

## Article

# Intra-Day Solar Power Forecasting Strategy for Managing Virtual Power Plants

Guillermo Moreno <sup>1</sup>, Carlos Santos <sup>2</sup>, Pedro Martín <sup>1</sup>, Francisco Javier Rodríguez <sup>1,\*</sup>, Rafael Peña <sup>2</sup> and Branislav Vuksanovic <sup>3</sup>

<sup>1</sup> Department of Electronics, University of Alcalá, Alcalá de Henares, 28805 Madrid, Spain; guillermo.morenob@uah.es (G.M.); pedro.martin@uah.es (P.M.)

<sup>2</sup> Department of Signal Theory and Communications, University of Alcalá, Alcalá de Henares, 28805 Madrid, Spain; carlos.santos@uah.es (C.S.); rafael.pena@uah.es (R.P.)

<sup>3</sup> School of Engineering, University of Portsmouth, Winston Churchill Ave., Portsmouth PO1 3HJ, UK; branislav.vuksanovic@port.ac.uk

\* Correspondence: franciscoj.rodriguez@uah.es; Tel.: +34-91-885-6561

**Abstract:** Solar energy penetration has been on the rise worldwide during the past decade, attracting a growing interest in solar power forecasting over short time horizons. The increasing integration of these resources without accurate power forecasts hinders the grid operation and discourages the use of this renewable resource. To overcome this problem, Virtual Power Plants (VPPs) provide a solution to centralize the management of several installations to minimize the forecasting error. This paper introduces a method to efficiently produce intra-day accurate Photovoltaic (PV) power forecasts at different locations, by using free and available information. Prediction intervals, which are based on the Mean Absolute Error (MAE), account for the forecast uncertainty which provides additional information about the VPP node power generation. The performance of the forecasting strategy has been verified against the power generated by a real PV installation, and a set of ground-based meteorological stations in geographical proximity have been used to emulate a VPP. The forecasting approach is based on a Long Short-Term Memory (LSTM) network and shows similar errors to those obtained with other deep learning methods published in the literature, offering a MAE performance of 44.19 W/m<sup>2</sup> under different lead times and launch times. By applying this technique to 8 VPP nodes, the global error is reduced by 12.37% in terms of the MAE, showing huge potential in this environment.

**Keywords:** power forecasting; long short-term memory recurrent neural network (LSTM-RNN); virtual power plant (VPP)



**Citation:** Moreno, G.; Santos, C.; Martín, P.; Rodríguez, F.J.; Peña, R.; Vuksanovic, B. Intra-Day Solar Power Forecasting Strategy for Managing Virtual Power Plants. *Sensors* **2021**, *21*, 5648. <https://doi.org/10.3390/s21165648>

Academic Editor: Hossam A. Gabbar

Received: 29 July 2021

Accepted: 20 August 2021

Published: 22 August 2021

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



**Copyright:** © 2021 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

Around the world, the full deployment of solar energy is being facilitated by several factors including, but not limited to, the reduced price of solar panels; environmental, political and social concerns; and solar energy undercutting utility prices, inter alia. According to [1] global installed capacity will double every two years; however, significant factors have been identified which impede the speed at which solar dominance can be achieved: (i) lack of investments in efficiency, (ii) insufficient government incentives, and (iii) regulatory constraints. Small-scale Photovoltaic (PV) installations such those in the residential sector benefit from self-consumption by shifting a load from hours when electricity prices are high to hours when the PV energy is being generated, thereby achieving electricity bill savings. Going one step further, the aggregation and coordination of several PV installations in the shape of a Virtual Power Plant (VPP) with the accurate forecasting of global production facilitates its integration into the network [2]. Consequently, the increasing PV penetration can lead to the increasing aggregation of PV systems into VPPs. However, these new business models are difficult to implement due to the previously mentioned regulatory constraints.

Power forecasting along with load demand and energy prices, for different time horizons and resolutions, are factored into the equation. For VPPs, spatial horizons should also be considered. Forecasting methods can be classified according to different factors, such as: the forecasted parameter (irradiance or power), the time horizon and resolution, the lead time, the model approach, and the nature of the forecasting statistic. Regarding the forecasted parameter, two different alternatives exist: direct [3] and indirect [4]. The direct method predicts the solar power through historical datasets of PV power generation and weather conditions. Indirect forecasting differs from the direct method in that it firstly predicts the solar irradiance and then, the solar power is calculated by using a performance model of the PV plant. As far as the time horizon is concerned, four categories can be found [5,6]: nowcasting (from 1 min to several minutes) which is used for real-time optimization in Energy Management Systems (EMSs); short-term forecast (from 1 h to several hours) used for intra-day market participation and for day-ahead operation optimization; medium term forecast (from 1 month to 1 year); and long-term forecast (up to several years). Time resolutions may range from 1 minute for real-time market operations, and 15-minute periods for load-shifting strategies and for optimizing Battery Energy Storage Systems (BESSs), to 1 hour for longer time horizons used by consumption monitoring, and a 1-week resolution for 1-year time horizons which can be used to identify consumption trends [7]. The lead time can be defined as the time difference between the instant when the forecast is launched and the occurrence of the forecasted value, considering the forecast horizon as the maximum forecast lead time. Forecast errors increase with forecast lead time due to the atmospheric motion. As for the model, the optimal method for solar irradiance prediction depends on the forecast lead time [8]. In this regard, four approaches have been widely used [9]: (a) time-series-based statistical models whose aim is to identify patterns between historical datasets and the output parameters; (b) machine learning (ML) models mainly based on artificial neural networks (ANNs), which use historical datasets to learn the dependency between the past and the future; (c) physical strategies which utilize Numerical Weather Prediction (NWP) and PV models for solar power forecasting; and (d) hybrid models which explore different algorithm combinations with the aim of improving forecast accuracy and reducing the computational burden of online forecasting applications [5]. The objective of all the models is to improve forecasting accuracy by minimizing some quality metrics, usually the sum of squared errors. The existence of different models raises the question of whether one method is better than the others. This is particularly true for statistical and ML models. Some studies conclude that statistical models outperform ML models [10] while others state the opposite [11,12]. However, this interpretation may appear to be fairly simplistic without taking into account the dataset size [13], the variable being forecast [14], the time horizon [15], or the computational load [16]. Although historically, the forecasts have been dominated by statistical methods, over the last decade there has been a significant shift toward ML strategies [17]. This comparative study is beyond the scope of the paper.

Regardless of the method used, the existence of forecasting errors poses a major challenge in optimizing the PV plant operation. While minor forecasting errors may not adversely affect the PV plant operation, larger errors can produce negative effects in the optimization models. Uncertainties hinder the performance in terms of accurately assessing the variables during the PV plant scheduling and operation. Forecast uncertainty quantification is, therefore, crucial. For this reason, considering the prediction intervals, which account for the uncertainty, provides additional accurate information about the expected values in terms of the range of plausible values and the probability assigned to each of them [17,18]. Another solution to the problem involves the aggregation of several PV sites for a unique forecasting strategy, since the error is significantly reduced as the number of installations increases. To prove this, in [19] the authors present an approach to forecast the PV power from irradiance prediction maps, obtaining the power forecast of 200 sites located in Germany. Results show that the error is reduced from a Root Mean Square Error (RMSE) of 0.11 kW/kW<sub>peak</sub> for single sites, to 0.06 kW/kW<sub>peak</sub>

for an area of  $220 \text{ km} \times 220 \text{ km}$  with multiple sites. The distance among sites is also an important factor which influences accuracy, since the error is significantly reduced when the distance between facilities increases. This strategy provides a powerful solution in the context of VPPs, since multiple systems or nodes are controlled, managing Distributed Generation (DG) units, Energy Storage Systems (ESSs), flexible loads and Information and Communication Technologies (ICTs) [20]. Regarding the types of DG units, PV systems can be considered as the easiest and most cost-effective Renewable Energy Sources (RESs) to exploit, mainly for households, where it is possible to turn PV installations into flexible VPP nodes [21].

Finally, as stated above, for indirect forecasting approaches, performance models of PV systems are required to obtain the prediction of solar power generation. To this end, a strategy that works under arbitrary conditions of irradiance and temperature must be adopted. Methods that exhibit these key characteristics are the Osterwald's method [22], which stands out by its simplicity, or similar studies from the literature that improve the performance of the Osterwald's method by adjusting the results under low irradiance levels [23,24]. When the operating point of the PV panels is known, alternative methods, such as those reported in [25,26], can improve accuracy, while other research uses parametrization models to simplify the process [27]. Sometimes, the irradiance of the site is measured on a horizontal plane, obtaining the Global Horizontal Irradiance (GHI). However, the panels are on a different plane. This is typical in satellite measurements but can also be the case in installations with multiple Maximum Power Point Trackers (MPPT) or PV panels with axis trackers. To solve this problem, a conversion process is needed, using: (i) different expressions to tackle the problem step-by-step by separating the global components into direct irradiance, diffuse irradiance, and albedo, modifying the angle of these components to obtain the global irradiance on the plane of the panel, estimating its losses to obtain the effective irradiance, or (ii) an approach that simplifies the process [28]. In this regard, it becomes crucial to reduce the complexity and the computational burden placed on the forecasting algorithms. With this in mind, this work makes use of the Osterwald's method to calculate the PV power, since low irradiance values ( $G < 125 \text{ W/m}^2$ ) are barely existent in the dataset and a generalization of the algorithm for VPP environments leads to better results. Satellite data are also required in this work since they offer information on the GHI, which is converted into irradiance on the tilted plane by following the steps stated above.

The forecasting strategy developed in this paper, uses long short-term memory recurrent neural networks (LSTM-RNNs) and is based on an indirect approach in which the irradiance is forecasted first and the output power is calculated by using the PV model. LSTM-RNNs have been used in several works, achieving satisfactory results on account of their recurrent architecture, which includes memory units [16]. These allow the ANN to identify temporal patterns from the historical data of the forecast variable, thereby reducing the forecast error in comparison to other alternatives. The authors in [29] propose a PV power forecasting strategy based on LSTM-RNN which is compared with other methods without memory units, showing their limitations in terms of not being able to model the dynamics of the PV output power data. In [30] a LSTM-RNN with only exogenous inputs, e.g., dry bulb and wet bulb temperatures, and relative humidity, is used to forecast the day-ahead solar irradiance.

The main contributions of this paper are summarized as follows: (i) the PV forecasting method is applied to a VPP environment to reduce the forecasting error, which is modelled as a function of two well-defined parameters called lead time and launch time; (ii) prediction intervals are used to model the forecast uncertainty as a function of not only the lead time and the launch time, but also the Cloud Cover Factor (CCF), which allows the type of day to be identified; (iii) the input data for the forecasting strategy are derived from free-of-charge open-access data sources, offering a viable and cost-effective solution; and (iv) a trade-off between accuracy and computational burden facilitates the application of multiple PV power forecasts at different locations, within the context of a VPP.

The remainder of this paper is organized as follows: Section 2 introduces the framework for the intra-day power forecasting strategy; the experimental results are presented in Section 3; and finally, some conclusions are drawn in Section 4.

