

Article

Prediction of a Grid-Connected Photovoltaic Park's Output with Artificial Neural Networks Trained by Actual Performance Data

Elias Roumpakias  and Tassos Stamatelos * 

Department of Mechanical Engineering, University of Thessaly, 383 34 Volos, Greece; roumpakias@uth.gr

* Correspondence: stam@uth.gr; Tel.: +30-24210-74067

Featured Application: The accuracy in long-term prediction of the performance of a photovoltaic (PV) park following several years of operation is essential for a number of reasons, including the extraction of reliable data for the economic analysis, the valid support of claims regarding PV panels' performance degradation and other issues. The specific work was motivated by the need to economically assess future investment in photovoltaic installations during an economic cycle with increasing interest rates. The monitoring data from the actual operation of a 100 kW PV park are employed in the refinement and validation of the prediction methodology. The results are applicable to the profitability analysis of future PV investments in the region of Thessaly.

Abstract: Increased penetration of grid-connected PV systems in modern electricity networks induces uncertainty factors to be considered from several different viewpoints, including the system's protection and management. Accurate short-term prediction of a grid-connected PV park's output is essential for optimal grid control and grid resilience. Out of the numerous types of models employed to this end during the last decade, artificial neural networks, (ANNs) have proven capable of handling the uncertainty issues of solar radiation. Insolation and ambient, or panel temperature, are most commonly employed as the independent variables, and the system's output power is successfully predicted within 3 to 5% error. In this paper, we apply a common type of ANN for the long-term prediction of a 100 kWp grid-connected PV park's output, by exploiting experimental data from the last 8 years of operation. Solar radiation and backsheet temperature were utilized for the ANN training stage. The performance metrics of this model, along with a standard linear regression model, are compared against the actual performance data. The capabilities of the ANN model are exploited in the effort to decouple the fluctuating effect of PV panel soiling which interferes with the efficiency degradation process. The proposed methodology aimed to quantify degradation effects and is additionally employed as a fault diagnosis tool in long-term analysis.

Keywords: photovoltaics; air mass; forecasting; neural networks



Citation: Roumpakias, E.; Stamatelos, T. Prediction of a Grid-Connected Photovoltaic Park's Output with Artificial Neural Networks Trained by Actual Performance Data. *Appl. Sci.* **2022**, *12*, 6458. <https://doi.org/10.3390/app12136458>

Academic Editors: Isabel Santiago Chiquero, Isabel María Moreno García and Rafael López Luque

Received: 11 June 2022

Accepted: 23 June 2022

Published: 25 June 2022

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



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

1. Introduction

The growth in investment for the production of renewable energy is remarkable and the recent severe threats to the energy safety of Europe due to the scarcity and soaring gas prices led to the increasing pace in the design of a future European electricity system characterized by a dominant share of renewable energy supply. This is in line with the stated targets of European governments and the official position of the EU, increasing its renewable energy target from 40 to 45%, while setting a 592 GWac (740 GWdc) target for solar in the European Union by 2030 [1]. In Greece, currently installed peak power is 4381 MWp of wind parks, 256 MWp of small-scale hydropower systems, 112 MWp of biomass systems, 3989 MWp of photovoltaic (PV) systems, and 352 MWp of rooftop PV systems on roofs [2]. Photovoltaics applications are widespread in the residential and commercial sector, assisted by the continuous drop in purchasing prices which currently range from 0.20–0.40 EUR/W_p for silicon-based photovoltaic modules with 20% efficiency [3]. Net-metering further

boosted the expansion of PV applications in the residential sector. Its application varieties include net-metering, virtual net-metering, and net-metering with battery systems in order to achieve zero energy induction in the public grid. Sizing these applications demands meteorological data, consumption data and profiles, and, most importantly, reliable solar potential data because of the uncertainty of the solar insolation. Thus, forecasting power output is an essential task to support the optimal sizing of these applications. The significant expansion of electric vehicles is another application that will necessitate a close synergy with rooftop and other PV installations [4]. Again, the system's optimization in these cases needs to be supported by reliable power forecasting because of the inherent uncertainty of the availability of solar energy and the randomness of the EV charging behavior [5]. Furthermore, power forecasting of the various forms of power production is also an important tool in electricity markets and energy markets [6,7].

Many researchers face this problem using different approaches. Hassan et al. proposed an empirical technique using experimental data and interpolated datasheets' information in order to produce a mathematical equation for power prediction under various weather conditions [8]. Huld et al. presented a method for energy rating using mathematical models for shallow-angle reflectivity, spectral sensitivity, module efficiency dependence on irradiance, and module temperature [9]. Zamo et al. studied PV forecasting for some power plants in mainland France using statistical methods exploiting outputs from numerical weather prediction (NWP) models. Forecasts were built without using technical information on the power plants. The results showed a root mean square error (RMSE) in the range of 9–12% [10]. Neural networks are another popular approach adopted by researchers in PV energy forecasting. Kim et al. [6] presented the most recent research works on deep neural networks applied in the short term (forecast horizon from 30 min to 1 week ahead) photovoltaic power forecasting. Graditi et al. conducted a comparative analysis among three models for energy forecasting: a phenomenological model, a multilayer perceptron (MLP) neural network, and one based on regression analysis, applied to a dataset of an existing PV power plant. The neural networks technique demonstrated superior accuracy in the results [11]. Sundaram et al. proposed an artificial neural network (ANN) for the long-term prediction of an existing 1MWp plant using three different options. Input variables were the module's ambient temperatures, wind speed, and global irradiance. The ANN model with the four inputs achieved more accurate results with a mean absolute percentage error (MAPE) of 1.68% [12]. Kardakos et al. proposed the application of the seasonal autoregressive integrated moving average (SARIMA) model and two ANNs for energy generation forecasting of grid-connected plants in Greece. Comparisons were conducted in terms of normalized root mean square error (nRMSE) [13]. López Gómez et al. created an ANN to predict power from a real PV plant in Italy. Three scenarios were investigated: data employed for both training and prediction, data employed just for training and prediction was compared with data from the global data assimilation system, and finally, data from GDAS were used for training and prediction. The results show that the lack of on-site weather measurements can be faced by means of a validated numerical weather model in combination with machine learning techniques for PV power forecasting [14]. Kothona et al. proposed a forecasting model based on the long short-term memory (LSTM) algorithm. The examined parameters were solar irradiance, PV module temperature, historical PV data, and clearness index. The results indicated that the inclusion of the clearness index as input can improve the performance of the forecaster [15]. Omar Nour-Eddine et al. studied PV power production data of a 5.94 kWp grid-connected PV plant in Morocco. The PV system comprised panels of three different PV technologies. The proposed models are an ANN and a persistence model that are applied to a data set with solar irradiance and temperature as inputs. The results were compared in terms of mean bias error (MBE), mean absolute error (MAE), MAPE, RMSE, and nRMSE. The ANN-based model demonstrated a 2.107% MAE and 2.645% RMSE against 2.406% and 5.185%, respectively, for the persistence model [16]. Akhter et al. proposed a model for an hour-ahead prediction on a yearly basis of three different PV plants, based on available data for wind speed, module, ambient

temperature, and solar irradiation employing a long short-term memory (LSTM) recurrent neural network (RNN) with a deep learning method, with the results compared with regression, hybrid Adaptive neuro-fuzzy inference system (ANFIS), and machine learning methods [17]. Natarajan et al. proposed a radial belied neural network (RBFN) with inputs from large-scale PV plants with evaluation metrics of RMSE, nRMSE, MBE, MAE, MaxAE, MAPE, and Kolmogorov–Smirnov test integral (KSI) and OVER metrics, skewness, and kurtosis and variability estimation [18]. Pan et al. proposed a support vector machine (SVM) that uses weather temperature, relative humidity, global horizontal radiation, diffuse horizontal radiation, wind direction, and sampling time as input parameters. The output of the model was active power and the results were a correlation coefficient R^2 up to 0.997, MSE and MAE values of 0.0349 and 0.1569, respectively [19]. Pasion et al. proposed a machine learning technique using latitude, month, hour, ambient temperature, pressure, humidity, wind speed, and cloud cover as independent variables and power production as output. A distributed random forest regression algorithm modeled the combined dataset with an R^2 value of 0.94 [20]. Jaber et al. presented a prediction model for comparing the performance of six different photovoltaic (PV) modules using artificial neural networks (ANNs), with the basic characteristics of the PV panels' manufacturer datasheets as cell temperature, irradiance, fill factor, short circuit current, open-circuit voltage, maximum power. Comparing the results to the actually measured data resulted in a 0.874% MAPE [21]. Kim et al. proposed a combination of bidirectional long short-term memory (BLSTM) with an ANN to predict power generation for a specific hour, using four types of historical input data: hourly PV generation (168 h ahead), hourly horizontal radiation, hourly ambient temperature, and hourly surface temperature. The results showed that the LSTM prediction model with the ANN estimation model using exponential moving average preprocessing exhibited higher accuracy [6].

Another related aspect of neural network application in the building's rooftop PV sector is consumption forecasting. This is a useful tool for the optimal sizing of PV systems for net-metering applications, including those with battery storage systems. Knowledge of the building's consumption profile is a necessary starting point for the improvement of the building's energy management and application of energy-saving actions. Villanueva et al. proposed a method for predicting the consumption of household appliances by evaluating statistical distributions (Kolmogorov–Smirnov and Pearson's χ^2 tests) [22]. Kalogirou et al. proposed a model for the prediction of energy consumption in a passive solar building in order to generate a mapping between the above easily measurable inputs and the desired output, i.e., the building's energy consumption [23].

