energies

Article

Time Series Forecasting for Energy Production in Stand-Alone and Tracking Photovoltaic Systems Based on Historical Measurement Data

Mateusz Sumorek and Adam Idzkowski *

Faculty of Electrical Engineering, Bialystok University of Technology, Wiejska 45D, 15-351 Bialystok, Poland
* Correspondence: a.idzkowski@pb.edu.pl

Abstract: This article presents a time series analysis for predicting energy production in photovoltaic (PV) power plant systems, namely fixed and solar-tracking ones, which were located in the north-east of Poland. The purpose of one-day forecasts is to determine the effectiveness of preventive actions and manage power systems effectively. The impact of climate variables affecting the production of electricity in the photovoltaic systems was analyzed. Forecasting models based on traditional machine learning (ML) techniques and multi-layer perceptron (MLP) neural networks were created without using solar irradiance as an input feature to the model. In addition, a few metrics were selected to determine the quality of the forecasts. The preparation of the dataset for constructing the forecasting models was discussed, and some ways for improving the metrics were given. Furthermore, comparative analyses were performed, which showed that the MLP neural networks used in the regression problem provided better results than the MLP classifier models. The Diebold–Mariano (DM) test was applied in this study to distinguish the significant differences in the forecasting accuracy between the individual models. Compared to KNN (k-nearest neighbors) or ARIMA models, the best results were obtained for the simple linear regression, MLPRegressor, and CatBoostRegressor models in each of the investigated photovoltaic systems. The R-squared value for the MLPRegressor model was around 0.6, and it exceeded 0.8 when the dataset was split and separated into months.

Keywords: renewable energy; energy forecasting; machine learning; deep learning; linear regression; neural networks; time series analysis

Citation: Sumorek, M.; Idzkowski, A. Time Series Forecasting for Energy Production in Stand-Alone and Tracking Photovoltaic Systems Based on Historical Measurement Data. *Energies* **2023**, *16*, 6367. https://doi.org/10.3390/en16176367

Academic Editors: Jacek Kusznier, Maciej Zajkowski and Paweł Węgierek

Received: 9 August 2023
Revised: 26 August 2023
Accepted: 31 August 2023
Published: 2 September 2023

1. Introduction

The increasing proportion of renewable energy sources (RES) in energy production creates a risk of temporary blackouts [1] or a decrease in energy quality [2]. In various countries, such blackouts are a common system failure. This is due to the fact that these sources (solar power plants, wind power plants) are dependent on weather conditions, which are stochastic in nature. It is therefore necessary to achieve a balance between energy consumption and production by temporarily increasing production at, for example, conventional or nuclear power plants.

Surpluses and deficits in energy production negatively affect the functioning of power grids. It is essential for the proper functioning of an energy system to adapt energy production to the current demand, which is currently practiced [3]. Although it is possible to adjust energy production in conventional power plants based on demand, it is not possible to adjust it in power plants based on renewable sources. For this purpose, it is necessary to accurately forecast energy production from renewable sources, especially when the share of such power plants is significant in relation to other sources of electricity. The additional knowledge of the production capacity of these energy sources, discussed in minute, hour, or even daily intervals, will allow for even more effective management and protection of power systems [4].

The constant development of the economy, changes in everyday life, social changes, or environmental changes require the modernization of the energy production profile. Increasing the share of renewable energy sources in the production of electricity brings new challenges [5]. Therefore, it is necessary to develop methods for forecasting electricity production by wind or solar power plants, depending on weather conditions [6].

The existing approaches to forecasting models can be classified into the following four categories: physical, statistical, machine learning, and hybrid [7]. Physical models can be employed without historical data sets. They are based on numerical weather predictions (NWP) or sky images. Conventional statistical techniques include autoregressive (AR), moving average (MA), autoregressive integrated moving average (ARIMA), seasonal ARIMA (SARIMA), and other variants of similar models [8,9]. Statistical models are typically linear forecasters, and as such, they are effective in areas where the frequency of data is low, such as those with weekly patterns. For hourly values, the nonlinear behavior of the data might be too difficult to predict. The use of machine learning methods is seen as an alternative to conventional linear forecasting methods. Hybrid models are designed to improve the performance of physical or statistical techniques. Recent studies have shown that the seq2seq architecture used in the transformer model is suitable for modeling complex relationships in sequence data and for multi-step time series forecasting. The transformer model is a powerful tool for multi-step time series forecasting, which is a difficult task for traditional statistical models, such as ARIMA and GARCH [10]. In the article in [11], the prediction performance of FusFormer, a transformer-based model for forecasting time series data, compared to that of the long short-term memory (LSTM) network, light-gradient-boosting machine (LightGBM), residual neural network (ResNet1D), transformer, and conv-transformer models was evaluated.

Forecasting is the process of making predictions based on past and present data. Deterministic and probabilistic forecasting is important in the daily operation of power systems. Given a set of input data, deterministic forecasting models provide a single-valued expectation series of the power output. Probabilistic methods give a wider view of possible power outputs, as the output of such models could be a percentile, a prediction interval, or an entire prediction distribution [12]. Using observation results with actual results, it is possible to determine the forecast quality (how it coincides with the actual state) and improve a developed model [13]. Nowadays, neural network models can be trained using different packages such as Keras, PyTorch, TensorFlow, and Sklearn [14].

Forecasting can be both a simple procedure and a very complex one, especially when there are many variables influencing the prediction model (multivariate data). A time series can be stationary or non-stationary. A trend behavior can be upward or downward, steep or not, and exponential or approximately linear. Many time series include trends, cycles, and seasonality. All of this, as well as data continuity and computational constraints, should be considered when choosing a modeling technique [15].

The progress in the field of machine learning (ML) and deep learning (DL) model development is still needed. Such models need a large training dataset and an optimal training algorithm. This is because it is necessary to specify more unique situations, especially considering that many problems are non-linear and there is a need for online re-training of the model to further improve its accuracy [16,17].

In particular, the contributions of this paper are summarized as follows:

- Neural-network-based energy yield models were developed and compared with other prediction models for two types of installed photovoltaic panels: a fixed system and a solar tracker. These systems were located in a humid continental climate.
- A large dataset was used for the training and test processes. It was shown that tuning the network's hyperparameters had a significant impact on the forecast accuracy and computational time.
- To assess the prediction performance of the regression algorithms (used with continuous data) and the classification algorithms (used with discrete data), quantitative metrics were adopted.

- It was shown that it is not advisable to rely on one metric (R-squared) as a universal indicator of forecast quality.
- This article focuses on forecasting electricity production in small photovoltaic systems. Consequently, an analysis of the climate variables that affect the forecasting of electricity production was also carried out, taking into account the availability of information in historical measurement data.

2. Literature Review

There are numerous publications in the field of forecasting electricity production in photovoltaic systems. Many of them have been created over the past few years, so it is a timely issue. These studies describe ultra-short-, short-, medium-, and long-term forecasts using artificial intelligence methods, most commonly artificial (ANN) or deep neural networks (DNN). Depending on the publication, different types of networks are proposed, such as MLP—multi-layer perceptron [18], MLP ABC—MLP network using the artificial bee colony algorithm [19], RNN—recurrent neural network [20], LSTM network [21], CNN—convolutional neural network [22,23], and hybrid models [16,24]. In addition, other machine learning algorithms are distinguished, such as k-nearest neighbors (KNN), decision trees (DT), LightGBM, CatBoost, and extreme gradient boosting (XGBoost) [25], among others.

There are several methods used for the direct prediction of PV power output that combine clustering, such as the K-means algorithm and prediction techniques. The main idea is to cluster the days based on their weather characteristics and then build separate prediction models for each cluster [26].

Depending on the publication, forecasts are made for different time horizons [27]. The dominant group is ultra-short-term (minute) forecasts; the next group is short-term (hourly) forecasts. Medium- and long-term forecasts are few. The forecasting period usually depends on the availability of information used for forecasting and the purpose of using these forecasts.

Examples of forecast uses include:

- Power system management [28–30];
- Event detection, e.g., covering panels with dust [31] or partial shading [32];
- Increasing the efficiency of photovoltaic systems by optimizing the operation of MPPT (maximum power point tracking) and intelligent-controlling-based MPPT systems [33,34].

The key to the development of solar production forecasts is the selection of the input variables. The literature indicates numerous weather factors that can be used. These include, among others: irradiance, temperature, humidity, atmospheric pressure, wind speed, wind direction, rainfall, dust accumulation, and cloudiness. Figure 1 shows the correlation coefficients among the various variables. There is a very strong correlation, equal to 0.99, between the PV power output and the irradiance. Furthermore, there is a significant correlation between the temperature of the PV module and the output power (0.57) [35]. The PV system in this example operates in Maceió, Brazil, which has a typical tropical climate.

Basing on another publication [4], Table 1 illustrates the correlation among individual quantities and the output power. The PV system in this example is located in the Qassim region of Saudi Arabia, which has a typical desert climate. Again, there is a very strong correlation, close to unity, between the output power and the irradiance. The R-squared coefficient of the module temperature is similar to the previous case and is equal to 0.59. In addition, a significant correlation between the wind speed and power output of 0.45 is demonstrated. In most publications, the correlation values between the irradiance and the output power are the same as the correlation values between the temperature and the output power [36]. Similarly, a negative correlation of humidity is clearly indicated [37].

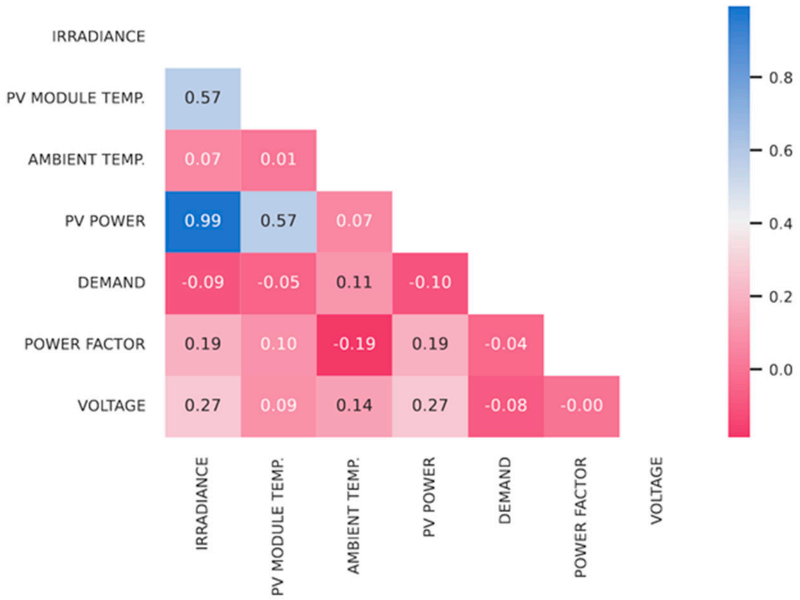

Figure 1. Correlation matrix showing correlation coefficients among sets of variables for a dataset from 2019 to mid-2022 in Maceió (Brazil) [35].

