{x}^*) = \frac{\sum_{i=1}^n \mu(\mathbf{x}^*, \mathbf{x}_t) \mathbf{y}_t}{\sum_{i=1}^n \mu(\mathbf{x}^*, \mathbf{x}_t)}, \quad (5)$$

where n is the number of patterns \mathbf{x}_t representing the same days of the week as pattern \mathbf{x}^* .

The patterns \mathbf{y}_t associated with patterns \mathbf{x}_t closer to the pattern \mathbf{x}^* have stronger influence on the formation of response \mathbf{y}^* , which is calculated as the mean of patterns \mathbf{y}_t weighted by the membership degrees $\mu(\mathbf{x}^*, \mathbf{x}_t)$.

In the case of the forecast pattern \mathbf{y}_t representing an untypical day (e.g. public holiday), it can deform the forecast pattern \mathbf{y}^* . To eliminate situations like this to each pair of $(\mathbf{x}_t, \mathbf{y}_t)$ a degree of confidence w_t is assigned. When pattern \mathbf{y}_t expresses an untypical day $w_t = 0$, otherwise $w_t = 1$. The modified equation (5) with the degrees of confidence has the form:

$$m(\mathbf{x}^*) = \frac{\sum_{i=1}^n w_t \mu(\mathbf{x}^*, \mathbf{x}_t) \mathbf{y}_t}{\sum_{i=1}^n w_t \mu(\mathbf{x}^*, \mathbf{x}_t)}. \quad (6)$$

For untypical days, a separate forecasting model has been designed, which is based on the analogies between patterns representing the same holidays in neighboring years [31].

Incorporating the weather factors into the model needs a definition of the forecast pattern context. The homogeneous factors correlated with loads represented by the pattern \mathbf{y} are called its context \mathbf{z} . The context \mathbf{z} is a vector of the factors or their functions, e.g. the temperature context can express the daily vector of hourly atmospheric temperatures. Different contexts can be defined, e.g. related to temperature, humidity, wind strength, etc. It is assumed that patterns \mathbf{y}_t with contexts \mathbf{z}_t which are more similar to the context \mathbf{z}^* of the forecasted pattern \mathbf{y}^* , are more informative and should have a stronger influence on the regression curve than patterns with distant contexts. The membership degree of the context \mathbf{z}_t to the context \mathbf{z}^* is modeled using the function of the same form as (4):

$$\mu(\mathbf{z}^*, \mathbf{z}_t) = \exp \left[- \left(\frac{d(\mathbf{z}^*, \mathbf{z}_t)}{\sigma_z} \right)^\alpha \right], \quad (7)$$

The estimator $m(x)$ in this case is of the form:

$$m(\mathbf{x}^*) = \frac{\sum_{t=1}^n \prod_{i=1}^m \mu(\mathbf{z}_i^*, \mathbf{z}_{i,t}) w_t \mu(\mathbf{x}^*, \mathbf{x}_t) \mathbf{y}_t}{\sum_{t=1}^n \prod_{i=1}^m \mu(\mathbf{z}_i^*, \mathbf{z}_{i,t}) w_t \mu(\mathbf{x}^*, \mathbf{x}_t)}, \quad (8)$$

where \mathbf{z}_i is the i th context.

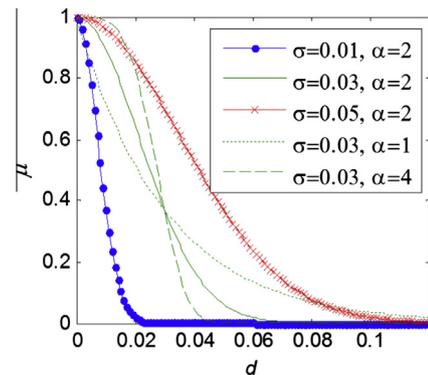


Fig. 2. Exemplary plots of membership function (4).

When all contexts of the t th pattern ($\mathbf{z}_{i,t}$) are similar to the contexts of the forecasted pattern \mathbf{z}_i^* , then the product of the membership degrees $\mu(\mathbf{z}_i^*, \mathbf{z}_{i,t})$, $i = 1, 2, \dots, m$ is high, close to 1. When some contexts $\mathbf{z}_{i,t}$ are not similar to \mathbf{z}_i^* , the product of $\mu(\mathbf{z}_i^*, \mathbf{z}_{i,t})$ decreases, and consequently the total weight of the t th pattern \mathbf{y}_t decreases. This results in the less influence of \mathbf{y}_t in the weighted mean (8). The influence level of the individual contexts on the forecast can be controlled by the membership function parameters – σ_z and α_z .

