Neuro-Fuzzy System for Medium-Term Electric Energy Demand Forecasting

Paweł Pełka(✉) and Grzegorz Dudek

Department of Electrical Engineering, Czestochowa University of Technology, Czestochowa, Poland
pavelle50@gmail.com, dudek@el.pcz.czest.pl

Abstract. Medium-term electric energy demand forecasting plays an important role in power system planning and operation as well as for negotiation forward contracts. This paper proposes a solution to medium-term energy demand forecasting that covers definition of input and output variables and the forecasting model based on a neuro-fuzzy system. As predictors patterns of the yearly periods of the time series are defined, which unify input data and filter out the trend. Output variable is encoded in tree ways using coding variables describing the process. For prediction of coding variables, which are necessary for postprocessing, ARIMA and exponential smoothing models are applied. The simplified relationship between preprocessed input and output variables is modeled using Adaptive-Network-Based Fuzzy Inference System. As an illustration, we apply the proposed time series forecasting methodology to historical monthly energy demand data in four European countries and compare its performance to that of alternative models such as ARIMA, exponential smoothing and kernel regression. The results are encouraging and confirm the high accuracy of the model and its competitiveness compared to other forecasting models.

Keywords: Neuro-Fuzzy systems · Medium-term load forecasting · Pattern-based forecasting

1 Introduction

Accurate medium-term electric energy demand forecasting plays an essential role for electric power system planning and operation, and offers significant benefits for companies operating in a regulated and deregulated energy markets. Generally, it is used to optimize energy production and transmission and improve power system reliability. It is necessary for scheduling and coordinate maintenance and production across a power system, negotiation fuel purchases for power stations, and optimization of renewable energy sources, such as wind farms.

Characteristic feature of the electricity demand time series is the yearly seasonality corresponding to climatic factors and weather variations. Also a trend is observed following the economic and technological development of a country, and random component disturbing the time series. These features should be included in the forecasting process to increase the prediction accuracy.

L. Borzemski et al. (eds.), *Information Systems Architecture and Technology: Proceedings of 38th International Conference on Information Systems Architecture and Technology – ISAT 2017*, Advances in Intelligent Systems and Computing 655, DOI 10.1007/978-3-319-67220-5_4

The medium-term load forecasting (MTLF) methods can be categorized into two general groups [1]. The first group, referred to as the conditional modeling approach, focuses on economic analysis, management and long term planning and forecasting of energy load and energy policies. It takes into account changes in socioeconomic conditions which impact energy demands. Input variables include historical load data and weather factors as well as economic indicators and electrical infrastructure measures. Efforts of researchers in this field are focused on definition of optimal set of input variables and construction of appropriate forecasting models. An example of a model of this type can be found in [2], where macroeconomic indicators, such as the consumer price index, average salary earning and currency exchange rate are taken into account as inputs.

The second group, referred to as the autonomous modeling approach, requires a smaller set of input information to forecast future electricity demand. Primarily past loads and weather variables. This approach is more suited for stable economies. The forecasting models used in this group include classical methods such as ARIMA or linear regression [3] as well as computational intelligence methods, such as neural networks [4, 5] and support vector machines [6].

The forecasting model proposed in this work can be classified to autonomous modeling approach. It uses neuro-fuzzy network which works on preprocessed data. Inputs are defined as patterns of yearly fragments of the demand time series, which are normalized version of demand vectors. Outputs are encoded demands. The proposed way of time series preprocessing unifies data and filters out a trend.

The remaining sections of the paper are organized as follows. In Sect. 2, time series representation is described. Section 3 presents in detail a neuro-fuzzy forecasting model. In Sect. 4, we evaluate the performance of the model in monthly electricity demand forecasting using real-world data. Finally, Sect. 5 is a summary of our conclusions.

2 Time Series Representation

Let us consider the task of prediction of the monthly electricity demand with horizon τ. The predicted time series point is $E_{i+\tau}$. As predictors we use preprocessed n points preceding the forecasted point, i.e. the time series fragment $X_i = \{E_{i-n+1}, E_{i-n+2}, \ldots, E_i\}$. This fragment is represented by input pattern $\mathbf{x}_i = [x_{i,1}\ x_{i,2}\ \ldots\ x_{i,n}]^T$. The components of this vector are defined as follows [7]:

$$x_{i,t} = \frac{E_{i-n+t} - \bar{E}_i}{D_i} \tag{1}$$

where: $t = 1, 2, \ldots, n$, $\bar{E}_i$ is the mean value of the points in sequence X_i, and $D_i = \sqrt{\sum_{j=1}^{n} (E_{i-n+j} - \bar{E}_i)^2}$ is a measure of their dispersion.