Table 1. PV power output and its correlation with the inputs based on the meteorological data of the Qassim region (KSA) [4].

Inputs	R-Squared Coefficient
Irradiance on module	0.998
Module temperature	0.587
Wind speed	0.447
Relative humidity	−0.362

In many articles, the results of forecasting models for different geographic locations are presented. Forecasts were developed for plants located in Brazil [35], Israel [19], Turkey [38], Italy [19,30], India [31], United States [23], Saudi Arabia [4], Australia [16,24], Scotland [25], and South Korea [39]. Based on the literature, it is clear that publications of photovoltaic systems are most common in places where their use is economically justified. This is due to the climatic conditions in the given places. Australia, India, Saudi Arabia, and Italy have higher numbers of sunny days than Poland or Scotland, which is directly related to the irradiance and ultimately to the production of electricity. It is also possible to notice clearer a periodicity and regularity of weather in different seasons in these countries.

The exemplary results of forecasts from selected publications are presented in Table 2. Although not all the results from the analyzed publications have been highlighted, the literature clearly supports the use of artificial neural networks to implement forecasts.

The majority of the forecasts achieved very good results, and the R^2 coefficient reached values above 0.85, even tending to almost one (for minute forecasts). The linear regression method yielded satisfactory results, as shown in Table 2 (publication [21]), with an R^2 coefficient of 0.78. The value of the R^2 coefficient was worse for the same data and amounted to 0.69. The observed differences may be due to a variety of factors, including the selection of the network hyperparameters, the forecasting features, or the stochastic nature of the learning process. When considering publication [21], the KNN classification method was mentioned, and the R^2 coefficient was 0.35. The data from the same publication suggested that by bringing the forecast horizon closer (shortening the time ahead of the forecasts), better results could be obtained. When the time was shortened to 2 and 4 h, R^2 coefficients of 0.60 and 0.48 were obtained, respectively. As shown in publication [25], it was possible to obtain satisfactory results using the KNN method, if appropriate forecast assumptions were made. In Table 2, publication [19] showed that the separation of forecasting for cloudy and sunny conditions could improve the quality of the models. The literature suggests the creation of separate models for each month or season [40]. Some models bring better results in early spring, late autumn, and winter [41].

Table 2. The R-squared coefficient between the actual and modeled output power.

Publication	Model	R^2	Comments	Features	PV Power, Location	Datasets
[19]	MLP ABC	0.95	Separate forecasting for cloudy and sunny conditions	Total power, total irradiance, ambient temperature, and humidity	3.2 kWp, Tehran, Iran, elevation 1548 m	6895 of sunny, 3090 of cloudy, 680 for testing
	MLP ABC	0.83	Forecasting under all conditions			
[21]	LR	0.78	Forecast 8 h ahead	Deterministic and stochastic power, stochastic irradiance, ambient temperature, and panel temperature	1 MWp, Eni Energy Company, Italy	103,740 samples, 80% for training, 20% for testing
	MLP	0.69	Forecast 8 h ahead			
	KNN	0.35	Forecast 8 h ahead			
[25]	KNN	0.88	Forecast 1 h ahead	Year, month, day, hour, present energy, and energy 1, 2, and 3 h(s) ago	50 kWp, Cononsyth, Scotland	54,000 samples
	MLP	0.87	Forecast 1 h ahead			
	LSTM	0.85	Forecast 1 h ahead			
[39]	DNN	0.93	Forecast 1 d ahead	Weather forecast data from the Korean Meteorological Administration	2.448 kWp, Seoul, South Korea, rooftop PV system	3798 entries, 3000 for training, 798 for testing

3. Materials and Methods

The analysis in this study was performed using data from power plants located in a warm-summer, humid continental climate (Dfb in Köppen classification). In Bialystok, Poland, the annual average temperature is 7.7 °C (46.0 F). There are, on average, around 1755 sunshine hours per year and around 1365 sunshine hours from April until September.

Using the available data from the hybrid power plant, the following assumptions were made:

- Forecasts were made for a fixed-tilt system and a solar-tracking system due to the identical technical parameters of both PV plants;
- Forecasts were intended to specify day-ahead energy production expressed in kWh;
- Forecasts were made on the basis of the daily maximum and minimum temperatures, atmospheric pressure, wind speed, and an integer timestamp (a day of the year).

The variables used to predict energy production were chosen because this information is available in short-term and medium-term weather forecasts.

In this publication, a statistical method is used, which was based on the concept of a stochastic time series. Historical data were used for the learning process of the created model. The forecasting models were developed on the basis of data from 2015 to 2021, which covered the period from 1 April to 30 September, each year. Data collected in 2022 were used to test the developed forecasting models. Forecasts were made for different variants of model parameters, and the best results were presented graphically. For all forecasts, metrics defining their quality were determined. On their basis, a comparison of the results was performed.

The hybrid power plant (Figure 2) was located on the campus of the Bialystok University of Technology, Bialystok, Poland. The components of the power plant were two wind turbines and four PV micro-power plants. Among the plants, one can distinguish the following [42]:

- A fixed-tilt system with panels in the optimal direction, with a nominal power of 3.0 kWp (PV1);
- A solar-tracking system, with a nominal power of 3.0 kWp (PV3);

- A fixed-tilt system with panels that face the south-east, with a nominal power of 1.5 kWp (PV2a);
- A fixed-tilt system with panels that face the south-west, with a nominal power of 1.5 kWp (PV2b).

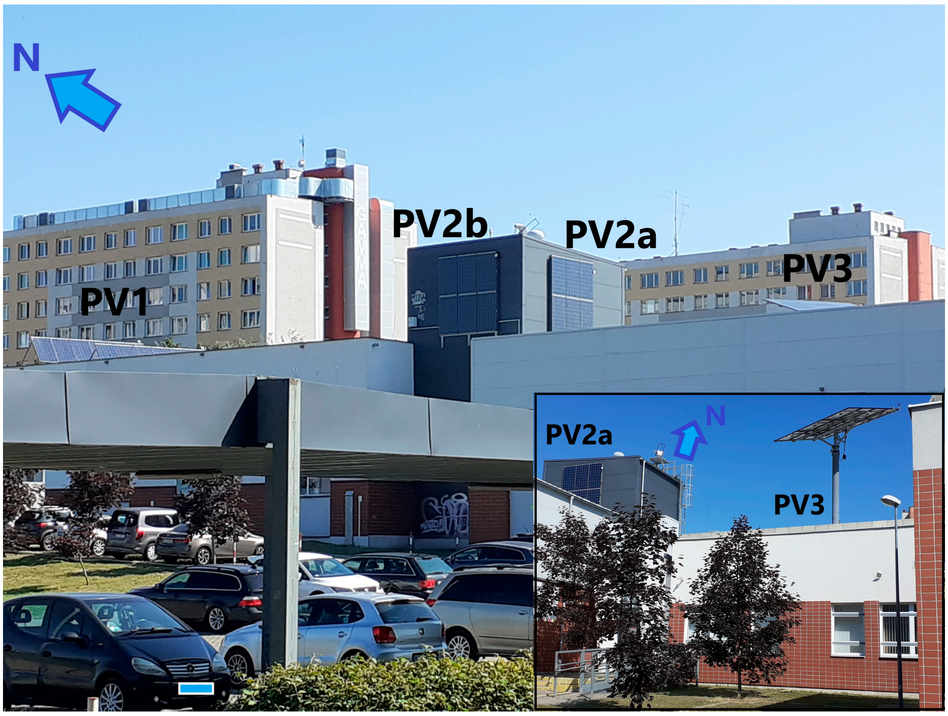

Figure 2. General view of the hybrid power plant, PV1—a fixed-tilt system with optimally directed panels, PV2a and PV2b—the panels on the facades, PV3—a solar tracking system.

The power plant had a telemetry system that gathered information about the electricity generated by each unit and the atmospheric parameters (Table 3). In this study, two photovoltaic systems were selected for the analysis. The first one was a solar tracking system (PV3) and the second one was a fixed-tilt system with optimally positioned panels (PV1). Two installations with identical power ratings were selected.

Table 3. Technical parameters of the weather station WS501-UMB [43,44].

Measured Quantity	Method	Measurement Performance
Speed and wind direction	Ultrasonic	Wind direction
		Range: 0–359.9° Accuracy: RMSE < 3° at speed > 1 m/s
		Wind speed
		Range: 0–75 m/s Accuracy: ±0.3 m/s or ±3% (0–35 m/s) ±5% (>35 m/s) RMS
Air temperature	NTC	Range: from −50 °C to +60 °C Accuracy: ±0.2 °C (−20–+50 °C), ±0.5 °C (>−30 °C)
Relative humidity	Capacitive	Range: 0–100% RH Accuracy: ±2% RH
Atmospheric pressure	MEMS	Range: 300–1200 hPa Accuracy: ±0.5 hPa (0–+40 °C)
Irradiance	Pyranometer	Range: 2000 W/m^2 (300–2800 nm)

The first step in any data analysis is the proper preparation of the data. The measurement data from the hybrid power plant were downloaded from the server as CSV files (http://elektrownia.pb.edu.pl, accessed on 1 February 2023). For a given year, each file was separate, and the data included measurements from 1 April to 30 September. The system recorded values with a period of about 10 s, which translated into over 2.1 million records for a given year in the analyzed period. Therefore, it was necessary to reduce the amount of data in order to create appropriate models.

A Python script was developed for this purpose. Its task was to load a selected CSV file and make a file with a single dataset from a given day, containing:

- The date converted to a day number from the analyzed period (timestamp);
- The energy produced by the PV1 unit (energy pv1);
- The energy produced by the PV1 unit- classified (energy pv1-class);
- The energy produced by the PV3 unit (energy pv3);
- The energy produced by the PV3 unit- classified (energy pv3-class);
- The maximal temperature (temp-max);
- The minimal temperature (temp-min);
- The atmospheric pressure (pressure);
- The average wind speed (wind-speed).

As a result, 183 items were obtained from a given year, each corresponding to one of the days of the analyzed period in the format as presented in Table 4. Regression algorithms were used with continuous data, and classification algorithms were employed with discrete data. The discretization of the value of the energy produced by the individual photovoltaic units, PV1 and PV3, relied on converting each numerical value to the nearest integer value, from 0 upwards, with a width of 1 kWh. In subsequent computations, the width of the discretization was 0.5 kWh. The data from 2022, which were used to evaluate the model, were prepared similarly. The last stage of the data preparation for the model calculation was to combine all the data into one dataset for the period 2015–2021.

Table 4. The format of the data used to develop the models.