In the application example described in the next section only the temperature context is used in the model, because from a number weather factors (temperature, wind strength, clouds, humidity, rainfall and snowfall) it affects the load most [31]. The temperature context \mathbf{z}_T is a vector of the mean hourly atmospheric temperatures in the period t' : $\mathbf{z}_{T,t} = [T_{t',1} T_{t',2} \dots T_{t',24}]$. If the temperature forecasts on the day $t + \tau$ are available, the temperature context can be formed from them.

Application examples

To demonstrate the effectiveness of the proposed approach, the load forecasting of the Polish power system was carried out using historical quarter-hourly load and weather (temperature) data from the period 1997.05.01 – 2009.11.16. The dataset was divided into two subsets: training and test. The first sequences of the time series (from 1997.05.01 to 2007.11.19) were included in the training set, and the latest sequences were included in the test set. The training set was used to optimize the model (in leave-one-out procedure) and the test set – to the assessment of model performance.

The earlier experiments have shown that the shape parameter α is not as important as the width parameter σ [8]. So we assumed the constant value of $\alpha = 2$ and we adjust the value of σ . The Euclidean distance was used in (4) and (7). Untypical days (13 days in the year) were forecasted using a separate model [31].

The model quality was evaluated using the mean average percentage error (MAPE), which is conventionally used in STLF. Table 1 presents the mean error (MAPE), the daily peak load error (MAPE_{dp}), the daily second peak load error (MAPE_{d2p}), and the daily valley load error (MAPE_{dv}). MAPE for DCP is presented in Fig. 3, and the cumulative distribution function $F(\text{MAPE})$ is presented in Fig. 4.

Fig. 5 shows how the forecast error for CDCP changes depending on the moment of the forecast preparation, which is shifted in time with regard to the moment of DCP forecast preparation (i.e. 12 noon of the day t , that is 48 quarter) for 1, 2, ..., 143 quarters ahead. The shift for one quarter means the forecast for all quarters of the day $t + 1$ which is prepared at the 49th quarter of the day t . The shift for 143 quarters means the forecast for the last quarter of the day $t + 1$, which is prepared at the 95th quarter of this day.

The lowest errors are observed for the workdays Tuesday – Friday, whereas the higher ones for Mondays and Sundays. It is worth to notice that the mean forecast error for daily peaks is smaller than the mean error for all quarters, independently on the forecast horizon τ . Fig. 3a) shows the lower errors for valley loads. This fact is probably related with the lower load variance during the nights than during the days, especially in peak periods.

Table 1
Forecast errors for DCP, IDCP and TCDB.

Plan	τ	MAPE	MAPE _{dp}	MAPE _{d2p}	MAPE _{dv}
DCP	1	1.67	1.62	1.95	1.65
	2	1.99	1.89	2.17	2.07
TCDB	3	2.24	2.08	2.35	2.43
	4	2.37	2.18	2.45	2.65
	5	2.48	2.30	2.51	2.79
	6	2.44	2.24	2.41	2.78
	7	2.55	2.38	2.51	3.00
	8	2.63	2.47	2.64	3.08
	9	2.69	2.51	2.69	3.08

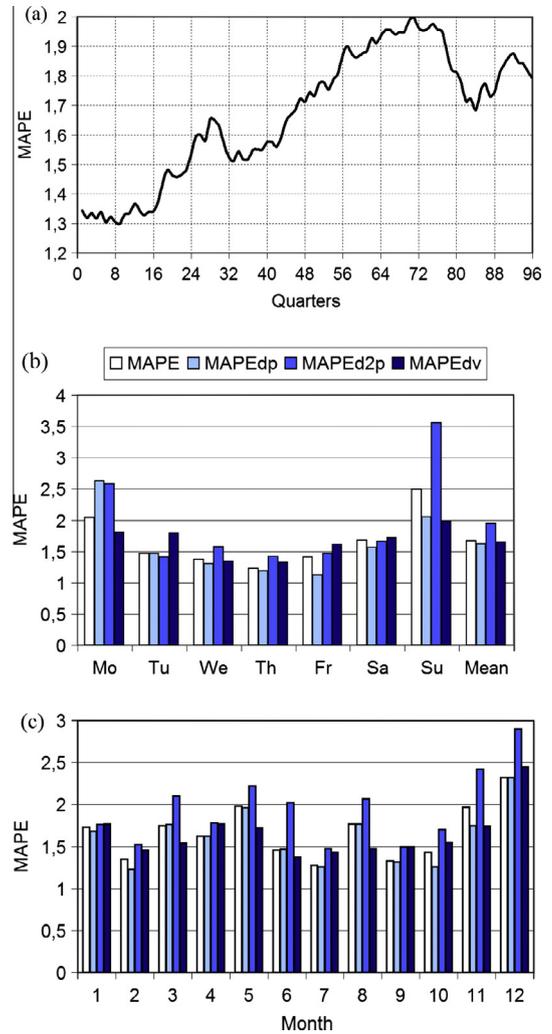


Fig. 3. Forecast errors for DCP for: successive quarters of the day (a), day types (b) and months (c).

