Iuliia KUIEVDA
Serhii BALIUTA
Petro ZINKEVICH
Oleksandr STOLIAROV

National University of Food Technologies, 68 Volodymyrska str., 01601 Kyiv, Ukraine

Corresponding author: e-mail: julika@gmail.com

DOI: 10.53412/jntes-2021-1.5

FORECASTING THE ELECTRICITY GENERATION OF PHOTOVOLTAIC PLANTS

Abstract: *Due to the need in accordance with Ukrainian legislation to submit a day-ahead hourly forecast of electricity generation of solar power plants, the problem of forecasting model quality becomes very important. In the study it is proposed a method of choosing the optimal structure and sensitivity assessment of ANFIS-based forecasting model. In the model the input is solar irradiance, the output is solar panel generation power. The method is based on computational procedures using MATLAB software. For the data set, used in the study, the results, optimal for normalized mean absolute error (NMAE), were achieved on 5 triangular input member functions (trimf), while the error varied within 0.23% depending on number and shape of input member functions. According to the calculations of input error sensitivity of the forecasting model with 5 input trimf membership functions, the increasing of input error up to 8.19% NMAE leads to the raising of the output error in the testing sample up to 5.78%, NMAE. The rather low sensitivity of the model to the input data error allows us to conclude that forecasted meteorological data with a pre-known fixed forecast error can be used as input data.*

Keywords: *solar power plants; forecasting; ANFIS; fuzzy inference system.*

Introduction

Since 2019, the electricity market in Ukraine has moved to a new model of operation. Producers from Renewable energy sources (RES) make up an increasing share of the market. According to Ukrenergo [2], in March 2021 the installed capacity of RES in Ukraine was 6.97 GW, of which the largest share falls on solar power plants (SPP) – 5.51 GW, namely 10% of the total installed generation capacity in the country. RES producers sell electricity at a "green tariff", which is the highest among all others in Ukraine. From the beginning of 2021, RES producers began to compensate for the imbalance of electricity in the market a day in advance relative to the day-ahead hourly generation forecast that they provide before the start of the market day.

All these circumstances have increased the importance of the tasks of forecasting the generation of electricity from RES producers, in particular SPP, both at the level of the power system of Ukraine as a whole and at the level of producers.

The day-ahead hourly forecast of SPP generation is referred to as short-term forecasts and it is used physical, statistical, and intelligent methods [1]. Among the intelligent forecasting methods, the most popular are models based on the Artificial Neural Networks (ANN) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS), as well as combinations thereof. For example, in [5] and [6] new combined methods are presented, which include ANN and ANFIS for short-term forecasting of SPP generation.

The above publications describe the classical and new methods of forecasting and include assessment of their accuracy, but it is paid not enough attention to the study of the properties of the models themselves, in particular their structure and sensitivity to input error. The latter is especially true for models that use intelligent technologies, such as ANN and ANFIS, where it is impossible to explicitly express the relationship between input and output variables.

The aim of this work is to form an approach to the study of the forecast error dependence on the number of input membership functions and their form, as well as the sensitivity to input data error of the ANFIS-based day-ahead hourly forecast SPP generation model. The sensitivity of the model to input data error is particularly important because the input data for generation forecasting can come from meteorological forecast providers, which indicate the predefined forecast error.

Materials and Methods

In the study it was built a numerical model of the dependence of solar panel generation power on current solar irradiance based on ANFIS 3. This model can be used for hourly day-ahead generation forecast based on solar irradiance forecast values from weather forecast providers, considering the cloud opacity.

The ANFIS methodology is based on a network of special structure that allows to create and configure a set of fuzzy rules such as Takagi-Sugeno to approximate the relationship between multiple inputs and a single output. The author of ANFIS in [3] showed that it is a universal approximator of continuous functions of several variables defined on compact sets.

Data from the open dataset Photovoltaic (PV) Solar Panel Energy Generation from UK Power Networks from the London Datastore repository were used to train and test the model [4]. The input data of the model – current solar radiation (W/m^2), measured at the local weather station, and the output data – the power of the solar panel (kW) from the location of Forest Road. For the research, the sample was divided into training and testing data. In Figure 1 it is given the graph of input and output data and shown training sample in blue, while testing sample in red.

