IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 25, 2020, accepted October 9, 2020, date of publication October 15, 2020, date of current version November 5, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3031439

PV Plant Power Nowcasting: A Real Case Comparative Study With an Open Access Dataset

SONIA LEVA, (Senior Member, IEEE), ALFREDO NESPOLI, (Graduate Student Member, IEEE), SILVIA PRETTO, MARCO MUSSETTA, (Senior Member, IEEE), AND EMANUELE GIOVANNI CARLO OGLIARI, (Member, IEEE)
Department of Energy, Politecnico di Milano, 20156 Milano, Italy
Corresponding author: Alfredo Nespoli (alfredo.nespoli@polimi.it)

ABSTRACT Energy systems around the world are undergoing substantial changes, with an increasing penetration of Renewable Energy Sources. For this reason, the availability of a pool of suitable forecasting models specific for the needed time horizon and task is becoming crucial in the grid operation. In addition, nowcasting techniques aiming at provideing the power forecast for the immediate future, are more often investigated due to the spread of micro-grids and the need of facing changing electrical market environments. In this paper a novel comprehensive methodology aiming at computing the PV power forecast on different time horizons and resolutions is introduced. Moving from the 24-hours ahead prediction provided by the Physical Hybrid Artificial Neural Network (PHANN), a technique to refine the power forecast for the following 3 hours with an hourly granularity is analyzed, leveraging on newer information available during the operations. Moreover, in order to provide the power forecast for the following 30 minutes on a minutely basis, an innovative modification of a statistical technique is proposed, the robust persistence. The proposed comprehensive approach allowed to greatly reduce the overall error committed when compared with the benchmark models. Finally, the proposed methodology is validated and tested on a freely available database consisting on different parameters recorded at both the meteorological and photovoltaic test facility at SolarTech^{LAB}, Politecnico di Milano, Milan.

INDEX TERMS Photovoltaic systems, renewable energy sources, power forecast, nowcasting, real case study.

I. INTRODUCTION

Renewable energy sources (RES), are mainly variable along time due to the intrinsic nature of the primary source and, as a consequence, they are considered Non-Programmable Renewable Sources (NPRS) [1]. The constantly increasing installed capacity from NPRS leads to urgent problems of energy and power management, to be addressed on different time horizons [2]. In addition, electric energy markets in several countries are carried out continuously, with the need to accurately predict the energy production from RES, not only in the day-ahead perspective, but also for the immediate future [3], [4].

One of the main issues related to the forecasting topic, is the choice of the more suitable (accurate) prediction model to be applied according both to the given time horizon and the task to be accomplished [5], [6]. In fact, when PV modules are part of the generation units of a complex energy system, an integrated forecasting approach is useful to address several scopes. From the Energy Management System (EMS) point of view, the 24 hours ahead forecast is needed to initially perform a strategic optimization of the system management, with the goal of, for example, minimizing the operating cost, reducing the fuel consumption over the following 24 hours and properly manage and size the battery energy storage system (BESS) [7], [8]. Approaching the operating time, additional refinements are useful to further correct the original forecast and tune the dispatching schedule in the so called

The associate editor coordinating the review of this manuscript and approving it for publication was Dongbo Zhao.

intraday or very-short term forecast [9]. Eventually, during the operations, an integration of the previous forecast for the following minutes is required for the predictive power management [10], [11].

In order to address the 24 hours ahead forecast, several methodologies are adopted in literature. These approaches are usually divided into 3 main categories: physical, statistical and hybrid models [12]. Physical methodologies rely on the mathematical modeling of the power plant, hence describing how it is able to convert the primary source into electrical power. Statistical models, on the other hand, are able to retrieve information from a statistical analysis performed on the different input variables. Therefore, the past time-series data are fundamental. Among the statistical approaches, regression models and Machine Learning (ML) techniques are the most diffused. The former move from the ARMA model, which is a combination of autoregressive (AR) and moving average (MA) models [13]. In the wide group of ML, worth mentioning are the Artificial Neural Networks (ANN) [14]. Among all, ANN is more suitable compared with classical statistical methods when non liner and complicated correlation exists between the data and when no prior assumption is formulated [15]. In particular, ANN proved to be a highly reliable method due to its ability of modeling RES power generation in presence of non-linearity, as for example between meteorological data and PV power production [16]. Finally, hybrid models are any combination of the previous two and are able to leverage on the strengths of each method they are composed by, while mitigating their weaknesses [17].

In order to increase the reliability of the PV power forecast, some corrections to the 24-hours ahead predictions are proposed in literature. In [18] different intraday methodologies are studied and compared. In particular, SARIMA and SARIMAX are considered to the scope, concluding that exogenous input can be omitted without influencing the forecast accuracy. In [19], three ML methods, namely Lasso, SVR and MLP, are compared, concluding that SVR and MLP outperform the Lasso model. From the EMS perspective, in [20], a correction of the unit commitment is introduced on the basis of the refinement of the PV generation forecast and its importance is stated. Nevertheless, no clear information on the methodology are given.

Further reducing the forecast horizon, the main issue is related to the identification of the sudden spikes and drops in the PV power production. These fluctuations are mainly related to the presence of moving clouds which hinder the solar radiation. For this reason, most of the work found in literature, relies on the adoption of satellite images and whole sky images such as [21], [22]. Despite providing an accurate result in terms of forecast accuracy, those systems require high investment cost for the equipment and a higher computational burden due to the image processing techniques. Other works on the other hand, rely on ML and Deep Learning algorithm to serve the scope [23]. All the above mentioned approaches are usually compared with a benchmark which is represented by the naive persistence. Persistence forecasting method consists in imposing the next value of the forecast parameter, equal to the last measured one [24]. In the very short term, this method achieves very good results, especially in stationary conditions and presents the advantage of a negligible computational burden. Nevertheless, the validity of persistence loses effectiveness as the forecast horizon increases [25].

In [26], a comparison among different time horizons is provided, while in the present work, three time horizons are managed and inspected in an integrated procedure for a continuous refinement of the prediction. In fact, a comprehensive methodology aiming at continuously improving the PV power forecast provided by the 24-hours ahead logic is here presented. Moving from this, with a rolling horizon approach, the power prediction for the following 3 hours with a hourly resolution is given in the intraday refinement. Both these predictions are computed through the adoption of ML techniques. Moreover, the forecast for the following 30 minutes (nowcasting) is carried out every ten minutes with a minutely granularity and through an innovative adjustment of the persistence approach. The intraday refinement and nowcasting methodologies setup is therefore unique, as well as the forecasting process.

Additionally, the lack of a comprehensive and open-source datasets, can lead to the difficulty, and sometimes inability, of selecting the proper methodology for the case under study, since the site specific conditions (data) used to validate the procedure are unknown and this can hinder a fair comparison and choice. For this reason, it is decided to provide a freely online distributed dataset of the PV related measurements acquired in the SolarTech^{LAB} facility installed at Politecnico di Milano, Milan, Italy [27]. The provided dataset also includes meteorological measurements to allow the integration and refinement of the adopted forecasting methods. The here proposed analysis leverages on this dataset, with real-time data from the experimental setup.

The paper is structured as follows: in section II, authors aim at presenting the novel comprehensive methodology, describing the more suitable forecasting models according to the specific time horizon and resolution considered. Moreover, some improvements in the existing forecasting models are provided. In section III, the performance metrics adopted to evaluate the proposed forecasting methods are presented. Section IV, describes in details the SolarTech^{LAB} test facility located at Politecnico di Milano, Milan and the characteristics of the freely available provided data. Furthermore, the cleaning process of the dataset is detailed in section V. Finally, in section VI the validation and comparison of the described forecasting methods is discussed.

