Optimisation of Generation Models for Clusters of Photovoltaic Plants

Optimisation of Generation Models for Clusters of Photovoltaic Plants

Alessandro Ciocia, Gianfranco Chicco, Filippo Spertino
Dipartimento Energia "Galileo Ferraris"
Politecnico di Torino
Torino, Italy
{alessandro.ciocia, gianfranco.chicco, filippo.spertino}@polito.it

***Abstract*— The present work analyses the actual production profiles of a group of tens of thousands of PV plants. Actual PV generation profiles are represented as hourly average powers gathered from the energy meters measured at the point of common coupling with the grid. After filtering, data cleaning and statistical analysis, reference PV plants are selected for the improvements of the PV generation models. Production profiles are calculated by using literature generation models, with weather data as inputs. An optimisation is performed on the parameters of the literature models to minimise the differences between the cumulative distribution functions of calculated profiles and measured data. The study compares the performance of the different models and shows how the optimisation increases the quality of the calculated profiles.**

Keywords— photovoltaic systems, modelling, optimisation.

I. INTRODUCTION

The integration of renewable energy systems into the electricity grid is one of the key points in the energy transition. For example, in Italy more than one million of photovoltaic (PV) plants are currently operating in 2022, leading to new challenges in the grid management. To properly manage the intermittent generation of huge groups of actual PV plants, adequate and realistic models for their cumulative production are necessary. The literature presents many works with PV generation models [1][2]. Nevertheless, in these models there are two main issues. The first issue is that the models typically refer to well-working ideal PV plants, which do not represent the real situation for the installed PV generators. For example, many of the actual PV plants do not have adequate design and installation, so that components are not correctly working [3], and shadows often occur [4]. The second issue is the correct calculation of the temperature of the photovoltaic cells, which is a function of type of installation and weather conditions. To solve these issues and obtain more accurate simulated profiles, the models can be improved using new parameters in the original model, and by changing some of the default parameters to better match experimental data. The values of the new parameters are defined by an optimisation method described in the present work. The optimisation is performed on a cluster of selected plants, which represents all the generators in a wide geographic area. A stratified sampling technique is used to check if the selected plants statistically represent the whole population of generators.

II. MODELLING OF PRODUCTION FROM PHOTOVOLTAIC GENERATORS

A. *Active power production model*

The active power production from a generic photovoltaic plant can be calculated by a model proportional to irradiance G and dependent on the temperature of the PV modules (5):

$$P_{AC} = G \cdot \eta_{STC} \cdot \eta_{low,G} \cdot C_T \cdot \eta_{mix} \cdot \eta_{conv} \cdot \eta_l \quad (1)$$

where:

- η_{STC} is the rated efficiency of the plant defined in the Standard Test Conditions (STC, irradiance G=1000 W/m^2, cell temperature T = 25°C and air mass 1.5) [6].
- $\eta_{low,G}$ is the efficiency used to take into account the nonlinearity effects of semiconductor technology as irradiance G changes:

$$\eta_{lowG} = 1 - G/G_0 \quad (2)$$

- G_0 is defined as the irradiance value for which modules do not produce due to loss effects within modules. This value is in the range 10÷50 W/m^2; as a result, with G <500 W/m^2 the efficiency of the PV modules is strictly non-linear, while it is almost linear for higher G values. Obviously, in case of $G \leq G_0$, the PV production P_{AC} is null [6].
- C_T is the parameter that describes the linear dependence of production on the cell temperature:

$$C_T = 1+\gamma_T(T-T_{STC}) \quad (3)$$

When the cell temperature T is above T_{STC} = 25°C, the production decreases proportionally. The thermal coefficient of power γ_T is in the range -(0.3÷0.5) %/°C for crystalline silicon technology. Regarding the correlation between the cell temperature and the weather data, a detailed description is present in the next paragraphs.

- $\eta_{mix} = \eta_{life}\ \eta_{dirt}\ \eta_{refl}\ \eta_{mis}\ \eta_{cable}$ is the overall performance of the PV plant. Production decreases due to the reflection effect and dirt deposited on the glass of the modules, the mismatch of *I-V* curves, and the joule losses in cables. These experimental loss parameters are considered constant [7].
- The loss by ageing $\eta_l = 1\text{-}\gamma_{life}\cdot n_{life}$ is proportional to the age of the plant n_{life} by the annual loss coefficient γ_{life}. According to the specifications in the datasheets, the PV modules are certified to reduce their efficiency down to 80% of the initial value after 20 years (γ_{life}=-1%/year). Nevertheless, experimental results in [8] and in [9] show that the efficiency reduction is lower and proportional to the parameter γ_{life}=-0.5%/year.
- η_{conv} is the non-linear performance of the AC/DC converter, which into account Maximum Power Point Tracking (MPPT) and inverter. This efficiency is determined by a quadratic model, based on experimental data [10]. First, the model includes the no-load losses P_0 (W) due to the supply of auxiliary circuits. Then, the linear term of the losses is expressed by using the parameter C_L (W^{-1}) that takes into account conduction of diodes and IGBT, and switching losses. Finally, the quadratic term of the losses is expressed by using the parameter C_Q (W^{-2}) for the conduction of MOSFETs and the resistive contribution. The maximum efficiency of the converter is about 98%. The formula of the converter efficiency is:

