Monitoring performance and efficiency of photovoltaic parks

Federico Bizzarri [a], Angelo Brambilla [a, *], Lorenzo Caretta [b], Carlo Guardiani [b]

[a] Politecnico di Milano, Dipartimento di Elettronica, Informazione e Bioingegneria, p.za. Leonardo da Vinci, n. 32, I20133 Milano, Italy
[b] Renience Srl, v.le G. Marconi 52, Desenzano del Garda, Italy

Received 28 August 2014
Accepted 2 January 2015

1. Introduction

In the last decade the interest in electrical energy production from renewable sources has largely increased. In many countries it has reached a significant fraction of the total energy produced by non-renewable sources [1]. Since in many countries the target renewable vs. non-renewable ratio was reached, public contributions are now being withdrawn and the cost per kWh produced by new solar plants must significantly decrease thus becoming competitive with that by non-renewable sources. Photovoltaic (PV) plants will no longer be allowed to work far from their optimal efficiency condition since the economic baseline gain often imposes a plant efficiency above 90% [2,3].

A typical industrial PV park is composed of a large array of PV panels, usually mounted on static support frames (i.e. not considering solar trackers) with no moving mechanical parts. This may lead to the wrong conclusion that maintenance requirements for such simple system is very limited and to expectations of almost constant level of energy production for several years without the need of system monitoring. Actually this is very far from being true and it is now common practice to plan and implement a significant amount of resources for the monitoring and maintenance of industrial-strength PV parks. From this there is the need for more efficient, accurate and reliable methodologies and tools to monitor and rapidly identify potential failures with the goal of optimizing energy production and minimizing maintenance costs [4–6].

It is possible to define two types of failures – *catastrophic* or *hard* ones such as those due to blown fuses, broken cables or disconnections from the power-grid caused by inverter failures, etc, and – *soft* ones such as those caused by dust or dirt on the surface of the PV panels, hot-spot formation, partial shading, etc. In general, hard failures lead to drastic and fast variation of the electrical variables. Usually monitoring of hard failures is trivial since they can be easily identified; on the contrary, detecting and diagnosing soft failures may be very difficult. Their effect is often mistaken for the daily power production fluctuation or it builds up very slowly, thus gradually decreasing the efficiency of the PV plant without rising any red flag.

A key objective of PV plant monitoring is the generation of alarm messages that are automatically sent to the O&M (operation and maintenance) engineers. It is very important to minimize the generation of false alarms that may make the monitoring system useless since human processing is needed to select those that need attention. Direct raw data comparison logged from the PV park may not always be the best approach. It is often better to introduce suitable post-processing functions, such as figure of merits (FOMs) and aggregate efficiency indices, that yield indirect measures of the actual performance of the PV plant and trigger alarms in case of

* Corresponding author.
E-mail addresses: federico.bizzarri@polimi.it (F. Bizzarri), angelo.brambilla@polimi.it (A. Brambilla), lorenzo.caretta@renience.com (L. Caretta), carlo.guardiani@renience.com (C. Guardiani).

malfunction. In this paper we exclusively consider PV parks (i.e., industrial PV plants) providing an accurate description of them and of the FOMs and aggregate indices used in the proposed monitoring system.

The definition of a model suitable to be used for establishing a reference baseline for PV power production was tackled in previous works [7–11], that mostly focused on the definition of analytical models of power production and efficiency of a PV panel as functions of irradiance and module temperature. These models are not based on the physics of the pn junction that forms the basis of c-Si (Crystalline Silicon) PV cells, do not model the electro-thermal characteristic of the PV panels and often do not take into account the layout of electrical interconnections. For instance the PV model described in Ref. [11] and used in the monitoring system presented in Ref. [12] links the maximum power point (mpp) to temperature and irradiance but does not consider electrical variables. This implies that all PVs panels are assumed working with the same branch voltage and that the maximum power point tracker (MPPT) always performs in the best way, i.e., keeping all strings working at their mpp even if they show mismatch in the electrical characteristics.

In this paper we present a completely different approach to implement a large, industrial-strength PV parks monitoring system leveraging on the simulation model of PV parks described in Refs. [13,14]. The electrical and thermal equations governing the behaviour of panels and the resistance of interconnecting cables derived from the layout of the actual PV park are considered, thus leading to a compact and numerically efficient model that is capable to account for the interdependency of the thermal and electrical equations. This accurate physical electro-thermal model was implemented in a simulation environment also including behavioural modelling language capabilities, which are very useful e.g. to account for the specific MPPT algorithm that is characteristic of the specific inverter used in the PV park. Furthermore this model is capable of coping with often neglected effects, such as the ohmic losses in cables and suboptimal operating conditions e.g. in the case of MPPT flaws. The proposed approach allows to easily and directly compare measured and simulated electrical data, such as for example DC/AC inverter power and string currents and the simulated ones. This is a unique and novel feature of the proposed methodology.

A prototype implementation of this monitoring system has been used for several months to track the performance of 10 PV parks, both rooftop and ground mounted, with power capabilities ranging from 100 kW to 1 MW and installed in several different locations in Italy, for a total of 8.6 MW of nominal electrical power. Some application examples that exploit the already available historical data set are used in this paper as case studies to prove the effectiveness of the proposed approach. Since the proposed methodology requires the definition of a significant number of constants, parameters and variables, a reference for the most important of them is collected in Table 1.

Table 1
Symbol reference.

Plant data	
P_m	Measures instantaneous power of the park
P_s	Simulated instantaneous power of the park
P_{mo}	Power by the ideal park at standard test conditions (STC)
P_{mpp}	Power at the mpp working condition
i_{sc}	Panel short circuit current
v_{oc}	Panel open circuit voltage
v_{mpp}	Panel voltage at the MPP working condition
i_{mpp}	Panel current at the MPP working condition
Environmental data	
T_{amb}	Ambient temperature
T_{stc}	STC temperature (298.15 K [15])
T_{noct}	Normal operating cell temperature (NOCT)
T_{pv}	Surface temperature of panels
R_m	Instantaneous solar irradiance
R_{sc}	Irradiance at STC (1 kWm^{-2})
R_{noct}	Irradiance at NOCT (800 W/m^2)
R_e	Theoretical irradiance on the panel surface at STC reduced from 1367 Wm^{-2} to 1 kWm^{-2} [16]
t_r	Sunrise time
t_s	Sunset time
β	Temperature coefficient at P_{mpp} with panel operating at STC
ν	Wind speed (ms^{-1})
ν_0	Nominal wind speed (1 ms^{-1})
h	Convection coefficient (6.62) [17,18]

2. Definition of FOM metrics

Our goal is to define a set of FOMs that can be successfully and reliably used to trigger alarms in case of PV system failures, both hard and soft. An ideal set of FOM metrics should be necessary and sufficient, in the sense that it should never create false alarm while they should always set off when the system has a problem which requires attention. In our systems FOMs are defined on a per-field basis, having in mind that PV parks are usually composed of several PV fields.[1]

In the sequel, grounding on the well known *performance ratio* FOM, an effective and reliable aggregate index named $\mathscr{T}_4$, that is used to trigger alarms, is defined following an incremental approach, i.e., going through the definition of the $\mathscr{T}_1$, $\mathscr{T}_2$ and $\mathscr{T}_3$ FOMs.

