STUDIA INFORMATICA
Nr 1-2 (24) Systems and information technology 2020

Dariusz RUCIŃSKI[1]

[1] Siedlce University of Natural Sciences and Humanities
Faculty of Exact and Natural Sciences
Institute of Computer Science
ul. 3 Maja 54, 08-110 Siedlce, Poland

Neural modelling of electricity prices quoted on the Day-Ahead Market of TGE S.A. shaped by environmental and economic factors

DOI: 10.34739/si.2020.24.02

Abstract. The paper contains the results of research on the impact of the number of factors used to build the Day-Ahead Market model at Polish Power Exchange S.A. Five models with a different number of factors influencing the model were tested. To test the quality of models according to the adopted evaluation criteria, i.e., mean square error and the coefficient of determination for the weighted average prices sold in a given hour of the day, the influence of weather factors, socio-economic factors and energy demand were adopted. The results obtained from the analysis show a relatively high correctness of the simplest of the adopted models, which differs slightly from the best model.

Keywords. Polish Power Exchange, Day Ahead Market, Artificial Neural Network, System Modelling, MATLAB.

1. Introduction

Modelling, with the use of mathematical methods, of specific real systems functioning in a given field of economic life, etc., usually causes many problems related to the selection of factors that have a fundamental impact on its functioning. The essence of modelling is a generalization and simplification of the real system due to the computational capabilities, access to valued data or the level of knowledge or cognitive capabilities of the mechanisms of the real system itself. In the case of the Day-Ahead Market (DAM) operating since mid-2002 on 'Polish Power Exchange S.A.' possess a similar nature.

The purpose of this article is to investigate the impact of the above-mentioned factors on the DAM system model and determination of the minimum number of parameters for which the DAM system model is possible to build.

MATLAB environment was used as a tool to carry out the above-mentioned research. The criteria used for evaluating examining models are mean square error (MSE) and the coefficient of determination R^2.

After analysing the functioning of the DAM system, some economic conditions were selected that have a large impact on the economic situation and, consequently, the demand for energy, which in turn affects the price, the following factors were selected: the level of inflation, the amount of debt, the balance of state expenditure and the money supply. The next group of factors analysed were environmental factors such as temperature, air humidity, cloud cover and wind force.

2. Methodology of the study

The process of selecting factors for model construction is the subject of many studies, e.g. [3] where multi-variational models were focused, in order to estimate the number of significant factors. A methodology for estimating the number of factors was proposed, using the adopted criteria of convergence, in order to estimate them, or a comprehensive statistical analysis of classification issues, correlation between selected factors, representativeness of the examined data set, etc. [17]. The aim of this paper is not a methodological analysis of the construction of the model, but only, on the basis of the developed methods, to determine the number of factors needed for the construction of the model of the Day-Ahead Market on the Polish Power Exchange S.A.

The process of selecting the number of criteria consists in selecting them in such a way that they are independent of each other, i.e., that there is no correlation between them. The natural direction of the search is to minimize the number of factors sufficient to build the model. When building a model, it is necessary to select an appropriate number of non-dependent factors, where:

1) too few factors cause the model to be undefined, it is usually over fitted,
2) too small number of factors causes that the models are usually less precise,

3) with a well-chosen number of factors, the models maintain a balance between the ability to generalize knowledge and its accuracy.

An important factor, in addition to linear independence, is also the weight of the influence of a given factor on the model.

3. Description of the modelled system

Day-ahead market modelling using both different versions of ANN [1, 12, 20] as well as other modelling methods [5, 8, 9, 10, 11, 13] is not a new issue and has been descripted in many papers. They referred both to finding the better model than the previous ones for classic power plants as well as for photovoltaic or wind farms. In the paper [18] the ANN model was proposed with the use of a correction factor which is temperature as a factor determining energy consumption. According to the authors, this allows for taking into account short-term changes in demand in the market model. Another proposal is presented in the paper [6], where an attempt was made to find sensitive parameters, and then reduce the number of factors for ANN. In turn, the article [16] gives a way to look for an improvement in the market model in the form of ANN by modifying the model using the method of backward propagation of error in order to make it resistant to unusual, extreme and boundary data.

An important issue in each of the cited papers is the selection of the number of factors modelling the system. This paper focuses on the analysis of the impact of their number on the quality of the obtained model.

