



Article

Advanced Forecasting Methods of 5-Minute Power Generation in a PV System for Microgrid Operation Control

Paweł Piotrowski ^{1,*}, Mirosław Parol ¹, Piotr Kapler ¹ and Bartosz Fetliński ²

¹ Institute of Electrical Power Engineering, Faculty of Electrical Engineering, Warsaw University of Technology, Koszykowa 75 Street, 00-662 Warsaw, Poland; mirosław.parol@pw.edu.pl (M.P.); piotr.kapler@pw.edu.pl (P.K.)

² Institute of Microelectronics and Optoelectronics, Faculty of Electrical Engineering, Warsaw University of Technology, Koszykowa 75 Street, 00-662 Warsaw, Poland; bartosz.fetlinski@pw.edu.pl

* Correspondence: pawel.piotrowski@pw.edu.pl; Tel.: +48-22-234-7314

Abstract: This paper concerns very-short-term (5-Minute) forecasting of photovoltaic power generation. Developing the methods useful for this type of forecast is the main aim of this study. We prepared a comprehensive study based on fragmentary time series, including 4 full days, of 5 min power generation. This problem is particularly important to microgrids' operation control, i.e., for the proper operation of small energy micro-systems. The forecasting of power generation by renewable energy sources on a very-short-term horizon, including PV systems, is very important, especially in the island mode of microgrids' operation. Inaccurate forecasts can lead to the improper operation of microgrids or increasing costs/decreasing profits for microgrid operators. This paper presents a short description of the performance of photovoltaic systems, particularly the main environmental parameters, and a very detailed statistical analysis of data collected from four sample time series of power generation in an existing PV system, which was located on the roof of a building. Different forecasting methods, which can be employed for this type of forecast, and the choice of proper input data in these methods were the subject of special attention in this paper. Ten various prognostic methods (including hybrid and team methods) were tested. A new, proprietary forecasting method—a hybrid method using three independent MLP-type neural networks—was a unique technique devised by the authors of this paper. The forecasts achieved with the use of various methods are presented and discussed in detail. Additionally, a qualitative analysis of the forecasts, achieved using different measures of quality, was performed. Some of the presented prognostic models are, in our opinion, promising tools for practical use, e.g., for operation control in low-voltage microgrids. The most favorable forecasting methods for various sets of input variables were indicated, and practical conclusions regarding the problem under study were formulated. Thanks to the analysis of the utility of different forecasting methods for four analyzed, separate time series, the reliability of conclusions related to the recommended methods was significantly increased.

Keywords: microgrids; operation control; power generation; PV system; very-short-term forecasting; machine learning; interval type-2 fuzzy logic system



Citation: Piotrowski, P.; Parol, M.; Kapler, P.; Fetliński, B. Advanced Forecasting Methods of 5-Minute Power Generation in a PV System for Microgrid Operation Control. *Energies* **2022**, *15*, 2645. <https://doi.org/10.3390/en15072645>

Academic Editor: Gabriele Grandi

Received: 12 March 2022

Accepted: 1 April 2022

Published: 4 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Microgrids are autonomous energy micro-systems that can operate in both the synchronous (parallel) mode with distribution system operators' grids and the island mode. Control of the microgrid operation in both modes, in particular in the island mode, is a very important issue. Forecasts of power generated from renewable energy sources and forecasts of power demand, in a very-short-term horizon, affect the proper microgrid operation, especially in the island mode. Because of this, forecasts are more and more important. Very-short-term power-generation forecasts, if imprecise, can cause increased costs/decreased profits for microgrids operators or improper operation of energy micro-systems.

It is expected that microgrids will undergo management of electrical power and energy in a very-short-term horizon. All active components of the microgrid, e.g., controllable

microsources, energy storage units, and controllable loads, take part in the management process. For the electrical power and energy-management process to proceed correctly, a lot of detailed data is needed. These data include, among others: data on current and forecast loads, data on current and forecast values of power and energy generated by nondispatchable sources (among them, renewable energy sources), and data on the current and forecasted prices of the electrical energy market. These data enable the correct control process of the above-mentioned active components of the microgrid. Obtaining accurate forecasts of power generated in PV systems in a very-short-term horizon is therefore very important from the point of view of power and energy management in the microgrid.

1.1. Related Works

The first part of the literature review refers to the very-short-term forecasting. Within this field, we distinguish between load-demand forecasts and power-generation (wind power and photovoltaic power) forecasts.

The problem of forecasting power demand in a very-short-term horizon is presented in several publications, e.g., in [1–4]. The authors of [4] describe the 10 s forecasting of power demand in the case of highly variable loads. In turn, paper [5] includes a very comprehensive overview of load forecasting methods in short-term and very-short-term horizons. Topics such as the different areas and locations to which this type of forecast can be applied (smart buildings, microgrids, small cities), along with forecasted time horizons, are described in this overview.

Various methods (models) can be applied to prepare forecasts of wind power generation in very short time horizons. In [6], a model of wavelet decomposition and weighted random forests for very-short-term wind power forecasts is presented. The authors of [7] describe hybrid empirical mode decomposition and team empirical mode decomposition models for the needs of wind power forecasts. The authors of [8] present various approaches: neuro fuzzy systems, a support vector regression, and a regression tree in the case of forecasting 1 h wind power. The authors of [9] address, in turn, different approaches for forecasts of wind power in minute horizons. The fuzzy model of Takagi–Sugeno applied to very-short-term forecasts of wind power is presented in [10]. In [11], models based on a discrete-time Markov chain for very-short-term wind power forecasting are described.

Another aspect to be considered is photovoltaic power forecasting in very-short-term or short-term horizons. Two methods, including smart persistence and random forests for the needs of forecasts of PV energy production, are presented in [12]. The authors of [13] address a team model for short-term PV power forecasts. In [14], a complex model for solar power forecasts is described. The model combines wavelet transform, ANFIS, and hybrid firefly and PSO algorithms. The authors of [15] discuss a physical hybrid ANN for 24 h-ahead PV power forecast in microgrids. A very comprehensive review and evaluation of different methods (models) for PV power forecasting are included in [16]. A review of various methods concerning power-generation forecasting in PV systems is also presented in [17]. Paper [18] includes an extensive comparison of different physical models, which can be used for the needs of forecasts concerning PV power generation. In turn, the impact of the availability of design data on the exactness of power-generation forecasts in PV systems, based on physical models, is described in [19].

The second part of the literature review specifically refers to microgrids.

The topic of microgrids was discussed intensively in the literature. In [20,21], a formal definition of microgrids is presented. The idea of microgrids was described in many other publications, e.g., [22,23]. A lot of books and papers address the topic of microgrids' operation control [22–29]. In [24,28,29], a very comprehensive overview of works relating to optimum control (centralized control and decentralized (distributed) control) in microgrids is presented. The authors of [25,28] describe the centralized control logic. In turn, the distributed control logic in microgrids is discussed in [26,28,29]. The authors of [25] present the model of predictive control in microgrids. The operational

control in the microgrid island mode is addressed in [25,27]. In [30], a fault detection, localization, and categorization method in the case of a PV-fed DC microgrid is described.

In the analyzed works, different issues concerning photovoltaic power forecasting in a very-short-time horizon and microgrids were considered. The main aim of this paper is to provide a very comprehensive review of the various possible methods of very-short-term photovoltaic power-generation forecasting for the needs of low-voltage microgrids operation, as well as select the best methods among those considered.

1.2. Objective and Contribution

The following are the main objectives of this paper:

- Carry out an analysis of the statistical properties of a time series of the measured values of 5 min power generation in a PV system;
- Verify the usefulness of the available input variables—perform a validity analysis (the time series of solar irradiance, air temperature, PV module temperature, wind direction, and wind speed) using four different methods and select eight sets of input variables to make forecasts using various methods;
- Check the efficiency of 5 min horizon power-generation forecasts by means of ten forecasting methods, including machine learning, hybrid, and ensemble methods (several hundred various models with different set values of parameters/hyperparameters have been verified for this purpose);
- Point out forecasting methods that are the most effective for this 5 min power-generation time series depending on the number of input variables used.
- The selected contributions of this paper are as follows:
- The research concerns unique data—a time series of 5 min power-generation values in a small, consumer PV system. In the case of such small PV systems and such a short forecast horizon (5 min), meteorological forecasts are usually not used in the forecasting process due to difficulties in obtaining them, which makes it problematic to obtain forecasts with very high accuracy;
- We provide a detailed description of the performance of photovoltaic systems regarding the main environmental parameters;
- We performed extensive statistical analyses of the available time series (including an analysis of the importance of the input variables);
- We used tests of ten different prognostic methods (including hybrid and team methods);
- We developed a new, proprietary forecasting method—a hybrid method using three independent, MLP-type neural networks;
- We indicate the most favorable prognostic methods for various sets of input variables (from 3 input variables to 15 input variables) and formulate practical conclusions regarding the problem under study, e.g., from the point of view of microgrids' operation.
- We provide a broad comparative analysis of forecasting methods of a very-short-term horizon for power generation in PV systems that can be connected to low-voltage microgrids.

After completing our studies, we can state that there are efficient, very-short-term forecasting methods for PV power generation, which are suitable for practical use in microgrids' operation.

The organization of this paper is as follows: Section 2 describes the influence of the main environmental parameters on the performance of photovoltaic systems. Section 3.1 includes an analysis of the statistical properties of the time series of PV power generation data investigated in this paper. The analysis leading to the choice of proper input data (explanatory variables) for various prognostic methods is shown in Section 3.2. Section 4 addresses the forecasting methods applied in this paper. In turn, Section 5 discusses criteria employed to evaluate the quality of the forecasting models considered. A broad comparative analysis of forecasting methods of a very short time horizon for power generation in PV systems is presented in Section 6. Section 7 includes the main conclusions resulting from our studies. A list of references ends the paper.

2. Performance of Photovoltaic Systems

The two main environmental parameters affecting the performance of photovoltaic (PV) systems are solar irradiance and cells' temperature [31,32]. Changes in solar irradiance result in a generally proportional shift of the I–V (current–voltage) curve along the current axis, along with a relatively much smaller voltage change. Under low-irradiance conditions, such as those during overcast weather, the maximum power of the PV module tends to be further decreased due to the higher significance of the parallel resistance, which results in a slight decrease in current with an increasing voltage. This effect is highly dependent on PV cells' technology. The current changes resulting from the changing irradiance are instantaneous from the point of view of PV system energy yields. The PV system power output is primarily dependent on the available irradiance.