II. PV POWER FORECAST: A COMPREHENSIVE METHODOLOGY

In this section, a novel comprehensive methodology aiming at assessing the PV power forecast on different time horizons and resolutions is proposed. The contribution of the presented

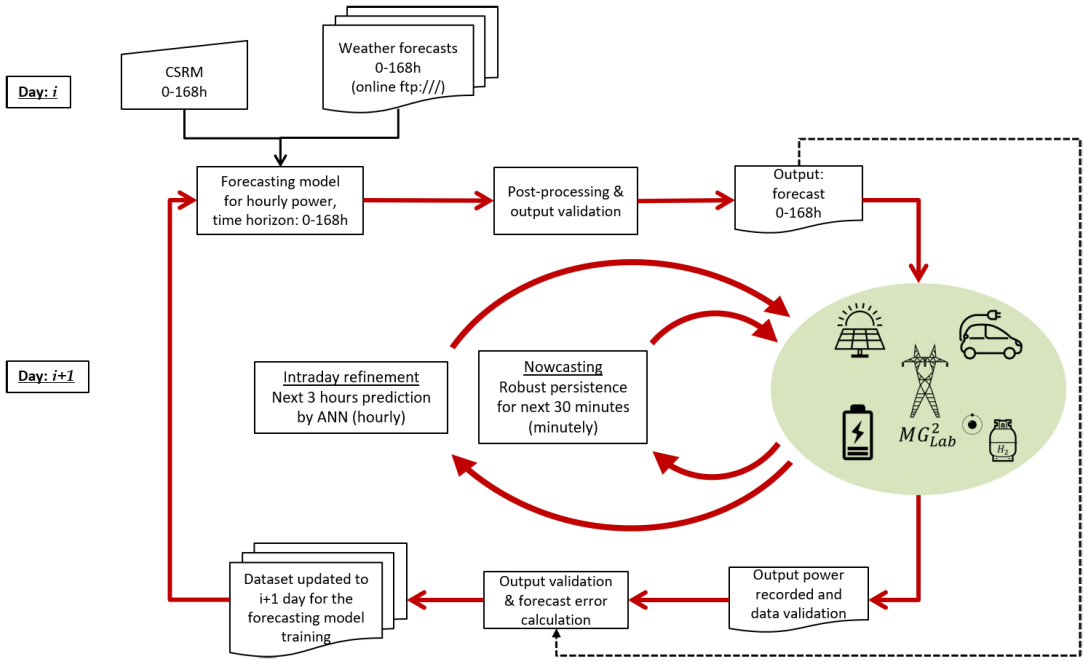


FIGURE 1. Novel comprehensive forecasting scheme: 24-hours ahead forecast, intraday refinement and nowcasting update.

procedure is the integration of three subsequent steps with the scope of refining the initial forecast. In fact, the methodology continuously improves the prediction accuracy achieved through the 24-hours ahead logic with newer information acquired in the operations. The algorithms composing the system, are specifically developed and tailored for the current scope and represent a novelty on their own.

When PV modules are integrated in broader systems such as microgrids, approaching the considered time, a finer granularity is needed to further tune the dispatchable units control strategy. Therefore, as function of the proximity to the objective time, three forecast horizons with different time granularity are here proposed. The presented approach can be also employed in different scenarios, such as for developing a bidding strategy, microgrid management, dispatch and control and for prognostic activity [28].

In Fig. 1, the proposed approach is presented and detailed. The outer loop, as described in [29], provides the 168 hours ahead forecast on a hourly basis, leveraging on the weather predictions for the subsequent hours. Here, the 24-hours forecasts are considered for the sake of comparison with other methods. The two additional loops, on the other hand, allow refining the aforementioned predictions when, during the operation, new data are acquired. In particular, the intermediate cycle addresses the issue of correcting the next three hours outcome with a hourly granularity, representing the intraday refinement. Moreover, the inner loop serves the scope of providing the next 30 minutes predictions on a minute basis through the robust persistence procedure. In Table 1 a summary of the whole approach is reported. In the following subsections, the methods to carry out each step of the novel process are presented. The proposed methodologies are

TABLE 1. Forecast horizon and resolution considered.

Forecast	Horizon	Resolution	Forecasting technique
24-h ahead	24 h	1 h	PHANN
Intraday refinement	3 h	1 h	ANN
Nowcasting	30 min	1 min	Robust persistence

implemented in the PV plant facility located at SolarTech^{LAB} at Politecnoco di Milano. The overall procedure is compared and technically validated on real case studies and real-time data from the experimental setup. Finally, the dataset, consisting on different parameters recorded, is freely available online.

A. 24-HOURS AHEAD POWER FORECAST

The first forecasting step, for example dealing with the energy management system in microgrids, is the 24-hours ahead power forecast with a hourly resolution [2]. This step allows some initial optimization procedures with the ultimate aim of attaining the highest exploitation of renewable energy resources, planning the usage of the energy storage in view of possible instantaneous fluctuations.

More in general, in many different contexts concerning the 24-hours ahead PV power forecast, several studies have been recently carried on [1]. It is common to find in literature that the best and most promising results were often obtained through the deployment of Artificial Intelligence (AI) techniques which, relying on available historical data, are able to generalize trends [13]. In addition, Machine Learning techniques infer the inherent relationships existing between the input parameters and the output variable, acquiring the characteristics of the system. For example, in the PV power forecast, ANN are able to infer, from historical datasets

TABLE 2. Training parameters for the intra-day forecast refinement.

Parameter	UoM
DoY_{t+h}	-
$Hour_{t+h}$	-
$GHI_{f,t+h,24}$	(W/m^2)
$GPOA_{f,t+h,24}$	(W/m^2)
$P_{f,t+h,24}$	(W)
$E_{f,t}$	(|W|)
$P_{m,t}$	(W)
$T_{m,t}$	(°C)
$GPOA_{m,t}$	(W/m^2)

provided during the training step, all the peculiarities of the PV plant such as near shadings, aging and dust on the PV modules which are possibly affecting the PV plant power output along with time.

Moreover, as it has been previously said, hybrid methods have usually well-established that they are capable to join the forces of the native methods by mitigating weaknesses of the singles. In the research field of the 24-hours ahead PV power forecast, it is proved that hybrid models more often score better performances rather than the others [9], [30].

Therefore the Physic Hybrid Artificial Neural Network (PHANN) forecasting method is adopted. It combines together both the deterministic Clear Sky Solar Radiation Algorithm (CSRM) and the stochastic Artificial Neural Network (ANN) method in order to enhance the 24-hours PV power forecast. In fact PHANN is here employed, to partly overcome the inaccuracies of the weather forecasts. In previous works PHANN has been compared to physical [31] and purely stochastic models [32], successfully overcoming both of the previous two single methods.

Besides, the reliability of this hybrid method has been tested on the forecasts of different PV plants. The role of several training sets varying in the amount of data and number of trials, which should be included in the "ensemble forecast" was assessed [33].

B. INTRADAY REFINEMENT

Once the 24-hours ahead power forecast are available, further refinement is required to update the dispatching strategy of the available units. During the plant operation, newer information are continuously collected, that could be valuable to the scope. The updated data could be either in accordance with those expected or not, due, for example, to unpredicted changes in the environmental conditions.

In the following part of this section, a new method aiming at performing the refinement of the power forecast provided with the 24-hours ahead logic is presented. In particular, given the current time t, the refinement $P_{f,t+h}$ is provided for the following 3 hours ($h \in [1, 2, 3]$). The algorithm relies on Feed Forward Neural Network (FFNN), and the required inputs needed to produce $P_{f,t+h}$ are given in Table 2.

Where:

- DoY_{t+h}, $Hour_{t+h}$ are the day of the year and hour of the corresponding sample to be forecast;

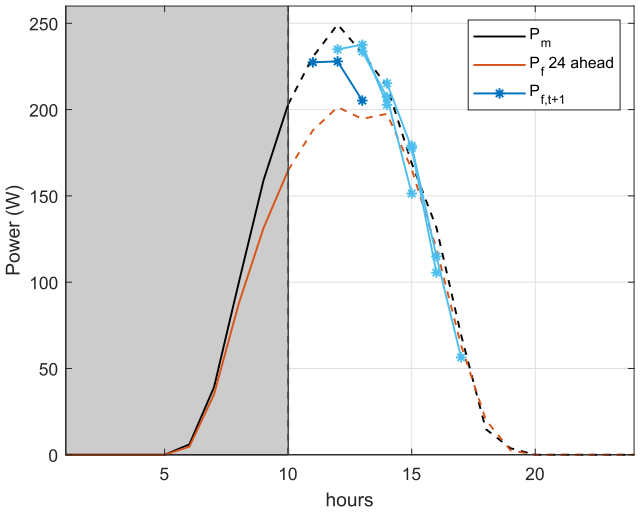


FIGURE 2. Intraday refinement, actual and forecast power. Dashed lines refer to future values, while solid lines are past values.

- $GHI_{f,t+h,24}$ and $GPOA_{f,t+h,24}$ are the predicted irradiation level (both on the horizontal and on the plane of the array) collected the previous day and already used to compute the forecast power with the 24-hours ahead logic;
- $P_{f,t+h,24}$ is the forecast power obtained with the 24-hours ahead logic;
- $E_{f,t}$ is the last forecast error available;
- $P_{m,t}$, $T_{m,t}$ and $GPOA_{m,t}$ are the current measurements for the weather parameters.

In Fig. 2, an exemplifying picture is presented to further clarify the methodology. In this example the shaded grey area represents the consolidated period of time. In black, the power measurements are given, where the dashed line represents future observations, not available at the time. Similarly, the 24-h ahead forecast is displayed in orange. The blue starred line is representative of the following 3 hours forecast refinement, while in light blue, by way of example, the new intraday power forecast refinement computed in the following hours is reported.