$$\eta_{CONV} = \frac{P_{AC}}{P_{AC}+P_0+C_L\cdot P_{AC}+C_Q\cdot {P_{AC}}^2} \qquad (4)$$

In order to use the same AC/DC converter model for plants with different sizes, the previous formula is normalised.

B. Calculation of photovoltaic module temperature

Regarding the use the PV model described in (1), it is necessary to calculate the temperature T_c of the PV cells.

The model in [11] is based on the Normal Operating Cell Temperature (*NOCT*) and the environmental data used as inputs are the air temperature T_{air} and irradiance G:

$$T_c = T_a + \frac{NOCT-25°C}{G_{NOCT}}\cdot G \qquad (5)$$

The Normal Operating Cell Temperature (NOCT) is the experimental temperature indicated by the manufacturer in the datasheet of the PV module. It is measured in steady-state conditions, i.e., NOCT = 20 °C, G_{NOCT} = 800 W/m^2 with module inclined at 45° in stable open circuit conditions and wind speed of 1 m/s. As a result, the NOCT of commercial crystalline silicon modules is between 42÷50°C [12].

Another possible solution is the use of a formula based on a measurement campaign of weather data and PV modules temperature, as in the work presented in [13]. In this case, the cell temperature is calculated as a function of air temperature T_a, irradiance and wind speed w_s (m/s):

$$T_c = a_T\cdot T_a + b_T\cdot G - c_T\cdot w_s + d_T \qquad (6)$$

where the first parameter a_T=0.943 is dimensionless, b_T is equal to 0.028 °C/W, c_T =1.528 °C·s/m and d_T=4.3 °C.

With respect to the above-described models, the work presented in [14] proposes a different calculation of cell temperature (from now, this model will be called "MATTEI"). It is based on a simplified energy balance, in which radiative thermal exchange is considered negligible. Moreover, the temperature difference between the glass (cover) and the cell is neglected, and the temperature on the surface of the PV module is considered uniform. By including the phenomena of solar energy conversion in electricity, the calculation of the temperature of the module is expressed as follows:

$$T_c = \frac{U_{pv}T_a+\phi[(\alpha\tau)-\eta_r-\mu T_r]}{U_{pv}-\mu\phi} \qquad (7)$$

where η_r is equal to the efficiency of the module at STC, τ is the transmittance of materials, while α is the absorption coefficient. Both coefficients are the subject of study, and in the literature, there are different estimations. For example, the product ατ is equal to 0.9 in [15], while ατ=0. 875 in [16]. U_{pv} is a parameter that takes in consideration the heat exchange with air; thus, it strongly depends on the wind speed w_s. Generally, a simplifying hypothesis is that the heat exchange is present on the faces of the module neglecting the small lateral surfaces. In [16] there are different formulations of U_{pv}, deriving from different assumptions. In every case, the linear equations include the two terms e_T and f_T:

$$U_{pv} = e_T + f_T\cdot w_s \qquad (8)$$

Finally, another model for the calculation of the PV module temperature is presented in [17]. In this model (from now it will be called KING), the temperature is proportional to the ratio between the reference irradiance G_{STC}=1000 W/m^2 and the incident irradiance G on the surface of the PV module:

$$T_c = T_m + \frac{G}{G_{STC}} \cdot g_T \qquad (9)$$

where T_m (°C) is the temperature of the backside of the module. The cell temperature is slightly different from the temperature of the backside of the module. Nevertheless, the two temperatures are correlated according to the simple relationship shown above, based on the assumption of one-dimensional heat exchange [17]:

$$T_m = E \cdot \left[e^{h_T + i_T \cdot w_s}\right] + T_a \qquad (10)$$

As in the previous models, the wind direction is not considered, because it does not lead to improvements in the results of the model. The parameters g_T, h_T and i_T are empirical coefficients, obtained for modules made of different materials and related thicknesses, as a function of different installation conditions. All the sets of parameters are available in [17].