Note, that the x-pattern defined using the above equation is a normalized vector $[E_{i-n+1}\ E_{i-n+2}\ \ldots\ E_i]^T$. It has the unity length and the mean value equal to zero. Moreover, x-patterns representing different fragments have the same variance. Thus, the time series, which is nonstationary, is represented by unified patterns, having the same mean and variance. The trend is filtered out. When $n = 12$ the x-pattern carries information about the shape of the yearly cycle.

The output variable, $E_{i+\tau}$, is encoded in three ways. In the first approach (C1) the forecasted value is encoded as follows:

$$y_i = \frac{E_{i+\tau} - \bar{E}_i}{D_i} \tag{2}$$

where $\bar{E}_i$ and D_i are determined from the sequence X_i.

Note, that $\bar{E}_i$ and D_i are known at the moment of making the forecast (moment i) and can be used for calculation the forecast of demand based on the forecast of y_i returned by the forecasting model. We use for this the transformed Eq. (2):

$$\widehat{E}_{i+\tau} = \widehat{y}_i D_i + \bar{E}_i \tag{3}$$

In the second approach (C2) $\bar{E}_i$ and D_i are determined from the annual period following the period X_i, i.e. the period including time series fragment $\{E_{i+1}, E_{i+2}, \ldots, E_{i+12}\}$. This approach is used for forecast horizon $\tau \in \{1, 2,\ldots, 12\}$. Note, that in this case coding variables $\bar{E}_i$ and D_i are not available at the time of making the forecast. Thus, they should be forecasted. In the experimental part of the work we use ARIMA and exponential smoothing (ES) for forecasting the coding variables.

In the third approach (C3), which is used only for one-step ahead forecasts ($\tau = 1$), the coding variables $\bar{E}_i$ and D_i are determined from the annual period including time series fragment $\{E_{i-n+2}, E_{i-n+3}, \ldots, E_{i+1}\}$. In this case coding variables cannot be calculated from time series elements because the value of E_{i+1} is not known. Thus, $\bar{E}_i$ and D_i should be predicted. Just like in the case of C2, we use for this ARIMA and ES.

3 Neuro-Fuzzy Forecasting System

The proposed forecasting model is based on Adaptive-Network-Based Fuzzy Inference System (ANFIS) developed by Jang [8]. This is a multi-input, single-output quasi-nonlinear model consisting of a set of linguistic if-then rules. Its architecture is functionally equivalent to a Sugeno type fuzzy rule base. In Fig. 1 ANFIS architecture in application to the energy demand forecasting is shown. Squares in this figure indicate adaptive nodes, whereas circles indicate fixed nodes (without parameters).

The functions of ANFIS nodes in subsequent layers are described below.

Layer 1. A node represents a membership function. In our case this is a Gaussian function of the form:

$$\mu_{A_k^m}(x_k) = \exp\left[-\left(\frac{x_k - c_k^m}{\sigma_k^m}\right)^2\right] \tag{4}$$

where $m = 1, 2, \ldots, M$ is the fuzzy set number, $k = 1, 2, \ldots, 12$ is the x-pattern component number, A_k^m is the fuzzy set describing linguistically the input component x_k, c_k^m and σ_k^m are premise parameters: center and spread, respectively.

An output of the node expresses a membership degree of x_k in the fuzzy set A_k^m. The number of nodes is determined by the number of linguistic labels M and the x-pattern length.

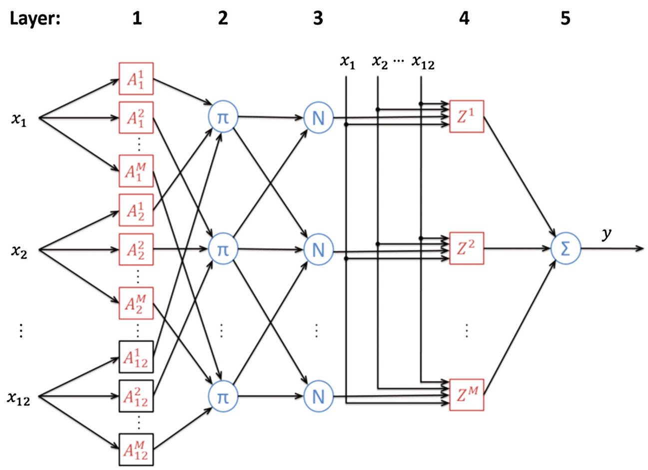

Fig. 1. ANFIS architecture.

