Stochastic model for prediction of microgrid photovoltaic power generation

Cite as: AIP Conference Proceedings **2333**, 090020 (2021); https://doi.org/10.1063/5.0041825
Published Online: 08 March 2021

Silvia Baeva, Rad Stanev, Stoyan Popov, and Nikolay Hinov

Stochastic Model for Prediction of Microgrid Photovoltaic Power Generation

Silvia Baeva [a)], Rad Stanev [b)], Stoyan Popov [c)] and Nikolay Hinov [d)]

Technical University of Sofia, Bulgaria

[a)] Corresponding author: sbaeva@tu-sofia.bg
[b)] rstanev@tu-sofia.bg
[c)] stoian_popov95@abv.bg
[d)] hinov@tu-sofia.bg

Abstract. In this article, a stochastic model for prediction of microgrid photovoltaic power generation, using statistical and stochastic methods is presented. The study is performed in the following steps: Processing of a large database of historical data (Data Mining); Construction of a stochastic forecasting model; Reporting a symmetric mean absolute percentage error in forecasting.

INTRODUCTION

Forecasting tasks are solved in a variety of areas of human activity - science, economics, manufacturing and many others. Forecasting is an essential element of the organization of the management of both the individual business entity and the economy as a whole.

The forecasting task is considered to be one of the most complex tasks in the field of Data Mining. It requires careful analysis of large volumes of data, as well as taking into account the influence of many related and dynamic factors on the parameters of the forecast model. Forecasting is aimed at determining the trends of the dynamics of a particular object or event based on retrospective data. In some cases, problems in which there is no dynamics of observations over time are presented as forecasting tasks, and an adequate description of the relationship between the values of dependent parameters and the influencing values of independent quantities (factors, predictors) is sought.

The methods for solving the forecasting task are based exclusively on the statistical approach. In all cases, as a solution to the problem, a substantiated statistical assessment of the estimated value is sought, combined with the construction of its confidence area (that area in which the probability of the assessment is high enough). Popular methods for solving the problem are the following:

- Regression analysis, usually linear - its use is applied to find a statistical quantitative relationship between the values of the predicted quantity and the predictors (respectively time). Despite its widespread application, the use of this approach is inexpedient for forecasting, as extrapolating the forecast values outside the observed time interval leads to a rapid accumulation of error.
- Time series - statistical methods specially designed to study trends in series of observed values. Widely used methods with the ability to report trends, periodic and seasonal fluctuations, etc.
- Neural networks - have the same qualities as the methods of the time series with the ability to adapt to changes in the characteristics of the series of observations.

Applications of Mathematics in Engineering and Economics (AMEE'20)
AIP Conf. Proc. 2333, 090020-1–090020-10; https://doi.org/10.1063/5.0041825
Published by AIP Publishing. 978-0-7354-4077-7/$30.00

The authors in [1] compare the diurnal forecast characteristics for energy production, obtained by using two methods, which are most often used to predict the generated power of photovoltaic systems. Both methods are based on artificial neural networks (ANNs), which are trained using the same data set. Such a statement makes it possible to make the necessary homogeneous comparison, for which there is currently no published data in the available literature. For the purposes of the study, the data set consists of an hourly series of climatic and PV parameters of the system, which cover a time period of one year. This period is grouped to distinguish sunny from cloudy days, thus being chosen for ANN training. One forecasting method is powered only by the available data set, while the other uses a hybrid method that uses the daily weather forecast. As a result, the following conclusions are made: on sunny days, the first method shows very good and stable prognostic results, with practically constant normalized mean absolute error, NMAE%, in all cases (1% <NMAE% <2%); On the other hand, the hybrid method gives even better indicators (NMAE% <1%) for two of the days considered in this analysis, but is generally characterized by less stable indicators (NMAE%> 2% and up to 5.3 % for all other cases). On cloudy days, the accuracy of forecasting by both methods usually decreases, and the accuracy is quite stable for the method that does not use weather forecasts. On the other hand, the accuracy of the hybrid method varies considerably for the days considered for the analysis.

In [3], a model for predicting the yield of a photovoltaic generator (PV) using multiple repetitive networks (GRUs) is proposed, which are known to effectively improve the prediction accuracy on the one hand, and also significantly they also reduce training time compared to a typical GRU network. In addition, other popular machine learning algorithms have often been used for the purpose of predicting stochastic quantities, namely artificial retrieval neural network (ANN), vector regression (SVR) support, and K. Nearest neighbors (KNN) are implemented for comparison with the proposed model. The accuracy of each model was assessed by Normalized Root Mean Squared Error (NRMSE). For the proposed GRU model, submission of ANN, SVR and KNN there are NRMSE results are obtained respectively: 9.64%, 10.53%, 11.62%, 11.45% and 11.89%.