The PV cell temperature is the second most important factor influencing the energy output of a PV system, as demonstrated by analyses utilizing the performance ratio (PR) parameter to model PV systems' operations [33,34]. The increase in the PV cell temperature results in a decrease in the PV device's open-circuit voltage, along with a minor increase in the short-circuit current. The PV output power temperature coefficients of silicon-based solar cells are of the order of $-0.45\%/K$ [35]. Due to heat capacity of PV modules being heavily dependent not only on the materials and structure of the module itself but also on its mounting structure, tilt angle, and surrounding ground, the rate of response of the module's temperature to the environmental conditions (the irradiance, velocity, and direction of wind and the ambient temperature) varies significantly and must be assumed to be an individual property of the particular system under analysis. The literature provides numerous similar approximations of the influence of temperature on PV systems' efficiency and output power, often using empirically established coefficients [36]. Direct measurement of the temperature of laminated solar cells is difficult, and temperature sensors are usually attached to the rear backsheet of the module. The significant temperature gradient between different points of a single module exposed to sunlight (due to proximity of the frame or mounting structure attachment) makes reliable module temperature challenging, with guidelines suggesting the use of up to four temperature sensors on a single module to model the temperature correctly [37,38]—an effort rarely undertaken, even in research-oriented test systems, and even more so in commercial systems.

Spectral effects, related to the mismatch between the spectral response of a PV module (which primarily depends on PV cell technology) and the spectrum of the incident irradiance (which consists of the direct and diffused component of the solar spectrum and light reflected from the surrounding objects—particularly important for bifacial and multijunction modules), primarily contribute towards varying irradiance effects. However, the spectral mismatch also impacts thermalization and sub-bandgap losses, which result in PV module heating. These factors are difficult to quantify in the analysis of PV systems' performance, as their inclusion would require long-term monitoring of the solar spectra in the location of the system under analysis. Their impact is also highly specific to PV cell technology [39]. The size and layout of a PV system may also impact both the degree and pace of the change of its power output due to external factors, which is particularly important in the case of large-area systems [40].

3. Data

3.1. Statistical Analysis of the Time Series of Power-Generation Data

The installed power of the analyzed PV system is 3.2 kW. The power output of the analyzed system was monitored using the built-in capability of the system's inverter, type SMA Sunnybox SB3000. The built-in measurement system records parameters, such as AC and DC side power, voltages, and currents. The data points are recorded in 5 min intervals. These electrical data are then merged with the data gathered by the meteorological station. Statistical analysis is based on fragmentary time series, including 4 full days. Each day is from a different season. The daily time series includes 288 periods of 5 min. The total number of 5 min periods of power generation in watts is 1152. Before statistical analysis

was performed, the data “cleaning” process was performed. Wrong data were identified and replaced with data most relevant to their location (e.g., in the case of non-zero power-generation values in the period between sunset and sunrise or zero-generation values in the period when solar irradiance was non-zero).

Table 1 shows selected statistical measures of the time series of power generation in the PV system. As much as 50% of power generation from the time series was due to small values, below 46.914 (W) (which is more than 68 times less than the installed power of the PV system).

Table 1. Descriptive statistics of time series of power generation.

Statistical Measures	PV System Data
Mean	635.61 (W)
Percentage ratio of mean power to installed power	19.86%
Standard deviation	930.63 (W)
Minimum	0.00 (W)
Maximum	3114.81 (W)
Range	3114.81 (W)
Coefficient of variation	146.41%
The 10th percentile	0.00 (W)
The 25th percentile (lower quartile)	0.00 (W)
The 50th percentile (median)	46.91 (W)
The 75th (upper quartile)	1127.19 (W)
The 90th percentile	2368.10 (W)
Variance	866,074.80 (-)
Skewness	1.23 (-)
Kurtosis	-0.04 (-)

Figure 1 shows the daily time series of power generation for every season of the year (actual measurement data). The whole spring day was cloudless (generation was close to the rated power, very smoothed time series). The opposite of a spring day was a winter day with a much shorter power-generation period and a significantly smaller generation compared to a spring day and a summer day. Dynamic changes in the quantity of generation during the summer and autumn days are evidence of the high variability of cloud cover on these days.

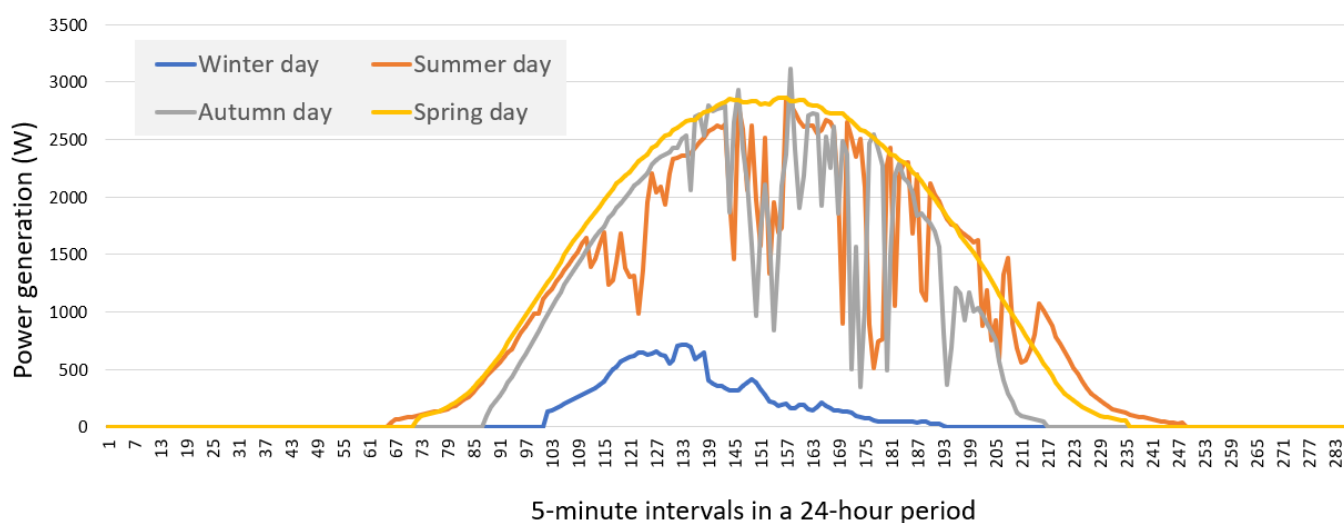


Figure 1. Daily time series of power generation for every season of the year.

For the time series of power generation, the autocorrelation coefficient (ACF) slowly decreased from 0.974 (one period back, e.g., 5 min) to 0.892 (twelve periods back, e.g., 1 h) (see Figure 2). All autocorrelation coefficients are statistically significant (5% significance

level). The use of several past values of the forecasted time series of power generation as input data for forecasting models seems justified.

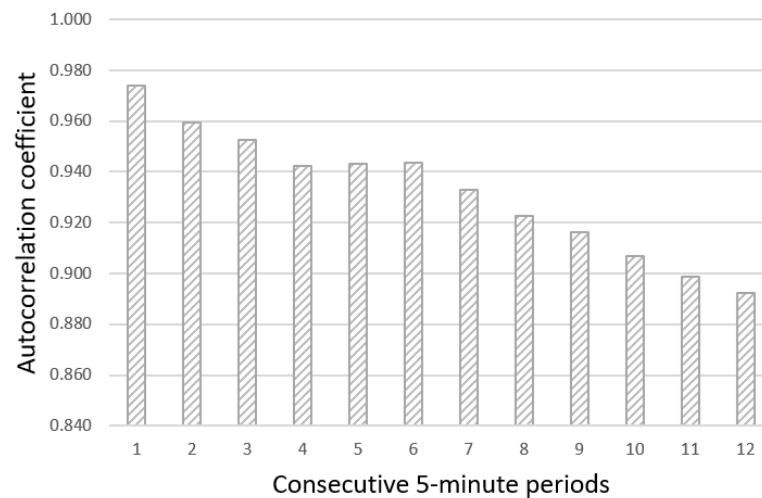


Figure 2. Autocorrelation function (ACF) of the analyzed time series of power generation.

3.2. Analysis of Potential Input Data for Forecasting Methods

The forecasted output is the power generation in the PV system (generation in the DC part of the system). Five additional time series (measured, real values) are available for analysis as potential input data. There are no forecasts of these time series. The following time series are available:

- Solar irradiance (W/m^2);
- Air temperature ($^{\circ}\text{C}$);
- PV module temperature ($^{\circ}\text{C}$);
- Wind direction (degrees);
- Wind speed (m/s).

Only the past values of the five exogenous explanatory variables and the past values of the dependent variable (endogenous variable) can be selected as input data for the forecasting methods. Furthermore, a weighted averaging of the time series of power-generation values can be performed. This activity should reduce the random component of this time series. The selected past values of such transformed time series may be a valuable set of input data. They can even potentially replace the past values of the forecasted time series as input data in the forecasting model. The values of the smoothed time series of power generation were calculated from Equation (1).

$$P_t^{\text{smoothed}} = P_{t-1} \cdot w_{t-1} + P_{t-2} \cdot w_{t-2} + P_{t-3} \cdot w_{t-3}, \sum_{k=1}^{k=3} w_{t-k} = 1 \quad (1)$$

where P_t^{smoothed} is the smoothed value of power generation for period t and P_{t-k} is the value of power generation for period $t-k$, $w_{t-1} = 0.6$, $w_{t-2} = 0.3$ and $w_{t-3} = 0.1$.

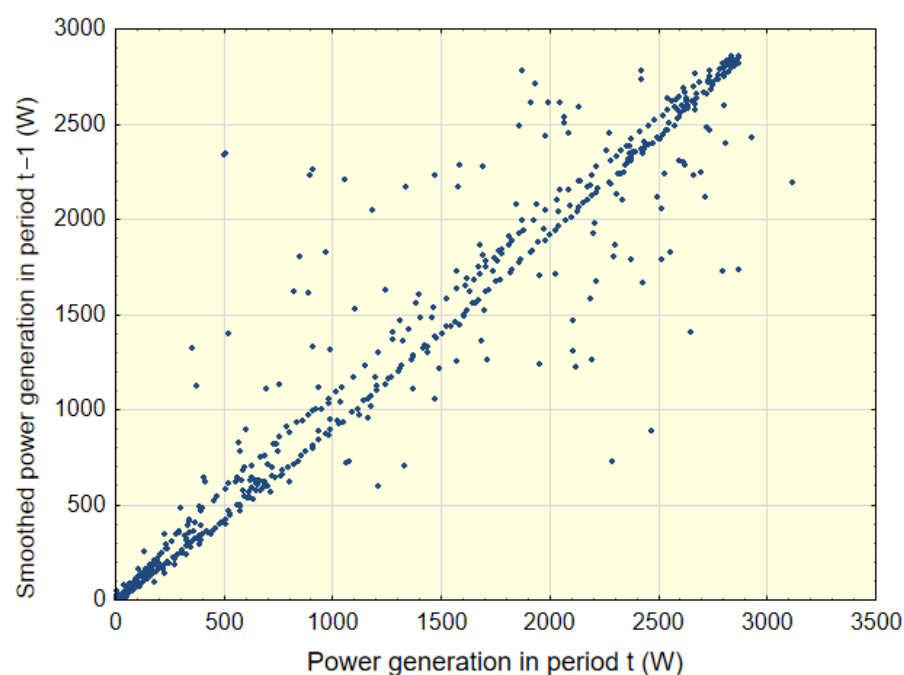
Table 2 shows Pearson linear correlation coefficients (R) between the 5 min power generation and the potential explanatory variables considered. All correlation coefficients are statistically significant (5% level of significance). The number of expertly proposed past values (from one to three withdrawals) for each explanatory variable results from the value of the Pearson correlation coefficient (the higher the value of the Pearson coefficient, the greater the significance of the variable) and the independence of information contained in a given explanatory variable (the small value of the Pearson coefficient comparing the analyzed explanatory variable and other explanatory variables).

