



Research Article

Forecasting of Energy Production for Photovoltaic Systems Based on ARIMA and ANN Advanced Models

Laurentiu Fara ^{1,2}, Alexandru Diaconu,¹ Dan Craciunescu ¹ and Silvian Fara¹

¹Department of Physics, Faculty of Applied Sciences, Polytechnic University of Bucharest, 060042, Romania

²Academy of Romanian Scientists, 030167 Bucharest, Romania

Correspondence should be addressed to Laurentiu Fara; lfara@renerg.pub.ro

Received 21 April 2021; Revised 2 June 2021; Accepted 15 July 2021; Published 9 August 2021

Academic Editor: Alberto Álvarez-Gallegos

Copyright © 2021 Laurentiu Fara et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Accurate forecasting of solar energy is essential for photovoltaic (PV) plants, to facilitate their participation in the energy market and for efficient resource planning. This article is dedicated to two forecasting models: (1) ARIMA (Autoregressive Integrated Moving Average) statistical approach to time series forecasting, using measured historical data, and (2) ANN (Artificial Neural Network) using machine learning techniques. The main contributions of the authors could be synthesized as follows: (1) analysis and discussion of the experimental and simulated results regarding solar radiation forecast, as well as energy production prediction and forecasting based on ARIMA and ANN models for two case studies: (a) laboratory BIPV system developed at the Polytechnic University of Bucharest and (b) large PV park placed in a specific site of the south of Romania. A variability index of solar radiation was introduced for the model improvement; (2) comparison between the ARIMA and ANN results to highlight the ARIMA model which is more efficient than the ANN one; (3) optimized method defined by the GMDH model (Group Method of Data Handling) proposed to provide a software program for calculation of the PV energy production.

1. Introduction

Solar energy is one of the most promising sources of power generation for residential, commercial, and industrial applications [1–5]. Photovoltaic solar energy, based on PV systems, has increased progressively in recent years due to its advantages of being abundant, inexhaustible, clean, and environmentally friendly [6–8].

Reliable and accurate forecasts play a key role in improving PV solar power plants [1–3, 9–12]. The main challenging problem in the production of solar energy is the intermittent electricity generation using PV systems, due to weather conditions. A variation in temperature and solar irradiance can have a profound impact on the quality of electricity production, leading to a decrease of more than 20% in the PV energy production provided by real PV installations. This limits the integration of PV systems into the grid. Therefore, an accurate short-term forecast of photovoltaic energy is very useful for the efficient daily/hourly management of electricity production and storage in the grid [13].

Accurate forecasting of solar energy is essential for PV plants, in order to facilitate their participation in the energy market and for efficient resource planning [1–3]. Various methods have been reported in the literature for the forecast of PV energy [2, 12]. These methods can be divided into four classes: (i) statistical approaches to time series forecasting, using measured historical data (ARIMA) [14]; (ii) machine learning techniques, in particular Artificial Neural Networks (ANN), based on machine learning approaches [15]; (iii) physical models based on numerical weather prediction and satellite imagery [16]; and (iv) hybrid approaches that represent a combination of the first three methods [17].

The beauty of the ARIMA model lies in its simplicity and can be applied only to stationary time series [14, 18, 19]. Therefore, our data from the time series, which are seasonal and nonstationary, are transformed into a stationary series for the application of the ARIMA model. The model can be developed using sophisticated statistical techniques [20]. The optimal approach is selected and validated using the Akaike Information Criterion (AIC) and the residual sum of squares (SSE).

Another statistical model is based on the seasonal analysis of the ARIMA time series (SARIMA) and is further improved by incorporating short-term solar radiation predictions derived from the NWP (numerical weather prediction) model [19]. Such a model can be used in energy forecast simulation software for buildings, such as Energy Plus, TRNSYS, or eQuest, both for existing buildings and in the design of new buildings [20, 21].

On the other hand, data-based models are determined by mathematical models and measurements. They do not require detailed knowledge of buildings or equipment. Their forecasts are mainly based on historical data, which are available using the control systems implemented in buildings for (1) building automation systems (BAS) and (2) building energy management systems (BEMS). The accuracy of these models depends on the quality of the selected forecast model, as well as the quality and quantity of available data. Such models are easy to adapt to changing conditions and are relatively easy to be used. In most cases, the relationship between the predicted variable and its physical function is not clearly established [17, 22].

In recent decades, researchers have been dedicated to improving efficiency and building energy use through various techniques and strategies. Energy forecasting in an existing building is essential for a variety of applications, such as demand response, fault detection and diagnosis, predictive model control, energy optimization, and management. Energy estimation models are a promising field of research, and this is true given the new advances in artificial intelligence and machine learning. Such models have been widely applied to both building energy systems and HVAC (heating, ventilation, and air conditioning) systems, as they can help with a variety of tasks. Most conventional approaches for solar energy forecasting are limited to discover the data correlation but are not able to analyze them in depth and discover relevant information. With many data in the modern power system, the use of conventional approaches is not appropriate to guarantee an accurate forecast [15, 20, 23].

Recently, Deep Learning (DL) approaches have emerged as powerful machine learning tools that allow shape recognition and regression analysis, as well as prediction applications [24–26]. DL approaches are becoming increasingly popular due to their real ability to describe time series data dependencies. Recently, many Deep Learning models have been proposed, including Boltzmann machines, Deep Belief Networks (DBNs), and Recurrent Neural Networks (RNNs). RNN is a type of neural network that exploits the sequential nature of input data. RNNs are used to model time-dependent data and give good results, which have proven to be successful in several application fields [15, 23, 27]. Short-Term Memory Networks (STMN) are a type of RNNs, which are able to cope with the storage of information for much longer periods of time [16, 28–30]. They are considered to define one of the most used RNN models for predicting data from time series, which perfectly correspond to the problems of PV energy production.

Zhao and Magoulés published a review study in 2012, focusing on the main approaches to energy prediction and forecasting in buildings [20]. Specifically, the authors

compared physical models with those based on machine learning, as well as with statistical ones. The authors noted that machine learning-based models had the highest accuracy and flexibility, especially compared to statistical models. Vector support machines have been developed to outperform ANN models; one of the areas recommended for future investigations is that of optimized applications.

In 2017, Wang and Srinivasan [23] explored the use of models based on artificial intelligence (AI) and overall models for predicting and forecasting the energy use of buildings. The authors provided a breakdown of how artificial intelligence as a whole was applied to predict the energy of the buildings. Most AI-based works were applied to an entire building load with hourly data. The authors also explored how overall methods were applied in predicting the energy of buildings. It was noted that such assemblies have been widely applied to fields outside the energy of buildings, and the results showed improved performance compared to unique forecast models.

Akhter et al. [31], Van Deventer et al. [32], Das et al. [33], and Seyedmahmoudian et al. [34] analyzed and compared different forecasting models for PV energy production, including the hybrid ones; their strengths and weaknesses were highlighted.

At the same time, the benefits of model optimization were discussed.

The main objective of this study was to apply the advanced techniques of ARIMA and ANN models for two cases: (1) a lab BIPV system and (2) a large PV park placed in the south of Romania. The obtained simulated and experimental results were compared and allowed to have interesting information regarding the energy production forecast on short term, as well as on long term; it could be very useful for improved performance of energy prediction and forecasting for buildings based on statistical approaches and artificial intelligence tools.

2. Conceptualization and Methodology

2.1. The Advanced Statistical ARIMA Model. The ARIMA (Autoregressive Integrated Moving Average) model establishes a remarkable method used in short-term solar radiation forecasting [14]. The functionality of the ARIMA model is represented in Figure 1 [35, 36]. The software used to make the forecast, respectively, IBM SPSS, is a predictive analysis program, which offers a wide range of statistical procedures: linear regressions, Monte Carlo simulations, geospatial analysis, etc.

The model chosen in this article for the solar radiation forecast is of the ARIMA (1,0,0) or (1,1,0) type. This choice is due to the Akaike Informational Criterion (AIC) [19]. The AIC is used to evaluate the model, depending on the entered data; the aim is to determine the quality of the statistical model on a dataset, establishing whether this model is more efficient than other models. The ARIMA model provides a relative estimate of the information used to determine the processes that could generate the data.

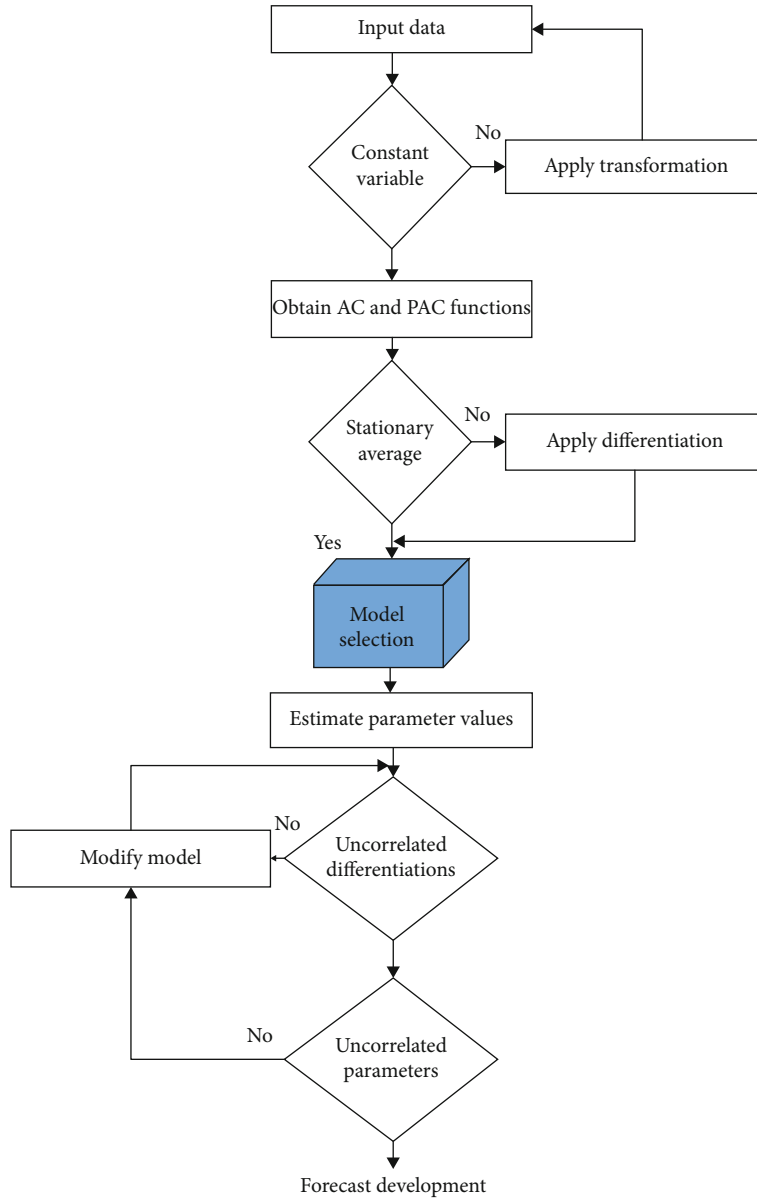


FIGURE 1: Operation of the ARIMA model for the energy production forecast of a photovoltaic system; AC and PAC autocorrelation functions are used.