2.1. $\mathscr{T}_1$: a first improvement of the performance ratio

The first and perhaps best known FOM is the performance ratio [19].

$$\mathscr{T}_0(t) = \mathscr{P}_r(t) = \frac{P_m(t)}{R_m(t)} \frac{R_{sc}}{P_{mo}}. \tag{1}$$

The metric $\mathscr{P}_r$[2] is widely adopted primarily for its simplicity since it requires only a few and readily available measured data as input. It suffers of some serious drawbacks limiting its effectiveness to identify the efficiency of a PV field. In fact the power yield of panels largely depends on their working temperature which in turn depends on irradiance and convection coefficients (i.e. the parameters that characterize the rate of heat transfer from the panel to the environment). $\mathscr{P}_r(t)$ can thus be improved as

$$\mathscr{T}_1(t) = \frac{P_m(t)}{R_m(t)} \frac{R_{sc}}{P_{mo} + \beta\big(T_{\mathrm{pv}}(t) - T_{\mathrm{stc}}\big)}. \tag{2}$$

When the surface temperature of panels is not available, since for example sensors are not installed, according to [17] – by considering also wind velocity – we can derive $T_{\mathrm{pv}}(t)$ through

$$T_{\mathrm{pv}}(t) = T_{\mathrm{amb}}(t) + \frac{R_m(t)(T_{\mathrm{noct}} - T_{\mathrm{stc}})}{R_{\mathrm{noct}} + h(\nu(t) - \nu_o)(T_{\mathrm{noct}} - T_{\mathrm{stc}})}. \tag{3}$$

Note that T_{noct} is measured with the panel working in open circuit conditions [2,20–23]. If the anemometer is not available, h

[1] A "PV field" is defined as a subset of PV strings connected to the same inverter.

[2] Actually the dependence of $\mathscr{T}_0$ on time t is not explicit and one would better write $\mathscr{T}_0(P_m(t), R_m(t))$. This is true even for the other FOMs defined in the sequel. Whenever the clarity of the presentation is not jeopardized, we omit the dependence of a given FOM on specific quantities, such as $P_m(t)$ and $R_m(t)$ for $\mathscr{T}_0$, but simply specify its time-varying nature.

has to be set to 0. The suggestion to enhance $\mathscr{P}_r(t)$ is not novel. Something similar was done in Ref. [24] and in Ref. [8] the authors derived an expression for the power generated by the PV panel depending on irradiance and *module temperature* (modelled as in Ref. [25]) which we assume to be the panel surface temperature.

By recalling that our goal is to use the FOMs to automatically trigger alarms in case of performance drops, we report in Fig. 1 the $\mathscr{P}_r(t)$ performance ratio and $\mathscr{F}_1(t)$, evaluated during a sunny day of one of the two fields composing the PV park that we refer to as $\mathscr{S}1$ in the following. $\mathscr{S}1$ delivers 777 kW total power at STC. For completeness we report also $T_{\mathrm{pv}}(t)$ and $R_m(t)$. All the panels used in $\mathscr{S}1$ are `STP185S-24/Ad` type, with $\beta = -0.0048$ Wk^{-1}. As it can be seen $\mathscr{P}_r(t)$ shows a minimum of 76.2% at 330.95 K maximum surface temperature of panels, almost perfectly aligning with the 1080 Wm^{-2} irradiance peak. If we consider $\mathscr{F}_1(t)$ we see that it shows a minimum of 92.6% and a maximum of 99%. This clearly shows that $\mathscr{P}_r(t)$ underestimates the PV field efficiency. This is mainly due to the fact that $\mathscr{P}_r(t)$ does not depend on working temperature of panels; by observing Fig. 1 we see that the minimum of $\mathscr{P}_r(t)$ is located about at noon when there is the temperature peak. At the same time this data also show that the daily variation of both $\mathscr{P}_r(t)$ and $\mathscr{F}_1(t)$ (even though slightly lesser than that of $\mathscr{P}_r(t)$) can prevent their use for detecting PV field soft failures.

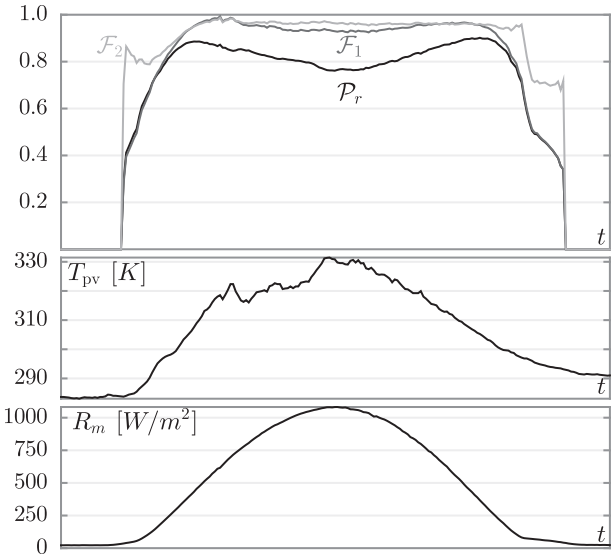

Fig. 1. (Upper panel) The $\mathscr{P}_r(t)$ (black), $\mathscr{F}_1(t)$ (grey) and $\mathscr{F}_2(t)$ (light grey) FOMs computed during a working day of the $\mathscr{S}1$ park delivering 777 kW total nominal power at STC. To allow comparison between these FOMs we report also surface temperature of panels (centre panel) and irradiance (lower panel).