The authors in [7] used a model to predict the yield of a photovoltaic system using a periodic neural network (RNN) in a cascade model combined with hierarchical clustering. Thus, the overall accuracy of PV energy yield forecasting is considered to be improved. The proposed model is compared with other similar neural network training algorithms, namely: Sending an artificial neural network (FFNN), GRU supporting vector regression (SVR) and K Nearest Neighbors (KNN) using cluster data from K -Means Clustering and Hierarchical Clustering, which have the lowest average NRMSE of 8.88%, due to the use of hierarchical clustered data. According to the presented results, it is concluded that the Hierarchical clustering is more suitable for the needs of preparation of PV yield forecast than K-means clustering.

Manuscript [4] presents a prediction of the output power of a photovoltaic system for an hour ahead, using a combination of wave conversion techniques (WT) and artificial intelligence (AI). This is achieved by taking into account the interactions of the photovoltaic system with data from solar radiation and temperature. In the proposed method, WT is used to exclude unscrupulous photovoltaic data from time series, as the AI techniques used capture the nonlinear fluctuation of PV yield in a better way.

In [8], for the purposes of forecasting, an experimental database of the instantaneous values of the output energy of the photovoltaic generator, solar radiation, air temperature and the temperature of a single photovoltaic module is used. The manuscript uses data from the office of a building electrified with green energy in Malaysia, the Taichung TPP in Taiwong and the National University of Pengu. Based on the historical data on the power obtained from PV and the time set in the described numerical experiment, all the factors influencing the photovoltaic energy yield are considered. In addition, five types of forecasting models have been developed and applied, the main task being to predict the output energy produced by PV for one hour ahead. The models include ARIMA, SVM, ANN, ANFIS, as well as combined models using the GA algorithm. The results of the forecast prove the applicability of this approach, which is characterized by its high accuracy and efficiency.

The purpose of the research presented in [2] is initially to make a detailed review of conventional and modern techniques for predicting the yield of photovoltaic generator (PV) and then on this basis to highlight the predominant factors that affect the generation of photovoltaic energy. This document also examines the principle of operation and applicability of different methods for forecasting the energy yield from PV. To confirm the results, the main factors influencing energy generation are assessed using data from real objects.

In [5] a mathematical multilinear regression model of a power electronic device in a photovoltaic power plant is shown. The model is based on the efficiency of the electronic energy converter and for modeling purposes uses the input DC voltage and the input DC power of the PV inverter as two independent variables that can be determined by the parameters of the solar installation and the ambient temperature. The developed model is implemented in MatLab environment. The analysis of the accuracy of the model and its adequacy are performed on the basis of comparison with real data from a photovoltaic system. The results are shown a combination of high computational efficiency and good accuracy. In addition, the approach used provides good prospects for further improvement and development.

The authors in [6] have performed a comprehensive and in-depth review of the methods used to directly predict the production of photovoltaic energy. The significance of the correlation between the input-output data and the pre-processing of the input data required for the use of the model is considered. The review covers the analysis and comparison of the efficiency of several models for predicting the power generated by PV. Particularly useful for researchers is the critical analysis of recent manuscripts on statistical methods for forecasting and the use of machine learning for forecasting based on historical data. Based on the analysis, the strengths and weaknesses of the various forecasting models, including innovative and classical ones, are examined. The potential benefits of the optimization of the developed model are also considered.

In all cases, the application of forecasting models should be combined with an analysis of their accuracy and adequacy, which is done by statistical means. It should be noted some similarity between the classification tasks and those for forecasting, as in both cases the aim is to predict the probable state (respectively class affiliation) of the studied object based on the values of the variables describing it. Properly conducted research and adequately composed forecasting model can be extremely useful in many practical activities of life.

In this article, a stochastic model for prediction of microgrid photovoltaic power generation is presented, using statistical and stochastic methods. The study is performed in the following steps:

1. Processing of a large database of historical data (Data Mining);
2. Construction of a stochastic forecasting model;
3. Reporting a symmetric mean absolute percentage error in forecasting.

DESCRIPTION OF STUDY METHODOLOGY

Processing of a Large Database of Historical Data (Data Mining)

The database includes for the last few years for each day of each minute reported - day, hour, point, generated power, power consumption, etc.

The data are processed as follows:

1. The data records for the generated power for the given years for each day and every minute are extracted from the database;
2. The total generated power for each day of the given years have been found and designated as $k_{y,m,d}$, where y is the corresponding year, m is the corresponding month of the year y, and d is the corresponding day of that month;
3. The total generated power for each month of the given years are found and noted as $K_{y,m}$, where y is the corresponding year, m is the corresponding month of the year y.

Construction of a Stochastic Forecasting Model

Let the generated power for each month of the given years be set with the following matrix:

$$K = \begin{pmatrix} K_{y_1,1} & K_{y_2,1} & \dots & K_{y_n,1} \\ K_{y_1,2} & K_{y_2,2} & \dots & K_{y_n,2} \\ \dots & \dots & \dots & \dots \\ K_{y_1,12} & K_{y_1,12} & \dots & K_{y_n,12} \end{pmatrix} = \left\{K_{y_i,j}\right\}_{12\times n} \tag{1}$$

and for each year $y_i, i = 1, \dots, n$:

$$K_{y_i} = \begin{pmatrix} K_{y_i,1} \\ K_{y_i,2} \\ \dots \\ K_{y_i,12} \end{pmatrix}. \tag{2}$$

Weights $w_{y_i}, i = 1, \dots, n$, such that

$$\sum_{i=1}^{n} w_{y_i} = 1, \tag{3}$$

which characterize the stochastic nature of the data, in this case the generated power, over the years.