The Akaike statistical criterion for the ARIMA model can be calculated based on the following relation [35, 36]:

$$AIC = 2k - 2 \ln(L), \quad (1)$$

where L represents the maximum value given by an estimation function (MLE (Maximum Likelihood Estimation)) and k represents the number of estimated parameters. The minimum value of the AIC result defines that the corresponding model is the most efficient.

The SPSS program [35, 36] automatically calculates the AIC, which is a parameter for model fitting.

If the forecast process contains seasonal fluctuations, as in this case, the process becomes SARIMA $(p, d, q) (P, D, Q) s$,

where p is the order of the AR process, d represents the differentiation term, q is the order of the moving average, P represents the order of AR seasonal processes, Q represents the MA order, D is the order of seasonal differentiation, and s represents the length of the seasonal period [19].

The peculiarities of the model used in this forecast are highlighted by defining a backshift operator B of the time series [9]: $Bz_t = z_{t-1}$, where z_t and z_{t-1} are two consecutive observations of the time series. Then, B^j can be defined as $B^j z_t = z_{t-j}$. Using standard notation, $\phi(B)$ is the autoregression operator, represented as a polynomial in the backshift operator: $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$, and $\theta(B)$ the mobile operator is represented by a polynomial in the operator backshift: $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 \dots - \theta_q B^q$. Having d —the

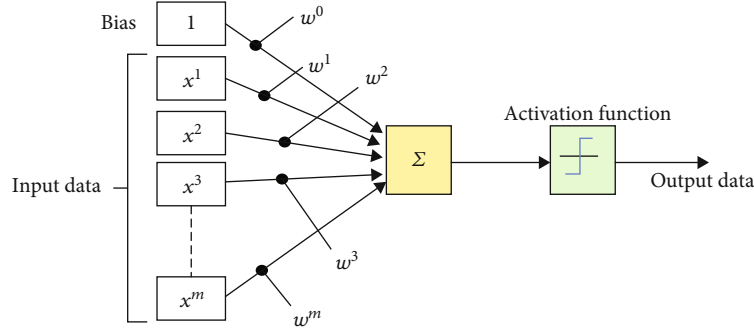


FIGURE 2: Functionality of the ANN model for the energy production forecast of a photovoltaic system.

degree of nonseasonal differentiation and at the white noise, from a mathematical point of view, the ARIMA model could be written as follows:

$$\phi(B)W^d z_t = \theta(B) a_t, \quad (2)$$

where W is the differentiation operator which is the equivalent of result 1- B .

We call this model ARIMA (p, d, q) , where p , d , and q are process orders. In our model, the self-regression delay is determined separately for each month represented by the number of lighting hours, and the hours with zero irradiation (night hours) are neglected. As a consequence, we have used the ARIMA model $(x, 0, 0)$, where x represents the number of hours in a day with solar radiation of those months (for example, for January it is 10, and for July it is 16).

ARIMA $(x, 0, 0)$ was used to avoid the SARIMA model (seasonal model) [19, 35, 36]. The seasonality component is taken into account by the autoregression component, where the autoregression is equal to the period. 10 in January represents the fact that the morning value from 8 o'clock correlates with the value of the next day from 8 o'clock, which is in the 10th place from the first value, since we have used only 10 values each day. In other words, in the used experimental data string, there is an autocorrelation of 10 orders.

2.2. The Advanced ANN Model Based on Artificial Intelligence (AI). Neural networks are a set of processing elements that have been developed separately from standard regression techniques. The ability to “mimic” natural intelligence by learning from experience makes this technique very attractive for solar radiation predictions. Functioning like a biological brain, a neural network is made up of a large number of interconnected neurons. Two main classes of neural network architecture can be identified, namely, (1) *the architecture with the propagation of information from the input data to the output data*, also called *feed-forward architecture*, and (2) *the architecture of recurrent networks*. The functionality of the neural network model is represented in Figure 2 (see [15, 25]).

The ANN model is also used to estimate global solar radiation at an hourly interval, based on the following meteorological parameters: air temperature (T), relative humidity

(RH), atmospheric pressure (P), wind speed (W), cloudiness (C), and hours of the day (H):

$$R_G = f(T, RH, P, W, C, H). \quad (3)$$

The main parameters of the ANN model used in this study are as follows [35, 36]: the number of neurons in the hidden layer is 1, the initial weight is 0.3, the learning rate is 0.3, and it used the Sigmund logarithmic function. The 10 days of meteorological data are included in the learning process.

3. The PV Energy Production Forecast for a Lab BIPV (Building Integrated Photovoltaic) System

3.1. Results and Discussion: Solar Radiation and Energy Production Forecast Based on the ARIMA and ANN Models for a Lab BIPV System

- (a) The BIPV system from the Polytechnic University of Bucharest is a demo system connected to the grid; it was put into operation in July 2008 (see Figure 3) [11]. It consists of six PV panels, an inverter, equipment for monitoring and storing data, and a laboratory station for monitoring meteorological parameters. The power delivered by the PV panels is 615 Wp; they are semitransparent and have a power of 85 Wp each (type 1), respectively, and 120 Wp each (type 2). The system consists of three modules of type 1 and three modules of type 2

Also, to complete the BIPV system and make it functional, there is an inverter that makes the conversion between dc and ac, as well as monitoring and data storage equipment. The inverter is of the Sunny Boy SB700 type and achieves a high efficiency of conversion of direct current into alternating one, which it then delivers to the electricity distribution grid.

With the mentioned technical resources of the studied BIPV system, data regarding the brightness and performance of the PV system were collected, in terms of power (Pac) for a period of five days. Three days were considered for the forecasting process, and the other two days were used to validate the results. Figure 4 illustrates the



FIGURE 3: BIPV system developed within the laboratory of nanostructures, solar cells, and PV systems from the Department of Physics, Faculty of Applied Sciences, Polytechnic University of Bucharest (inside/outside view).

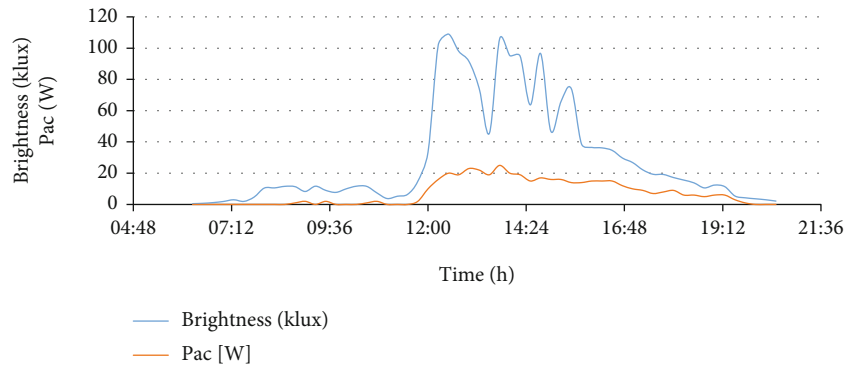


FIGURE 4: Time evolution of power (W) and brightness (klux) parameters.

evolution of the power and brightness parameters, over a period of one day of the chosen interval.

The detailed comparison, which provides data every 15 minutes, allows us to obtain important information about system performance. An unexpected decrease in power can be seen in the second half of the day. The explanations can be the following:

- (i) The photovoltaic window is oriented to the east, so that the incident solar radiation is maximum in the first part of the day
- (ii) The system is shaded in the second part of the day, due to the architecture of the building

The solar radiation data, obtained with the help of the lab weather station, were used for the short-term solar radiation forecast, using various methods, in the location where the BIPV system is placed [11].

The results of the ARIMA model are represented by the input dataset for 13.06-15.06 2012, and the model is checked by the data for 16.06-18.06 2012 [11, 35, 36] (see Table 1).

TABLE 1: The accuracy of the ARIMA model.

Statistical fitting	Average
R^2	0.811
RMSE	18,367.868
MAPE	199.410
MaxAPE	3,838.473
MAE	10,359.011
MaxAE	54,998.512

R^2 represents how much of the variation of the dependent variable is determined by the estimated quantity, RMSE (Root Mean Squared Error) represents the difference between the predicted values and the observed ones, MAPE (Mean Absolute Percentage Error) represents the accuracy of the forecasted model, and MaxAPE (Maximum Absolute Percentage Error) is the largest predicted error, expressed as a percentage. This error is useful to determine the efficiency of the forecast. MAE (Mean Absolute Error) represents the degree of variation of the measured data compared to the forecast. MaxAE (Maximum Absolute Error) is the largest error in the forecast, like MaxAPE, also useful in determining the efficiency of the forecast at the point with maximum error.

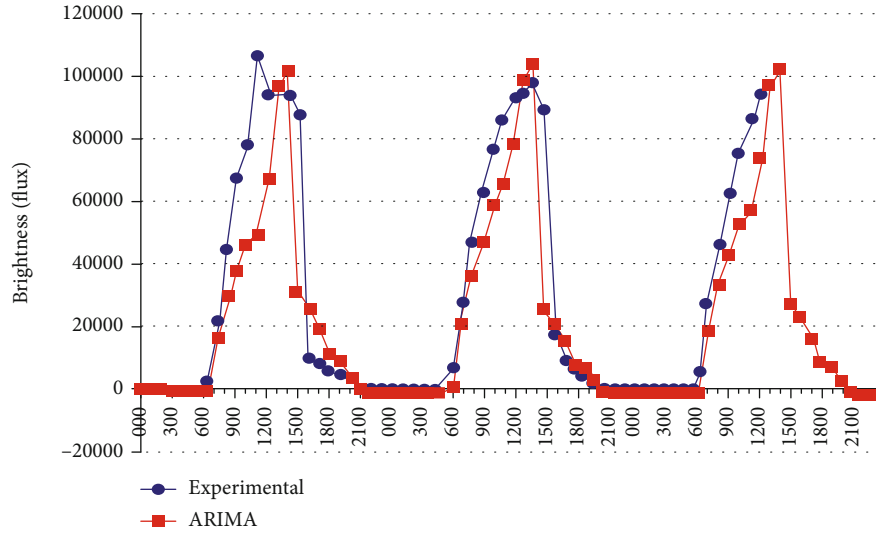


FIGURE 5: Adjusting the forecasted ARIMA data with the measured ones.