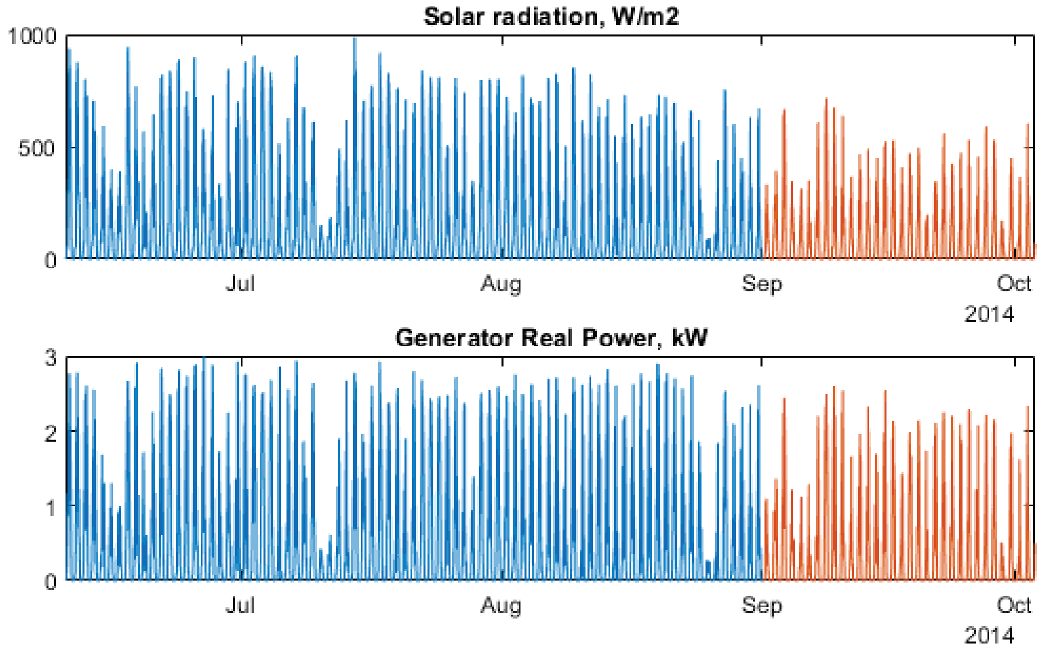

Figure 1. *Data from the location of Forest Road (open dataset Photovoltaic (PV) Solar Panel Energy Generation), training sample is blue, testing sample is red*

The research was performed using numerical simulation in MATLAB with the use of Fuzzy Logic Toolbox. The first part of the research was devoted to the choice of the number of input membership functions and their form. In the second part of the study, the sensitivity of the model to the input data error was determined for the selected number and form of membership functions from the first part. The results were evaluated by errors root mean square error (RMSE), mean absolute error (MAE) and normalized mean absolute error (NMAE):

$$\mathrm{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{y}_i-y_i\right)^2} \tag{1}$$

$$\mathrm{MAE}=\frac{1}{N}\sum_{i=1}^{N}\left|\hat{y}_i-y_i\right| \tag{2}$$

$$\mathrm{NMAE}=\frac{100}{N}\sum_{i=1}^{N}\frac{\left|\hat{y}_i-y_i\right|}{y_{\max}} \tag{3}$$

where:

y_i – are observation values;

$\hat{y}_i$ – predicted values;

$y_{\max}$ – maximum value of the observed variable.

Results and Discussions

It was calculated the table of dependence of RMSE, MAE and NMAE errors on the number of input membership functions and their forms for training, testing and the whole sample. The number of input membership functions varied from 2 to 30, and the type of membership functions (MF) was chosen from the set (MATLAB notation [7]): Generalized bell-shaped MF (gbellmf), Gaussian MF (gaussmf), Gaussian combination MF (gauss2mf), Triangular MF (trimf), Trapezoidal MF (trapmf), Difference between two sigmoidal MF (dsigmf), Product of two sigmoidal MF (psigmf), Pi-shaped MF (pimf).

To determine the optimal number and form of input membership functions, the NMAE error of the testing sample was used. Figure 2 shows graph of NMAE errors which were calculated from identified on training sample ANFIS models with 8 types of input membership functions from the set, mentioned above. Each line in the graph corresponds to specific type of membership function, *x*-axis shows the number of membership functions.