Then the projected generated power for each month of the forecast year y_{n+1} is:

$$K = \begin{pmatrix} \sum_{i=1}^{n} w_{y_i}.K_{y_i,1} \\ \sum_{i=1}^{n} w_{y_i}.K_{y_i,2} \\ \dots \\ \sum_{i=1}^{n} w_{y_i}.K_{y_i,12} \end{pmatrix}. \tag{4}$$

Reporting a Symmetric Mean Absolute Percentage Error in Forecasting

To assess the quality of the forecasting model, an average absolute percentage error is introduced, which is:

$$SMAPE = \sum_{j=1}^{12} \frac{2.\left|K_{y_n,j} - \sum_{i=1}^{n} w_{y_i}.K_{y_i,j}\right|}{\left(K_{y_n,j} + \sum_{i=1}^{n} w_{y_i}.K_{y_i,j}\right)}.100. \tag{5}$$

NUMERICAL REALIZATION

The numerical realization is performed using an autonomous microgrid system (figure 1), consisting of:

- PV generator: 2,3kW; Hydro Generator: 3kW; Diesel Generator: 12kW;
- Hybrid Inverter- Charger -Transfer device: 48V DC, 230VAC, 8kVA;
- Battery storage: 43,2 kWh;
- Supplied AC load: 230V.

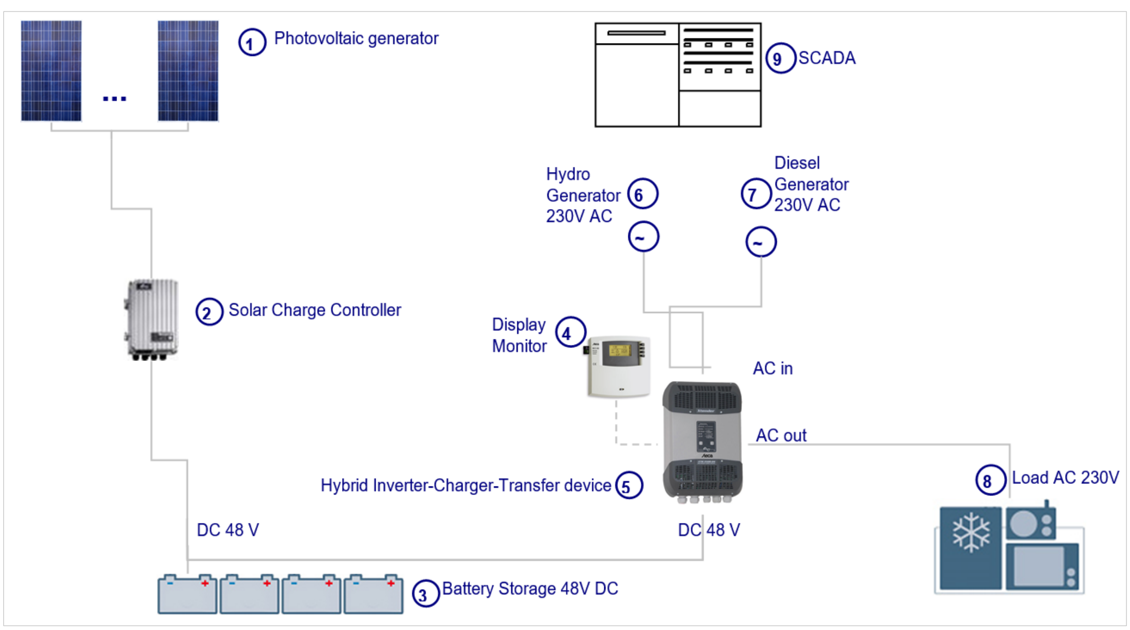

FIGURE 1. Scheme of the studied autonomous microgrid system

Input numerical data and processing

The input data are processed according to the methodology described above and the total generated power [kW] for each month of the given years are found (Table 1).

TABLE 1. Input data [kW]

Year / Month	2017	2018	2019
Jan	31438.29	33194.28	31363.11
Feb	23161.94	32326.97	32891.81
Mar	33655.99	27508.68	25154.47
Apr	25112.98	31201.57	29167.51
May	33346.76	29107.47	25135.57
June	24731.51	30077.35	18510.85
July	27936.58	23800.73	16028.65
Aug	26190.65	22470.06	24345.07
Sep	26204.67	20674.59	22834.23
Oct	29785.00	27196.13	24978.64
Nov	25922.28	25169.70	26658.02
Dec	35941.20	28622.31	31068.19
Σ	343427.85	331349.80	308136.10

Numerical Results

Six cases of forecasting depending on weights are realized and the numerical results are presented in Table 2.