## 2. Intra-Day Power Forecasting Framework

The proposed intra-day power forecasting strategy is depicted in Figure 1. It consists of four main blocks, namely: (i) input data; (ii) data preprocessing; (iii) model design and forecasting; and (iv) VPP coordination. The input data, which come from different sources, are fed to the preprocessing stage. The preprocessing step prepares the data as required by the training and forecasting models. Finally, the output of the forecasting algorithms is used as the input of the EMS of the VPP. In the following, the different parts are explained in detail.

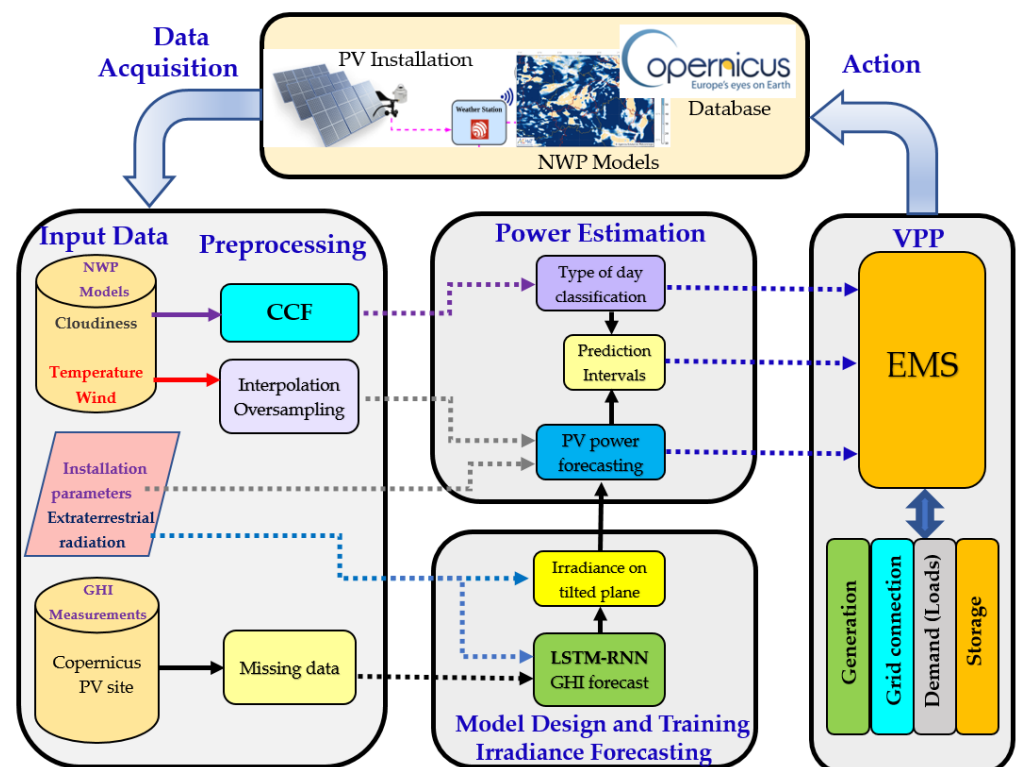


Figure 1. Forecasting framework.

### 2.1. Input Data

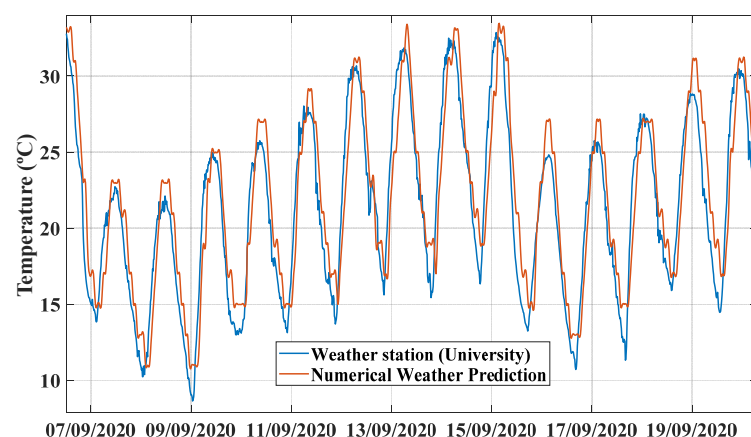
The input data consist of three specific categories according to the source and type of the information provided. The first category includes cloudiness and temperature, which are obtained from forecast maps, at different spatial and temporal scales, generated and regularly published by the Spanish agency of meteorology AEMET, via NWP [31]. The cloudiness dataset is used to define the Cloud Cover Factor (CCF), which indicates to what extent a cloud area on the NWP-based cloudiness maps creates shadows on the PV installation. This parameter is used to define the type of day: sunny, cloudy, and overcast. This allows the dataset to be split in different groups to create prediction intervals. Temperature data, on the other hand, are used to estimate the cell temperature of the solar panel at the prediction instant [32]. NWP-based weather maps are of great interest since some useful weather variables might not be available in solar installations. The deviation in the estimation of the cell temperature is then assessed by using the data obtained from the experimental setup, which is located at the Polytechnic School of the University of Alcalá (Spain) and consists of a 2.97 kW<sub>p</sub> PV facility with a meteorological station that gathers information of GHI, temperature and cell temperature [33]. The dataset, obtained

from the PV facility, is taken during the period between 1 June 2020 and 31 May 2021, with a resolution of 15 min. In the second category, the Global Horizontal Irradiance (GHI) measurements are obtained from two sources: (i) a pyrometer, which is installed in the experimental setup and 30-second GHI measurements are taken and stored on the cloud (ThingSpeak) [34]; and (ii) the Copernicus Atmosphere Monitoring Service (CAMS), which provides a free historical dataset of the incoming surface solar irradiance that can be used for any purpose. The data accuracy is ensured by a regular quality control against information from in situ systems such as ground stations [35]. At the PV facility, the Mean Absolute Error (MAE) committed for the temperature with respect to NWP maps is 2.12 °C. Likewise, the MAE obtained between the CAMS and the PV station is 46.97 W/m<sup>2</sup>, for the whole year of measurements. This database is used to provide the forecasting models with a large GHI dataset for training purposes. Finally, the third category comprises non-stochastic data, such as sun position, used for the CCF calculation to determine the type of day; the extraterrestrial radiation for generating the forecasts and working out the irradiance on the tilted plane of the PV modules; and the installation parameters which are required for the PV power forecasting, as is explained in the following sections.

## 2.2. Data Preprocessing

The information obtained from the NWP-based weather forecasts must be transformed into numerical values. The forecasting time resolution is set to 15 min, mainly to follow the European Electricity Market Directive to be implemented in the coming years, which sets 15-minute energy matching periods. However, the AEMET only generates the weather maps hourly. This poses the inherent problem of merging time series with different time steps. For instance, for the PV power forecasting, the cell temperature (based on the ambient temperature) and the irradiance on the tilted plane are required. Since the latter has a time resolution of 15 min, so too should the time resolution of the time series for the cell temperature. To this end, quadratic interpolation is performed to create an oversampling of the NWP time series. Changes in the ambient temperature are usually smooth and it is assumed that the measurements shown in the NWP maps are defined with their intermediate values, since the Darboux property [36] is accomplished.

To prove the accuracy of this approach, Figure 2 depicts the ambient temperature obtained from the AEMET forecasts with respect to the values measured by a weather station located in the PV installation. The remarkable accuracy of the weather forecast for the temperature is noticeable.



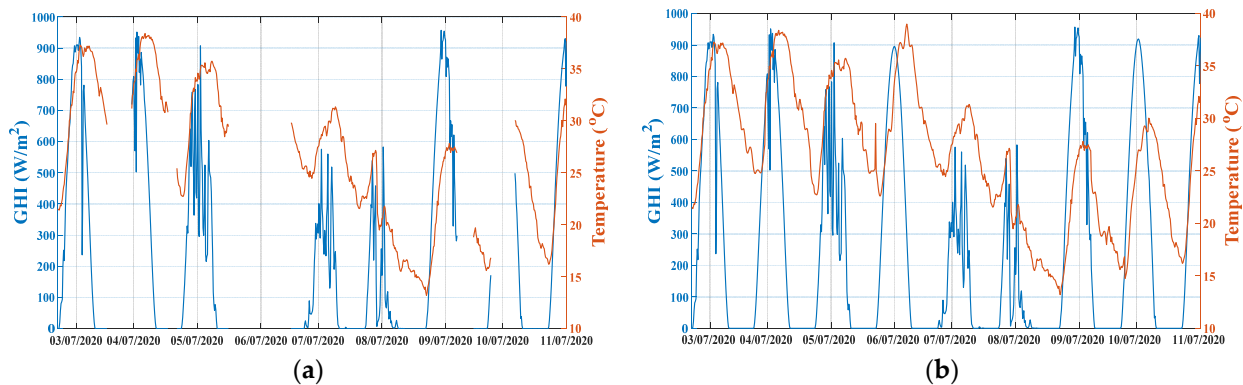
**Figure 2.** A comparison between the ambient temperature measured at the station and the temperature obtained from the AEMET website.

The CCF, on the other hand, is obtained by processing cloudiness information from weather maps. This parameter, which allows the type of the day to be defined, is used to identify those periods of time for which the presence of clouds can alter the PV power



generation over a region through blocking the sun's radiation. The CCF is obtained using a similar method as the work presented in [37], which provides a detailed description of how to calculate this parameter; mainly by detecting cloud-contaminated pixels in the weather maps that interfere between the sun and the installation.

Finally, missing data can negatively affect the accuracy of the forecasts. To fill the missing gaps in the temperature and GHI datasets obtained from the weather station in the PV installation, GHI satellite data and the data from the NWP-based weather forecasts are used. Figure 3 shows an example of the reconstruction of missing data for the temperature and irradiance time series.



**Figure 3.** (a) Missing data of the GHI and the temperature on the site; (b) Time series reconstruction of the GHI and the temperature on the site.

### 2.3. Model Design and Irradiance Forecasting

The third part in the forecasting framework deals with the LSTM-RNN-based model design and the forecasting itself, which aims to: (a) predict the mean PV power for a particular day with a 15-minute time step at the experimental PV facility, and (b) compute prediction intervals intended to show the likely uncertainty in the forecasting outcome [17]. This information constitutes an important input for the EMS in the VPP.