In addition to the above issues, photovoltaic applications must be based on special techniques for fault detection, degradation investigation, and other effects such as soiling and potential induced degradation. Neural networks offer great help in studying these effects. Laurino et al. proposed a four-layered feed-forward artificial neural network that learns the correlation of the IV curves with irradiance and temperature. This approach has been also been applied to detect anomalous increases in series resistance from a large experimental set of curves [24]. Chine et al. proposed an ANN that uses a given set of working conditions parameters—solar irradiance and photovoltaic (PV) module's temperature—to predict a number of attributes such as current, voltage, and number of peaks in the I–V curve of the PV strings. The simulated attributes were then compared with real field measurements, leading to the identification of possible faulty operating conditions [25]. Massi Pavan et al. studied the effect of soiling on large-scale photovoltaic plants using two different techniques: Four Bayesian neural network (BNN) models have been developed in order to calculate the performance at standard test conditions (STCs) of two plants installed in Southern Italy and compare against the results of a regression model. The results indicate that the losses due to dust accumulation on poly-crystalline Si PV modules' surfaces range from roughly 1 to 5% after 1 year of operation [26], which agrees with observations from other researchers [27].

Neural networks are a significant help in solar radiation forecasting which is necessary for the planning, operation, and maintenance procedures of PV plants. Behrang et al. [28] employed different ANN techniques to predict daily global radiation on a horizontal surface, based on the daily mean air temperature, relative humidity, and sunshine hours. De O. Santos et al. proposed a heterogeneous ensemble dynamic selection model, named HetDS, to forecast solar irradiance, choosing the most suitable forecasting model from a pool of seven well-known literature methods: ARIMA, support vector regression (SVR), multilayer perceptron neural networks, extreme learning machine (ELM), deep belief network (DBN), random forest (RF), and gradient boosting (GB). The experimental evaluation showed that the proposal can overcome the single models in almost all scenarios with a small dispersion [29].

It can be observed that the various approaches differentiate regarding the type of Neural Network, the selection of inputs, outputs, and the kind of data employed for the training and validation. Another important selection parameter is the type of forecasting, which may be classified into three main categories according to Das et al. [30]: (i) short-term forecasting of PV power generation that is applied in scheduling and dispatching of electrical power, or in the design of a PV-integrated energy management system, (ii) medium-term forecast, carried out for periods from one week to one month for power system and maintenance scheduling, and (iii) long-term PV power forecasting for periods from one month to one year, which is helpful for the planning of electricity generation and distribution [30]. As expected, long-term PV power forecasting is the most demanding task for the use of machine learning methods, which need a lot more research to understand their complications and transform them into efficient computational tools [31]. The experience from the application of deep learning in other related fields is useful in this context. Nguyen et al. [32] constructed a long short-term memory (LSTM) RNN for making predictions of French nuclear power plants' steam generator output over a long-term horizon, in which the network hyperparameters are automatically optimized by a tree-structured Parzen estimator (TPE) algorithm. RNNs include feedback connections from the hidden/ output layer to the preceding layers, to catch the dynamics of sequential data and reproduce previous patterns in the future prediction. Similarly, RNNs are applied in long-term forecasting of lithium-ion battery performance [33], membrane degradation in fuel cells [34], rolling bearings remaining useful life [35], and machine health monitoring [36].

The current work belongs to long-term power forecasting. It focuses on the exploitation of actual data collected from monitoring grid-connected PV systems and has three objectives. The first objective is the comparison of an ANN model with a simple linear regression in PV energy forecasting using actual data. The second is the evaluation of the data in order to be useful either for forecasting or for fault diagnosis and degradation analysis. The third is to create a methodology for the performance evaluation of a PV system and reliable estimation of degradation rates, which is essential information for the creation of energy forecasting models.

The rest of the manuscript is organized as follows: Section 2.1 describes the photovoltaic park's details and the experimental setup employed for its monitoring. Section 2.2 describe the pre-processing and Section 2.3 introduces the forecasting methods, whereas the results are presented and compared with actual data in Section 3. In Section 4, a comparative discussion of the performance of the alternative models and training approaches is carried out, and finally, conclusions are drawn on the long-term prediction and fault diagnosis capabilities.

2. Materials and Methods

The experimental setup employed for the monitoring of the PV park and the specific methods implemented for power forecasting are described in this section.

2.1. Experimental Setup

A grid-connected 99.84 kWp PV park in central Greece is monitored since 2013. The PV park comprises 416 PV panels on the park, mounted in a fixed south-facing position at 25 degrees tilt angle. A total of 8 inverters are employed in the DC/AC transformation, with technical characteristics presented in Appendix A. The following parameters are monitored at 15 min intervals: solar irradiance, back panel temperature, ambient temperature, DC voltage to the inverter, and AC power output from the inverter. The data employed in the specific work refer to the period 1 January 2013 to 30 June 2021. The system's technical data and monitoring and measurement equipment are summarized in Appendix A. No system for automatic cleaning of the panels' surface is included.

2.2. Pre-Processing Methodology

The first step in the pre-processing methodology is the application of a quality assurance procedure to the available dataset. This is a regular practice of outlier removal and missing data imputation techniques dictated by experience in the processing of large time series of measurements [32,37]. Specific criteria are defined for cleaning up the data, with regard to irradiance, airmass, and inverter power output. Recordings with zero values of inverter power output are rejected and the same is done with those with irradiance values $< 50 \text{ W/m}^2$. Airmass is another important indicator that should be taken into account with data preprocessing [38]. According to previous experience, data records with $AM > 10$ are also rejected.

2.3. Forecasting Methodology

Following the preprocessing, the first step in the forecasting methodology is the application of the first three years' (2013–2015) data in training (i) a linear regression model and (ii) a feed-forward artificial neural network (FF ANN) with one hidden layer (MLP). The input parameters for both models are in-plane irradiance and backsheet temperature and the single output parameter is AC inverter power output.

The structure of the FF ANN is shown in Figure 1. It is implemented by the use of the open-source machine learning framework: Tensorflow (Python programming language) [39]. The input layer comprises two input nodes, namely, irradiance (G) and backsheet temperature (T_p) (Figure 1). The hidden layer comprises 20 nodes, each one containing a sigmoid activation function [40]. The Levenberg–Marquardt optimization algorithm is employed for training [41,42].

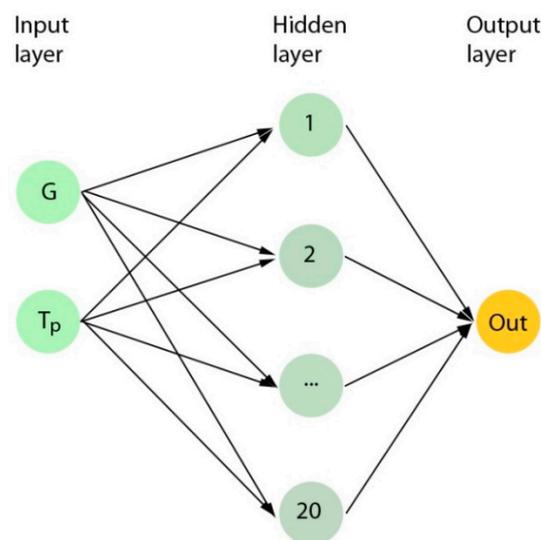


Figure 1. Structure of the feed-forward artificial neural network with one hidden layer of 20 nodes, employed in the data processing.

Training the neural network can be envisaged as the fitting of a complex curve through the training data, until it can approximate it with the highest possible accuracy, by employing loss minimization algorithms and kernels (weights) and biases optimization. Since this complex fitting will be employed in the prediction of the future behavior of the PV park, it is very important to exclude erroneous or noisy information from the dataset. This explains the required careful quality assurance procedure included in the preprocessing of the datasets. The performance of these models is assessed based on the measurements of the last five and a half years in the dataset (2016–2021). As a next step in the forecasting methodology, training of the FF ANN was carried out based on the first five years' (2013–2017) data and the performance of the neural network was assessed based on the measurements of the last three and a half years in the dataset (2018–2021).

The statistical metrics employed for the assessment of the accuracy of prediction are the root mean square error (RMSE), normalized root mean square error (nRMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), mean bias error (MBE), and Pearson correlation coefficient (COR), which are the most common metrics employed in the literature comparisons [28,43,44]. They are defined as follows:

$$\text{MAE} = \frac{\sum_{i=1}^N |\text{Pm}_i - \text{Pf}_i|}{N} \quad (1)$$

$$\text{MAE} = \frac{\sum_{i=1}^N |\text{Pm}_i - \text{Pf}_i|}{N} \quad (2)$$

$$\text{nRMSE} = \frac{\sqrt{\sum_{i=1}^N (\text{Pm}_i - \text{Pf}_i)^2}}{\sqrt{\sum_{i=1}^N (\text{Pm}_i)^2}} \quad (3)$$

$$\text{MAPE} = \frac{100\%}{N} \sum_{i=1}^N \frac{|\text{Pm}_i - \text{Pf}_i|}{\text{Pm}_i} \quad (4)$$

$$\text{MBE} = \frac{\sum_{i=1}^N (\text{Pm}_i - \text{Pf}_i)}{N} \quad (5)$$

$$\text{COR} = \frac{\text{COV}(\text{P}_{\text{mi}}, \text{P}_{\text{fi}})}{\sqrt{V(\text{P}_{\text{mi}})} \sqrt{V(\text{P}_{\text{fi}})}} 100\% \quad (6)$$

where N is the number of samples in the time series, Pm_i and Pf_i denote the measured and predicted value of the modeled variable at the time step i ($i = 1, N$). In the definition of the Pearson correlation coefficient, (COR), COV denotes the covariance of the two variables (measured and modeled values), and V is the variance of each respective variable.

