Determining solar cell parameters and degradation rates from power production data

Joseph Chakar, Marko Pavlov, Yvan Bonnassieux, Jordi Badosa

Energy Conversion and Management: X 15 (2022) 100270

Contents lists available at ScienceDirect

Energy Conversion and Management: X

journal homepage: www.sciencedirect.com/journal/energy-conversion-and-management-x

Determining solar cell parameters and degradation rates from power production data

Joseph Chakar [a,*], Marko Pavlov [b], Yvan Bonnassieux [a], Jordi Badosa [c]

[a] Laboratoire de Physique des Interfaces et des Couches Minces (LPICM), Centre National de la Recherche Scientifique (CNRS UMR 7647), École Polytechnique, Institut Polytechnique de Paris (IP Paris), Palaiseau 91120, France
[b] Feedgy, Paris 75009, France
[c] École Polytechnique, Institut Pierre-Simon Laplace (IPSL), IP Paris, Ecole Normale Supérieure (ENS), Sorbonne Université, CNRS, Laboratoire de Météorologie Dynamique (LMD), Palaiseau 91120, France

ARTICLE INFO

Keywords:
Solar cell modeling
Single-diode model
Optimization
Photovoltaic degradation
Power production data

ABSTRACT

Practical but accurate methods that can assess the performance of photovoltaic (PV) systems are essential to all stakeholders in the field. This study proposes a simple approach to extract the solar cell parameters and degradation rates of a PV system from commoditized power generation and weather data. Specifically, the teaching-learning-based optimization algorithm was used to estimate the single-diode model parameters of a monocrystalline silicon PV module from a handful of power production data points that capture the operating current and voltage under real working temperatures and irradiance levels. These parameters can reproduce the solar panel's actual behavior under all operating conditions and provide insights into its underlying degradation mechanisms. The results were validated by site measurements as well as a sensitivity analysis, thus offering exciting possibilities for the future of PV performance analysis, power forecasting, and remote fault detection for real-life applications.

1. Introduction

As the International Energy Agency [1] puts it, solar is the "powerhouse of growth" in renewable energy. Researchers, companies, and countries are all racing to build high-efficiency low-cost panels and make way for large-scale solar deployment. But as the global PV fleet grows in installed capacity and age, more and more importance is being accorded to Operation & Maintenance (O&M) and repowering services, particularly in Europe, where PV plants have been operating for years [2]. In this context, PV performance modeling and monitoring play an important role in evaluating the technical and economic viability of a PV plant over its lifetime.

The cornerstone of PV modeling is the choice of a reliable equivalent circuit model. Once the model is selected – typically the single- or double-diode model – different techniques exist to estimate its parameters. The accuracy of these parameters is critical because it delimits the accuracy of the simulation results. Generally speaking, the diode model parameters are not provided by the manufacturer and cannot always be measured experimentally. Nevertheless, several methods can estimate these parameters from basic current-voltage (IV) characteristics.

Analytical methods use implicit algebraic equations to solve for solar cell parameters from key IV curve points or datasheet properties. These parameters are usually extracted at standard test conditions and can then be adjusted to the operating conditions based on their physical dependence on temperature and irradiance [3-8]. They have the advantage of being straightforward, fast, and suitable for practical applications. However, their calculation strategy relies on specific points and often involves approximations that may compromise accuracy, especially if one or more of these points is not properly specified [9].

Instead of relying on select points, numerical methods can derive solar cell parameters using iterative algorithms that typically feed on larger amounts of data. They do so by solving the diode model equation in an explicit form and fitting it to this data, thus turning the parameter extraction problem into an optimization problem that can be solved using deterministic or heuristic techniques. Deterministic methods such as the Newton-Raphson [10] and least-squares [11] methods have been proposed to extract solar cell parameters, but they are sensitive to the initialization process and may fall into local optima [12]. An increased interest is thus being given to more robust population-based (meta) heuristic techniques, which are well suited for this type of problem. A

* Corresponding author.
E-mail address: joseph.chakar@polytechnique.edu (J. Chakar).

https://doi.org/10.1016/j.ecmx.2022.100270
Received 25 May 2022; Received in revised form 28 June 2022; Accepted 14 July 2022
Available online 16 July 2022

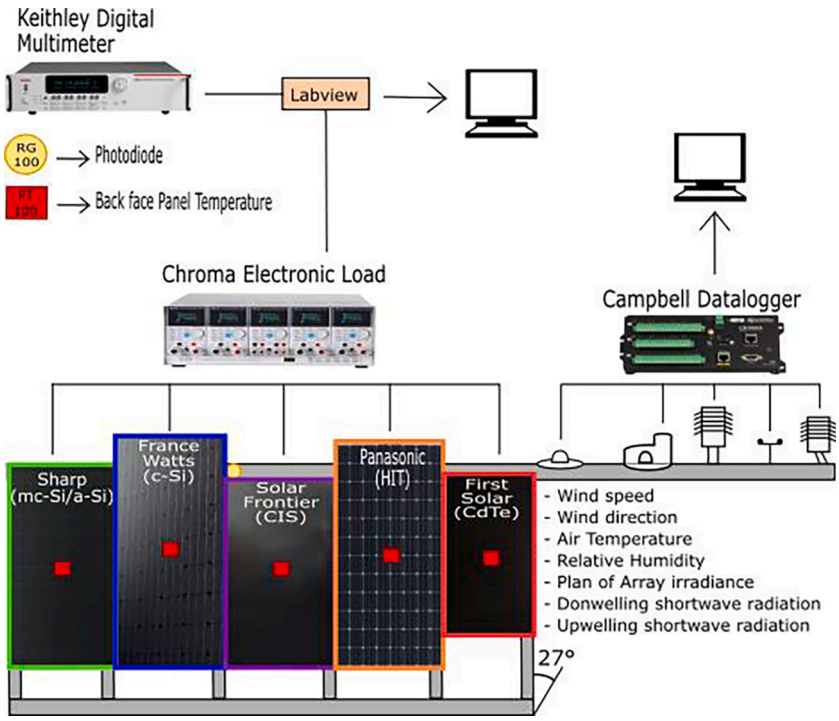

Fig. 1. Schematic of the SIRTA PV test bench. In this study, only the *FranceWatts* c-Si module is considered.

plethora of heuristic approaches have been applied to extract solar cell parameters, including particle swarm optimization, genetic algorithms, and teaching-learning-based optimization among many others [12-20]. Some authors also improve on established heuristic methods with additional features [21,22], while others resort to hybrid algorithms that merge two or more of these methods [23-26].