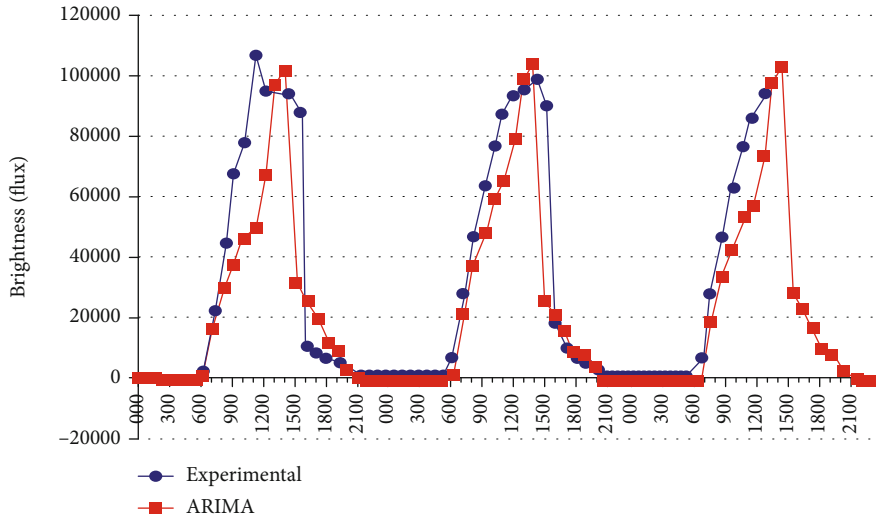


FIGURE 6: Adjusting the forecasted ANN data with the measured ones.

Figure 5 shows the adjustment of the ARIMA model between the forecasted and measured data.

- (b) Taking into account five meteorological variables, namely, air temperature, relative humidity, atmospheric pressure, wind speed, and sunshine duration, the forecast was made for 16.06.2012, using the method of neural networks (ANN) (see Figure 6) [11, 35, 36]

A statistical relationship is defined between global solar radiation and energy production, taking into account the technical parameters of the system. The losses of the system due to its location were also taken into account.

The values of the predicted data for P_{ac} were compared with the measured ones; this situation is represented in Figure 7. The error calculated between the measured and the predicted values of P_{ac} using the ANN model is 8.89%.

3.2. Improving the Forecast Quality by the Variability Index of Solar Radiation. The meteorological data used in the study of the solar radiation forecast for a lab BIPV system come from the Archive of the National Meteorological Authority [9]. The parameters were measured in the Bucharest-Afumatzi weather station ($44^{\circ} 30'N$ $26^{\circ} 13'E$, $h = 91$ m, 8 km away from Bucharest, WMO code 15421) and represent data obtained every hour during 2008-2009. The meteorological tools and observations correspond to the WMO instructions (Global Observation System Manual, 2010). The meteorological variables included in this study are global solar irradiance in kJ/m^2 , average air temperature in $^{\circ}\text{C}$ at a height of two meters, atmospheric pressure in mbar, relative humidity, wind speed in m/s, and cloudiness in oktas.

At the Bucharest-Afumatzi meteorological station, the global solar irradiance is automatically measured on a horizontal surface, with the pyranometer CM11 Kipp&Zonen-

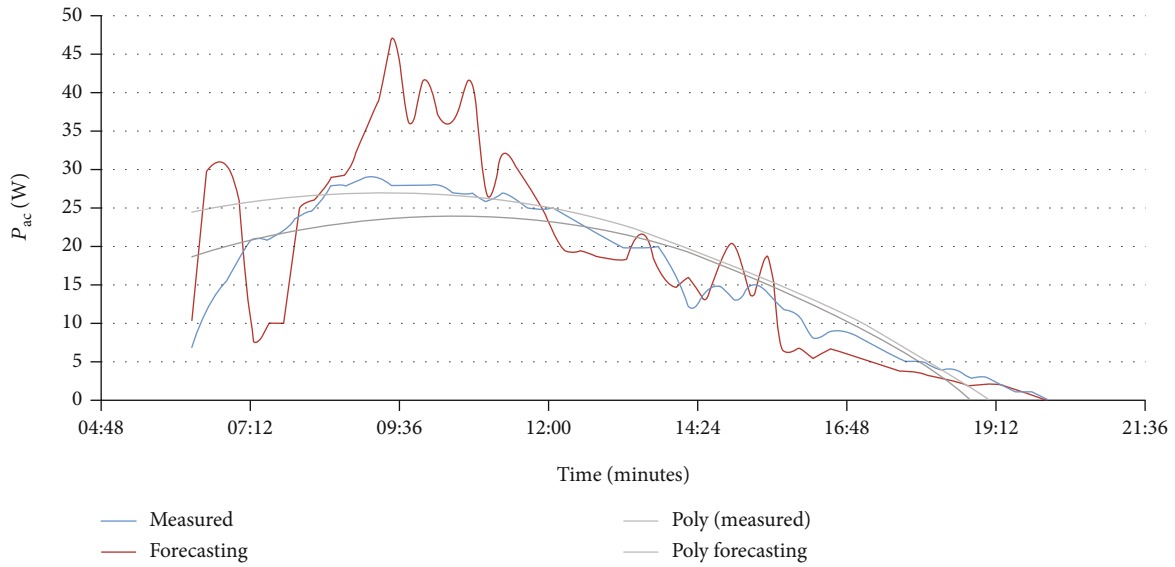


FIGURE 7: Comparison between measured and forecasted values of P_{ac} for 17.06.2012.

Delft. The pyranometer sensor is a thermopile protected from external effects by a double glass dome. It has the following features: a response time (95%) of 15 seconds, spectral range 305-2800 nm, sensitivity 4-6 $\mu V/W/m^2$, nonlinearity $< \pm 0.6\%$ ($< 1000 W/m^2$), this being installed inside the meteorological platform, facing south, at a height of 2 meters.

In the solar radiation forecast of the BIPV lab system from the Polytechnic University of Bucharest, both studied models were used, namely, ARIMA and ANN [9, 11]. The error of the solar radiation forecast depends on its daily variation. At the same time, the daily variation of solar radiation showed an important correlation with clouding; thus, cloud information will be needed to increase the accuracy of the forecast.

However, this type of calculation requires a large amount of input data and a huge computational ability. In most cases, acceptable results cannot be obtained. In order to take into account the influence of the clouding factor but at the same time reduce the number of input data, it is found that the affected situations are determined by an empirical index, as an input parameter, namely [11],

$$V_{DR} = \frac{\sigma_{DR}}{\chi_{MR}}, \quad (4)$$

where V_{DR} is the variability index of the solar irradiance, χ_{MR} is the average of the monthly global solar irradiance, and σ_{DR} is the standard deviation of the daily global solar irradiance.

In fact, this index quantifies the variation of solar irradiance for a given day, compared to the general situation in a given month; the variation of solar irradiance in the previous days is taken into account indirectly by the monthly average. This daily index provides more detailed information, as well as the classification based on synoptic situations, the latter variant including a longer period with various clouding factors. By default, the index provides an important correlation with the clouding factor, $r = 0.66$ in

2008 and $r = 0.61$ in 2009, both being considered at a probability level of 9.99%. The average number of days with cloudiness, which denotes $V_{DR} < 0.5$, is 7.93 in 2008 and 7.69 in 2009, the standard deviation being 0.79 and 0.95, respectively. At the same time, the average number of days with cloudiness, showing a $V_{DR} > 0.5$, is 4.49 in 2008 and 4.57 in 2009, the standard deviation being 2.27 and 2.53, respectively. As a consequence, the value of V_{DR} indicates changes in global solar radiation related to the following situations:

- (a) $V_{DR} < 0.5$: fog or cloud cover situations
- (b) $V_{DR} > 0.5$: situations with partly cloudy or clear sky

In our study, four different situations are analyzed based on the cloudiness quantified by the V_{DR} index, and then, the forecasts of solar radiation values are developed independently of each other. The main reason for this classification is to reduce the synoptic situations represented by the two classes of cloud cover types.

The four situations include changes in the amounts of cloudiness on the previous day and on the forecast day and are as follows (see Figure 8) [9]:

- (1) Cloudy/foggy day on the forecast day and partly cloudy/clear sky on the previous day $V_{DR} < 0.5$ (previous day $V_{DR} > 0.5$)
- (2) Cloudy/foggy situation on the forecast day and cloudy/foggy sky situation on the previous day $V_{DR} < 0.5$ (previous day $V_{DR} < 0.5$)
- (3) Partly cloudy/clear sky on the forecast day and cloudy/foggy sky on the previous day $V_{DR} > 0.5$ (previous day $V_{DR} < 0.5$)
- (4) Partly cloudy/clear sky on the forecast day and partly cloudy/clear sky on the previous day $V_{DR} > 0.5$ (previous day $V_{DR} > 0.5$)

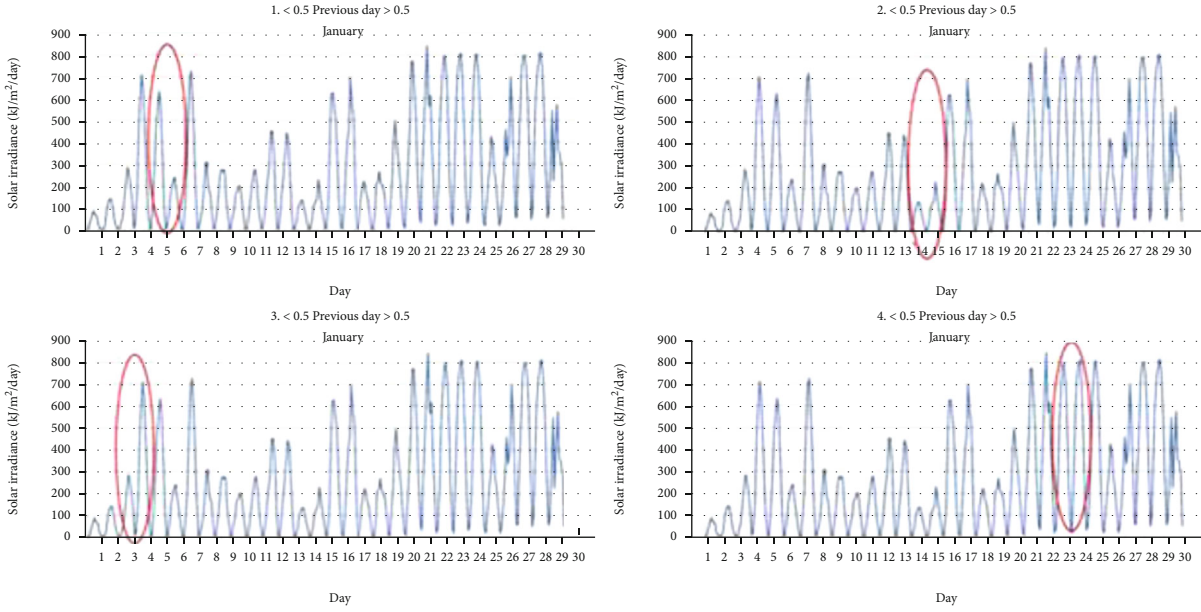


FIGURE 8: The four cloudy change situations quantified by the V_{DR} index, based on solar irradiance data measured at the Bucharest-Afumatzi Meteorological Station ($\text{kJ}/\text{m}^2/\text{day}$).