III. Procedure for the Analysis of a Cluster of PV Plants

The procedure for the analysis of a big cluster of PV plants requires the collection of the data about the plants, the measured production profiles, and the weather data. After that, production profiles are filtered to remove wrong data, and a classification of the database of plants is performed. The classification is necessary to check the statistical validation of the data remaining after the filtering with respect the whole cluster of plants. After the definition of the group of plants representing the whole cluster, the energy models are used to calculate the production profiles, which are compared with measured data. Finally, the energy deviations are minimised by the optimisation procedure. The details about the main step of the procedure are presented in the next subparagraphs and in Section IV.

A. Import of weather data and PV plants information

The measured production profiles of photovoltaic plants can be analysed, and simulation can be performed, only if same basic information is available. First at all, the rated power is necessary to analyse the measured profiles, and to perform simulations. As explained in the next subparagraph, rated powers are also the variables used to classify the plants. The site coordinates are essential to obtain meteorological data [18,19]. The tilt and azimuth of the PV modules can be obtained by the layout of the plants, by inspections, or by satellite/street images [20]. This info is necessary to obtain the irradiance on their surface at each considered time step. Finally, the installation year is necessary to calculate the degradation losses [21].

B. Production profiles filtering

Production profiles obtained by distributed meters can be affected by issues. In many cases, it is not possible to define the source of the wrong data; errors can occur in data measurement, transmission, and storage. Profiles with unexpected trends could be the results of plant shutdowns and failures of the components. Thus, an accurate filtering process of the data is necessary. The procedure presented in [22] deals with the filtering of thousands of profiles, without the possibility of accessing detailed information of the plants. The available information is only the basic information described in the previous sub-paragraph. The first criterion is the check of the absence of production during night hours, as such production is obviously physically impossible. The second criterion is the check of the absence of days of data, due to the failure of the monitoring infrastructure. In fact, the profiles with missing days, weeks, and months of data affect the calculation of the performance of the plant. The third criterion is the check of the typical territorial range productivities. In particular, if the yearly specific production in kWh/kW/year is out of an accepted range, the performance is too low or too high, and the profile cannot be considered acceptable. Productivity maps, such as those available in the PVGIS database [18] can be used to define the boundaries of the acceptable performance ranges, depending on the site location.

C. Data classification based on rated powers of the plants

Inferential statistical methods allow us generalising the results of a whole population of elements with an acceptable confidence, by analysing a small sample of elements. These methods can be used in the present work, i.e., in case of large clusters of PV systems, where most of the data about the plants (the power profiles) are affected by errors, resulting not useable. Thus, the elements of the population are the sizes of the PV plants, and the Neyman's Stratified Sampling (SS) technique [23] is applied. As described in [22], the SS techniques indicate the production profiles surviving the data filtering that are significant enough to statistically represent the whole population of PV plants.

D. Power profiles comparison and energy deviations

The analysis of the results is carried out through the hourly comparison of the estimated profile with the measured production. The estimated profiles are calculated by using the energy model (1) and one of the temperature models described in Section II. The error between the different profiles is assessed in energy terms:

$$\Delta E_{\%_{s,i}}^{\Delta t} = 100 \cdot \frac{E_{s,i}-E_{m,i}}{E_{m,i}}\Big|_{\Delta t} \quad (11)$$

where $E_{s,i}$ is the simulated energy in the s^{th}-model s = 1,2…,5 and $E_{m,i}$ is the measured energy for the i^{th} plant, for the same timeframe Δt equal to one year. To evaluate errors for shorter periods, such as one day, it may be useful to calculate the residual ratio for each configuration:

$$R_{s,i}^{\Delta t} = \frac{E_{s,i}}{E_{m,i}}\Big|_{\Delta t} \quad (12)$$

The error calculated with (11) and (12) is positive or greater than unity when the model overestimates production with respect to the measurements.

IV. Optimisation Procedure of PV Model Parameters

A. Improvements in the production model

The main issue of the models described in the previous paragraphs is that (as in most of the models in the literature) these models are used to calculate the production in case of ideally fully functional PV systems. Nevertheless, the main goal the present work is to better match the actual production of working PV plants. Thus, the model can be improved by using new parameters in the original model and by changing some of the default values to better match the experimental data. Regarding the introduction of new parameters, an adjustment coefficient of power C_a is multiplied to the rated efficiency of the system η_{STC} to consider all the phenomena that are not taken into account in the original model. For example, a change of C_a with respect to the unitary value could lead to a high mismatch in current-voltage characteristics of the PV modules. The high mismatch can be due to a non-uniform installation condition (different tilts and azimuths), or to the use of PV modules with different specifications. AC_P can decrease due to the failure of components (e.g., AC/DC converters), due to the disconnection of portions of the plants or due to large amounts of dirt and bird dropping.