Figure 4 shows the flowchart of the model design and forecasting. The forecasting process starts with the LSTM-RNN model definition based on an iterative approach. Five years of GHI measurements from the Copernicus databases are utilized in the training process. The LSTM-RNN architecture depends on the characteristics of the input and output data and the cross-validation process. When creating the LSTM-RNN, 10% of the training set is used as the cross-validation set, optimizing the number of hidden layer units, mini-batch sizes, regularization factors, learn rate, and epochs (Table 1). Once these parameters are defined, the algorithm is extended to be used for future forecasts. The error in the training process is minimized by computing the RMSE, taking into account not only the proper convergence of the system but the computational time of the process. Squared errors lead the convergence in the LSTM-RNN as they are responsible for avoiding atypical errors, which have remarkable importance in energy management tasks. The architecture is composed of two input layers, one recurrent hidden layer (based on fifty memory blocks), and one output layer (Table 1). The memory block includes one or more self-connected memory cells along with four multiplicative gates (input, output, update, and forget gates). These gates provide the mechanism whereby the information can be stored and accessed over long periods of time, thereby avoiding the vanishing and exploding gradient problem posed by the conventional RNNs [38], e.g., the activation of the cell can be delayed, providing that the input gate remains closed to new inputs which can later become available by opening the output gate. The purpose of LSTM-RNN is, therefore, to model long-range dependencies. When training with sequential data, Gated Recurrent Unit (GRU), LSTM-RNN, and the Convolutional Neural Network (CNN)-LSTM are predominant in the literature [16]. As for CNN-LSTM models, they ensure higher accuracies for predictions based on more features which significantly compromise the

computational time. It is worth noting that only two variables are used in this work. In [39] the authors show that these deep learning techniques ensure a higher accuracy than conventional ANNs or Support Vector Machines (SVMs) in GHI short-term forecasting. Consequently, LSTM-RNNs are used in this paper for the forecasting process. LSTM-RNNs achieve remarkable forecast accuracy with different prediction intervals, on account of their ability to memorize long historical data and determine the optimal time lags for the time series. These features are fundamental in the context of irradiance forecasting since there is no previous knowledge of the relationship between forecasts and the length of the historical dataset.

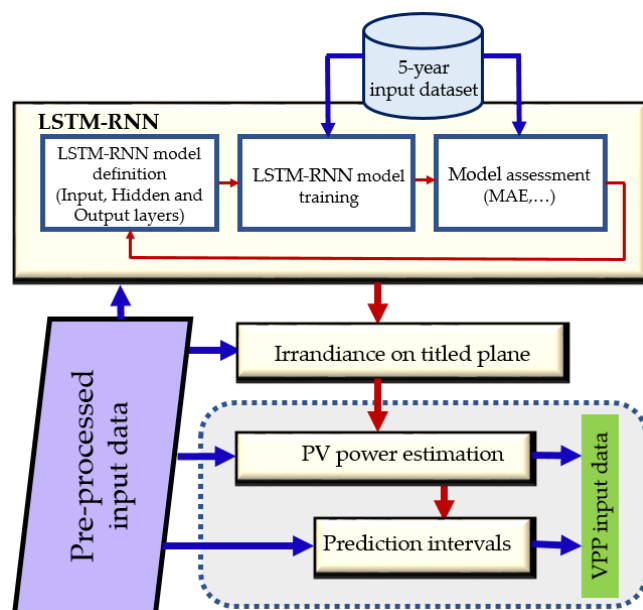


Figure 4. Flowchart for the LSTM-RNN-based forecasting model design.

Table 1. Parameters selected in the LSTM-RNN.

Number of Features	2 (GHI, Extra-Terrestrial Radiation)
Hidden layer units	50
Number of responses	1
Mini-batch size	256
Regularization factor	$5 \times 10^{-4}$
Optimizer	Adam ( $\beta_1 = 0.9, \beta_2 = 0.999, \epsilon = 1 \times 10^{-8}$ )
Initial learn rate	0.01
Learn rate schedule	Piecewise (periodically)
Learning drop	0.5 every 20 epochs
Epochs	70
Limited gradient	1

Once the LSTM-RNN model has been devised, the GHI prediction is made, followed by the estimation of the effective irradiance on the tilted plane of the PV module. Firstly, the calculation of the effective irradiance uses information from the two components of irradiance in the horizontal plane (direct and diffuse, since the albedo is zero in this case), calculated as a function of the clarity index ( $k_{th}$ ), to obtain the diffuse fraction ( $k_{dh}$ ) [40]. Once this information is obtained, the conversion into the tilted plane is estimated with the diffuse irradiance [41] and the albedo:

$$albedo = r_o ghm_0 (1 - \cos\beta)/2 \quad (1)$$

where  $r_o$  is the albedo coefficient, considering that a value of 0.2,  $ghm_0$  is the GHI and  $\beta$  is the tilted angle of the panels. Finally, the effective irradiance is determined by considering angular [42] and spectral [43] losses for p-Si modules and a typical moderate dust degree of  $DT = 0.97$  for the installation.

The Osterwald's model [22] is used to convert the effective irradiance into PV power:

$$P_{DC} = SF \eta_{DC} P_{peak} \frac{G_{panel}}{G_{STC}} (1 + \delta P_m (T_{cell} - T_{cell,STC})), \quad (2)$$

where  $P_{DC}$  is the PV power forecasted;  $SF$  represents the shading losses due to the surroundings of the installation, determined in Section 3.2 for this particular case;  $\eta_{DC} = 0.927$  includes wiring losses, module tolerances and mismatch losses;  $P_{peak} = 2.97$  kW is the peak power of the installation;  $G_{panel}$  is the effective irradiance of the panels previously calculated;  $G_{STC} = 1$  kW/m<sup>2</sup> is the irradiance under Standard Test Conditions; (STC),  $\delta P_m = -0.4\%/^{\circ}\text{C}$  is the temperature coefficient of the PV panels of the installation;  $T_{cell}$  is the cell temperature; and  $T_{cell,STC}$  is the cell temperature under STC.

The cell temperature can be determined with the following expression, assuming the wind speed is negligible, since it can be considered as a nonsignificant effect complex to model because the wind does not affect each panel in the facility equally:

$$T_{cell} = \frac{T_{cell,NOCT} - T_{amb,NOCT}}{G_{NOCT}} G_{panel} + T_{amb}, \quad (3)$$

where  $T_{cell,NOCT} = 45$  °C is the cell temperature under Normal Operating Cell Temperature (NOCT) conditions;  $T_{amb,NOCT} = 20$  °C is the ambient temperature under NOCT conditions;  $G_{NOCT} = 0.8$  kW/m<sup>2</sup> is the irradiance under NOCT conditions; and  $T_{amb}$  is the ambient temperature, obtained from NWP forecasts.

Then, with the historical dataset of PV power forecasts, it is possible to compute prediction intervals for new forecasts. A prediction interval is an interval estimate for an unknown future value [17] which can be regarded as a random variable at the time when the prediction is made. In this paper, statistical prediction intervals are employed based on the work presented in [44], considering a Laplacian distribution model for the error as a function of the lead time, the launch time, and the type of day. Figure 5 shows the intervals for a specified day with 90% confidence, providing additional, valuable information from the forecast. PV power generation strongly depends on the weather conditions, the latter varying according to the season. This greatly hinders the ability of the forecasting algorithms to deliver accurate predictions, causing some degree of uncertainty which should be evaluated. Prediction intervals constitute the tool that can be used to express the degree of uncertainty of point forecasts which add a given confidence level. Additional details about the definition of the intervals, such as group selection and accuracy, are further explained in Section 3.3.



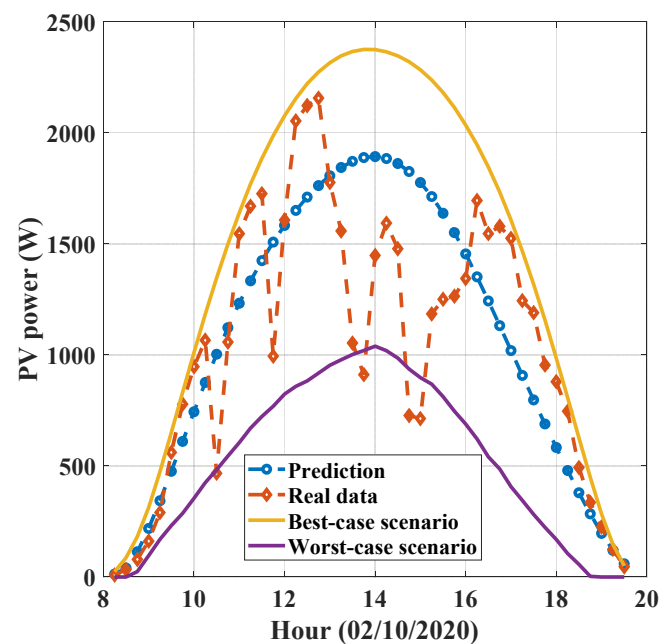


Figure 5. Prediction intervals with respect to the PV power.

### 3. Results

This section presents the results obtained by the proposed intra-day forecasting strategy for VPP, which is divided into different steps: (a) GHI forecasting for a real VPP node and for an emulated VPP; (b) PV power estimation from the GHI forecasting output; (c) the quantitative assessment of prediction intervals; and (d) VPP scheduling. Firstly, the results are validated for a real PV installation, which plays the role of a VPP node. The PV installation is located in the Polytechnic School, at the University of Alcalá (Madrid). Secondly, the strategy is developed for an emulated VPP, by using several ground-based meteorological stations uniformly spread over the Community of Madrid [33]. In order to evaluate the effectiveness of the model, a performance comparison in terms of accuracy/error, with respect to other methods proposed in literature, is also performed.

#### 3.1. LSTM-RNN-Based GHI Forecasting for a Real VPP Node

The LSTM-RNN-based GHI forecasting for the real VPP node is performed by using measurements of irradiance taken in the PV facility located at the Polytechnic School of the University of Alcalá (Spain). The initial training dataset is based on a 5-year period of irradiance values obtained from the CAMS dataset, since RNNs require a large amount of data for the learning process and GHI measurements are scarce in new installations. However, the test dataset is based on real measurements taken during the period from 1 June 2020 to 31 May 2021. Therefore, a whole year of real GHI values under different seasonal weather conditions are used to assess the accuracy of the forecasting approach. With a resolution of 15 min, the forecasting process starts at sunrise and ends at sunset. Furthermore, a new prediction is launched every 15 min and the dataset of irradiance is then updated, which ensures the accuracy of the results obtained. The network is trained with new measurements every day, during the night, to yield the best results. The GHI forecasts are given as a function of both the launch time and the lead time, parameters which are further defined, with the aim of computing the prediction intervals.

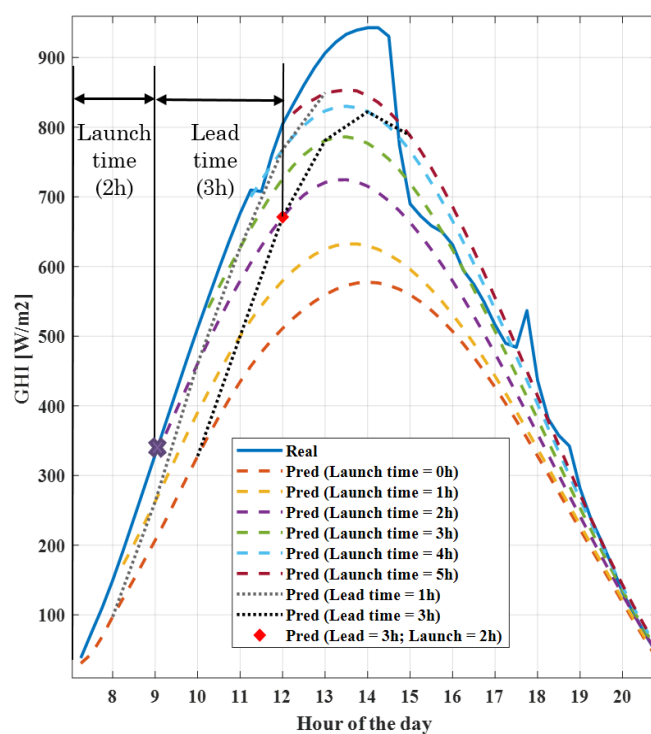
As far as the error assessment is concerned, this work relies on two types of metrics: (i) scale-dependent metrics such as the MAE and the Root Mean Square Error (RMSE); (ii) percentage-error metrics, such as the relative Mean Absolute Error (rMAE); and (iii) the relative Root Mean Square Error (rRMSE). Absolute values provide information about the average forecasting whereas the quadratic values are more sensitive to outliers, the combined analysis of the two allows for a thorough study of the results. Error percentage

values, on the other hand, provide an intuitive understanding of the error committed, which allows for a fair comparison to be conducted since the dependence on the magnitude is removed. However, when these values are near zero, scale-dependent metrics constitute the preferred option. The error metrics are summarized in Table 2, where  $Y_t$  is the measured data at time  $t$ ,  $\hat{Y}_t$  is the forecast value at time  $t$ , and  $T$  is the length of the time series used to assess the accuracy of the algorithm.