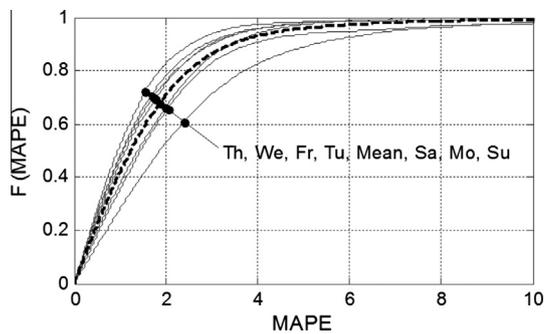


Fig. 4. The cumulative distribution function $F(\text{MAPE})$ for DCP.

The forecasts for CDCP become more and more accurate with the shift of the forecast preparation moment, but in some cases errors increase with the shift, e.g. for quarters from 60 to 72.

In the next study we compare the proposed STLF model based on fuzzy estimators (FE) with the classical STLF models: ARIMA and exponential smoothing (ES), as well as with the machine learning models: neural networks and random forest. The first three models are described in detail in [11] and the last one is described in [12]. We use hourly electrical load data of the Polish power system from the period 2002–2004 (this data can be downloaded from the website <http://gdudek.el.pcz.pl/varia/stlfdata>). Our goal

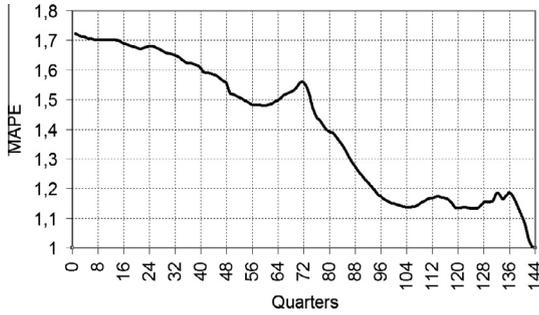


Fig. 5. The forecast error for CDCP depending on the shift of the forecast preparation moment.

is to forecast the load curve for the next day ($\tau = 1$). The moment of forecasting is the last hour of the day preceding the day of forecast. The test set includes 30 days from January 2004 (without untypical 1 January) and 31 days from July 2004. The forecast errors and their interquartile ranges in Table 2 are shown. As we can see from this table errors for FE, ANN and RF are similar and lower than errors for ARIMA and ES (that was confirmed by the Wilcoxon signed-rank test). Notice that FE model is much less complex than comparative models. In the basic variant, which has been used in this study, it has only one free parameter (σ) which is easy to estimate. So the learning of the model is fast.

Forecast models supporting monthly coordination plans

For the demands of monthly coordination plans a number of experimental tests for many forecasting models, based on different mathematical tools, have been developed. The most successful ones in terms of accuracy and efficiency are two models: the adjusted predictor of standard load curve (SLC) and the analogue model. Below the basic assumptions concerning the SLC model are presented.

Generally, the model forecasts the load time dependencies at first as typical ones (averages from the past) for a given kind of day. These dependencies are next corrected using the procedures based on auxiliary explanatory variables. The presented model may be classified as the so-called “standard load curve”.

Routine and procedures implementing the SLC model

Determination of the static load dependencies in the input data set is carried out within the following steps [6]:

- Elimination of the details pertaining to season variability from the data set by means of polynomial approximation of average week power using the least square method. Fig. 6 shows such approximation.
- Determination of typical 24-h load profiles for every day of a week in individual months of the year from the input data set, which is devoid of excessive 24 h day sets (those include public holidays, as well as days preceding and following them). Typical profiles are determined as arithmetical averages of static dependencies devoid of season variability.

Table 2
Forecast errors for benchmark models.

Model	January		July		Mean	
	MAPE	IQR	MAPE	IQR	MAPE	IQR
FE	1.22	1.30	0.96	0.89	1.08	1.06
ARIMA	2.64	2.34	1.21	1.24	1.91	1.67
ES	2.35	1.88	1.19	1.30	1.76	1.56
ANN	1.32	1.30	0.97	1.01	1.14	1.15
RF	1.42	1.39	0.92	0.98	1.16	1.17

- Clustering of days excluded in the preceding stage into classes: holidays occurring from Monday to Friday, holidays occurring on Saturdays and Sundays, days preceding holidays from Monday to Friday, days following holidays from Monday to Friday, days preceding and following holidays occurring on Saturdays and Sundays.
- Determination of the polynomial approximations (2nd–3rd order) $\rho = F(\tau)$ for every of the above-described groups for pairs of quantities (τ, ρ_q), where:

$$\tau = \cos\left(n_d \cdot \frac{2\pi}{365}\right), \quad (9)$$

$$\rho_q = \frac{P_q^{statdes}}{P_q^{pprof}}, \quad (10)$$

and n_d is the number of the day in the year, $P_q^{statdes}$ the true static load, devoid of static variability, in the 15-min period q of the considered 24-h period, P_q^{pprof} is the power during the q 15-min period of the typical profile pertaining to the considered 24-h period.