C. NOWCASTING

As shown in literature [25], when greatly reducing the forecast horizon to several minutes ahead, the most effective techniques are image-based or statistical. The first one, though providing more promising results, is highly expensive from a computational point of view and requires a higher initial investment with respect to the latter due to the additional cost of the instrumentation. A trade-off is thus needed among the computational burden and the accuracy. It is worth highlighting that the final purpose is to provide a useful tool to assess the PV power nowcast to be later included in a power and/or energy management system.

To the scope of assessing the PV production in the next 30 minutes ($s \in [1, 2, \ldots, 30]$), in this work, two methodologies leveraging on statistical models are presented, the first

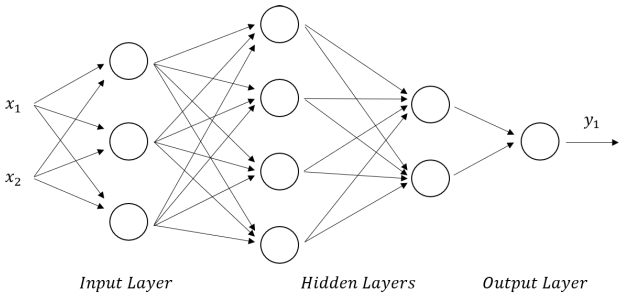


FIGURE 3. FFNN general structure, with two hidden layers.

one based on persistence models while the other relying on AI techniques. Both of the implemented methodologies provide a forecast resolution of 1 minute as summarized in Table 1.

1) ROBUST PERSISTENCE

The simplest way of forecasting future power production is represented by the naive persistence [15], generally employed as a benchmark for accuracy evaluation. It assumes the future power production at time t to be the same of the last registered value, at time $t - s$, where s is the time horizon, as detailed in (1).

$$P_t = P_{t-s} \tag{1}$$

Some improvement can though be obtained. In fact, sun irradiation is mainly dependent on two factors, the first one purely deterministic given by the sun position and the second mainly stochastic due to atmospheric conditions. While the latter cannot be inferred through a statistical analysis on such a short horizon, the first one can be instead addressed. In order to correct the irradiation level due to the sun position, the following formulation is proposed:

$$P_t = P_{t-s} \cdot \frac{\alpha_t}{\alpha_{t-s}} \tag{2}$$

In (2), the persistence forecast is strengthen by the introduction of a correction based on the solar altitude angle α. This method is referred as Robust persistence.

2) ANN BASED

The second methodology relies on the adoption of an AI technique, namely FFNN. The same inputs described in Table 2 are provided to the neural network. The architecture adopted in this work has the following settings and its general structure is depicted in Fig. 3:

1) 2 hidden layers;
2) 10 and 6 neurons;

These settings, found according to the methodology explained in [34], leverage on a sensitivity analysis that allowed to identified the optimal ANN structure in terms of number of layers and number of neurons in each layer in order to minimize the forecasting error.

III. FORECASTING PERFORMANCE METRICS

In order to asses the performance of the described methodologies, suitable indexes are needed to evaluate forecating errors depending on the forecast horizon. These indicators are usually based on the difference between the measured power ($P_{m,t}$) and the forecast one ($P_{f,t}$), referred to the t-th time sample.

Regarding the nowcast techniques, it is common to refer the accuracy to the well-known *Mean Absolute Error* (MAE). As far as the day ahead and the intra-day refinement are concerned, additional indicators are usually employed to evaluate daily errors: *Normalized MAE* (NMAE), which normalizes MAE by the nominal capacity P_n of the plant, *Weighted MAE* (WMAE), which is based on the measured production of energy in the considered time horizon, and *Normalized Root Mean Square Error* (nRMSE), whose definition is based on the maximum measured power [35].

These metrics are largely used to evaluate the accuracy of predictions and trend estimations. However, often outliers occurs when small power values are considered: in such cases, WMAE and nRMSE could result very large and biased, showing values above 100%, for instance when the forecast significantly overestimates the real energy production, which therefore has a very low value.

Based on these consideration, recently developed additional indicators are here adopted, aimed to provide a reliable evaluation of the forecasting accuracy in the range between 0 and 100% [33].

The first indicator is the *Envelope-Weighted MAE* (EMAE), defined as:

$$\mathrm{EMAE} = \frac{\sum_{t=1}^{N} |P_{m,t} - P_{f,t}|}{\sum_{t=1}^{N} \max(P_{m,t}, P_{f,t})} \cdot 100 \tag{3}$$

where the numerator is the same as WMAE, while the denominator is the sum of the maximum between the forecast and the measured power, thus avoiding values above 100%. This indicator can be suitably used for any time series forecasting evaluation and is not dependent on the specific PV application.

The second indicator, instead, was specifically defined as a new indicator focused on PV production: in fact, it is based on the definition of the Performance Ratio coefficient ($PR_{\%}$), as it is expressed in the IEC 61724 norm [36]. The *Objective Mean Absolute Error* (OMAE) is therefore defined as:

$$\mathrm{OMAE} = \frac{\sum_{t=1}^{N} \left|P_{m,t} - P_{f,t}\right|}{\sum_{t=1}^{N} G^{\mathrm{cs}}_{\mathrm{POA},t}} \cdot \frac{G_{\mathrm{STC}}}{P_n} \cdot 100 \tag{4}$$

where $G^{\mathrm{cs}}_{\mathrm{POA},t}$ is the theoretical solar irradiance on the plane of the array, computed by the clear sky solar irradiance model (CSRM), as described in [37], and G_{STC} is the solar irradiance at standard test conditions (*i.e.* 1000 W/m^2).

From (4) it is possible to find a relationship of OMAE with the NMAE indicator:

$$\mathrm{OMAE} = \mathrm{NMAE} \cdot \frac{G_{\mathrm{STC}}}{\sum_{t=1}^{N} G^{\mathrm{cs}}_{\mathrm{POA},t}} \cdot N \tag{5}$$

FIGURE 4. SolarTech^{LAB} test facility at Politecnico di Milano, Milano.

Indeed, this relation highlights how the newly proposed indicator is able to improve the NMAE metric by considering the performance of the forecast with respect to a modulated value, *i.e.* the theoretical irradiation, more descriptive than the mere nominal power P_n, which is attainable only in the hours of maximum irradiation.

Once all these error metrics are computed for each day, in order to analyze the overall indicator for the period under study, their respective average values are then evaluated.

IV. EXPERIMENTAL CAMPAIGN

The proposed forecasting techniques have been implemented and validated on a real case study dataset available at [38]. The test facility is located at Politecnico di Milano, Milan, on the rooftop of the Department of Energy, at the SolarTech^{LAB} (STL) [27], at latitude N 45° 30' 10.3'' and longitude E 9° 9' 23.66''. Fig. 4 shows a panoramic image of the test facility.

A. SOLARTECH^{LAB} TEST FACILITY: PV MODULES

At the SolarTech^{LAB}, several technologies are simultaneously installed and tested, like traditional monocrystalline modules, multi-crystalline, thin film, concentrated PV and concentrated PV with thermal recovery. In order to test the modules under different meteorological conditions, each of them is singularly controlled by a dedicated micro inverter, which carries out the connection to the grid and optimizes the DC output through an MPPT DC-DC controller. This configuration allows to finely control each module and ultimately increase the overall production. All PV modules are oriented with an azimuth γ equal to $-6°30'$, assuming 0° is the south positive west, and a tilt θ of 30°. Anyway, it is possible to modify the tilt of the modules and the distance among arrays.

In order to perform the following analysis, a single monocrystalline module of nominal power of 245 W_p is considered. The DC power recordings of the year 2017 are adopted. The measures are on a minutely basis but, due to various occurrences such as faults and disconnections, not all the days are available.

TABLE 3. PV module and inverter datasheets.

PV module	
P_{MPP}	245 W
Classification Range	-0/+4.99 W
Measurement Accuracy (P_{MPP})	-3/+3 %
U_{MPP}	31.30 V
I_{MPP}	7.84 A
U_{OC}	37.10 V
I_{SC}	8.48 A
U_{max}	1000 V
Inverter	
V_{DC} max	65 V
$V_{DC,MPP}$	12- 60 V
V_{DC}, Full Power	30-50V
$I_{DC,max}$	10.5 A
$I_{SC,max}$	12.5 A
$V_{AC,nom}$	230 V
f_{nom}	50 Hz
$P_{AC,nom}$ $\cos\phi = 1$	250 W @ 65 $°C$ amb.
$I_{AC,max}$	1.3 A

The PV module and inverter datasheets are reported in Table 3. In the online available dataset, the DC PV power measurements are given with a minute resolution.