B. Optimisation problem formulation

The optimisation problem attempts to minimise the differences, over a defined time frame, between the measured production profiles and the calculated production profiles. The quality of the match between these profiles is quantified by calculating the sum of their root mean square difference at each time step of the measurement data. In case of an analysis of a group of j plants, the average quadratic deviation σ_P is calculated on the difference between the estimated power values P_{AC} and the measured values $P_{AC,m}$. For each plant, this quantity is normalised by the P_{PV} rated power of the plant. The calculation is performed over the time optimisation T_{opt} discretised with k elements.

The optimisation problem is formulated as follows:

$$\begin{cases} \min\limits_{x} \sigma(\mathbf{x}) \\ s.t: \ lb \le \mathbf{x} \le ub \end{cases} \quad (13)$$

$$\sigma(\mathbf{x}) = \sum_{j=1}^{J} \frac{\sqrt{\frac{1}{T_{opt}} \sum_{k=1}^{K} \left(P_{AC-k}^{j}(\mathbf{x}) - P_{AC,m-k}^{j}\right)^2}}{P_{PV,j}} \quad (14)$$

where $\mathbf{x}$ is the vector that includes the variables of the optimisation problem defined in (13). For the sake of clarity, $\mathbf{x}$ can be divided in two sub-vectors of variables $\mathbf{x}=[x_1, x_2]$. The first sub-vector x_1 includes the thermal factor $\gamma_{T\%}$, the low irradiance level G_0, and the new adjustment coefficient C_A:

$$x_1 = [\gamma_{T\%}, G_0, C_A] \quad (15)$$

The second sub-vector x_2 includes the variables to be modified for better calculation of the cell temperature T. For example, in case of the model in (5), the NOCT is optimised. The reason is that NOCT is calculated by manufacturers as a steady-state temperature, created in laboratory. Such equilibrium condition is strongly affected by the installation condition, which often does not permit an optimal heat dissipation in actual plants. It happens particularly in case of

building applied or building integrated plants, in which modules have actual steady-state temperatures at T_{NOCT} = 20 °C, G_{NOCT} = 800 W/m^2 higher than the declared NOCT.

$$x_2 = x_{2,NOCT} = [NOCT] \quad (16)$$

In case of the model defined in (6), all the four parameters can be modified. The change of the parameter a_T quantifies the effect of air temperature on the module temperature; the same approach is individually applied by considering b_T and c_T for the effect of irradiance and temperature, respectively. In every case, the parameter d_T is also changed. The analysis of individual parameters can be done to compare the variation of each one on the module temperature calculation. If all the four parameters are changed, the vector of optimised parameters will be:

$$x_2 = x_{2,w} = [a_T, b_T, c_T, d_T] \quad (17)$$

Following the same criteria, similar vectors of variables can be created to optimise the other temperature models. In case of the MATTEI model (7), the sub-vector x_2 is:

$$x_2 = x_{2,Mattei} = [e_T, f_T, \alpha\tau] \quad (18)$$

In case of the KING model (9), the sub-vector x_2 is:

$$x_2 = x_{2,King} = [g_T, h_T, i_T] \quad (19)$$

In every simulation, the optimisation is constrained. The terms *lb* and *ub* are respectively the lower and upper bounds vector of the inequality constraint, which limits the values that the variables can take. In case of parameters referring to technical aspects of the PV plants, there are ranges defined in the literature. For example, in case of the thermal coefficient of power, it is in the range -(0.3÷0.5) %/°C for crystalline Silicon modules. When referring to the adjustment coefficient C_A, this parameter could be close to unity, but a specific range cannot be defined at priori. In any case, the limits are imposed to avoid affecting the optimisation results.

V. Case Study

The group of PV plants under analysis includes plants installed in Lombardia, a region in northern Italy. A total of 125,483 plants are installed in this area: the sum of the rated powers of the whole cluster is ≈2.3 GW. Only 1% of these plants has power higher than 200 kW$_p$, but these generators reach about 70% of the total installed power. 85% of the plants have polycrystalline silicon modules, while 14% have m-Si modules. The remaining 1% is made of plants with amorphous silicon modules. The Italian Transmission System Operator (TSO) [24] provided the general data about the plants, i.e., coordinates of the installation sites, rated powers, and the installation year. The hourly production profiles are not available for all the plants; they are provided for a subgroup of 20,471 plants. A clustering procedure [22] is applied on the 125,483 values of rated power. The result is the definition of 10 classes of plants, as shown in Table I.