2.2. $\mathscr{F}_2$: the relative power error

Although $\mathscr{F}_1(t)$ can be easily determined through Eq. (3) even if surface temperature of panels is not available, an important accuracy improvement is offered by electro-thermal simulation of the PV park to derive the simulated counterpart electrical quantities. As a starting point we can take as reference the $P_s(t)$ instantaneous power produced by the simulated PV field, that works with the same environmental conditions of the real one but with ideal panels, i.e. not degraded by ageing, dust or formation of hots-spots, ideal MPPT and inverter. We thus introduce the simple but effective

$$\mathscr{F}_2(t) = \frac{P_m(t)}{P_s(t)}$$

relative power error FOM. To give a quick idea on the advantages of having the reference frame by simulation, $\mathscr{F}_2(t)$ referring to the same working day of the $\mathscr{S}1$ park is shown in Fig. 1. As we can see it is more "regular" exhibiting an almost constant value of 94% during the portion of the day with high irradiance level, i.e. above 300 Wm^{-2}, and a less relevant drop at irradiance levels below this value. As it can be seen $\mathscr{F}_1(t)$ is in good agreement with $\mathscr{F}_2(t)$ at relatively high irradiance that corresponds to high energy production.[3]

The $\mathscr{F}_2(t)$ FOM gives a per-sample measure of the field efficiency. In general one sample is recorded every few minutes, therefore this FOM can be difficultly used to trigger alarms on possible malfunctions of the field. In fact a time-local large fluctuation, that does not practically cause any appreciable variation of the total energy produced in the working day, can in turn trigger an alarm. In general these possible time-local large fluctuations are due to the fact that.

- the irradiance sensors are located in a specific point of the PV park, for example near the inverter room. This means that scattered clouds can cover a large number of panels without covering the irradiance sensors and vice versa. This causes erroneous fluctuations of $\mathscr{F}_2(t)$;
- data-loggers get data from the PV field often by adopting a "polling" schema thus misaligning the time instants at which the instantaneous power, irradiance and panel surface temperature are recorded. Despite the application of interpolating functions this can result in erroneous fluctuations of $\mathscr{F}_2(t)$ along some recording points.

[3] It will be discussed in the sequel why the $\mathscr{F}_2(t)$ drops are limited at sunrise and sunset, i.e. when the inverter is driven by low power.

2.3. $\mathscr{F}_3$: the integral version $\mathscr{F}_2$

Drawbacks of $\mathscr{F}_2$ suggest to low pass filter it. We thus introduce the *discrete time* relative energy error function

$$\mathscr{F}_3(n) = \begin{cases} \alpha_m & \text{if} \quad \widetilde{\mathscr{F}}_3(n) < \alpha_m \\ \alpha_M & \text{if} \quad \widetilde{\mathscr{F}}_3(n) > \alpha_M \\ \widetilde{\mathscr{F}}_3(n) & \text{elsewhere} \end{cases}, \tag{4}$$

where

$$\widetilde{\mathscr{F}}_3(n) = \frac{\int_{t_r(n)}^{t_s(n)} P_m(t)dt}{\int_{t_r(n)}^{t_s(n)} P_s(t)dt}, \tag{5}$$

$\alpha_m \in \mathrm{R}_{>0}$ lower limits $\mathscr{F}_3(n)$, $\alpha_m < \alpha_M \in R^+$ upper limits $\mathscr{F}_3(n)$, $n \in \mathrm{N}$ is the day index. $t_r(n)$ and $t_s(n)$ can be determined by using for example the SOLPOS tool [16]. In general we choose $\alpha_m = 0.1$ and $\alpha_M = 1.2$, since values of $\widetilde{\mathscr{F}}_3(n)$ even less than 0.1 already give evidence of malfunctions and above 1.2 show that there is something wrong in data sampling of $P_m(t)$ and/or $P_s(t)$. $\mathscr{F}_3(n)$ is thus undersampled with respect to $\mathscr{F}_2(t)$ since it generates one sample per day.

By limiting the integrals in Eq. (4) from $t_r(n)$ to $t_s(n)$ we prevent the computation of $\mathscr{F}_3(n)$ during night. Apparently this seems useless since PV parks do not produce power in the night time and integral contributions are null. However we often experienced the recording of small radiation intensity in the night due to small

offset of the irradiance sensors and/or night illuminations by city lamps. If this happens the FOM is incorrect since there is an excess of integrated $P_s(t)$.

The $\mathscr{F}_3(n)$ FOM does not derive the efficiency of the field independently from the irradiance intensity. More in detail, consider efficiency in two very different days, for example the first one sunny with peak irradiance above 1 kWm^{-2} and the second one cloudy with peak irradiance below a few hundreds of Wm^{-2}. In these two days energy productions are very different and small absolute differences between simulated and measured instantaneous power can lead to large drops of $\mathscr{F}_3(n)$ that can be mistaken for PV field malfunctions.

2.4. The $\mathscr{F}_4$ aggregate index

We suggest to modify $\mathscr{F}_3(n)$ by weighting efficiency according to irradiance, obtaining the *weighted relative energy error*

$$\mathscr{F}_4(P_s(t),P_m(t),n) = 1 - \frac{\int_{t_r(n)}^{t_s(n)} w(t)(P_s(t)-P_m(t))dt}{\int_{t_r(n)}^{t_s(n)} P_s(t)dt}, \tag{6}$$

where $w(t)$ is the weighting function defined as

$$w(t) = \begin{cases} \beta_m & \text{if} \quad \widetilde{w}(t) < \beta_m \\ \beta_M & \text{if} \quad \widetilde{w}(t) > \beta_M \\ \widetilde{w}(t) & \text{elsewhere} \end{cases}, \tag{7}$$

$\widetilde{w}(t) = R_m(t)/R_e(t)$, β_m and β_M have the same meaning of α_m and α_M in Eq. (4). The $R_e(t)$ irradiance can be determined by considering the tilting and azimuth of each panel and the geographical coordinates of the field [16,26]. The $\widetilde{w}(t)$ function scales the difference between the simulated and measured powers according to the irradiance intensity. Above the reduced theoretical irradiance this difference is enhanced and below reduced. We choose $\beta_m = 0.1$ and $\beta_M = 1.2$ since "attenuations" or "amplifications" by $w(t)$ lesser and larger than these values, respectively, are meaningless.

In sunny days, for a good performing park, $\mathscr{F}_4(P_s(t),P_m(t),n)$ is very close to $\mathscr{F}_3(n)$ since $w(t)$ acts only at low irradiance levels. In case of a badly performing park, differences in simulated and measured instantaneous power are amplified and this leads to differences between $\mathscr{F}_4$ and $\mathscr{F}_3$. During cloudy days $\mathscr{F}_4$ can be larger than $\mathscr{F}_3$ since $w(t)$ attenuates possible differences between measured and simulated powers. The key aspect is that $\mathscr{F}_4$ is an aggregate index and not a FOM and gives more reliable results on a per day basis and it was used to check possible malfunctions and to trigger automatic messages.