B. SOLARTECH^{LAB} TEST FACILITY: METEOROLOGICAL DATA

Installed next to the photovoltaic field, the meteorological station allows to collect environmental parameters of interest. It is equipped with two secondary standard pyranometers to detect the total solar irradiance on horizontal and tilted plane (30°). The latter measurement is directly implemented to define the solar irradiation on the modules, since they have the same tilt angle. Additionally, a thermohygrometer for temperature and humidity measurements is installed. The main characteristics of the above mentioned sensors are reported in Table 4. Remaining sensors are a combined speed-direction anemometer for wind detection, a rain collector and two additional pyranometers, one for diffuse radiation and the other one for global radiation measurement, as detailed in Fig. 5. The meteorological station performs ambient conditions measurements every ten seconds. The average, maximum, minimum and standard deviation of the values measured by the sensors are calculated with a minute frequency and these values are stored into the database. The minutely averages acquired by the station are available online.

V. DATASET RECORDING AND PROCESSING

The test facility installed in Politecnico di Milano, allows to record and store all the variables described in sections IV-A and IV-B. More into detail, it is possible to freely download a comprehensive dataset at [38], that includes both the production of a single PV module and the meteorological variables. In VII, the description of the dataset is provided. Worth highlighting is that the provided data must undergo a cleaning and validation procedure since some measurements might be affected by some common

TABLE 4. Solar irradiance and temperature sensor characteristics.

Irradiance sensors	Pyranometer (LSI, DPA252)
Standard	Secondary standard ISO 9060
Measurements range (W/m^2)	<2000
Spectral range (nm)	300-3000
Total achievable daily uncertainty	<2%
Directional response (W/m^2)	$< \pm 5.4$
Thermal drift	<2%
Temperature and humidity sensor	LSI, DMA 875
Temperature sensor	Pt100 1/3 B (DIN EN 60751)
Measurements range	[-30°C, +70°C]
Uncertainty	0.2°C (at 0°C)
Resolution	0.04°C
Response time (T90)	3 min

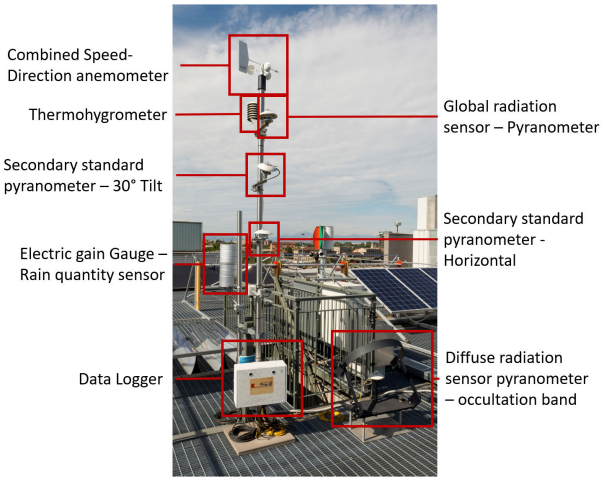


FIGURE 5. SolarTech^{LAB} weather station [27].

problems as single missing values. These operations should be performed every time real field measurements are adopted.

A cleaned version of the dataset has been implemented to carry out the here reported analysis. The first procedure that the data must undergo is to reconstruct the portion of the daily curve that is not recorded, the night. In order to do so, the PV measurements are cross correlated with the solar altitude [39], [40] of the location. This procedure allows to properly detect the night period and erratic values here registered due to instantaneous logging problems.

Furthermore, during some days there are some clear evidence of misleading power profile, when the power is really close to zero, despite appreciable level of radiation. Those days must be discarded since they are not interesting to our scope.

Moreover, some minor error in the communication between the PV inverter and the central monitoring station might lead to the loss of some isolated samples.

In this work, once the cleaning process is performed, a single day of measurements is considered reliable and hence used, only if it has at least 90% of the theoretically expected numerosity. Once all the reliable days are identified, in order to deal with missing values ($\leq$ 10%), linear interpolation is performed.

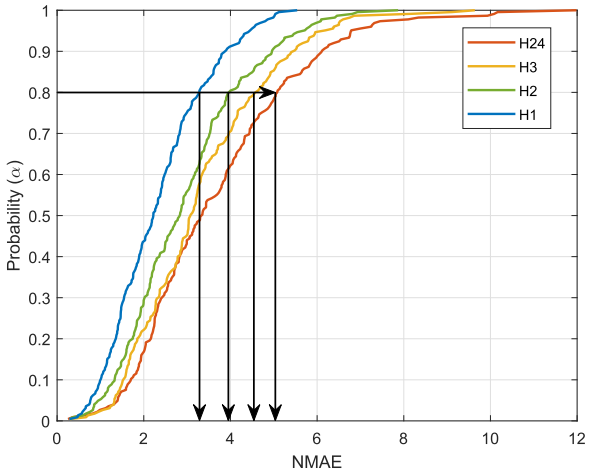


FIGURE 6. Intraday refinement: NMAE cumulative distribution functions (CDF) for the 24-hours ahead forecast and the following refinements.

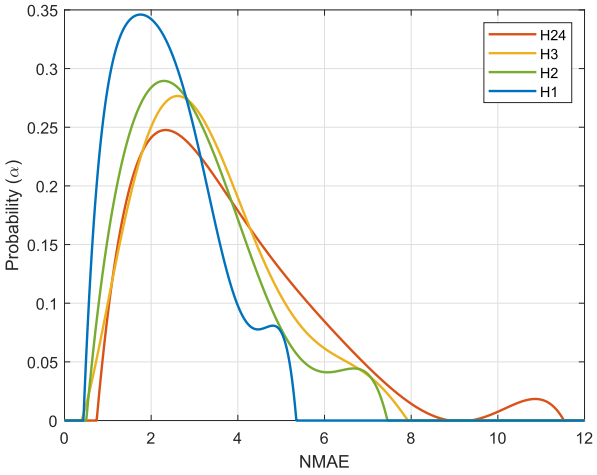


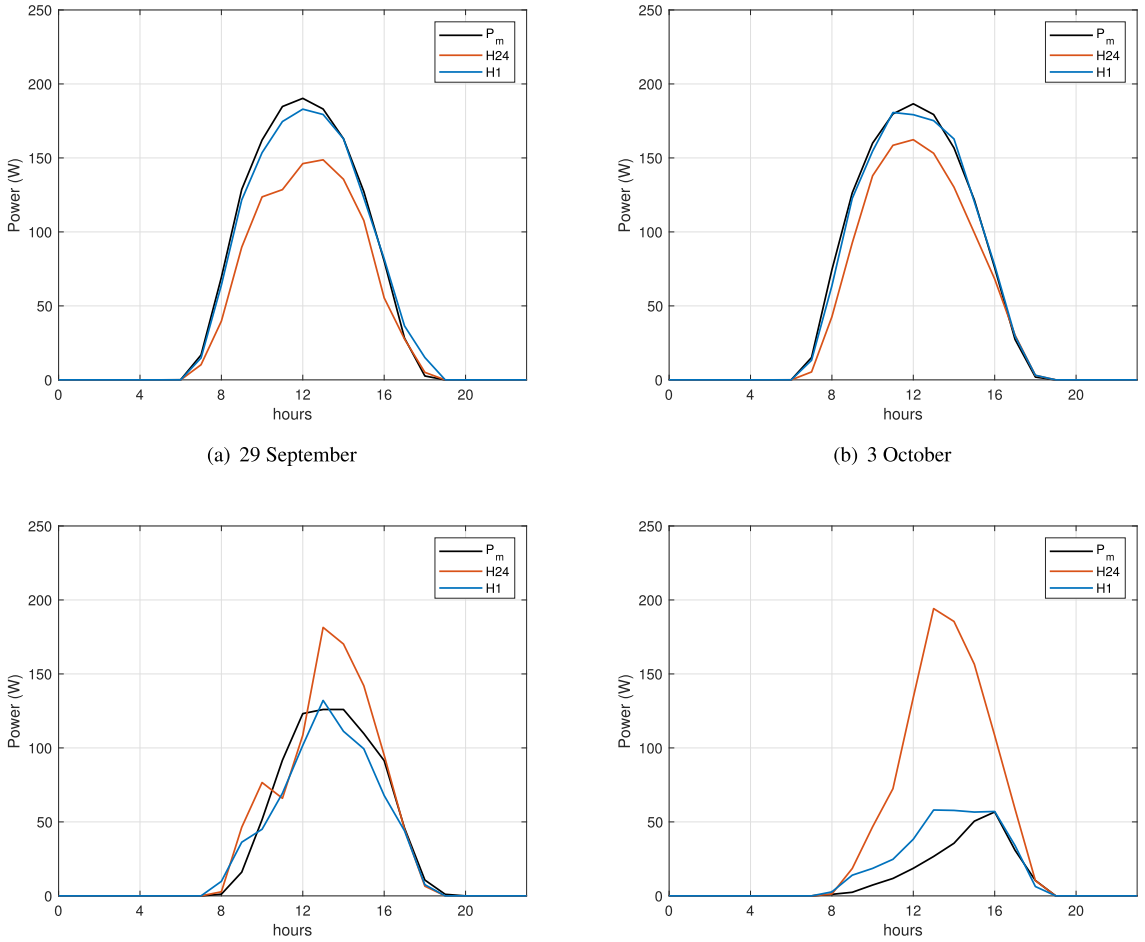
FIGURE 7. Intraday refinement: probability density function (PDF) for the 24-hours ahead forecast and the following refinements.