Table 2. Values of Pearson linear correlation coefficients between 5 min power generation and the explanatory variables considered.

Code of Variable	Potential Explanatory Variables Considered	R
SPG(T-1)	Smoothed power generation in period $t-1$	0.9756
PG(T-1)	Power generation in period $t-1$	0.9744
PG(T-2)	Power generation in period $t-2$	0.9601
PG(T-3)	Power generation in period $t-3$	0.9536
SI(T-1)	Solar irradiance in period $t-1$	0.9661
SI(T-2)	Solar irradiance in period $t-2$	0.9587
SI(T-3)	Solar irradiance in period $t-3$	0.9385
AT(T-1)	Air temperature in period $t-1$	0.4261
PV_MT(T-1)	PV module temperature in period $t-1$	0.7134
WD(T-1)	Wind direction in period $t-1$	-0.2475
WS(T-1)	Wind speed in period $t-1$	0.1825

The three past values of power generation and the three past values of solar irradiance have very large and similar values of the Pearson's coefficient related to the dependent variable (output data)—power generation. PV module temperature in period $t-1$ has a significantly greater R value than the air temperature in period $t-1$. The smallest R values have wind direction in period $t-1$ and wind speed in period $t-1$. All R values except wind direction in period $t-1$ are positive.

Figure 3 presents dispersion diagrams—relationships between power generation in period t and smoothed power generation in period $t-1$. The relationship is close to linear. The strongest linear relationship is visible for values close to the extremes (power generation close to zero and power generation close to the rated power). The few points significantly deviating from the linear relationship can be interpreted as a change in cloud cover over a period of 5 min. The Pearson linear correlation coefficient between the output data (the power generation in period t) and proposed new input data (the smoothed power generation in period $t-1$) is equal to 0.9756. This R value for the smoothed power generation in period $t-1$ is the biggest of all potential input data.

**Figure 3.** Relationship between power generation in period t and smoothed power generation in period $t-1$.

In order to determine the importance of the potential input data, the following methods of selecting variables were additionally used, using all possible 11 inputs and 1 output:

- C&RT decision trees algorithm for the selection of variables in regression problems—for each potential predictor (input data), the coefficient of determination R^2 is calculated;
- Analysis of variances (F statistics)—this method calculates the quotient of the intergroup variance to the intragroup variance (the dependent variable) in predictor intervals (the number of quantitative predictor classes is determined before the analysis);
- Global Sensitivity Analysis (GSA statistics) for multilayer perceptron (MLP) neural network. A neural network with one hidden layer and four neurons in this layer was used for the analysis. The training algorithm is BFGS, the activation function in the hidden layer is the hyperbolic tangent, and the activation function in the output layer is linear. The value of the importance factor for input data number k is the quotient of the RMSE error of the forecasts of the trained MLP network using the remaining 10 input data and the input data number k is replaced by its mean value from the total data to the RMSE error of the forecasts using all 11 sets of input data. The greater the value of the importance factor for the given input data, the greater their significance. Results below 1 for a given input data mean that these input data can probably be eliminated because the MLP network without these input data has a lower RMSE error in the forecasts;
- The importance of input data using the random forest (RF) algorithm is the many decision trees (DCs). The importance of the given input data is measured by checking to what extent nodes (in all decision trees) using the input data reduce the impurity Gini indicator, with the weight of each node being equal to the number of associated training samples [38]. It was assumed for the analysis that each decision tree would have 6 randomly selected sets of input data from the total of 11.

The results of the input data selection with the C&RT decision tree algorithm are shown in Figure 4. The values of the coefficient of determination are sorted in descending order. The most important explanatory variable according to this method is smoothed power generation in period $t-1$.

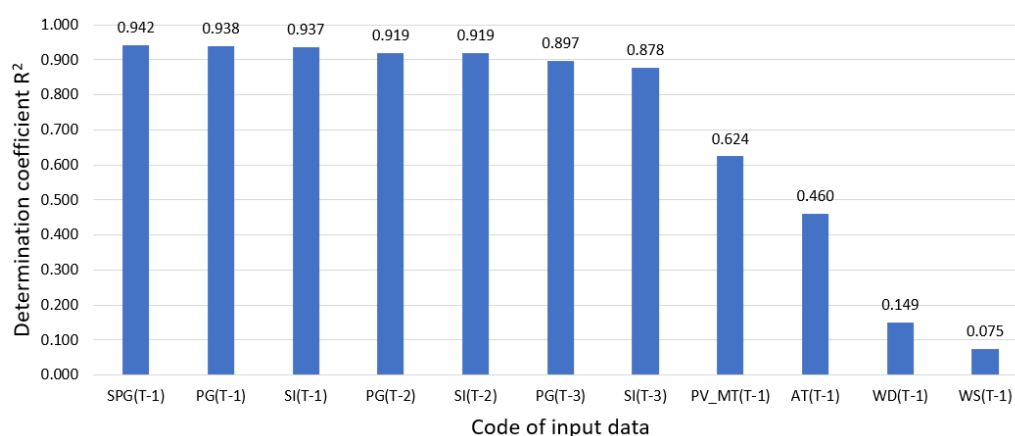


Figure 4. The results of the input data selection with the C&RT decision tree algorithm.

In Figure 5, the results of the input data selection with the use of the analysis of variance (F statistics) are presented. The F values are sorted in descending order. Power generation in period $t-1$ is the most important explanatory variable according to this method.

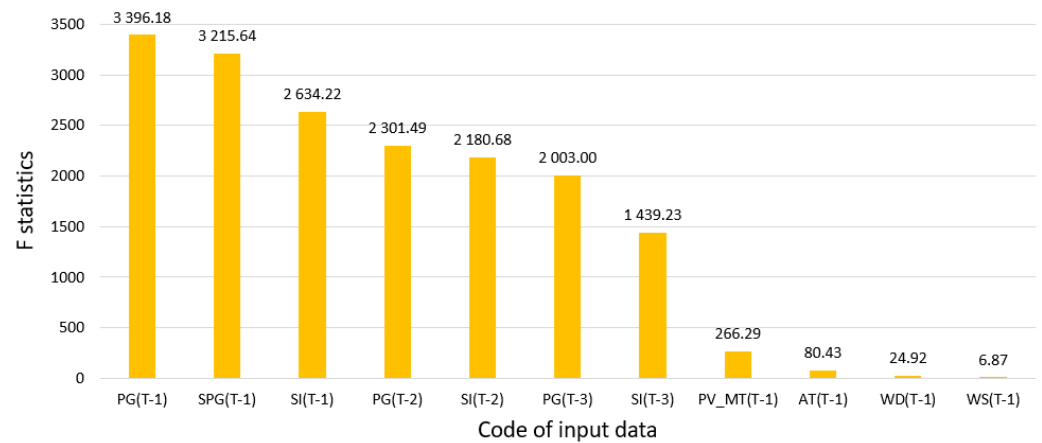


Figure 5. The results of the input data selection using analysis of variance (F statistics).

The results of input data selection using the Global Sensitivity Analysis for the multilayer perceptron (MLP) neural network are shown in Figure 6. The importance factor values are sorted in descending order. The most important explanatory variable according to this method is solar irradiance in period $t-1$.

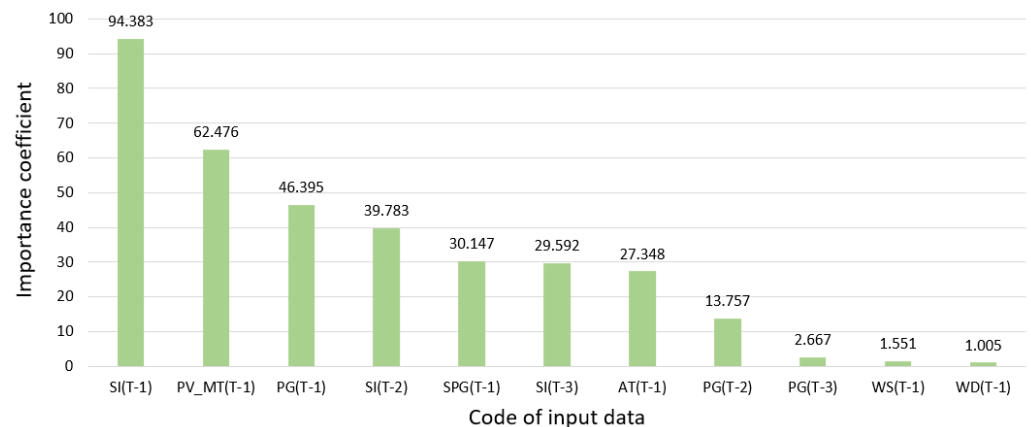


Figure 6. Input data selection results using Global Sensitivity Analysis for MLP-type neural network.

In Figure 7, the results of the input data selection with the use of the random forest algorithm are presented. The importance values are sorted in descending order. Smoothed power generation in period $t-1$ is the most important explanatory variable according to this method.

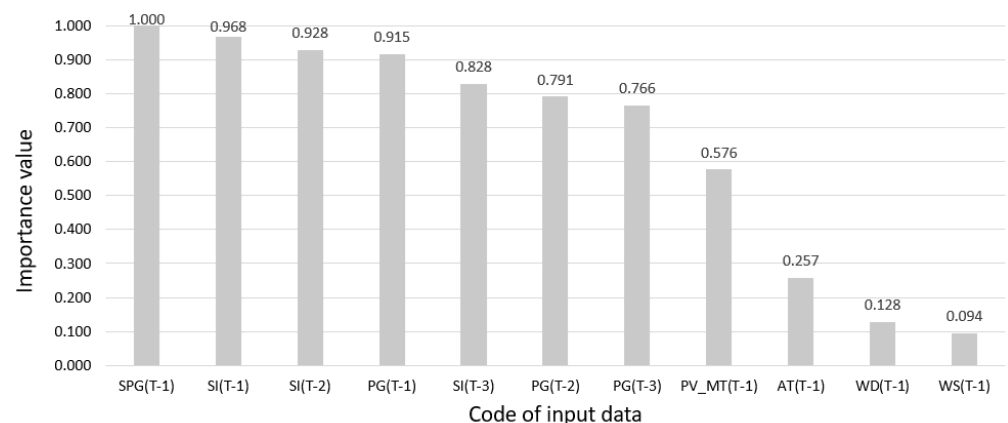


Figure 7. The results of the input data selection using the random forest algorithm.