Layer 2. A node expresses a firing strength of the mth rule. It is calculated as the product t-norm:

$$\alpha^m = \prod_{k=1}^{12} \mu_{A_k^m}(x_k) \tag{5}$$

The firing strength of the mth rule depends on the membership degree of each x-pattern component in the relevant fuzzy set. It has the highest value if the x-pattern components coincide with the centers of the membership functions.

Layer 3. A node expresses a normalized firing strength for the mth rule:

$$\bar{\alpha}^m = \frac{\alpha^m}{\sum_{j=1}^{M} \alpha^j} \tag{6}$$

Layer 4. A node expresses a conclusion of the mth rule. Conclusion is determined using the Takagi-Sugeno-Kang method, where the output membership functions are either linear or constant (first or zeroth order Sugeno-type systems). Each rule weights its output level by the firing strength of the rule. The node function for the first order system is of the form:

$$z^m = \bar{\alpha}^m \left(\sum_{k=1}^{12} a_k^m x_k + b^m \right) \tag{7}$$

where a_k^m and b^m are consequent parameters.

Layer 5. A node computes the overall system output as the sum of all incoming signals:

$$y = \sum_{m=1}^{M} z^m = \frac{\sum_{m=1}^{M} \alpha^m \left(\sum_{k=1}^{12} a_k^m x_k + b^m \right)}{\sum_{j=1}^{M} \alpha^j} \tag{8}$$

Note, that the output of each rule is a linear combination of inputs and the final output of the system is the weighted average of all rule outputs. Because the membership functions are nonlinear in our case, the weights (firing strengths) are dependent on inputs nonlinearly, and the final output is nonlinear.

I our case the rule base is of the form:

$$\begin{aligned}
&\text{If } x_1 \text{ is } A_1^1 \text{ and}\ldots\text{and } x_{12} \text{ is } A_{12}^1 \text{ then } z^1 = \sum_{k=1}^{12} a_k^1 x_k + b^1 \\
&\text{If } x_1 \text{ is } A_1^2 \text{ and}\ldots\text{and } x_{12} \text{ is } A_{12}^2 \text{ then } z^2 = \sum_{k=1}^{12} a_k^2 x_k + b^2 \\
&\ldots \\
&\text{If } x_1 \text{ is } A_1^M \text{ and}\ldots\text{and } x_{12} \text{ is } A_{12}^M \text{ then } z^M = \sum_{k=1}^{12} a_k^M x_k + b^M
\end{aligned} \tag{9}$$

The premise part of each rule defines a fuzzy region for the linear model included in the consequence part. The inference mechanism interpolates smoothly between each local model to provide a global model.

Before the training, an initialization of ANFIS is required. Initial positions of the membership functions in the premise parts of rules are determined using fuzzy c-means clustering on the input data. The number of clusters corresponding to the numbers of rules M is selected in leave-one-out cross-validation procedure. This parameter decides about the bias–variance tradeoff. Increasing M we increase the model variance and decrease its bias.

For ANFIS training, i.e. estimation of the premise and consequent parameters, a hybrid learning algorithm is applied. It uses a combination of the least-squares and backpropagation gradient descent methods to model the training data set. The error measure minimized during training is defined by the sum of the squared difference between actual and desired outputs.

4 Application Example

In order to assess the performance of the proposed forecasting model to obtain generalized conclusions, we use it to forecast monthly electricity demand for four European countries: Poland (PL), Germany (DE), Spain (ES) and France (FR). The data used for the experiments were retrieved from the ENTSO-E repository (www.entsoe.eu). The datasets contain monthly electricity demand from the period 1998-2015 for PL and 1991-2015 for other countries. The forecasts are made for 2015, using data from previous years as training data. The only model parameter is M (number of rules in ANFIS). It was selected for each ANFIS model from the range 2-13 in the leave-one-out cross-validation.

The forecasts were generated in two procedures. In the first procedure (A) the forecasts for successive 12 months of 2015 are generated by 12 ANFIS models. Each model gets the same input pattern representing time series fragment from January to December 2014, and produces a forecast for kth month of 2015 (k = 1, 2,…, 12). Thus, the forecast horizon for the model for January 2015 is $\tau = 1$, for the model for February 2015 is $\tau = 2$, etc. The output variable is encoded using C1 or C2 approach. In the latter case coding variables $\bar{E}_i$ and D_i for 2015 are predicted using ARIMA and ES on the basis of their historical values, i.e. values determined for 12 months of the successive years. The results of forecasting in Fig. 2 are shown.