**Table 2.** Metrics used to evaluate the model performance.

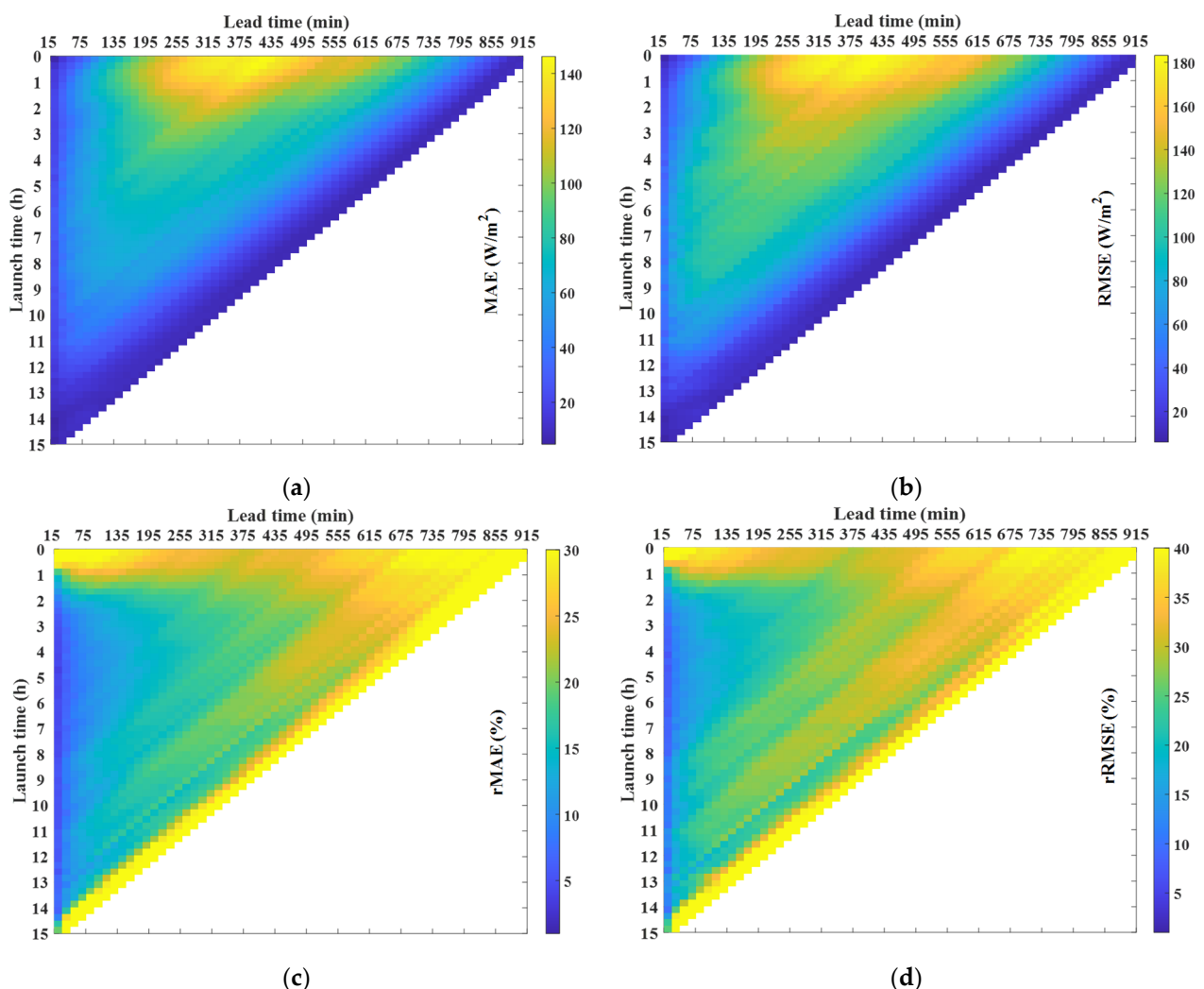
Metrics	Scaled (W/m <sup>2</sup> )	Percentage (%)
<b>Absolute</b>	$MAE = \frac{1}{T} \sum_{t=1}^T  Y_t - \hat{Y}_t $	$rMAE = \frac{\frac{1}{T} \sum_{t=1}^T  Y_t - \hat{Y}_t }{\frac{1}{T} \sum_{t=1}^T Y_t} \times 100$
<b>Quadratic</b>	$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (Y_t - \hat{Y}_t)^2}$	$rRMSE = \frac{\sqrt{\frac{1}{T} \sum_{t=1}^T (Y_t - \hat{Y}_t)^2}}{\frac{1}{T} \sum_{t=1}^T Y_t} \times 100$

The value of  $Y_t$  denotes GHI at a specific hour of the day,  $t$ , and  $\hat{Y}_{t',t}$  is the prediction of  $Y_t$  at  $t'$ . The initial time,  $t_0$ , is fixed for each day and corresponds to the sunrise. To assess the error, two parameters are defined: lead time and launch time. Lead time corresponds to  $(t' - t)$  and is the difference between the time instant of the prediction and the moment when the prediction is launched. Launch time, on the other hand, is denoted by  $(t' - t_0)$  and is the difference between the current time and sunrise. Launch and lead time for the predictions of a particular day are better explained in Figure 6. When the launch time is fixed and the lead time is used as a parameter, a vector of predictions is obtained. However, when both parameters are set to a value, a single point forecast is obtained (red diamond in Figure 6).



**Figure 6.** Real measurements of a selected day and its predictions. Dashed lines are the predictions for different launch times and dotted lines correspond to different lead times. Both parameters can be specified for a single day, obtaining the point forecast plotted with a red diamond.

The 3D plot in Figure 7 depicts the errors as a function of the lead time and the launch time which leads to the following conclusions. Firstly, for the scaled error, a high error rate is observed for short launch times under medium lead times. It is expected that the scaled error is large under the previous conditions since the radiation is high. However, as the launch time increases, this error significantly decreases. Secondly, it was clear that the lower the radiation, the smaller the scaled error; however, for percentage errors, the opposite is the case; when the launch time is small (less than 1 h), the percentage error is high, irrespective of the lead time. These plots give some insight into the prediction behavior and become particularly useful in enhancing confidence in the prediction with respect to other forecasting techniques. In this particular case, the intra-daily prediction is used when the mean error is smaller than the day-ahead prediction [37]. Finally, prediction intervals are derived from the MAE, assuming a particular distribution and splitting the predictions into groups as a function of the lead time, the launch time and the type of day, being very useful when a high degree of accuracy is required for the prediction.



**Figure 7.** Error matrices obtained from GHI real measurements and GHI forecasted values, as a function of the launch time and the lead time: (a) MAE; (b) RMSE; (c) rMAE; and (d) rRMSE.

Finally, the predictions obtained by the LSTM-RNN used in this work are compared with those available in the literature, which are depicted in Table 3. It is worth noting that this comparative analysis should not be strictly considered, since each dataset can have a relative influence on the performance. Nevertheless, some preliminary conclusions can be drawn from the study. Firstly, taking into account other widely used techniques

from [45], the forecast error obtained in this work, in terms of the  $rMAE$ , is much smaller under short lead times (15 min), increasing until a similar value of the error is obtained under large lead times (6 h). A good performance under small forecast horizons is also obtained when comparing the results with [46] for a statistical AutoRegressive Integrated Moving Average (ARIMA) model, in terms of the MAE, obtaining a similar error to that of traditional RNNs, and a higher error with respect to a similar LSTM-based approach presented in [46], despite considering other inputs highly correlated with the irradiance. Finally, comparing the strategy presented in this paper with respect to the deep learning techniques (GRU, LSTM-RNN, and CNN-LSTM) from [39,47–49], a similar performance can be observed. To conclude, for small lead times, the forecasting approach introduced in this paper yields better results than those obtained by traditional methods. However, the forecasting error of the proposed LSTM-RNN-based method increases for higher lead times, until a similar performance is obtained with respect to the traditional methods compared from the literature. It is also observed that an increase in the number of inputs seems to slightly improve the performance of the forecast approach. Adding exogenous inputs to the forecast process is an alternative which is often used by researchers but negatively affects the performance when those resources are not available.