Based on the analysis of the selection of variables using these four methods, the following conclusions can be drawn:

- For all analyzed methods of selecting variables, the significantly least important input data are wind direction in period $t-1$ and wind speed in period $t-1$. In the vast majority of cases, the last, least-important input data are (somewhat surprisingly) wind speed in period $t-1$;
- The best input data include smoothed power generation in period $t-1$, power generation in period $t-1$, and solar irradiance in period $t-1$.
- The results of the individual selection methods were quite similar, except for the input-data-selection method using global sensitivity analysis for the MLP-type neural network. In this case, the most important input data—solar irradiance in period $t-1$ —are significantly more important than the second (surprisingly) in order input data, the PV module temperature in period $t-1$. This method also obtained validity results with the greatest diversification of numerical values;
- For all analyzed methods of selection of input data, the PV module temperatures in period $t-1$ are more important input data than the air temperature in period $t-1$;
- The results of the input-data-selection method with the C&RT decision tree algorithm (values of the coefficient of determination) are very similar to the values of Pearson’s linear correlation (Table 2), both in relation to the order of input data in the ranking as well as the values of the coefficients.

Table 3 shows the input datasets that will be applied to forecasts using various methods, including hybrid methods and team methods. The input datasets proposed for the forecast quality tests assume the use of all data nominated on the basis of the selection made using four methods, as well as the use of a limited number of inputs for a given method (e.g., maximum of four sets of input data—this is the limitation of the Interval Type-2 Fuzzy Logic System method due to the computational time consumption). Thanks to the construction of many sets with a different number of input data, it will be possible to verify whether it is reasonable to limit data to those that selection methods indicate as the most important input data or whether it is better to use all available input data that are statistically significant. The persistence model only uses the last known value of the forecast time series for the prediction (set 0 (1 input)). This model is a reference point for other more advanced methods, the forecasts of which should have lower error measures.

Table 3. Sets of input data selected for forecasting methods.

Name of Set	Codes of Input Data and Additional Comments
Set 0 (1 input)	PG(T-1)
SET I (3 inputs)	PG(T-1), PG(T-2), PG(T-3)
SET II A (4 inputs)	PG(T-1), SI(T-1), PV_MT(T-1), AT(T-1)
SET II B (4 inputs)	SPG(T-1), SI(T-1), PV_MT(T-1), AT(T-1)
SET II C (3, 3, 4 inputs)	PG(T-1), PG(T-2), PG(T-3)—inputs for predicting $PG_{forecast}(T)$ SI(T-1), SI(T-2), SI(T-3)—inputs for predicting $SI_{forecast}(T)$ PG _{forecast} (T), SI _{forecast} (T), PV_MT(T-1), AT(T-1)—inputs for predicting PG(T)
SET III (11 inputs)	SPG(T-1), PG(T-1), PG(T-2), PG(T-3), SI(T-1), SI(T-2), SI(T-3), AT(T-1), PV_MT(T-1), WD(T-1), WS(T-1)
SET IV (3, 3, 13 inputs)	PG(T-1), PG(T-2), PG(T-3)—inputs for predicting $PG_{forecast}(T)$ SI(T-1), SI(T-2), SI(T-3)—inputs for predicting $SI_{forecast}(T)$ SPG(T-1), PG(T-1), PG(T-2), PG(T-3), SI(T-1), SI(T-2), SI(T-3), AT(T-1), PV_MT(T-1), WD(T-1), WS(T-1), PG _{forecast} (T), SI _{forecast} (T)—inputs for predicting PG(T)
SET V (15 inputs)	SPG(T-1), PG(T-1), PG(T-2), PG(T-3), SI(T-1), SI(T-2), SI(T-3), AT(T-1), AT(T-2), AT(T-3), PV_MT(T-1), PV_MT(T-2), PV_MT(T-3), WD(T-1), WS(T-1)

One of the sets (set I (three inputs)) assumes the use of only three retracted values of the forecast time series. This is to compare the quality of forecasts based only on the time series without the use of exogenous input variables with the forecasts using additional exogenous input variables.

Set II C (three, three, and four inputs) and set IV (three, three, and thirteen inputs) are sets for the hybrid method. The first model forecasts power generation in period t using the last three values of the time series. The second model forecasts solar irradiance in period t using the last three values of the time series. The third model that generates the correct final forecast of power generation in period t uses the forecasts from the first model and the second model as input data.

Set V uses all available statistically significant data, including the last three sets' previous values of the following variables: power generation, solar irradiance, PV module temperature, and air temperature.

4. Forecasting Methods

This section describes the methods employed in this paper. Forecasts are made using single methods, ensemble methods, and hybrid methods. In total, ten prognostic methods were used. Figure 8 presents a general diagram of subsequent activities related to the forecasting process.

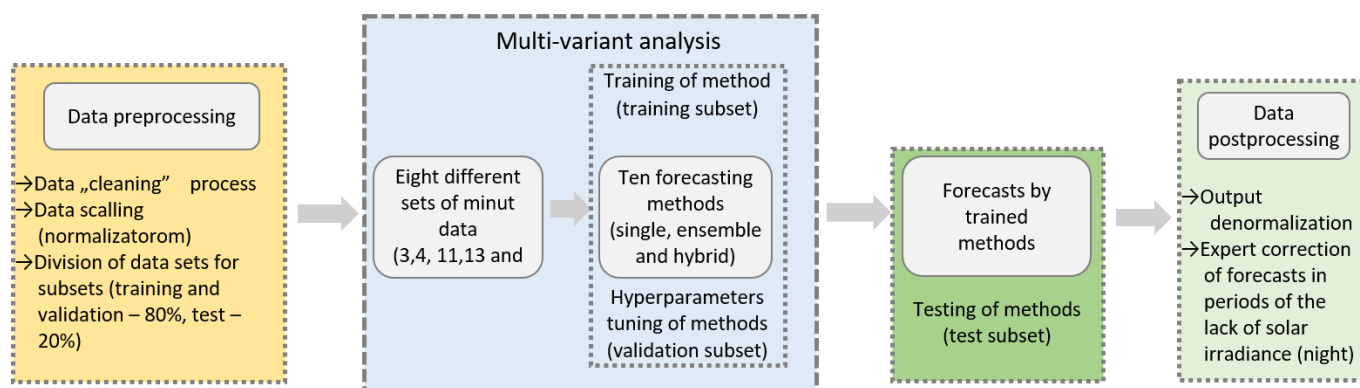


Figure 8. A general diagram of the consecutive steps in the forecasting process.

In the first step, data were preprocessed. In the beginning, before the process of data scaling (normalization) and data processing into the appropriate sets (input data and output data, the process of data “cleaning” was performed. Next, the data from the time series of the PV system’s power generation were normalized to relative units (one relative unit is equal to the installed power). The other time series of data (exogenous input variables) were normalized using min–max scaling. The data, including 1152 periods of 5 min, were divided into three subsets: training, validation, and test subsets, respectively. Training and validation subsets consisted of 80% of the time series chosen randomly (division into training and validation parts, which are different depending on the forecasting method used). The test subset comprised the remaining 20% of the time series via random selection. Estimation of model parameters was performed with the training subset. The validation subset was used for tuning the hyperparameters of the selected methods. The last one—test subset—was applied to find the final results of errors in the forecast methods used. The choice of the training and validation subsets from 80% of the data of the time series was made with the usage of gradient-boosted trees (GBT) along with the bootstrap technique. The multiple linear regression model (LR) only used the training subset (80% of the data of the time series) without a validation subset—this model had no hyperparameters, only parameters determined during a one-time parameter-optimization process.

Next, multivariate analysis was performed—using the predictive methods on eight different input datasets in the training subset and the selection of appropriate hyperparameters of the methods in the validation subset. An example of the selected hyperparameters and the scope of their searches for the selected methods is included in Appendix A—Table A1.

Then, the final predictions for the subset test were made for all methods with the selected hyperparameters.

Postprocessing was performed in the last step. The values of the generated forecasts were scaled (de-normalized) to natural values (watts). An expert forecast correction was performed—non-zero power-generation values from the periods between sunset and sunrise were reset (power generation is impossible) in these time periods.

Following is a brief description of the proposed predictive methods. The persistence model was a benchmark for the quality of other, more advanced forecasting methods.

Persistence model. The naive model was the simplest to implement. It assumes that the forecast generation value is equal to the actual power-generation value obtained from the period 5 min before. Forecasts were calculated by Equation (2):

$$\hat{y}_t = y_{t-1} \quad (2)$$

where \hat{y}_t —forecast power generated by the PV system in a 5 min period t and y_{t-1} —power generation in a period lagged by $t-1$ from forecast period t .

Multiple linear regression model. This is a linear model that adopts a linear association among the input variables and the single output variable [41,42]. The input data are particular lags of the forecasted output variable. The other input explanatory variables (including their particular lags) are correlated to the output variable. The least-squares approach was used to fit the model.

K-Nearest Neighbors Regression. This technique is a non-parametric method used for regression and classification tasks [42,43]. The input consists of the k nearby training examples from the feature space. When using the KNN regression, the output is the property value for the object. This value represents the average of the values of the k nearest neighbors. The number of nearest neighbors is treated as the main hyperparameter for the tuning process. Models with a very low k value of 1 or 2 are most likely to suffer from overfitting. Along with increasing the value of k , this model should work more efficiently, but it may also lead to an increase in the load on the model and the occurrence of underfitting. The distance metric is the second hyperparameter.

MLP-type artificial neural network. This is a group of feedforward artificial neural networks (ANNs). MLP is an effective and popular linear or non-linear (depending on the kind of activation function in hidden layer/layers and output layer) universal approximator [44,45]. It consists of one input layer, which typically has one or two hidden layers, and one output layer. It often uses the backpropagation algorithm for the supervised learning process. The number of neurons in the hidden layer(s) is usually the main hyperparameter in the tuning task. Another selectable hyperparameter is the activation function in the hidden layer(s) and in the output layer. The Broyden–Fletcher–Goldfarb–Shanno (BFGS) method used for solving unconstrained non-linear optimization problems was chosen as a learning algorithm for the neural network.

Support Vector Regression. SVM for regression of the Gaussian kernel converts the classification process into regression by specifying the width ε tolerance region around the destination [46]. The learning process for SVR is diminished to the quadratic optimization problem and depends on several hyperparameters, such as tolerance ε , regularization constant C , and width parameter s of the Gaussian kernel.

Interval Type-2 Fuzzy Logic System. Type-2 fuzzy sets (T2 FSs) are used in type-2 fuzzy logic systems (T2 FLSs). Type-2 fuzzy sets are an expansion of type-1 fuzzy sets (T1 FSs). Investigations on T2 FSs were performed by Zadeh, Karnik, Mendel, and Liang [47–49]. Membership functions with three dimensions (MFs), including a footprint of uncertainty (FOU), are features of T2 FSs [50]. The structure of T2 FLSs was presented, e.g., in [4]. Typical blocks include the fuzzification block, the fuzzy inference block, the base of fuzzy rules, the type reduction block, and the defuzzification block as components of T2 FLSs. In the type reduction block, the transformation of T2 FS to T1 FS occurs. Usually, for type reduction, the Karnik–Mendel (KM) algorithm is employed [48].

Interval type-2 fuzzy logic systems (IT2 FLSs) (see, e.g., [50]) are often used in practice because of the computational complexity of T2 FLSs [51]. Among the different IT2 FLSs, the IT2 TSK FLS (the IT2 FLS with the inference model of Takagi–Sugeno–Kang [50]), or

the IT2 S FLS (the IT2 FLS with the Sugeno inference model), can be distinguished. IT2 TSK FLS and IT2 S FLS require a lower number of model parameters than the standard IT2 FLS. Genetic algorithms (GAs) or PSO algorithms are often used in the training process of the IT2 FLSs (in the determination of their parameters' values).