Most of these heuristic algorithms are easy to implement and successfully extract solar cell parameters from IV curve data. However, even though IV curves are an effective way to assess PV performance, they are rarely measured in residential, commercial, or utility-scale solar farms, which makes it difficult to apply these studies to real-life applications. In contrast, power production data (i.e. time-series that list the operating current and voltage of deployed PV systems at various temperatures and irradiance levels) is an often overlooked but information-rich resource that has become increasingly accessible – as it can be measured and recorded by commercial PV equipment such as inverters and data loggers.

This article shows how power production data can be used to determine the solar cell parameters and degradation rates of a PV system. First, the single-diode model is selected to simulate the performance of a monocrystalline PV module under given operating conditions. Next, the Teaching-Learning-Based Optimization (TLBO) algorithm [27] is chosen to find a set of model parameters that can reproduce the panel's actual behavior using just a few power production data points. These parameters are then translated into IV characteristics to determine the module degradation rates. Different filters are suggested to clean the electrical and environmental data, and the possibility of using weather data in the absence of on-site temperature and irradiance measurements is also validated (see Supplemental Information). The modeling results are backed by site measurements, and a sensitivity analysis (see Supplemental Information) is included to check their robustness.

2. Methodology

2.1. PV model

The Single-Diode Model (SDM) is a simple yet effective model that is widely-used to describe a solar cell. Derived from physical principles, it relates the cell output current to the voltage at its terminals using the following governing equation [28]:

$$I = I_L - I_0[exp(\frac{V + IR_s}{nV_T}) - 1] - \frac{V + IR_s}{R_{sh}} \quad (1)$$

For a given set of operating conditions (i.e. solar irradiance and cell temperature), the SDM depends on five primary parameters: the light-induced current (I_L), diode reverse saturation current (I_0), series resistance (R_s), shunt resistance (R_{sh}), and ideality factor (n, typically between 1 and 2 [28]). V_T is the thermal voltage given by $V_T = \frac{k_B T_c}{q}$, where k_B is Boltzmann's constant, T_c the cell temperature, and q the elementary charge.

Even though the PV cell is the primary power generation unit, solar panel specifications are more accessible, and module-level modeling significantly reduces the computational cost. If we assume that all the cells of a solar panel comprising N_s cells in series are identical and under uniform and equal irradiance and temperature (i.e. they generate the same current and voltage), we get $I_M = I_{cell}$ and $V_M = N_s \times V_{cell}$, where I_M and V_M represent the module current and voltage, respectively. The governing equation becomes [29]:

$$I_M = I_L - I_0\left[exp\left(\frac{V_M + I_M N_s R_s}{a}\right) - 1\right] - \frac{V_M + I_M N_s R_s}{N_s R_{sh}} \quad (2)$$

The modified ideality factor a is simply the product of the ideality factor n, number of cells in series N_s, and thermal voltage V_T: $a = n \times N_s \times V_T$.

Here, PV modeling is done at the module level using the pvlib Python package [30]. In other terms, a PV module is represented by a set of five

Table 1
Measured vs. estimated ideal energy production of the *FranceWatts* module during the analysis period.

Measured Energy Output	Ideal Energy Output
92.93 kWh	99.31 kWh

The DC energy produced by the FranceWatts c-Si module is taken as the product of its output current and voltage at the MPP, integrated over time. The analysis period is set from 07/01/2021 to 30/09/2021. For a fair comparison, the same timestamps are used in both cases. The only data cleaning tools used here are the uptime and outlier filters, which are described in the Supplemental Information.

SDM parameters at Standard Test Conditions (STC)[1]. After adjusting these parameters to the operating conditions using the De Soto model [4], the module's IV curve can be simulated to find the DC output current, voltage, and power at the Maximum Power Point (MPP). The full procedure is described in the Supplemental Information and can be adapted to any PV system.

2.2. PV test bench

This study is based on the data collected at the SIRTA (Site Instrumental de Recherche par Télédétection Atmosphérique) observatory's [31] PV test bench [32] located in Palaiseau, France (48.7 N, 2.2E) on the Ecole Polytechnique campus and portrayed in the schematic of Fig. 1. The test bench was installed in 2014 and hosts five commercial solar panels of different technologies. The panels are installed in a free-standing configuration, facing South with a 27° tilt.

Agilent DC electronic loads (6060B) measure IV curves roughly every 40 s, from sunrise to sunset. These curves are translated into time-series of the short-circuit current I_{sc}, open-circuit voltage V_{oc}, fill factor *FF*, cell efficiency η, series resistance R_s, and shunt resistance R_{sh} for each panel. Due to a voltage drop along the junction line and the transient mode of measurement, an uncertainty of $\pm$ 7.2% should be considered for power measurements [32].

In terms of environmental measurements, four-wired class A platinum sensors (PT100) measure the module backside temperatures using the same time interval. The measurement uncertainty is $\pm$ 0.4 °C, and the difference between the cell temperature and the glued probe on the back is always less than 2 °C [32]. These sensors are read by a digital multimeter (Tektronix DMM 3700A). The global in-plane irradiance (GPOA) is also measured every 5 s using a a second class solar radiometer (Hukseflux SR01) installed in the same plane of the modules. This device is compliant with the ISO 9060 s class specifications and is expected to have a daily uncertainty $\leq$ 1% [32]. Several other parameters are measured on site, but they are beyond the scope of this study.