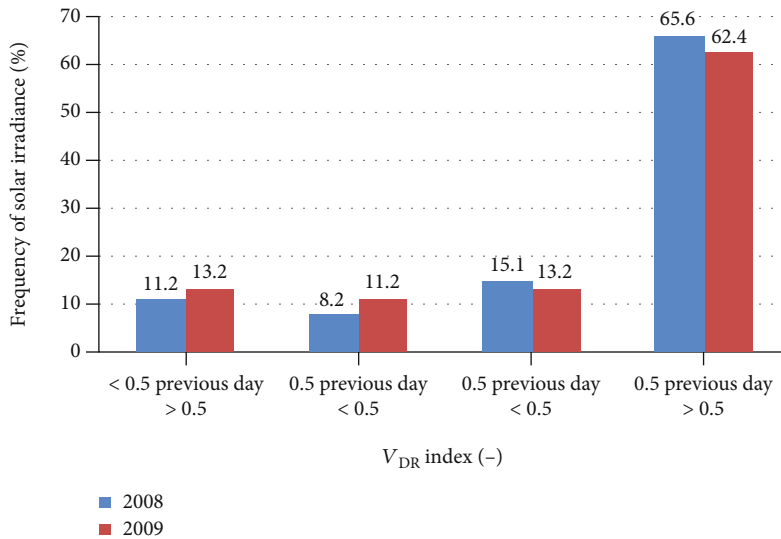


FIGURE 9: Relative frequency of solar irradiance/clouding situations quantified by the V_{DR} index, using the Bucharest/Afumatzi weather station.

We have analyzed the accuracy of the solar irradiance forecast developed in the four situations using a time series for the ARIMA and ANN models in order to quantify the significance of the forecast error in each case [22, 27]. Considering the frequency of situations during the investigated period, the most common situation is the fourth, which represents 64% of cases. The least common situation is the second, with a frequency of occurrence of only 8-11% of cases. The results of the analysis are represented in Figure 9. By this method, the errors of daily solar radiation forecasts can be quantified even during the forecasting process, separating only three sit-

uations, namely, the cloudy situation, the fog situation, and the one with partly overcast sky/clear sky conditions.

3.3. The Solar Radiation Forecast on Short and Long Term

3.3.1. Short-Term Solar Radiation Forecast with Application for a BIPV Lab System. Using the forecast models discussed in Section 3.1, respectively, ARIMA and ANN, we have made a short-term forecast, taking into account the synoptic situations of the days quantified by the V_{DR} index. For each situation, a random example is analyzed for each month, in 2008

TABLE 2: rRMSE values in the case of the four quantified situations of the V_{DR} index. Processed data from the 2008 year for a forecasting period of 10 days.

Month	<0.5 (the day before) >0.5		<0.5 (the day before) <0.5		>0.5 (the day before) <0.5		>0.5 (the day before) >0.5	
	ARIMA	ANN	ARIMA	ANN	ARIMA	ANN	ARIMA	ANN
1	168.6	194.7	35.6	56.9	79.2	90	9.2	15.9
2	197.9	191.6	58.6	40.5	103.7	115.6	9.9	21.8
3	232.7	209.1	53.3	77.6	47.6	64.4	39.4	42.6
4	191.4	184.4	101.9	77.4	67.4	92.1	22.2	26.3
5	136.6	127	53.5	59.9	51	66.5	12.9	17.2
6	101	95.3	46.8	66.7	51	69.4	16.3	22.9
7	132.3	134.1	85.1	72.5	44.8	69.2	25.6	36.4
8	No data	No data	No data	No data	No data	No data	23.8	20.9
9	No data	No data	No data	No data	No data	No data	6.9	10.9
10	423.3	347.3	No data	No data	100.4	108	12.4	32.6
11	105.9	68.2	70.9	26.6	99.7	105.5	30	24.5
12	136.0	183.2	58.7	58.2	87.4	89.7	35.6	48.2
Average	182.6	173.5	62.7	59.6	73.2	87	20.4	26.7

TABLE 3: rRMSE values in the case of four quantified situations of the V_{DR} index. Processed data from the 2009 year for a forecasting period of 10 days.

Month	<0.5 (the day before) >0.5		<0.5 (the day before) <0.5		>0.5 (the day before) <0.5		>0.5 (the day before) >0.5	
	ARIMA	ANN	ARIMA	ANN	ARIMA	ANN	ARIMA	ANN
1	291.1	356.1	55.2	65.1	57	76.5	25.7	33.3
2	67.4	111.3	56.2	42.2	86.2	104.8	28.9	18.8
3	299.5	367.3	51	40.6	69.6	100.2	26.8	23.8
4	141.3	108.5	131.8	50.2	103	113	20.8	17
5	237.3	232	60.1	70	55.2	87.5	9.8	14
6	189.1	168.3	No data	No data	69.5	84.6	15	24.2
7	207.6	204.1	No data	No data	87.7	108.5	21.5	28.8
8	327.9	233	No data	No data	61.6	67.7	6.2	19.3
9	469.9	440.4	232.3	66.5	75.8	84.6	15	29
10	No data	No data	131.4	35.4	83.9	94.1	53.4	20.3
11	443.8	551.1	98.7	53	82.1	95.6	16.7	24
12	344.6	334.5	54.5	50.2	87.1	113.8	51.4	33.5
Average	274.5	282.4	96.8	52.6	76.6	94.2	24.3	23.8

and 2009. For both models, the analyzed time series is 10 days. The accuracy of the forecasts is quantified by calculating the mean square relative error, using the relation:

$$\text{rRMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (F_t - A_t)^2} * \frac{100}{G_n}, \quad (5)$$

where F_t represents the predicted value, A_t represents the current value (measured), n is the number of predicted points, and G_n represents the daily average of the current (measured) values.

Tables 2 and 3 contain the results of the forecasts corresponding to the four situations. It is noted that the best results are obtained in the case of partly cloudy/clear sky

on the forecast day and partly cloudy/clear sky on the previous day, where the monthly rRMSE varies between 6.2 and 53%, and the annual one is less than 26.7%. In the case of cloudy/foggy sky on the forecast day and cloudy/foggy sky the day before, as well as partly cloudy/clear sky on the forecast day and partly cloudy/foggy sky on the previous day, the accuracy of the forecasts is relatively similar, ranging between 52.6% and 96.8%. In the case of cloudy/foggy sky on the forecast day with partly cloudy/clear sky on the previous day, the forecasts show a huge error of over 250%.

Comparing the two models, ARIMA and ANN, it was observed that the overall amplitude of errors is relatively similar in each case, but at the same time, the ARIMA model offers better results by 2.9% in the fourth case. We can

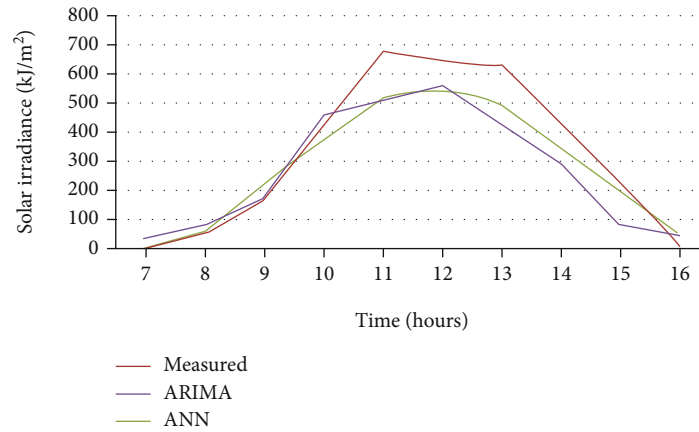


FIGURE 10: Solar radiation forecast for 20.01.2009 prepared with ARIMA and ANN models.

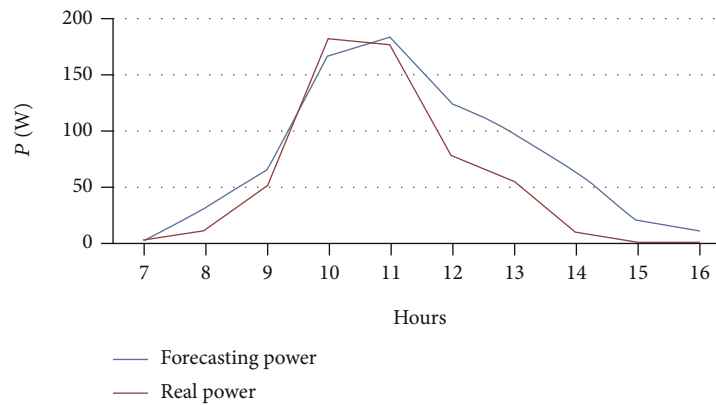


FIGURE 11: Predicted (forecast) and measured (real) power for 20.01.2009.

conclude that the accuracy of the forecasts depends closely on the daily variation of the solar radiation controlled by the clouding factor, and the weather series forecast provides acceptable results in the case of partly cloudy/clear sky on the forecast day with partly cloudy/clear sky the previous day (see [9, 15]).