**Table 3.** Comparison between the research results from this paper and those from other articles in the literature.

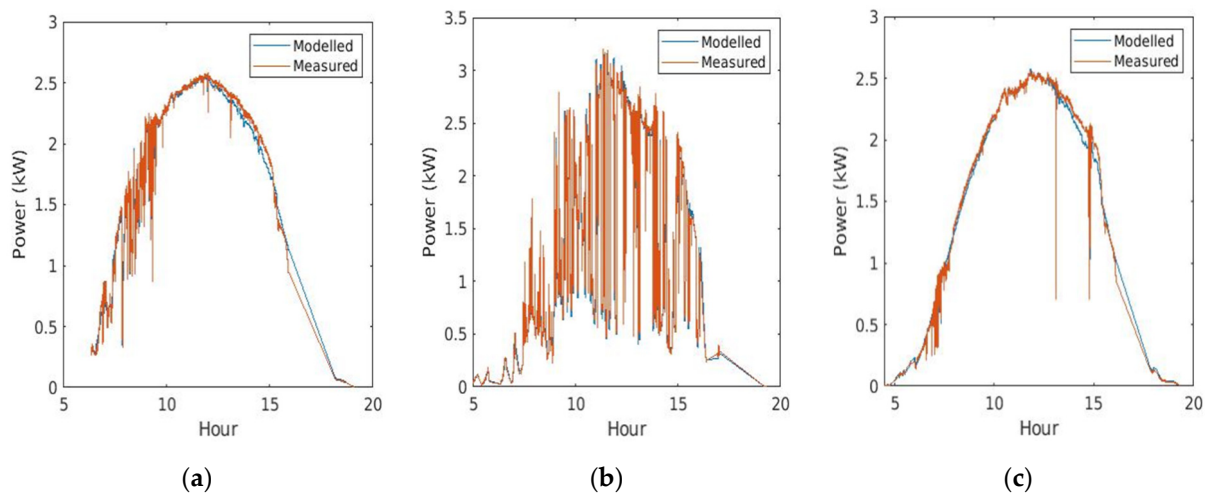
Model [Article]	Error	Forecast Horizon	Time Interval	Inputs	Results from This Paper
Smart pers. [45] CIAD Cast [45] Satellite [45] WRF-Solar [45] SVM-Radial [45]	$rMAE = (8\text{--}18)\%$ $rMAE = (11\text{--}20)\%$ $rMAE = (10.5\text{--}19.5)\%$ $rMAE = (12\text{--}18)\%$ $rMAE = (7.5\text{--}15.5)\%$	6 h	15 min	GHI, Clear Sky GHI, Cloud index maps, Cloud top height maps, ...	$rMAE = (4.17\text{--}17.73)\%$
ARIMA [46] RNN [46] LSTM [46]	$MAE = 71.48\text{ W/m}^2$ $MAE = 41.83\text{ W/m}^2$ $MAE = 31.86\text{ W/m}^2$	1 h	1 h	GHI, Clear Sky GHI, Cloud type, Temperature, Humidity, Precipitation, Wind, ...	$MAE = 41.88\text{ W/m}^2$
CNN-LSTM [39] CNN-LSTM [39] CNN-LSTM [39] CNN-LSTM [39]	$MAE = 41.88\text{ W/m}^2$ $RMSE = 78.17\text{ W/m}^2$ $rMAE = 10.58\%$ $rRMSE = 19.75\%$	1 h	1 h	GHI, Temperature, Wind, Precipitation, Humidity, Azimuth, ...	$MAE = 41.88\text{ W/m}^2$ $RMSE = 72.54\text{ W/m}^2$ $rMAE = 8.72\%$ $rRMSE = 15.1\%$
LSTM [47] LSTM [47]	$RMSE = (77\text{--}143)\text{ W/m}^2$ $rRMSE = (18.4\text{--}33)\%$	8 h	1 h	GHI, Humidity, Cloudiness, Temperature, Extra-terrestrial	$RMSE = (72\text{--}124)\text{ W/m}^2$ $rRMSE = (15.1\text{--}29.2)\%$
GRU [48] LSTM [48]	$RMSE = 67.29\text{ W/m}^2$ $RMSE = 66.57\text{ W/m}^2$	1 h	1 h	GHI, Zenith, Humidity, Temperature	$RMSE = 72.54\text{ W/m}^2$
GRU [49] LSTM [49]	$RMSE = 58\text{ W/m}^2$ $RMSE = 55.29\text{ W/m}^2$	30 min	1 min	GHI	$RMSE = 55.78\text{ W/m}^2$

### 3.2. PV Power Estimation from the Forecasted GHI

The following step consists of estimating the power delivered by the PV modules from the GHI forecasts. To this end, the following parameters are required: (i) the prediction time instant; (ii) the site location in terms of latitude, longitude, and altitude; (iii) the installation characteristics, which include the orientation and inclination of the panels, rated parameters of the PV models available in datasheets, and losses associated with each part of the installation; and (iv) the ambient temperature, obtained from NWP maps. As stated above, analytical techniques exist to achieve this goal and, as a result, it is possible to quantify the error committed in the procedure.

This section focuses on two different approaches. Firstly, real measurements of PV power are compared against the estimated values of PV power obtained from real measurements of GHI at the site. Secondly, the PV power is estimated from the forecasted values of GHI, evaluating the errors associated with the whole process. The GHI conversion searches for a reduced value of the error to maintain a similar performance to that obtained in the previous section, using the errors to construct the prediction intervals (Section 3.3).

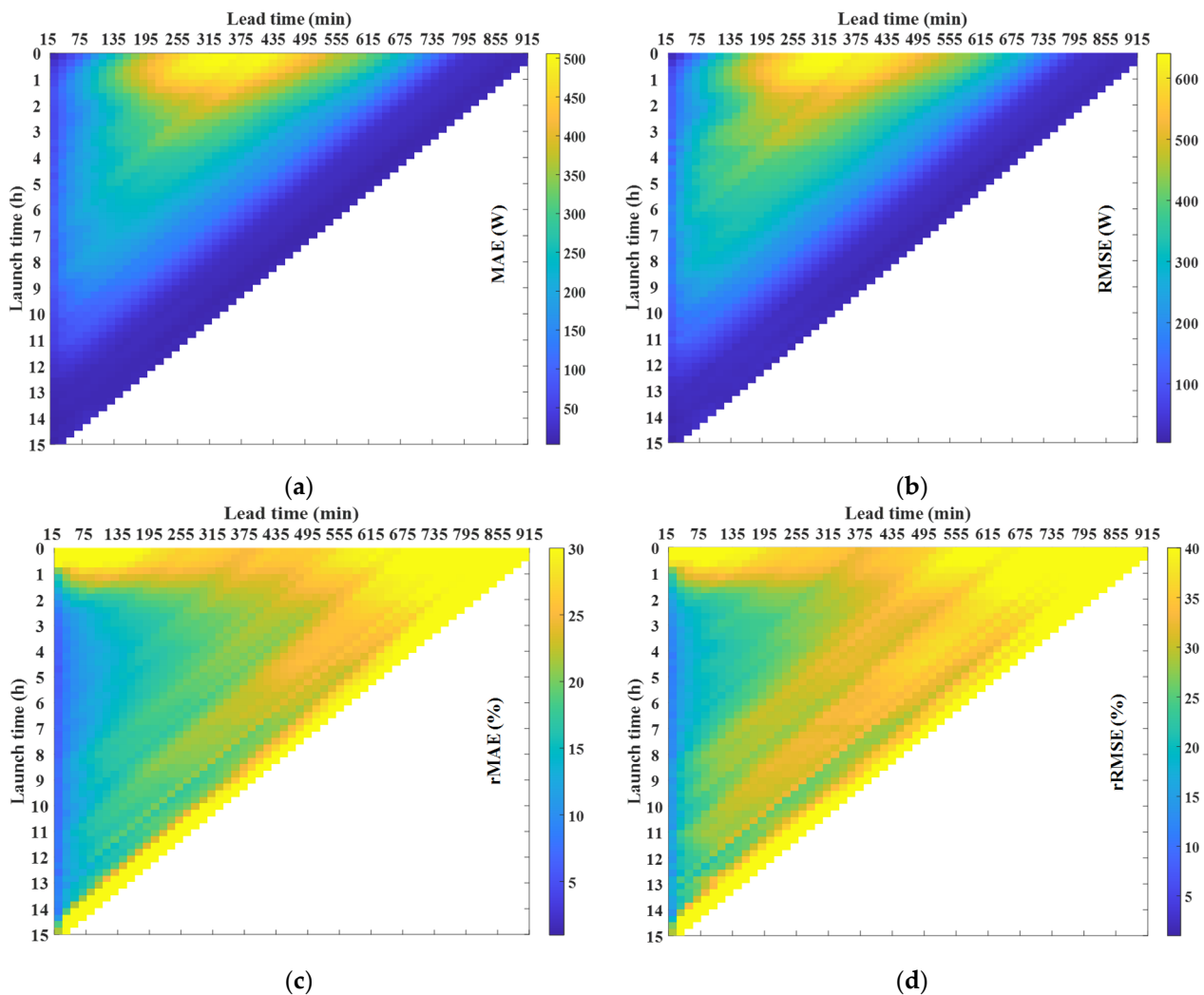
Figure 8 depicts the comparison between the measured values of PV power at the site with respect to the PV power estimation obtained from real GHI measurements taken at the site. Three types of days have been selected: a cloudy day, an overcast day and a sunny day. The  $x$  axis is expressed in solar time. It is worth noting that the experimental setup at the site location has a building near the PV panels that generates partial shadows on some of them, starting from 16:36 and continuing until sunset. This event is also modelled in Equation (2), assuming a linear variation of this effect with respect to time (in Figure 8  $SF = 0.95$  at 16:36, decreasing until  $SF = 0.4$  at sunset), and it also varies depending on the season of the year. Results show a reduced value for the error similar to that reported in other works [28], obtaining an  $rMAE = 2.54\%$  for sunny days, an  $rMAE = 3.04\%$  for partially cloudy days, and an increased value of  $rMAE = 4.03\%$  for overcast days. In terms of the squared error, values range from  $rRMSE = 3.44\%$  on sunny days and  $rRMSE = 3.90\%$  on partially cloudy days, to  $rRMSE = 5.95\%$  for overcast days. The transient characteristic of the inverter MPPT controller reveals that, in the presence of passing clouds, the inverter operating point becomes unstable. This is the reason why the error increases on these days. However, this does not pose any problem for the forecasting process since the time interval is 15 min, which considerably mitigates this negative effect.



**Figure 8.** Comparison between the measured values of PV power with respect to values obtained from the conversion of real GHI measurements at the site. The selected days are: (a) a partially cloudy day (17 May 2021:  $rMAE = 3.04\%$   $rRMSE = 3.90\%$ ), (b) an overcast day (1 June 2021:  $rMAE = 4.03\%$   $rRMSE = 5.95\%$ ) and (c) a sunny day (4 June 2021:  $rMAE = 2.54\%$   $rRMSE = 3.44\%$ ).

Finally, Figure 9 depicts the forecast error in terms of the difference between the measured and estimated PV power as a function of the lead time and the launch time. The shapes of the figures are similar to the previous section, with similar percentage errors. Therefore, from the figure, the same conclusions reached by analyzing Figure 7 can be drawn: (i) the scaled error is high for short launch times and medium lead times but decreases significantly as the launch time increases; (ii) for launch times of less than an hour the percentage error is high, irrespective of the lead time; and (iii) the percentage error is high at lead times higher than approximately 7 h. The forecast error, which is dependent on the lead time and the launch time, is used to generate the prediction intervals in the following section.



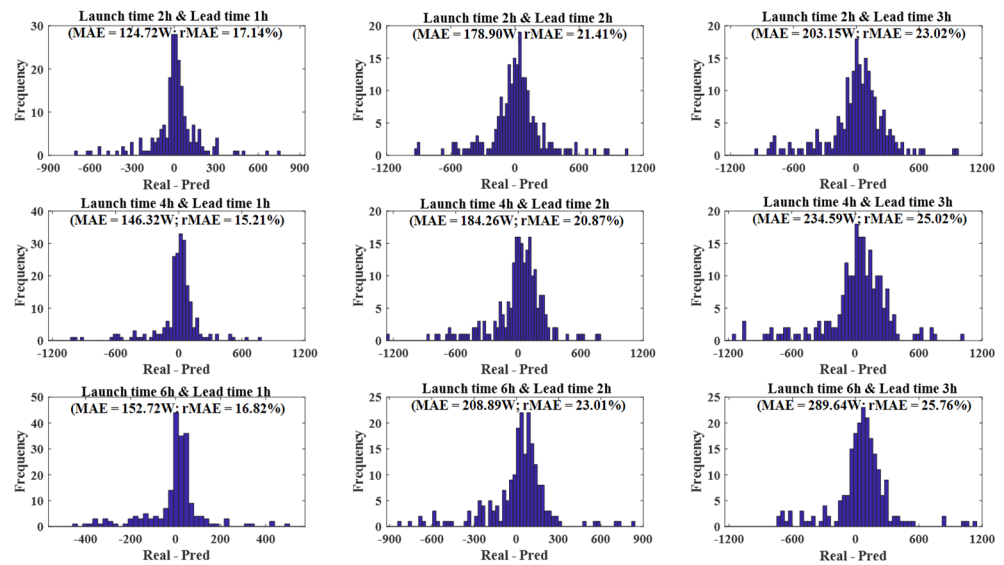


**Figure 9.** Error matrices obtained from PV power real measurements and PV power estimated values, as a function of the launch time and the lead time: (a) MAE; (b) RMSE; (c) rMAE; and (d) rRMSE.