TABLE I. Classes of Plants and Related Power Ranges

Class	Power ranges (kW)	Number of plants
1	$P \leq 3.5$	49290
2	$3.5 < P \leq 6.5$	49605
3	$6.5 < P \leq 12.5$	7482
4	$12.5 < P \leq 25$	8118
5	$25 < P \leq 70$	4693
6	$70 < P \leq 120$	3316
7	$120 < P \leq 500$	2021
8	$500 < P \leq 1200$	696
9	$1200 < P \leq 3600$	54
10	$3600 < P$	9

In parallel to the creation of the classes, a data filtering procedure is applied on the 20,471 power profiles. Following the criteria defined in [21], wrong profiles are removed, and a sample of plants is selected for the optimisation of the parameters of the models. Due to the presence of a high number of measured profiles affected by issues (errors in data measurement, transmission, and storage, or poorly functioning generators) most of the profiles are removed. The sample of plants remaining for the optimisation is composed of 139 generators. The criteria are applied in a very restrictive way, by removing the profiles even in case of only one missing or incorrect data. Table II shows the distribution of the 139 plants in the classes (the classes that are not present in the table are empty):

TABLE II. Distribution of Plants after Filtering in the Classes

Class	4	5	6	7	8	9	10
Number of plants after filtering	1	10	26	25	66	9	2

Finally, to statistically validate the subgroups of 139 PV plants with respect the whole cluster of 125,483 plants, the stratified sampling procedure is applied, as defined in Section III. The result is that the statistical validation is confirmed mainly for the classes including the biggest PV plants (classes 6, 8, 9 and 10), as shown in Table III. The first consideration is that bigger PV plants have better maintenance and check of the energy production; thus, production profiles are adequate (and not removed during filtering of data). Secondly, the monitoring and data collection infrastructure have to be strongly improved, especially for small PV plants.

TABLE III. DISTRIBUTION OF PLANTS AFTER FILTERING IN THE CLASSES

Class	4	5	6	7	8	9	10
Min. number of plants for the statistical validation	5	13	10	51	31	9	2
Validation	no	no	yes	no	yes	yes	yes

A. Comparison of the temparature calculations

For every PV plant in the group, the cell temperature is calculated hour by hour by using the models presented in Section II. For the sake of simplicity, a single PV plant is used as an example. The selected PV plant is in class 9 with a nominal power of 3.2 MW. Table IV shows the average cell temperature calculated for winter and summer. The results are presented for the four models after the optimisation of their parameters.

TABLE IV. ANALYSIS OF THE TEMPERATURE MODELS

Model	Average module temperature in winter [°C]	Average module temperature in summer [°C]
NOCT - optm.	11.2	29.4
WIND model - optm.	8.7	30.1
MATTEI - optm.	10.2	27.2
KING - optm.	10.9	29.5

The optimised WIND model evaluates the highest PV cell temperature in the summer period, while, in winter it presents a low average value. This model underestimates the PV cell temperature, since the optimisation increases the dissipation parameter correlated to the wind speed, thus increasing heat removal. On the contrary, the optimised MATTEI model presents relatively low temperatures in the summer, and high in the winter. Finally, the optimised KING model calculates relatively high temperatures in every season (that are lower only than MATTEI values). To better understand the temperature trends calculated with the different models, two examples of sunny days in the summer are presented. The main difference is the wind speed: during the first day (June 1st), the wind speed is relatively high, ranging from 3.8 to 5.2 m/s, with an average value of 4.3 m/s.

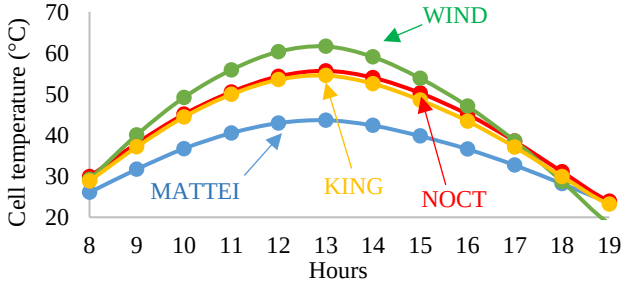

Fig. 1. Comparison of the cell temperatures calculated with the different models – day with high wind speed Jun, 1st

During the second day under analysis (June 16th), the wind speed is relatively low, ranging from 0.7 to 2.1 m/s, with an average value of 1.2 m/s. Calculated temperatures (hourly values) are presented in Fig. 1 and Fig. 2 for the two days, respectively.

The two graphs show that in every condition there are differences between the PV cell temperatures calculated with the four optimised models. In fact, the maximum deviation between the models reaches 15° C in both days. During the windiest day (June 1st), the WIND model calculates a temperature of ≈60°C, while the KING and NOCT models result in similar temperature ≈54°C. The MATTEI model has the lowest value corresponding to ≈43°C.