Random Forest Regression. RF is a collaborative method based on numerous single decision trees (the same type of models). In the regression process, the prediction in a single decision tree consists of the average target value of all instances related to the single leaf node [4]. The final prediction is the average value of all n single decision trees. Random forests are created on the basis of quite deep trees—forecasts using this method are characterized by a low load along with quite a large variance. The regularization hyperparameters depend on the algorithm used but generally restricted, are among others, are factors such as the minimum number of data points placed in a node before the node is split, the maximum number of levels in each decision tree, the maximum depth of a single decision tree, the minimum number of data points allowed in a leaf node, and the maximum number of nodes. The number of predictors for each of the n single decision trees is made by the random choice of k predictors from all available n predictors [4,41]. The overfitting problem, in this case, is usually related to redundant decision trees in the random forest.

Gradient-Boosted Trees for Regression. Gradient boosting refers to an ensemble method that can combine several weak learners into a strong learner [4]. GBT ensures the minimization of variance and bias in relation to single prognostic models. On the other hand, the algorithm is more susceptible to outliers than, for example, simple decision tree models. The GBT algorithm sequentially adds predictors (the same type of models) to the ensemble, each one correcting its own predecessor. This technique tries to fit the new predictor into the residual errors made by the previous predictor. The final prediction consists of the average value from all n single decision trees. In comparison with random forest, GBT has one additional hyperparameter—the learning rate, which is used for scaling the contribution of each tree [41,52]. The problem of overfitting is most often associated with too many trees in the ensemble.

Weighted Averaging Ensemble. This is an integration of the results of selected predictors into the final verdict of the ensemble. The final forecast is defined as the average of the results produced by all n predictors organized in an ensemble [42,46]. The final prediction result is calculated by Equation (3).

$$\hat{y}_i = \frac{1}{n} \sum_{j=1}^n \hat{y}_i^j \quad (3)$$

where i is the prediction point, \hat{y}_i is the final predicted value, \hat{y}_i^j is the value predicted by predictor number j , and n is the number of predictors in the ensemble. Note: all weights are equal to $1/n$ in this case.

This formula makes use of the stochastic distribution of the predictive errors. The process of averaging reduces the final error of forecasting. The averaging of the forecast results is an established method of reducing the variance of forecast errors. An important condition for including the predictor in the ensemble is independent operation from the others and a similar level of prediction error [42,46]. The choice of predictors (forecasting methods) is based on the smallest RMSE error on the validation subset, and only predictors of different types are selected for the ensemble.

Hybrid method—connection of three MLP models. As an element of the prognostic problem decomposition, separate forecasts of selected exogenous variables for the forecast of the power-generation period can be made. This procedure creates new input explanatory exogenous variables (forecasts) that may be valuable for power generation in PV system forecasting methods. In the first step, MLP no. 1 forecasts power generation in period t . On the other hand, MLP no. 2 forecasts solar irradiance in period t . In the second step, neural network MLP no. 3 forecasts the final value of power generation in period t based

on forecasts from the neural network of MLP no. 1 and no. 2 and other endogenous and exogenous variables (4 or 13 depending on the variant). For each of the three MLP neural networks, their appropriate hyperparameters are selected (the number of neurons in the hidden layer and activation functions in the hidden layer as well as in the output layer). Figure 9 shows a general diagram of the developed, proprietary hybrid method.

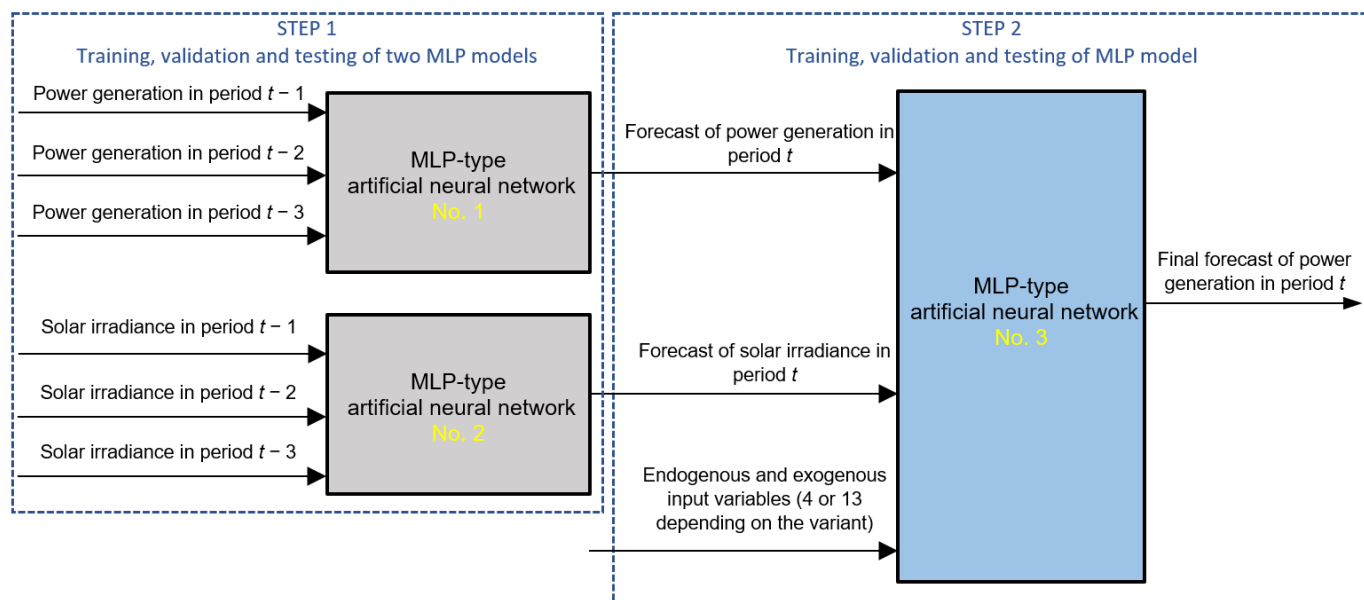


Figure 9. General scheme of the developed hybrid method with the use of three MLP neural network models.

Table 4 shows tested input datasets for each method and the codes of the methods. One reason for organizing data into such sets was to verify the influence of the type and number of variables on the forecast accuracy.

Table 4. Tested input datasets for each method and the codes of the methods.

The Name of Method	The Method Code	Complexity of Method/Type	Tested Sets of Input Data
Persistence model	NAIVE	Single/linear	Set 0
Multiple linear regression model	LR	Single/linear	SET I, SET II A, SET II B, SET III, SET V
K-Nearest Neighbors Regression	KNNR	Single/non-linear	SET I, SET II A, SET II B, SET III, SET V
MLP-type artificial neural network	MLP	Single/non-linear	SET I, SET II A, SET II B, SET III, SET V
Support Vector Regression	SVR	Single/non-linear	SET I, SET II A, SET II B, SET III, SET V
Interval Type-2 Fuzzy Logic System	IT2FLS	Single/non-linear	SET I, SET II A, SET II B
Random forest regression	RF	Ensemble/non-linear	SET I, SET II A, SET II B, SET III, SET V
Gradient-Boosted Trees for regression	GBT	Ensemble/non-linear	SET I, SET II A, SET II B, SET III, SET V
Weighted Averaging Ensemble	WAE (p_1^*, \dots, p_m)	Ensemble/non-linear	SET I, SET IIB, SET III
Hybrid method—connection of three MLP models	MLP&MLP→MLP	Hybrid/non-linear	SET II C, SET IV

Remark: * denotes first predictor in ensemble of m predictors.

5. Evaluation Criteria

In order to have a broader view of the quality of individual forecasting models, four evaluation criteria were used, including RMSE, nMAPE, nAPEmax, and MBE. The RMSE error was adopted as the most important measure due to the greater sensitivity to large partial errors. In all three tables (presented later) with performance measures of proposed methods, the results are sorted by this error measure. On the other hand, the second measure in the order of importance is the nMAPE error. The nAPEmax and MBE measures, in turn, are only auxiliary.

The Root Mean Square Error is calculated by Equation (4). The RMSE measure is typically used for power-generation forecasts from RES, including PV systems.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4)$$

where \hat{y}_i is the predicted value, y_i is the actual value, and n is the number of prediction points.

The Normalized Mean Absolute Percentage Error is determined by Equation (5). Due to the zero values occurring in the power-generation time series, it is impossible to use the popular and recommended measure of the MAPE error. Therefore, the nMAPE measure was used, in which the real power-generation value presented in the denominator of the MAPE formula was replaced with the value representing the normalizing factor (the installed power of PV system).

$$\text{nMAPE} = \frac{1}{n} \sum_{i=1}^n \frac{1}{c_{norm}} |y_i - \hat{y}_i| \cdot 100\% \quad (5)$$

where c_{norm} is the normalizing factor (installed power).

The Normalized Maximum Absolute Percentage Error is calculated by Equation (6). The nAPEmax error is the largest partial error of all individual n nAPE errors.

$$\text{nAPEmax} = \max_{i=1, \dots, n} \frac{1}{c_{norm}} |y_i - \hat{y}_i| \cdot 100\% \quad (6)$$

The Mean Bias Error (MBE) captures the average bias in the prediction and is defined by Equation (7). The forecasting method underestimates values if the nMBE < 0 or overestimates values if the nMBE > 0 . The MBE error of a properly functioning prognostic method should be equal to or very close to zero.

$$\text{MBE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \quad (7)$$

6. Results and Discussion

This section presents a wide comparative analysis of very-short-term forecasting methods for power generation in PV systems.

Table 5 shows performance measures of the proposed methods (on test subset) using three sets of input data—SET I (three inputs). This is the most basic set of input data using only the last three retracted values of the forecast time series of power generation. The study was completed to verify the quality of the forecasts; in this case, it was worse compared to forecasting methods that also use exogenous input variables with a similar amount of input data. Furthermore, the Table 5 shows forecast errors for the simplest reference method—persistence methods (NAIVE), using only one set of input data. Tabular results are ordered by ascending RMSE error values. Table A1 in Appendix A shows the results of hyperparameter tuning for the proposed methods using only three sets of input data.

Table 5. Performance measures of proposed methods (on test subset) using three input data.

Method Code	Input Data Set	RMSE (W)	nMAPE (%)	nAPEmax (%)	MBE (W)
MLP	SET I (3 inputs)	122.558	1.474	32.781	−4.539
WAE [MLP,RF]	SET I (3 inputs)	129.491	1.527	30.847	−2.623
RF	SET I (3 inputs)	133.931	1.674	28.439	−6.832
IT2FLS	SET I (3 inputs)	135.965	1.773	29.808	−6.536
KNNR	SET I (3 inputs)	137.828	1.533	33.802	−5.291
LR	SET I (3 inputs)	140.989	1.617	34.214	9.711
SVR	SET I (3 inputs)	142.441	1.481	34.426	−3.364
GBT	SET I (3 inputs)	154.257	1.948	29.019	6.427
NAIVE *	SET 0 (1 input)	165.783	1.975	40.199	−9.030

Remarks: The best fitting result for each fitting measure is printed in bold in blue. The worst fitting result is printed in red. * Reference model.