### 3.3. Prediction Intervals of the Forecasted PV Power

Prediction intervals provide additional information about the plausible range of PV energy that will be generated at the site, for a defined confidence level selected by the user. Prediction intervals also indicate the degree of uncertainty in point forecasts. This could avoid unexpected energy shortages or, by contrast, energy surpluses, which are less critical than the former since the inverter can change its operational point to produce only the energy needed, despite wasting an exploitable energy resource.

In this paper, prediction intervals are obtained based on the work carried out in [44]. Previous results show how dependent the forecast accuracy is on the lead time and the launch time. This fact is used to split the dataset of predictions and create groups, assuming a specific distribution which is built based on the MAE. Therefore, each group is defined by selecting a launch time and a lead time, obtaining 365 samples per group, since a whole year is forecasted on this research. Figure 10 shows different error distributions for launch time values of 2, 4, and 6 h, and lead time values of 1, 2, and 3 h. In all of them, a Laplacian distribution is considered, similar to the work carried out in [37] but as a function of the CCF. Prediction intervals ( $E_{15m} \pm p_s$ ) for each subset can be defined in terms of the MAE under this assumption: for a Laplacian distribution, a percentile  $p$  of probability  $(1 - s)$  has an interval of  $p_s = \pm MAE \cdot \ln(2s)$ .



**Figure 10.** Error distribution for different subsets. A Laplacian distribution is assumed to create prediction intervals in terms of the MAE.

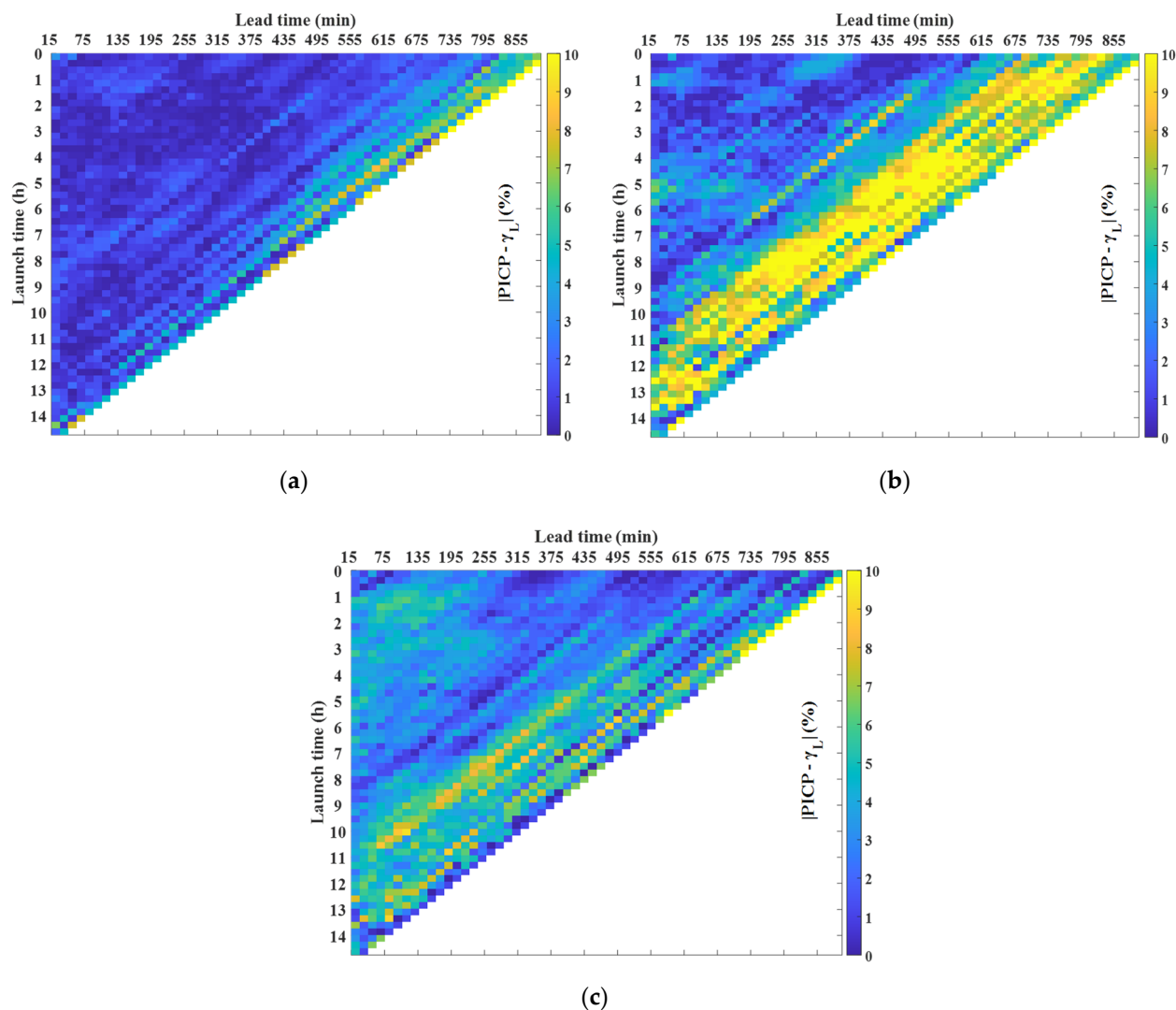
More detailed distributions can be determined provided that the selected groups are also created as a function of the CCF. However, by considering 10 groups as presented in [37], the number of samples of each group is not sufficient to create a proper error distribution. To overcome this drawback, the number of CCF groups is reduced to three, using the type of day classification criteria (e.g., sunny, cloudy, and overcast). The CCF parameter has an hourly resolution, its value is 0 when the sun is not covered by clouds and 1 when the sunlight is totally blocked. The type of day is classified evaluating the CCF during the daylight hours, with an hourly weighting of the amount of energy produced during the day. After that, the k-nearest neighbors (k-NN) method is used to form the groups, since it allows the dataset to be split in a simple way, offering an independent solution for each site in the VPP.

The assumption of a Laplacian distribution for each new selected subset carries an error that is necessary to quantify. The Prediction Interval Coverage Probability (*PICP*) [50], in Equation (4), indicates the percentage of predicted values that are inside the interval selected, and it must be close to the confidence level ( $\gamma_L$ ). The confidence level selected in this research is  $\gamma_L = 80\%$ , although this parameter can be modified depending on the operational risks that the site can handle: the higher the risks, the higher the benefits from the installation:

$$PICP = \frac{1}{T} \sum_{t=1}^T \epsilon_t, \text{ where } \epsilon_i = \begin{cases} 1 & \text{if } x_i \in [L_i, U_i] \\ 0 & \text{if } x_i \notin [L_i, U_i] \end{cases} . \quad (4)$$

Figure 11 depicts the absolute difference between the confidence level and the *PICP* for each type of day, being an effective method when this difference is close to zero. On sunny days, the *PICP* is close to the confidence level across the whole area, except for high lead times under small launch times where the difference increases. On cloudy days, the *PICP* is quite different from the confidence level during sunset. Nevertheless, the difference is acceptable in the rest of the area. In this case, the forecast has a lesser value during sunset since the energy produced is significantly reduced. Hence, prediction intervals also offer valuable information on cloudy days. Finally, for overcast days, the difference between the *PICP* and the confidence level increases with respect to sunny days, but the magnitude is acceptable and the prediction intervals are still valuable. To conclude, there are some zones with a high difference between the *PICP* and the confidence level. However, these scenarios correspond to small PV power measurements with bad forecasting performance

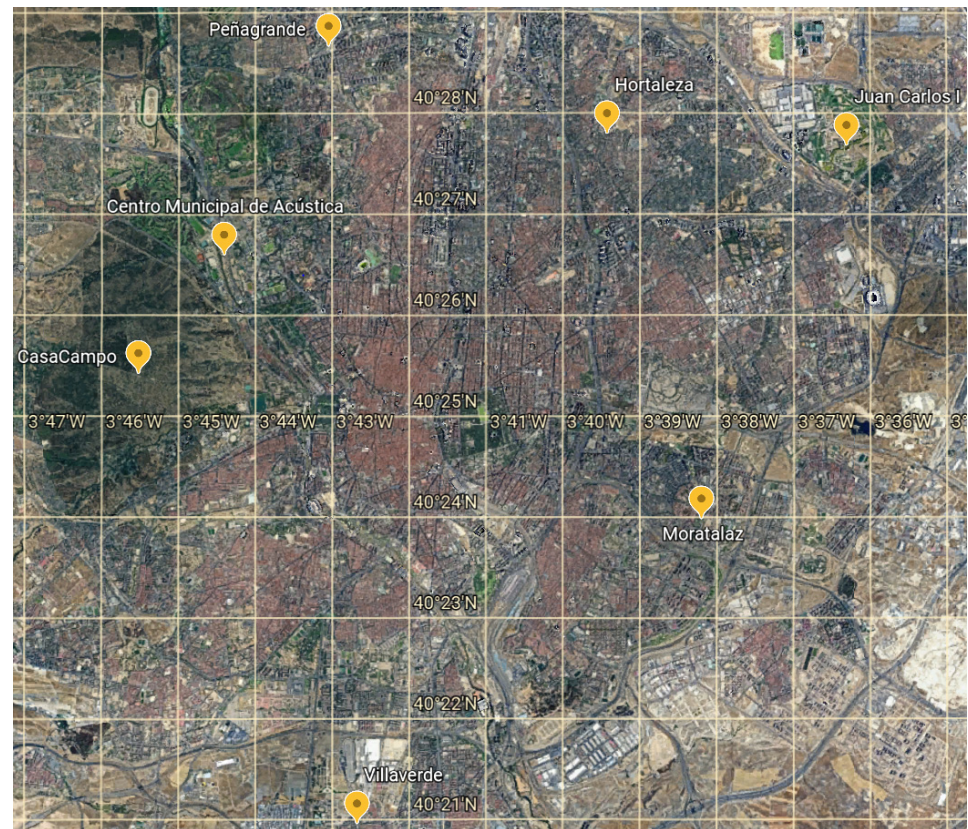
(Figure 9). Therefore, prediction intervals are of little value for these points, since the strategy presented in this paper does not focus on those cases.



**Figure 11.** Absolute difference between the PICP and the confidence level for every subset selected on the prediction intervals for different types of day: (a) sunny; (b) cloudy; and (c) overcast.