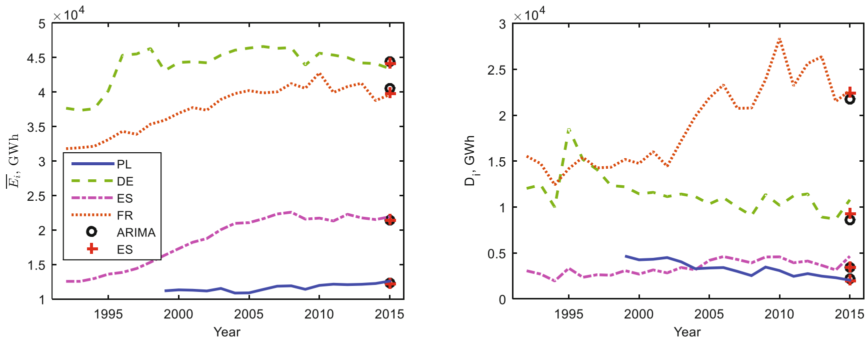

Fig. 2. Forecasts of coding variables.

In the second procedure of forecasting (B) one-step ahead forecasts are generated ($\tau = 1$) for successive months of 2015. The input patterns for the models represent 12 preceding months, i.e. the model for January 2015 gets input pattern representing time series fragment from January to December 2014, the model for February 2015 gets input pattern representing time series fragment from February 2014 to January 2015, etc. The output variable is encoded using C1 or C3 approach. The latter case needs the coding variables $\bar{E}_i$ and D_i to be predicted. We use for this ARIMA and ES, as in the A procedure.

The real and forecasted values of monthly demand are presented in Figs. 3 and 4, and errors for each month of the test period in Figs. 5 and 6. Forecast errors for validation and test samples in Tables 1 and 2 are presented. In these tables results of comparative models are also shown: ARIMA, ES and Nadaraya-Watson estimator (N-WE) [7]. As can be seen from the figures and tables, it is hard to select the best model variant. For PL data the lowest errors gives variant C1 for both A and B forecasting procedures, and the worst variant is C2-ES. C1 is also the best for DE, variant B. For three out of eight cases the variants using ES, i.e. C2-ES and C3-ES, turned out to be the most accurate among the proposed ones. And variants using ARIMA were the best in two cases. When comparing errors of all models, it should be noted that the classical ES model outperformed all other models in five of eight cases.

Comparing results for A and B procedures we can conclude that variant B which generates one-step ahead forecasts, usually provides better results than variant A. An exception is FR data, where higher errors in variant B are observed. Due to significant contribution of the random component in the time series the errors for successive months are very varied. Different level of heteroscedasticity in time series for different countries and also occurrence of nonlinear trend in some of them cause deterioration in the model accuracy.

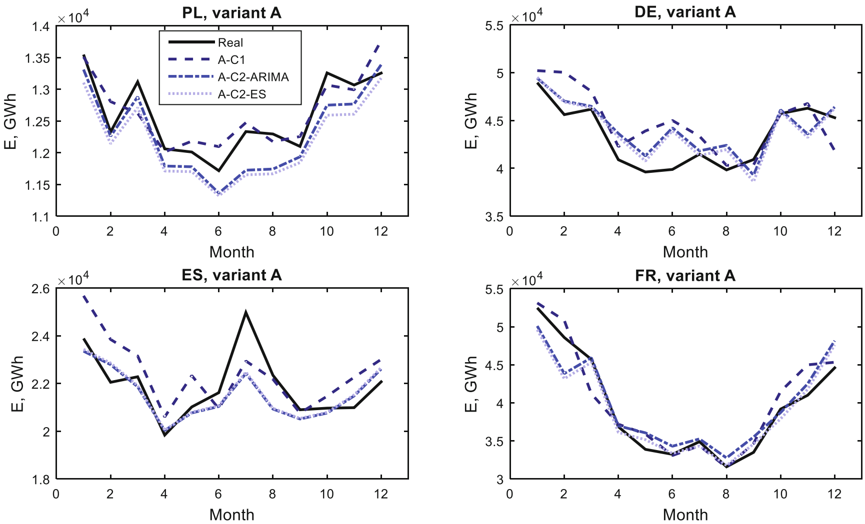

Fig. 3. Real and forecasted monthly demand for 2015, variant A.