4. Factors subject to examination

In each case, the basis for constructing the model is the knowledge of the factors affecting the subject being studied. Selecting factors is a process that requires detailed analysis and knowledge of the nature of a given phenomenon. There is also a need to parameterize or estimate them so that the impact of a given factor on the behaviour of the object of analysis can be quantitatively estimated. An important issue is also the appropriate selection of the learning data set so that it is representative, eliminating the so-called disturbances at the stage of initial analyst [14]. In the case of DAM, the factors affecting both price and power demand partly overlap and partly different as a result of market play. Nevertheless, it can be assumed that, in addition to factors affecting only price or volume, there are such factors as, for example, temperature, which are reflected in both price and volume of electricity demand, which, due to the main purpose of this dissertation, is not the subject of a deeper analysis. Nevertheless, an

Ishikawa diagram was developed, which presents selected factors of influence on demand and volume and price (Fig. 1, 2).

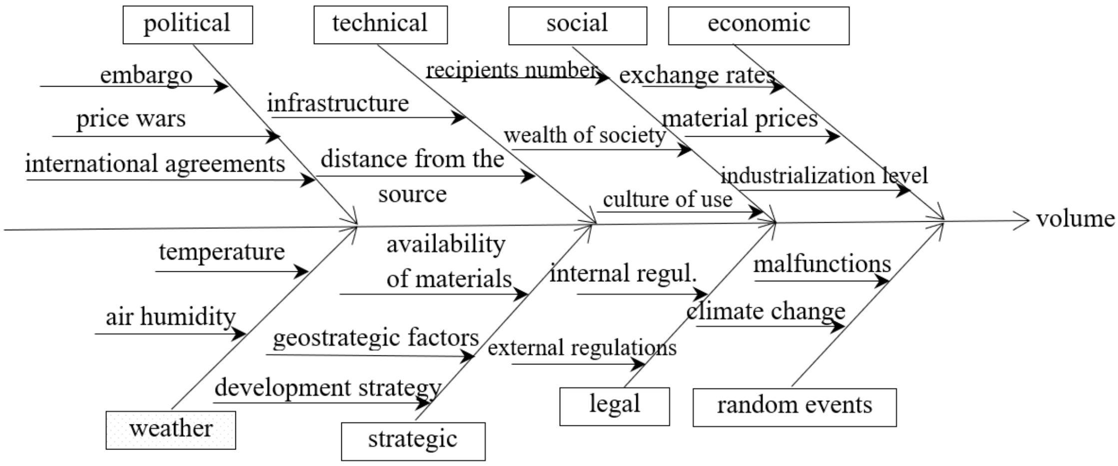

Figure 1. Diagram of the influence of factors on the volume of electricity demand. Source: own study based on [7].

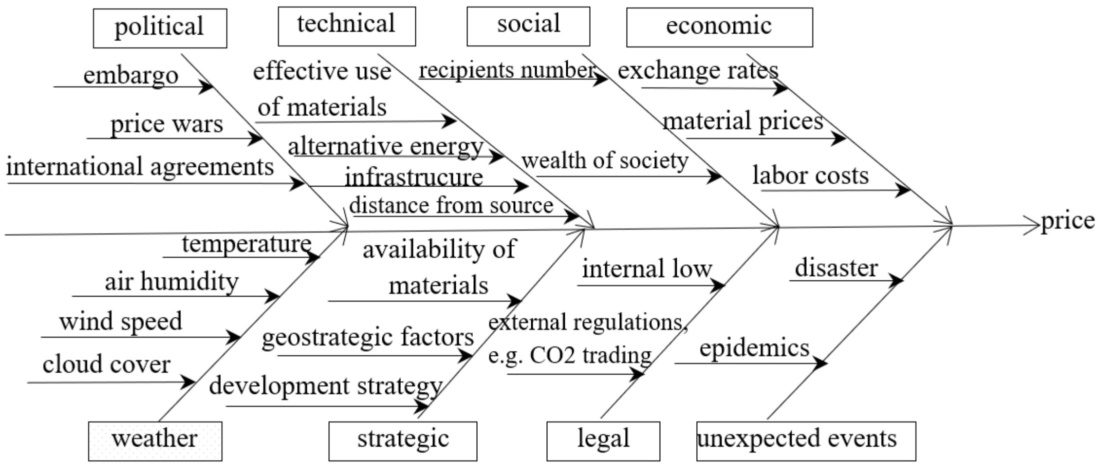

Figure 2. Diagram of various influences of factors on price. Source: own study based on [7].

Examining the impact of all the above-mentioned factors creates technical difficulties both in the terms of their parameterization and availability. For the purposes of this paper, the focus was on two main factors, i.e., economic and weather.