Using the models of solar radiation forecast, and establishing the most frequent situation of the days from the meteorological point of view, the solar radiation forecast for January 20, 2009, was made and is represented in Figure 10.

With the help of this forecast, another forecast was elaborated, that of the power delivered by the lab BIPV system, represented in Figure 11.

The error calculated in this case was 26.9% and is consistent with the errors calculated for the solar radiation forecast, namely, 26.7%. This result is due to the improvement of the solar radiation forecast by integrating the data related to different meteorological parameters (temperature, cloud cover, etc.), as well as by separating the synoptic situations according to the variability index.

3.3.2. Long-Term Solar Radiation Forecast with Application for a BIPV Lab System. We have also analyzed the decade variation of solar radiation as long-term changes to be taken into account in solar energy applications. For the analysis of the change trend, we have applied the linear

regression model. The annual data come from the World Radiation Data Center database (<http://wrdc.mgo.rssi.ru>) and represent data corresponding to the period 1975-2006 for the Bucharest weather station. The linear trend is significant at a probability level of 95%, thus identifying an increase in solar radiation. The magnitude of the absolute changes is $36.5 (\pm 2.43) \text{ J/cm}^2 \text{ day}^{-1}/\text{decade}$ (Figure 12).

Based on the analysis of multiannual changes in solar radiation ($\text{J/cm}^2 \text{ day}^{-1}$), the power forecast of the BIPV system for 2013 was made (see Figure 13).

It was observed that the maximum power production of the BIPV system is achieved in May. The total production forecast for 2013 is 163 kW, a result that is in line with the power achieved in the previous year, of 157 kW. We can conclude that the best forecast result can be obtained in situation (4) of partly cloudy/clear sky on the forecasted day and partly cloudy/clear sky on the previous day $V_{\text{DR}} > 0.5$, rRMSE—the annual error being 22.3 in the case of ARIMA and 25.3 in the case of ANN. In this case, ARIMA gives the best result in line with Reikard [9, 23] who showed that at 60-minute time resolution, the ARIMA model forecast dominates all time series forecasting methods in four from the six test stations used in the study. In the other synoptic situations, the errors are large. Given the frequency of occurrence of the four synoptic

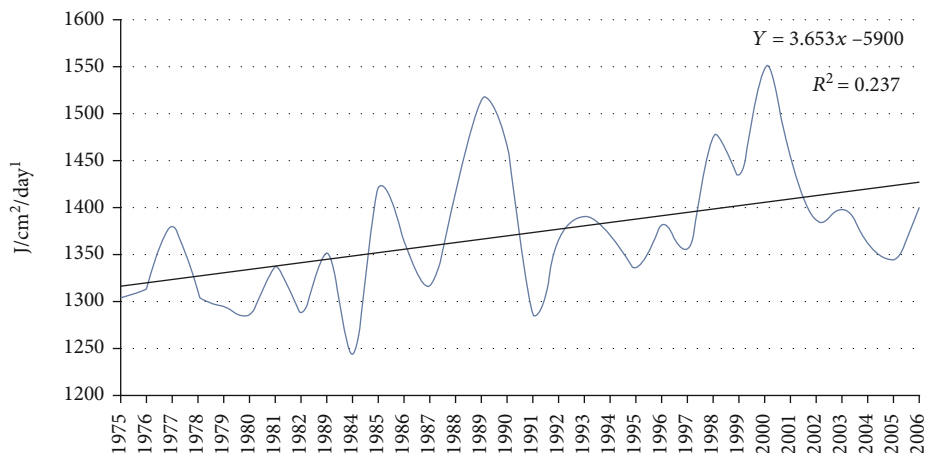


FIGURE 12: Multiannual variation of solar radiation in Bucharest during 1975-2006.

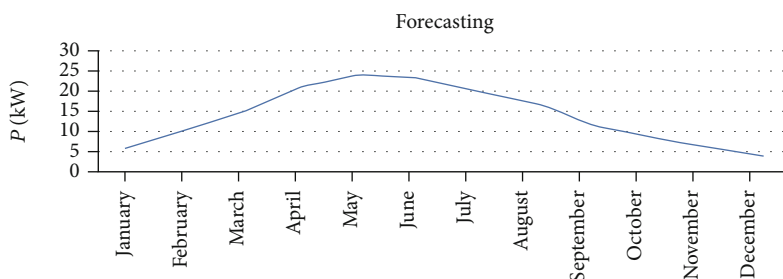


FIGURE 13: Power forecasted by the BIPV system for 2013, based on the estimation of multiannual solar radiation.



FIGURE 14: View of the PV park from Grojdibodu.

situations presented above, the situation in which we obtain the best results appears in 65.56% of the days of a year, so acceptable forecasts can be made in 65% of that year. In the case of the other situations, namely, (1), (2), and (3), physical and/or satellite models are required [24].

The results of the study contributed to the improvement of the forecasts of the photovoltaic systems, which will be used for the analysis of an experimental PV park [31, 32].

4. The Energy Production Forecast for a PV Park Using the ARIMA and ANN Models

4.1. Short Presentation of an Experimental PV Park. The experimental PV park from Grojdibodu (located in southern Romania) has in its composition 1931 strings, each string having 21 Suntech PV modules of 245 W. In total, there are 40,551 PV modules totaling an installed power of 9,934 kW. Solar

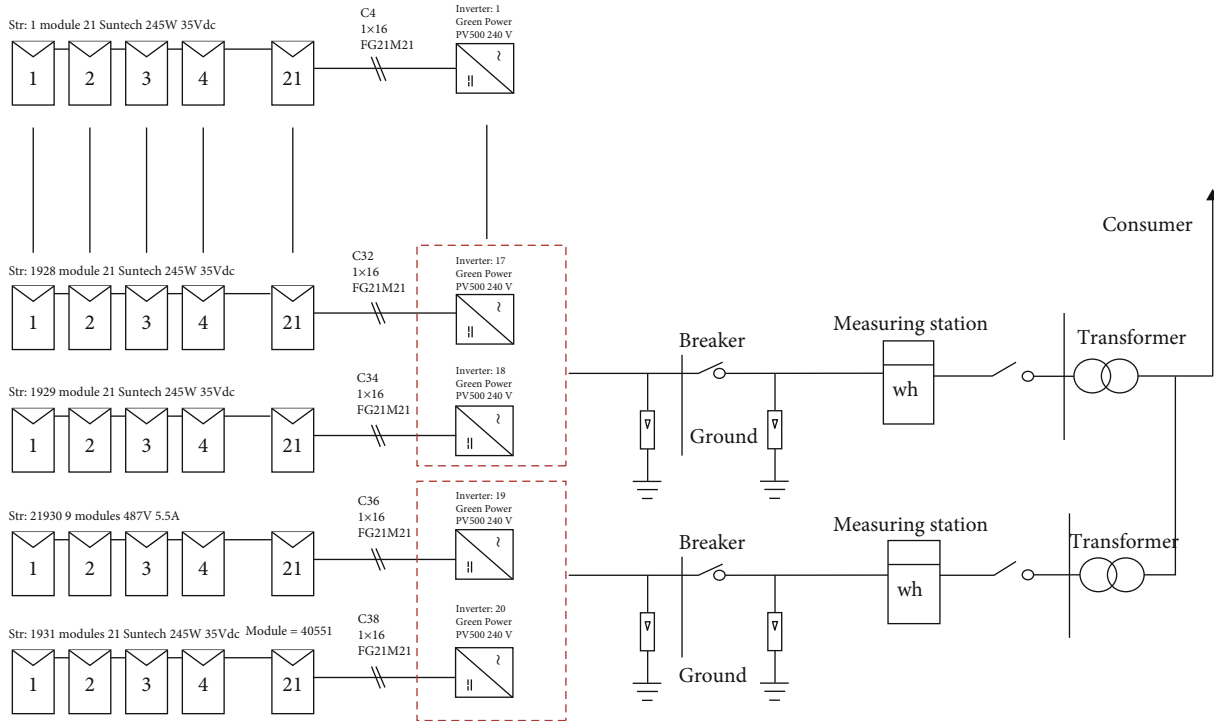


FIGURE 15: Electrical diagram of the studied PV park.

radiation data are acquired by 2 pyranometers and 10 calibration solar cells. The pyranometers are located in the meteorological station within the PV park, and the solar calibration cells are placed at each transformation center, thus being arranged on the entire surface of 33 ha of the park. The pyranometers and calibration cells are placed in a plane that has the same inclination to the horizontal as PV modules. There are a total of 20 Green Power PV500 inverters. The energy produced is measured at the connection point, and the furthest inverter is located at 400 m from this connection point. Each transformation center contains two inverters. The PV module is made of 72 polycrystalline silicon solar cells.

A view of the PV park from Grojdibodu is presented in Figure 14, and its electrical diagram is shown in Figure 15 [35, 36].

The main meteorological parameters of the Grojdibodu location defining the PV park are indicated in Table 4 [35, 36].

The technical parameters of the studied PV system/PV module, used in the forecast, are presented in Tables 5(a) and 5(b). Figure 16 shows the I - V (current-voltage) characteristics for the Suntech Power PV module used in the component of the studied PV park.

The technical performance indicators of the Grojdibodu PV system for one year are designated in Table 6.

4.2. Results and Discussion: Energy Production Forecast for the Experimental PV Park. The purpose of the forecast was to size an experimental PV park located in the south of Romania, at Grojdibodu. Thus, a comparison was made for the results obtained by numerical modeling, with those measured in the PV park, as well as a comparative analysis of the programs used in terms of their efficiency. For simulation, we

TABLE 4: Meteorological parameters for Grojdibodu location.

Solar radiation	Annual average temp.	Annual average wind speed	Albedo
1.78 kWh/m ² /day	11.5°C	3.2 m/s	0.2

TABLE 5

(a) Technical parameters of the PV park from Grojdibodu

Installed power	9,934 kW
Power of the PV module	245 W Suntech Power 245
Total number of PV modules	40,551
Number of PV modules per string	21
Number of strings	1931
Number of inverters	20—Green Power PV500
Used surface	33 ha

(b) Technical parameters of the PV modules—main components of the PV experimental park

PV module parameters	
Nominal efficiency	13.9098%
Maximum power	269.850 W _{cc}
Maximum voltage	35 V _{cc}
Maximum current	7.7 Acc
Open circuit voltage	44.5 V _{cc}
Short circuit current	8.2 Acc

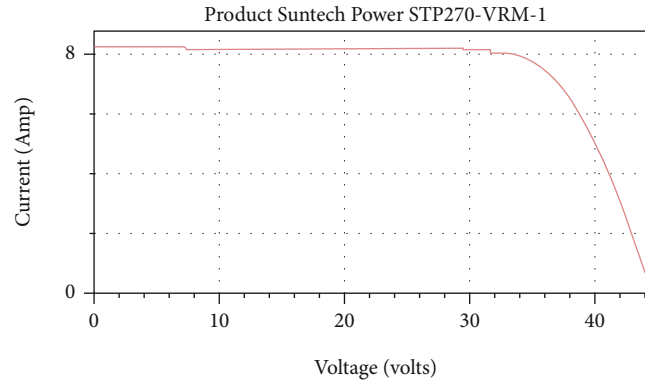
FIGURE 16: *I-V* characteristics of Suntech Power PV module.