By evaluating $\mathscr{F}_4(I_{s,k}(t),I_{m,k}(t),n)$, where $I_{s,k}(t)$ and $I_{m,k}(t)$ are the simulated and measured currents, respectively, of the k-th string, it is also possible to detect malfunctions of subsections in the PV field. This aggregate index is applied to each string current to compute the corresponding efficiency.[4] A possible drop of efficiency of a single string can be rarely detected by observing the total power generated by the field, since the relative drop of $\mathscr{F}_4$ is often marginal. Therefore application of $\mathscr{F}_4$ to string currents is an effective tool to detect both global (as it will be shown in Sec. 4) and local malfunctions.

[4] The current of each string is measured and recorded by data loggers and it is acceptable to consider currents and not string powers since the field voltage is shared among all string that are connected in parallel as shown in the schematic of Fig. 2.

3. Working and monitoring flows

3.1. Data-base organisation

The working data from the PV park are organised in a data-base (DB), that also stores the following items.

1. The layout of each field, i.e. interconnections of panels to form strings and connections of strings in the junction boxes as shown in Fig. 2. In the schematic reported in Fig. 2, the Rc1…Rcn and Rm resistors model resistances of interconnecting cables. Blocks labelled as PV models groups of PV cells (in general 20) that form panels that in turn are connected in series to form strings. Each of these blocks has a bypass diode that allows current to flow through the rest of the string if a PV is damaged. Each string is connected through the Dcn diode and Rcn resistor modelling losses in interconnecting cables to the MPPT block that acts on the entire PV field. The Dc1…Dcn diodes prevent currents to flow back to the corresponding strings, i.e. avoid that a malfunctioning string becomes a load of the PV field and dissipates power. The block labeled as MPPT implements an ideal MPPT that sets voltage at node M to maximize power delivered by the PV field. The resulting topological structure is a tree with panels as leaves and the PV park as root. In the first level nodes there are the fields and thus the inverters. Each inverter can be equipped with more than one MPPT.
2. Available sensors, types of the irradiance sensors, of the panel surface temperature sensors and of the ambient temperature ones.
3. Type and orientation of each panel; each panel is logically linked to a subset of sensors.
4. Types, sections and length of connecting cables.
5. The nominal electrical characteristics of each panel of the PV park as reported in the data-sheets, i.e., i_{sc}, v_{oc}, P_{mpp}, v_{mpp} and i_{mpp} (see Table 1). These values are grouped in two sets, one related to STC and the other to NOCT (e.g, P^{stc}_{mpp} and P^{noct}_{mpp}). Often a unique panel type is used for the entire PV park. If flash tests of panels are available, they are inserted and related to the corresponding panels.
6. Data logged from the PV park; these data can be environmental ones, such as for example the working temperature of the inverter transistors and CPU, and electrical variables, such as the instantaneous DC power from the PV park, the voltage level set by the MPPT and the AC power delivered to the grid. In general in conventional application loggers elaborate data and for example store irradiance values averaged over 30 min. In our case we do not perform any averaging and data are stored every 5 min.

3.2. Model fitting

Since we use the simulation model presented in Refs. [13,14], we have to fit model parameters to monitor the efficiency and yield of a

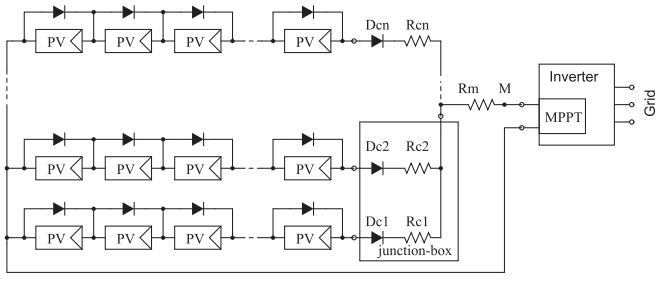

Fig. 2. Block diagram schematic of the electrical layout of a considered PV field.

PV park. The electro-thermal characteristic of panels are fitted in STC and NOCT. The use of these two characteristics at different temperatures allows the fitting of parameters that define the thermal behaviour of the panel. The fitting process is off-line and performed once-and-for-all during the start-up phase of the PV park. The electrical characteristics of panels are read from the DB; fitted model parameters are stored in the DB after fitting.

We perform a constrained optimisation [27] of the electro-thermal characteristics to minimize the error among the values of the given panel parameters listed in item 5 of Sec. 3.1 and those from the model. The objective function is

$$\begin{aligned} y(x) = {} & \alpha_1 \left| P^{\text{noct}}_{\text{mpp}} - P_{\text{mpp}}(x) \right| + \alpha_2 \left| v^{\text{noct}}_{\text{mpp}} - v_{\text{mpp}}(x) \right| + \alpha_3 \left| v^{\text{noct}}_{\text{oc}} \right. \\ & \left. - v_{\text{oc}}(x) \right| + \alpha_4 \left| i^{\text{noct}}_{\text{sc}} - i_{\text{sc}}(x) \right| + \alpha_5 \left| P^{\text{stc}}_{\text{mpp}} - P_{\text{mpp}}(x) \right| \\ & + \alpha_6 \left| v^{\text{stc}}_{\text{mpp}} - v_{\text{mpp}}(x) \right| + \alpha_7 \left| v^{\text{stc}}_{\text{oc}} - v_{\text{oc}}(x) \right| + \alpha_8 \left| i^{\text{stc}}_{\text{sc}} \right. \\ & \left. - i_{\text{sc}}(x) \right|, \end{aligned} \tag{8}$$

where $x \in \mathrm{R}^Q$ is the vector of the Q model parameters and $\alpha_1,\ldots,\alpha_8 \in \mathrm{R}^+$ are weighting coefficients. As an application item consider the STP185S-24/AD panel by SUNTECH, which equips $\mathscr{S}1$. Its electrical characteristics at STC and NOCT are reported in Table 2 where it can be appreciated that fitting leads to very good matching of parameter values, with a maximum relative error within the manufacturing tolerances of the panel.

3.3. PV park simulation

The PV park simulation is organised in three different steps described in the following. These three actions are automatically daily performed starting at a specific time instant (in our case at midnight). All simulations were performed with our PAN circuit simulator [32].

On average this simulation flow elaborates 6 samples per second on an I5-2400@3.10 GHz computer running LINUX; a sample is acquired from the PV parks every 5 min, therefore the simulation flow elaborates 30 min working per second.

3.3.1. The first step

In the first step the netlister reads the electrical layout of the PV park from the DB together with the electro-thermal models of panels. It generates a hierarchical netlist organised as the tree describing the layout of the PV park. The netlister also handles parasitic resistances of interconnections, the thermal network and the auxiliary code to elaborate results and load them in the DB at the end of the simulation. The auxiliary code implements also the MPPT feature through the "amoeba" optimisation method. In practice an objective function expressing the power transferred from the PV field to the inverter is maximised [27].