- Creation of secondary input set consisting of proper typical profiles P_q^{pprof} and of products $F_q(\tau) \cdot P_q^{pprof}$ for excessive 24-h time periods.
- Creation of the training set for the radial-basis-function neural network (RBF) from the secondary input set. The components of the input vector – the so-called vector of feature space [27] are the explanatory variables. In the context of this work these are: time (day of year), day of week and kind of day (according to the aforementioned classification of excessive days). Other load determinants, e.g. the forecast weather and incidental factors could also be included here. The output vector is the series of 96 values of correction coefficients ρ'_q for each quarter of the day:

$$\rho'_q = \frac{P_q^{statdes}}{P_q^{pprof}}, \quad (11)$$

where P_q^{pprof} is an appropriate curve from the secondary set.

- For RBF networks it is important to code properly the components of features. A trivial time representation using e.g. the number of month could mislead RBF network, that December, represented with “12” is distant from January, represented with “1” in the feature space. Similarly in the case of days of week, Friday shall be presented to the RBF network as lying closer to Saturday and Sunday than to Monday or Tuesday.
- In the presented model, time is coded using the aforementioned τ variable (approximately proportional to the duration length of night in the considered season of year, whereas day of week and kind of day are coded using long-term average load (static, devoid of season variability).
- The RBF network learning to map the feature space into output vector for setting the prediction of correction coefficients ρ'_q . During learning for determination of centers in the Gaussian basis functions [27] the k -means clusterization method is used, whereas for weight determination of the output layer the ridge regression method with iterative optimization of smoothing parameter [27] is used:

$$\mathbf{w}_q = \mathbf{A}^{-1} \mathbf{H}^T \rho_q, \quad (12)$$

where \mathbf{w}_q is the weight vector of the q th neuron of the output layer, ρ_q the vector of demanded outputs of the q th neuron, \mathbf{H}^T the transposed Jacobean of errors of q th neuron of the network output layer with respect to its weights (for the RBF network with linear output layer, for its every q th neuron, \mathbf{H} is the activation matrix of neurons of the radial layer), \mathbf{A} is the hessian of network errors approximated with the Equation:

$$\mathbf{A} = \mathbf{H}^T \mathbf{H} + \lambda \mathbf{I}, \quad (13)$$

where \mathbf{I} is the unit matrix and λ is the smoothing parameter; its re-estimation formula is

$$\lambda_{i+1} = \frac{\rho_q^T \mathbf{P}^2 \rho_q \text{diag}(\mathbf{A}^{-1} - \lambda_i \mathbf{A}^{-2})}{\mathbf{w}_q^T \mathbf{A}^{-1} \mathbf{w}_q \text{diag}(\mathbf{P})}, \quad (14)$$

where:

$$\mathbf{P} = \mathbf{I} - \mathbf{H} \mathbf{A}^{-1} \mathbf{H}^T, \quad (15)$$

$\text{diag}(\cdot)$ – the sum of elements of the main diagonal.

- Assignment of respective profiles typical for every forecast day.
- Correction of the above-given profiles for the excessive days with the functions $F_q(\tau)$.
- Prediction of correction coefficients ρ_q using the RBF network and their use for definite profile adjustment.
- Carrying out the reverse activities to the procedures of removal of season variability and long term linear trends.
- Incorporation of auxiliary control variables into the model inputs. The correlation analysis is useful for identification of the correct form of these variables and their application scopes.

In order to create the monthly coordination plans, OTS requires information on one month-ahead forecasts of the maximum peak load, secondary peak load and minimum load in every day. The quality assessment of the SLC model is made using the validation set, covering final 24 months in the data set, i.e. from November 2008 to October 2009. The earlier data from January 2000 to October 2008 are used for construct and optimize the model. It is assumed that the forecast procedure is run on the 25th day of the month preceding the forecast period (on 23rd day in the case of forecasts for March) and for making up the forecast the pieces of data concerning system load until 24th (or 22nd for February) day of month, preceding the month under forecast. The average MAPE errors obtained with the SLC model for individual forecast months are depicted in Figs. 7 and 8.