Based on the results from Table 5, the following preliminary conclusions can be drawn regarding the proposed methods using only three sets of input data:

- The smallest RMSE and nMAPE errors were obtained by the MLP method among the seven tested methods (including two group methods), and this method can be considered the preferred one. The RF and IT2FLS methods obtained a slightly higher RMSE error;
- The difference in the quality of forecasts between the best MLP method and the worst NAIVE is quite large;
- The SVR method obtained an RMSE error significantly higher than the best MLP method (according to the RMSE measure), while the nMAPE error was almost identical to the MLP method.

Table 6 shows performance measures of the proposed methods (on test subsets) using four sets of input data (SET II A (four inputs), SET II B (four inputs), and SET II C (three, three, and four inputs)). In this case, the amount of input data is limited to only the most relevant input data, both endogenous and exogenous. The study was completed to verify the quality of the forecasts; in this case, it was worse compared to the prognostic methods using all available statistically significant endogenous and exogenous input variables. Another goal of this research was to verify which of the two sets of input data (SET IIA, SET IIB) obtains smaller forecasting errors using different forecasting methods. In addition, the quality of the proprietary hybrid model was verified in relation to other forecasting methods. Furthermore, the table shows forecast errors for the simplest reference method—the persistence method (NAIVE), using only one set of input data. Tabular results are ordered by ascending RMSE error values.

Based on the results from Table 6, the following preliminary conclusions can be drawn regarding the proposed methods using only four different sets of input data (including exogenous variables):

- The use of exogenous variables for forecasts made it possible to reduce the RMSE error of all the methods used;
- The use of smoothed power generation in period $t-1$ (in SET 2B) as an input variable instead of power generation in period $t-1$ (in SET 2A) turned out to be beneficial—all tested methods obtained a lower RMSE error;
- The smallest RMSE error and nMAPE error were obtained by the original, proprietary hybrid method (MLP&MLP->MLP). On the other hand, the MLP method obtained RMSE and nMAPE errors that were slightly higher;
- The largest RSME error, significantly greater than other methods, was obtained by the reference method—the NAIVE method, while the GBT method was the method with the second-greatest RSME error;

- The SVR method using four sets of input data (including exogenous variables) significantly reduced the RMSE error compared to the forecasts using three sets of input data (only the last two withdrawn values of the forecast process)—see Table 5.

Table 6. Performance measures of proposed methods (on test subset) using four input data.

Method Code	Input Data Set	RMSE (W)	nMAPE (%)	nAPEmax (%)	MBE (W)
MLP&MLP→MLP	SET II C (3, 3, 4 inputs)	98.113	1.336	19.110	−0.192
MLP	SET 2B (4 inputs)	99.761	1.393	18.100	−6.335
WAE[SVR,MLP]	SET 2B (4 inputs)	101.825	1.423	17.743	−0.098
SVR	SET 2B (4 inputs)	104.942	1.506	16.329	−0.638
WAE[IT2FLS,MLP]	SET 2B (4 inputs)	105.274	1.535	18.784	−0.116
IT2FLS	SET 2B (4 inputs)	110.461	1.540	19.223	4.956
KNNR	SET 2B (4 inputs)	116.924	1.837	19.328	6.615
MLP	SET 2A (4 inputs)	116.949	1.715	26.040	−2.637
RF	SET 2B (4 inputs)	117.383	1.632	23.882	−0.142
KNNR	SET 2A (4 inputs)	118.985	1.506	27.016	4.145
IT2FLS	SET 2A (4 inputs)	121.590	1.661	27.173	7.511
LR	SET 2B (4 inputs)	123.803	1.961	19.513	−8.993
SVR	SET 2A (4 inputs)	124.312	1.695	30.977	6.775
GBT	SET 2B (4 inputs)	124.747	1.914	17.422	−9.287
RF	SET 2A (4 inputs)	129.469	1.676	28.658	11.728
LR	SET 2A (4 inputs)	132.943	2.076	27.148	−2.051
GBT	SET 2A (4 inputs)	134.457	1.914	28.318	−4.637
NAIVE *	SET 0 (1 input)	165.783	1.975	40.199	−9.030

Remarks: The best fitting result for each fitting measure is printed in bold in blue. The worst fitting result is printed in red. * Reference model.

Table 7 shows, in turn, performance measures of the proposed methods (on test subsets) using 11, 13, and 15 sets of input data. This study aimed to verify whether the use of as many available and statistically significant endogenous and exogenous input variables would improve the quality of forecasts compared to a limited number of input data (three or four sets). In addition, the quality of the proposed proprietary hybrid model and the original “Weighted Averaging Ensemble” models compared to other forecast methods was verified. Furthermore, the Table 7 shows forecast errors for the simplest reference method—the persistence method (NAIVE), using only one set of input data. Tabular results are ordered by ascending RMSE error values.

Table 7. Performance measures of proposed methods (on test subset) using 1, 11, 13, and 15 sets of input data.

Method Code	Input Data Set	RMSE (W)	nMAPE (%)	nAPEmax (%)	MBE (W)
MLP&MLP→MLP	SET IV (3,3,13 inputs)	61.633	0.805	13.918	1.196
MLP	SET V (15 inputs)	63.092	0.848	12.375	−3.498
MLP	SET III (11 inputs)	64.794	0.809	16.173	3.664
WAE (SVR,MLP)	SET V (15 inputs)	65.391	0.832	12.397	−0.053
SVR	SET V (15 inputs)	71.618	0.824	12.629	−6.088
WAE (LR,MLP)	SET III (11 inputs)	71.884	0.826	15.038	−0.048
SVR	SET III (11 inputs)	90.379	1.416	16.065	3.570
LR	SET V (15 inputs)	90.491	1.015	14.982	−0.049
LR	SET III (11 inputs)	91.674	1.030	21.215	−3.760
KNNR	SET V (15 inputs)	104.505	1.360	18.228	−4.819
RF	SET V (15 inputs)	111.269	1.577	22.362	−2.483
RF	SET III (11 inputs)	116.497	1.619	23.848	3.406
KNNR	SET III (11 inputs)	118.490	1.565	23.348	7.507
GBT	SET V (15 inputs)	118.547	1.523	20.386	−0.179
GBT	SET III (11 inputs)	122.569	1.587	25.624	−2.750
NAIVE *	SET 0 (1 input)	165.783	1.975	40.199	−9.030

Remarks: The best fitting result for each fitting measure is printed in bold in blue. The worst fitting result is printed in red. * Reference model.

Based on the results from Table 7, the following preliminary conclusions can be drawn regarding the proposed methods using different numbers of sets of input data ranging from 11 to 15 (including exogenous variables):

- The use of a larger number of input data, from 11 to 15 for forecasts (including exogenous variables), allowed for a significant reduction in the RMSE error of all methods used compared to the use of only four sets of input data;
- The smallest RMSE error and nMAPE error were obtained by the original, proprietary hybrid method (MLP&MLP->MLP), and it is the recommended method. On the other hand, the MLP method obtained RMSE and nMAPE errors that were slightly higher;
- The largest RSME error, significantly greater than other methods, was obtained by the reference method—the NAIVE method, while the GBT method was the method with the second-greatest RSME error;
- The SVR method using 15 sets of input data (including exogenous variables) was one of the best methods, but the use of 11 sets of input data proved to be less favorable;
- Team methods with different types of predictors in the team (WAE (SVR,MLP) and WAE (LR,MLP)) were also among the best methods—the RMSE error of these methods was slightly greater than the second MLP method in the list;
- For all tested methods, it was more advantageous to use 15 sets of input data than 11 sets of input data.

Figure 10 shows the RMSE error, for each of the eight tested datasets, obtained by the best prognostic method for the test range. The MLP neural network method (yellow) is definitely the most common method for various input datasets. On the other hand, the smallest RMSE error (green) was obtained by the proprietary developed hybrid model (MLP&MLP->MLP). The highest RMSE error (gray) was achieved by the persistence (naïve) model as the simplest one, using only one set of input data. It should be noted that the quality of forecasts increases significantly with the increasing number of input data used. Thus, it can be concluded that by providing the predictive model at the input with more information related to the predicted process, in particular, with more than just one retracted value of a given explanatory variable (both exogenous and endogenous), smaller forecast errors can be expected.

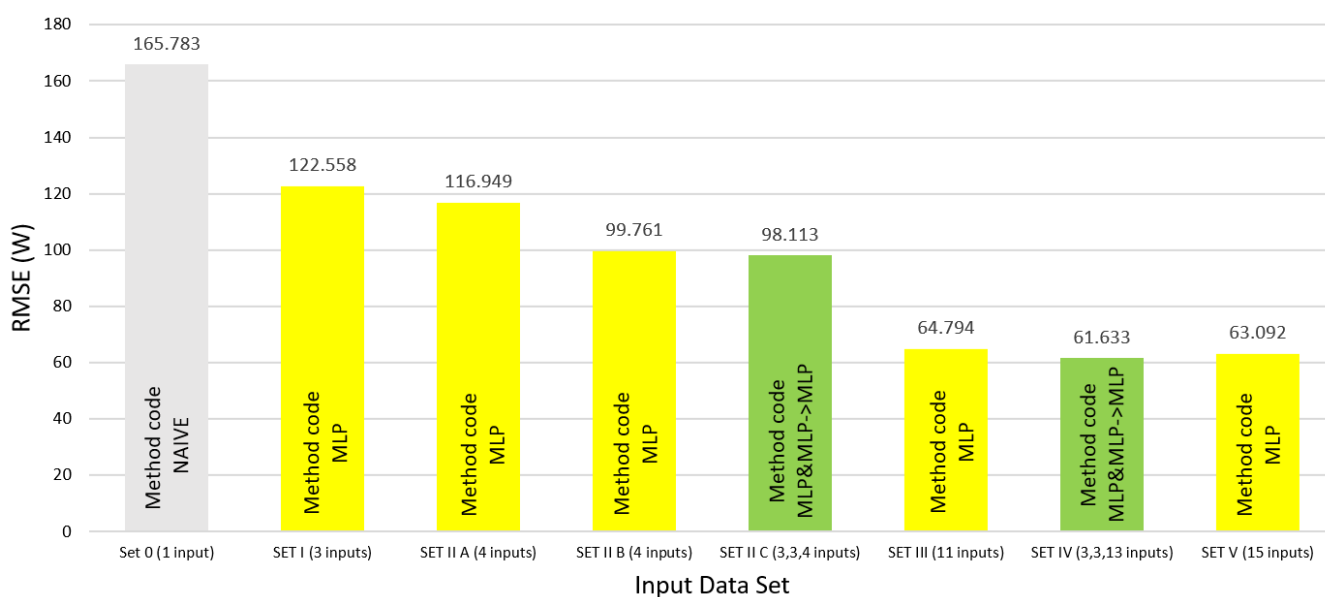


Figure 10. Summary of RMSE error values for the best predictive method depending on the dataset.