Table 2
Fitted electrical characteristic of the STP185S-24/AD panel.

Par.	Fit.	Value	Rel. err %
$v^{\text{stc}}_{\text{mpp}}$	35.487	36.400	−2.50
$P^{\text{stc}}_{\text{mpp}}$	185.000	185.000	0.00107
$v^{\text{stc}}_{\text{oc}}$	44.171	45.000	−1.84
$i^{\text{stc}}_{\text{sc}}$	5.431	5.430	0.03
$v^{\text{noct}}_{\text{mpp}}$	33.414	33.200	0.64
$P^{\text{noct}}_{\text{mpp}}$	138.738	137.000	1.26
$v^{\text{noct}}_{\text{oc}}$	41.297	41.300	−0.0049
$i^{\text{noct}}_{\text{sc}}$	4.345	4.390	1.01

Par.: parameter name, Fit.: fitted value of the parameter, Value: parameter value from the data-sheet, Rel. err: relative error of fitted parameter.

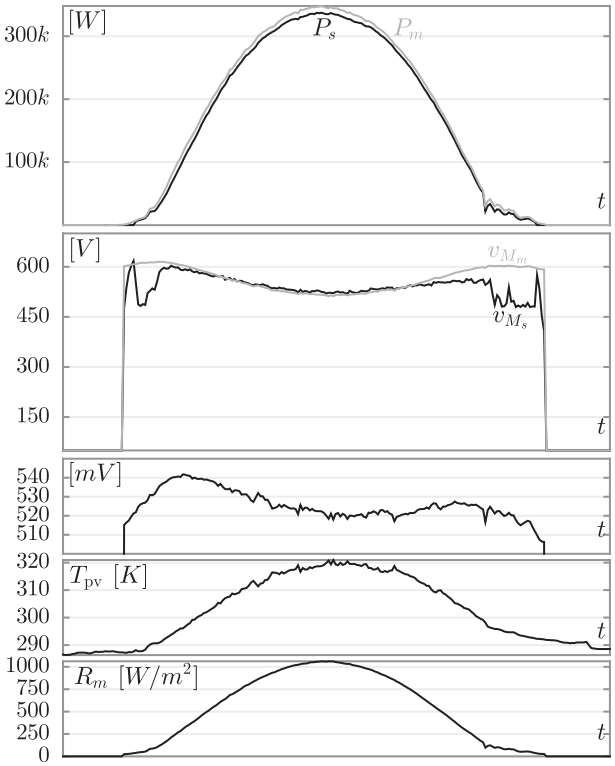

Fig. 3. From top to bottom. Panel 1: simulated (black) and measured (grey) powers along one working day of $\mathscr{S}1$. Panel 2: simulated (black) and measured (grey) voltages at the M node of the circuit shown in Fig. 2. Panel 3: simulated voltage across a single cell in a PV panel. Voltage fluctuations are mainly due to temperature variations and minimally due to irradiance variation. Panel 4: measured surface temperature of panels. Panel 5: measured irradiance.

3.3.2. The second step

In the second step the simulator is launched; it reads the netlist, accesses the PV park DB and retrieves data. More precisely the simulator retrieves the vectors of measured irradiance and measured panel surface temperature. In general panel temperature is referred to specific (one or two) panels of the field on which the temperature sensors are mounted. In industrial parks temperature can be considered uniform with good accuracy; clouds lead to non uniformity of temperature that however on average can be neglected. Irradiance is read from one or two sensors and the same considerations about measuring of temperature hold. The working temperature of the PV cell junction is determined by exploiting the electro-thermal characteristic. This differentiate the proposed approach from other ones that derive it by considering for example irradiance, wind speed and convection coefficients [28,29,12]. This has also the benefit to take in consideration possible overheating/cooling of panels when they are mounted for example on the roofs of industrial buildings. The simulator performs an electro-thermal *static* analysis on the discrete time mesh of measured input data and computes power produced by the simulated PV park together with current flowing through each string of the PV park.[5] The

[5] With the term static we underline that a set of algebraic equations and not differential ones is solved [30], i.e. electrical capacitors and inductors together with lumped thermal capacitors are neglected. This assumption makes sense since (i) the time constants of the thermal networks modelling the PV cell are smaller than the sampling time interval and (ii) the surface temperature of panels, whose dynamics can be extremely slow, is not simulated. The surface temperature is considered as a time-varying input of the reference frame; it is measured by the sensors and provided to the simulator.

"amoeba" optimisation method is launched at each time sample; it performs several DC analyses of the PV field and varies the field working voltage till reaching the maximum power transfer to the inverter. This action is repeated at each time sample. Simulation results are uploaded in the DB.

As an example, in Fig. 3 we reports some simulated and measured electrical quantities along one working day of one field of $\mathscr{S}1$.[6] In this simulation the characteristic of panels is assumed identical (no flash tests were available) therefore we report one simulated string current that constitutes the reference. The voltage is set by the MPPT, the string currents are thus determined by their i–v characteristics. In the simulated case the MPPT is *ideal* and it finds the unique maximum of the power versus voltage characteristic. As it can be noted there is a very good agreement between the simulated and measures instantaneous powers. Fig. 3 shows also a good matching between v_{M_m} and v_{M_s} (i.e., respectively, the measured and simulated field voltages set by the MPPT at node M in the circuit of Fig. 2) but at sunrise and sunset, i.e. at lower irradiance values where the *real* MPPT sets a too lower field voltage with respect to the optimal one by simulation. Making reference to Fig. 1, this is the reason of the fast drops of $\mathscr{P}_r(t)$ and $\mathscr{T}_1(t)$ at sunrise and sunset and of the best behaviour of $\mathscr{T}_2(t)$ which is computed using the ideal MPPT for the evaluation of $\mathscr{P}_s(t)$.

Furthermore, the field voltage lowers as irradiance increases and reaches its minimum at the maximum value of irradiance, that is, it shows a concave parabola shape. Examining Fig. 3 it can be noticed that such voltage remains almost constant versus irradiance and PV cell temperature variations showing fluctuations of ±20 mV around its average value of about 525 mV. In an approximated way, the photo-generated current can be assumed as linearly dependent on the irradiance and the PV cell voltage as constant when irradiance is above a lower threshold. The losses in cables and metal stripes connecting cells quadratically increase with respect to field current, i.e. with respect to irradiance. Voltage drop at node M is then due to power losses of interconnects, i.e. the MPPT has to lower voltage at the *M* node in the schematic shown in Fig. 2 to compensate the voltage drop across Rc1, Rc2, Rcn to extract the maximum power. In this case we computed a peak power loss by interconnections equal to 17.6 kW versus a corresponding 337 kW measured peak power. The relative energy loss is 3.1% versus 10.4 MJ of daily produced energy.