TABLE 2. Estimated values of generated power [kW]

Stochastic forecast for 2019						
Weights / Months	$w_{17}=0.2$ $w_{18}=0.8$	$w_{17}=0.25$ $w_{18}=0.75$	$w_{17}=0.3$ $w_{18}=0.7$	$w_{17}=0.35$ $w_{18}=0.75$	$w_{17}=0.4$ $w_{18}=0.6$	$w_{17}=0.5$ $w_{18}=0.5$
Jan	32843.08	32755.28	32667.48	32579.68	32491.88	32316.29
Feb	30493.96	30035.71	29577.46	29119.21	28660.96	27744.46
Mar	28738.14	29045.51	29352.87	29660.24	29967.60	30582.34
Apr	29983.85	29679.42	29374.99	29070.56	28766.13	28157.28
May	29955.33	30167.29	30379.26	30591.22	30803.19	31227.12
June	29008.18	28740.89	28473.60	28206.31	27939.01	27404.43
July	24627.90	24834.69	25041.49	25248.28	25455.07	25868.66
Aug	23214.18	23400.21	23586.24	23772.27	23958.30	24330.36
Sep	21780.61	22057.11	22333.61	22610.12	22886.62	23439.63
Oct	27713.90	27843.35	27972.79	28102.23	28231.68	28490.57
Nov	25320.22	25357.85	25395.47	25433.10	25470.73	25545.99
Dec	30086.09	30452.03	30817.98	31183.92	31549.87	32281.76
Σ	333765.44	334369.34	334973.24	335577.14	336181.04	337388.85
Average	27813.79	27864.11	27914.44	27964.76	28015.09	28115.74

The average absolute percentage error in the six cases is calculated and presented in Table 3.

TABLE 3. Average absolute percentage error [%]

Weights	$w_{17}=0.2$ $w_{18}=0.8$	$w_{17}=0.25$ $w_{18}=0.75$	$w_{17}=0.3$ $w_{18}=0.7$	$w_{17}=0.35$ $w_{18}=0.75$	$w_{17}=0.4$ $w_{18}=0.6$	$w_{17}=0.5$ $w_{18}=0.5$
$\boldsymbol{SMAPE}$ [%]	7.98	8.17	8.35	8.53	8.71	9.06

The numerical results for the six cases are presented graphically in Figures 2÷7.