For the purposes of the experiment, five models were built:

1) containing one factor, energy demand, (Fig. 3),
2) containing all factors,

3) containing economic factors,
4) containing weather factors,
5) containing weather factors and selected economic factors (Fig. 4).

A diagram of two selected models is presented below.

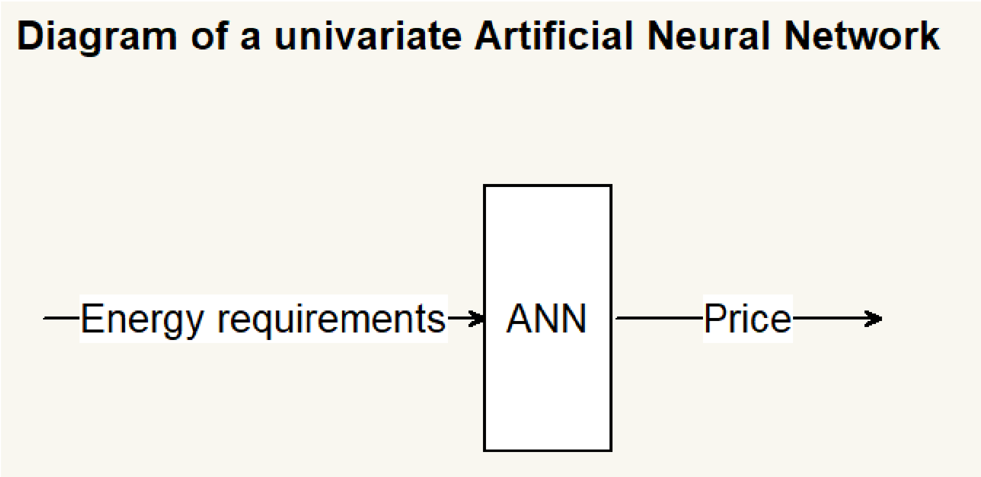

Figure3. Model that takes into account only energy demand and price as the output of the model.
Source: own study.

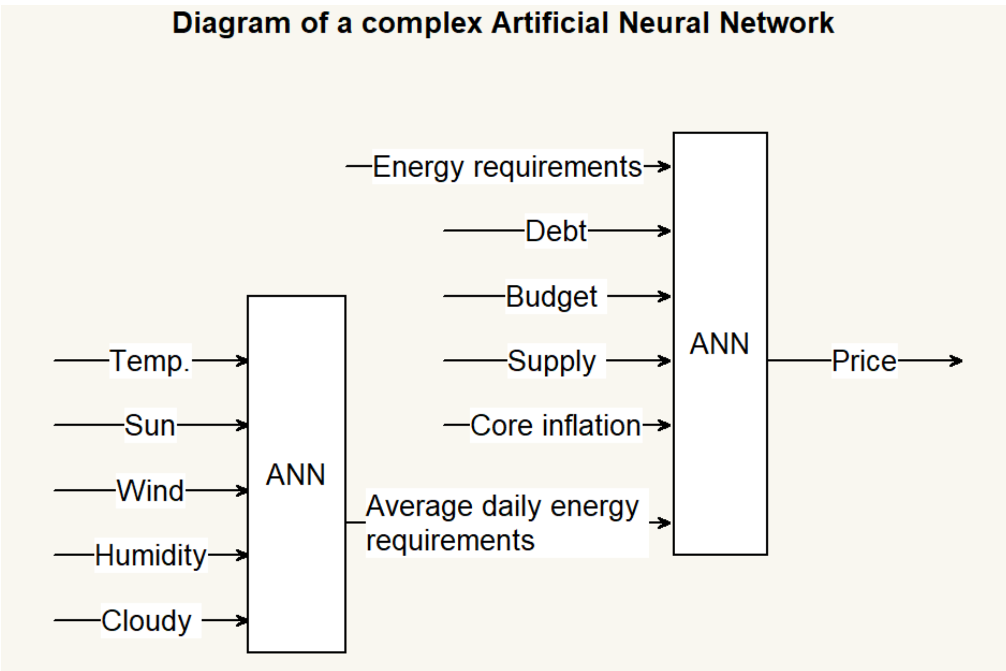

Figure 4. Model that includes weather, economic, demand and price factors as the model output.
Source: own study.