A straightforward economic analysis showed the need to lower power losses of cables by increasing their sections in next solar park realizations.

This shows the versatility and flexibility of the proposed monitoring approach and tool that allows us to easily simulate and elaborate electrical quantities.

3.3.3. *The third step*

In the third step of the monitoring flow the performance analyzer reads experimental and simulated data, such as power and string currents and computes the FOMs of the PV park thus providing some overall indices of the efficiency and yield. Results by the performance analyzer are uploaded in the DB.

4. Experimental results

In this section we present and discuss some results obtained by applying $\mathscr{T}_4(\cdot,\cdot,n)$ to detect malfunctions of a set of industrial pv parks located in Italy. All results refer mainly to the 2013 year and apply on historical data stored in the data-base.

[6] Each of the figures in this article refers, in general, to different time intervals but the panels, which make up each figure, refer to the same time period.

In Fig. 4 we report $\mathscr{T}_4(P_s(t),P_m(t),n)$ of one field of the $\mathscr{S}1$ PV park along several working days (upper panel) and the measured, simulated powers (lower panel). This field runs in an acceptable way with efficiency close to 90%. It can be also appreciated that $\mathscr{T}_4(P_s(t),P_m(t),n)$ has limited fluctuations when there is a large variation of delivered power due to irradiance drop in cloudy days. These days can be identified in the lower panel of Fig. 5 since they correspond to fast variations of the instantaneous power with respect to *t*.

A field of this park is composed of 140 strings, the application of $\mathscr{T}_4(\cdot,\cdot,n)$ to string currents showed that 137 strings were running correctly and 3 were unconnected due to fuse interruption. This was an "abrupt" failure easily detected in one day that reduced the total efficiency of about 2.5%. This estimation was performed by integrating the simulated energy produced by the reference string in the simulation model. It was decided not to plan an extra maintenance intervention and to wait the scheduled one to restore string functionality.

The second considered park ($\mathscr{S}2$) is locate near $\mathscr{S}1$ and was designed to deliver 510 kW at STC (two fields). It was equipped with the same panel type as $\mathscr{S}1$. In Fig. 5 we report $\mathscr{T}_4(P_s(t),P_m(t),n)$ and the measured, simulated powers along some working days. Fig. 5 shows a well known and easily identifiable efficiency drop; during the fourth reported working day, the measured power drops to 0 due to a disconnection from the grid. Reconnection takes place in the late afternoon. In this day $\mathscr{T}_4(P_s(t),P_m(t),n)$ drops to about 60%. During the reported days there were some scattered clouds that caused variations of power but not of $\mathscr{T}_4(P_s(t),P_m(t),n)$ that kept at about 90%.

In Figs. 6 and 7 we still consider $\mathscr{S}2$ and report the same quantities as in Fig. 5 but in subsequent working intervals. By observing $\mathscr{T}_4(P_s(t),P_m(t),n)$ in Fig. 6, we see that it lowers from day to day and that on the right portion of the figure, it goes below 80%. If we compare the measured (black thick curve) and simulated (grey thin curve) powers in the lower panel of Fig. 6, we see that there is an appreciable difference between their peak values and mainly that the measured power does not go above 200 kW peak value. This drawback is much more evident in the right portion of Fig. 7 where peaks of simulated power further increased with respect to those of Fig. 6 due to increasing seasonal irradiance, while measured power was evidently clipped. Efficiency continued to drop and reached almost 65% (see the right portion of Fig. 7).

We applied $\mathscr{T}_4(I_s(t),I_{m,k}(t),n)$ (string current aggregate index), where $k=1,\ldots,98$ (see Fig. 8). We see that all the strings of the field performed in a bad way with low index values that varied from about 45% to 65%. By observing string currents we see that they overlapped the reference current in a very good way from sunrise during increasing irradiance till a value where they showed a sudden decrease. The decrease continued till few hours before sunset when string currents returned to overlap in a good way the

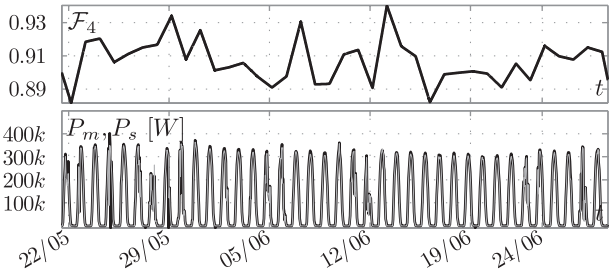

Fig. 4. Upper panel: the $\mathscr{T}_4(P_s(t),P_m(t),n)$ aggregate efficiency index during several working days of $\mathscr{S}1$. Lower panel: measured (black tick curve) and simulated (grey thin curve) powers.

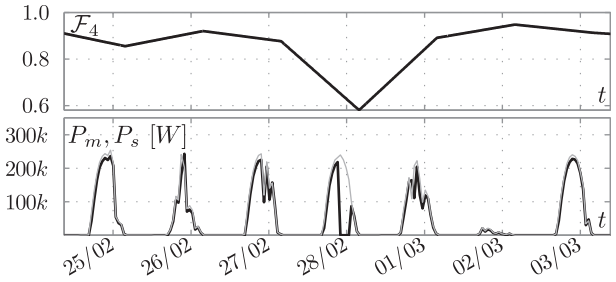

Fig. 5. Upper panel: the $\mathscr{F}_4(n)$ aggregate efficiency index of $\mathscr{S}2$. Lower panel: simulated (grey thin curve) and measured (black thick curve) powers.

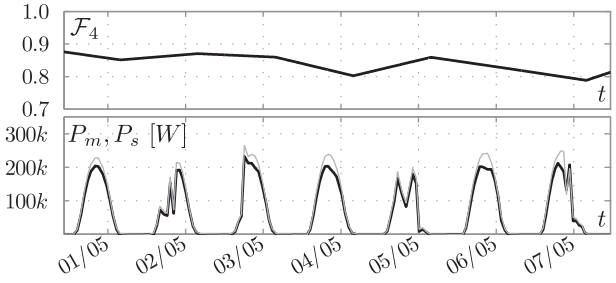

Fig. 6. Upper panel: the $\mathscr{F}_4(n)$ aggregate efficiency index of $\mathscr{S}2$. Lower panel: simulated (grey thin curve) and measured (black thick curve) powers.