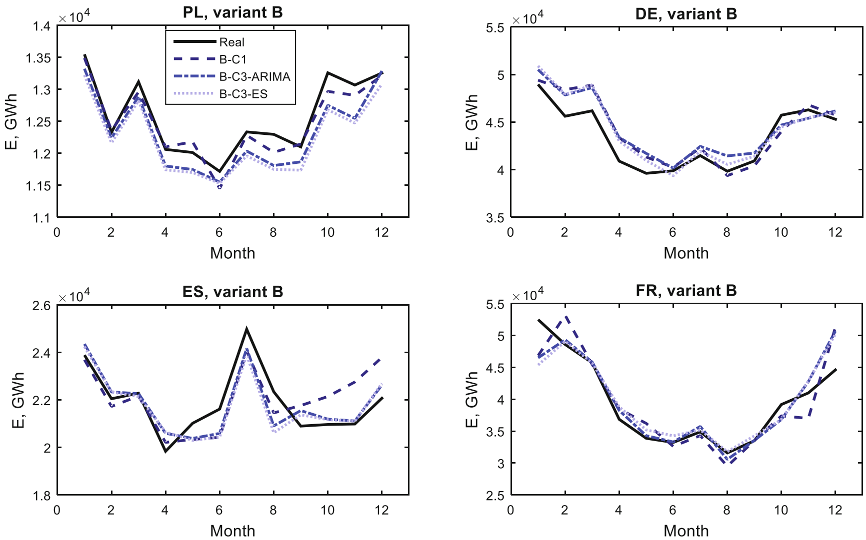

Fig. 4. Real and forecasted monthly demand for 2015, variant B.

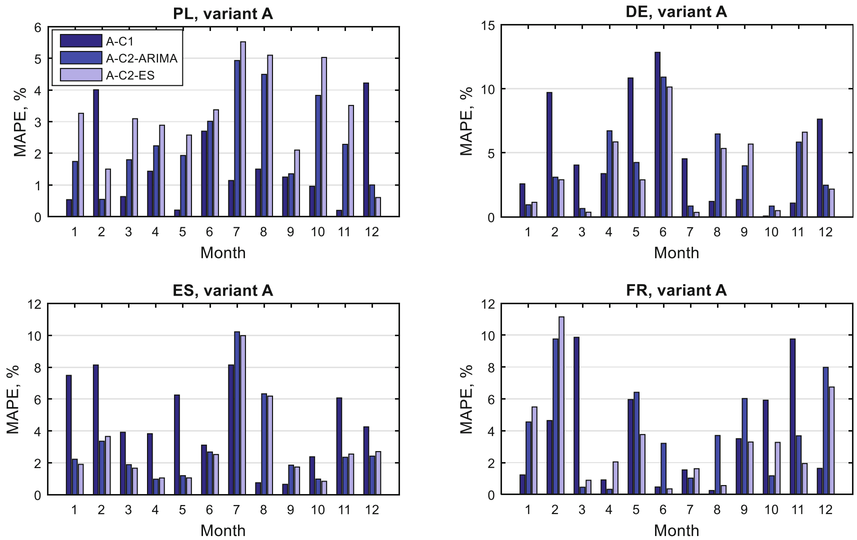

Fig. 5. Errors for consecutive months of 2015, variant A.

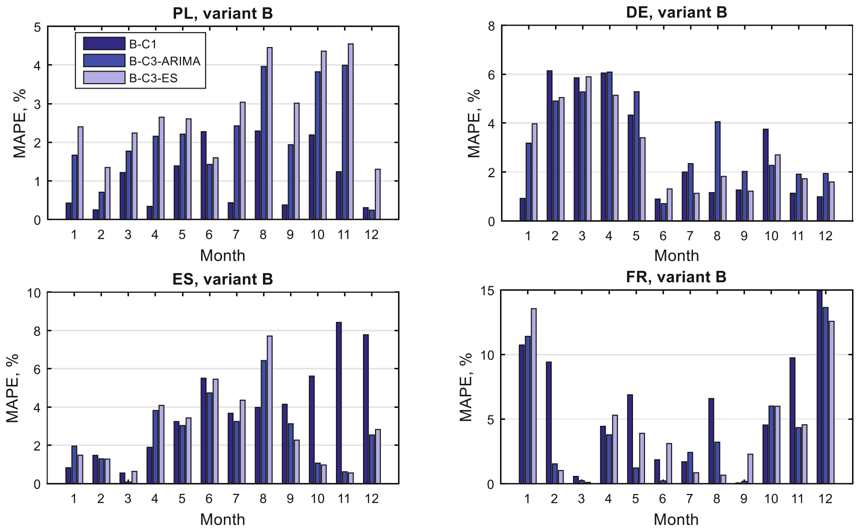

Fig. 6. Errors for consecutive months of 2015, variant B.