5. Description of the research

The time intervals were selected for the period corresponding to the previously tested models in order to compare the quality of the obtained model with those obtained previously. Due to the fact that most of the data was obtained in the daily system and the demand data is in the hourly system, the data from a given day was treated as data for each of the 24 hours. The average hourly demand on a given day was taken as a starting point. Average values are taken due to the fact that the input data is also on a daily basis.

Data not normalized due to their different values, e.g. degrees Celsius or wind speed, have been normalized. The normalization method was the quotient of a given quantity by the sum of these quantities in a given period, in this case 181 days. Choosing a learning method

The choice of learning method was adopted according to the accepted methodology shown in the paper [14-15, 19] and it came down to the following steps:

1) designing the ANN architecture, in this case using the feedforwardnet (hiddenSizes, trainFcn) functions,
2) selection of the neuron activation function for individual layers, in this case the functions were respectively: for the first layer of neurons - tansig and for the second layer of neurons – purelin,
3) selection of ANN learning parameters, i.e. division of the input and output data set into training, testing and validation data and the method of initialization of the initial weight and bias values,
4) the evaluation function in the learning process was the Mean Squad Error (MSE), an additional evaluation index (but not during the learning process) of the model was the coefficient of determination R^2,
5) each of the five models was taught the DAM system twenty times, i.e., the selected model was taught twenty times (with the appropriate set of data),
6) experiments.

The first experiment was the simplest possible model, implemented, among others, in the paper [14-15, 19], with a one factor, i.e., energy demand, the ideological scheme is shown in Fig. 3. The second experiment was model taking into account weather and economic factors, energy demand and price as the output of the model, the diagram is shown in Fig. 4.

Subsequent experiments, i.e.:

1) containing only economic factors,

2) containing only weather factors,
3) containing weather factors and selected economic factors.

All experiments were carried out according to the same method described in point 5.1, the selected results and the course of MSE and R^2 values for all experiments are presented in graphs.

Table 1. Selected MSE results for model No. 1 and No. 2. Source: own study.

Trial	1	2	3	4	5	6	7	8	9	10
MSE complex input	2.62E-05	1.58E-05	2.06E-05	2.17E-05	2.24E-05	2.46E-05	2.01E-05	3.25E-05	2.16E-05	3.40E-05
MSE simple input	2.52E-05	3.12E-05	2.49E-05	2.70E-05	2.26E-05	2.55E-05	2.96E-05	2.84E-05	2.42E-05	2.39E-05
Trial	11	12	13	14	15	16	17	18	19	20
MSE complex input	1.50E-05	2.33E-05	1.90E-05	1.56E-05	2.04E-05	2.09E-05	3.42E-05	1.84E-05	1.92E-05	1.45E-05
MSE simple input	2.76E-05	3.94E-05	2.32E-05	2.60E-05	2.25E-05	3.22E-05	2.75E-05	2.49E-05	2.93E-05	2.34E-05

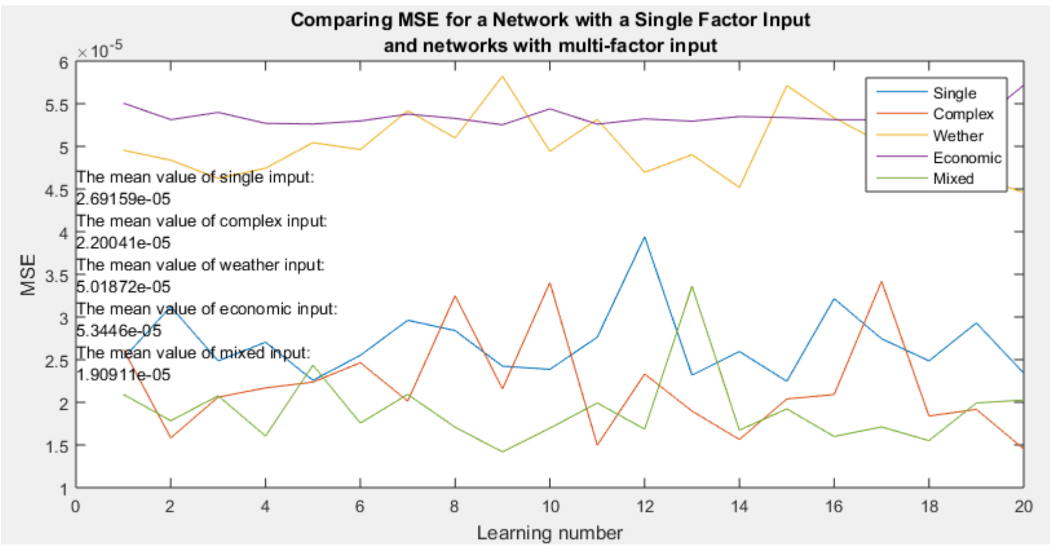

Figure 5. Comparison of MSE waveform values for all models. The designations x-axis - ordinal value of subsequent attempts to train the model, oy - MSE value for a given sample. Source: own study in the environment of MATLAB [4].