Analyzing the results of forecasts with the SLC method it has to be stated that in this model, the lowest errors are observed for work days from Monday till Friday, whereas the increased ones are observed for Saturdays and Sundays (see Fig. 7). During the year the lowest errors occur in September ($MAPE < 1.5\%$), whereas the highest ones in November and December (see Fig. 8). The average MAPE, calculated for a 2-years-long period (since November 2008 until October 2009), obtained for forecast of the daily peaks were equal to 2.5%, for secondary peaks – 2.9%, while the highest ones, for forecasts of daily minima, were equal to 3.3%.

Forecast model supporting annual coordination plans

The control quantities in the forecast model supporting annual coordination plans are the forecasts made three years ahead of the annual load peaks in the Polish Power System and of annual energy gross.

On the basis the forecasts the average load level in the power system is calculated from the formula:

$$\hat{m}_{rj} = \frac{\hat{A}_j}{T_{rj} \hat{P}_{r \max j}}, \quad (16)$$

where \hat{A}_r is the forecast of annual energy gross, T_r the number of hours in the year (8760 h), $\hat{P}_{r \max}$ the forecast of annual peak load, j is the subsequent number of the year under forecast.

It is also necessary to determine the levels of load variation for monthly peaks, monthly minima, averages from peak days in month for individual months during the year on the basis of the process history, which should be considered the longest possible.

For the experiment verification, the data from the range 1990–2009 were used.

Having determined the quantity of the average annual load level \hat{m}_{rj} for the given year, the forecasts of variation indicators for 36 months ahead concerning the monthly peaks, the monthly minima as well as the averages from peak days during the month are determined from hybrid functions of the form:

$$\hat{\sigma}'_{ijk} = \alpha_{ik} f_i(m_{rj})^{LIN1} w_1 + \beta_{ik} f_i(m_{rj})^{LIN2} w_2 + \chi_{ik} f_i(m_{rj})^{LIN3} w_3, \quad (17)$$

where j is the subsequent number of forecast year $j = 1, 2, \dots, 3$, i the month under forecast $i = 1, 2, \dots, 12$, k -type of the forecast: 1 for peak, 2 for minimum and 3 for average of peaks during the month, $f_i(m_{rj})^{LIN1}$ the linear form of the approximation function determined from the process history for the initial time interval of approximation, $f_i(m_{rj})^{LIN2}$ the linear form of the approximation function determined from the process history for the time interval between the initial approximation instant and the final approximation instant, $f_i(m_{rj})^{LIN3}$ the linear form of the approximation function determined from the process history for the final time interval of approximation, α, β, χ the percentage share coefficients, w_1, w_2, w_3 is the weight coefficients of information aging.

Using any optimization tool the values of parameters α, β, χ for the process history should be determined. In the considered tests the minimum of fitting error MAPE was used as the optimization criterion. As far as the weight coefficients of information aging are concerned, the arbitrary values for relevant intervals from the process history are assumed: $w_1 = 0.1$, $w_2 = 0.5$ and $w_3 = 1.0$. The final individual power values are calculated from the formula:

$$\hat{P}_{mkij} = \hat{\sigma}'_{ijkA} \hat{P}_{r \max j}, \quad \text{for } i = 1, \dots, 12, j = 1, \dots, 3, k = 1, \dots, 3, \quad (18)$$

The afore-described forecasting method for the annual coordination plan was verified on the Polish power system data from the period 1991–2009. The accuracies obtained were as follows for 17 forecast steps, each 36 months long:

- For the forecasts of monthly peaks, the average MAPE error was 1.75% (from 1.26% up to 2.19% in 17 steps).
- For the forecasts of monthly minima the average MAPE error was 2.98% (from 1.97% up to 3.97% in 17 steps).
- For the forecasts of monthly average peaks in 24-h period the average MAPE error was 1.50% (from 0.96% up to 1.89% in 17 steps) (see Figs. 6–8).

Figs. 9–11 depict the obtained fitting accuracy for the forecast values of peak power, minimal power and average value of the peak from peaks coming from working days, obtained with the use of the afore-described model. An ex-post step prediction for each of the examined quantities was carried out for the years 1991–2009. For 17 forecasting steps, 36 month-long each, we obtained 612 MAPE for each quantity.

Summary

In this article we present the forecasting methods for preparing the daily, monthly and annual coordination plans. The similarity-based fuzzy forecasting model, applied to the daily coordination plans, is characterized by simplicity and high accuracy. This model turns out to be one of the best in comparison to other methods tested by us earlier, including neural networks (multilayer perceptron, Kohonen, counterpropagation, RBF, GMDH), regression trees, random forests, neuro-fuzzy nets, cluster analysis methods, naïve methods, nearest neighbor method, kernel estimators and artificial immune system. A comparison of many similarity-based methods is reported in [8].