As can be seen in Figure 2, NMAE varies in the range from 3.92% to 4.15%, and minimum is achieved on 5 triangular trimf membership functions (4) with tunable parameters *a*, *b*, *c*, which shapes are shown in Figure 3. Let us call this model ANFIS5trimf.

$$f(x)=\max\left(\min\left(\frac{x-a}{b-a},\frac{c-x}{c-b}\right),0\right) \tag{4}$$

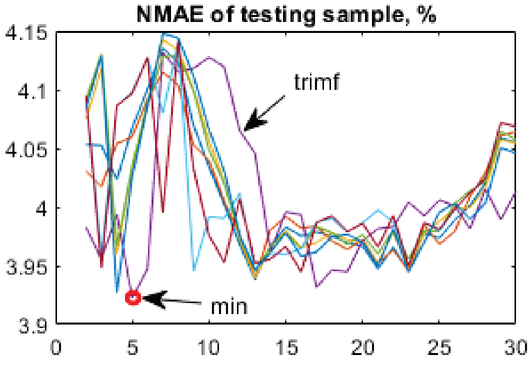

Figure 2. *Minimum of NMAE on testing sample*

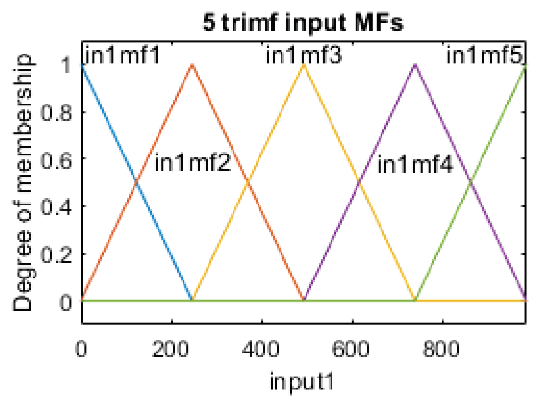

Figure 3. *5 input triangular membership functions*

The structure of ANFIS5trimf is presented in Figure 4. As shown in the figure, the model consists of 1 input – solar radiation in (W/m^2), 1 output – power of the solar panel (kW), 3 layers of 5 nodes, which represent trimf input, linear output membership functions and fuzzy logical operations, and 1 summarizing layer to produce the output value.

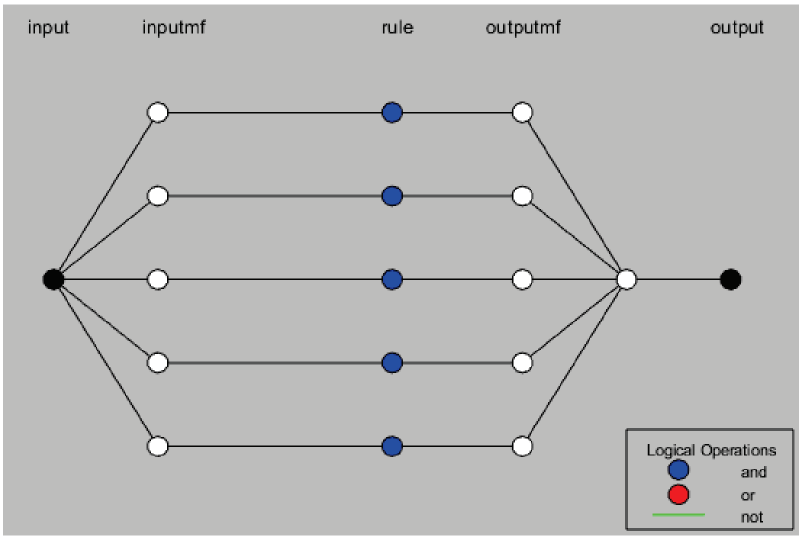

FIGURE 4. *The structure of ANFIS model with 5 input membership functions*

Figure 5 demonstrates scatterplot of dependence of solar panel generation power on measured solar irradiance from the Forest Road data set, mentioned above. The training sample consists of blue dots, and testing sample – of orange dots. The red curve stands for the output surface of trained ANFIS5trimf, while the green curve shows the result of linear regression. For comparison of ANFIS5trimf and linear regression models on the testing sample it was calculated RMSE (1): 0.2444 and 0.2641 correspondingly, which means that ANFIS5trimf better fits the data than linear regression.