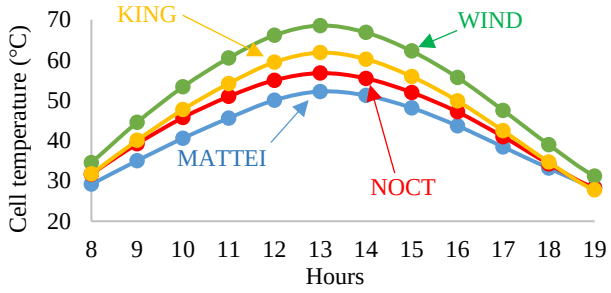

Fig. 2. Comparison of the cell temperatures caculated with the different models – day with low wind speed June 16th

During the least windy day (June 16st), the models follow a similar ranking; the main difference is that the KING and NOCT models have more different results with temperature deviation of ≈2°C at midday.

B. Comparison of the power profiles

Fig. 3 shows the power profiles calculated with the different models during the example day with high wind speed (June 1st). The profiles refer to the same plant used as an example in the previous sub-paragraph with rated power of 3.2 MW. To better show the power deviations, the figure zooms on powers higher than 1 MW; the deviations early in the morning and in the late afternoon are negligible. As visible in Fig. 3, the models underestimate the power production during the sunny days. The model with the best performance is KING with a deviation of -8% with respect the measurement at midday. The NOCT and MATTEI models have a similar deviation of -9.5%, while the WIND model calculates -13% less power.

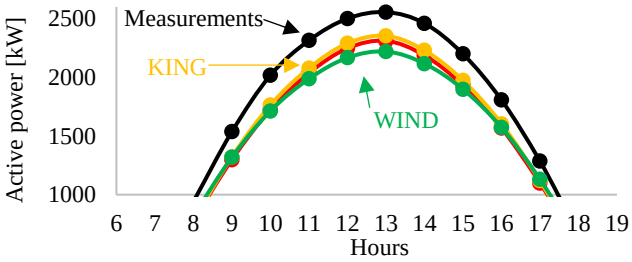

Fig. 3. Comparison of the power profiles calculated with the different models – day with high wind speed June 1st

The results obtained in the second example of day (June 16st) are similar, confirming a general underestimation in the summer production for all the models. The reason is that the optimisation process has the goal to decrease the deviation between measurements and calculate values during all the year (not only during summer), for the whole group of plants under analysis. Thus, energy deviations in daily trends for each single PV plant are acceptable and consistent with the goal of calculating the production of the whole cluster of PV plants.

Table IV shows the monthly energy deviations between the results from measurements and models without optimisation. The results refer to the whole group of PV plants. The results show that non-optimised models underestimate the PV production during all months of the year, with peaks in summer (deviations in the range -12÷-14%). During December the production can be very variable, and the models overestimate it with deviations between 11% and 13%. Nevertheless, the energy production in December is low with respect the yearly production; thus, the overestimations have negligible effect on the yearly results. The best model is MATTEI, with yearly underestimation of -8%, while KING and WIND models have -9%, and NOCT calculates -10% of production.

TABLE V. MONTLY ENERGY DEVIATIONS BETWEEN MEASUREMENTS AND MODELS RESULTS WITHOUT OPTIMISATION

	NOCT	WIND	MATTEI	KING
Jan	-7%	-6%	-6%	-6%
Feb	-5%	-5%	-4%	-4%
Mar	-6%	-5%	-4%	-5%
Apr	-12%	-10%	-9%	-10%
May	-11%	-10%	-9%	-10%
Jun	-13%	-11%	-10%	-11%
Jul	-14%	-12%	-12%	-13%
Aug	-14%	-12%	-11%	-12%
Sept	-11%	-9%	-9%	-10%

	NOCT	WIND	MATTEI	KING
Oct	-10%	-9%	-8%	-9%
Nov	-5%	-5%	-5%	-4%
Dec	11%	11%	12%	13%
Year	-10%	-9%	-8%	-9%

The optimisation of the parameters leads to significant improvements in all the models. Deviations decrease down to 1÷2% on an annual basis, with higher variations on a monthly basis. In particular, there is a generalised underestimation of the production in summer, compensated during winter. From June to August the underestimation is lower than -1%, while during winter the overestimations of the production are in the range 20÷25% in December. The best models are WIND and MATTEI: excluding December, the absolute values of the deviations are always less than 4% per month.