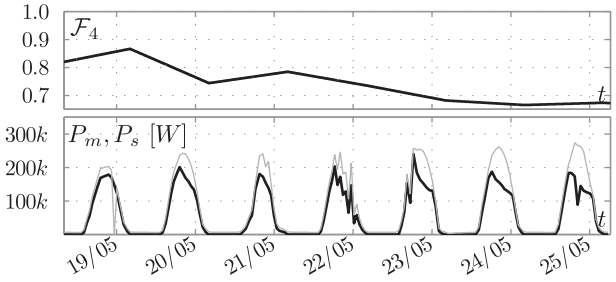

Fig. 7. Upper panel: the $\mathscr{F}_4(P_s(t), P_m(t), n)$ aggregate efficiency index of one field of $\mathscr{S}2$. Lower panel: simulated (grey thin curve) and measured (black thick curve) powers.

reference current. The important aspect is that this behaviour is homogeneous to the entire field and thus the malfunction is not caused by a specific field component. The conclusion is that the malfunctions had to be caused by the MPPT or the inverter. Direct

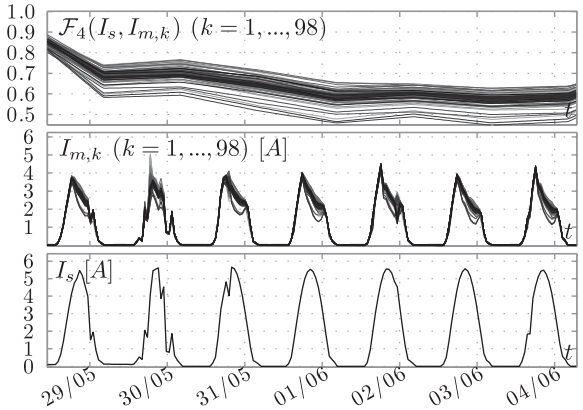

Fig. 8. Upper panel: the $\mathscr{F}_4(I_s, I_{s,k}, n)$ aggregate efficiency index of one field of $\mathscr{S}2$. Central panel: $I_{s,k}$ measured string currents. Lower panel: the I_s reference string current (in this case it is the same for all the strings).

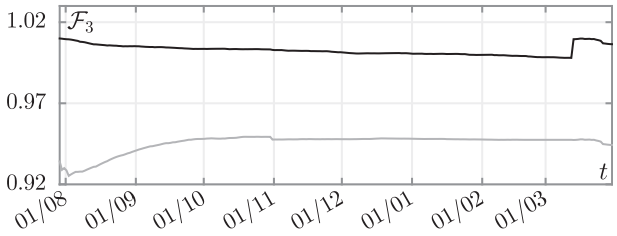

Fig. 9. The $\mathscr{F}_3$ FOM computed by using the park voltage imposed by the *real* MPPT (black curve) and the $\mathscr{F}_3$ FOM computed by using the *ideal* MPPT.

comparison of measured and simulated working voltages of the field, with observation of the inverter room temperature, allowed to identify the malfunction. Park inspection showed a failure of the air-conditioner that was unable to sufficiently cool the inverter room. Overheating of the inverter room increased according to seasonal irradiance and ambient temperature increase leading to a stop of the MPPT functionality and to a reduction of power conversion by the inverter. This behaviour was slowly worsening and identifiable by observing the slope of $\mathscr{F}_4(P_s(t), P_m(t), n)$. Note that the loss of control of the field voltage by the MPPT leads to the spread of the string currents and $\mathscr{F}_4(I_s(t), I_{m,k}(t), n)$ visible in Fig. 8.

5. Discussion

The main advantage is granularity in determining possible failures and accuracy. To highlight this aspect we considered a park that generates 404.8 kW at STC. The electrical characteristics of the 88 strings (20 panels per string) forming the park were measured and fitted with an average relative accuracy better than a fraction of %. The voltage operating point set by the *real* MPPT was used in the simulations. We computed $\mathscr{F}_3$ along 244 working days as shown in Fig. 9. As it can be seen $\mathscr{F}_3$ is slightly greater than 1 which means that the simulated PV park underestimates the performances of the real one. $\mathscr{F}_3$ is slowly decreasing and an apparent maximum mismatch of 1.2% with an abrupt decrease between the simulated model and the real one happened on 12/03/2014. In this very day panels were washed. It implies better performances, thus the observed mismatch was due to soiling. This accuracy can not be obtained by approaches based on instantaneous irradiance derived from satellite measurements [31] or a mix of satellite measures and terrestrial ones such as for example that described in Ref. [12], that was applied to domestic plants, since they suffer of an unacceptable uncertainty (up to several tens of %). In figure Fig. 9 we report also $\mathscr{F}_3$ computed through the simulation done with the *ideal* MPPT. In this case efficiency is always less than 94% showing that the *real* MPPT severely underperforms and lowers the economic return of the park. This was considered unacceptable as already partially shown by considering $\mathscr{S}_1$ and $\mathscr{S}_2$. Once more we underline that this result can be achieved only with a versatile and accurate model that is also a key aspect in soft failures identification.

The main disadvantage of the proposed approach is the need of a very accurate description of the electrical characteristics of the park, the need of an accurate measure of the environmental variables and storing of a relatively large amount of data per day. This does not pose any problem in industrial parks but can potentially limit its extension to domestic plants. In this context, it is authors' opinion that good results may possibly be obtained with domestic plants by installing irradiance sensors in a reduced number of plants located in the same relatively small area, e.g. a city. For example these irradiance sensors may be installed on the roof of public buildings. Data collection and elaboration can be done by dedicated and remote servers. Data logging from plants and

transmission to servers can be done once per day. Nowadays ADSL connections are available in almost all private houses and data loggers cost less than 100€.

6. Conclusions

An approach to accurately monitor the performance and thus efficiency of large (hundreds of kW), grid-connected PV parks has been proposed. This approach grounds on an index suitable to compute the per-day performance of the park and of the strings composing it. It exploits an accurate electro-thermal model of the PV panels and is capable of taking into account all relevant environmental variables. The essence of this novel index is to be a *weighted relative energy error*, i.e., it scales the difference between the simulated and measured powers according to the irradiance intensity. The reliability of the proposed approach allows to use it not only to monitor the overall performances of a given PV park but also to check possible malfunctions and to trigger automatic alarm messages. The proposed approach was and it's applied to monitor several industrial PV parks located in Italy.

References

[1] Tech. rep. Annual energy outlook 2009 with projections to 2030. Eia.doe.gov; 2012

[2] Pai F, Chao R, Ko S, Lee T. Performance evaluation of parabolic prediction to maximum power point tracking for PV array. Sustain Energy IEEE Trans 2011;2(1):60–8. http://dx.doi.org/10.1109/TSTE.2010.2078844.