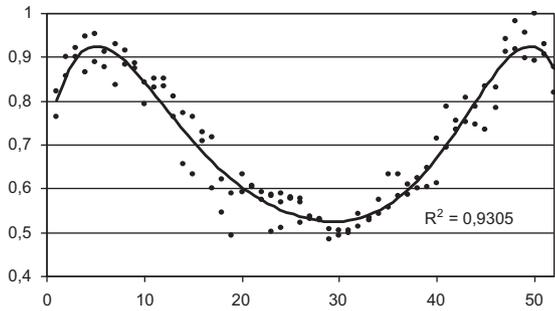


Fig. 6. Season variability of static average week power.

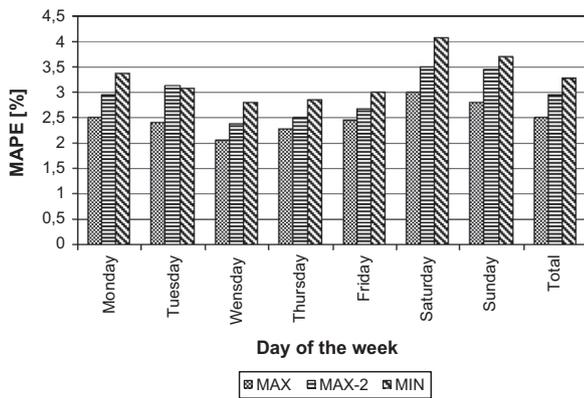


Fig. 7. The average MAPE errors obtained with the SLC model, partitioned into days of week and weekly averages.

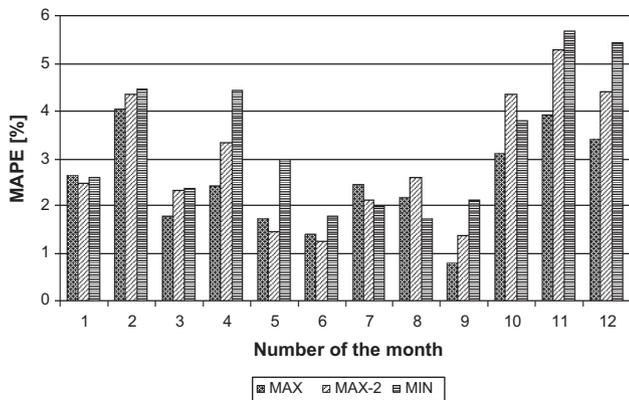


Fig. 8. The average MAPE errors obtained using the SLC model, partitioned into months.

The model based on fuzzy estimators of the regression function has only two user defined parameters – the widths of the membership functions. These parameters are easy to estimate, and the model sensitivity to their deviation from the optimal values is limited [8]. Models with a fewer number of parameters have better generalization properties. The additional input information, e.g. weather factors, can be incorporated to the model with the help of contexts. The model demonstrates low sensitivity to incomplete and noisy input information [8].

The advantages of the SLC method used for the monthly coordination plans is its scalability. It can be used successfully for systems of all sizes and for a forecast horizon from days to months.

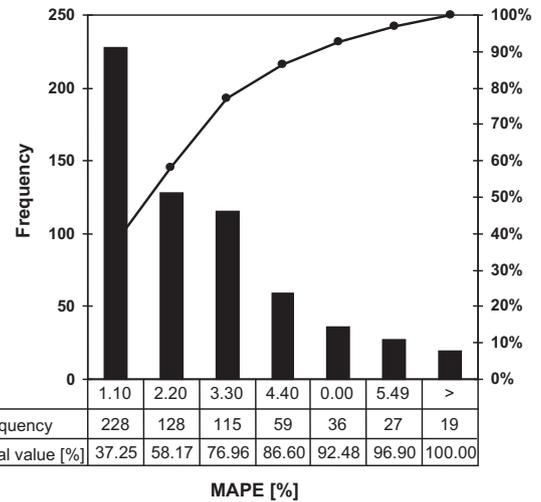


Fig. 9. Error distribution of forecast peak power for subsequent 36 months.

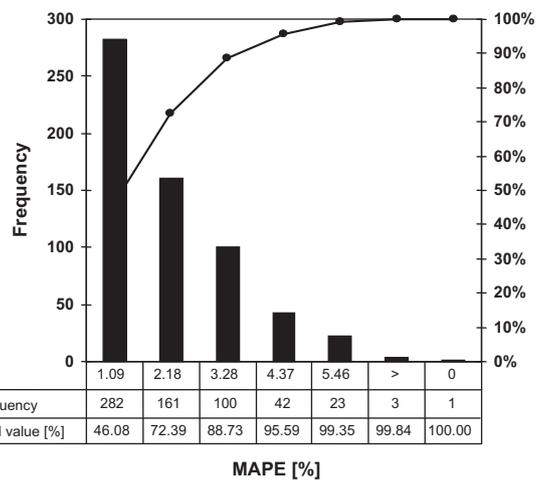


Fig. 10. Error distribution of forecast average peak days for subsequent 36 months.