TABLE 6: Performance indicators of the PV park from Grojdibodu.

Performance indicators (in the first year)	
Energy produced annually	13,069,573 kW
Capacity factor	15.0%
Energy efficiency	1,316 kWh/kW
Performance ratio	0.83

have used the data measured by the meteorological station within the PV park, considering the optimization of the PV system parameters, the increase of its energy efficiency, and the improvement of its global efficiency [26, 35, 36].

With the help of the ARIMA and ANN models, the forecast of surface solar radiation (SSR) was made, using data obtained from the PV park. Forecasts for 10 days have been developed based on the previous time series. The periods were chosen randomly, and the forecasts were checked for clear or partially clear days, which have the variability index of direct solar radiation $V_{DR} > 0.5$. The variability index refers to the classification of synoptic situations (clear or cloudy). These situations are not included in the forecasts made. The index helps to validate forecasts; thus, statistical forecasts are more effective in the case of partly cloudy/clear sky on the forecast day and partly cloudy/clear sky on the previous day $V_{DR} > 0.5$. In the case of “cloudy” situations (1), (2), and (3), the rRMSE has large errors; for this reason, the statistical/empirical method for forecasting is not useful, and physical models are needed to improve it.

The correlation between the variability index and nebulosity was calculated in order to validate the index. The index shows the cloudiness on that day, without using cloudy data. The aim was to obtain information about nebulosity from solar radiation data. The validation of the index was performed for a shorter period; for example, in cases where the index had a value > 0.5 , we checked whether the nebulosity was low or not, and vice versa. Following the verifications, the correlation was significant; as a result, we have used V_{DR} instead of nebulosity data.

Short-term solar radiation forecasts were developed using the ARIMA and ANN models. The efficiency of the forecasted time series is quantified by rRMSE. The

model was improved by separating the days with clear sky from the days with a high degree of cloudiness and using the variability index of solar radiation, V_{DR} . The separation of synoptic situations was performed on time series, not on predicted values. It has been established that the most common cases for the use of the V_{DR} index are days with clear sky on the forecast day and the day before. These days represent more than 75.4% of the total number of days of the year.

For the rest of the days when there are situations of cloudy sky or high fog, the forecast provides very large errors. The forecasted results conclude that the ARIMA model is more efficient than the ANN model. The statistics are significant in the case of ARIMA (1,0,14), this being the reason why this variant was chosen for analysis. Also, comparing the results of the forecast with the measured values, we noticed that the ARIMA model (1,0,14) is more efficient than the ARIMA (1,0,7). The ARIMA (1,0,7) and ARIMA (1,0,14) models were selected after performing several iterations. The statistical test used to identify the most significant model is Box Ljung Statistics. This test is applied in the case of ARIMA on the residues of a fitted model, not on the original series, and verifies that these residues do not show autocorrelation.

The results of the measurements and forecasted values for August and September 2013 are presented in Figures 17 and 18. The accuracy of the predictions is quantified using relative absolute mean error (rMAE) and relative mean square error (rRMSE):

$$\begin{aligned} \text{rMAE} &= \frac{1}{n} \sum_{i=1}^n |F_t - A_t| \times \frac{100}{G_n}, \\ \text{rRMSE} &= \sqrt{\frac{1}{n} \sum_{i=1}^n (F_t - A_t)^2 \times \frac{100}{G_n}}, \end{aligned} \quad (6)$$

where the following notations are introduced: F_t is the predicted value, A_t is the measured value, n is the number of predicted points, and G_n is the daily average of the measured values. Table 7 shows the measured and forecasted values for 10 days in August 2013, while Table 8 shows the same data in September 2013 for 10 days.

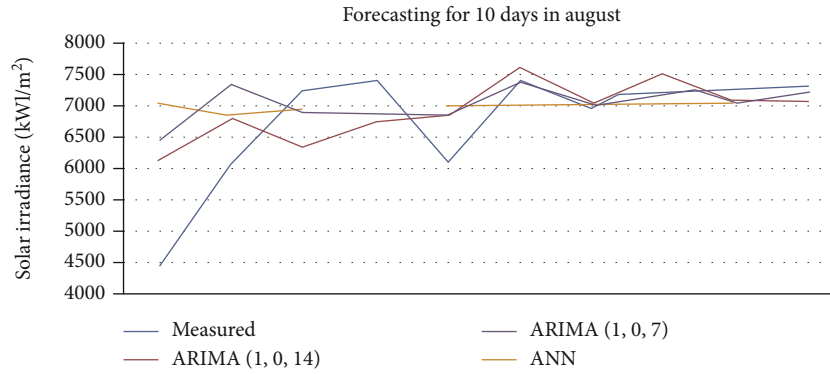


FIGURE 17: Solar irradiance forecast for Grojdibodu PV park, for 10 days, in August 2013, using the ARIMA and ANN models.

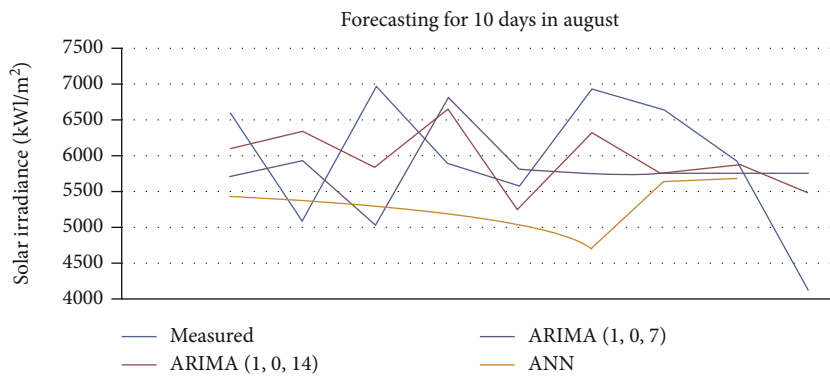


FIGURE 18: Solar irradiance forecast for Grojdibodu PV park, for 10 days, in September 2013, using the ARIMA and ANN models.

TABLE 7: Measured and forecasted values for 10 days, August 2013, based on ARIMA and ANN models. This table is reproduced from [35] under the permission of the editor.

Year	Month	Day	ARIMA (1,0,14) (Wh/mp)		ARIMA (1,0,7) (Wh/mp)		ANN (Wh/mp)	
			Measured	Forecasted	Measured	Forecasted	Measured	Forecasted
2013	7	31	4,451.3	6,138.76	4,451.3	6,459.48	4,451.3	7,044
2013	8	1	6,103.6	6,822.83	6,103.6	7,354.92	6,103.6	6,844
2013	8	2	7,240.3	6,353.43	7,240.3	6,916.88	7,240.3	6,963
2013	8	3	741.6	6,754.42	7,416.6	6,878.25	—	—
2013	8	4	6,091	6,878.99	6,091	6,894.17	6,091	7,024
2013	8	5	7,388.08	7,598.88	7,388.08	7,393.5	7,388.08	7,019
2013	8	6	7,101.34	7,006.33	7,101.34	6,974.19	7,101.34	7,002
2013	8	7	7,214.13	7,512.72	7,214.13	7,200.2	7,214.13	7,101
2013	8	8	7,285.4	7,055.09	7,285.4	7,069.23	7,285.4	7,098
2013	8	9	7,325.3	7,109.01	7,325.3	7,210.03	—	—

The two models offer good efficiency in terms of relative errors, but forecasts with errors below 40% are present only in the summer periods, especially in July, August, and September characterized by high values of solar radiation. In the case of low SSR, this methodology produces very large errors.

Based on the predicted values of solar radiation, the electrical power that would be inserted in the national power distribution grid could be calculated taking into account the

efficiency of the PV park. Most likely, a global relationship for estimating the energy generated was [35, 36]

$$E = A * H * PR, \tag{7}$$

where E represents the delivered energy (kWh), A represents the total surface area covered by the PV modules, H represents the average annual solar irradiance on the inclined

TABLE 8: Measured and forecasted values for 10 days, September 2013, based on ARIMA and ANN models. This table is reproduced from [35] under the permission of the editor.

Year	Month	Day	ARIMA (1,0,14) (Wh/mp)		ARIMA (1,0,7) (Wh/mp)		ANN (Wh/mp)	
			Measured	Forecasted	Measured	Forecasted	Measured	Forecasted
2013	9	19	—	—	—	—	—	—
2013	9	20	6,624.2	6,123.75	6,624.2	5,721.44	6,624.2	5,431
2013	9	21	5,078	6,379.47	5,078	5,973.61	5,078	5,421
2013	9	22	6,997.1	5,857.37	6,997.1	5,038.02	6,997.1	5,315
2013	9	23	5,907	6,697.06	5,907	683.91	5,907	5,210
2013	9	24	5,595.3	5,228.81	5,595.3	5,847.36	5,595.3	5,019
2013	9	25	6,976.2	6,379.57	6,976.2	5,745.71	6,976.2	4,696
2013	9	26	6,661.8	5,745.95	6,661.8	5,755.16	6,661.8	5,668
2013	9	27	5,983.8	5,882.45	5,983.8	5,756.6	5,983.8	5,706
2013	9	28	4,089.2	5,472.23	4,089.2	5,755.73	—	—

TABLE 9: Main results of the energy production by the experimental PV park from Grojdibodu forecasted for the period 01-09.08.2013.