VI. RESULTS AND DISCUSSION

In this section, the main results which have been obtained through the implementation in SolarTech^{LAB} of the above-mentioned procedures are presented. Some significant examples providing the effectiveness of this new comprehensive forecasting methodology, on different time horizons, are here detailed and critically discussed giving a promising outlook for future implementation.

A. INTRADAY REFINEMENT

When the methodology explained in section II-B is implemented, significant improvement can be observed in the prediction outcome with respect to the 24-hours ahead logic. In Fig. 6, the obtained results of the intraday refinement are presented. In this graph, four cumulative distribution functions (CDF) for the error committed in terms of NMAE are provided, one for each time horizon considered. From the right, the orange line is representative of the 24-hours ahead



(a) 29 September (b) 3 October

(c) 19 October (d) 20 October

FIGURE 8. Intraday refinement: comparison of 24-hours ahead forecast, 1 hour ahead intraday refinement and measured power for four exemplifying days.

TABLE 5. Percent point in terms of NMAE (%) for three selected values of the probability α.

Horizon	α=0.80	α=0.90	α=1
H24	5.08	6.09	11.99
H3	4.57	5.40	9.64
H2	3.96	4.95	7.87
H1	3.28	3.85	5.54

TABLE 6. Forecast comparison between the 24 hours ahead logic and the 1 hour ahead refinement in terms of mean and standard deviation of the error committed.

Metric	Mean H24	Mean H1	Std H24	Std H1
NMAE (%)	4.29	2.72	2.19	1.3
WMAE (%)	62.51	34.43	147.42	84.19
EMAE (%)	32	22.13	19.67	14.38
nRMSE (%)	26.15	14.75	82.53	47.03
OMAE (%)	16.48	10.02	9.31	4.33

forecast while the yellow, green and blue lines (H3, H2 and H1 respectively) correspond to the forecast adjustment provided 3, 2 and 1 hour in advance. Moving from the y-axis, as shown by the arrows, the evaluation of the probability of committing an error lower or equal than the corresponding value on the x-axis can be made. This can be expressed in terms of percent point, hence assessing the probability that the variable X is less than or equal to x for a given α, as in (6).

$$F(x) = Pr[X \leq x] = \alpha \tag{6}$$

In Table 5, three values of x (NMAE) for different choices of α (0.80, 0.90 and 1) are given. As can be seen, the decrease of the forecast horizon allows to reduce both the maximum error committed ($\alpha = 1$) as well as the overall number of days in which the error committed exceeds a given threshold.

In Fig. 7, on the other hand, the Probability Density Function (PDF) of the error committed (NMAE), corresponding to the CDF above, are provided. As it is possible to see, reducing the time horizon, the error committed decreases accordingly. Moreover, with respect

TABLE 7. Performance comparison between the 24-hours ahead and intraday refinement methodology for the four selected days.

		NMAE (%)	WMAE (%)	EMAE (%)	nRMSE (%)	OMAE (%)
29 September	H24	5.5	24.2	24.1	11.7	19.2
	H1	1.2	5.3	5.2	2.6	4.2
3 October	H24	3.9	17.6	17.5	8.3	13.9
	H1	0.8	3.6	3.5	1.8	2.8
19 October	H24	4.1	30.1	24.2	15.0	16.2
	H1	2.4	17.6	16.9	8.0	9.5
20 October	H24	12.5	291.8	74.5	104.3	50.3
	H1	2.1	49.4	33.4	17.5	8.5

to the 24 hours ahead prediction, the overall dispersion around the mean value is lowered as well, as detailed in Table 6 for a single forecast horizon (H1). Worth highlighting is that all the performance metrics reported show the same improving pattern I, evaluated according to (7), almost halving both the standard deviation and the mean value.

$$I = \frac{E_b - E_p}{E_b} \cdot 100 \tag{7}$$

In this equation, E_b stands for the benchmark performance while E_p is the error associated to the proposed method. For example, an improvement I of 36.6% is faced in terms of mean value and of 40.6% in terms of standard deviation for the NMAE metric.

In Fig. 8, an example of the improvement provided by the presented technique is given. In the two upper graph for the 29^{th} of September and the 3^{rd} of October, the 24-hours ahead forecast tends to underestimate the power production from the solar source. Due to the ambient measurements acquired in the current day, the intraday algorithm was able to correct the underestimation, returning a more accurate outcome. The opposite situation can be observed in the two lower graphs, corresponding to the 19^{th} and 20^{th} of October. The 24-hours ahead prediction, due to inaccuracies in the weather forecasts, tends to overestimate the real production, while the intraday correction, relying on updated information, is able to capture the real trend and correct the outcome.

The same conclusion can be drawn inspecting Table 7, where the performance indicators are reported. For these exemplifying days, the adoption of the intraday refinement allows to greatly reduce forecast inaccuracies for all the analyzed metrics. For example, considering the NMAE, an improvement of 78.2%, 79.5%, 41.5% 83.2% is assessed for the 29^{th} of September, 3^{rd} of October, 19^{th} of October and 20^{th} of October respectively.

B. NOWCASTING TECHNIQUES COMPARISON

When the proposed methodologies described in section II-C are applied for the forecast of the following 30 minutes, the graph in Fig. 9, is produced. In this graph, the performances (MAE) of the naive persistence model is presented in blue, while in yellow and orange the smart persistence and ANN are shown respectively. On the x-axis the forecast horizon can be found. In the zoomed box, an insight into the first 3 minutes ahead is provided. As can be seen, the two persistence-based models provide a similar result when reducing the forecast horizon (≤ 3 min), outperforming the ANN model. On the other hand, increasing the forecast horizon (≥ 5 min), the naive persistence model shows the worst performance accuracy due to the neglected variation of the solar position, hence of the theoretically available radiation. When considering ANN results, it is possible to observe that are not as stable as the other two, due to

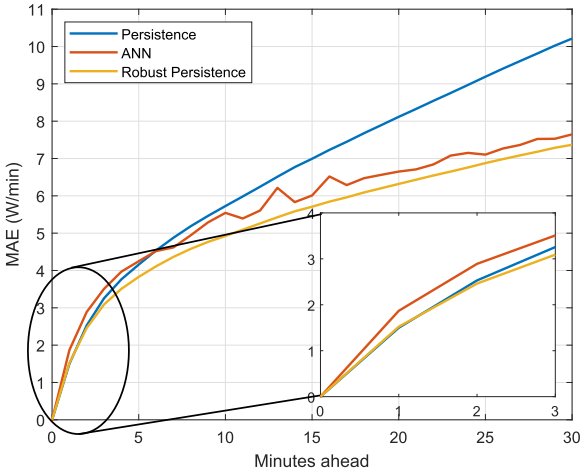
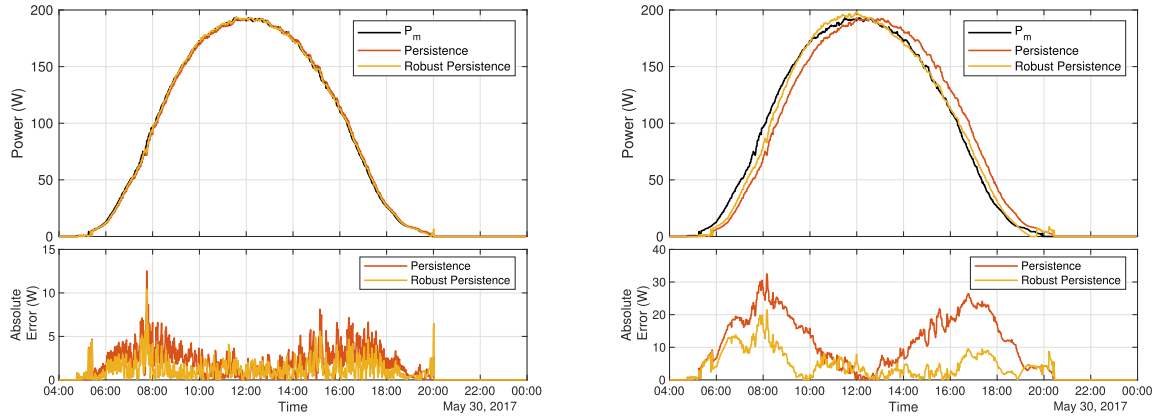


FIGURE 9. Nowcasting: comparison among the proposed methodologies, namely Persistence, ANN and Robust Persistence. In the zoomed box, the behavior for the first 3 minutes is given.