### 3.4. Evaluation of the GHI Forecasting for an Emulated VPP

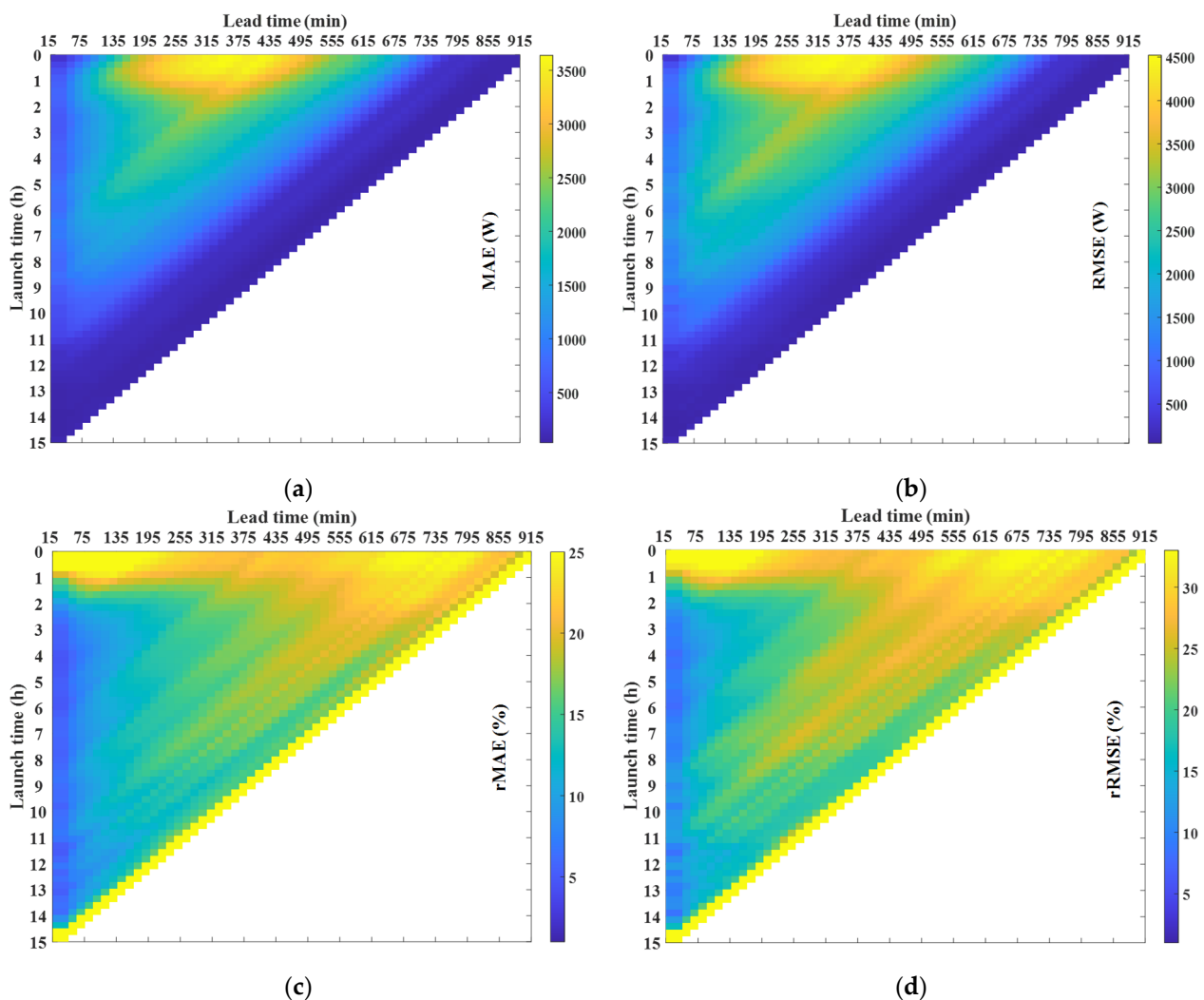
The effectiveness of the whole forecasting process has been demonstrated for a single PV installation, which plays the role of a VPP node, along with its limitations with respect to the launch time and the lead time. The next step consists of assessing the algorithm performance for a set of PV facilities, forming a VPP. There are, however, no additional PV installations available in the study. Therefore, seven ground-based meteorological stations located in the Community of Madrid, apart from the PV facility at the university, are used to emulate the VPP nodes. Their locations are depicted in Figure 12. These ground-based stations are equipped with GHI sensors which allow the GHI forecasts to be generated. As for the power conversion, the characteristics of the PV installation from the university are used to obtain the power estimation for each emulated VPP node (peak power,  $P_{peak} = 2.97$  kW, temperature coefficient,  $\delta P_m = -0.4\%/^{\circ}\text{C}$ , and the performance of the equipment).



**Figure 12.** Location of the ground-based stations in the Community of Madrid used in the research.

The same results as those shown in Figure 9 are used to quantify the accuracy of the prediction. However, in this case, the PV power forecast for each station is individually evaluated and the sum of power forecasts of the stations represents the PV power generated by the VPP, whose forecast error is depicted in Figure 13. By doing so, the PV power obtained at each station can be compared with respect to the total PV power forecasted. It can be observed that the scaled values of the error (MAE and RMSE) are higher than those in Figure 9. However, there is an 8-fold increase in the peak power with respect to a single facility. As a result, by looking at the relative values of the error (rMAE and rRMSE) it can be noted that the performance of the prediction increased for the VPP. The accuracy improvement of the PV power forecast can be expressed as the difference between the VPP forecast error and the sum of the error on each installation, dividing that value by the mean error committed on a single installation, obtaining a mean value of 12.37% with respect to the MAE, and 11.84% with respect to the RMSE. The shapes of the figures lead to identical conclusions to those reached by the analysis in Figure 9. Therefore, the prediction intervals maintain their potential value for error forecasting in the case of a VPP.





**Figure 13.** Error matrices of PV power forecasts from the VPP emulated in the research, represented as the sum of 8 different VPP nodes located in the Community of Madrid, as a function of the launch time and the lead time: (a) MAE; (b) RMSE; (c)  $rMAE$ ; and (d)  $rRMSE$ .

#### 4. Discussion and Conclusions

The technical development of VPPs must be supported by EMSs, for which PV power forecasting is an essential part. By knowing the energy produced by each VPP node, usually based on renewable resources such as solar technologies, it is possible to optimize the expected profit generated by energy exchanges with the grid operator. However, it is difficult to obtain PV power forecasts when it is necessary to gather information from several nodes scattered throughout a wide area, especially when the input data, required for the predictions, incur costs. This research presents a way of accomplishing this objective, using an LSTM-RNN-based strategy to, firstly, forecast the GHI by using a dataset of irradiance values derived from satellite data freely obtained from the CAMS, and secondly, estimate the solar power by utilizing a PV model of the installation. The forecast is updated during the day to achieve the highest accuracy, and prediction intervals are estimated as a function of the MAE. This provides a useful framework to understand the behavior of each installation that composes the VPP.

The first results provided are related to the GHI forecast for the installation and are based on the lead time and the launch time, which allow zones with a reduced error and a high level of confidence to be created in the shape of prediction intervals which depend on the type of day. The GHI error, as a function of the lead time and the launch time, shows



a low performance when the launch time is lower than 1.5 h, corresponding to sunrise. To avoid this, the forecasting process can begin at 1.5 h after sunrise; before this time, this research can rely on the day-ahead prediction made in [37] to obtain the irradiance forecast. To assess the accuracy of the intraday forecast, the results have been compared with those in the literature, achieving similar results to those obtained from deep learning algorithms and outperforming traditional techniques. The distinction between the lead time and the launch time means it possible to create better comparisons with respect to the literature, but also means it is difficult to summarize the research with only one value. The *MAE* committed, without considering the lead time and the launch time, is of  $44.19 \text{ W/m}^2$ , which is coherent with other studies.

Once the irradiance is forecasted, the conversion to PV power is analytically calculated, minimizing the error, which ranges from 2.54% to 4.03% in terms of the *rMAE* and from 3.44% to 5.95% in terms of the *rRMSE*. The error committed in this case is similar to the errors found in other articles [26,28]. The shapes of the error matrixes show similar results to those presented above. Therefore, similar conclusions can be drawn. The global *MAE* committed in this case is  $137.21 \text{ W}$  in a PV facility of  $2.97 \text{ kW}_p$ .

Prediction intervals are selected once the PV power forecast is available, which allow a range of plausible values of point forecasts to be obtained. The method considers a Laplacian distribution of the error and distinguishes between the lead time, the launch time and the type of day, which is selected with a k-NN algorithm as a function of the CCF. To verify whether the boundaries maintain the associated level of confidence, the *PICP* is calculated, obtaining values close to the selected confidence level of  $\gamma_L = 80\%$ . In this case, results reveal a noticeable difference between the *PICP* and the confidence level on cloudy days close to sunset. However, the predictions at those hours have minor importance. It can be concluded that the selected prediction intervals are of great relevance.

Finally, the PV power forecast is created, and the prediction intervals are selected for the PV facility so that conclusions under a VPP environment can be drawn. In this case, a real PV facility and seven ground-based weather stations in the Community of Madrid are selected to emulate the VPP, obtaining an improvement in the accuracy of 12.37% with respect to the *MAE*, and 11.84% with respect to the *RMSE*. Similar conclusions can be reached regarding the error as a function of the lead time and the launch time. Therefore, the whole strategy can be applied under different scenarios for launch times higher than 1.5 hours, relying on the day-ahead prediction prior to this. For this case, the error matrixes also indicate the best moments to obtain the predictions of the nodes, making it possible to increase the reliability of the VPP operation.

The major limitation of this study is related to the information of temperature and cloudiness freely obtained in Spain from NWP maps. In locations where this information is not available forecasts cannot be provided. Future works will focus on the application of this strategy along with a day-ahead time horizon strategy to schedule the operation of a VPP, creating a software that simplifies the process.

**Author Contributions:** G.M., C.S., P.M. and F.J.R. developed the forecasting framework of the research, generated the predictions of the installation, created the prediction intervals and emulated the VPP; R.P. and G.M. developed the model to obtain the PV power system from GHI; G.M., C.S., P.M., R.P. and B.V. carried out the experimental tests and wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Ministry of Science, Innovation and University of the Government of Spain under the HELIOS Sharing project (RTC-2017-6231-3) and Autonomous Community of Madrid under projects P2018/EMT-4366 and Y2020/EMT-6368.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The datasets generated for this study are available on request to the author Guillermo Moreno.