TABLE VI. MONTLY ENERGY DEVIATIONS BETWEEN MEASUREMENTS AND OPTIMISED MODELS

	NOCT	WIND	MATTEI	KING
Jan	5.1%	1.6%	1.8%	5.4%
Feb	5.2%	1.8%	2.6%	5.2%
Mar	5.0%	2.7%	3.9%	5.2%
Apr	0.1%	-0.5%	0.9%	0.4%
May	0.8%	-0.1%	0.8%	0.7%
Jun	-0.3%	-0.4%	0.7%	-0.3%
Jul	-0.7%	-1.1%	-0.1%	-1.0%
Aug	-0.4%	-0.7%	0.0%	-0.6%
Sept	1.6%	1.3%	2.1%	1.4%
Oct	1.9%	1.2%	1.4%	2.2%
Nov	8.5%	3.8%	3.1%	8.8%
Dec	24.6%	20.0%	21.6%	24.7%
Year	1.9%	0.9%	1.7%	1.9%

Table VII shows the optimised variables that are common for every model under analysis (e.g., the sub-vector x_1 containing the first part of the optimised parameters, as defined in Section IV). In every case, the thermal coefficient of power is higher than the default value equal to -0.4 %/°C. The reason is that the cumulative production includes PV plants with different installation conditions, such building integrated systems, in which the heat dissipation is lower than in the case of free-standing plants. The increased G_0 takes into account not only the losses due to low irradiance, but also all the phenomena that can occur when the PV production is low. Finally, the C_A parameter shows the most interesting change: it is higher than unity in every case. Thus, the original models underestimate the whole production of the cluster. The reason could be an overestimation of the losses due to the aging of the plants; this aspect will be studied in further works.

TABLE VII. OPTIMISED PARAMETERS FOR THE PV PRODUCTION MODELS

Model	$\gamma_{T\%}$	G_0	C_A
NOCT - optm.	-0.43%	28.0	1.09
WIND model - optm.	-0.44%	33.8	1.13
MATTEI - optm.	-0.42%	27.0	1.06
KING - optm.	-0.43%	34.9	1.12

Finally, the values for the other optimised parameters (different for each model) are the following. In the case of NOCT model, the optimal *NOCT* value is ≈47°C, while the initial value was 45°C. Regarding the wind model optimisation, the first parameter is a_T=0.994, b_T is equal to 0.040 (°C/W), c_T =1.524 (°C·s/m) and d_T=4.4 °C. All these values are higher than the default ones, demonstrating that the cell temperature was underestimated. In case of the MATTEI model, e_T=24.4, f_T=3.2, and $\alpha\tau$=0.82. Finally, in the KING model g_T=-3.19 (°C), h_T is equal to -0.058, and i_T=1.68 °(s/m). (DA CORREGGERE).

VI. CONCLUSIONS

The calculation of photovoltaic generation can be improved to better match the actual measured production from big groups of existing PV plants. The work presented in this paper demonstrates that the temperature models for the PV panels differ in the calculated profiles. Then, the energy models have deviations in the range -8÷10% on an annual basis. After the optimisation of the parameters of the analysed models, the calculation of the energy production improves, with deviations lower than 2% for the whole cluster of PV plants. On a monthly basis, deviations are ≤1% in summer, when there is most of the annual production. In some winter months the deviations can be still high

(≈20%) demonstrating that the models were initially created mainly for summer and clear sky days. Finally, the analysis of the data demonstrates that the actual monitoring infrastructure is not fully adequate. An improvement in the data quality collection will permit to repeat the proposed procedure with higher statistical validity for PV plants of any rated power.

REFERENCES

[1] S. Liu and M. Dong, "Quantitative research on impact of ambient temperature and module temperature on short-term photovoltaic power forecasting," In Proceedings of the 2016 International Conference on Smart Grid and Clean Energy Technologies (ICSGCE), Chengdu, China, 19–22 October 2016; pp. 262–266.

[2] 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. on Sust. Energy*, vol. 4, no. 2, pp. 278-285, 2013.

[3] F. Spertino, E. Chiodo, A. Ciocia, G. Malgaroli, and A. Ratclif, "Maintenance Activity, Reliability, Availability, and Related Energy Losses in Ten Operating Photovoltaic Systems up to 1.8 MW," *IEEE Transactions on Industry Applications*, vol. 57, no. 1, pp. 83-93, 2021.

[4] A. Ciocia, P. Di Leo, S. Fichera, F. Giordano, G. Malgaroli and F. Spertino, "A Novel Procedure to Adjust the Equivalent Circuit Parameters of Photovoltaic Modules under Shading," 2020 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM), 2020, pp. 711-715

[5] P. Di Leo, F. Spertino, S. Fichera, G. Malgaroli, and A. Ratclif, "Improvement of Self-Sufficiency for an Innovative Nearly Zero Energy Building by Photovoltaic Generators", Proc. IEEE PowerTech, Milano, Italy, 2019.