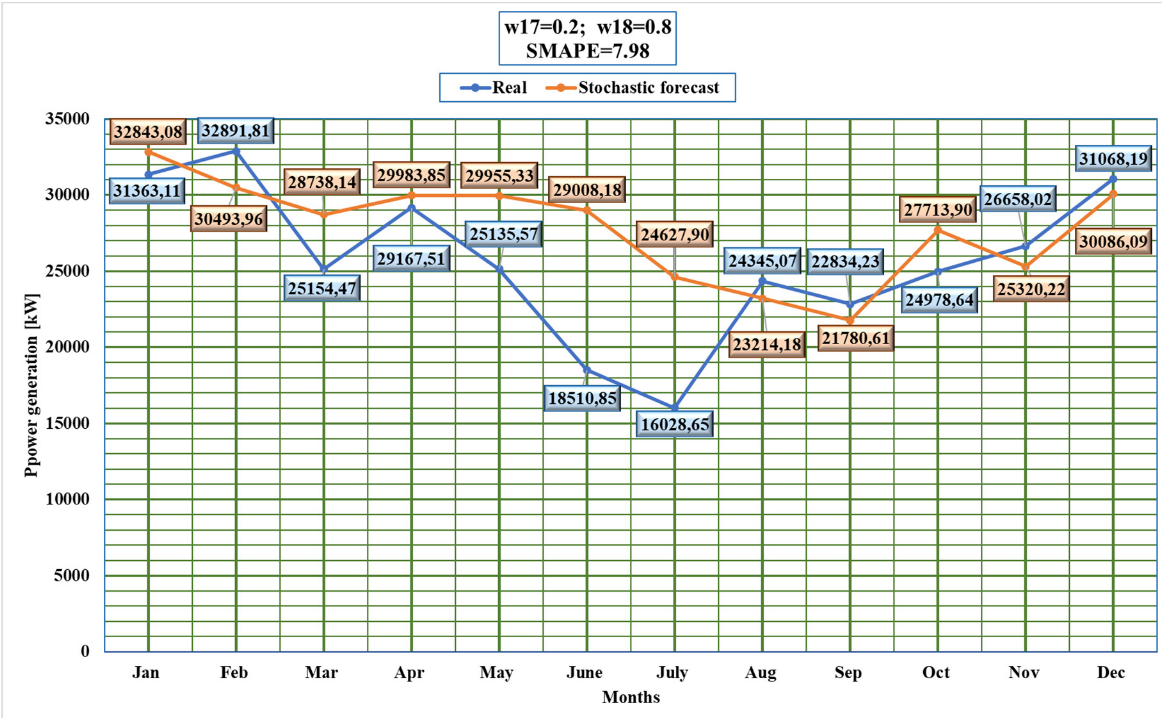

FIGURE 2. Numerical results in case $w_{17} = 0.2$; $w_{18} = 0.8$

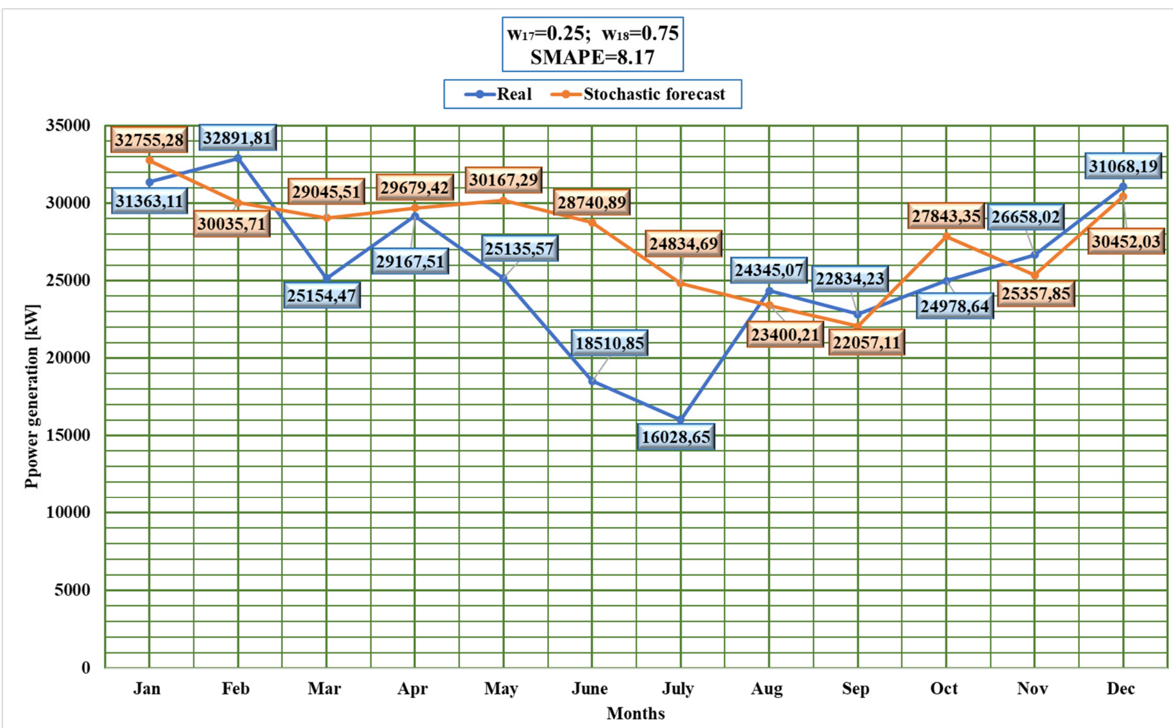

FIGURE 3. Numerical results in case $w_{17} = 0.25; w_{18} = 0.75$

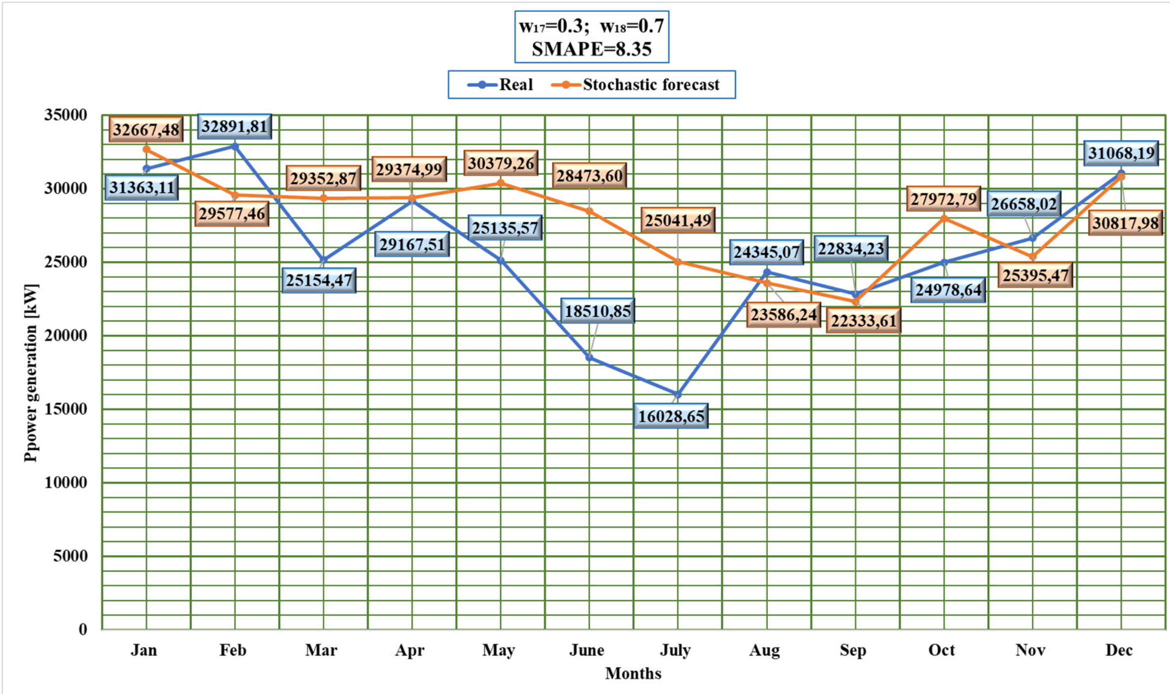

FIGURE 4. Numerical results in case $w_{17} = 0.3; w_{18} = 0.7$

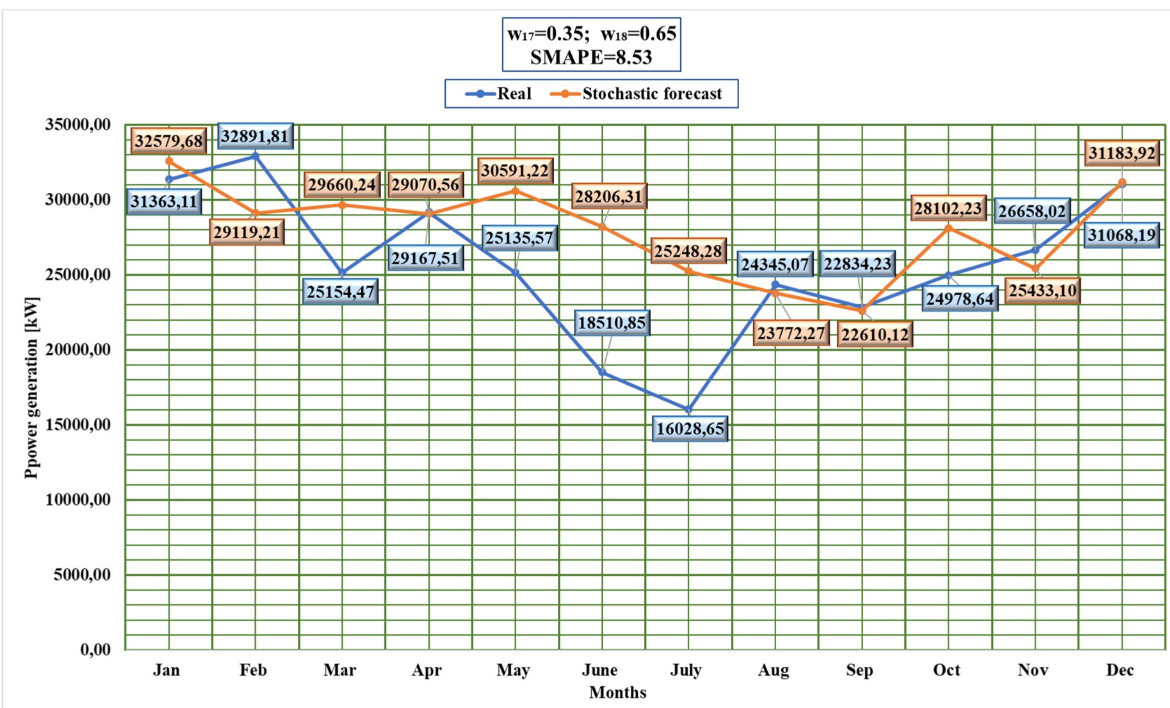

FIGURE 5. Numerical results in case $w_{17} = 0.35; w_{18} = 0.65$

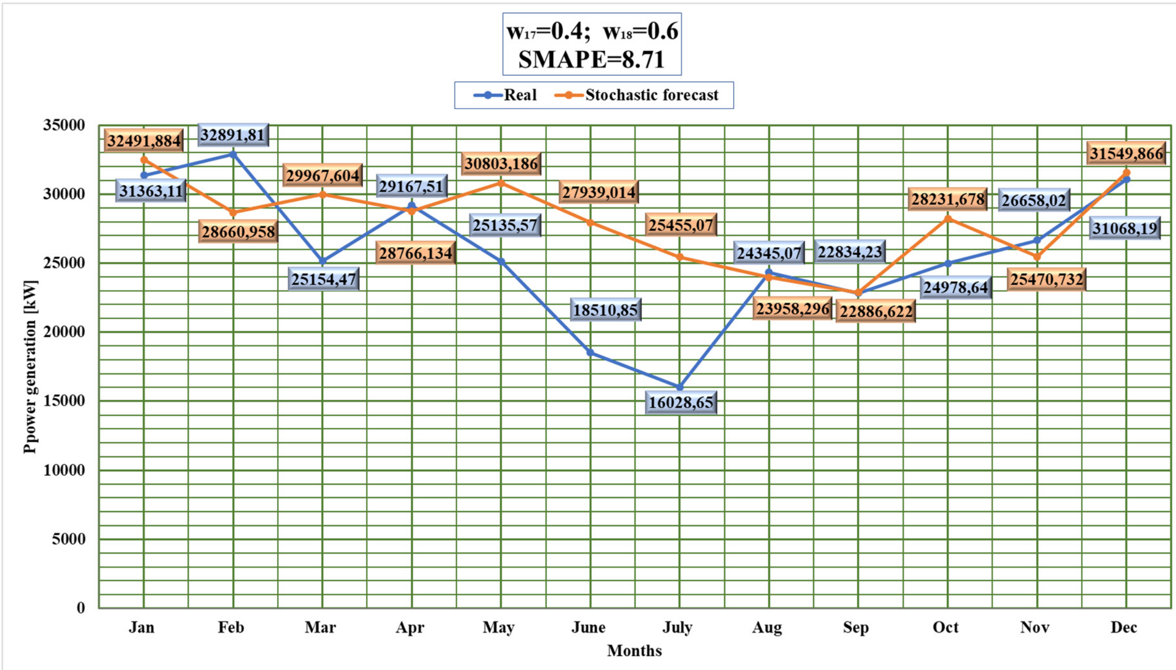