Another important metric that will be computed in our study is the yearly performance ratio (PR) which is defined in Appendix B.

3. Results

First, it is important to conduct a sensitivity analysis in order to correlate input parameters, irradiance, and back panel temperature with the output parameter (power output) and—consequently—energy production. An additional, important parameter that assists this sensitivity analysis procedure is the airmass [45], which defines the direct optical path length of the Sun's rays through the Earth's atmosphere, expressed as a ratio relative to the minimum path length vertically upwards at the solar zenith. This coefficient is useful for characterizing the modification of the solar spectrum after traveling through the atmosphere. It is also intensively investigated as a valuable predictor in meteorological models [46]. For our purposes, airmass is routinely calculated for the available system's performance data according to Appendix B. Figure 2 classifies the electricity produced by the PV park according to the respective, dimensionless airmass values, which was found convenient to divide among six classes (AM1–AM2, AM2–AM3, AM3–AM4, AM4–AM5, AM5–AM10, AM10, and beyond) [47].

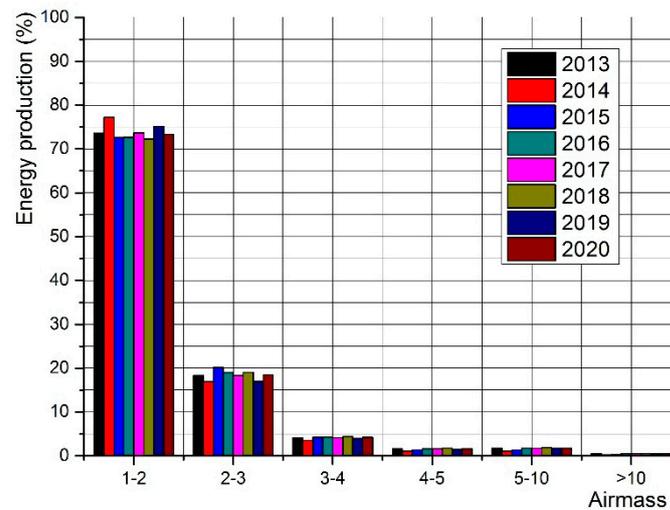


Figure 2. Percentage of electricity production in the six airmass classes defined, during the 8-year period.

According to the results of Figure 2, more than 90% of electricity production corresponds to conditions with Airmass values in the range of 1 to 4. Moreover, 99% of the production is achieved for airmass values <10. For this reason, data for higher values are rejected as discussed in Section 2. Data with values of irradiance smaller than 50 W/m² are also rejected. Figure 3 shows that only 7.5% of electrical energy production is achieved for irradiance values smaller than 200 W/m². As already mentioned in Section 2, records that correspond to irradiance values lower than 50 W/m² are rejected. These points represent in total, less than 1% of the annual electricity generation. Exclusion of the above-mentioned records with low information quality significantly improves the prediction accuracy of the neural networks since we avoid training the NN with noisy data with low information value.

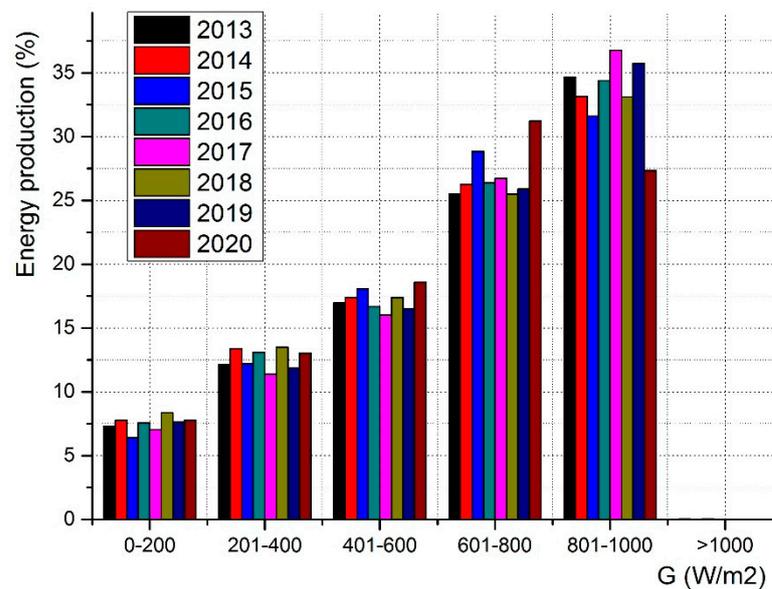


Figure 3. Percentage of energy production in the various irradiance classes during the 8-year period.

The percentages of electricity production in the different backsheet temperature classes defined in the pre-processing are presented in Figure 4. More than 99% of the total electricity production occurs in the four backsheet temperature classes between 15 and 55 °C.

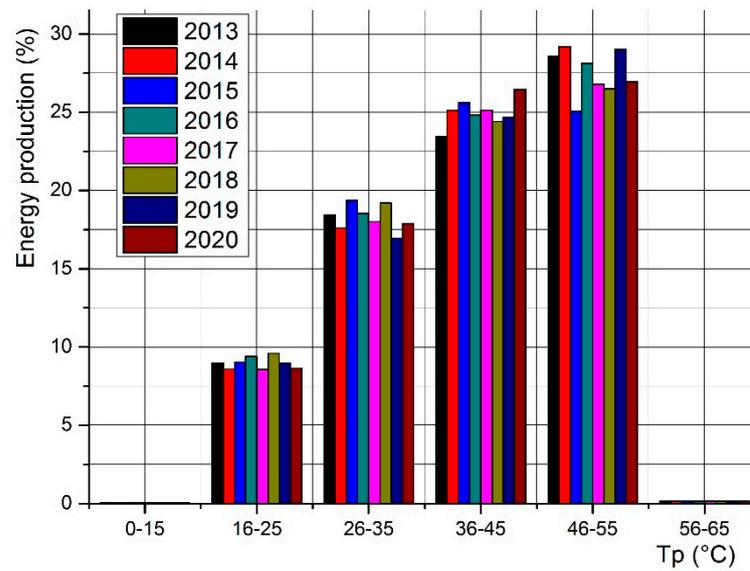


Figure 4. Percentage of energy production in the different backsheet temperature classes during the 8-year period.

Figure 5 shows the observed linear trend in the evolution of AC power output with irradiance. On the other hand, backsheet temperature has also a general linear trend with irradiance. However, according to the technical datasheet of the PV panels, there is a negative effect of backsheet temperature on the panel’s efficiency. These two conflicting trends produce the correlation in Figure 6 between power output and temperature, characterized by a wide margin of possible backsheet temperatures associated with each AC output level.

The situation is more clearly depicted in Figure 7, where the correlation between power output, temperature, and irradiance is apparent. The PV panels’ manufacturers describe the correlation of these three parameters with characteristic curves and temperature coefficients. STCs include irradiance (1000 W/m²), temperature (25 °C), and spectrum corresponding to AM 1.5. The lack of availability of spectral measurements is adjusted by the introduction of the airmass factor. Air mass quantifies the length of trail of the Sun’s beams through the Earth’s atmosphere [48].

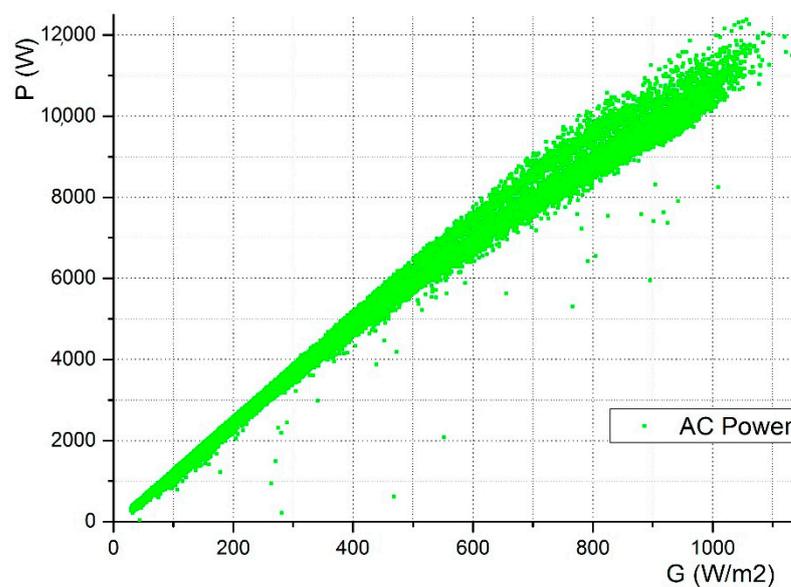


Figure 5. Correlation between irradiance and AC power output.