Table 2. Selected R^2 results for model No. 1 and No. 2. Source: own study.

Trial	1	2	3	4	5	6	7	8	9	10
Regression complex input	0,80445	0,756447	0,805817	0,789157	0,824628	0,799346	0,770874	0,776248	0,812272	0,819306
Regression simple input	0,8006	0,882303	0,842319	0,835333	0,83776	0,809407	0,846567	0,737195	0,835827	0,719567
Trial	11	12	13	14	15	16	17	18	19	20
Regression complex input	0,783912	0,665545	0,820037	0,809382	0,825949	0,752303	0,784567	0,808514	0,768042	0,818657
Regression simple input	0,888114	0,820638	0,858485	0,890899	0,843733	0,843197	0,72092	0,869515	0,861323	0,892547

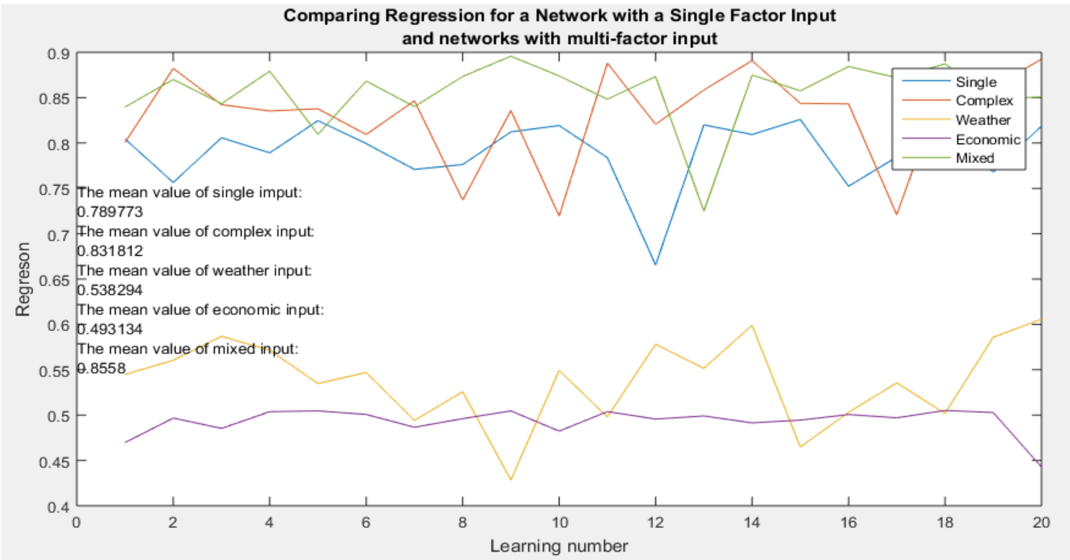

Figure 6. Comparison of mileage values R^2 for all models. The designations of the axis x- the ordinal value of subsequent attempts to train the model, axis y - the value of R^2 for a given sample. Source: own study in the environment of MATLAB [4].

6. Conclusions

As can be seen the mean values of the MSE index for:

1) containing one factor, i.e., demand, is - $2.691*10^{-5}$,

2) containing all factors, i.e., energy demand, economic factors such as the level of inflation, the amount of debt, the balance of state expenditure, money supply and environmental factors such as temperature, air humidity, cloud cover and wind force, it is - 2,200 * 10^{-5},
3) including economic factors, it is - 5.344 * 10^{-5},
4) containing weather factors, it is - 5.018 * 10^{-5},
5) containing weather factors and selected economic factors, i.e., the balance of state expenditure, amounts to - 1.909 * 10^{-5}.

The coefficient of determination R^2 is:

1) containing one factor, i.e., demand, amounts to - 0.789,
2) containing all factors, i.e., energy demand, economic factors such as the level of inflation, the amount of debt, the balance of state expenditure, money supply and environmental factors such as temperature, air humidity, cloud cover and wind force, it amounts to - 0.832,
3) including economic factors, is - 0.493,
4) containing weather factors, amounts to - 0.538,
5) containing weather and selected economic factors, i.e., the balance of state expenditure, amounts to - 0.856.