[3] Shayani R, de Oliveira M. A new index for absolute comparison of standalone photovoltaic systems installed at different locations. Sustain Energy IEEE Trans 2011;2(4):495–500. http://dx.doi.org/10.1109/TSTE.2011.2161678.

[4] Zhao Y, Lehman B, Ball R, Mosesian J, de Palma J-F. Outlier detection rules for fault detection in solar photovoltaic arrays. In: Applied power electronics conference and exposition (APEC), 2013 twenty-eighth annual IEEE; 2013. p. 2913–20. http://dx.doi.org/10.1109/APEC.2013.6520712.

[5] Zhao Y, Yang L, Lehman B, De Palma J-F, Mosesian J, Lyons R. Decision tree-based fault detection and classification in solar photovoltaic arrays. In: Applied power electronics conference and exposition (APEC), 2012 twenty-seventh annual IEEE; 2012. p. 93–9. http://dx.doi.org/10.1109/APEC.2012.6165803.

[6] Bagnasco A, Allasia G, Giannettoni M, Pinceti P, Parodi G. Innovative solutions for photovoltaic plants remote monitoring. In: Remote engineering and virtual instrumentation (REV), 2012 9th international conference on; 2012. p. 1–5.

[7] Huld T, Gottschalg R, Beyer HG, Topic M. Mapping the performance of PV modules, effects of module type and data averaging. Sol Energy 2010;84(2): 324–38.

[8] Huld T, Dunlop E, Beyer HG, Gottschalg R. Data sets for energy rating of photovoltaic modules. Sol Energy 2013;93(7):267–79.

[9] Schubert G. Modeling hourly electricity generation from pv and wind plants in europe. In: European energy market (EEM), 2012 9th international conference on the; 2012. p. 1–7. http://dx.doi.org/10.1109/EEM.2012.6254782.

[10] Tutsch M, Vojcinak P, Koziorek J, Skrepek M. Using automated evaluation of efficiency for photovoltaic power plant. In: Emerging technologies factory automation (ETFA), 2011 IEEE 16th conference on; 2011. p. 1–4. http://dx.doi.org/10.1109/ETFA.2011.6059189.

[11] Beyer H, Betcke J, Heinemann ADD, Lorenz E, Heilscher G, Bofinger S. Identification of a general model for the MPP performance of pv-modules for the application in a procedure for the performance check of grid connected systems. In: Nineteenth EUPVSEC; 2004. p. 3073–6.

[12] Drews A, de Keizer A, Beyer H, Lorenz E, Betcke J, van Sark W, et al. Monitoring and remote failure detection of grid-connected {PV} systems based on satellite observations. Sol Energy 2007;81(4):548–64.

[13] Bizzarri F, Brambilla A, Gruosso G, Guardiani C, Sangiovanni Vincentelli A, Storti Gajani G. Modeling and estimating yield and efficiency of photovoltaic solar parks. In: Industrial technology (ICIT), 2013 IEEE international conference on; 2013. p. 734–9. http://dx.doi.org/10.1109/ICIT.2013.6505763.

[14] Bizzarri F, Bongiorno M, Brambilla A, Gruosso G, Gajani G. Model of photovoltaic power plants for performance analysis and production forecast. Sustain Energy IEEE Trans 2013;4(2):278–85. http://dx.doi.org/10.1109/TSTE.2012.2219563.

[15] International standard IEC 61215. 2nd ed. 2005.

[16] Perez R, Pierre I, Seals R, Michalsky J, Stewart R. Modeling daylight availability and irradiance components from direct and global irradiance. Sol Energy 1990;44(5):271–89.

[17] Reis F, Brito M, Corregidor V, Wemans J, Sorasio G. Modeling the performance of low concentration photovoltaic systems. Sol Energy Mater Sol Cells 2010;94(7):1222–6.

[18] Jones A, Underwood C. A thermal model for photovoltaic systems. Sol Energy 2001;70(4):349–59.

[19] Photovoltaic system performance monitoring guidelines for measurement, data exchange and analysis. 1998.

[20] Mattei M, Notton G, Cristofari C, Musselli M, Poggi P. Calculation of polycrystalline PV module temperature using a simple method of energy balance. Renew Energy 2006;31(4):553–67.

[21] Shmilovitz D. On the control of photovoltaic maximum power point tracker via output parameters. Electr Power Appl IEE Proc 2005;152(2):239–48. http://dx.doi.org/10.1049/ip-epa:20040978.

[22] Brambilla A, Gambarara M, Torrente G. Perturb and observe digital maximum power point tracker for satellite applications. In: European space agency, (special publication) ESA SP (502); 2002. p. 263–8.

[23] Skoplaki E, Palyvos J. Operating temperature of photovoltaic modules: a survey of pertinent correlations. Renew Energy 2009;34(1):23–9.

[24] Kymakis E, Kalykakis S, Papazoglou TM. Performance analysis of a grid connected photovoltaic park on the island of Crete. Energy Convers Manag 2009;50(3):433–8.

[25] Skoplaki E, Boudouvis A, Palyvos J. A simple correlation for the operating temperature of photovoltaic modules of arbitrary mounting. Sol Energy Mater Sol Cells 2008;92(11):1393–402. http://dx.doi.org/10.1016/j.solmat.2008.05.016.

[26] Tuffner D, Hammerstrom J, Singh R. Incorporation of NREL solar advisor model photovoltaic capabilities with gridlab. Tech. rep. U.S. Department of Energy; 2012

[27] Nelder JA, Mead R. A simplex method for function minimization. Comput J 1965;7:308–15.

[28] Huang B, Yang P, Lin Y, Lin B, Chen H, Lai R, et al. Solar cell junction temperature measurement of PV module. Sol Energy 2011;85(2):388–92.

[29] Garcia A, Balenzategui J. Estimation of photovoltaic module yearly temperature and performance based on nominal operation cell temperature calculations. Renew Energy 2004;29(12):1997–2010.

[30] Brambilla A, Premoli A. Electrothermal oscillations of a pn junction operating in avalanche breakdown region. Electron Device Lett IEEE 1999;20(8):405–8. http://dx.doi.org/10.1109/55.778159.

[31] Journe M, Bertrand C. Improving the spatio-temporal distribution of surface solar radiation data by merging ground and satellite measurements. Remote Sens Environ 2010;114(11):2692–704.

[32] Bizzarri F, Brambilla A, Storti Gajani G, Banerjee S. Simulation of real world circuits: extending conventional analysis methods to circuits described by heterogeneous languages. IEEE Circuits Syst Mag 2014;14(4):51–70.