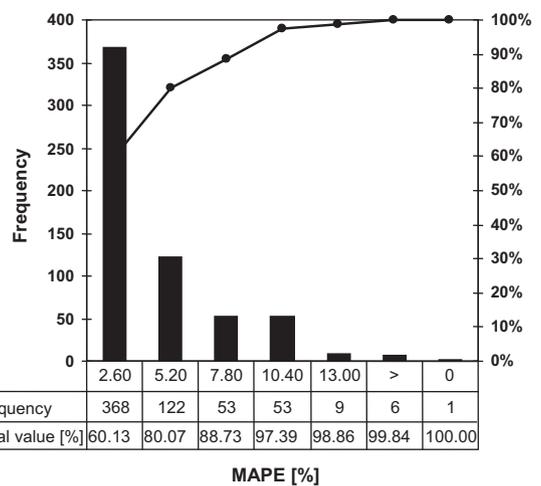


Fig. 11. Error distribution of forecast minimum load power for subsequent 36 months.

For the short horizons this method gives satisfactory results with relatively low computational cost. In addition, the algorithm can be simplified by omitting the whole stage for the use of RBF. Then

the SLC method becomes similar to the analog methods while maintaining the advantages for the untypical days. When we use SLC with RBF we can include additional explanatory variables to the model in an easy way.

The weakness of the SLC method is its high sensitivity to both the temporary and permanent disturbance of the process. In this case, additional procedures are required for transforming the historical data, which complicates the usage. In comparison with other methods adaptability of SLC is limited, and the method has a larger inertia.

The forecasting model used for the annual coordination plans shows a good fit to the historical data, what may be proven with charts shown in Figs. 9–11. Assuming that the relationships determined from the historical 19-year period do not change in the future, it may be expected that the forecasts for the successive years will be just as accurate.