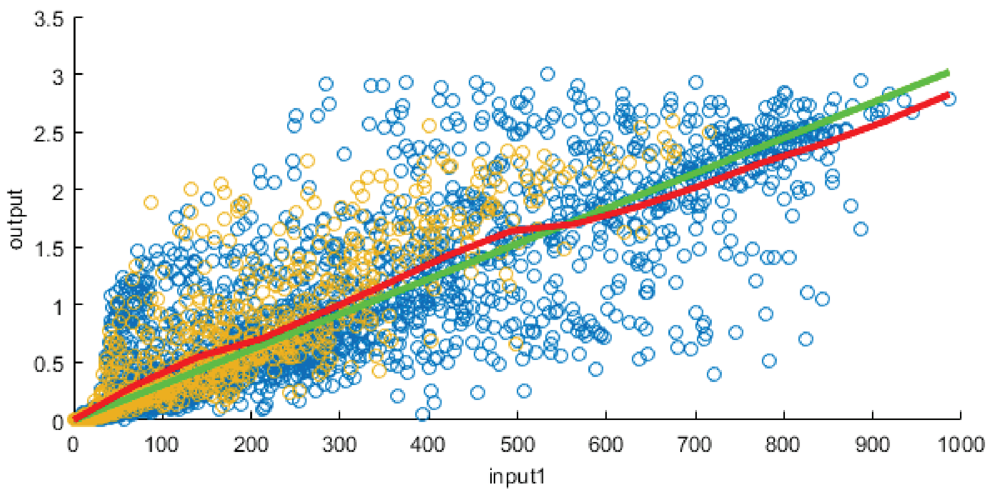

FIGURE 5. *Solar radiation in training (blue) and testing (orange) samples with ANFIS5trimf output surface (red) and linear regression curve (green)*

As can be seen in Figure 5 the data are rather widely spread out from the approximating curves. There are some possible reasons of that, which can be considered for both variables in data set. For measurements of solar radiation by solar sensor at the local weather station it can be mentioned that they are taken from close but different place than solar panels location, solar panels can take large area and be under different cloud conditions, furthermore the dirt and snow can influence its measurements. For generator power data there are also some additional influencing factors, such as uneven shading, precipitation, level of degrading of PV modules, and their availability, etc. All these factors can be considered to make the model more accurate for further researches.

Figure 6 shows the example of ANFIS5trimf forecast for one week of testing sample: purple line consists of forecasted values, and orange line consists of measured values.

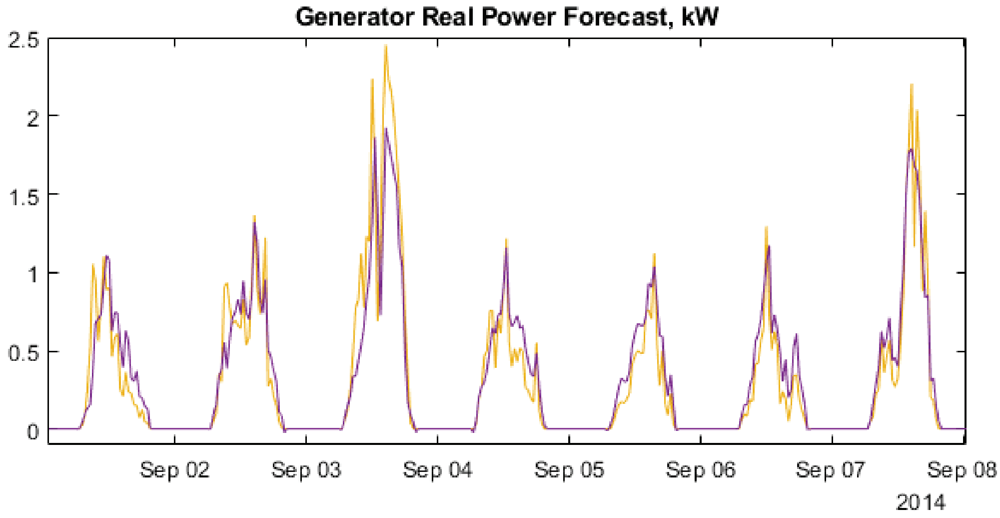

FIGURE 6. *ANFIS forecast of generator real power, where forecasted values are purple and measured values are orange*

The ANFIS5trimf model was tested for sensitivity to the error of the input data. The generated random sequences with different variances were sequentially added to the testing input data, which varied NMAE of the input data error from 1.81% to 8.19%. In the Figure 7 the blue line is input NMAE error and the red line is corresponding output NMAE error. The output error starts with a non-zero value because of initial error of ANFIS5trimf model for the testing sample. The NMAE error of the output data of the testing sample varied from 4.19% to 5.78%, on the basis of which we can conclude that the calculations showed a sufficiently low level of variation in the output values relative to the error of the input data.