Figure 11 shows a scatter plot between the actual power-generation values and the values obtained from the forecast using the best method—a proprietary hybrid model (MLP&MLP->MLP) for the test range. From the graph, it can be observed that the accuracy

of forecasts was the highest for small power-generation values below 750 W (where the installed power of a PV system is equal to 3200 W).

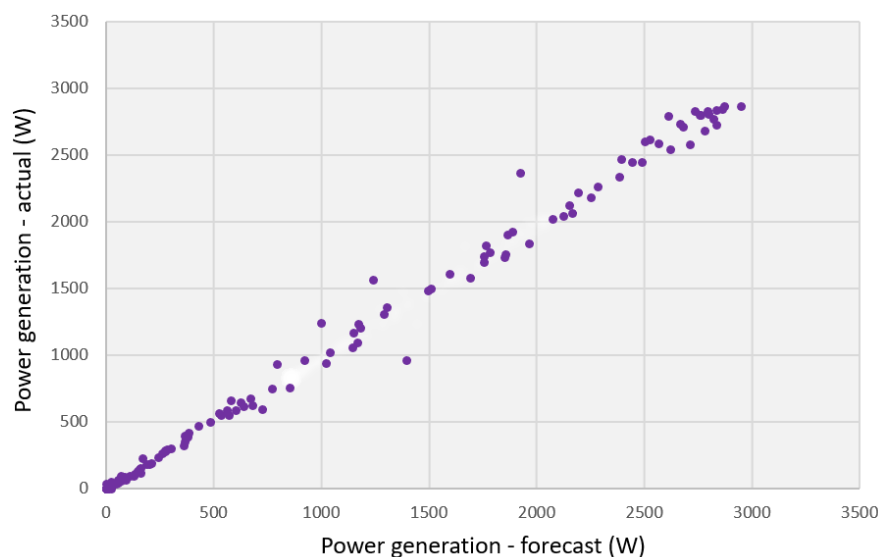


Figure 11. The scatter plot of the real power-generation values and the values obtained from the forecast with the best hybrid model.

7. Conclusions

The analysis of the available input variables with the use of four different methods of selecting input variables for forecasting models allowed us to identify the most important input variables. The most important input data include smoothed power generation in period $t-1$, power generation in period $t-1$, and solar irradiance in period $t-1$. The significantly least-important input data are wind direction in period $t-1$ and wind speed in period $t-1$.

The influence of the type and number of input variables on the quality of forecasts was investigated. The use of only three withdrawn values of power generation showed that this is the least-effective solution. Additionally, the use of other available exogenous variables (the selected historical values of solar irradiance, PV module temperature, wind direction, and wind speed) allowed us to reduce the RSME error of forecasts. An additionally valuable input variable is the smoothed value of power generation (see Equation (1)), a value calculated on the basis of the reverted values of the forecast process. The smallest forecast errors (RMSE) were obtained using a set of SET IV and SET V input variables, i.e., sets with the largest number of input variables.

The effectiveness of many prognostic methods, both single as well as team and hybrid, was verified. The smallest RMSE and nMAPE errors were obtained for the original, developed hybrid method using three MLP neural networks (method code MLP & MLP-MLP) using a set of SET IV input variables. Compared to the reference method (method code NAIVE), the hybrid method obtained an RMSE error 62.8% lower. However, compared to the best single method (the MLP method code) using the SET V input variable set, the RMSE error of the hybrid method was 2.3% lower. In the case of the number of input variables limited to four, the proprietary hybrid method also obtained the smallest RMSE error. Compared to the method code MLP, the RMSE error for the hybrid method was 1.7% lower. Among the single prognostic methods, the MLP neural network was the best method. Other machine learning techniques (RF, SVR, KNNR, and GBT) obtained slightly larger RMSE errors. The most advantageous of these four machine learning techniques was the SVR method with the SET V set of input variables. It is also advantageous to use the collective method (method code WAE (SVR, MLP)), which obtained an RMSE error slightly greater than the best method single (MLP).

In the authors' opinion, some of the forecasting methods investigated are effective and promising tools for practical applications, e.g., for very-short-term PV generation power forecasting. In turn, forecasts of this type are very useful for the needs of low-voltage microgrid operation control.

Research may be continued and expanded in the future. The proposed research directions include:

- Increasing the forecast horizon to 1 h (4 forecasts for consecutive 15 min periods);
- Using various techniques for decomposing the prognostic problem and examining their impact on the quality of forecasts (in the case of obtaining data from a period of several years);
- Examining the distribution of forecast errors during the day—verifying whether there is a relationship between the RMSE error rate and the time of day;
- Quality-testing forecasting models using additional solar irradiance, wind speed, and wind direction forecasts (in the case of obtaining such meteorological forecasts).

Author Contributions: Conceptualization, P.P. and M.P.; formal analysis, P.P.; methodology, P.P. and P.K.; investigation, P.P., P.K., M.P., and B.F.; supervision, M.P.; validation, P.P.; writing, P.P., P.K., M.P., and B.F.; visualization P.P.; project administration, P.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The dataset deployed for calculations in this paper was gathered by the meteorology station of the Photovoltaic Laboratory at the Institute of Microelectronics and Optoelectronics of Warsaw University of Technology. The authors thank this institute heartily for sharing these data.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACF	Autocorrelation Function
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
BFGS	Broyden–Fletcher–Goldfarb–Shanno algorithm
C&RT	Classification and Regression Trees algorithm
DCs	Decision Trees
EIASC	Enhanced iterative algorithm with stop condition
GA	Genetic algorithm
GBT	Gradient-Boosted Trees
GSA	Global sensitivity analysis
IASC	Iterative algorithm with stop condition
IT2FLS	Interval Type-2 Fuzzy Logic System
KM	Karnik–Mendel
KNNR	K-Nearest Neighbors Regression
LR	Linear Regression
MBE	Mean Bias Error
MLP	Multi-Layer Perceptron
nAPE _{max}	Normalized Maximum Absolute Percentage Error
nMAPE	Normalized Mean Absolute Percentage Error
PR	Performance ratio
PSO	Particle Swarm Optimization

PV	Photovoltaic
R	Pearson linear correlation coefficient
RES	Renewable Energy Sources
R ²	Determination coefficient
RF	Random forest
RMSE	Root Mean Square Error
SVM	Support Vector Machine
SVR	Support Vector Regression

Appendix A

Table A1 shows the results of hyperparameter tuning for the proposed methods using three sets of input data.

Table A1. Results of hyperparameter tuning for proposed methods using three sets of input data.

Method Code	Description of Method, Name, and Range of Values of Hyperparameters' Tuning and Selected Values
SVR	Regression SVM: Type-1, Type 2, selected: Type-1; kernel type: Gaussian (RBF); width parameter σ : 0.333; regularization constant C , range: 1–50 (step 1), selected: 2; tolerance ϵ , range: 0.01–0.2 (step 0.01), selected: 0.02.
KNNR	Number of nearest neighbours k , Distance metrics: Euclidean, Manhattan, Minkowski, selected: Euclidean; range: 1–50, selected: 13.
MLP	Learning algorithm: BFGS; the number of neurons in hidden layer: 2–10, selected: 3; activation function in hidden layer: linear, hyperbolic tangent, selected: hyperbolic tangent; activation function in output layer: linear.
IT2FLS	Interval Type-2: Sugeno FLS, Mamdani FLS, selected: Sugeno FLS; learning and tuning algorithm: GA, PSO, selected: PSO; initial swarm span: 1500–2500, selected: 2000; minimum neighborhood size: 0.20–0.30, selected: 0.25; inertia range: from [0.10–1.10] to [0.20–2.20], selected: [0.50–0.50]; number of iterations in the learning and tuning process: 5–20, selected: 20; type of the membership functions: triangular, Gauss, selected: Gauss; the number of output membership functions: 3–81, selected: 81; defuzzification method: Centroid, Weighted average of all rule outputs, selected: Weighted average of all rule outputs; AND operator type: min, prod, selected: min; OR operator type: max, probor, selected: probor; implication type: prod, min, selected: min; aggregation type: sum, max, selected: sum; the k-Fold Cross-Validation value: 1–4, selected: 4; window size for computing average validation cost: 5–10, selected: 7; maximum allowable increase in validation cost: 0.0–1.0, selected: 0.1; the type-reduction methods: KM, IASC, EIASC, selected: KM.
RF	The number of decision trees: 2–50, selected: 5; the number of predictors chosen at random: 1, 2, selected 2. Stop parameters: maximum number of levels in each decision tree: 5, 10, 20, selected 10; minimum number of data points placed in a node before the node is split: 10, 20, 30, 40, 50, selected 20; minimum number of data points allowed in a leaf node: 10; maximum number of nodes: 100.
GBT	Considered max depth: 2/4, selected depth: 2; trees number: 50/100/150/200/250, selected number: 100; learning rate: 0.1/0.01/0.001, selected: 0.1.

References

- Shamsollahi, P.; Cheung, K.W.; Chen, Q.; Germain, E.H. A neural network based very short term load forecaster for the interim ISO New England electricity market system. In Proceedings of the 2001 Power Industry Computer Applications Conference, Sydney, Australia, 20–24 May 2001; pp. 217–222.
- Parol, M.; Piotrowski, P. Very short-term load forecasting for optimum control in microgrids. In Proceedings of the 2nd International Youth Conference on Energetics (IYCE 2009), Budapest, Hungary, 4–6 June 2009; pp. 1–6.
- Parol, M.; Piotrowski, P. Electrical energy demand forecasting for 15 minutes forward for needs of control in low voltage electrical networks with installed sources of distributed generation. *Przegląd Elektrotechniczny Electr. Rev.* **2010**, *86*, 303–309. (In Polish)
- Parol, M.; Piotrowski, P.; Kapler, P.; Piotrowski, M. Forecasting of 10-Second Power Demand of Highly Variable Loads for Microgrid Operation Control. *Energies* **2021**, *14*, 1290. [[CrossRef](#)]