Timestamp	Energy PV1	Energy PV1 Class	Energy PV3	Energy PV3 Class	Temp Max	Temp Min	Pressure	Wind Speed
Day No.	kWh	kWh	kWh	kWh	°C	°C	hPa	m/s
1	4.534	5	3.772	4	5.7	1.9	978.3	5.7
2	4.087	4	3.376	3	5.5	−0.7	984.2	3.7
3	5.009	5	4.326	4	6.5	−1.1	989.2	2.9

Based on the assumptions defined so far, programs responsible for the preparation of the forecast models were written using MLP neural networks. The goal was to create either medium-term or long-term forecasts. The Python scripting and Scikit-learn module for machine learning were employed for this purpose. A fully connected MLPRegressor, implemented as a function with a set of user-defined hyperparameters, was compared with other ML algorithms such as the simple linear regression, KNN, and MLPClassifier algorithms [45–48]. The predictive values of the electrical energy were obtained using regression machine learning models based on the following algorithm libraries: the extreme-gradient-boosting XGBRegressor model of the XGBoost library and the gradient-boosting CatBoostRegressor model of the CatBoost library [49].

4. Model Performance Metrics and Statistical Tests

Forecasts are certain predictions based on observations or collected data that are supported by appropriate calculations with a certain level of accuracy. Estimated time series tend to differ from real ones to a greater or lesser extent. There are various deterministic metrics and statistical tests to determine the quality of a forecast. The most relevant metrics are presented in publications [16,22,36,43]:

- *MAE* (mean absolute error);
- *MSE* (mean squared error);
- *RMSE* (root mean squared error);
- The coefficient of determination, R^2.

The mean absolute error expresses the average value of the absolute error in a given set of samples. The absolute error expresses the difference between the estimated value and the actual value expressed as a percentage. The *MAE* is described by the following formula:

$$MAE(y,\hat{y}) = \frac{1}{n}\sum_{i=0}^{n-1}|y_i - \hat{y}_i|, \tag{1}$$

where: n—number of samples, y_i—i-th actual value, and $\hat{y}_i$—i-th estimated (predicted) value.

The mean squared error assesses the average squared difference between the actual and predicted values. When a model has no error, the *MSE* equals zero. The formula for *MSE* is as follows:

$$MSE(y,\hat{y}) = \frac{1}{n}\sum_{i=0}^{n-1}\left(y_i - \hat{y}_i\right)^2. \tag{2}$$

The root mean squared error measures the average difference between the predicted values and the actual values. Mathematically, it is the standard deviation of the residuals. Residuals represent the distance between the regression line and the data points. *RMSE* values can range from zero to positive infinity and use the same units as the dependent variable:

$$RMSE(y,\hat{y}) = \sqrt{\frac{1}{n}\sum_{i=0}^{n-1}\left(y_i - \hat{y}_i\right)^2}. \tag{3}$$

The R-squared (R^2) value determines the proportion of variance in the dependent variable that can be explained by the independent variable. The R-squared value is a statistical measure of how close the data are to the fitted regression line. It is also known as the coefficient of determination. It is always between 0 and 1. Generally, the larger the R-squared value is, the better the regression model fits the observations. The formula for the R-squared value is as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}, \tag{4}$$

where: $\overline{y}$—the mean of actual values.

The Diebold–Mariano (DM) test for predictive accuracy is often used to statistically identify forecast accuracy equivalence for two sets of predictions [50]. It is assumed that the difference, d_i, between the first list of predictions and the actual values, y_i, is e_1, and that of the second list of predictions and the actual values, y_i, is e_2. The parameter e can represent *MSE*, *MAE*, or *MAPE* (mean absolute percentage error). In this analysis, d_i (loss-differential time series) is defined based on the following criterion:

$$d_i = e_1^2 - e_2^2 \tag{5}$$

Under the null hypothesis, the Diebold–Mariano statistic (*DM*) follows a standard normal distribution. The null hypothesis, H0, is that the two models (A and B) have the same forecast accuracy (equal predictive ability), i.e., $E(d_{A,i}) = E(d_{B,i})$. The alternative hypothesis, H1, that one is better than the other is given as $E(d_{A,i}) \neq E(d_{B,i})$. For a smaller number of samples, n, it is better to use the Harvey, Leybourne, and Newbold (HLN) test [51].

5. Results and Discussion

5.1. Feature Selection

To select the irrelevant and correlated features, the correlations with the PV output energy and intercorrelations between two features (Pearson's correlation coefficients) were checked and are presented in Table 5.

Table 5. Matrix of Pearson's correlation coefficient, which presents values of correlations and intercorrelations.

Power Plant PV1 (PV3)						
	energy	temp max	temp min	pressure	wind speed	timestamp
energy	1	-	-	-	-	-
temp max	0.40 (0.45)	1	-	-	-	-
temp min	0.06 (0.14)	0.83	1	-	-	-
pressure	0.32 (0.29)	0.06	−0.09	1	-	-
wind speed	−0.08 (−0.07)	−0.12	−0.10	−0.07	1	-
timestamp	−0.11 (−0.13)	0.36	0.47	0.12	−0.08	1

The strongest correlations existed between the energy produced and the maximum temperature (for PV1: 0.40; for PV3: 0.45), the minimum and maximum temperature (for PV1: 0.83), and the minimum temperature and the timestamp (for PV1: 0.47). It should be noted that the solar irradiance was not taken into account as an input feature. This is rarely practiced in the literature [4,19,21,24,25].

The results of numerous tests led to the development of many forecast variants. For both the PV1 and PV3 systems, forecasts were made. In both cases, predictions were made using the same methods.

5.2. PV1—Prediction Models and Their Performance

All the algorithms listed in Table 6 (except for linear regression) have several parameters that can be tuned. For the fixed-tilt system, Table 6 summarizes the best metrics using individual prognostic models. The most optimal training parameters and hyperparameters of the models are listed in the table footer. For the LR model, the number of coefficients of the polynomial, m (the variable is referred to as a degree), varied in the range from 1 to 5. The LR model provided the highest R-squared value for $m = 2$. The k in the KNN model is a parameter that refers to the number of nearest neighbors. The k value creates an environment for the data points to understand its similarities based on proximity.

The CatBoostRegressor model has several parameters, including the number of iterations, learning rate, L2 leaf regularization, and tree depth. XGBRegressor is an implementation of the gradient-boosting decision trees algorithm, and it uses the extreme-gradient-boosting algorithm. RandomForestRegressor is based on decision tree learners. The estimator applies multiple decision trees to randomly extracted subsets of a dataset and averages their predictions.

The Scikit-learn API provides several online solvers, but for the MLPRegressor and MLPClassifier models, a batch solver called lbgfs (limited-memory Broyden–Fletcher–Goldfarb–Shanno) is the best choice. With the same criteria as when selecting the best variant (maximizing the R^2 value), it can be said that the MLP artificial neural network in the regression variant worked the best. As can be seen, the coefficient of determination did not necessarily correlate with the *MAE*. Its lower value was not necessarily associated with a better correlation between the forecast and the actual state. When examining the *RMSE* values in Table 6, it is clear that the smaller *MSE* was, the higher the R^2 coefficient was.

Table 6. Summary of the best forecasts for individual methods—PV1.

Model		*MAE* [kWh]	*MAE* [%]	*MSE* [kWh^2]	*RMSE* [kWh]	R^2
Linear regression [1]		2.65	42.05	11.96	3.46	0.544
ARIMA		3.79	75.21	22.25	4.72	0.152
KNN [2]		4.07	59.96	25.88	5.09	0.014
XGBRegressor		2.90	53.08	13.19	3.63	0.497
RandomForestRegressor		2.85	48.32	13.11	3.62	0.500
CatBoostRegressor [5]		2.64	41.38	11.79	3.43	0.551
MLP	Regressor [3]	2.43	37.48	3.22	1.79	0.605
	Classifier [4]	2.87	34.94	3.80	1.95	0.450

1—variant for m = 2; 2—variant for k = 12 and *weights* = 'uniform'; 3—variant for *activation* = 'tanh', *solver* = 'lbfgs', and hidden_layers_size = 5; 4—variant for *activation* = 'logistic', *solver* = 'lbfgs', and *hidden_layers_size* = 100; 5—variant for *iterations* = 1000, *loss function* = '*RMSE*', and *l2_leaf_reg* = 30, *depth* = 6.

Figures 3–6 illustrate the best forecasts. The blue line represents historical measurement data, and the red and the green dotted lines represent the predicted values. The values of the x-axis represent day numbers from the period that was studied, i.e., day number 1 is the first day of April 2022, and number 183 is the last day of September 2022. With the visual representations of forecasts, it can be said that each of the forecasts (Figures 3–6) was generally consistent with the measurements. However, the KNN model (Figure 3) provided a lower-quality forecast.

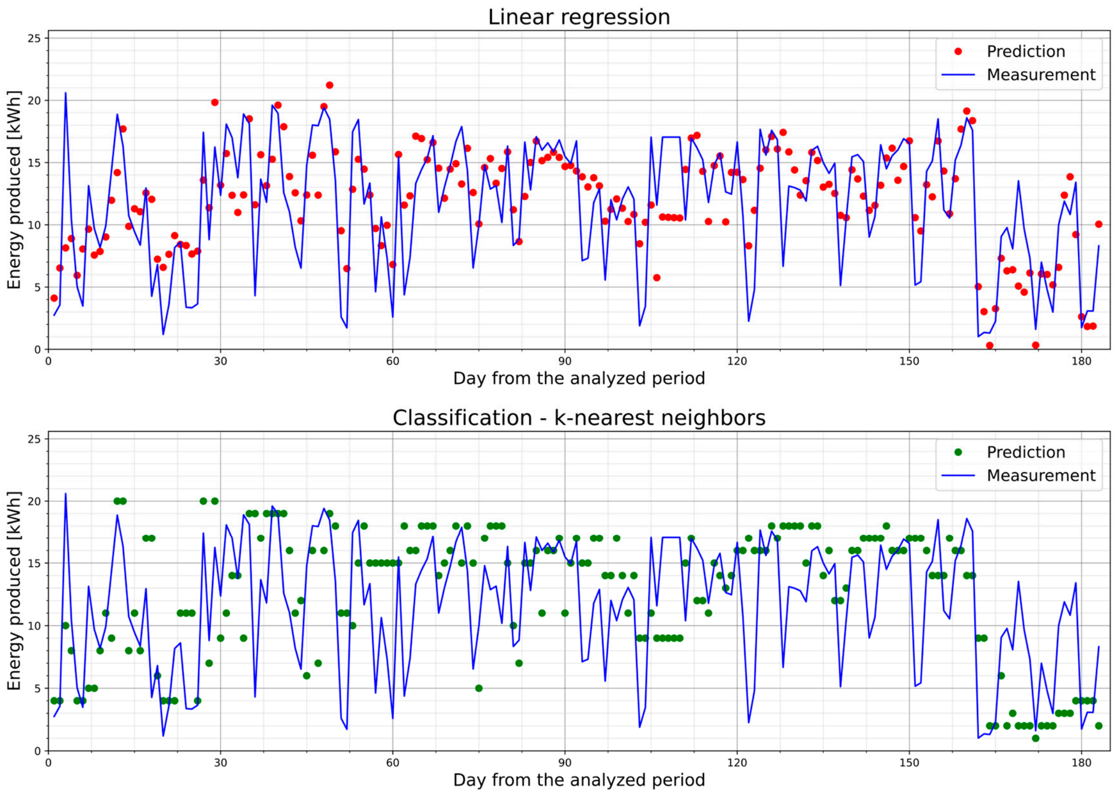

Figure 3. The energy produced in each day of the analyzed period and forecasts for the PV1 power plant using the linear regression model for m = 2 and the KNN model for k = 12 and *weights* = 'uniform'; the width of discretization is 1 kWh.