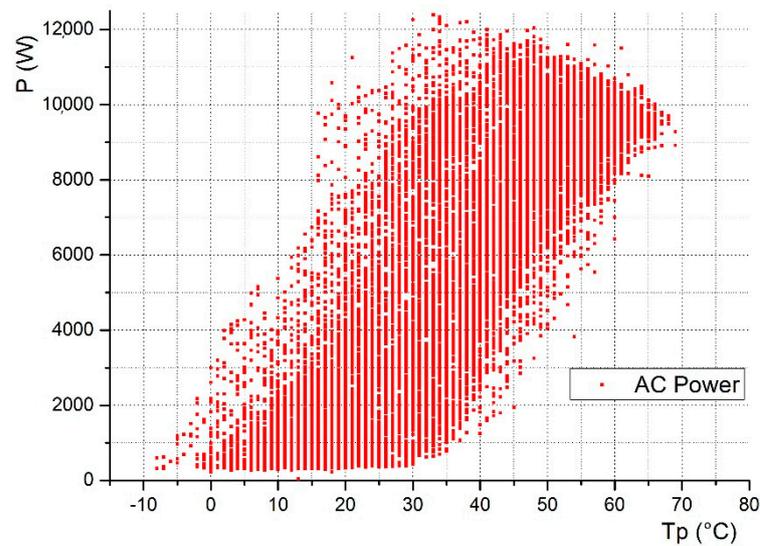


Figure 6. Correlation between backsheet temperature and AC power output.

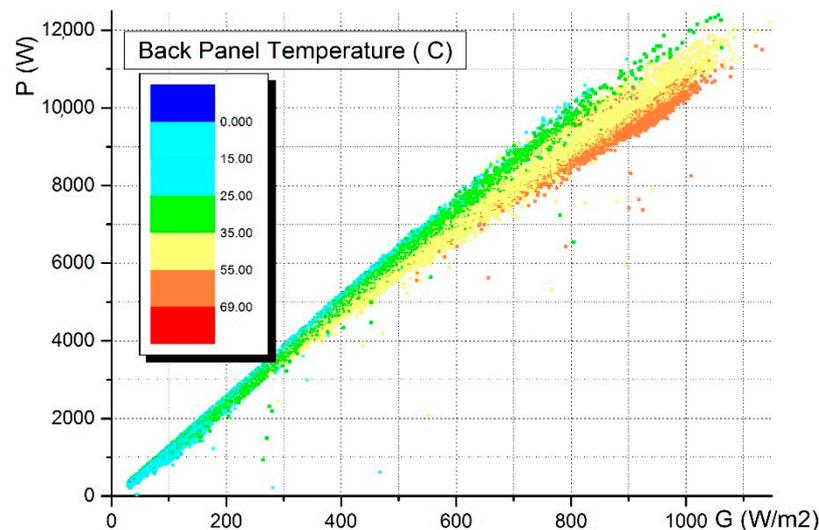


Figure 7. Correlation between AC power output, irradiance, and backsheet temperature.

Figure 8 shows the correlation between power output and airmass in different classes of irradiance. It is observed that power output has an exponential trend. Airmass factor calculation should operate as an indicator in order to compare values of time series in different seasons and hours when sunbeams have the same length of trail.

Next, we proceed to compare the results of the ANN and the linear regression model predictions to the actual performance during characteristic periods in the years 2016 to 2021.

Figure 9a compares the results of the linear regression model and the ANN with the actual performance data during three consecutive cloudy days in 2016. As expected, the prediction with the more complex (nonlinear) curve fitting of the ANN is always closer to the actual data. The same behavior is observed during the sunny days (Figure 9b).

The comparison continues with three consecutive cloudy days in 2017, followed by three sunny days (Figure 10). The performance of the ANN is more accurate in these examples from 2017. The linear regression model lacks accuracy in prediction, especially during the high insolation time during the noon hours of sunny days.

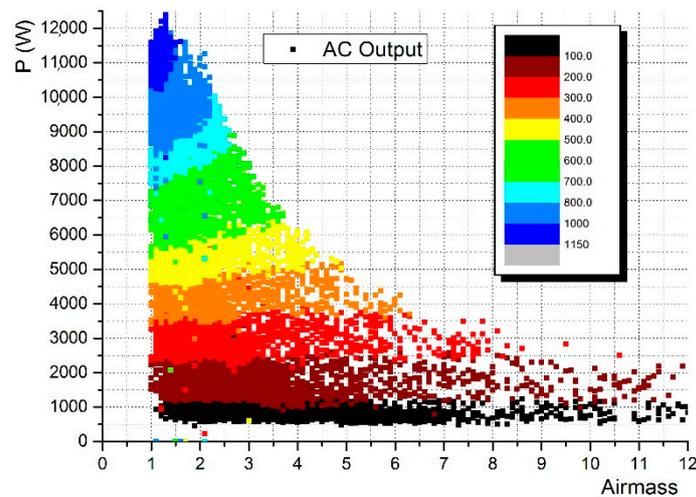


Figure 8. Correlation between AC power output, airmass, and irradiance.

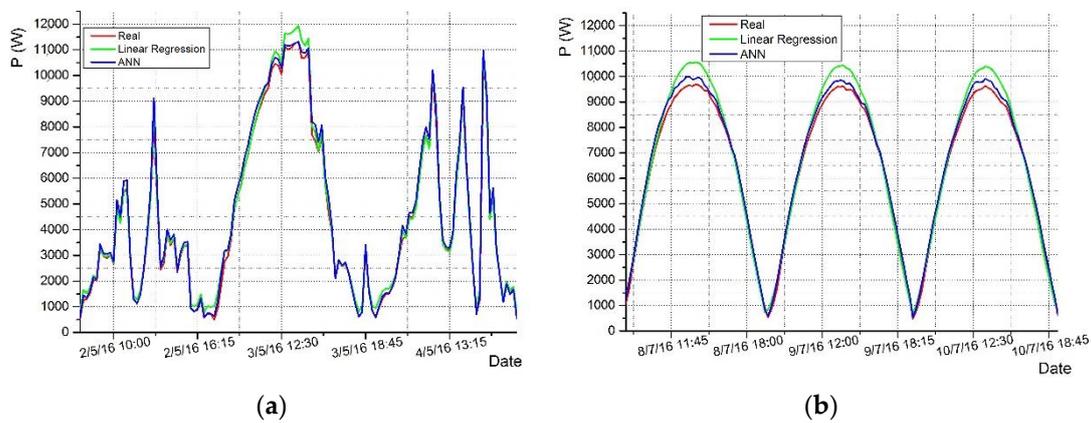


Figure 9. Comparison between actual power output and forecasts of the two alternative methods for three consecutive cloudy (a) and three sunny days (b), respectively, from 2016.

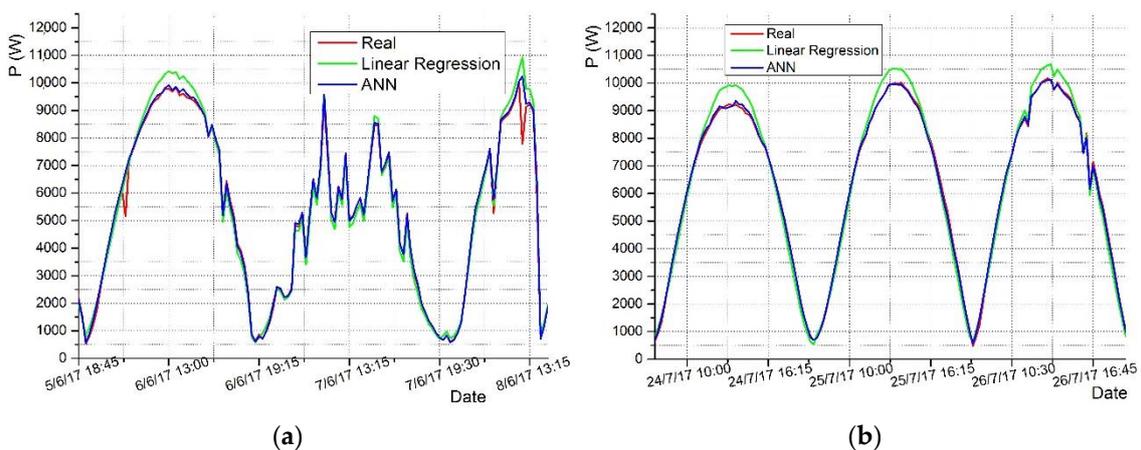


Figure 10. Comparison between real power output and two models forecasting for three cloudy (a) and three sunny days (b) for 2017.

The comparison continues in Figure 11 with three cloudy and three sunny days, with similar findings. The ANN prediction continues to be closer to the measurement, as seen by the qualitative comparison in Figure 11.

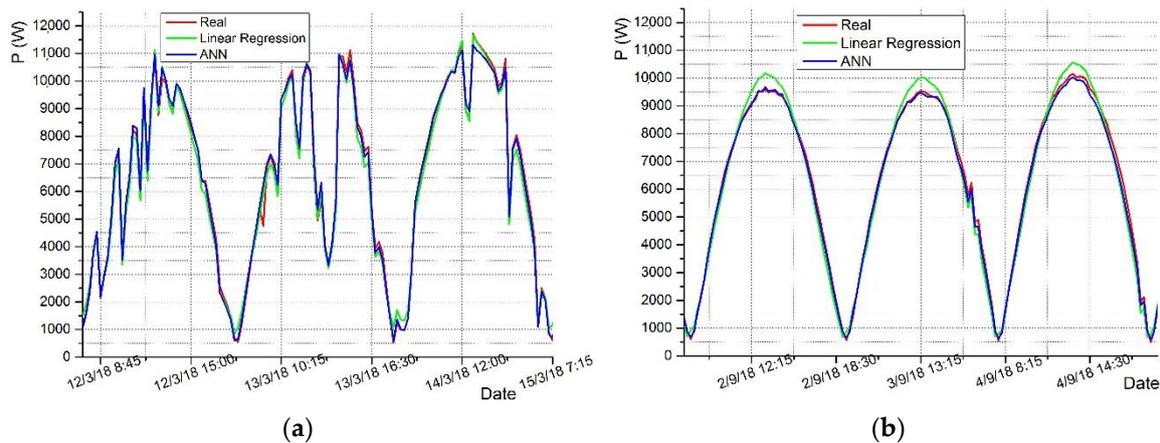


Figure 11. Comparison between real power output and two models forecasting for three cloudy (a) and three sunny days (b) for 2018.