5. Hernandez, L.; Baladron, C.; Aguiar, J.M.; Carro, B.; Sanchez-Esguevillas, A.J.; Lloret, J.; Massana, J. A Survey on Electric Power Demand Forecasting: Future Trends in Smart Grids, Microgrids and Smart Buildings. *IEEE Commun. Surv. Tutor.* **2014**, *16*, 1460–1495. [[CrossRef](#)]
6. Niu, D.; Pu, D.; Dai, S. Ultra-Short-Term Wind-Power Forecasting Based on the Weighted Random Forest Optimized by the Niche Immune Lion Algorithm. *Energies* **2018**, *11*, 1098. [[CrossRef](#)]
7. Bokde, N.; Feijóo, A.; Villanueva, D.; Kulat, K. A Review on Hybrid Empirical Mode Decomposition Models for Wind Speed and Wind Power Prediction. *Energies* **2019**, *12*, 254. [[CrossRef](#)]
8. Adnan, R.M.; Liang, Z.; Yuan, X.; Kisi, O.; Akhlaq, M.; Li, B. Comparison of LSSVR, M5RT, NF-GP, and NF-SC Models for Predictions of Hourly Wind Speed and Wind Power Based on Cross-Validation. *Energies* **2019**, *12*, 329. [[CrossRef](#)]
9. Würth, I.; Valldecabres, L.; Simon, E.; Möhrlein, C.; Uzunoğlu, B.; Gilbert, C.; Giebel, G.; Schlipf, D.; Kaifel, A. Minute-Scale Forecasting of Wind Power—Results from the Collaborative Workshop of IEA Wind Task 32 and 36. *Energies* **2019**, *12*, 712. [[CrossRef](#)]
10. Liu, F.; Li, R.; Dreglea, A. Wind Speed and Power Ultra Short-Term Robust Forecasting Based on Takagi–Sugeno Fuzzy Model. *Energies* **2019**, *12*, 3551. [[CrossRef](#)]
11. Carpinone, A.; Giorgio, M.; Langella, R.; Testa, A. Markov chain modeling for very-short-term wind power forecasting. *Electr. Power Syst. Res.* **2015**, *122*, 152–158. [[CrossRef](#)]
12. Tato, J.H.; Brito, M.C. Using Smart Persistence and Random Forests to Predict Photovoltaic Energy Production. *Energies* **2018**, *12*, 100. [[CrossRef](#)]
13. Zhu, R.; Guo, W.; Gong, X. Short-Term Photovoltaic Power Output Prediction Based on k-Fold Cross-Validation and an Ensemble Model. *Energies* **2019**, *12*, 1220. [[CrossRef](#)]
14. Abdullah, N.A.; Rahim, N.A.; Gan, C.K.; Adzman, N.N. Forecasting Solar Power Using Hybrid Firefly and Particle Swarm Optimization (HFPSO) for Optimizing the Parameters in a Wavelet Transform-Adaptive Neuro Fuzzy Inference System (WT-ANFIS). *Appl. Sci.* **2019**, *9*, 3214. [[CrossRef](#)]
15. Nespoli, A.; Mussetta, M.; Ogliari, E.; Leva, S.; Fernández-Ramírez, L.; García-Triviño, P. Robust 24 Hours ahead Forecast in a Microgrid: A Real Case Study. *Electronics* **2019**, *8*, 1434. [[CrossRef](#)]
16. Ahmed, R.; Sreeram, V.; Mishra, Y.; Arif, M.D. A review and evaluation of the state-of-art in PV solar power forecasting: Techniques and optimization. *Renew. Sust. Energ. Rev.* **2020**, *124*, 109792. [[CrossRef](#)]
17. Antonanzas, J.; Osorio, N.; Escobar, R.; Urraca, R.; Martinez-de-Pison, F.J.; Antonanzas-Torres, F. Review of photovoltaic power forecasting. *Sol. Energy* **2016**, *136*, 78–111. [[CrossRef](#)]
18. Mayer, M.J.; Gróf, G. Extensive comparison of physical models for photovoltaic power forecasting. *Appl. Energy* **2021**, *283*, 116239. [[CrossRef](#)]
19. Mayer, M.J. Influence of design data availability on the accuracy of physical photovoltaic power forecasts. *Sol. Energy* **2021**, *227*, 532–540. [[CrossRef](#)]
20. Study Committee: C6; CIGRÉ Working Group C6.22. *Microgrids 1: Engineering, Economics, & Experience-Capabilities, Benefits, Business Opportunities, and Examples-Microgrids Evolution Roadmap*; Technical Brochure 635; CIGRE: Paris, France, 2015.
21. Marnay, C.; Chatzivasileiadis, S.; Abbey, C.; Irvani, R.; Joos, G.; Lombardi, P.; Mancarella, P.; von Appen, J. Microgrid evolution roadmap. Engineering, economics, and experience. In Proceedings of the 2015 International Symposium on Smart Electric Distribution Systems and Technologies (EDST15), CIGRE SC C6 Colloquium, Vienna, Austria, 8–11 September 2015.
22. *Microgrids: Architectures and Control*; Hatziaargyriou, N.D. (Ed.) Wiley-IEEE Press: Hoboken, NJ, USA, 2014.
23. Baczynski, D.; Ksiezzyk, K.; Parol, M.; Piotrowski, P.; Wasilewski, J.; Wojtowicz, T. *Low Voltage Microgrids. Joint Publication Edited by Mirosław Parol*; Publishing House of the Warsaw University of Technology: Warsaw, Poland, 2013. (In Polish)
24. Li, Y.; Nejabatkhah, F. Overview of control, integration and energy management of microgrids. *J. Mod. Power Syst. Clean Energy* **2014**, *2*, 212–222. [[CrossRef](#)]
25. Olivares, D.E.; Canizares, C.A.; Kazerani, M. A Centralized Energy Management System for Isolated Microgrids. *IEEE Trans. Smart Grid* **2014**, *5*, 1864–1875. [[CrossRef](#)]
26. Morstyn, T.; Hredzak, B.; Agelidis, V.G. Control strategies for microgrids with distributed energy storage systems: An overview. *IEEE Trans. Smart Grid* **2018**, *9*, 3652–3666. [[CrossRef](#)]
27. Lopes, J.A.P.; Moreira, C.L.; Madureira, A.G. Defining Control Strategies for MicroGrids Islanded Operation. *IEEE Trans. Power Syst.* **2006**, *21*, 916–924. [[CrossRef](#)]
28. Parol, M.; Rokicki, Ł.; Parol, R. Towards optimal operation control in rural low voltage microgrids. *Bull. Pol. Ac. Tech.* **2019**, *67*, 799–812.
29. Parol, M.; Kapler, P.; Marzecki, J.; Parol, R.; Potecki, M.; Rokicki, Ł. Effective approach to distributed optimal operation control in rural low voltage microgrids. *Bull. Pol. Ac. Tech.* **2020**, *68*, 661–678.
30. Zakir, M.; Sher, H.A.; Arshad, A.; Lehtonen, M. A fault detection, localization, and categorization method for PV fed DC-microgrid with power-sharing management among the nano-grids. *Int. J. Electr. Power Energy Syst.* **2022**, *137*, 107858. [[CrossRef](#)]
31. Sharma, V.; Chandel, S.S. Performance and degradation analysis for long term reliability of solar photovoltaic systems: A review. *Renew. Sust. Energ. Rev.* **2013**, *27*, 753–767. [[CrossRef](#)]
32. Makrides, G.; Norton, M.; Georghiou, G.E. Performance of Photovoltaics Under Actual Operating Conditions. *Third Gener. Photovolt.* **2012**, 201–232. [[CrossRef](#)]

33. Dierauf, T.; Growitz, A.; Kurtz, S.; Cruz, J.L.B.; Riley, E.; Hansen, C. *Weather-Corrected Performance Ratio*; Technical Report; NREL/TP-5200-57991; National Renewable Energy Lab.(NREL): Golden, CO, USA, April 2013.
34. Nordmann, T.; Clavadetscher, L. Understanding temperature effects on PV system performance. In Proceedings of the 3rd IEEE World Conference on Photovoltaic Energy Conversion, Osaka, Japan, 11–18 May 2003; pp. 2243–2246.
35. Virtuani, A.; Pavanello, D.; Friesen, G. Overview of Temperature Coefficients of Different Thin Film Photovoltaic Technologies. In Proceedings of the 25th EUPVSEC Conference, Valencia, Spain, 6–10 September 2010. [[CrossRef](#)]
36. Skoplaki, E.; Palyvos, J.A. On the temperature dependence of photovoltaic module electrical performance: A review of efficiency/power correlations. *Sol. Energy* **2009**, *83*, 614–624. [[CrossRef](#)]
37. Makrides, G.; Theristis, M.; Bratcher, J.; Pratt, J.; Georghiou, G.E. Five-year performance and reliability analysis of monocrystalline photovoltaic modules with different backsheets materials. *Sol. Energy* **2018**, *171*, 491–499. [[CrossRef](#)]
38. Taylor, N. *Traceable Performance Measurements of PV Devices*; European Union: Luxembourg, 2010. [[CrossRef](#)]
39. Polo, J.; Alonso-Abella, M.; Ruiz-Arias, J.A.; Balenzategui, J.L. Worldwide analysis of spectral factors for seven photovoltaic technologies. *Sol. Energy* **2017**, *142*, 194–203. [[CrossRef](#)]
40. Rahmann, C.; Vittal, V.; Ascui, J.; Haas, J. Mitigation Control Against Partial Shading Effects in Large-Scale PV Power Plants. *IEEE Trans. Sustain. Energy* **2016**, *7*, 173–180. [[CrossRef](#)]
41. Géron, A. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 2nd ed.; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2019.
42. Piotrowski, P.; Baczyński, D.; Kopyt, M.; Gulczyński, T. Advanced Ensemble Methods Using Machine Learning and Deep Learning for One-Day-Ahead Forecasts of Electric Energy Production in Wind Farms. *Energies* **2022**, *15*, 1252. [[CrossRef](#)]
43. Dudek, G.; Pełka, P. Forecasting monthly electricity demand using k nearest neighbor method. *Przegląd Elektrotechniczny Electr. Rev.* **2017**, *93*, 62–65.
44. Piotrowski, P.; Baczyński, D.; Kopyt, M.; Szafranek, K.; Helt, P.; Gulczyński, T. Analysis of forecasted meteorological data (NWP) for efficient spatial forecasting of wind power generation. *Electr. Power Syst. Res.* **2019**, *175*, 105891. [[CrossRef](#)]
45. Dudek, G. Multilayer perceptron for short-term load forecasting: From global to local approach. *Neural Comput. Appl.* **2019**, *32*, 3695–3707. [[CrossRef](#)]
46. Osowski, S.; Siwek, K. Local dynamic integration of ensemble in prediction of time series. *Bull. Pol. Ac. Tech* **2019**, *67*, 517–525.
47. Karnik, N.; Mendel, J.; Liang, Q. Type-2 fuzzy logic systems. *IEEE Trans. Fuzzy Syst.* **1999**, *7*, 643–658. [[CrossRef](#)]
48. Karnik, N.N.; Mendel, J.M. Centroid of a type-2 fuzzy set. *Inf. Sci.* **2001**, *132*, 195–220. [[CrossRef](#)]
49. Mendel, J.; John, R. Type-2 fuzzy sets made simple. *IEEE Trans. Fuzzy Syst.* **2002**, *10*, 117–127. [[CrossRef](#)]
50. Khosravi, A.; Nahavandi, S.; Creighton, D.; Srinivasan, D. Interval Type-2 Fuzzy Logic Systems for Load Forecasting: A Comparative Study. *IEEE Trans. Power Syst.* **2012**, *27*, 1274–1282. [[CrossRef](#)]
51. Liang, Q.; Mendel, J.M. Interval type-2 fuzzy logic systems: Theory and design. *IEEE Trans. Fuzzy Syst.* **2000**, *8*, 535–550. [[CrossRef](#)]
52. Hastie, T.; Tibshirani, R.; Friedman, J.H. *The elements of Statistical Learning. Data Mining, Inference, and Prediction*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2009.