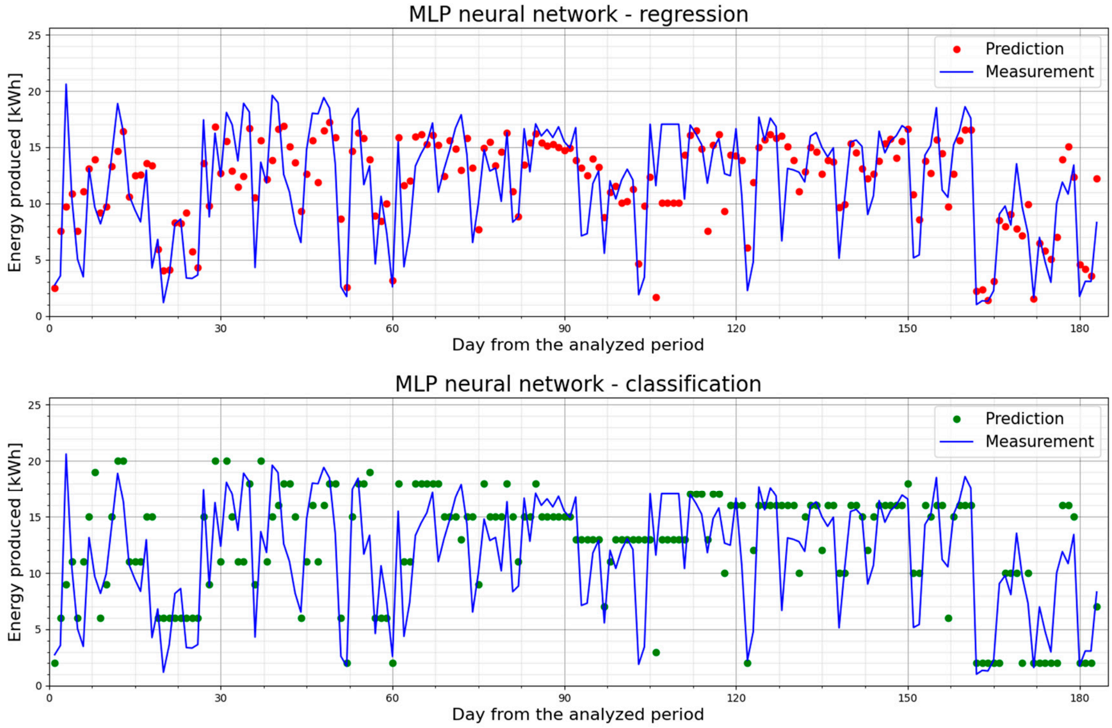

Figure 4. The produced energy in each day of the analyzed period and forecasts for the PV1 power plant using the MLPRegressor model and MLPClassifier model (a discretization width of 1 kWh).

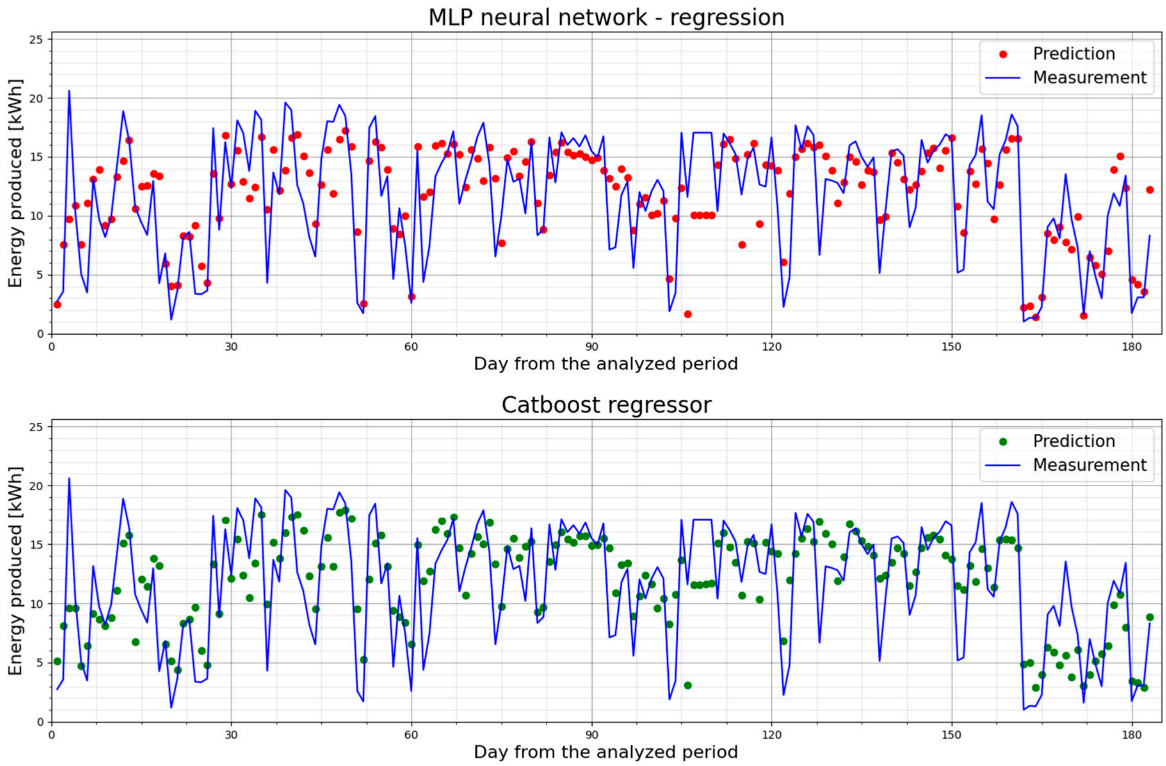

Figure 5. The produced energy in each day of the analyzed period and forecasts for the PV1 power plant using the MLPRegressor and CatBoostRegressor models.

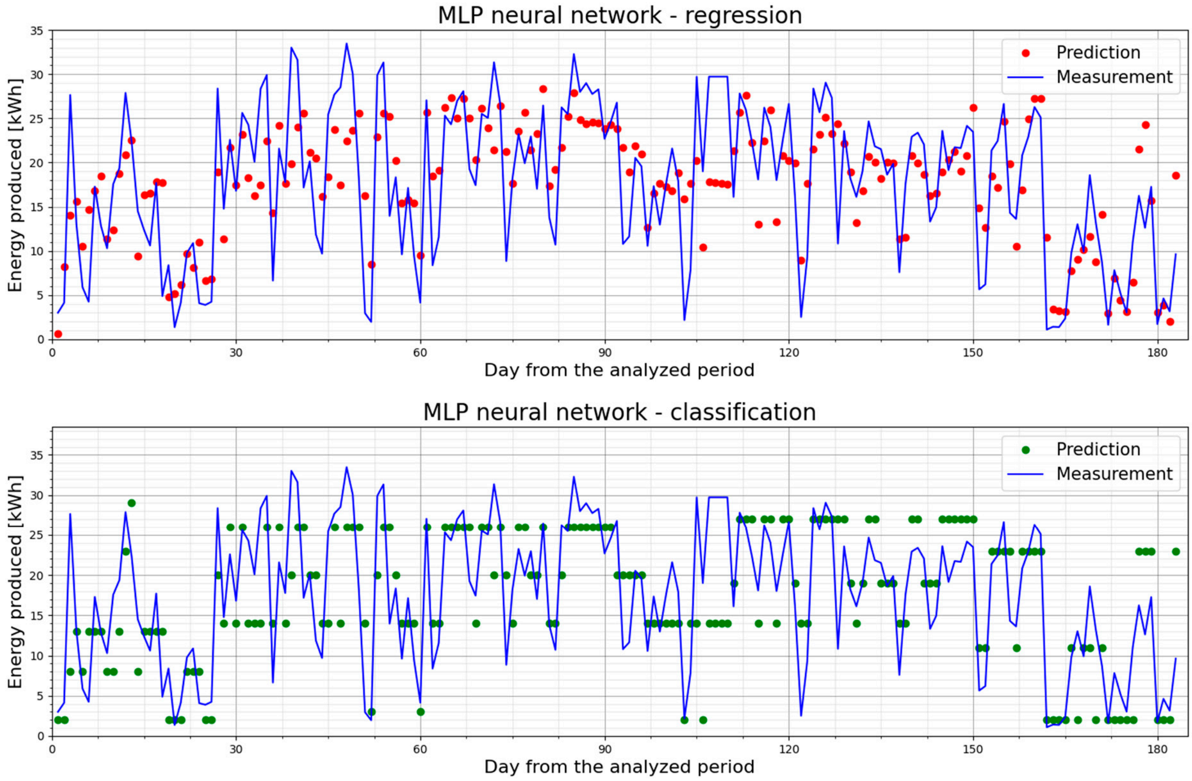

Figure 6. The energy produced in each day of the analyzed period and forecasts for the PV3 power plant using the MLPRegressor model and MLPClassifier model (with a discretization width of 1 kWh).

It was possible to see upward and downward trends, but the forecast points could be delayed in some segments. In September (Figure 3, days 153–183), when the weather conditions significantly deteriorated, the LR forecast was much better than the KNN method. The results of a deeper analysis of the KNN model are presented in Table 7. The reduction in the discretization width from 1 kWh to 0.5 kWh was not accompanied by an enhancement in terms of the metrics. The highest R-squared value of 0.014 was achieved for k = 12 nearest neighbors, with a mean computation time of 1.7 ms.

Table 7. The *RMSE*, R-squared, and computation time depending of the number of neighbors for two discretization widths (KNN model, weights = 'uniform').

	***Energy PV1 Class* with a Width of 0.5 kWh**			***Energy PV1 Class* with a Width of 1 kWh**		
Number of Neighbors *k*	***RMSE* [kWh]**	**R^2**	**Computation Time [s]**	***RMSE* [kWh]**	**R^2**	**Computation Time [s]**
4	6.044	−0.392	0.001	6.099	−0.418	0.002
5	6.075	−0.407	0.002	5.821	−0.291	0.002
6	6.396	−0.559	0.002	5.906	−0.329	0.001
7	6.229	−0.479	0.002	5.574	−0.184	0.002
8	6.224	−0.476	0.002	5.729	−0.250	0.002
9	6.007	−0.375	0.001	5.531	−0.166	0.002
10	5.882	−0.319	0.002	5.345	−0.089	0.002
11	5.524	−0.163	0.002	5.185	−0.025	0.002

Table 7. *Cont.*

	Energy PV1 Class with a Width of 0.5 kWh			*Energy PV1 Class* with a Width of 1 kWh		
Number of Neighbors *k*	***RMSE* [kWh]**	**R^2**	**Computation Time [s]**	***RMSE* [kWh]**	**R^2**	**Computation Time [s]**
12	5.440	−0.128	0.002	5.087	0.014	0.001
13	5.381	−0.103	0.002	5.191	−0.027	0.002
		mean	0.0017		mean	0.0017

Figures 4 and 5 present the measured energy and predictions using the MLPRegressor, MLPClassifier, and CatBoostRegressor models.