In summary, as regards the years 2016 to 2018, which are closer to the training period, the characteristic results of Figures 9–11 show that both models forecast the power output with a very good accuracy, the ANN model being generally superior. The nRMSE values fluctuated between 4.6 and 7.4% for the ANN model and 6.2–8.8% for the linear regression model. Respectively, MAPE values fluctuate between -2.27 and -1.32% for the ANN model and -1.81 and -1.13% for the linear regression model. It is important to note additionally that during these years the performance ratio (PR) was in the range of 0.86–0.88.

Now, as one moves at more distance from the training period, Figure 12 shows that both models forecast the power output with a remarkably higher nRMSE error value for 2019 compared to the previous years. nRMSE values are 11.2% for the ANN model and 12.5% for the linear regression model. The respective MAPE values are -5.97 for the ANN model and -5.75% for the linear regression model. It must be noted here that this year was characterized by a very low PR of 0.84. On the other hand, both models have an increase in RMSE error and especially 9.8% and 9.7% for the ANN and regression model, respectively.

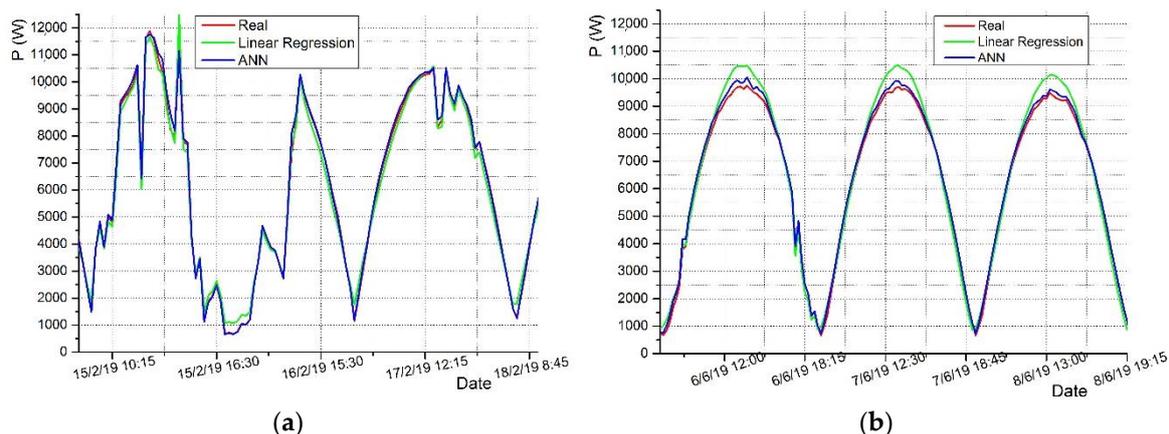


Figure 12. Comparison between actual power output and the one forecasted by the two models, for three cloudy (a) and three sunny days (b) for 2019.

Shifting attention to the year 2020 (Figure 13), the situation further deteriorates with the prediction accuracy of both models, which both overpredict the electricity output of the PV park. Obviously, this behavior must be related to the rate of efficiency deterioration of the PV panels. During the assessment of the first six years of operation of the specific PV park [47], the degradation rates of normalized efficiency for the first two AM classes, which have the highest impact on total energy production, varied between 1.28 and 6.92%. The degradation rate was variable during the years. Thus, the degradation embodied in

the first three years' data employed in the training of both models, cannot be expected to predict with high fidelity its evolution during the next several years.

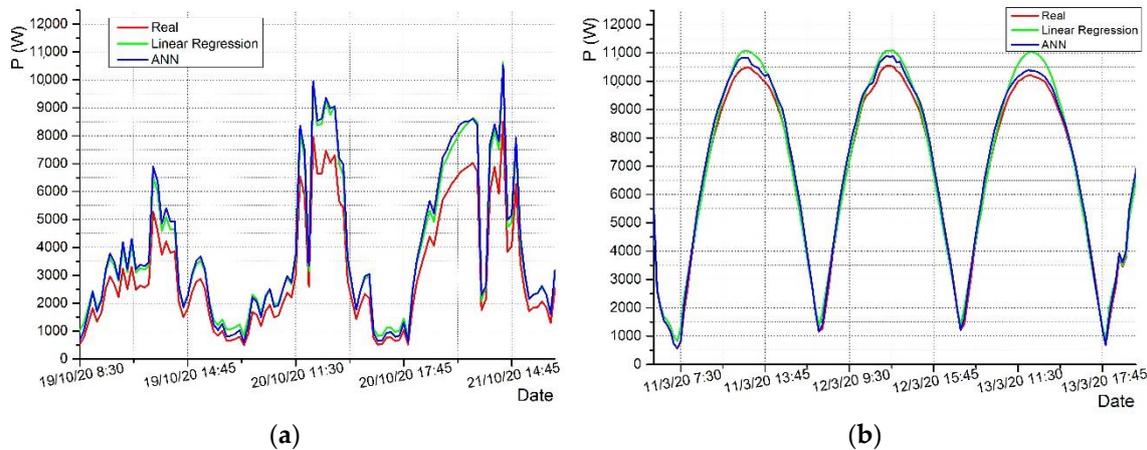


Figure 13. Comparison between real power output and two models forecasting for three cloudy (a) and three sunny days (b) for 2020.

Figure 14 shows that during the next year, 2021, for the first time an underestimation in forecasting values is observed by both models. This is further supported by the values of MAPE which are now positive. This fact would probably indicate an error in the irradiance sensor. To allow for a better assessment of the situation, a comparison of the calculated average PV plant efficiency for the period 2016–2021 is presented in Table 1.

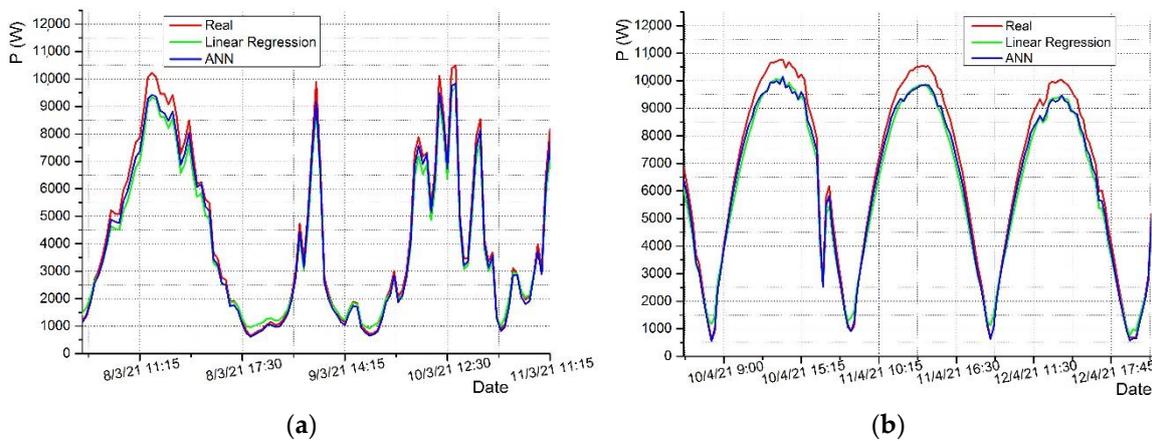


Figure 14. Comparison between real power output and two models forecasting for three cloudy (a) and three sunny days (b) for 2021.

Table 1. Statistical and energy metrics.

	2016	2017	2018	2019	2020	2021
Average η [%]	13.93	13.99	13.99	13.47	13.97	15.17
Standard deviation [%]	1.04	1.51	1.21	1.56	1.36	1.74

4. Discussion

The application of linear regression (M1) and ANN (M2) provides insight into the effectiveness of these models in understanding and predicting the operation of a PV plant over the period of a decade. It is clear that the ANN model achieves better statistical metrics than the regression model, as expected by the published experience of other researchers. However, as one moves further away from the training period, both models show a significant differentiation in performance, especially during the years 2019–2021. This

fact indicates a possible drop in the PV plant's performance that may be explained either by an expected degree of the degradation rate of the PV panel or by the soiling effect as there exists no systematic cleaning procedure in the plant. Another explanation could point to an error in the irradiance measurement system. The irradiance sensor described in Appendix A is by itself a reference PV cell, thus subjected to soiling with an adverse effect on the irradiance values recorded.

Energy generation metrics in the form of PR and Total PV performance, indicate a decreasing trend from 2016 to 2019 that is coincident with the manufacturer's warranties. On the contrary, PR and total PV performance show an increase during 2020–2021. Moreover, the average performance for the year 2021 is significantly higher than the years 2016–2018. This fact points to an erroneous irradiance sensor. Figure 15 points to this fact by the comparison of irradiance–airmass correlations of the years 2013 and 2021. The readings for 2021 indicate a systematic shift to lower values of insolation producing useful electricity in all the airmass ranging from 1–10. This refers to a drift of sensor output.