[6] A. Ciocia, J. Ahmad, G. Chicco, P. Di Leo, F. Spertino, "Optimal size of photovoltaic systems with storage for office and residential loads in the Italian net-billing scheme", Proc. of the 51st Intern. Universities Power Engineering Conference (UPEC), Coimbra, Portugal, 2016.

[7] S.B. Schujman, J.R. Mann, G. Dufresne, L.M. LaQue, C. Rice, J. Wax, D.J. Metacarpa, and P. Haldar, "Evaluation of protocols for temperature coefficient determination," Proc. IEEE 42nd Photovoltaic Specialist Conference (PVSC), New Orleans, LA, USA, June 2015.

[8] A. Carullo, A. Vallan, Outdoor experimental laboratory for longterm estimation of photovoltaic-plant performance, *IEEE Trans. on Instrum. and Meas.*, 61(5) (2012) pp. 1307-1314.

[9] Y.R. Golive et al., "Analysis of Field Degradation Rates Observed in All-India Survey of Photovoltaic Module Reliability 2018," *IEEE Journal of Photovoltaics*, vol. 10, no. 2, pp. 560-567, 2020.

[10] F. Spertino, A. Ciocia, F. Corona, P. Di Leo and F. Papandrea, "An experimental procedure to check the performance degradation on-site in grid-connected photovoltaic systems," 2014 IEEE 40th Photovoltaic Specialist Conference (PVSC), 2014, pp. 2600-2604.

[11] J.H. Bae, D.Y. Kim, J.W. Shin, S.E. Lee, and K.C. Kim, "Analysis on the Features of NOCT and NMOT Tests With Photovoltaic Module", *IEEE Access*, vol. 8, pp. 546-554, 2020.

[12] Schujman S. B.; Mann J. R.; Dufresne G.; LaQue L. M.; Rice C.; Wax J.; Metacarpa D. J.; Haldar P. Evaluation of protocols for temperature coefficient determination. In Proceedings of the 2015 IEEE 42nd Photovoltaic Specialist Conference (PVSC), New Orleans, LA, USA, 14–19 June 2015; pp. 1–4.

[13] G. TamizhMani, L. Ji, Y. Tang, L. Petacci and C. Osterwald, "Photovoltaic module thermal/wind performance: Long-term monitoring and model development for energy rating", Proc. NCPV Solar Program Rev. Meeting, pp. 936-939, 2003.

[14] M. Mattei, G. Notton, C. Cristofari, M. Muselli, P. Poggi, "Calculation of the polycrystalline PV module temperature using a simple method of energy balance," *Renewable Energy*, 2006, Vol. 31:4, pp. 553-567.

[15] Maria Brogren, Per Nostell, Björn Karlsson, Optical efficiency of a PV–thermal hybrid CPC module for high latitudes, *Solar Energy*, Vol. 69:6, 2001, pp. 173-185.

[16] Sandnes B, Rekstad J. A photovoltaic/thermal (PV/T) collector with a polymer absorber plate, experimental study and analytical model. *Solar Energy* 2002; Vol. 72(1), pp. 63-73.

[17] King, D.L., Boyson, W.E., Kratochvil,. Report: Photovoltaic Array Performance Model. Sandia National Laboratories, J.A., 2004. Available online: https://www.osti.gov/servlets/purl/919131sca5ep/, (accessed on Jan 2022).

[18] Photovoltaic Geographical Information System (PVGIS). Available online: https://ec.europa.eu/jrc/en/pvgis (accessed on Jan 2022).

[19] SODA Solar Radiation Data. Available online: http://www.soda-pro.com (accessed on 29 Jan 2022).

[20] Google Street View, Available online (Accessed on Jan 2022), https://en.wikipedia.org/wiki/Google_Street_View.

[21] Carullo, A.; Castellana, A.; Vallan, A.; Ciocia, A.; Spertino, F. In-field monitoring of eight photovoltaic plants: Degradation rate over seven years of continuous operation. *Acta IMEKO* 2018, *7*, 75–81.

[22] G. Alba, G. Chicco, A. Ciocia and F. Spertino, "Statistical Validation and Power Modelling of Hourly Profiles for a Large-Scale Photovoltaic Plant Portfolio," 2021 IEEE 6th International Forum on Research and Technology for Society and Industry (RTSI), 2021, pp. 18-23.

[23] J. Neyman, "On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection," *Journal of the Royal Statistical Society*, vol. 97, no. 4, pp. 558–625, 1934.

[24] Terna, Italian Transmission System Operator, Available online: www.terna.it/en (accessed on 29 Jan 2022)