Eleven trials of forecasting with the use of the MLPClassifier are presented in Table 8. Selected metrics and the computation time for two discretization widths are given. The R-squared values were higher than those obtained by the KNN method and the *RMSE* values were about three times smaller, but the calculation time was much longer (Table 7).

Table 8. The MLPClassifier—the values of metrics and the computation time for two discretization widths.

Trial no.	*RMSE* [kWh]	R^2	Computation Time [s] 0.5 kWh Width	Computation Time [s] 1 kWh Width
1	1.99	0.402	96.41	84.99
2	1.97	0.430	97.87	84.06
3	2.02	0.361	96.32	84.54
4	2.01	0.381	96.68	63.84
5	1.96	0.440	105.94	27.92
6	1.96	0.435	67.24	62.96
7	1.94	0.456	97.55	43.18
8	1.93	0.473	97.63	71.53
9	1.96	0.441	97.63	84.61
10	2.01	0.373	103.35	29.34
11	1.88	0.521	101.97	72.65
		mean	96.24	64.51
		max	105.94	84.99
		min	67.24	27.92

5.3. PV3—Prediction Models and Their Performance

In Table 9, the best results are shown for the solar-tracking system. It can be said that the best forecast was developed for the MLPRegressor model, using the same criteria as in the previous analyses. As previously stated, it was difficult to establish a clear correlation between the *MAE* and the R-squared coefficient. The PV3 results had higher *RMSE* values than the PV1 ones, which is worth noting.

The forecasts for the solar-tracking system presented in Figure 6 provided analogous observations, as in the case of the fixed-tilt system. The forecasts were generally consistent with the measurements, and it was possible to see upward and downward trends.

Table 9. Summary of the best forecasts for individual methods—PV3.

Model		*MAE* [kWh]	*MAE* [%]	*MSE* [kWh^2]	*RMSE* [kWh]	R^2
Linear regression [1]		4.38	47.95	32.48	5.70	0.572
ARIMA		6.16	90.51	57.37	7.57	0.244
KNN [2]		6.70	71.74	74.30	8.62	0.022
XGBRegressor		4.94	61.87	37.12	6.09	0.511
RandomForestRegressor		4.70	54.13	35.60	5.97	0.531
CatBoostRegressor [5]		4.33	46.38	31.75	5.63	0.582
MLP	Regressor [3]	4.35	46.73	5.65	2.38	0.576
	Classifier [4]	4.61	37.29	6.21	2.49	0.493

1—variant for $m = 2$; 2—variant for $k = 2$ and *weights* = 'distance'; 3—variant for *activation* = 'tanh', *solver* = 'lbfgs', and *hidden_layers_size* = 5; 4—variant for *activation* = 'logistic', *solver* = 'lbfgs', and *hidden_layers_size* = 100; 5—variant for *iterations* = 1000, *loss function* = '*RMSE*', and *l2_leaf_reg* = 30, *depth* = 6.

5.4. PV1 vs. PV3—A Comparison of Quantitative Metrics

When forecasting using neural networks, it is impossible to use the same procedure as in the case of the LR or KNN models to select their best parameters. This is due to the stochastic nature of the neural network learning process. Each time, a different model is created by using the same training dataset. Therefore, the selection of the most promising outcomes was accomplished using an experimental approach. Numerous trials were conducted using various combinations of hyperparameters. In Figure 7, the *RMSE* values depending on the number of nodes in the hidden layers are presented. The *RMSE* for the MLPRegressor model with five nodes was the lowest. For the MLPClassifier method, the *RMSE* was low for the hidden layers with 30 nodes, with the lowest being at 100 nodes. In Figure 8, the *RMSE* values depending on the number of iterations in the CatBoostRegressor model are presented. A large difference in the *RMSE* values between the PV1 and PV3 power plants was observed.

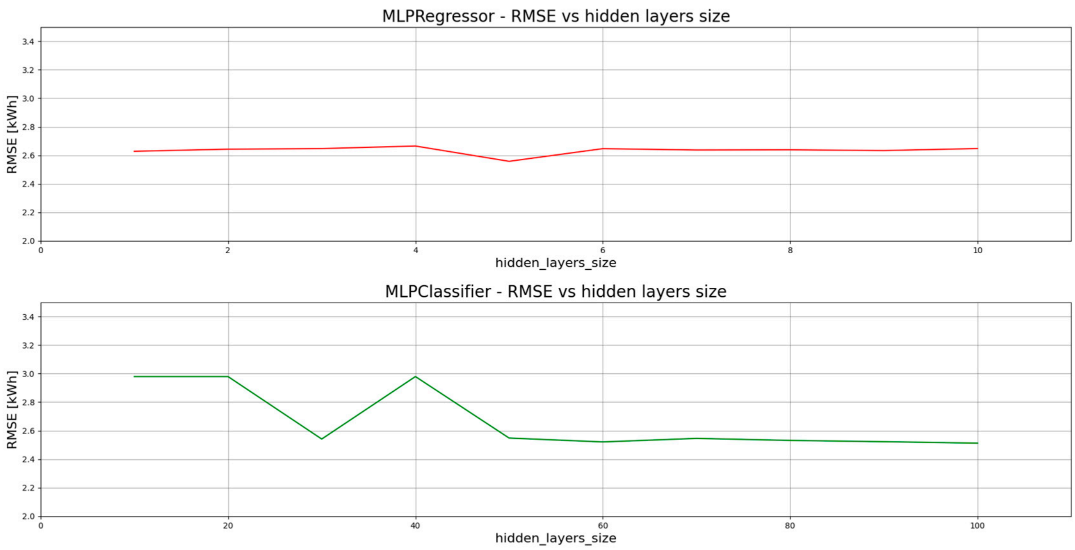

Figure 7. *RMSE* values depending on the number of nodes in the hidden layers for the MLPRegressor and MLPClassifier models.

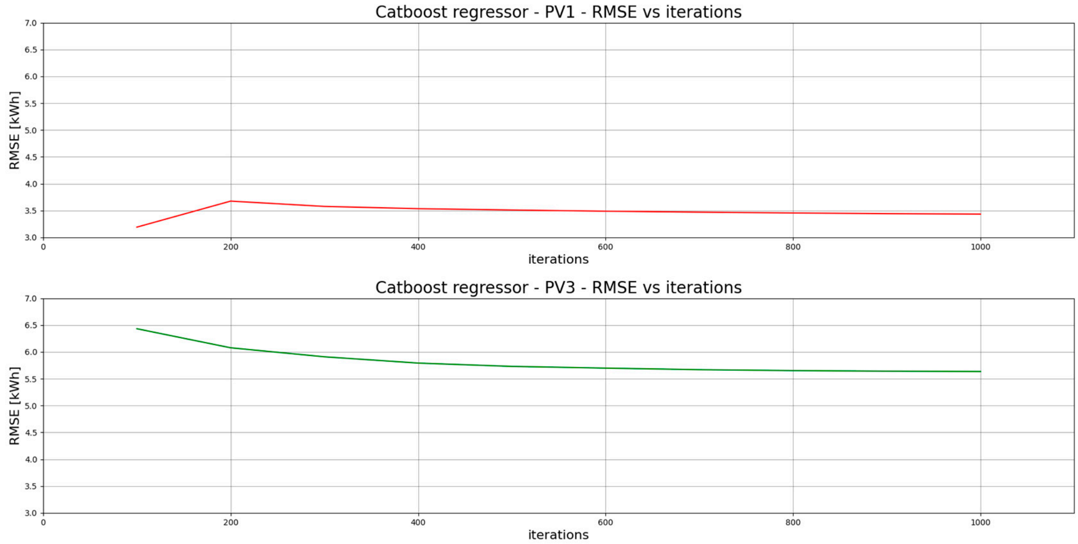

Figure 8. *RMSE* values depending on the number of iterations for the CatBoostRegressor models.

The mean time of computations was compared for the three models. As shown in Table 10, the MLPRegressor and CatBoostRegressor models performed predictions up to 2 s. In the case of the MLPClassifier model, a number of nodes above 30 gave results from 14 to 64 s (CPU: AMD RYZEN 5 2600, 32 GB RAM, Windows 10 22H2 system, single-core computations).

Table 10. The mean computation time depending of the number of nodes in the hidden layers or the number of iterations for three prediction models.

MLPRegressor		MLPClassifier		CatBoostRegressor	
Number of Nodes in the Hidden Layers	**Mean Time (10 Trials) [s]**	**Number of Nodes in the Hidden Layers**	**Mean Time (10 Trials) [s]**	**Number of Iterations**	**Mean Time (10 Trials) [s]**
1	0.006	10	1.535	100	0.490
2	0.031	20	6.737	200	0.584
3	0.105	30	3.121	300	0.743
4	0.057	40	14.346	400	0.920
5	0.210	50	13.466	500	1.028
6	0.701	60	18.697	600	1.123
7	0.616	70	30.275	700	1.261
8	0.784	80	40.475	800	1.433
9	1.858	90	37.015	900	1.535
10	2.297	100	64.509	1000	1.693

After conducting many trials, three of the best results for the MLPRegressor model were obtained and are compiled in Table 11 with the right combinations of hyperparameters that maximized the model's performance.

Table 11. The best results of regression using MLPRegressor model for the fixed-tilt system (PV1).

	MLPRegressor				
Variant	***MAE* [kWh]**	***MAE* [%]**	***MSE* [kWh^2]**	***RMSE* [kWh]**	**R^2**
1	2.44	36.62	3.24	1.800	0.600
2	2.50	35.11	3.31	1.819	0.582
3	2.63	39.64	3.45	1.857	0.547

Variant 1—*activation* = 'tanh', *solver* = 'lbfgs', *hidden_layers_size* = 5. Variant 2—*activation* = 'tanh', *solver* = 'lbfgs', and *hidden_layers_size* = 5. Variant 3—*activation* = 'tanh', *solver* = 'lbfgs', and *hidden_layers_size* = 5.

The most effective variants of the energy forecasting model for PV3 are compiled in Table 10. The optimal combinations of hyperparameters are shown in the table footer. The activation function of each of them was different (tanh, relu, identity).

5.5. DM Test to Distinguish the Significant Differences in the Forecasting Accuracy

Tables 11 and 12 show that similar R-squared values were obtained for the PV3 forecasting models, but all the other metrics had higher values. It was difficult to determine which forecast model variant was better in the sense of having a better predictive accuracy. One further step was to determine whether there was a significant difference among the forecasts presented in rows 1, 2, and 3 of Table 12 for the measurement data shown in Figure 6. For this purpose, the Diebold–Mariano (DM) test was used. The results for the two-tailed and one-tailed tests are shown in Table 13.