FIGURE 6. Numerical results in case $w_{17} = 0.4; w_{18} = 0.6$

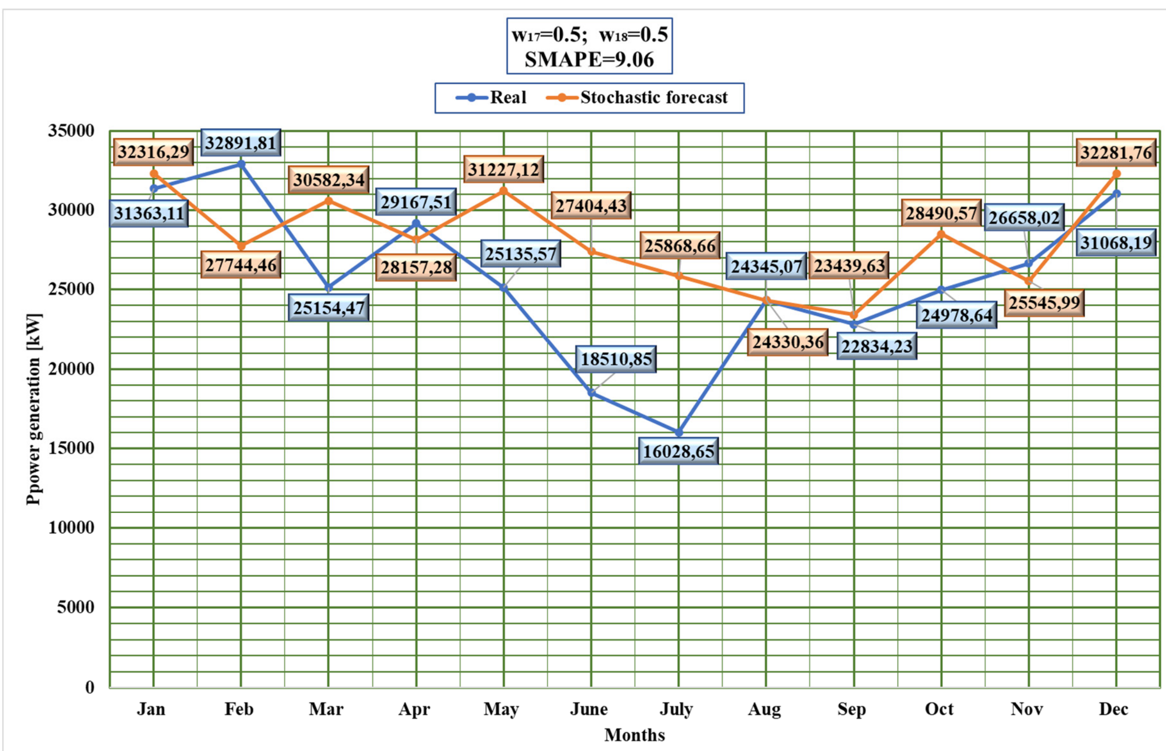

FIGURE 7. Numerical results in case $w_{17} = 0.5; w_{18} = 0.5$

For the numerical realization specialized software is used in software environments: Python, Maple, MS Excel.

Discussion on the Obtained Numerical Results

From the obtained numerical results in the proposed methodology for prediction of microgrid photovoltaic power generation the following conclusions can be made:

- When the weights in the model change, which characterize the stochastic nature of the forecast for the generated power for the respective years, it is observed that with increasing weight w_{17} and decreasing weight w_{18} the average absolute percentage error increases (Figure 8). This shows the greater influence of the generated power in the months of the previous year forecast than the others on the accuracy of the forecasting model.