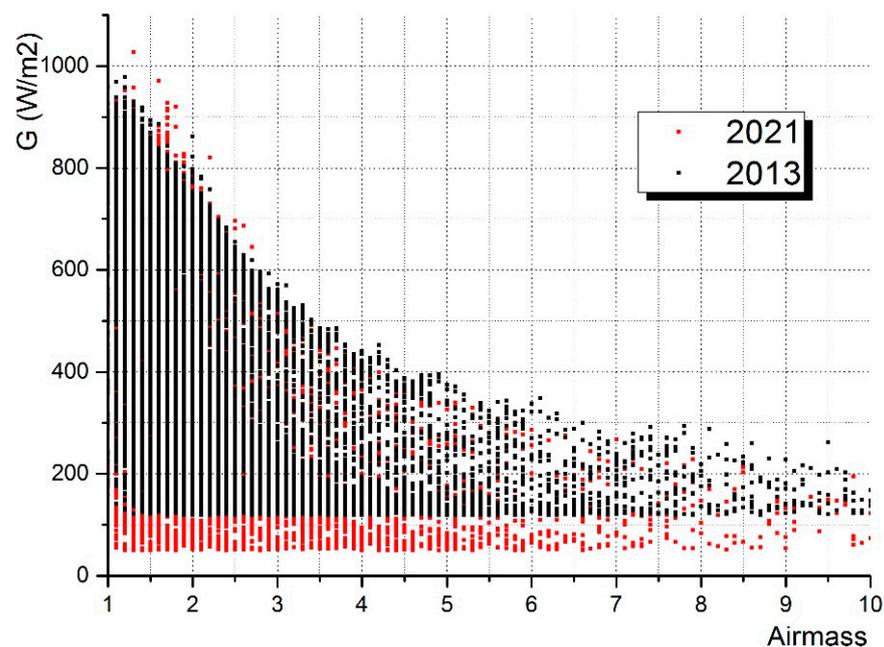


Figure 15. Comparison of measured Irradiance of 2013 and 2021 in correlation with airmass.

A closer look in Figure 16 reveals the problem with the soiled irradiance sensor. More specifically, the day of 26 April is a clear sky, sunny day; however, the sensor does not read correctly the maximum irradiance value of about 1000 W/m^2 . The same observation can be carried out for the next three days, which are days with intermittent clouds; however, the maximum irradiance in the sunny intervals is always recorded at reduced levels.

Table 2 summarizes the calculated statistical and energy metrics of models M1 (regression model) and M2 (FF ANN trained with three-year data), for the years 2016 to 2021.

Figure 17 shows the trend of MAPE and nRMSE of the two models' predictions with time. In general, an increase in error is observed over the years. However, this increase is not monotonous, as also observed in Table 2. Of course, one cannot expect the specific approach to be very accurate for long-term forecasting, because of the erratic nature of the efficiency degradation rate of an ensemble consisting of a large number of PV panels, which are not regularly cleaned from dust and soiling. However, it is employed here as a useful tool for fault detection and degradation analysis. Previous studies, which investigate the efficiency degradation of PV plants and especially in Greece, point to an average annual degradation rate in the range of 1–4% [47]. These findings converge—on average—with the general trends resulting from the proposed method (Figure 17). Furthermore, the effect of heavy soiling is another important factor that may decrease energy production by up

to 6.5% on an annual basis [49]. Heavy soiling was observed during the spring of 2019 in Central Greece which resulted in an average decrease of 5.6% [27] in energy generation. This fact may be related to the observed deterioration in the models' performance for 2019. One must keep in mind that the dust and soiling effects are reversible, e.g., by the short-term accumulation and the subsequent meltdown of snow on the panels' surface. This fact may partially explain the characteristic error fluctuations from year to year in Figure 17.

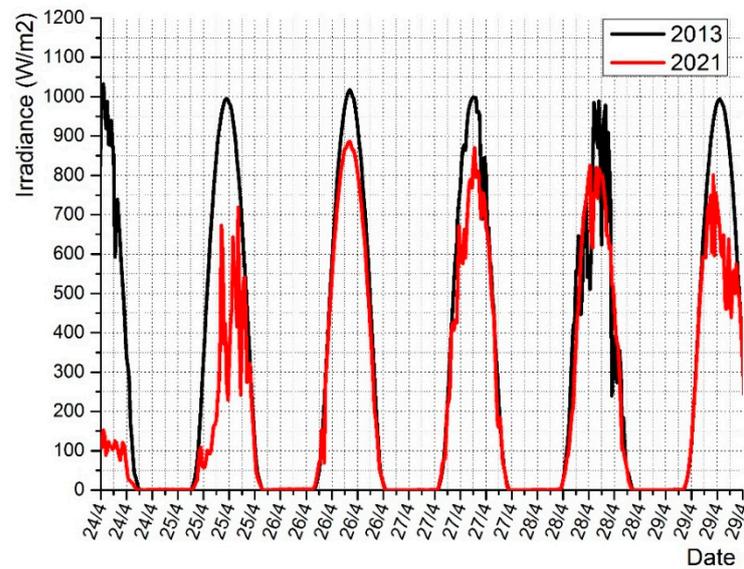


Figure 16. Comparison of the evolution of measured irradiance for six consecutive days of April between 2013 and 2021.

Table 2. Statistical and energy metrics for 2016–2021 (M1 and M2 versions).

	2016		2017		2018		2019		2020		2021	
	M1 ¹	M2 ²										
N [samples]	14,416		14,107		14,231		14,312		14,477		6233	6233
RMSE [W]	3.4	2.4	4.8	4.1	3.2	2.4	6.4	5.7	0.1	0.1	12.7	10.8
MBE [W]	97.6	122.5	64.1	74.6	90.1	108.4	302.6	314.5	43.4	83.7	−340.4	−270.3
MAPE [%]	−1.81	−2.27	−1.13	−1.32	−1.74	−2.10	−5.75	−5.97	−0.82	−1.59	6.28	4.99
MAE [W]	258.5	147.1	306.8	148.5	267.5	148.5	411.8	328.6	395.4	349.9	625.0	470.6
nRMSE [%]	6.6	4.6	8.8	7.4	6.2	4.6	12.5	11.2	9.7	9.8	15.9	13.5
COR	0.992	0.997	0.984	0.989	0.993	0.996	0.977	0.982	0.981	0.983	0.969	0.958
PR	0.87		0.88		0.86		0.84		0.89			
kWh/kWp	1530.67		1571.27		1459.5		1527.2		1545.89			
η_{average} [%]	13.93		13.99		13.99		13.47		13.97		15.17	
$\eta_{\text{standard deviation}}$ [%]	1.04		1.51		1.21		1.56		1.36		1.74	

¹ Regression model. ² Feed-forward artificial neural network trained with 3-years' data.

Now, degradation is an important factor that must be considered in the long-term prediction of a PV plant's performance. In the specific case study, due to the above-mentioned implications of fluctuating dust and soiling effects that overlap with the monotonous efficiency degradation, the short training period of three years may be the cause of the propagation of this instability. In order to further investigate this matter, it was decided to extend the training period of the ANN by two more years, to a total of five years. This is model version M3, which is added in Figure 17 and Table 3, for the years 2019 and 2020. Now, it is interesting to compare in Table 3 the prediction errors of this second scenario for

the remaining years 2019 and 2020. The improvement in the MAPE and nRMSE with the 5-year training of the ANN model is also apparent in Figure 17 (years 2019–2020).

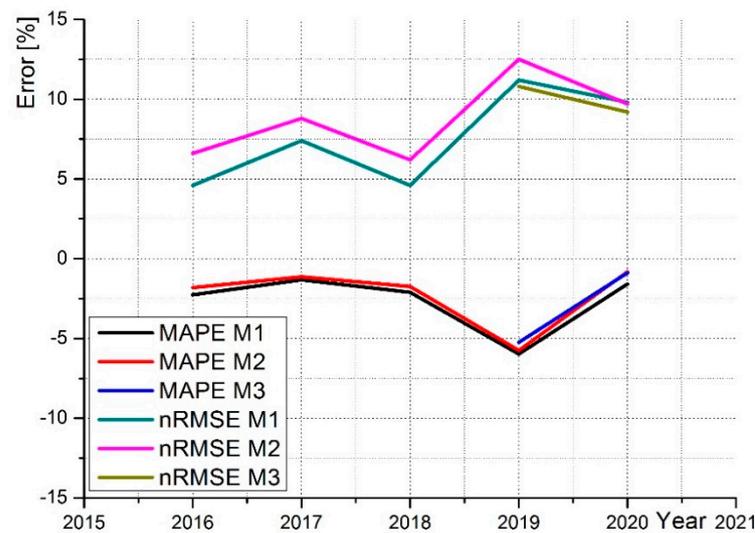


Figure 17. Evolution of nRMSE and MAPE over the years for the regression and ANN models.

Table 3. Statistical and energy metrics for 2019–2020 (all three model versions).

	2019			2020		
	M1 ¹	M2 ²	M3 ³	M1 ¹	M2 ²	M3 ³
N	14,312			14,477		
RMSE [W]	6.4	5.7	5.5	0.1	0.1	0.1
MBE [W]	302.6	314.5	276.2	43.4	83.7	46.6
MAPE [%]	−5.75	−5.97	−5.24	−0.82	−1.59	−0.89
MAE [W]	411.8	328.6	300.2	395.4	349.9	346.4
nRMSE [%]	12.5	11.2	10.8	9.7	9.8	9.2
COR	0.977	0.982	0.983	0.981	0.983	0.984
PR	0.84			0.89		
kWh/kWp	1527.2			1545.89		
η_{average} [%]	13.47			13.97		
$\eta_{\text{standard deviation}}$ [%]	1.56			1.36		

¹ Regression model. ² Feed-forward artificial network (3-year training period). ³ Feed-forward artificial network (5-year training period).