Table 12. The best results of regression using MLPRegressor neural network for the solar-tracking system (PV3).

	MLPRegressor				
Variant	***MAE* [kWh]**	***MAE* [%]**	***MSE* [kWh^2]**	***RMSE* [kWh]**	**R^2**
1	4.50	52.84	5.74	2.40	0.566
2	4.56	49.77	5.80	2.41	0.556
3	4.82	56.56	5.99	2.45	0.527

Variant 1—*activation* = 'tanh', *solver* = 'lbfgs', *hidden_layers_size* = 5. Variant 2—*activation* = 'identity', *solver* = 'lbfgs', and *hidden_layers_size* = 5. Variant 3—*activation* = 'relu', *solver* = 'adam', and *hidden_layers_size* = 5.

Table 13. The DM test for PV3 forecasting performance based on MPLRegressor models.

DM Test		**The DM Statistic (*DM*)**			**The *p*-Value (*p*)**		
two-tailed	Variant	1	2	3	1	2	3
	1	-	−0.3858	−1.5601	-	0.7000	0.1205
	2	0.3858	-	−1.9891	0.7000	-	0.0482
	3	1.5601	1.9891	-	0.1205	0.0482	-
one-tailed	Variant	1	2	3	1	2	3
	1	-	−0.3858	−1.5601	-	0.3500	0.0602
	2	0.3858	-	−1.9891	0.6500	-	0.0241
	3	1.5601	1.9891	-	0.9398	0.9760	-

The forecast horizon: $h = 1$; loss function: *MSE* (5); the long-run variance estimator: auto-correlation function (acf).

The null hypothesis can be rejected if the p-value is less than 0.05. Since the DM statistic converges to a normal distribution, the null hypothesis, H0, can also be rejected at the 5% level if $|DM| > 1.96$. The findings from the DM test statistics showed that the observed differences of the forecasting values between model variants 2 and 3 were significant at 5% ($p = 0.05$), as mentioned in Table 13, and indicated that variant 2 had a better forecasting efficiency than variant 3 (the alternative hypothesis, H1).

5.6. Improving Regression Model Performance Using a Target Transformation

Multiple linear regression is used to estimate the relationship between two or more independent variables (features) and one dependent variable. The relationship between these variables is functional, which is realized in a mathematical model. The variable whose value is predicted is called the target. In the Scikit-learn package, the class called TransformedTargetRegressor enables the transforming of the targets before fitting a regression model.

In this application, the transformation was realized in two ways:

- Function transformation—logarithmic, log(1 + x), and exponential, exp(x) – 1, functions were used to transform the targets before training a linear regression model and using it for prediction (Table 14);
- Feature scaling data—each feature was scaled to a 0–1 range (MinMaxScaler), and inverse transformation was used (Table 15).

Table 14. The effect of function transformation on the analysis using regression models (PV1).

	Log Transformation				
Regressor	***MAE* [kWh]**	***MAE* [%]**	***MSE* [kWh^2]**	***RMSE* [kWh]**	***R^2***
LinearRegression	2.89	41.46	13.33	3.65	0.492
HuberRegressor	2.86	47.84	12.74	3.57	0.514
Ridge	2.89	41.46	13.44	3.67	0.492
BayesianRidge	2.89	41.55	13.33	3.65	0.492
RidgeCV	2.89	41.47	13.33	3.65	0.492

Table 15. The effect of MinMaxScaler on the analysis using regression models (PV1).

	Feature Scaling Data				
Regressor	***MAE* [kWh]**	***MAE* [%]**	***MSE* [kWh^2]**	***RMSE* [kWh]**	***R^2***
LinearRegression	2.71	44.22	12.13	3.48	0.538
HuberRegressor	2.64	40.55	11.77	3.43	0.551
Ridge	2.76	46.42	12.50	3.54	0.523
BayesianRidge	2.71	44.36	12.15	3.49	0.537
RidgeCV	2.71	44.46	12.16	3.49	0.536

The effect of the target transformation is shown in Figure 9. The metrics shown in Tables 14 and 15 showed a slight reduction in the *MAE* and *RMSE* and an increase in the R-squared value. The HuberRegressor, a regression technique that is robust to outliers [52], used with MinMaxScaler had the best effect on boosting the metrics.

5.7. The Performance of the Selected Models after Splitting the Dataset

The dataset, described in Section 2, was split and separated into months. The computations were performed for each month separately. Table 16 presents the R-squared values that indicate how well both models fit the measured output energy in the separate months of 2022. The values of the R-squared coefficient for the MLPRegressor model were higher than for the CatBoostRegressor model. When datasets are smaller, it has been shown that MLPRegressor is superior to CatBoostRegressor. The MLPRegressor model delivered best results for September 2022 (PV1, R^2 = 0.827; PV3, R^2 = 0.818). The results in Table 16 illustrate the variability in the weather conditions in Poland during the most productive months.

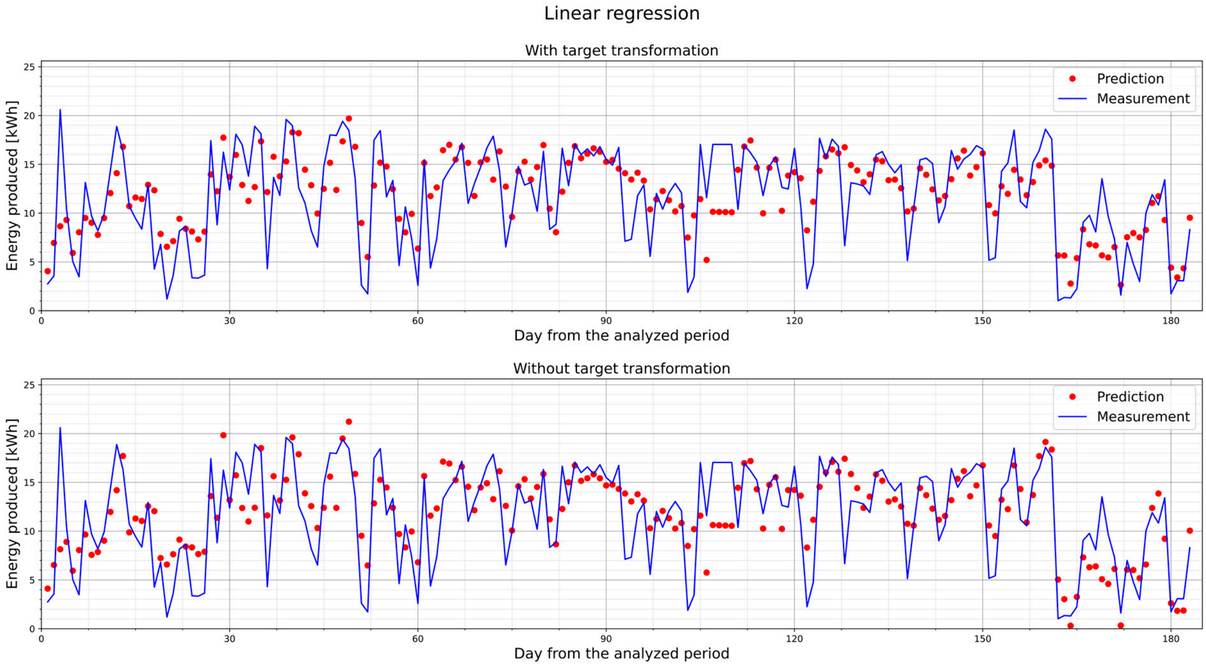

Figure 9. The energy produced in each day of the analyzed period and forecasts for the PV1 power plant using the linear regression model for m = 2 and predictions with and without target transformation (MinMaxScaler).

The energy produced in each day of the selected months in 2022 is presented in Figures 10 and 11, and the forecasts for the PV1 and PV3 power plants using the MLPRegressor and CatBoostRegressor models are compared.

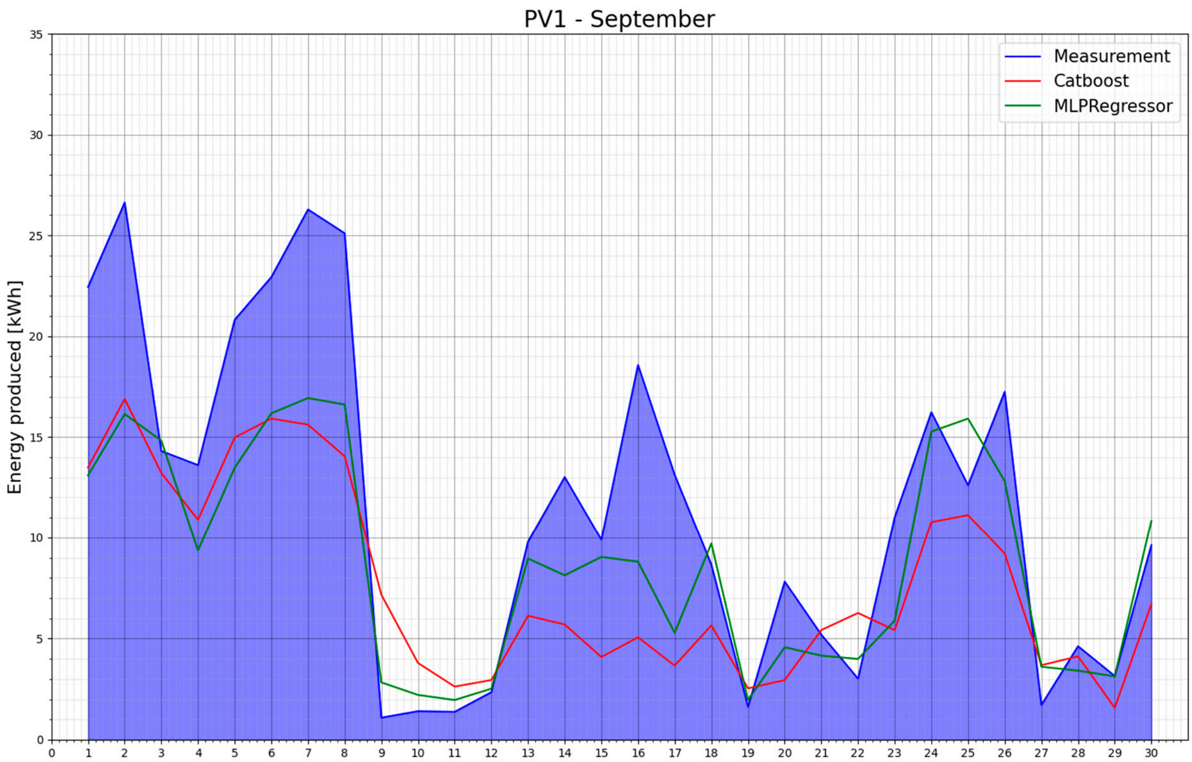

Figure 10. The energy produced in September 2022 and forecasts for the PV1 power plant using MLPRegressor and CatBoostRegressor models.