Day	1	2	3	4	5	6	7	8	9
Hours when the inverter is running (solar irradiance > 10 W/m ²)	13.75	13.25	13	13	13	13.25	13	13	13
Total hours * string	26,565	25,599	25,116	25,116	25,116	25,599	25,116	25,116	25,116
Total solar irradiance (Wh/m ²)	6,103.6	7,240.3	7,416.6	6,091	7,388.08	7,101.34	7,214.13	7,285.4	7,325.3
Exported active energy (MWh)	50.77	59.48	60.74	48.49	60.18	56.62	59	59.16	59.54
PR (%) calculated in SCADA	83.51	82.55	82.22	79.99	81.83	80.07	82.12	81.54	81.66
Real availability (%)	100	100	100	97.5	98.84	98.11	100	100	100
Availability given by the manufacturer (%)	100	100	100	97.5	98.84	100	100	100	100
Events in which energy production is disrupted	—	—	—	Inverter off 6 hours	Inverter: motherboard damaged (replaced)	Inverter off 15 min	—	—	—
*String hours with energy loss	—	—	—	627.9	291.34	483.82	—	—	—
Estimated energy losses (MWh)	—	—	—	1.8	0.4	1.75	—	—	—

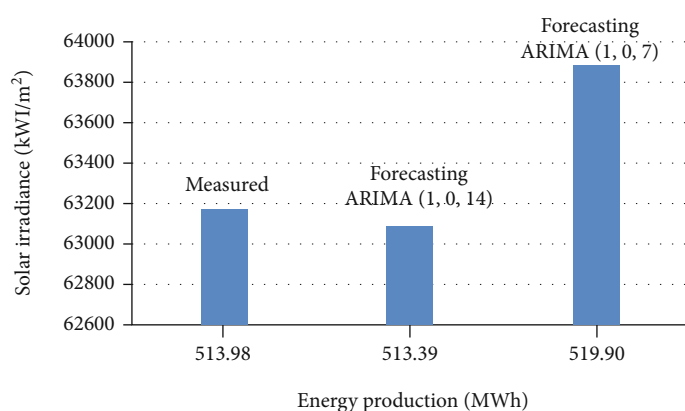


FIGURE 19: Difference between the measured values of solar irradiance and predicted ones based on the ARIMA model.

PV panels (Wh/m²), and PR (performance ratio) represents the efficiency of the PV system.

The comparison between the forecasted energy and the energy inserted in the grid is made between 01.08.2013 and 09.08.2013 for the analyzed experimental PV park (see Table 9).

Considering the data collected within the PV park, the monthly exported energy shows that in the period chosen for testing 01-09.08.2013, a quantity of 513.98 MWh was delivered in the distribution grid. The share of the measured solar irradiance, corresponding to this value of the delivered energy, was 63165.75 Wh/m².

TABLE 10: Comparison of different forecasting models.

Ref.	Year of publ.	Forecasting model	Error	Time horizon	Location
[34]	2018	ARIMA, GA, PSO, DE	VAR, WME	1 day	Canada
[33, 34]	2018, 2018	SVM	RMSE, MRE	15 min.	China
[31, 34]	2019, 2018	NWP	MAE, MBE, RMSE,	1 h, 3 h	USA—California state
[32, 34]	2019, 2018	ANN	NRMSE, NMAE, NBE	1 h	Italy
[34]	2018	SVM, NN, MLR, RT	RMSE	2 h	Australia
Present article: BIPV lab system	2021: to be published	ARIMA, ANN	RMSE, MAE, MAPE, MaxAPE, MaxAE	15 min, 10 days (time series)	South of Romania, Bucharest, BIPV lab system
Present article: experimental PV park	2021: to be published	ARIMA, ANN	rMAE, rRMSE	10 days (time series)	South of Romania, experimental PV park

Using the ARIMA model (1,0,14), the solar irradiance for this period was forecasted at 63091 Wh/m², respectively, and at 63891 Wh/m² using the ARIMA model (1,0,7). Taking into account the technical parameters of the system and the forecast for solar irradiance, the energy forecast with the ARIMA type model will be 513.39 MWh, respectively, and 519.90 MWh (Figure 19).

5. Comparative Analysis of Different Forecasting Models

A comparison of different forecasting models (algorithms) developed in various countries is presented in Table 10.

The results obtained in the present article were compared based on ARIMA and ANN tools with new advanced forecasting models (algorithms), like PSO, GA, SVM, MLR, and RT [31–34]. The accuracy was much better in these cases. However, the authors have studied two interesting PV installations, namely, a BIPV lab system and an experimental PV park; both were placed in the south of Romania. Based on our results, it was established that the ARIMA (1,0,14) model is more efficient than the ANN one and is very close to the experimental case (see Figure 19). Although our approach is dedicated especially to short-term forecast, the long-term solar radiation forecast was discussed too with application for a BIPV system developed at the Polytechnic University of Bucharest. This feature is different from the other more advanced forecasting models presented in Table 10.

6. Optimized Forecasting Method for Energy Production Based on GMDH Model: Case Study—PV Park in Romania

The GMDH (Group Method of Data Handling) represents a family of inductive algorithms for the mathematical modeling of multiparametric datasets that are able to automatically optimize the used models [37]. In the GMDH, continuous or discrete input variables can be introduced,

and depending on their type, optimal parametric or non-parametric algorithms can be identified.

In order to optimize the energy production developed by a Romanian PV park, a specific approach was conceived to render efficient solar radiation prediction on the surface of PV panels for a PV park. As the prediction method involves a large amount of data and a great computing power, Romania's territory was divided into 28 regions, considering that each region possesses at least one actinometric station.

For the created regions, important meteorological data will be provided for solar radiation prediction and will be introduced into advanced models for prediction and evaluation (ARIMA, ANN). The goal is to improve prediction efficiency in both effective results and computing time.

Alongside corresponding solar radiation data and associated factors, for an efficient prediction, it will be introducing also the PV park's characteristics [38]. System loss categories are present in all PV systems but differ from one system to another. It is important to mention the fact that the PV park loss is not an aggregate of particular losses. These losses affect the system individually, their impact being calculated for each component. The losses are due to the dust deposits on the module's surfaces, shadows within the emplacement zone, snow deposits, electrical installation's asymmetry, losses between module's connections (wire harness loss), connection loss, etc. Also, the degrading agent must be considered; it represents the production loss during the maintenance period. Considering the PV park's parameters and also the system loss, the energy produced by the PV park was determined both during the days 04, 08, and 14 of July for 2013 and 2014 and during the days 29, 30, and 31 of July 2013 and 2014. The results are presented in Table 11.

Forecasted values show acceptable results correlated with the errors for solar radiation forecasting. Since the forecasts were made using data from previous days without taking into account future forecasts, this method could be optimized by introducing additional parameters that

TABLE 11: PV park—comparison between measured and forecasted energy using PR of the PV system.

Date	PR (performance ratio) (%)	Measured exported energy (MWh)	Forecasted energy (MWh)	Percentage difference between measured and forecasted values (%)
04 July 2013	81.58	57.47	67.45	17%
04 July 2014	84.98	49.47	66.96	35%
08 July 2013	84.2	57.36	52.79	8%
08 July 2014	84.04	56.83	61.64	8%
14 July 2013	85.72	63.15	67.73	7%
14 July 2014	88.31	30.89	76.71	148%
29 July 2013	80.05	58.11	56.36	3%
29 July 2014	67.02	36.28	44.03	21%
30 July 2013	83.36	49.33	59.87	21%
30 July 2014	83.42	59.75	56.61	5%
31 July 2013	86.31	38.5	62.22	62%
31 July 2014	85.81	35.66	52.35	47%

meteorologically describe the following day. Research continues to identify correlations between meteorological parameters that could accurately be forecasted for a given day and their impact on solar radiation.

7. Conclusions and Prospects

Short-term solar radiation forecasts were developed using the ARIMA and ANN models. The efficiency of the forecasted time series is quantified by rRMSE. The models were improved by separating the days with clear sky from the days with a high degree of cloudiness using the variability index of solar radiation, V_{DR} . The separation of the synoptic situations was performed on time series, not on the predicted values. It has been established that the most common cases for the use of the V_{DR} index are days with clear sky on the forecast day and the day before, where the day “forecast” is represented by the day “tomorrow” and the day “before” is represented by “today” or the day on which the forecast is made. These days represent more than 75.4% of the total number of days of the year, so only these cases were chosen. For the rest of the days when there are situations of cloudy sky or high cloud cover, these forecast solutions offer very large errors. The results of the forecast attest to the fact that the ARIMA model is more efficient than the ANN model. The statistics is significant in the case of ARIMA (1,0,14), this being the reason why this variant was chosen for analysis. Also, comparing the results of the forecast with the measured values, we notice that the ARIMA model (1,0,14) is more efficient than the ARIMA (1,0,7). The ARIMA (1,0,7) and ARIMA (1,0,14) models were selected after performing several iterations. The statistical test used to identify the most significant model is Box Ljung Statistics. This test is applied for ARIMA on the residues of a fitted model, not on the original series, and verifies if these residues do not show autocorrelation. The Box-Ljung model, also known as the statistically modified Box-Pierce model, provides guidance on whether the model is specified correctly. A value significantly lower than 0.05 implies that within

the observed data there is a structure that is not relevant for the model. A value greater than 0.05, as determined in the case, indicates that the chosen model is correct (see [9, 35, 36]).

The future work will be based on a new optimized forecasting method for PV energy production, presented in Section 6. Taking into account that the optimized prediction and forecasting method would involve a large amount of data and great computing power, Romania’s territory could be divided into 28 regions; each region would possess one actinometric (solar radiation) station. A polynomial neural network based on the GMDH (Group Method of Data Handling) approach would be used. It would be represented by a family of inductive algorithms (parametric and nonparametric), which could optimize automatically the energy production of the PV park.

Abbreviations

PV:	Photovoltaic
ARIMA:	Autoregressive Integrated Moving Average
ANN:	Artificial Neural Network
BIPV:	Building Integrated Photovoltaics
GMDH:	Group Method of Data Handling
AIC:	Akaike Information Criterion
SARIMA:	Seasonal ARIMA
NWP:	Numerical weather prediction
TRNSYS:	Transient system simulation tool
eQuest:	Quick energy simulation tool
Energy Plus:	Open source software
BAS:	Building automation systems
BEMS:	Building energy management systems
HVAC:	Heating, ventilation, and air conditioning
DL:	Deep Learning
DBN:	Deep Belief Networks
RNN:	Recurrent Neural Networks
STMN:	Short-Term Memory Networks
AI:	Artificial intelligence
IBM SPSS:	Predictive analysis program
MLE:	Maximum Likelihood Estimation

AR:	Autoregressive
MA:	Moving average
AC and PAC:	Autocorrelation command and partial autocorrelation command
RH:	Relative humidity
RMSE:	Root Mean Squared Error
MAPE:	Mean Absolute Percentage Error
MaxAPE:	Maximum Absolute Percentage Error
MAE:	Mean Absolute Error
MaxAE:	Maximum Absolute Error
WMO:	World Meteorological Organization
WRDC:	World Radiation Data Center
V_{DR} :	Variability index of the solar irradiance
rMAE:	Relative mean absolute error
rRMSE:	Relative mean squared error
PR:	Performance ratio
GA:	Genetic algorithm
PSO:	Particle swarm optimization
DE:	Differential evolution
SVM:	Support vector machine
MRE:	Mean relative error
VAR:	Variance error
WME:	Weekly mean error
MBE:	Mean bias error
NRMSE:	Numerical root mean squared error
NMAE:	Numerical mean absolute error
NBE:	Numerical bias error
RT:	Regression tree
MLR:	Multiple linear regression
NN:	Neural network.