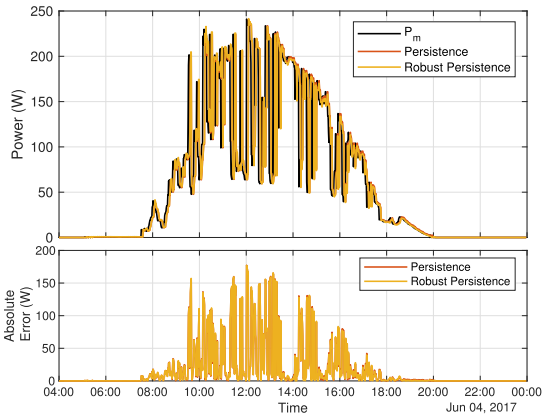
TABLE 8. Performance comparison among the nowcasting techniques in terms of MAE (W/min).

Horizon (min.)	Persistence	ANN	Robust Persistence
1	1.50	1.87	1.52
2	2.53	2.89	2.46
3	3.26	3.51	3.09
4	3.76	3.97	3.51
5	4.16	4.25	3.83
⋮	⋮	⋮	⋮
10	5.73	5.54	4.92
15	6.99	6.01	5.71
20	8.12	6.65	6.32
25	9.19	7.10	6.88
30	10.21	7.65	7.37

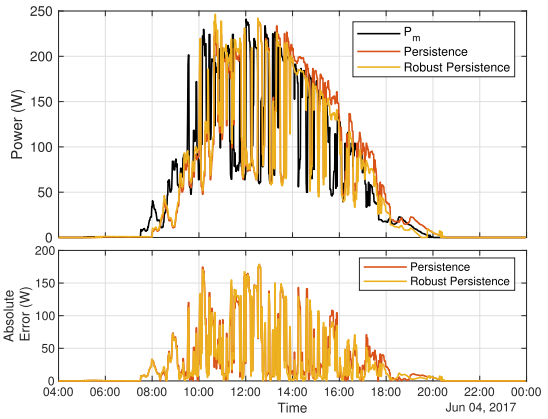


(a) 30 May: 5 minutes ahead forecast

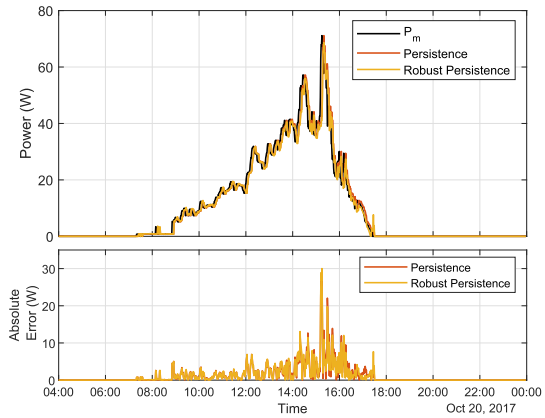
(b) 30 May: 30 minutes ahead forecast



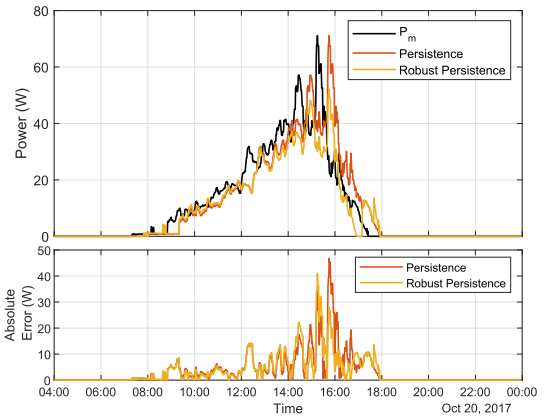
(c) 4 June: 5 minutes ahead forecast



(d) 4 June: 30 minutes ahead forecast



(e) 20 October: 5 minutes ahead forecast



(f) 20 October: 30 minutes ahead forecast

FIGURE 10. **Comparison among the measured power and the forecast obtained with the persistence and robust persistence models.**

variability and stochasticity in the training process. Worth mentioning is that ANNs, thanks to their learning process and to the ability of generalizing trends, take into account partial shading. Despite this, the overall error committed by this methodology is larger than the one committed by the robust persistence, allowing to conclude that, for the case under study, partial shading do not play a significant role and considering it would not provide any improvement. Moreover, from a computational point of view, the robust persistence methodology is much lighter, needing a single

TABLE 9. Nowcasting: power forecast accuracy evaluation by means of MAE (W).

		Persistence	Robust Persistence	% Improvement	Day Type
30 May	5 min	1.42	0.81	43.0%	Sunny
	30 min	8.00	3.30	58.8%	
4 June	5 min	14.28	14.16	0.8%	Rapidly changing
	30 min	21.38	19.77	7.5%	
20 October	5 min	1.01	0.98	3.0%	Cloudy
	30 min	2.71	2.55	5.9%	

algebraic operation compared with the whole neural training process. Therefore, even though robust persistence and ANN model show an asymptotic behaviour ($\geq$ 15 min), the former should be preferred in the nowcasting context. When further increasing the forecast horizon ($\geq$ 1h), as reviewed in [9], AI models, and in particular Physical Hybrid Artificial Neural Network (PHANN) [29] generally outperform all the others.

In Table 8 the performances obtained with the three presented methodologies are given in terms of MAE, considering different time horizons. The first five minutes are reported with a minute granularity to appreciate the discrepancies among the models performances. For the sake of readability, and being lower the variation from one minute to the other, the time gap is increased to 5 minutes. From these results, it is possible to grasp how the robust persistence outperforms the other methodologies from the second minute on.

In Fig. 10 the nowcast for three selected days, together with the error committed, is shown. The comparison is performed over two different time horizon, 5 and 30 minutes. The two upper graphs are referred to the 30^{th} May 2017, the two in the middle are from the 4^{th} of June, while the lower two are from the 20^{th} of October of the same year. In these graphs, the smart persistence is compared with a widely adopted benchmark, the naive persistence. The ANN methodology has been excluded from the present analysis due to its worse overall performances as it has been demonstrated in the previous paragraphs. In near-to-clear-sky weather conditions, as the case of the 30^{th} of May, the proposed method outperforms the naive persistence on both the 5 and 30 minutes time horizon. Furthermore, the beneficial of tacking into account the solar position correction, hence adopting the robust persistence, is particularly noticeable when the forecast horizon is increased. In case of rapidly changing weather conditions (e.g. the 4^{th} of June) or overcast weather conditions (e.g. the 20^{th} of October), the improvement given by the implementation of the described methodology is less relevant. In the first case, in fact, statistical methods, not relying on surrounding information such as sky images, fail at identifying quick power ramps and drops. In these situations, in fact, there is no evidence of correlation between the power production across several minutes. In the second one, on the other hand, being the diffuse component of the radiation predominant, the dependence on the incidence angle is less relevant when compared with the other factors playing a role in the production (e.g. cloud movement, composition, altitude and transmittance).

In Table 9, the error (MAE) committed for the three analyzed days is given for both the considered time horizons and the selected methodologies. As can be noticed, the robust persistence allows to greatly increase the forecast accuracy with respect to the naive persistence during sunny days, introducing a correction for the solar position. This allowed to approximately halve the error committed (reduction of the MAE of 43.0% and 58.8% for 5 and 30 minutes ahead respectively). On the other hand, in cloudy or rapidly changing conditions, the adoption of the proposed methodology, is less impacting on the overall result (MAE reduction comprised between 0.8% and 7.5%). Worth highlighting is that, despite providing a less marked improvement in this particular situations, the robust persistence does not worsen the overall forecast accuracy.

One of the main drawbacks regarding statistical methodologies, such as the ones presented in this work, is the inability of foreseeing ramps, both positive and negative. This may cause problems, for example, in the control strategy of microgrids. In fact, those occurrences are not particularly critical from the Energy Management System point of view, being marginal their contribution in terms of energy, but can cause severe problems from the instantaneous power management perspective.