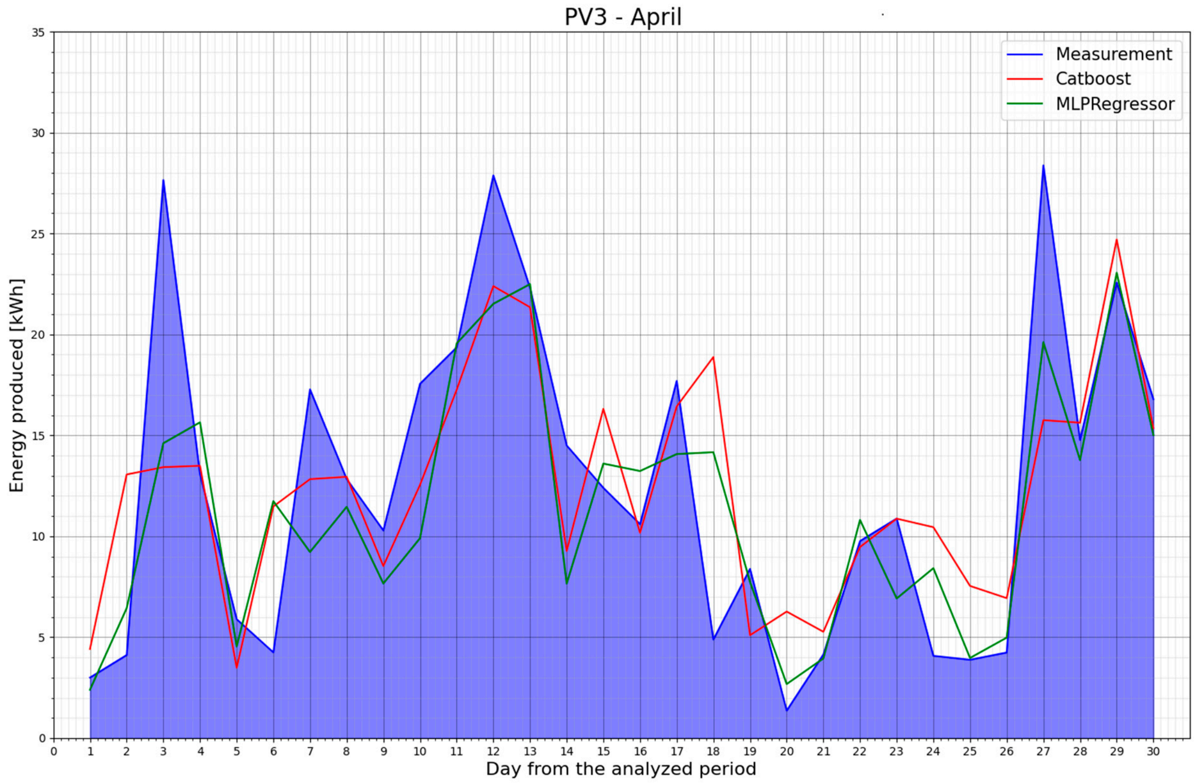

Figure 11. The energy produced in April 2022 and forecasts for the PV3 power plant using MLPRegressor and CatBoostRegressor models.

Table 16. The R-squared values that indicate how well both models fit the PV output energy in the separate months of 2022.

PV System	Model	April	May	June	July	August	September
PV1	MLPRegressor	0.564	0.603	0.463	0.474	0.422	0.827
PV1	CatBoostRegressor	0.502	0.445	0.211	0.402	0.267	0.664
PV3	MLPRegressor	0.626	0.564	0.550	0.449	0.432	0.818
PV3	CatBoostRegressor	0.494	0.284	0.446	0.381	0.223	0.682

6. Conclusions

The selected supervised learning algorithms were provided with historical data and used to find the relationship that had the best predictive power. Using a machine learning framework like Scikit-learn, it was easy to fit the different machine learning models on a predictive modeling dataset. In this article, traditional ML techniques were compared.

Considering all the models examined in this article, it can be concluded that, for the presented dataset and the selected features, the regression models provided more accurate forecasts than the classification models. In each of the photovoltaic units, the best results were obtained using the regression-based models, i.e., the simple linear regression, HuberRegressor, MLPRegressor, and CatBoostRegressor models. In general, the *RMSE* values were higher for the solar-tracker power plant, PV3, than for PV1.

When the classification-based models were examined (KNN, MLPClassifier), the forecasts were less accurate, considering the *RMSE* and R-squared metrics. The use of MLP neural networks improved the quality of the forecasts, particularly in terms of the *RMSE*. The results of these trials were significantly closer to those obtained using regression methods (Table 6, Table 9). Considering the time of computation, the KNN method was the fastest (2 ms), and the MLPClassifier method was the slowest (up to 64 s). The poor results for ARIMA were due to the fact that it has some limitations. The ARIMA contains a procedural part, which creates a stationary time series from a non-stationary one. It is designed to use univariate time series data. Therefore, this model would probably bring better results when the correlations among the variables are lower.

The detailed conclusions are as follows:

- The number of input variables utilized in the modeling can affect the forecasting performance. As evident from Table 5, the strongest intercorrelation existed between the minimum and maximum temperatures. The minimum and maximum temperatures had a considerable periodicity (intercorrelations with the timestamp). However, the absence of a minimum temperature in the MLPRegressor model had a negative impact on the forecasts, with a decrease in the value of the R-squared coefficient from 0.6 to about 0.4.
- A single metric was not sufficient to be a universal indicator of forecast quality (Tables 6 and 9).
- The DM test can reveal significant differences in the forecasting performance between two variants of a model. In the case of MLPRegressor neural network models, it can depend on the adopted activation function and solver (Tables 11 and 12).
- The performance of regression models can be improved by using a target transformation (Tables 14 and 15). This affects the metrics.
- Splitting the dataset can boost the metric values. Table 16 presents a comparison between the MLPRegressor and CatBoostRegressor models after splitting the dataset. The MLPRegressor model proved to be more effective for forecasting within individual data groups (monthly data). The MLPRegressor model yielded the highest R-squared value for September 2022, exceeding 0.8. This could mean that this month had the most repeated weather conditions in Poland compared to the months of September in the previous seven years.
- To enhance the model performance, the periodicity of light could be taken into consideration. Also, seasonality (weekly or monthly repeating patterns) could be detected and incorporated in the forecasting.

In addition to the traditionally used one-block models for PV power forecasting, there are currently developed state-of-the-art models based on the encoder–decoder approach [53]. Future research could focus on developing techniques for scheduling and forecasting renewable energy using novel robust algorithms [54], creating more advanced DL models with an attention mechanism, or using the newest auto-regressive encoder–decoder transformer models [55].

Author Contributions: Conceptualization, A.I.; methodology, M.S. and A.I.; software, M.S.; validation, M.S.; formal analysis, A.I.; investigation, M.S.; resources, M.S.; data curation, M.S. and A.I.; writing—original draft preparation, M.S. and A.I.; writing—review and editing, A.I.; visualization, M.S.; supervision, A.I.; project administration, A.I.; funding acquisition, A.I. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by the Ministry of Education and Science in Poland at the Bialystok University of Technology under research project no. WZ/WE-IA/7/2023.

Data Availability Statement: Not applicable.

Acknowledgments: This research could take place thanks to the earlier realization of two projects: a hybrid system of small wind and photovoltaic power designed to supply electricity to the Research and Teaching Center of the Electrical Faculty at Zwierzyniecka Street 10, the project No. WND-RPPD. 05. 02. 00-20-034/12, entitled 'Improving the energy efficiency of the infrastructure of the Bialystok University of Technology using renewable energy sources', Priority Axis V. Development of infrastructure for environmental protection, Measure 5.2 Development of local environmental infrastructure, co-financed by the European Regional Development Fund under the Regional Operational Programme for Podlaskie Voivodeship 2007–2013. Additionally, the operation of the production system in urban (urbanized) conditions was examined in the framework of project No. WND-RPPD. 01. 01. 00-20-015/12, entitled 'Study of the effectiveness of active and passive methods of improving the energy efficiency of infrastructure using renewable energy sources', Priority Axis I. Increase of innovation and support of entrepreneurship in the region, Measure 1.1. Creating conditions for the development of innovation, co-financed by the European Regional Development Fund and the state budget under the Regional Operational Programme for Podlaskie Voivodeship 2007–2013.

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

References

1. Vita, V.; Fotis, G.; Pavlatos, C.; Mladenov, V. A New Restoration Strategy in Microgrids after a Blackout with Priority in Critical Loads. *Sustainability* **2023**, *15*, 1974. [CrossRef]
2. Soto, E.A.; Bosman, L.B.; Wollega, E.; Leon-Salas, W.D. Analysis of Grid Disturbances Caused by Massive Integration of Utility Level Solar Power Systems. *Eng* **2022**, *3*, 236–253. [CrossRef]
3. Paska, J.; Surma, T.; Terlikowski, P.; Zagrajek, K. Electricity Generation from Renewable Energy Sources in Poland as a Part of Commitment to the Polish and EU Energy Policy. *Energies* **2020**, *13*, 4261. [CrossRef]
4. Alaraj, M.; Kumar, A.; Alsaidan, I.; Rizwan, M.; Jamil, M. Energy Production Forecasting from Solar Photovoltaic Plants Based on Meteorological Parameters for Qassim Region, Saudi Arabia. *IEEE Access* **2021**, *9*, 83241–83251. [CrossRef]
5. ElNozahy, M.S.; Salama, M.M.A. Technical impacts of grid-connected photovoltaic systems on electrical networks-a review. *J. Renew. Sustain. Energy* **2013**, *5*, 032702. [CrossRef]
6. Yin, L.; Cao, X.; Liu, D. Weighted fully-connected regression networks for one-day-ahead hourly photovoltaic power forecasting. *Appl. Energy* **2023**, *332*, 120527. [CrossRef]
7. Ulbricht, R.; Fischer, U.; Lehner, W.; Donker, H. First steps towards a systematical optimized strategy for solar energy supply forecasting. In Proceedings of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECMLPKDD 2013), Prague, Czech Republic, 23–27 September 2013.
8. Ssekulima, E.B.; Anwar, M.B.; Al Hinai, A.; El Moursi, M.S. Wind speed and solar irradiance forecasting techniques for enhanced renewable energy integration with the grid: A review. *IET Renew. Power Gener.* **2016**, *10*, 885–898. [CrossRef]
9. Fara, L.; Diaconu, A.; Craciunescu, D.; Fara, S. Forecasting of energy production for photovoltaic systems based on ARIMA and ANN advanced models. *Int. J. Photoenergy* **2021**, *2021*, 6777488. [CrossRef]
10. Lezmi, E.; Xu, J. *Time Series Forecasting with Transformer Models and Application to Asset Management*; SSRN: Rochester, NY, USA, 2023. [CrossRef]
11. Yang, Y.; Lu, J. A Fusion Transformer for Multivariable Time Series Forecasting: The Mooney Viscosity Prediction Case. *Entropy* **2022**, *24*, 528. [CrossRef]
12. Petropoulos, F.; Apiletti, D.; Assimakopoulos, V.; Babai, M.Z.; Barrow, D.K.; Ben Taieb, S.; Bergmeir, C.; Bessa, R.J.; Bijak, J.; Boylan, J.E.; et al. Forecasting: Theory and practice. *Int. J. Forecast.* **2022**, *38*, 705–871. [CrossRef]
13. Mohamad Radzi, P.N.L.; Akhter, M.N.; Mekhilef, S.; Mohamed Shah, N. Review on the Application of Photovoltaic Forecasting Using Machine Learning for Very Short- to Long-Term Forecasting. *Sustainability* **2023**, *15*, 2942. [CrossRef]
14. Raschka, S.; Patterson, J.; Nolet, C. Machine Learning in Python: Main Developments and Technology Trends in Data Science, Machine Learning, and Artificial Intelligence. *Information* **2020**, *11*, 193. [CrossRef]
15. Petropoulos, F.; Makridakis, S.; Assimakopoulos, V.; Nikolopoulos, K. 'Horses for Courses' in demand forecasting. *Eur. J. Oper. Res.* **2014**, *237*, 152–163. [CrossRef]
16. Li, G.; Wei, X.; Yang, H. Decomposition integration and error correction method for photovoltaic power forecasting. *Measurement* **2023**, *208*, 112462. [CrossRef]
17. Sarmas, E.; Strompolas, S.; Marinakis, V.; Santori, F.; Bucarelli, M.A.; Doukas, H. An Incremental Learning Framework for Photovoltaic Production and Load Forecasting in Energy Microgrids. *Electronics* **2022**, *11*, 3962. [CrossRef]
18. Colak, M.; Yesilbudak, M.; Bayindir, R. Daily Photovoltaic Power Prediction Enhanced by Hybrid GWO-MLP, ALO-MLP and WOA-MLP Models Using Meteorological Information. *Energies* **2020**, *13*, 901. [CrossRef]
19. Khademi, M.; Moadel, M.; Khosravi, A. Power Prediction and Technoeconomic Analysis of a Solar PV Power Plant by MLP-ABC and COMFAR III, considering Cloudy Weather Conditions. *Int. J. Chem. Eng.* **2016**, *2016*, 1031943. [CrossRef]
20. Abdel-Nasser, M.; Mahmoud, K. Accurate photovoltaic power forecasting models using deep LSTM-RNN. *Neural Comput. Appl.* **2019**, *31*, 2727–2740. [CrossRef]