As can be seen, both MSE and R^2 values are better for models 5, 3, 1 and weaker for models 2 and 4. An important observation may be the result of model 5, containing selected factors, the mean values of which are better than model 2, containing all analysed factors, which shows that an increase in the number of factors modelling a given system does not necessarily improve its model.

The results of model 1 are also interesting from the implementation point of view, which indicate that in this case the simplest model based only on energy demand relatively correctly reflects the operation of the system, which in this case is the Day-Ahead Market.

References

1. Amjady N., (2006), Day-ahead price forecasting of electricity markets by a new fuzzy neural network, IEEE Transactions on Power Systems, Vol. 21, No. 2, pp. 887-896.

2. Anderson, T. W. (2003), An Introduction to Multivariate Statistical Analysis, Wiley, New York.

3. Bai J., Ng S., (2002), Determining the Number of Factors in Approximate Factor Models Econometrica, Vol. 70, No. 1, pp. 191–221.

4. Beale M. H., [et al.] (1992-2020), Neural Network Toolbox™ User's Guide, by The MathWorks, Inc.

5. Conejo A. J., Plazas M. A. , et all, (2005), Day-ahead electricity price forecasting using the wavelet transform and ARIMA models, IEEE Transactions on Power Systems, Vol. 20, No. 2, pp. 1035-1042, doi: 10.1109/TPWRS.2005.846054.

6. Contreras J., Espínola R. [et al.] (2003), ARIMA models to predict next-day electricity prices, IEEE Transactions on Power Systems, Vol. 18(3), pp. 1014-1020.

7. Kowalik K., Klimecka-Tatar D.,(2018), The process approach to service quality management. Production Engineering Archives, 18, pp. 31-34.

8. Li G., Liu C., et all, (2007), Day-Ahead Electricity Price Forecasting in a Grid Environment, IEEE Transactions on Power Systems, Vol. 22, No. 1, pp. 266-274, doi: 10.1109/TPWRS.2006.887893.

9. Maciejowska, K., Weron, R. (2015), Forecasting of daily electricity prices with factor models: utilizing intra-day and inter-zone relationships. Comput Stat 30, pp. 805–819. https://doi.org/10.1007/s00180-014-0531-0.

10. Mahler V., Girard R, et all. (2019), Simulation of day-ahead electricity market prices using a statistically calibrated structural model. 16th European Energy Market Conference, IEEE PES, Sep 2019, Ljubljana, Slovenia.

11. Marszałek A., Burczyński T. (2021), Journal Smart Environ Green Computing 2021.

12. McHugh C., Coleman S., et all, (2019), Forecasting Day-ahead Electricity Prices with a SARIMAX Model, 2019 IEEE Symposium Series on Computational Intelligence, pp. 1523-1529.

13. Mingming G., Jianjing Let all, (2019), Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM, Energy, Vol. 187.

14. Ruciński D., Tchórzewski J. (2016). Neural modeling of the electric power stock market in usage of MATLAB and Simulink tools for the day ahead market data, Information System in Management, Vol. 5 (2), pp. 215-226.

15. Ruciński D., (2019), The Influence of the Artificial Neural Network Type on the Quality of Learning on the Day-Ahead Market Model at Polish Electricity Exchange Join-Stock Company, Studia Informatica. System and Information Technology, Vol. 1-2(23), pp. 77-93.

16. Rusiecki, A., L., (2006), Robust Learning Algorithm with the Variable Learning Rate, ICAISC 2006, Artificial Intelligence and Soft Computing, pp. 83-90.

17. Senjyu T., Takara H. [et al.] (2002), One-Hour-Ahead Load Forecasting Using Neural Network, IEEE Transaction on Power Systems, Vol. 17, No. 1, pp. 113-118.

18. Stankovic A. M., Saric A. T., and Milosevic M. (2003), Identification of Nonparametric Dynamic Power System Equivalents with Artificial Neural Networks, IEEE Transactions on Power Systems, Vol. 18, No. 4, pp. 1478-1486.

19. Tchórzewski J. (2021) Metody sztucznej inteligencji i informatyki kwantowej w ujęciu teorii sterowania i systemów, Wydawnictwo UPH, Siedlce.

20. Wang F., Zhang Z, et all. (2019), Generative adversarial networks and convolutional neural networks based weather classification model for day ahead short-term photovoltaic power forecasting, Energy Conversion and Management, Volume 181, pp. 443-462.