Here, only the *FranceWatts* crystalline silicon (c-Si) PV module (second one from the left in Fig. 1) is considered. The analysis period[2] is set from 07/01/2021 to 30/09/2021.

2.3. Performance analysis

The "ideal" or initial reference SDM parameters of the *FranceWatts* c-Si PV module (i.e. $I_{L,i}$, $I_{0,i}$, $R_{s,i}$, $R_{sh,i}$, and a_i at STC) can be estimated from the panel datasheet properties using the California Energy Commission (CEC) model and estimation method [6]. These parameters can then be used to simulate the module's theoretical behavior during the analysis period. Table 1 compares this ideal PV performance with the measured production data in an effort to detect and quantify any system underperformance.

Note that this simulated ideal energy output does not account for the natural degradation covered by the *FranceWatts* performance warranty, which in this case guarantees 90% of production capability after 12 years. These results thus indicate that the c-Si module is functioning normally after roughly 7 years of operation and during the entirety of the analysis period – which is not surprising considering that it is part of a well-ventilated and isolated PV system.

While they can help quantify energy losses, the ideal SDM parameters alone do not reveal the root cause of this underperformance – regardless of whether it is due to normal or premature solar panel degradation. This study shows how a metaheuristic algorithm such as the Teaching-Learning-Based Optimization (TLBO) can be used to extract the "real" or actual SDM parameters of a PV system from widely available power generation and weather data. Along with their ideal reference values, these parameters can provide more insight into PV degradation.

2.4. Dataset

In order to accurately model the PV module, the SDM parameters need to be extracted from power production data points that reflect the true PV performance to the best extent possible. After filtering out cloudy days and outliers from the system uptime, the complete power production dataset of this PV module amounts to 21,359 distinct data points that capture the module performance under a variety of temperatures and irradiance levels. Each data point consists of a set of electrical (the module DC output current, voltage, and power at the MPP) and environmental measurements (the GPOA irradiance and module temperature) having the same timestamp. See the Supplemental Information for more details on the data cleaning procedures and dataset characteristics.

Nonetheless, this study reveals that only a tiny fraction of this production history is needed to extract the real SDM parameters of a PV system. Here, the training set is built by grouping the data into respective irradiance and temperature intervals of 100 W/m^2 and 5 °C, and then selecting five random points from each bin – leading to a total of 130 distinct data points. This represents only 0.6% of the filtered power production data. Next, the SDM parameters of the *FranceWatts* PV module are extracted from these data points using the TLBO, which is chosen from experience and for its relative simplicity and effectiveness.

2.5. Teaching-learning-based optimization

As the name implies, the Teaching-Learning-Based Optimization [27] is an optimization technique based on the teaching-learning process of a classroom. It is a population-based algorithm in which a randomly sampled group of learners learn from a teacher and pass on knowledge to one another in order to reach an optimal solution.

Based on their target values, learners are assessed through an objective function called the fitness function $F(X)$. This function can be thought of as an error value that should be minimized (and ideally reach zero). The performance of a given learner X_i can thus be determined by the value of its fitness function, also referred to as the fitness of the learner. The fitness of the total population is improved through the transfer of knowledge that occurs during the Teacher and Learner phases described below.

In the Teacher stage of the TLBO, the algorithm tries to improve the fitness of every learner by moving its value towards that of the classroom teacher $X_{teacher}$, which is simply the learner with the smallest fitness function. The teacher modifies each learner using the mean value of the population X_{mean} according to the following equation:

$$X_{i,new} = X_i + r(X_{teacher} - T_F X_{mean}) \quad (3)$$

where $i = 1, 2 \cdots N$, with N being the number of learners in the classroom.

[1] Typically, at an irradiance of 1000 W/m^2 and cell temperature of 25°C.

[2] The analysis period was set solely based on the data availability at the time of the project execution. The impact of selecting this period is considered at the end of the Results & Discussion section.

Table 2
Baseline parameter search space.

Single-Diode Model Parameter	Ideal Reference Value	Minimum Allowable Value	Maximum Allowable Value
I_L	$I_{L,i} = 8.73$ A	$\frac{1}{2}I_{L,i}$	$1.1 \times I_{L,i}$
I_0	$I_{0,i} = 3.45 \times 10^{-10}$ A	$0.9 \times I_{0,i}$	$I_{0,i} \times 20$
R_s	$R_{s,i} = 0.31\ \Omega$	$0.9 \times R_{s,i}$	$R_{s,i} \times 4$
R_{sh}	$R_{sh,i} = 523.31\ \Omega$	$\frac{1}{20}R_{sh,i}$	$R_{sh,i} \times 1.1$
a	$a_i = 1.57\ \ (n \sim 1)$	$N_s \times V_T\ \ (n = 1)$	$N_s \times V_T \times 2$

This table describes the baseline parameter search space of the Teaching-Learning-Based Optimization. The single-diode model parameters are ordered as follows: light-induced current (I_L), diode reverse saturation current (I_0), series resistance (R_s), shunt resistance (R_{sh}), and modified diode ideality factor (a). The ideal reference values at Standard Test Conditions (STC) listed in the second column are estimated from the panel datasheet properties using the California Energy Commission model and estimation method [6]. They are assumed to be the best possible values of each parameter (with a ± 10% margin to allow for change in both directions) and are used to define the parameter bounds.