References

- [1] Bashir ZA, El-Hawary ME. Applying wavelets to short-term load forecasting using PSO-based neural networks. *IEEE Trans Power Syst* 2009;24(1):20–7.
- [2] Božić M, Stojanović M, Stajić Z, Floranović N. Mutual information-based inputs selection for electric load time series forecasting. *Entropy* 2013;15:926–42.
- [3] Chang P-C, Fan C-Y, Lin J-J. Monthly electricity demand forecasting based on a weighted evolving fuzzy neural network approach. *Int J Electr Power Energy Syst* 2011;33(1):17–27.
- [4] Christianse W. Short term load forecasting using general exponential smoothing. *IEEE Trans PAS* 1971;900–910:1971.
- [5] Darbellay GA, Slama M. Forecasting the short-term demand for electricity—do neural networks stand a better chance? *Int J Forecast* 2000;16:71–83.
- [6] Dobrzańska I, Dąsał K, Łyp J, Popławski T, Sowiński J. Forecasting in power energy. Selected problems. Publisher Czestochowa University of Technology. Czestochowa; 2002.
- [7] Dudek G. Artificial immune system for short-term electric load forecasting. In: Rutkowski L, Tadeusiewicz R, Zadeh L, Zurada J, editors. *Lecture notes in artificial intelligence – proceedings of the 9-th international conference on artificial intelligence and soft computing ICAISC 2008*, vol. 5097. Springer; 2008. p. 1007–17.
- [8] Dudek G. Similarity-based approaches to short-term load forecasting. In: *Forecasting models: methods and applications*. iConcept Press; 2010a. p. 161–78. <<https://www.iconceptpress.com/download/paper/100917020141.pdf>>.
- [9] Dudek G. Short-term load forecasting based on kernel conditional density estimation. *Przegląd Elektrotech (Electr Rev)* 2010;86(8):164–7.
- [10] Dudek G. Short-term electrical load forecasting using fuzzy clustering method. *Electr Rev* 2006;9:26–8 [in Polish].
- [11] Dudek G. Forecasting time series with multiple seasonal cycles using neural networks with local learning. In: Rutkowski L, et al., editors. *Artificial intelligence and soft computing, ICAISC 2013, LNCS 7894*; 2013. p. 52–63.
- [12] Dudek G. Short-Term load forecasting using random forests. *Advances in Intelligent Systems and Computing, IS 2014, LNCS*; 2014 [in press].
- [13] Engle RF, Mustafa C, Rice J. Modeling peak electricity demand. *J Forecast* 1992;11:241–51.
- [14] Goia A, May C, Fusai G. Functional clustering and linear regression for peak load forecasting. *Int J Forecast* 2010;26:700–11.
- [15] Gonzalez-Romera E, Jaramillo-Moran MA, Carmona-Fernandez D. Monthly electric energy demand forecasting with neural networks and Fourier series. *Energy Convers Manage* 2008;49:3135–42.
- [16] Gross G, Galiana FD. Short-Term Load Forecasting. *Proc IEEE* 1987;75(12):1558–73.
- [17] Härdle W. *Applied nonparametric regression*. Cambridge University Press; 1994.
- [18] Ho K-L, Hsu Y-Y, Chen C-F, Lee T-E, Liang C-C, Lai T-S, et al. Short-term load forecasting of taiwan power system using a knowledge based expert system. *IEEE Trans Power Syst* 1990;5:1214–21.
- [19] Hong W-C. Application of chaotic ant swarm optimization in electric load forecasting. *Energy Policy* 2010;38:5830–9.
- [20] Khotanzad A, Afkhami-Rohani R, Lu TL, Davis MH, Abaya A, Maratukulam D. ANNSTLF – a neural network based electric load forecasting system. *IEEE Trans Neural Netw* 1997;8(4):835–46.
- [21] Kiartzis SJ, Bakirtzis AG. A fuzzy expert system for peak load forecasting. Application to the greek power system. In: *10-th Mediterranean electrotechnical conference MELECON 2000*; 2000. p. 1097–100.
- [22] Lee C-M, Ko C-N. Short-term load forecasting using lifting scheme and ARIMA models. *Expert Syst Appl* 2011;38:5902–11.
- [23] Łyp J. Methodology of the electric load analysis and forecasting for the local power systems. Ph.D. Dissertation, Dept. Elect. Eng., Czestochowa University of Technology; 2003 [in Polish].
- [24] Mamlook R, Badran O, Abdulhadi E. A fuzzy inference model for short-term load forecasting. *Energy Policy* 2009;37:1239–48.
- [25] Metaxiotis K, Kagiannas A, Askounis D, Psarras J. Artificial intelligence in short term load forecasting: a state-of-the-art survey for the researcher. *Energy Convers Manage* 2003;44:1525–34.
- [26] Niu D, Wang Y, Wu DD. Power load forecasting using support vector machine and ant colony optimization. *Expert Syst Appl* 2010;37:2531–9.
- [27] Orr M. Introduction to radial basis function networks. Edinburgh; 1996. <<http://www.anc.ed.ac.uk/rbf/papers/intro.ps.gz>>.
- [28] Papalexopoulos AD, Hesterberg TC. A regression based approach to short term load forecasting. *IEEE Trans Power Syst* 1990;5:1535–47.
- [29] Peng M, Hubele NF, Karady GG. Advancement in the application of neural networks for short-term load forecasting. *IEEE Trans Power Syst* 1992;7:250–7.
- [30] Popławski T, Dąsał K. Canonical distribution in short-term load forecasting. In: Szkutnik J, Kolcun M, editors. *Technical and economic aspect of modern technology transfer in context of integration with European Union*, Mercury-Smekal Pub House, Kosice, Slovak Republic; 2004. p. 147–153.
- [31] Popławski T, Dudek G, Łyp J, Sowiński J, Dąsał K, Starczynowska E. Development and validation of models for forecasting demand for electrical power in the Polish power system for development at PSE Operator SA coordination plans. Technical report, Center for Advanced Technology Application (CATA), Warsaw; 2010.
- [32] PSE Operator S.A. Instruction for operation and maintenance of the transmission network; 2014. <<http://www.pse-operator.pl>> [in Polish].
- [33] Rahman S, Bhatnager R. An expert system based algorithm for short-term load forecast. *IEEE Trans Power Syst* 1988;3:392–9.
- [34] Song K-B, Ha S-K, Park J-W, Kweon D-J, Kim K-H. Hybrid load forecasting method with analysis of temperature sensitivities. *IEEE Trans Power Syst* 2006;21(2):869–76.
- [35] Taylor JW, Menezes LM, McSharry PE. A comparison of univariate methods for forecasting electricity demand up to a day ahead. *Int J Forecast* 2006;22(1):1–16.
- [36] Taylor JW. Triple seasonal methods for short-term electricity demand forecasting. *Eur J Oper Res* 2010;204:139–52.
- [37] Ünler A. Improvement of energy demand forecasts using swarm intelligence: the case of Turkey with projections to 2025. *Energy Policy* 2008;36:1937–44.
- [38] Yun Z, Quan Z, Caixin S, Shaolan L, Yuming L, Yang S. RBF neural network and ANFIS-based short-term load forecasting approach in real-time price environment. *IEEE Trans Power Syst* 2008;23(3):853–8.