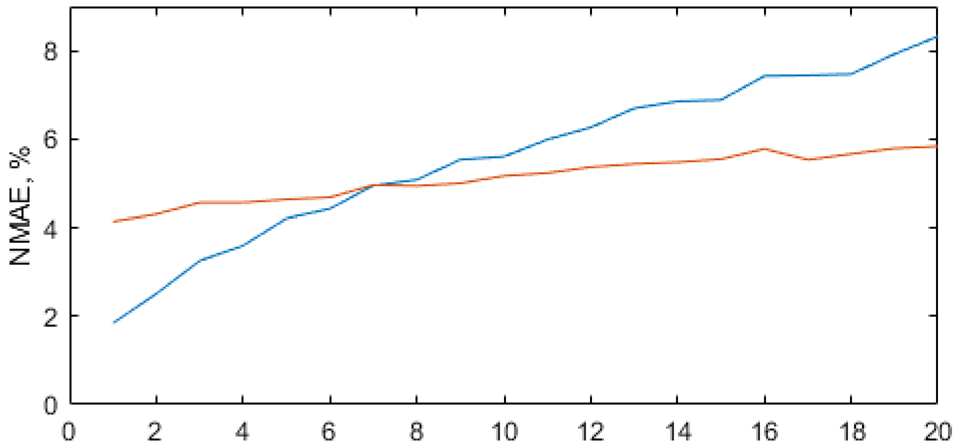

FIGURE 7. *The input data NMAE error (blue line) and corresponding output NMAE error (red line)*

Conclusions

The study demonstrates an algorithm according to which it is proposed to study the structure and sensitivity of ANFIS-based models for the problem of approximating the dependence of SPP generation power on solar irradiance. In this study, the best type of membership function was trimf, the number and shape of functions had little effect on the result – within 0.23% of NMAE.

Regarding the sensitivity of the model to the input error, it can be concluded that for 5 input trimf membership functions, increasing the input error to 8.19% NMAE leads to an increase in the output error in the testing sample to 5.78%, NMAE. The rather low sensitivity of the model to the input data error allows us to conclude that it can be used for forecast meteorological data with a pre-known fixed forecast error.

Conflicts of Interest: The author declares no conflict of interest.

References

[1] Akhter M.N., *ANFIS: Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques*. Akhter M.N., Mekhilef S., Mokhlis H., Shah N.M., IET Renewable Power Generation 2019, Vol. 13, Iss. 7, pp. 1009-1023.

[2] Installed capacity of the IPS of Ukraine values. Available at: https://ua.energy/vstanovlena-potuzhnist-energosystemy-ukrayiny (Accessed: 30 April 2021).

[3] Jang J.-S.R., *ANFIS: Adaptive-Network-Based Fuzzy Inference System*. Jang J.-S.R., IEEE Transactions on Systems, Man, and Cybernetics 2018, Vol. 23, No. 3, pp. 665-685.

[4] Photovoltaic (PV) Solar Panel Energy Generation data. Available at: https://data.london.gov.uk/dataset/photovoltaic--pv--solar-panel-energy-generation-data (Accessed: 30 April 2021).

[5] Semero Y.K., *PV Power Forecasting Using an Integrated GA-PSO-ANFIS Approach and Gaussian Process Regression Based Feature Selection Strategy*. Semero Y.K., Zhang J., Zheng D., CSEE Journal of Power and Energy Systems 2018, Vol. 4, No. 2, pp. 210-218.

[6] Wu Y.-K., *A Novel Hybrid Model for Short-Term Forecasting in PV Power Generation*. Wu Y.-K., Chen C.-R., Rahman H.A. International Journal of Photoenergy 2014, 9 p.

[7] MATLAB documentation. Available at: https://www.mathworks.com/help (Accessed: 30 April 2021).