VII. CONCLUSION

In the current work, a unique and comprehensive methodology aiming at providing a more accurate forecast on different time scales: from the following 24 hours to few minutes ahead. The proposed methodology has been assessed and validated on a freely available dataset recorded at SolarTech^{LAB}, which includes real data from an existing PV plant and from a meteorological station installed in the same location. More into detail, moving from the 24-hours ahead forecast, the predictions are continuously updated including in the analysis newer information collected in real time. The correction takes place on two different scales: the first one, 1- to 3-hours ahead, constitutes the intraday refinement, while the second one is related to the forecast of the following 30 minutes on a minutely basis (nowcasting).

The day ahead power forecast is assessed through the adoption of PHANN, a hybridized machine learning method effectively adopted and already presented in previous works here referenced. In order to further refine these predictions a purely stochastic forecasting method (ANN) is adopted for the next 3 hours. This technique allows to greatly improve the accuracy of the PV power previously estimated. For the year

TABLE 10. Sample chunk of the dataset available at [38].

Timestamp	P_m (W)	T_{air} (°C)	GHI (W/m^2)	$GPOA$ (W/m^2)	W_s (m/s)	W_d (°)
⋮	⋮	⋮	⋮	⋮	⋮	⋮
25/02/2017 10:00	158,7	8,97	432	684	0,58	153,6
25/02/2017 10:01	156,6	9,03	435	687	0,63	167,8
25/02/2017 10:02	156,6	8,97	435	686	1,54	58,7
25/02/2017 10:03	156,6	9,23	438	693	0,14	258,1
25/02/2017 10:04	160,0	9,40	442	698	0,79	349,8
25/02/2017 10:05	160,0	9,35	445	702	2,18	54,6
25/02/2017 10:06	159,5	9,33	444	704	2,39	58,0
25/02/2017 10:07	160,2	8,90	446	703	1,43	56,6
25/02/2017 10:08	164,2	9,01	447	707	1,59	68,8
25/02/2017 10:09	164,2	8,69	461	726	2,23	71,6
25/02/2017 10:10	164,6	8,45	461	725	2,28	76,8
25/02/2017 10:11	166,3	8,56	459	724	0,79	152,3
25/02/2017 10:12	166,3	9,28	462	731	1,65	71,0
25/02/2017 10:13	166,3	9,03	467	737	0,47	138,4
25/02/2017 10:14	166,6	9,15	468	737	1,86	124,8
25/02/2017 10:15	166,6	8,58	468	736	0,95	65,4
25/02/2017 10:16	166,2	8,52	468	736	0,47	54,2
25/02/2017 10:17	164,9	8,76	468	735	0,63	4,8
25/02/2017 10:18	164,9	8,95	462	727	0,00	0,1
25/02/2017 10:19	164,9	9,39	464	730	0,68	10,7
25/02/2017 10:20	167,4	9,53	470	738	1,01	35,6
⋮	⋮	⋮	⋮	⋮	⋮	⋮

2017 under analysis, considering the NMAE, a reduction of 36.6% is achieved over the whole period.

As regards nowcasting techniques, it is decided to implement a novel statistical methods, namely the Robust Persistence, which corrects the naive persistence with information related to the solar position. This implementation allowed to further minimize the uncertainty associated with the deterministic variation of the solar radiation. This methodology is particularly useful under near-to-clear-sky conditions, allowing to achieve a reduction of the forecast error with respect to the benchmark up to 58.8%. However, it should be highlighted that, in all the other cases, the robust persistence provides slight enhancements and does not worsen the overall forecast accuracy.

Finally, future steps will include the daily update of the available dataset in order to validate the here proposed methodology on a wider case study.

APPENDIX
DATASET STRUCTURE OVERVIEW

The dataset of PV measurements employed in this paper has been made freely available on the SolarTech$^{\text{LAB}}$ website [38] and in IEEE Dataport [41], for scientific research purpose and further data validation. In particular, the dataset is composed of the following variables and specifics, with a time resolution of 1 minute:

- *Timestamp*: column with time recordings; the data format is "dd-MM-yyyy hh:mm:ss", with the time always expressed in Central European Time (CET), *i.e.* UTC+01:00, both in winter and in summer, thus DST offset is not recorded.
- P_m: power recordings from the PV module (W); module tilt: 30°.
- T_{air}: Ambient temperature (°C) measured by the weather station described in section IV–B.
- *GHI*: Global horizontal irradiance (W/m^2) measured by the weather station described in section IV–B.
- *GPOA*: Global irradiance measured on the tilted plane (30°).
- W_s: Wind speed (m/s) measured by the weather station described in section IV–B.
- W_d: Wind direction (°), assuming 0° east, positive south.

It is worth noticing that the dataset, provided in.csv format, includes original measurements, *i.e.* it did not undergo the cleaning process described in section V, thus these raw data can be used for any additional post-processing, validation and further research. When a value is missing in the original measurements recording, a "NaN" is reported. An example of the recordings, available at [38], is given in Table 10.

REFERENCES

[1] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Sol. Energy*, vol. 136, pp. 78–111, Oct. 2016.

[2] L. Moretti, S. Polimeni, L. Meraldi, P. Raboni, S. Leva, and G. Manzolini, "Assessing the impact of a two-layer predictive dispatch algorithm on design and operation of off-grid hybrid microgrids," *Renew. Energy*, vol. 143, pp. 1439–1453, Dec. 2019.

[3] D. T. Ho, J. Frunt, and J. M. A. Myrzik, "Photovoltaic energy in power market," in *Proc. 6th Int. Conf. Eur. Energy Market (EEM)*, 2009, pp. 7–11.

[4] F. Nicolli and F. Vona, "Energy market liberalization and renewable energy policies in OECD countries," *Energy Policy*, vol. 128, pp. 853–867, May 2019.

[5] L. Gigoni, A. Betti, E. Crisostomi, A. Franco, M. Tucci, F. Bizzarri, and D. Mucci, "Day-ahead hourly forecasting of power generation from photovoltaic plants," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 831–842, Apr. 2018.

[6] F. Bizzarri, M. Bongiorno, A. Brambilla, G. Gruosso, and G. S. Gajani, "Model of photovoltaic power plants for performance analysis and production forecast," *IEEE Trans. Sustain. Energy*, vol. 4, no. 2, pp. 278–285, Apr. 2013.

[7] S. Lee, H. Jin, L. F. Vecchietti, J. Hong, and D. Har, "Short-term predictive power management of PV-powered nanogrids," *IEEE Access*, vol. 8, pp. 147839–147857, 2020.

[8] T. Khalili, A. Jafari, M. Abapour, and B. Mohammadi-Ivatloo, "Optimal battery technology selection and incentive-based demand response program utilization for reliability improvement of an insular microgrid," *Energy*, vol. 169, pp. 92–104, Feb. 2019.

[9] A. Mellit, A. M. Pavan, E. Ogliari, S. Leva, and V. Lughi, "Advanced methods for photovoltaic output power forecasting: A review," *Appl. Sci.*, vol. 10, no. 2, p. 487, Jan. 2020.

[10] X. Chen, Y. Du, E. Lim, H. Wen, and L. Jiang, "Sensor network based PV power nowcasting with spatio-temporal preselection for grid-friendly control," *Appl. Energy*, vol. 255, Dec. 2019, Art. no. 113760.

[11] T. Khalili, S. Nojavan, and K. Zare, "Optimal performance of microgrid in the presence of demand response exchange: A stochastic multi-objective model," *Comput. Electr. Eng.*, vol. 74, pp. 429–450, Mar. 2019.

[12] R. Ulbricht, U. Fischer, W. Lehner, and H. Donker, "First steps towards a systematical optimized strategy for solar energy supply forecasting," in *Proc. 1st Int. Workshop Data Anal. Renew. Energy Integr. (DARE)*, 2013, pp. 14–25.

[13] D. Yang, J. Kleissl, C. A. Gueymard, H. T. C. Pedro, and C. F. M. Coimbra, "History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining," *Sol. Energy*, vol. 168, pp. 60–101, Jul. 2018.

[14] D. Su, E. Batzelis, and B. Pal, "Machine learning algorithms in forecasting of photovoltaic power generation," in *Proc. Int. Conf. Smart Energy Syst. Technol. (SEST)*, Sep. 2019, pp. 10–15.

[15] U. K. Das, K. S. Tey, M. Seyedmahmoudian, S. Mekhilef, M. Y. I. Idris, W. Van Deventer, B. Horan, and A. Stojcevski, "Forecasting of photovoltaic power generation and model optimization: A review," *Renew. Sustain. Energy Rev.*, vol. 81, pp. 912–928, Jan. 2018.