The teaching factor T_F determines the impact of the population average X_{mean}, whereas the variable r represents the learning rate or fraction of knowledge being shared during the teacher-learner interaction. The limit values of 0 and 1 of r represent the two extreme cases where a learner either learns nothing or everything from the teacher. The values of T_F and r affect the amount of knowledge transferred as well as the speed of convergence. A small learning rate ensures a more thorough exploration of the search space but may come at the expense of the computation time. However, if this rate is too high, the training may not properly converge. In machine and deep learning applications, the training typically starts with a relatively large learning rate to quickly "educate" the randomly generated parameters. This rate is then slowly decreased during the training to fine-tune them.

The modified learner $X_{i,new}$ replaces the current learner X_i only if it is found to be "better" (i.e. if it has a smaller fitness function).

During the Learner phase, learners interact with each other to share knowledge between them. Every learner ($X_i,\ i = 1, 2 \cdots N$) interacts with another randomly selected learner ($X_j,\ j = 1, 2 \cdots N$) in the population, where $i \neq j$. If X_i is better than X_j (i.e. $F(X_i) < F(X_j)$), the value of X_j shifts towards that of X_i, and vice versa. A mathematical representation of this interaction is shown below:

$$if\ F(X_i) <\ F(X_j) : X_{i,new} = X_i + r(X_i - X_j) \tag{4}$$

$$if\ F(X_i) >\ F(X_j) : X_{i,new} = X_i + r(X_j - X_i) \tag{5}$$

Similarly here, X_i is replaced by $X_{i,new}$ only if the latter is better than the former.

In this study, each TLBO learner learns five subjects (i.e. I_L, I_0, R_s, R_{sh} and a at STC). In other terms, each learner X_i is defined as a set of SDM parameters that can simulate the electrical performance of the module under given operating conditions, using the PV modeling approach described earlier and detailed in the Supplemental Information.

The fitness function is defined as the average Mean Absolute Percentage Error (MAPE) between the modeled DC output current, voltage, and power and their target values (i.e. the values measured under the same operating conditions). During the teacher and learner phases of the TLBO, learners are continuously updated such as to minimize this fitness function. At the end of the training process, the set of SDM parameters of the best learner (i.e. the learner with the smallest fitness function) is selected.

Note that the TLBO algorithm is controlled by five main hyper-parameters: the population size (i.e. the number of learners), number of iterations, search space (i.e. the minimum and maximum allowable values for each subject), teaching factor T_F, and learning rate r. Here, 50 learners are trained over 150 iterations. During the initialization stage of the learner population, the set of ideal SDM parameters extracted from the panel datasheet is included to accelerate the training process. These parameters are also used to define the bounds of the parameter search space presented in Table 2. In Patel et al. [21], the teaching factor and learning rate are not fed as inputs to the algorithm. Based on the findings of Rao et al,[28] the authors randomly set T_F to either 1 or 2 during the training. Similarly, they randomly set the value of r between 0 and 1 on the fly. In this study however, T_F and r are taken as fixed inputs to the algorithm. Their values are set to 2 and 0.2, respectively.

This article describes the results obtained when the above baseline configuration and training set are used for the TLBO. The sensitivity of this iterative algorithm to the control variables and input data is also reported.

Table 3
Single-diode model parameters estimated from power production data.

Single-Diode Model Parameter	Estimated Value at STC
I_L	8.25 A
I_0	5.82×10^{-10} A
R_s	0.32 Ω
R_{sh}	178.58 Ω
a	1.64

The single-diode model parameters correspond to the best learner (i.e. the learner with the lowest fitness function) at the end of the training. The parameters are estimated at Standard Testing Conditions (STC) and ordered as follows: light-induced current (I_L), diode reverse saturation current (I_0), series resistance (R_s), shunt resistance (R_{sh}), and modified diode ideality factor (a).

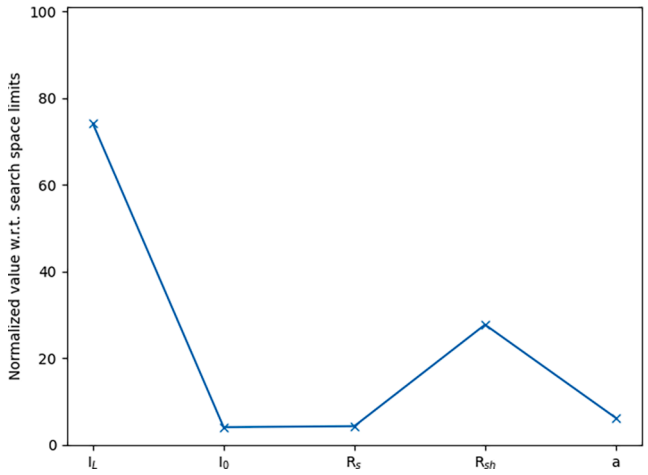

Fig. 2. Position of the fitted parameters w.r.t. their search space limits. A value of 0% means that the Single Diode Model (SDM) parameter is equal to its lower search space limit, and a value of 100% means that it is equal to its upper limit. The parameters are ordered as follows: light-induced current (I_L), diode reverse saturation current (I_0), series resistance (R_s), shunt resistance (R_{sh}), and modified diode ideality factor (a).

3. Results & Discussion

3.1. Parameter extraction

Table 3 lists the SDM parameters that correspond to the best learner obtained at the end of the training. To account for the stochastic nature of the algorithm, three consecutive TLBO runs are performed, and the average parameter values are reported below. Additional statistical information about the different runs is provided in the Supplemental Information.

Table 4
Estimated current-voltage parameters and degradation rates.