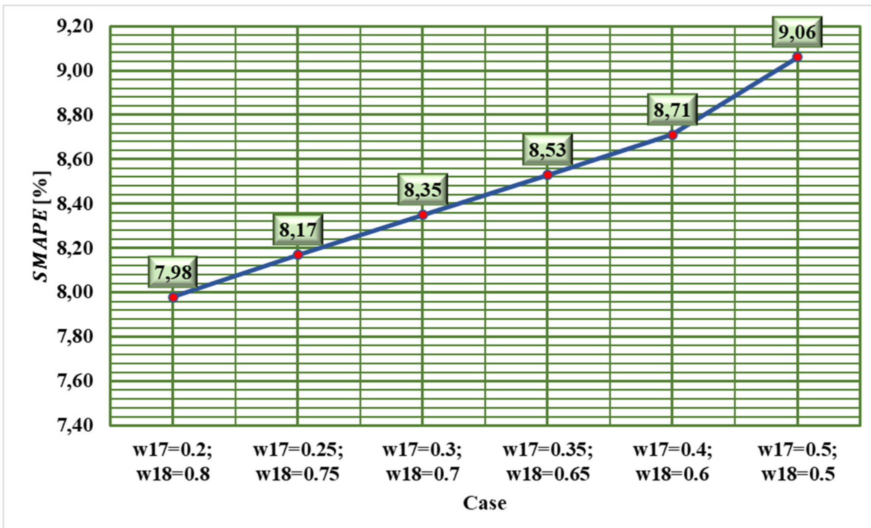

FIGURE 8. Dependence between the changes of w_{17} and w_{17} and the value of $SMAPE$ [%]

- The average absolute percentage error obtained in the numerical realization of the proposed model is not so good ($SMAPE = 7.98\% \div 9.06\%$). This is due to the fact that the stochastic forecast of the generated power is made only on the basis of historical data on the generated power for the specified period. In this case, the error estimated is good. If other factors that directly affect the process of power generation are taken into account such as intensity of solar radiation, daily duration of sunshine, ambient temperature, etc. and given their respective values in the period considered, it is assumed that the average absolute percentage error will be much lower than that obtained in this study. This will be considered in our further research works.

CONCLUSION AND FUTURE WORK

The relevance and significance of the presented research comes from the fact that in all modern activities the collection, storage, processing of information and extraction of useful results from its analysis is a key process. The presented stochastic model for prediction of microgrid photovoltaic power generation shows a possibility for prediction, which will be important for the specific object in terms of planning its further activities. Our future research will focus on forecasting capacity generation by adding additional factors that influence the process.

ACKNOWLEDGMENTS

This work was supported by the European Regional Development Fund within the Operational Program “Science and Education for Smart Growth 2014 - 2020” under the Project CoE “National center of mechatronics and clean technologies” - BG05M2OP001-1.001-0008.

REFERENCES

1. Alfredo Nespoli, Emanuele Ogliari, Sonia Leva, Alessandro Massi Pavan, Adel Mellit, Vanni Lughi and Alberto Dolara, *Day-Ahead Photovoltaic Forecasting: A Comparison of the Most Effective Techniques*, Energies 2019, 12, 1621, http://dx.doi.org/10.3390/en12091621, 2019.
2. Nahid-Al-Masood, Md. Iqbal Asif, Ahmed Manavi Alam, Shohana Rahman Deeba, Tareq Aziz, *Forecasting of Photovoltaic Power Generation: Techniques and Key Factors*, 2019 IEEE Region 10 Symposium (TENSYMP), https://doi.org/10.1109/TENSYMP46218.2019.8971337, 2019.
3. Nattawat Sodsong, Kun Ming Yu, Wen Ouyang, *Short-Term Solar PV Forecasting Using Gated Recurrent Unit with a Cascade Model*, 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), https://doi.org/10.1109/ICAIIC.2019.8668970, 2019.
4. Paras Mandala, Surya Teja Swarroop Madhira, Ashraf Ul haquec, Julian Mengc, Ricardo L. Pineda, *Forecasting Power Output of Solar Photovoltaic System Using Wavelet Transform and Artificial Intelligence Techniques*, Procedia Computer Science 12 (2012) 332 – 337, https://doi.org/10.1016/j.procs.2012.09.080, 2012.
5. Tanyo Ivanov, Rad Stanev, *Mathematical model of photovoltaic inverters*, IEEE, https://ieeexplore.ieee.org/document/9030705, 2019.
6. Utpal Kumar Das, Kok Soon Tey, Mehdi Seyedmahmoudian, Saad Mekhilef, Moh Yamani Idna Idris, Willem Van Deventer, Bend Horan, Alex Stojcevski, *Forecasting of photovoltaic power generation and model optimization: A review*, Renewable and Sustainable Energy Reviews, Volume 81, Part 1, January 2018, Pages 912-928,http://dx.doi.org/10.1016/j.rser.2017.08.017, 2018.
7. Wen Ouyang, Kun-Ming Yu, Nattawat Sodsong, Ken H. Chuang, *Short-term solar PV forecasting based on recurrent neural network and clustering*, Proceedings Volume 11321, 2019 International Conference on Image and Video Processing, and Artificial Intelligence; 113212U (2019), https://doi.org/10.1117/12.2550322, 2019.
8. Yuan-Kang Wu, Chao-Rong Chen, Hasimah Abdul Rahman, *A Novel Hybrid Model for Short-Term Forecasting in PV Power Generation*, Solar Energy and Clean Energy: Trends and Developments 2014, Research Article | Open Access, Volume 2014 |Article ID 569249 | 9 pages, https://doi.org/10.1155/2014/569249, 2014.