[16] S. Al-Dahidi, M. Louzazni, and N. Omran, "A local training strategy-based artificial neural network for predicting the power production of solar photovoltaic systems," *IEEE Access*, vol. 8, pp. 150262–150281, 2020.
[17] S. Sobri, S. Koohi-Kamali, and N. A. Rahim, "Solar photovoltaic generation forecasting methods: A review," *Energy Convers. Manage.*, vol. 156, pp. 459–497, Jan. 2018.
[18] S. I. Vagropoulos, G. I. Chouliaras, E. G. Kardakos, C. K. Simoglou, and A. G. Bakirtzis, "Comparison of SARIMAX, SARIMA, modified SARIMA and ANN-based models for short-term PV generation forecasting," in *Proc. IEEE Int. Energy Conf. (ENERGYCON)*, Apr. 2016, pp. 1–6.
[19] A. Catalina, A. Torres-Barrán, C. M. Alaíz, and J. R. Dorronsoro, "Machine learning nowcasting of PV energy using satellite data," *Neural Process. Lett.*, vol. 52, no. 1, pp. 97–115, Aug. 2020.
[20] T. Masuta, D. Kobayashi, H. Ohtake, and N. H. Viet, "Evaluation of unit commitment based on intraday few-hours-ahead photovoltaic generation forecasts to reduce the supply-demand imbalance," in *Proc. 8th Int. Renew. Energy Congr. (IREC)*, Mar. 2017, pp. 1–5.
[21] J. M. Bright, S. Killinger, D. Lingfors, and N. A. Engerer, "Improved satellite-derived PV power nowcasting using real-time power data from reference PV systems," *Sol. Energy*, vol. 168, pp. 118–139, Jul. 2018.
[22] A. Catalina, C. M. Alaiz, and J. R. Dorronsoro, "Combining numerical weather predictions and satellite data for PV energy nowcasting," *IEEE Trans. Sustain. Energy*, vol. 11, no. 3, pp. 1930–1937, Jul. 2020.
[23] G. Almonacid-Olleros, G. Almonacid, J. I. Fernandez-Carrasco, M. E. Estevez, and J. M. Quero, "A new architecture based on IoT and machine learning paradigms in photovoltaic systems to nowcast output energy," *Sensors*, vol. 20, no. 15, pp. 1–16, 2020.
[24] S. Dutta, Y. Li, A. Venkataraman, L. M. Costa, T. Jiang, R. Plana, P. Tordjman, F. H. Choo, C. F. Foo, and H. B. Puttgen, "Load and renewable energy forecasting for a microgrid using persistence technique," *Energy Procedia*, vol. 143, pp. 617–622, Dec. 2017.
[25] A. Kumler, Y. Xie, and Y. Zhang, "A physics-based smart persistence model for intra-hour forecasting of solar radiation (PSPI) using GHI measurements and a cloud retrieval technique," *Sol. Energy*, vol. 177, pp. 494–500, Jan. 2019.
[26] Z. Li, S. Rahman, R. Vega, and B. Dong, "A hierarchical approach using machine learning methods in solar photovoltaic energy production forecasting," *Energies*, vol. 9, no. 1, p. 55, Jan. 2016.
[27] SolarTech Lab. *SolarTech Lab Test Facility*. Accessed: Oct. 10, 2020. [Online]. Available: http://www.solartech.polimi.it/
[28] S. Leva, M. Mussetta, and E. Ogliari, "PV module fault diagnosis based on microconverters and day-ahead forecast," *IEEE Trans. Ind. Electron.*, vol. 66, no. 5, pp. 3928–3937, May 2019.
[29] A. Nespoli, M. Mussetta, E. Ogliari, S. Leva, L. Fernández-Ramírez, and P. García-Triviño, "Robust 24 hours ahead forecast in a microgrid: A real case study," *Electronics*, vol. 8, no. 12, pp. 1–13, 2019.
[30] M. Bouzerdoum, A. Mellit, and A. M. Pavan, "A hybrid model (SARIMA–SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant," *Sol. Energy*, vol. 98, pp. 226–235, Dec. 2013.
[31] E. Ogliari, A. Dolara, G. Manzolini, and S. Leva, "Physical and hybrid methods comparison for the day ahead PV output power forecast," *Renew. Energy*, vol. 113, pp. 11–21, Dec. 2017.
[32] A. Dolara, F. Grimaccia, S. Leva, M. Mussetta, and E. Ogliari, "A physical hybrid artificial neural network for short term forecasting of PV plant power output," *Energies*, vol. 8, no. 2, pp. 1138–1153, Feb. 2015.
[33] A. Dolara, F. Grimaccia, S. Leva, M. Mussetta, and E. Ogliari, "Comparison of training approaches for photovoltaic forecasts by means of machine learning," *Appl. Sci.*, vol. 8, no. 2, p. 228, Feb. 2018.
[34] F. Grimaccia, S. Leva, M. Mussetta, and E. Ogliari, "ANN sizing procedure for the day-ahead output power forecast of a PV plant," *Appl. Sci.*, vol. 7, no. 6, p. 622, Jun. 2017.
[35] A. Nespoli, E. Ogliari, A. Dolara, F. Grimaccia, S. Leva, and M. Mussetta, "Validation of ANN training approaches for day-ahead photovoltaic forecasts," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2018, pp. 1–4.
[36] *Photovoltaic System Performance—Part 1: Monitoring*, NSAI Standard, Irish Standard IEC 61724, 2017.
[37] R. E. Bird and C. Riordan, "Simple solar spectral model for direct and diffuse irradiance on horizontal and tilted planes at the earth's surface for cloudless atmospheres," *J. Climate Appl. Meteorol.*, vol. 25, no. 1, pp. 87–97, Jan. 1986.
[38] *SolarTech Lab Available Dataset*. Accessed: Oct. 10, 2020. [Online]. Available: http://www.solartech.polimi.it/activities/forecasting/dataset
[39] G. W. Hughes, "Engineering astronomy," Sandia Nat. Lab., Albuquerque, NM, USA, Tech. Rep. 5323, 1985.
[40] W. F. Holmgren, C. W. Hansen, and M. A. Mikofski, "Pvlib Python: A Python package for modeling solar energy systems," *J. Open Source Softw.*, vol. 3, no. 29, p. 884, Sep. 2018.
[41] S. Leva, A. Nespoli, S. Pretto, M. Mussetta, and E. Ogliari, "Photovoltaic power and weather parameters," *IEEE Dataport*, 2020. Accessed: Oct. 26, 2020. [Online]. Available: http://dx.doi.org/10.21227/42v0-jz14

SONIA LEVA (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the Politecnico di Milano, Italy, in 2001. She is currently a Full Professor of electrical engineering with the Department of Energy, Politecnico di Milano. She is also the Director of the Solar Tech Laboratory (SolarTech^{LAB}) and the Laboratory of MicroGrids (MG^{2}LAB), Politecnico di Milano. She is a Senior Member of the IEEE Power and Energy Society and a member of the IEEE Working Group Distributed Resources Modeling and Analysis.

ALFREDO NESPOLI (Graduate Student Member, IEEE) received the B.Sc. and M.Sc. degrees in energy engineering from the Politecnico di Milano, in 2016 and 2018, respectively, where he is currently pursuing the Ph.D. degree, in November 2019. He was a Temporary Researcher with the Electrical Division, Energy Department, Politecnico di Milano. His research interests include AI and renewable energy sources.

SILVIA PRETTO received the B.Sc. degree in energy engineering and the M.Sc. degree in energy engineering and management engineering from the Politecnico di Milano, in 2016 and 2019, respectively. She is currently with the Politecnico di Milano as a Researcher Assistant. Her research interest includes RES power forecast.

MARCO MUSSETTA (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the Politecnico di Milano, Italy, in 2007. He is currently an Associate Professor of electrical engineering with the Department of Energy, Politecnico di Milano. He is a Senior Member of the IEEE IES. He is the Chair of the IEEE Computational Intelligence Society (CIS) Task Force on Fuzzy Systems in Renewable Energy and Smart Grid.

EMANUELE GIOVANNI CARLO OGLIARI (Member, IEEE) received the M.Sc. and Ph.D. degrees in electrical engineering from the Politecnico di Milano, Italy, in 2016. His thesis was Novel Methods for Photovoltaic Energy Forecasting of the Day-Ahead Hourly Power Curve. He has been working on photovoltaic power plant design and their optimization since 2010 and has also been on RES expected power by means of computational intelligence techniques since 2012. He is currently an Assistant Professor with the Department of Energy, Politecnico di Milano, where he also teaches electrical engineering.