	Datasheet Value	Estimated Actual Value	Degradation Rate
I_{MPP}	8.21 A	7.63 A	1.00 %/year
V_{MPP}	30.52 V	31.00 V	−0.22 %/year
P_{MPP}	250 W	236.55 W	0.80 %/year
I_{sc}	8.64 A	8.23 A	0.68 %/year
V_{oc}	37.67 V	38.23 V	−0.21 %/year

This table compares the datasheet reference values and estimated current values of the *FranceWatts* module's current-voltage (IV) properties at Standard Testing Conditions (STC). The parameters are ordered as follows: output current, voltage, and power at the maximum power point (I_{MPP}, V_{MPP}, and P_{MPP} respectively), short-circuit current (I_{sc}), and open-circuit voltage (V_{oc}). A negative degradation rate implies improvement.

To better understand these values, the relative position of each SDM parameter w.r.t to its respective search space bounds is plotted in Fig. 2. A value of 0% means that the extracted parameter is equal to its lower limit, whereas a value of 100% means that it is equal to its upper limit.

Based on these results, it appears that the dark saturation current I_L, series resistance R_s and ideality factor n have not changed over time. However, there seems to be a deterioration of the light-induced current I_L, along with a considerable drop in the shunt resistance R_{sh}.

Using these SDM parameters, the PV module's actual current-voltage properties can be estimated at STC. Given that IV curves are not always provided for individual panels, these properties can be compared with the ones specified on the datasheet to estimate their linear degradation rates, as shown in Table 4.

As expected, the deterioration of the light-induced current and shunt resistance translates into lower short-circuit and MPP currents, and consequently a lower power output. Moreover, considering that PV performance typically drops in the first year of operation (1% to 3%) before steadily decreasing thereafter (0.5% to 1% per year) [33], the estimated power degradation rate of this overall healthy module is consistent with the equivalent rate of its warranty (roughly 0.83%/year). Though they may be exaggerated due to the modeling inaccuracies explained below, the voltage and current degradation rates all fall in the ranges reported by Jordan et al [34] for mono-crystalline silicon solar cells. They also provide insight into the performance analysis and reveal that this rather justified degradation is related to two physical parameters that mostly affect the output current.

To verify their ability to generalize to unseen data, the SDM parameters extracted from the training set are used to simulate the panel output current, voltage, an d power under all the 21,359 operating conditions present in the dataset. The training set is included in this validation stage since its size is negligible. Using a color map based on the frequency of data values, the plots of Fig. 3 show the match between the simulated values and measured data. The Mean Absolute Percentage Error (MAPE) and the Coefficient of Determination or R-squared (R2) are also chosen to evaluate the goodness-of-fit. Simply put, the first metric measures how far the modeled values fall from the observed data

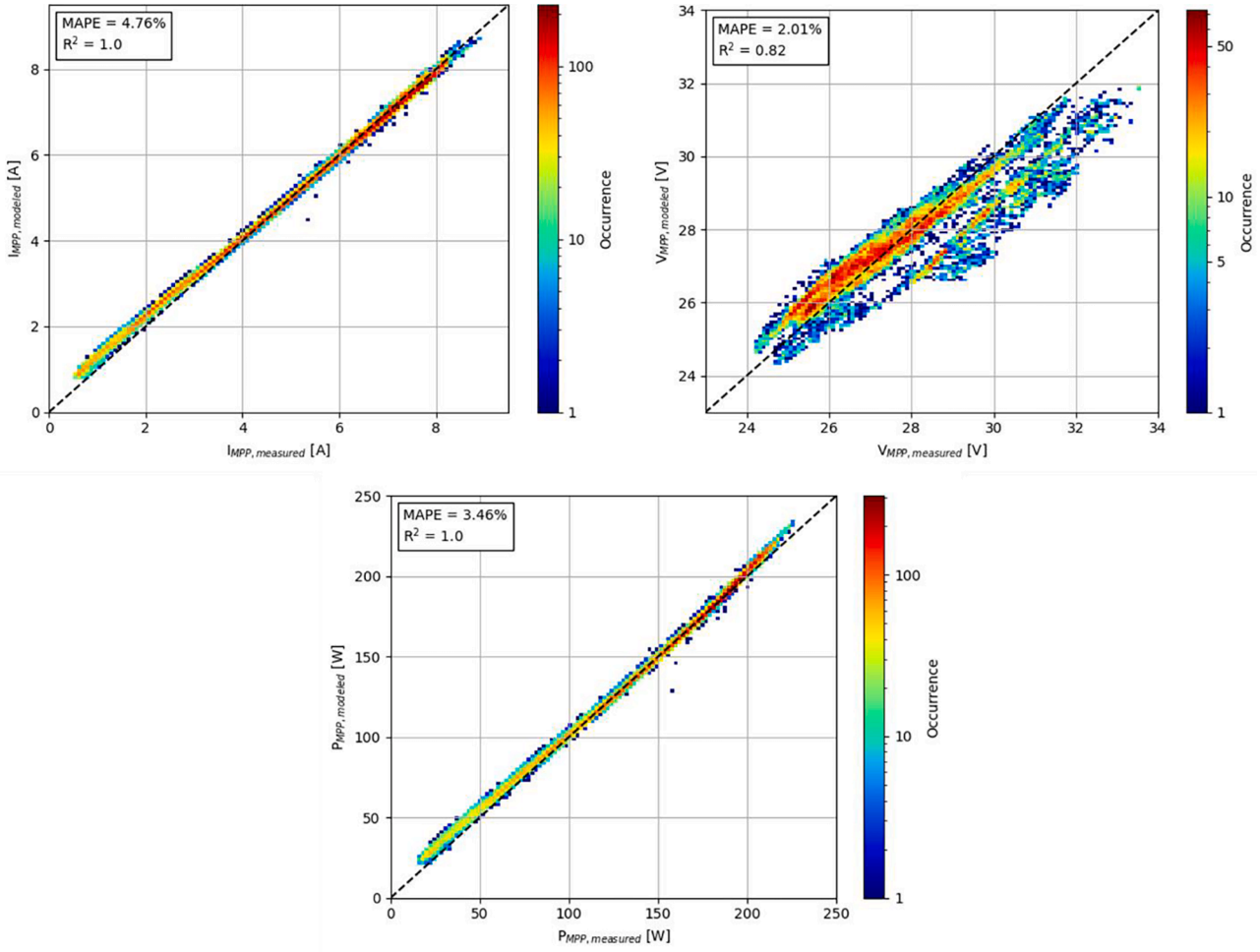

Fig. 3. Modeled vs. measured values of the panel output current, voltage, and power (on the clear-sky days of the analysis period). These plots compare the values of the *FranceWatts* module's output current, voltage, and power at the Maximum Power Point (MPP) that are measured on-site with those simulated using the extracted single-diode model parameters. The color map reflects the number of data points represented in each point block.