Table 3 compares the performance prediction for 2019 and 2020 for the regression model and FF ANN with 3-year training, and the FF ANN with 5-year training. The ANN models' performance is improved with the extended training; however, there is not yet available a fair ground for comparison between the two ANN training scenarios: the monitoring of the PV park must be extended for at least two more years in order to assess whether the extension of the training period improves the performance for long-term analysis.

5. Conclusions

Artificial Neural Networks along with traditional regression methods were applied in the long-term power output forecasting of a PV park, aimed at the reliable assessment of PV panels' efficiency degradation rate and fault diagnosis. The models employed are a linear regression model and a feed-forward ANN model. Available data from the 8-year monitoring period were employed. Initially, data from the first three years of operation

were used for training the ANN and feeding the regression model. As expected, the ANN model achieves smaller NRMSE errors than the regression model in the validation process. However, both model's prediction deteriorates over the years, presumably due to the fluctuating effect of PV panel soiling on the PV park's performance. Extension of the training period from 3 to 5 years led to an improvement in the neural network performance in short-term prediction. However, at least two more years of monitoring of the PV park would be required to assess the performance for long-term analysis. An additional outcome from the analysis for 2021 data was the spotting of an irradiance sensor's fault which was confirmed by the energy metrics, performance ratio, and yearly averaged PV performance. In its current state of development, the proposed methodology is appropriate for short-term analysis and as a fault diagnosis tool in long-term analysis. Further development in the methodology will be necessary to improve its long-term predictive capacity by decoupling the soiling effects from normal efficiency degradation.

Author Contributions: Conceptualization, E.R. and T.S.; methodology, E.R. and T.S.; software, E.R.; validation, E.R.; formal analysis, E.R. and T.S.; writing—original draft preparation, E.R.; writing—review and editing, T.S.; supervision, T.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

Nomenclature

AM	Air mass
ANFIS	Adaptive neuro-fuzzy inference system
ANN	Artificial neural network
BLSTM	Bidirectional long short-term memory
COR	Pearson correlation coefficient
DBN	Deep belief network
ELM	Extreme learning machine
FFANN	Feed-forward artificial neural network
GB	Gradient boosting
GDAS	Global data assimilation system
HETDS	Heterogeneous dynamic selection
KSI	Kolmogorov–Smirnov integral
LSTM	Long short-term memory
MAPE	Mean absolute percentage error
MBE	Mean bias error
MAPE	Mean absolute percentage error
MAE	Mean absolute error
MaxAE	Maximum absolute error
MLP	Multilayer perceptron
NN	Neural network
nRMSE	Normalized root mean square error
NWP	Numerical weather prediction
PR	Performance ratio
PV	Photovoltaics
RBFN	Radial basis function neural network
RF	Random forest
RMSE	Root mean square error
RNN	Recurrent neural network
STC	Standard test conditions
SARIMA	Seasonal autoregressive integrated moving average
SVM	Support vector machine
SVR	Support vector regression
TPE	Tree-structured Parzen estimator

Appendix A

Table A1. Technical data of the PV modules.

Yingli 60 Cell YGE SERIES			
Module Type		YL240P-29b	
		STC	NOCT
Power output	W	240	174.3
Module efficiency	%	14.7	13.3
Voltage at P_{max}	V	29.5	26.6
Current at P_{max}	A	8.14	6.56
Open-circuit voltage	V	37.5	34.2
Short-circuit current	A	8.65	7.01
Normal operating cell temperature (NOCT)	°C	46 ± 2	
Temperature coefficient of P_{max}	%/°C	−0.45	
Temperature coefficient of V_{oc}	%/°C	−0.33	
Temperature coefficient of I_{sc}	%/°C	0.06	
Temperature coefficient of V_{mpp}	%/°C	−0.45	
Dimensions(L/W/H)	Mm	1650/990/40	

STC: 1000 W/m² irradiance, 25 °C cell temperature, AM = 1.5, G spectrum according to EN 60904-3. Average relative efficiency reduction of 5% at 200 W/m² according to EN 60904-3. NOCT: open-circuit module operation temperature at 800 W/m² irradiance, 20 °C ambient temperature, 1 m/s wind speed.

Table A2. Technical characteristics and efficiency ratings of inverters.

Fronius IG Plus 150 V-3			
$P_{DC,MAX}$	W	12,770	
$I_{DC,MAX}$	A	55.5	
$U_{DC,MIN}$	V	230	
$U_{DC,START}$	V	260	
$U_{DC,R}$	V	370	
$U_{DC,MAX}$	V	600	
$P_{AC,R}$	W	12,000	
$I_{AC,MAX}$	A	17.4	
$U_{AC,R}$	V	3-NPE 400/230	
Maximum efficiency η_{inv}	%	95.9	
n_{inv} at 5% $P_{AC,R}$ (230 V/370 V/500 V)	%	91.8/92.5/91.1	
n_{inv} at 20% $P_{AC,R}$ (230 V/370 V/500 V)	%	94.7/95.1/94.6	
n_{inv} at 50% $P_{AC,R}$ (230 V/370 V/500 V)	%	95.3/95.9/95.3	
n_{inv} at 100% $P_{AC,R}$ (230 V/370 V/500 V)	%	94.0/95.2/95.1	
$P_{DC,MAX}$	W	12,770	
$I_{DC,MAX}$	A	55.5	
$U_{DC,MIN}$	V	230	
$U_{DC,START}$	V	260	

Table A3. Irradiance sensor characteristics.

Sensor	Mono-Crystalline Si-Sensor
Sensor voltage	75 mV at 1000 W/m ²
Accuracy	±5% (yearly average)
Ambient temperature	−40 to +85 °C
Design	Sensor mounted on z-shaped aluminum profile
Dimensions	55 × 55 × 10 mm
Fronius Product Nr.	43.0001.1189

Table A4. Temperature sensor characteristics.

Sensor	PT 100
Measuring range	−40 °C to +188 °C
Accuracy	±0.8 °C (in the range −40 °C to +100 °C)
Design	Sensor on adhesive film for surface measurements
Dimensions	Apendix32 × 32mm
Fronius Art. Nr.	43.0001.1190

Appendix B

Performance ratio calculation [50]:

$$Y_F = \frac{E}{P_{STC}} \left(\frac{\text{kWh}}{\text{kW}} \right)$$

$$Y_R = \frac{H}{G_{STC}} \left(\frac{\text{kWh}}{\text{kW}} \right)$$

$$PR = \frac{Y_F}{Y_R}$$

Airmass calculation [51]:

$$AM = \cos(z_s)^{-1} \frac{P}{P_0}$$

$$\frac{P}{P_0} = \exp^{(-0.0001184h)}$$

References

- Santos, B. EU wants rooftop PV mandate for public, commercial buildings by 2027, residential by 2029. *PV Magazine*, 18 May 2022.
- Monthly Report of Renewable Energy Sources and Cogeneration. March. 2022. Available online: <https://www.dapeep.gr/dimosieuseis/sinoptiko-pliroforiako-deltio-ape/> (accessed on 30 May 2022).
- Werner, J.H. How much photovoltaic efficiency is enough? *Solar* **2022**, *2*, 12. [CrossRef]
- Stamatellos, G.; Zogou, O.; Stamatelos, A. Interaction of a house's rooftop PV system with an electric vehicle's battery storage and air source heat pump. *Solar* **2022**, *2*, 11. [CrossRef]
- Hao, Y.; Dong, L.; Liang, J.; Liao, X.; Wang, L.; Shi, L. Power forecasting-based coordination dispatch of PV power generation and electric vehicles charging in microgrid. *Renew. Energy* **2020**, *155*, 1191–1210. [CrossRef]
- Kim, Y.; Seo, K.; Harrington, R.J.; Lee, Y.; Kim, H.; Kim, S. High accuracy modeling for solar PV power generation using noble BD-LSTM-based neural networks with EMA. *Appl. Sci.* **2020**, *10*, 7339. [CrossRef]
- Manasis, C.; Assimakis, N.; Vikias, V.; Ktena, A.; Stamatelos, T. Power generation prediction of an open cycle gas turbine using kalman filter. *Energies* **2020**, *13*, 6692. [CrossRef]
- Hassan, O.; Zakzouk, N.; Abdelsalam, A. Novel photovoltaic empirical mathematical model based on function representation of captured figures from commercial panels datasheet. *Mathematics* **2022**, *10*, 476. [CrossRef]
- Huld, T.; Amillo, A.M.G. Estimating PV module performance over large geographical regions: The role of irradiance, air temperature, wind speed and solar spectrum. *Energies* **2015**, *8*, 5159–5181. [CrossRef]
- Zamo, M.; Mestre, O.; Arbogast, P.; Pannekoucke, O. A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production, part I: Deterministic forecast of hourly production. *Sol. Energy* **2014**, *105*, 792–803. [CrossRef]
- Graditi, G.; Ferlito, S.; Adinolfi, G. Comparison of photovoltaic plant power production prediction methods using a large measured dataset. *Renew. Energy* **2016**, *90*, 513–519. [CrossRef]
- Sundaram, S. Improved ANN model for predicting the AC energy output of a realistic photovoltaic grid connected PV system. In Proceedings of the 2019 IEEE 1st International Conference on Energy, Systems and Information Processing (ICESIP), Chennai, India, 4–6 July 2019.
- Kardakos, E.G.; Alexiadis, M.C.; Vagropoulos, S.I.; Simoglou, C.K.; Biskas, P.N.; Bakirtzis, A.G. Application of time series and artificial neural network models in short-term forecasting of PV power generation. In Proceedings of the 48th International Universities Power Engineering Conference, Dublin, Ireland, 2–5 September 2013; pp. 1–6. [CrossRef]
- Gómez, J.L.; Martínez, A.O.; Pastoriza, F.T.; Garrido, L.F.; Álvarez, E.; García, J.O. Photovoltaic power prediction using artificial neural networks and numerical weather data. *Sustainability* **2020**, *12*, 10295. [CrossRef]