Table 1. Forecast errors, variant A.

	PL		DE		ES		FR	
	$MAPE_{val}$	$MAPE_{tst}$	$MAPE_{val}$	$MAPE_{tst}$	$MAPE_{val}$	$MAPE_{tst}$	$MAPE_{val}$	$MAPE_{tst}$
A-C1	2.27	1.57	2.96	4.93	2.63	4.58	3.57	3.81
A-C2-ARIMA	1.57	2.43	1.82	3.92	2.00	3.04	2.64	4.02
A-C2-ES	1.57	3.21	1.82	3.66	2.00	2.99	2.64	3.43
ARIMA	–	2.08	–	2.54	–	2.67	–	4.02
ES	–	1.92	–	2.32	–	2.17	–	3.02
N-WE	–	2.03	–	3.12	–	2.08	–	3.56

Table 2. Forecast errors, variant B.

	PL		DE		ES		FR	
	$MAPE_{val}$	$MAPE_{tst}$	$MAPE_{val}$	$MAPE_{tst}$	$MAPE_{val}$	$MAPE_{tst}$	$MAPE_{val}$	$MAPE_{tst}$
B-C1	2.04	1.06	2.70	2.87	2.33	3.92	3.14	5.95
B-C3-ARIMA	1.67	2.19	2.32	3.33	1.92	2.66	2.64	4.01
B-C3-ES	1.67	2.80	2.32	2.91	1.92	2.92	2.64	4.49
ARIMA	–	2.02	–	2.56	–	2.18	–	3.91
ES	–	1.92	–	2.32	–	2.16	–	2.98
N-WE	–	1.35	–	2.72	–	3.42	–	3.99

5 Conclusion

This work presents an ANFIS model which is used for medium-term electric demand forecasting. The model works on preprocessed time series fragments - patterns of yearly cycles. The patterns unify data and reduce nonstationarity. The novelty of this work is that output variable is encoded in three ways using coding variables determined from history or forecasted using ARIMA or exponential smoothing. The advantage of the ANFIS is that despite the complex structure, there is only one parameter to be tuned – the number of rules.

In the light of the experimental study, it can be concluded that neuro-fuzzy inference models have been proven to be useful in medium-term load forecasting. Their accuracy depend on time series features. For PL data set ANFIS model in its basic variant (C1) provided the best results. But for other data sets other models generated better results. In the future work, we are going to test the ANFIS forecasting model thoroughly in medium-term electric demand forecasting for other European countries.

References

1. Ghiassi, M., Zimbra, D.K., Saidane, H.: Medium term system load forecasting with a dynamic artificial neural network model. Electr. Power Syst. Res. **76**, 302–316 (2006)
2. Gavrilas, M., Ciutea, I., Tanasa, C.: Medium-term load forecasting with artificial neural network models. In: 16th International Conference and Exhibition on Electricity Distribution (2001)
3. Hor, C.L., Watson, S., Majithia, S.: Analyzing the impact of weather variables on monthly electricity demand. IEEE Trans. Power Syst. **20**, 2078–2085 (2005)
4. González-Romera, E., Jaramillo-Morán, M., Carmona-Fernández, D.: Monthly electric energy demand forecasting based on trend extraction. IEEE Trans. Power Syst. **21**, 1946–1953 (2006)
5. Chen, J.F., Lo, S.K., Do, Q.H.: Forecasting monthly electricity demands: an application of neural networks trained by heuristic algorithms. Information **8**, 31 (2017)
6. Hu, Z., Bao, Y., Chiong, R., Xiong, T.: Mid-term interval load forecasting using multi-output support vector regression with a memetic algorithm for feature selection. Energy **84**, 419–431 (2015)
7. Dudek, G., Pełka, P.: Medium-term electric energy demand forecasting using Nadaraya-Watson estimator. In: 18th International Scientific Conference on Electric Power Engineering (EPE), pp. 1–6, (2017)
8. Jang, J.S.R.: ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man Cybern. **23**(3), 665–685 (1993)