Data Availability

Previously reported data were used to support this article and were based on prior studies cited at relevant places within the text of the paper as references [9, 11, 35].

Disclosure

A PhD thesis related to this subject was successfully concluded at the Polytechnic University of Bucharest.

Conflicts of Interest

The authors declare no conflict of interest.

Authors' Contributions

Conceptualization was performed by L.F.; methodology was performed by L.F. and A.D.; formal analysis was performed by L.F.; investigation was performed by L.F. and A.D.; software was secured by A.D. and D.C.; validation was performed by L.F., A.D., DC, and S.F.; resources were secured by A.D. and S.F.; data curation was performed by A.D. and S.F.; writing (original draft preparation) was performed by L.F.; writing (review and editing) was performed by L.F.; visualization was performed by L.F., D.C., and S.F.; supervision was performed by L.F.; and project administration was performed by L.F.

Acknowledgments

The authors are grateful to Dr. Cristian Oprea who supplied the solar radiation data from Bucharest-Afumati Meteorological Station that were used in this article. At the same time, the authors appreciate the role of the persons in charge from the PV park from Grojdibodu, south of Romania, who allowed the utilization of some technical information related to this park, as well as specific meteorological data of the site within the present paper.

References

- [1] S. Sobri, S. Koochi-Kamali, and N. A. Rahim, "Solar photovoltaic generation forecasting methods: a review," *Energy Conversion and Management*, vol. 156, pp. 459–497, 2018.
- [2] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Solar Energy*, vol. 136, pp. 78–111, 2016.
- [3] H. M. Diagne, M. David, P. Lauret, and J. Bolan, "Solar irradiation forecasting: state-of-the-art and proposition for future developments for small-scale insular grids," in *Proceedings of the WREF 2012-World Renewable Energy Forum*, pp. 65–76, Denver, Colorado, 2012.
- [4] E. G. Kardakos, M. Alexiadis, S. Vagropoulos, and A. G. Bakirtzis, "Application of time series and artificial neural network models in short-term forecasting of PV power generation," in *2013 48th International Universities' Power Engineering Conference (UPEC)*, pp. 1–6, Dublin, Ireland, 2013.
- [5] L. Fara, A. Diaconu, and F. Dragan, "Trends, challenges and opportunities in advanced solar cells technologies and PV market," *Journal of Green Engineering*, vol. 5, no. 4, pp. 157–186, 2016.
- [6] L. Fara, C. Cincu, R. Mitroi et al., *Physics and technology of solar cells and photovoltaic systems*, AOSR Publication, Bucharest, 2009.
- [7] M. Paulescu, L. Fara, and E. Tulcan-Paulescu, "Models for obtaining daily global solar irradiation from air temperature data," *Atmospheric Research*, vol. 79, no. 3-4, pp. 227–240, 2006.
- [8] F. Iacobescu and V. Badescu, "The potential of the local administration as driving force for the implementation of the National PV systems Strategy in Romania," *Renew Energy*, vol. 38, no. 1, pp. 117–125, 2012.
- [9] L. Fara, B. Bartok, A. Galbeaza Moraru et al., "New results in forecasting of photovoltaic systems output based on solar radiation forecasting," *Journal of Renewable and Sustainable Energy*, vol. 5, no. 4, 2013.
- [10] M. Brabec, V. Badescu, and M. Paulescu, "Nowcasting sunshine number using logistic modeling," *Meteorology and Atmospheric Physics*, vol. 120, no. 1-2, pp. 61–71, 2013.
- [11] L. Fara, A. G. Moraru, P. Sterian, A. P. Bobei, A. Diaconu, and S. Fara, "BIPV (building integrated photovoltaic) systems in Romania. Monitoring, modelling and experimental validation," *Journal of Optoelectronics and Advanced Materials*, vol. 5, no. 1-2, pp. 125–130, 2013.
- [12] M. Paulescu, O. Mares, E. Paulescu et al., "Nowcasting solar irradiance using the sunshine number," *Energy Conversion and Management*, vol. 79, pp. 690–697, 2014.

- [13] L. Fara, D. Craciunescu, L. Fara et al., "Estimation of electricity production for a photovoltaic park using specialized advanced software," in *Proceedings of the EuroSun 2016- International Conference of Solar Energy for Buildings and Industry*, Palma de Mallorca, 2016.
- [14] S. Atique, S. Noureen, V. Roy, V. Subburaj, S. Bayne, and J. Macfie, "Forecasting of total daily solar energy generation using ARIMA: a case study," in *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 0114–0119, Las Vegas, NV, USA, 2019.
- [15] J. Runge and R. Zmeureanu, "Forecasting energy use in buildings using artificial neural networks: a review," *Energies*, vol. 12, no. 17, p. 3254, 2019.
- [16] X. Qing and Y. Niu, "Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM," *Energy*, vol. 148, pp. 461–468, 2018.
- [17] Y.-K. Wu, C.-R. Chen, and H. Abdul Rahman, "A Novel Hybrid Model for Short-Term Forecasting in PV Power Generation," *International Journal of Photoenergy*, vol. 2014, Article ID 569249, 9 pages, 2014.
- [18] M. Cococcioni, E. D'Andrea, and B. Lazzarini, "24-hour-ahead forecasting of energy production in solar PV systems," in *2011 11th International Conference on Intelligent Systems Design and Applications*, pp. 1276–1281, Cordoba, Spain, 2011.
- [19] S. Vagropoulos, G. Christos Chouliaras, E. G. Kardakos, and A. G. Bakirtzis, "Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting," in *2016 IEEE International Energy Conference (ENERGYCON)*, pp. 1–6, Leuven, Belgium, 2016.
- [20] H. X. Zhao and F. Magoulès, "A review on the prediction of building energy consumption," *Renewable and Sustainable Energy Reviews*, vol. 16, no. 6, pp. 3586–3592, 2012.
- [21] M. Paulescu, V. Badescu, and M. Brabec, "Tools for PV (photovoltaic) plant operators: nowcasting of passing clouds," *Energy*, vol. 54, pp. 104–112, 2013.
- [22] A. Mellit, M. Benghaneim, A. H. Arab, and A. Guessoum, "A simplified model for generating sequences of global solar radiation data for isolated sites: using artificial neural network and a library of Markov transition matrices approach," *Solar Energy*, vol. 79, no. 5, pp. 469–482, 2005.
- [23] Z. Wang and R. S. Srinivasan, "A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models," *Renewable and Sustainable Energy Reviews*, vol. 75, pp. 796–808, 2017.
- [24] A. T. Eseye, J. Zhang, and D. Zheng, "Short-term photovoltaic solar power forecasting using a hybrid Wavelet-PSO-SVM model based on SCADA and Meteorological information," *Renewable Energy*, vol. 118, pp. 357–367, 2018.
- [25] K. Wang, X. Qi, and H. Liu, "A comparison of day-ahead photovoltaic power forecasting models based on deep learning neural network," *Applied Energy*, vol. 251, 2019.
- [26] F. Harrou, B. Taghezouit, and Y. Sun, "Robust and flexible strategy for fault detection in grid-connected photovoltaic systems," *Energy Conversion and Management*, vol. 180, pp. 1153–1166, 2019.
- [27] M. A. Mat Daut, M. Y. Hassan, H. Abdullah, H. A. Rahman, M. P. Abdullah, and F. Hussin, "Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: a review," *Renewable and Sustainable Energy Reviews*, vol. 70, pp. 1108–1118, 2017.
- [28] M. K. Behera, I. Majumder, and N. Nayak, "Solar photovoltaic power forecasting using optimized modified extreme learning machine technique," *Engineering Science and Technology- an International Journal*, vol. 21, no. 3, pp. 428–438, 2018.
- [29] S. Srivastava and S. Lessmann, "A comparative study of LSTM neural networks in forecasting day-ahead global horizontal irradiance with satellite data," *Solar Energy*, vol. 162, pp. 232–247, 2018.
- [30] M. C. Mukkamala and M. Hein, "Variants of RMSProp and Adagrad with logarithmic regret bounds," in *Proceedings of the 34th International Conference on Machine Learning*, pp. 2545–2553, Sydney, Australia, 2017.
- [31] M. N. Akhter et al., "Review on forecasting of photovoltaic power generation based on machine learning and metaheuristic techniques," *IET Renewable Power Generation*, vol. 13, no. 7, pp. 1009–1023, 2019.
- [32] W. VanDeventer, E. Jamei, G. S. Thirunavukkarasu et al., "Short-term PV power forecasting using hybrid GASVM technique," *Renewable Energy*, vol. 140, pp. 367–379, 2019.
- [33] U. K. Das, K. S. Tey, M. Seyedmahmoudian et al., "Forecasting of photovoltaic power generation and model optimization: a review," *Renewable and Sustainable Energy Reviews*, vol. 81, Part 1, pp. 912–928, 2018.
- [34] M. Seyedmahmoudian, E. Jamei, G. S. Thirunavukkarasu et al., "Short-term forecasting of the output power of a building-integrated photovoltaic system using a metaheuristic approach," *Energies*, vol. 11, no. 5, 2018.
- [35] L. Fara, A. Diaconu, and D. Craciunescu, "Forecasting of energy production and operational optimization for photovoltaic systems," *Annals of the Academy of Romanian Scientists, Series on Physics and Chemistry Sciences*, vol. 2, no. 1, pp. 31–60, 2017.
- [36] A. Diaconu, *Contributions to the operational optimization of photovoltaic systems in applications*, [Ph.D. thesis], Polytechnic University of Bucharest, Bucharest, 2017.
- [37] <https://www.gmdhshell.com/time-series-forecasting>.
- [38] <http://photovoltaic-software.com/PV-solar-energy-calculation.php>.