15. Kothona, D.; Panapakidis, I.P.; Christoforidis, G.C. An hour-ahead photovoltaic power forecasting based on LSTM model. In Proceedings of the 2021 IEEE Madrid PowerTech, Madrid, Spain, 28 June–2 July 2021; pp. 1–5. [CrossRef]
16. Nour-Eddine, I.O.; Lahcen, B.; Fahd, O.H.; Amin, B.; Aziz, O. Power forecasting of three silicon-based PV technologies using actual field measurements. *Sustain. Energy Technol. Assess.* **2020**, *43*, 100915. [CrossRef]
17. Akhter, M.N.; Mekhilef, S.; Mokhlis, H.; Almohaimeed, Z.M.; Muhammad, M.A.; Khairuddin, A.S.M.; Akram, R.; Hussain, M.M. An hour-ahead PV power forecasting method based on an RNN-LSTM model for three different PV plants. *Energies* **2022**, *15*, 2243. [CrossRef]
18. Natarajan, Y.; Kannan, S.; Selvaraj, C.; Mohanty, S.N. Forecasting energy generation in large photovoltaic plants using radial belief neural network. *Sustain. Comput. Inform. Syst.* **2021**, *31*, 100578. [CrossRef]
19. Pan, M.; Li, C.; Gao, R.; Huang, Y.; You, H.; Gu, T.; Qin, F. Photovoltaic power forecasting based on a support vector machine with improved ant colony optimization. *J. Clean. Prod.* **2020**, *277*, 123948. [CrossRef]
20. Pasion, C.; Wagner, T.; Koschnick, C.; Schuldt, S.; Williams, J.; Hallinan, K. Machine learning modeling of horizontal photovoltaics using weather and location data. *Energies* **2020**, *13*, 2570. [CrossRef]
21. Jaber, M.; Hamid, A.S.A.; Sopian, K.; Fazlizan, A.; Ibrahim, A. Prediction model for the performance of different PV modules using artificial neural networks. *Appl. Sci.* **2022**, *12*, 3349. [CrossRef]
22. Villanueva, D.; San-Facundo, D.; Miguez-García, E.; Fernández-Otero, A. Modeling and simulation of household appliances power consumption. *Appl. Sci.* **2022**, *12*, 3689. [CrossRef]
23. Kalogirou, S.A.; Bojic, M. Artificial neural networks for the prediction of the energy consumption of a passive solar building. *Energy* **2000**, *25*, 479–491. [CrossRef]
24. Laurino, M.; Piliouguine, M.; Spagnuolo, G. Artificial neural network based photovoltaic module diagnosis by current-voltage curve classification. *Sol. Energy* **2022**, *236*, 383–392. [CrossRef]
25. Chine, W.; Mellit, A.; Lughi, V.; Malek, A.; Sulligoi, G.; Pavan, A.M. A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks. *Renew. Energy* **2016**, *90*, 501–512. [CrossRef]
26. Pavan, A.M.; Mellit, A.; De Pieri, D.; Kalogirou, S. A comparison between BNN and regression polynomial methods for the evaluation of the effect of soiling in large scale photovoltaic plants. *Appl. Energy* **2013**, *108*, 392–401. [CrossRef]
27. Roumpakias, E.; Stamatelos, T. Surface dust and aerosol effects on the performance of grid-connected photovoltaic systems. *Sustainability* **2020**, *12*, 569. [CrossRef]
28. Ahmed, R.; Sreeram, V.; Mishra, Y.; Arif, M. A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization. *Renew. Sustain. Energy Rev.* **2020**, *124*, 109792. [CrossRef]
29. De O. Santos, D.S.; de Mattos Neto, P.S.G.; de Oliveira, J.F.L.; Siqueira, H.V.; Barchi, T.M.; Lima, A.R.; Madeiro, F.; Dantas, D.A.P.; Converti, A.; Pereira, A.C.; et al. Solar irradiance forecasting using dynamic ensemble selection. *Appl. Sci.* **2022**, *12*, 3510. [CrossRef]
30. 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]
31. Kazem, H.A.; Yousif, J.H.; Chaichan, M.T.; Al-Waeli, A.H.; Sopian, K. Long-term power forecasting using FRNN and PCA models for calculating output parameters in solar photovoltaic generation. *Heliyon* **2022**, *8*, e08803. [CrossRef]
32. Nguyen, H.-P.; Liu, J.; Zio, E. A long-term prediction approach based on long short-term memory neural networks with automatic parameter optimization by Tree-structured Parzen Estimator and applied to time-series data of NPP steam generators. *Appl. Soft Comput.* **2020**, *89*, 106116. [CrossRef]
33. Zhang, Y.; Xiong, R.; He, H.; Pecht, M.G. Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries. *IEEE Trans. Veh. Technol.* **2018**, *67*, 5695–5705. [CrossRef]
34. Ma, R.; Yang, T.; Breaz, E.; Li, Z.; Briois, P.; Gao, F. Data-driven proton exchange membrane fuel cell degradation prediction through deep learning method. *Appl. Energy* **2018**, *231*, 102–115. [CrossRef]
35. Guo, L.; Li, N.; Jia, F.; Lei, Y.; Lin, J. A recurrent neural network based health indicator for remaining useful life prediction of bearings. *Neurocomputing* **2017**, *240*, 98–109. [CrossRef]
36. Cheng, Y.; Zhu, H.; Wu, J.; Shao, X. Machine health monitoring using adaptive kernel spectral clustering and deep long short-term memory recurrent neural networks. *IEEE Trans. Ind. Inform.* **2018**, *15*, 987–997. [CrossRef]
37. Konstantas, G.; Stamatelos, A. Quality assurance of exhaust emissions test data. *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.* **2004**, *218*, 901–914. [CrossRef]
38. Roumpakias, E.; Zogou, O.; Stamatelos, A. Correlation of actual efficiency of photovoltaic panels with air mass. *Renew. Energy* **2015**, *74*, 70–77. [CrossRef]
39. Tensorflow. An Open Source Machine Learning Framework for Everyone. Available online: <https://github.com/tensorflow/tensorflow> (accessed on 30 May 2022).
40. Keras. Deep Learning API Written in Python: Layer Activation Functions. Available online: <https://keras.io/api/layers/activations/> (accessed on 30 May 2022).
41. Marquardt, D.W. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* **1963**, *11*, 431–441. [CrossRef]
42. Di Marco, F. Tensorflow Implementation of Levenberg-Marquardt Training Algorithm. Available online: <https://github.com/fabiodimarco/tf-levenberg-marquardt> (accessed on 30 May 2022).

43. El-Baz, W.; Tzscheutschler, P.; Wagner, U. Day-ahead probabilistic PV generation forecast for buildings energy management systems. *Sol. Energy* **2018**, *171*, 478–490. [[CrossRef](#)]
44. Wang, F.; Xuan, Z.; Zhen, Z.; Li, K.; Wang, T.; Shi, M. A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework. *Energy Convers. Manag.* **2020**, *212*, 112766. [[CrossRef](#)]
45. Kasten, F.; Young, A.T. Revised optical air mass tables and approximation formula. *Appl. Opt.* **1989**, *28*, 4735–4738. [[CrossRef](#)]
46. Sauter, T.; Venema, V. Natural three-dimensional predictor domains for statistical precipitation downscaling. *J. Clim.* **2011**, *24*, 6132–6145. [[CrossRef](#)]
47. Roumpakias, E.; Stamatelos, A. Performance analysis of a grid-connected photovoltaic park after 6 years of operation. *Renew. Energy* **2019**, *141*, 368–378. [[CrossRef](#)]
48. Piliouguine, M.; Elizondo, D.; Mora-López, L.; Sidrach-De-Cardona, M. Modelling photovoltaic modules with neural networks using angle of incidence and clearness index. *Prog. Photovolt. Res. Appl.* **2014**, *23*, 513–523. [[CrossRef](#)]
49. Kaldellis, J.; Kokala, A. Quantifying the decrease of the photovoltaic panels' energy yield due to phenomena of natural air pollution disposal. *Energy* **2010**, *35*, 4862–4869. [[CrossRef](#)]
50. Marion, B.; Adelstein, J.; Boyle, K.; Hayden, H.; Hammond, B.; Fletcher, T.; Canada, B.; Narang, D.; Kimber, A.; Mitchell, L.; et al. Performance parameters for grid-connected PV systems. In Proceedings of the Conference Record of the Thirty-first IEEE Photovoltaic Specialists Conference, Lake Buena Vista, FL, USA, 3–7 January 2005; pp. 1601–1606. [[CrossRef](#)]
51. Wang, H.; Muñoz-García, M.; Moreda, G.; Alonso-Garcia, C. Seasonal performance comparison of three grid connected photovoltaic systems based on different technologies operating under the same conditions. *Sol. Energy* **2017**, *144*, 798–807. [[CrossRef](#)]