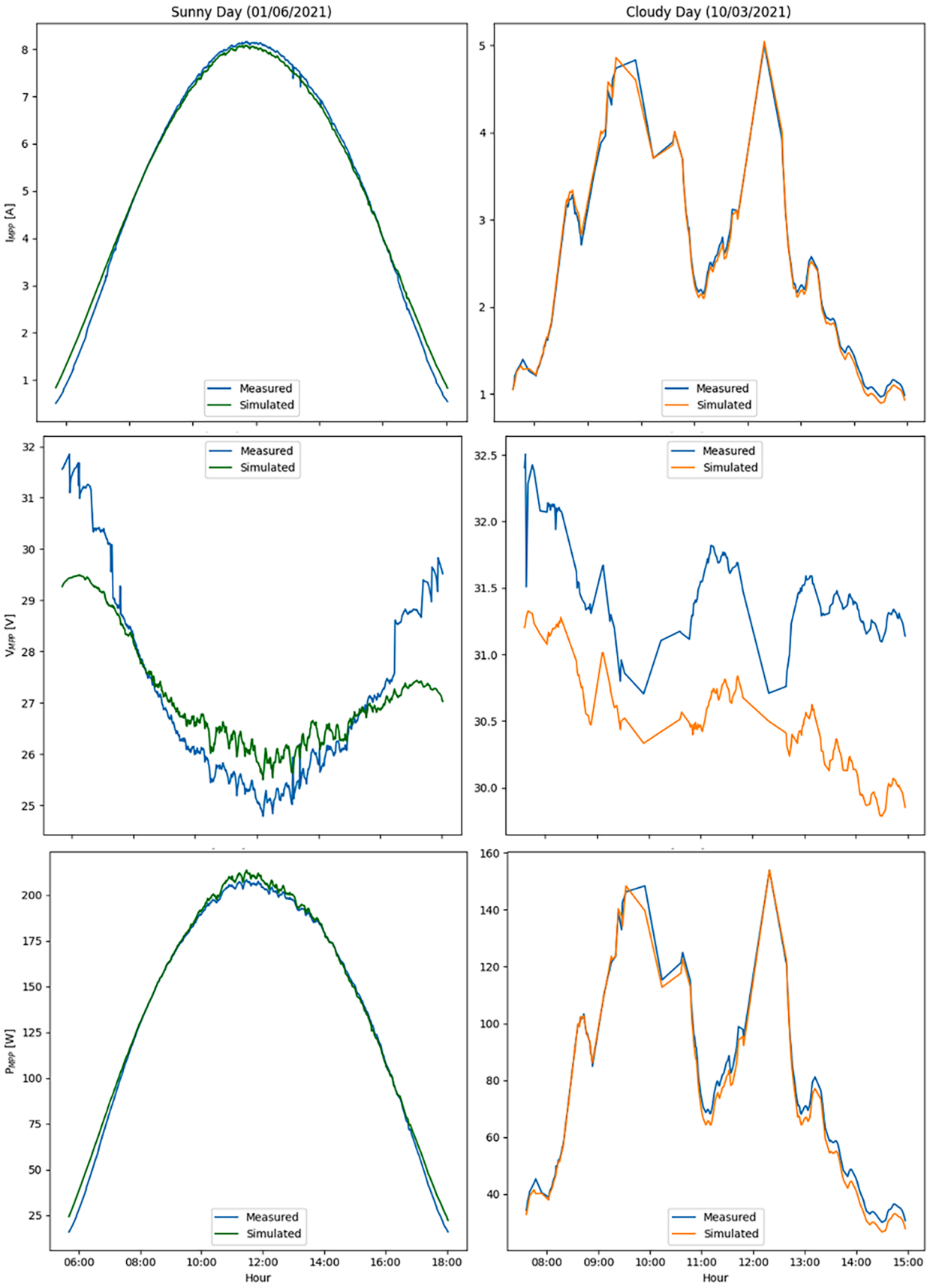

Fig. 4. Simulated vs. measured daily profiles of the module output current, voltage, and power on sample sunny and cloudy days. These plots compare measured daily profiles of the *FranceWatts* module's output current, voltage, and power (I_{MPP}, V_{MPP}, and P_{MPP}, respectively) with those simulated using the single-diode model parameters determined from the power production data.