**Acknowledgments:** The authors thank the company CLYSEMA SA for their collaboration in this research.

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

## References

1. Goodstein, E.; Lovins, L.H. A pathway to rapid global solar energy deployment? Exploring the solar dominance hypothesis. *Energy Res. Soc. Sci.* **2019**, *56*, 101197. [[CrossRef](#)]
2. Nosratabadi, S.M.; Hooshmand, R.A.; Gholipour, E. A comprehensive review on microgrid and virtual power plant concepts. *Renew. Sustain. Energy Rev.* **2017**, *67*, 341–363. [[CrossRef](#)]
3. 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](#)]
4. Kudo, M.; Takeuchi, A.; Nozaki, Y.; Endo, H.; Sumita, J. Forecasting electric power generation in a photovoltaic power system for an energy network. *Electr. Eng. Jap.* **2009**, *167*, 16–23. [[CrossRef](#)]
5. Raza, M.Q.; Nadarajah, M.; Ekanayake, C. On recent advances in PV output power forecast. *Sol. Energy* **2016**, *136*, 125–144. [[CrossRef](#)]
6. Wan, C.; Zhao, J.; Song, Y.; Xu, Z.; Lin, J.; Hu, Z. Photovoltaic and solar power forecasting for smart grid energy management. *CSEE J. Power Energy Syst.* **2015**, *1*, 38–46. [[CrossRef](#)]
7. Massidda, L.; Marrocu, M. Smart meter forecasting from one minute to one year horizons. *Energies* **2018**, *11*, 3520. [[CrossRef](#)]
8. McCandless, T.C.; Haupt, S.E.; Young, G.S. A regime-dependent artificial neural network technique for short-range solar irradiance forecasting. *Renew. Energy* **2016**, *89*, 351–359. [[CrossRef](#)]
9. Diagne, M.; David, M.; Lauret, P.; Boland, J.; Schmutz, N. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. *Renew. Sustain. Energy Rev.* **2013**, *27*, 65–76. [[CrossRef](#)]
10. Makridakis, S.; Spiliotis, E.; Assimakopoulos, V. Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLoS ONE* **2018**, *13*, e0194889. [[CrossRef](#)] [[PubMed](#)]
11. Wang, F.; Yu, Y.; Zhang, Z.; Li, J.; Zhen, Z.; Li, K. Wavelet decomposition and convolutional LSTM networks based improved deep learning model for solar irradiance forecasting. *Appl. Sci.* **2018**, *8*, 1286. [[CrossRef](#)]
12. Li, G.; Wang, H.; Zhang, S.; Xin, J.; Liu, H. Recurrent neural networks based photovoltaic power forecasting approach. *Energies* **2019**, *12*, 2538. [[CrossRef](#)]
13. Cerqueira, V.; Torgo, L.; Soares, C. Machine learning vs statistical methods for time series forecasting: Size matters. *arXiv* **2019**, arXiv:1909.13316.
14. Voyant, C.; Nottou, G.; Kalogirou, S.; Nivet, M.L.; Paoli, C.; Motte, F.; Fouilloy, A. Machine learning methods for solar radiation forecasting: A review. *Renew. Energy* **2017**, *105*, 569–582. [[CrossRef](#)]
15. Kostylev, V.; Pavlovski, A. Solar power forecasting performance—towards industry standards. In Proceedings of the 1st International Workshop on the Integration of Solar Power into Power Systems, Aarhus, Denmark, 24–25 October 2011; pp. 1–8.
16. Rajagukguk, R.A.; Ramadhan, R.A.; Lee, H.J. A Review on Deep Learning Models for Forecasting Time Series Data of Solar Irradiance and Photovoltaic Power. *Energies* **2020**, *13*, 6623. [[CrossRef](#)]
17. Antonanzas, J.; Osorio, N.; Escobar, R.; Urraca, R.; Martinez-de-Pison, F.J.; Antonanzas-Torres, F. Review of photovoltaic power forecasting. *Sol. Energy* **2016**, *136*, 78–111. [[CrossRef](#)]
18. Chatfield, C. Prediction intervals for time-series forecasting. In *Principles of Forecasting*; Springer: Boston, MA, USA, 2001; Volume 30, pp. 475–494. [[CrossRef](#)]
19. Lorenz, E.; Hurka, J.; Karampela, G.; Heinemann, D.; Beyer, H.G.; Schneider, M. Qualified forecast of ensemble power production by spatially dispersed grid-connected PV systems. In Proceedings of the 23rd European PV Solar Energy Conference and Exhibition (EU PVSEC), Valencia, Spain, 1–5 September 2008; pp. 3285–3291.
20. Ghavidel, S.; Li, L.; Aghaei, J.; Yu, T.; Zhu, J. A review on the virtual power plant: Components and operation systems. In Proceedings of the 2016 IEEE International Conference on Power System Technology (POWERCON), Wollongong, NSW, Australia, 24 November 2016; pp. 1–6. [[CrossRef](#)]
21. Nakamura, M.; Takeno, K.; Hisamitsu, R.; Shoyama, M. Bi-directional Multiport Converter for Utilizing Green Base Stations as Virtual Power Plant. In Proceedings of the 8th International Conference on Renewable Energy Research and Applications, Brasov, Romania, 3–6 November 2019; pp. 137–141. [[CrossRef](#)]
22. Osterwald, C. Translation of device performance measurements to reference conditions. *Sol. Cells* **1986**, *18*, 269–279. [[CrossRef](#)]
23. Menicucci, D. Photovoltaic array performance simulation models. In Proceedings of the PV and Isolation Measurements Workshop, Vail, Colombia, 1 July 1985; pp. 383–392.
24. Marion, B. Comparison of predictive models for photovoltaic module performance. In Proceedings of the 33rd IEEE Photovoltaic Specialists Conference, San Diego, CA, USA, 11–16 May 2008; pp. 1–6. [[CrossRef](#)]
25. Marion, B.; Rummel, S.; Anderberg, A. Current–voltage curve translation by bilinear interpolation. *Prog. Photovolt.* **2004**, *12*, 593–607. [[CrossRef](#)]
26. Peña, R.; Diez-Pascual, A.; Diaz, P.; Davoise, L. A new method for current–voltage curve prediction in photovoltaic modules. *IET Renew. Power Gener.* **2021**, *15*, 1331–1343. [[CrossRef](#)]

27. Heydenreich, W.; Müller, B.; Reise, C. Describing the world with three parameters: A new approach to PV module power modelling. In Proceedings of the 23rd European PV Solar Energy Conference and Exhibition (EU PVSEC), Valencia, Spain, 1–5 September 2008; pp. 2786–2789.
28. Takilalte, A.; Harrouni, S.; Yaiche, M.; Mora-López, L. New approach to estimate 5-min global solar irradiation data on tilted planes from horizontal measurement. *Renew. Energy* **2020**, *145*, 2477–2488. [CrossRef]
29. Abdel-Nasser, M.; Mahmoud, K. Accurate photovoltaic power forecasting models using deep LSTM-RNN. *Neural Comput. & Applic.* **2019**, *31*, 2727–2740. [CrossRef]
30. Husein, M.; Chung, I.Y. Day-ahead solar irradiance forecasting for microgrids using a long short-term memory recurrent neural network: A deep learning approach. *Energies* **2019**, *12*, 1856. [CrossRef]
31. Spanish Agency of Meteorology (AEMET) Numerical Weather Predictions. Available online: [http://www.aemet.es/es/eltiempo/prediccion/modelosnumericos/harmonie\\_arome\\_ccaa?opc2=mad](http://www.aemet.es/es/eltiempo/prediccion/modelosnumericos/harmonie_arome_ccaa?opc2=mad) (accessed on 27 July 2021).
32. Schwingshackl, C.; Petitta, M.; Wagner, J.E.; Belluardo, G.; Moser, D.; Castelli, M.; Zebisch, M.; Tetzlaff, A. Wind effect on PV module temperature: Analysis of different techniques for an accurate estimation. *Energy Procedia* **2013**, *40*, 77–86. [CrossRef]
33. Tradacete, M.; Santos, C.; Jiménez, J.A.; Rodríguez, F.J.; Martín, P.; Santiso, E.; Gayo, M. Turning Base Transceiver Stations into Scalable and Controllable DC Microgrids Based on a Smart Sensing Strategy. *Sensors* **2021**, *21*, 1202. [CrossRef] [PubMed]
34. The MathWorks, I. ThingSpeak for IoT Projects. Available online: <https://thingspeak.com/> (accessed on 27 July 2021).
35. Commission, E. Copernicus Europe’s eyes on Earth. Available online: <https://www.copernicus.eu/en> (accessed on 27 July 2021).
36. Malý, J. The Darboux property for gradients. *Real Anal. Exch.* **1996**, *22*, 167–173. [CrossRef]
37. Moreno, G.; Martin, P.; Santos, C.; Rodríguez, F.J.; Santiso, E. A Day-Ahead Irradiance Forecasting Strategy for the Integration of Photovoltaic Systems in Virtual Power Plants. *IEEE Access* **2020**, *8*, 204226–204240. [CrossRef]
38. Bengio, Y.; Simard, P.; Frasconi, P. Learning long-term dependencies with gradient descent is difficult. *IEEE Trans. Neural Netw.* **1994**, *5*, 157–166. [CrossRef]
39. Zang, H.; Liu, L.; Sun, L.; Chen, L.; Wei, Z.; Sun, G. Short-term global horizontal irradiance forecasting based on a hybrid CNN-LSTM model with spatiotemporal correlations. *Renew. Energy* **2020**, *160*, 26–41. [CrossRef]
40. Berrizbeitia, S.E.; Jadraque Gago, E.; Muneer, T. Empirical Models for the Estimation of Solar Sky-Diffuse Radiation. A Review and Experimental Analysis. *Energies* **2020**, *13*, 701. [CrossRef]
41. Davies, J.; Hay, J. Calculations of the solar radiation incident on an inclined surface. In Proceedings of the First Canadian Solar Radiation Data Workshop, Toronto, ON, Canada, 17–19 April 1978; pp. 32–58.
42. Martin, N.; Ruiz, J. Calculation of the PV modules angular losses under field conditions by means of an analytical model. *Sol. Energy Mater. Sol. Cells* **2001**, *70*, 25–38. [CrossRef]
43. King, D.L.; Boyson, W.E.; Kratochvil, J.A. *Photovoltaic Array Performance Model*; Sandia National Laboratories: Albuquerque, NM, USA, 2003; pp. 1–43.
44. Fonseca Junior, J.G.D.S.; Oozeki, T.; Ohtake, H.; Takashima, T.; Kazuhiko, O. On the use of maximum likelihood and input data similarity to obtain prediction intervals for forecasts of photovoltaic power generation. *J. Electr. Eng. Technol.* **2015**, *10*, 1342–1348. [CrossRef]
45. Huertas-Tato, J.; Aler, R.; Galván, I.; Rodríguez-Benítez, F.; Arbizu-Barrena, C.; Pozo-Vázquez, D. A short-term solar radiation forecasting system for the Iberian Peninsula. Part 2: Model blending approaches based on machine learning. *Sol. Energy* **2020**, *195*, 685–696. [CrossRef]
46. Yu, Y.; Cao, J.; Zhu, J. An LSTM short-term solar irradiance forecasting under complicated weather conditions. *IEEE Access* **2019**, *7*, 145651–145666. [CrossRef]
47. Chen, X.; Huang, X.; Cai, Y.; Shen, H.; Lu, J. Intra-day Forecast of Ground Horizontal Irradiance Using Long Short-term Memory Network (LSTM). *J. Meteorol. Soc. Jpn.* **2020**, *5*, 945–957. [CrossRef]
48. Wojtkiewicz, J.; Hosseini, M.; Gottumukkala, R.; Chambers, T.L. Hour-ahead solar irradiance forecasting using multivariate gated recurrent units. *Energies* **2019**, *12*, 4055. [CrossRef]
49. Yan, K.; Shen, H.; Wang, L.; Zhou, H.; Xu, M.; Mo, Y. Short-term solar irradiance forecasting based on a hybrid deep learning methodology. *Information* **2020**, *11*, 32. [CrossRef]
50. Van der Meer, D.W.; Widén, J.; Munkhammar, J. Review on probabilistic forecasting of photovoltaic power production and electricity consumption. *Renew. Sust. Energ. Rev.* **2018**, *81*, 1484–1512. [CrossRef]