21. Sala, S.; Amendola, A.; Leva, S.; Mussetta, M.; Niccolai, A.; Ogliari, E. Comparison of Data-Driven Techniques for Nowcasting Applied to an Industrial-Scale Photovoltaic Plant. *Energies* **2019**, *12*, 4520. [CrossRef]
22. Wang, H.; Yi, H.; Peng, J.; Wang, G.; Liu, Y.; Jiang, H.; Liu, W. Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network. *Energy Convers. Manag.* **2017**, *153*, 409–422. [CrossRef]
23. Zhu, T.; Guo, Y.; Li, Z.; Wang, C. Solar Radiation Prediction Based on Convolution Neural Network and Long Short-Term Memory. *Energies* **2021**, *14*, 8498. [CrossRef]
24. Trabelsi, M.; Massaoudi, M.; Chihi, I.; Sidhom, L.; Refaat, S.S.; Huang, T.; Oueslati, F.S. An Effective Hybrid Symbolic Regression–Deep Multilayer Perceptron Technique for PV Power Forecasting. *Energies* **2022**, *15*, 9008. [CrossRef]
25. Cabezón, L.; Ruiz, L.G.B.; Criado-Ramón, D.; Gago, E.J.; Pegalajar, M.C. Photovoltaic Energy Production Forecasting through Machine Learning Methods: A Scottish Solar Farm Case Study. *Energies* **2022**, *15*, 8732. [CrossRef]
26. Wang, Z.; Koprinska, I.; Rana, M. Clustering based methods for solarpower forecasting. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; pp. 1487–1494. [CrossRef]
27. Klyuev, R.V.; Morgoev, I.D.; Morgoeva, A.D.; Gavrina, O.A.; Martyushev, N.V.; Efremenkov, E.A.; Mengxu, Q. Methods of Forecasting Electric Energy Consumption: A Literature Review. *Energies* **2022**, *15*, 8919. [CrossRef]
28. Sunthornnapha, T. Utilization of MLP and Linear Regression Methods to Build a Reliable Energy Baseline for Self-benchmarking Evaluation. *Energy Procedia* **2017**, *141*, 189–193. [CrossRef]
29. Purlu, M.; Turkay, B.E. Estimating the Distributed Generation Unit Sizing and Its Effects on the Distribution System by Using Machine Learning Methods. *Elektron. Elektrotech.* **2021**, *27*, 24–32. [CrossRef]
30. Dellino, G.; Laudadio, T.; Mari, R.; Mastronardi, N.; Meloni, C.; Vergura, S. Energy production forecasting in a PV plant using transfer function models. In Proceedings of the 2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC), Rome, Italy, 10–13 June 2015; pp. 1379–1383. [CrossRef]
31. Khilar, R.; Suba, G.M.; Kumar, T.S.; Samson Isaac, J.; Shinde, S.K.; Ramya, S.; Prabhu, V.; Erko, K.G. Improving the Efficiency of Photovoltaic Panels Using Machine Learning Approach. *Int. J. Photoenergy* **2022**, *2022*, 4921153. [CrossRef]
32. Olabi, A.G.; Abdelkareem, M.A.; Semeraro, C.; Radi, M.A.; Rezk, H.; Muhaisen, O.; Al-Isawi, O.A.; Sayed, E.T. Artificial neural networks applications in partially shaded PV systems. *Therm. Sci. Eng. Prog.* **2023**, *37*, 101612. [CrossRef]
33. Bhatnagar, P.; Nema, R.K. Maximum power point tracking control techniques: State-of-the-art in photovoltaic applications. *Renew. Sustain. Energy Rev.* **2013**, *23*, 224–241. [CrossRef]
34. Kermadi, M.; Berkouk, E.M. Artificial intelligence-based maximum power point tracking controllers for Photovoltaic systems: Comparative study. *Renew. Sustain. Energy Rev.* **2017**, *69*, 369–386. [CrossRef]
35. Andrade, C.H.T.d.; Melo, G.C.G.d.; Vieira, T.F.; Araújo, Í.B.Q.d.; Medeiros Martins, A.d.; Torres, I.C.; Brito, D.B.; Santos, A.K.X. How Does Neural Network Model Capacity Affect Photovoltaic Power Prediction? A Study Case. *Sensors* **2023**, *23*, 1357. [CrossRef] [PubMed]
36. Das, U.K.; Tey, K.S.; Seyedmahmoudian, M.; Mekhilef, S.; Idris, M.Y.I.; Van Deventer, W.; Horan, B.; Stojcevski, A. Forecasting of photovoltaic power generation and model optimization: A review. *Renew. Sustain. Energy Rev.* **2018**, *81*, 912–928. [CrossRef]
37. Zhong, J.; Liu, L.; Sun, Q.; Wang, X. Prediction of Photovoltaic Power Generation Based on General Regression and Back Propagation Neural Network. *Energy Procedia* **2018**, *152*, 1224–1229. [CrossRef]
38. Icel, Y.; Mamis, M.S.; Bugutekin, A.; Gursoy, M.I. Photovoltaic Panel Efficiency Estimation with Artificial Neural Networks: Samples of Adiyaman, Malatya and Sanliurfa. *Int. J. Photoenergy* **2019**, *2019*, 6289021. [CrossRef]
39. Son, J.; Park, Y.; Lee, J.; Kim, H. Sensorless PV Power Forecasting in Grid-Connected Buildings through Deep Learning. *Sensors* **2018**, *18*, 2529. [CrossRef]
40. Eseye, A.T.; Zhang, J.; Zheng, D. Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and Meteorological information. *Renew. Energy* **2018**, *118*, 357–367. [CrossRef]
41. Paterova, T.; Prauzek, M. Estimating Harvestable Solar Energy from Atmospheric Pressure Using Deep Learning. *Elektron. Elektrotech.* **2021**, *27*, 18–25. [CrossRef]
42. Kusznier, J.; Wojtkowski, W. Impact of climatic conditions on PV panels operation in a photovoltaic power plant. In Proceedings of the 2019 15th Selected Issues of Electrical Engineering and Electronics (WZEE), Zakopane, Poland, 8–10 December 2019. [CrossRef]
43. Munir, M.A.; Khattak, A.; Imran, K.; Ulasyar, A.; Khan, A. Solar PV Generation Forecast Model Based on the Most Effective Weather Parameters. In Proceedings of the 2019 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), Swat, Pakistan, 24–25 July 2019. [CrossRef]
44. WS501-UMB Smart Weather Sensor. Available online: https://www.lufft.com/products/compact-weather-sensors-293/ws501-umb-smart-weather-sensor-1839/ (accessed on 1 February 2023).
45. Machine Learning in Python. Available online: https://scikit-learn.org/stable/index.html (accessed on 1 February 2023).
46. Geron, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, Concepts, Tools, and Techniques to Build Intelligent Systems*; O'Reilly Media Inc.: Sevastopo, CA, USA, 2019.
47. Albon, C. *Machine Learning with Python Cookbook, Practical Solutions from Preprocessing to Deep Learning*; O'Reilly Media Inc.: Sevastopo, CA, USA, 2018.
48. Murtagh, F. Multilayer perceptrons for classification and regression. *Neurocomputing* **1991**, *2*, 183–197. [CrossRef]

49. Xiang, W.; Xu, P.; Fang, J.; Zhao, Q.; Gu, Z.; Zhang, Q. Multi-dimensional data-based medium- and long-term power-load forecasting using double-layer CatBoost. *Energy Rep.* **2022**, *8*, 8511–8522. [CrossRef]
50. Diebold, F.X.; Mariano, R.S. Comparing predictive accuracy. *J. Bus. Econ. Stat.* **1995**, *13*, 253–264.
51. Harvey, D.; Leybourne, S.; Newbold, P. Testing the equality of prediction mean squared errors. *Int. J. Forecast.* **1997**, *13*, 281–291. [CrossRef]
52. Huber, P.J. Robust Estimation of a Location Parameter. *Ann. Math. Stat.* **1964**, *35*, 73–101. [CrossRef]
53. Kharlova, E.; May, D.; Musilek, P. Forecasting Photovoltaic Power Production using a Deep Learning Sequence to Sequence Model with Attention. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 17–24 July 2020; pp. 1–7. [CrossRef]
54. Yang, G. Asymptotic tracking with novel integral robust schemes for mismatched uncertain nonlinear systems. *Int. J. Robust Nonlinear Control* **2023**, *33*, 1988–2002. [CrossRef]
55. Khan, Z.A.; Hussain, T.; Baik, S.W. Dual stream network with attention mechanism for photovoltaic power forecasting. *Appl. Energy* **2023**, *338*, 120916. [CrossRef]