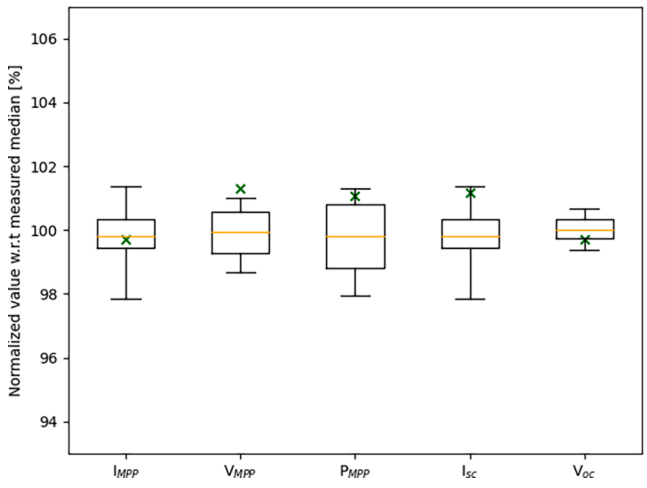

Fig. 5. Modeled vs. measured current-voltage parameters. For each parameter, the box extends from the first quartile to the third quartile of the measured data, with an orange line at the mean. The whiskers extend to the minimum and maximum of the measured data, and the green mark corresponds to the modeled value. For visualization purposes, all plotted values are normalized w.r.t. the median of the measured data. The parameters are ordered as follows: output current, voltage, and power at the maximum power point (I_{MPP}, V_{MPP}, P_{MPP} and respectively), short-circuit current (I_{sc}), and open-circuit voltage (V_{oc}). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and should be as close to zero as possible, while the second statistical measure indicates the degree of linear correlation between the simulated and target values and should ideally be close to 1.

Overall, there is a strong match between the simulated and observed data. In all cases, the MAPE is below 5% and remains within the measurement uncertainty. The R^2 value touches 1 in the case of the current and power, albeit the diode model seems to be overestimating them at lower ranges. The power output also tends to be slightly overestimated at higher ranges. At first sight, there seems to be a striking negative (positive) bias at higher (lower) voltage ranges, but the low MAPE value indicates that this error is relatively small and comparable with the precision of the equipment used on site.

Naturally, when cloudy days are added (49,781 data points in total), the MAPE increases slightly – most likely because of the lower signal-to-noise ratio. The R^2 value is unaffected in the case of the current and power, but it marginally decreases in the case of the voltage. Moreover, the negative (positive) bias at higher (lower) voltages (currents) is more pronounced (see Fig. S3). This is also expected since PV modules typically operate at higher (lower) voltages (currents) on cold days due to their negative (positive) voltage (current) temperature coefficient.

It is thus safe to say that the SDM parameters extracted from the production data can reproduce the module's behavior on both sunny and cloudy days, as illustrated in Fig. 4.

A closer look at the daily profiles reveals that these observed discrepancies occur throughout the entire analysis period, although they are accentuated at higher voltages and low irradiance levels during early mornings and late afternoons. This mismatch is largely due to the modeling inaccuracies associated with the solar cell parameters' dependence on the different operating conditions (see Supplemental Information).

3.2. Validation with on-site measurements

The IV curves measured on site can be used to further validate the modeling results. As STC have not been met for this module on any clear-sky day during the analysis period, data points with a clear-sky GPOA irradiance of 1000 ± 1% W/m^2 and module backside temperature of 50 °C (±0.4 °C) are selected instead (total of 34 points). The below box and whisker plot shows that the IV properties modeled using the extracted SDM parameters all fall within their respective range of measurements (Fig. 5) (see also Table S1).

These results are consistent with the previous observations. V_{MPP} is slightly overestimated here, but the relative error is still below 2% and well within measurement uncertainties for all of the IV properties. As shown in Fig. 6, the comparison between the modeled and measured IV curves at different temperatures and irradiance levels can provide additional insight into the ability of the PV model and SDM parameters to reproduce the full module behavior.

At higher irradiance levels, the SDM parameters extracted from the production data can successfully model the PV module's IV curves, especially near the MPP. It is also important to keep in mind that the IV curves collected on site are measured under real and thus variable conditions, rather than in a controlled lab environment. Nonetheless, the misalignment at the curve tail as well as the overestimation of V_{oc} and I_{sc} at lower irradiance and temperature levels both indicate potential flaws in the way the SDM parameters are adjusted according to the operating conditions (see Supplemental Information).

The proposed method also correctly captures the evolution of the series and shunt resistances, despite having different initial estimates of their STC values. Indeed, the numerical results show that as opposed to the series resistance, the shunt resistance has drastically deteriorated over time (from 523 Ω to 178 Ω). Similarly, the measurements collected during the analysis period suggest that the respective STC values of the series and shunt resistances are around 0.59 Ω and 45 Ω, compared with 0.58 Ω and 249.43 Ω when the panel was purchased and initially characterized using a PASAN SPROD IV curve Tester.

Being able to accurately model the full IV curve is critical for PV modeling, especially for bigger and more complex systems that can suffer from mismatch. This issue is also important because the nature of the deviations in an IV curve can provide valuable clues about potential performance problems [35]. As a relevant example, the shunt resistance can be used as a means to detect Potential Induced Degradation (PID) [36].

3.3. Sensitivity analysis

The SDM is more sensitive to some of its parameters than others (e.g. the ideality factor), and different combinations of these parameters exist to model the studied PV system. This trade-off between them is also affected by the stochastic nature of the TLBO and its configuration (see Supplemental Information). However, regardless of the configuration used (to a reasonable extent, of course), the end conclusion remains the same: there is mainly a current degradation due to a drop in the light-induced current and shunt resistance. Considering that the algorithm running time on the average computer is around one hour for this specific application, optimizing its hyperparameters can significantly lower the computational cost and convergence time. Moreover, resorting to alternative parameter estimation techniques such as Bayesian inference can help find the probability distributions of the SDM parameters and unravel the correlations between them.

It is also interesting to note that the proposed methodology is not sensitive to the precision of the input working conditions. This being said, the same end conclusion is obtained when less accurate but freely available weather data and a thermal model (see Supplemental Information) are used instead of on-site irradiance and temperature measurements – which is practical for real-life applications.

Given that only a handful of production data points is needed to extract the SDM parameters, this type of study can be repeated regularly using shorter analysis periods to observe how these parameters change over time and detect any potential degradation early-on. However, since the accuracy of the single-diode and De Soto models drops at low irradiance and temperature levels, a more robust PV model must be selected to get reliable periodic results across the different seasons. In this case,

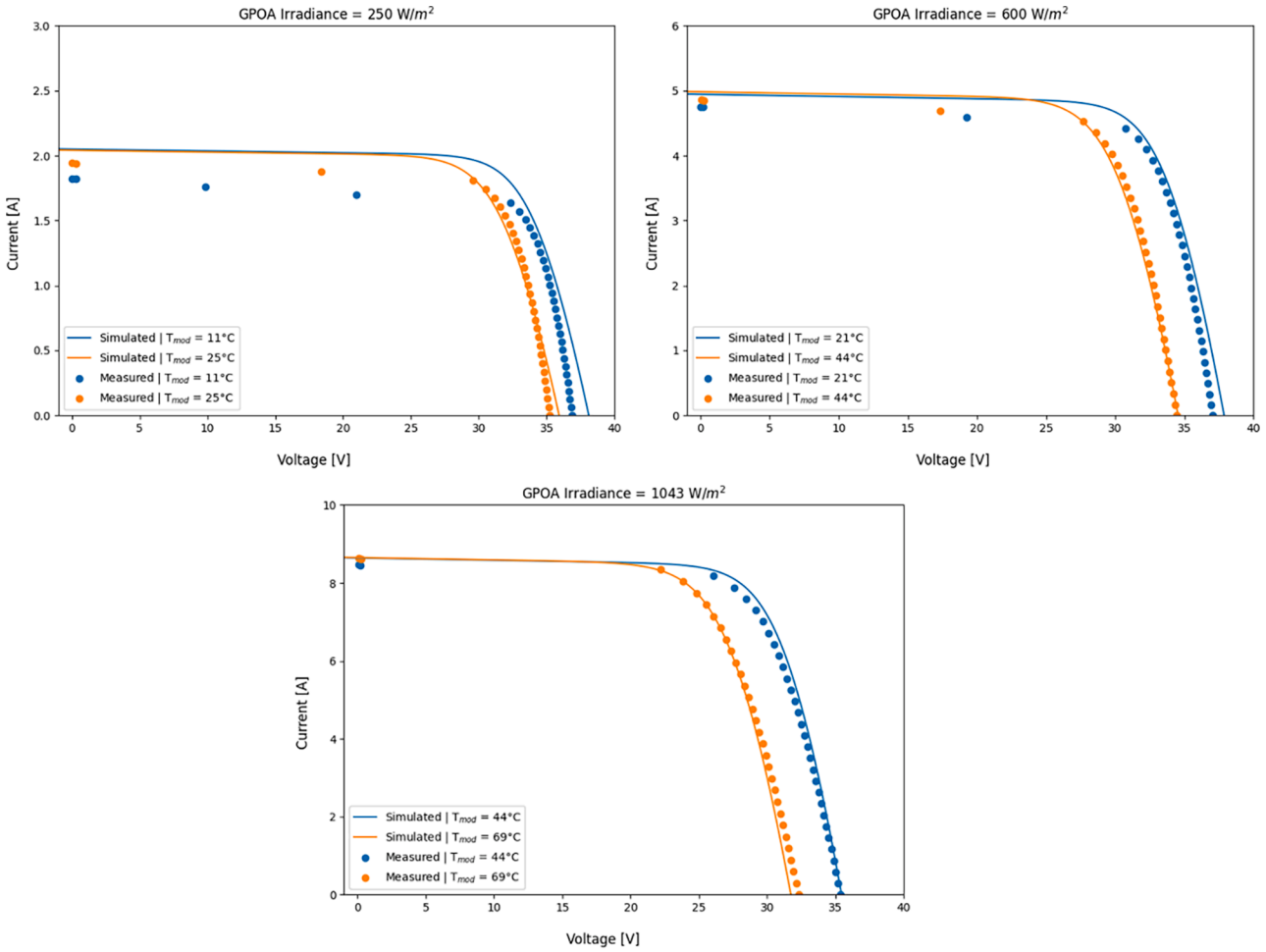

Fig. 6. Modeled vs. measured current-voltage curves at different temperatures and irradiance levels. The simulated current-voltage (IV) curves are represented by full lines and are calculated at different module temperatures and global plane-of-array (GPOA) irradiance levels. All of the discrete IV curve points measured on site are represented by dots.

the collected dataset captures PV performance across a wide range of temperature and illumination levels – as the PV module was operating normally during the 9 months of the analysis period – and most likely compensates for these modeling inaccuracies.

4. Conclusion

This article demonstrates the exciting possibility of using PV power generation data to determine solar cell parameters, simulate IV curves, understand PV degradation, and identify faults. It shows how detailed information on the electrical performance of a crystalline silicon PV module can be extracted using a simple metaheuristic algorithm and just a fraction of the production history. The promising results are validated with site measurements, and the proposed approach can be scaled up to conduct performance analysis of PV systems remotely, in real-time, and without any field surveys. These initial findings will hopefully encourage researchers to further investigate solar cell models, data processing techniques, and algorithms that can leverage the full power of the information hidden in the rather overlooked production data of different PV systems and technologies.

Data availability

The data that support the findings of this study are openly available in in the following repository https://gitlab.in2p3.fr/energy4climate/public/sirta-pv1-data. The codes are also available upon request from the corresponding author.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors acknowledge PhD candidate Moira Torres for helping access the data used in this study. This project was first launched at Feedgy, a PV performance analysis and repowering company, and is being further developed at Ecole Polytechnique as part of an ongoing collaboration. This research was also produced within the framework of Energy4Climate Interdisciplinary Center (E4C) of IP Paris and Ecole des Ponts ParisTech. It was supported by the 3rd Programme d'Investissements d'Avenir [ANR-18-EUR-0006-02] and by the Foundation of Ecole Polytechnique (Chaire "Défis Technologiques pour une Énergie Responsable" financed by Total Energies).

Appendix A. Supplementary Data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecmx.2022.100270.

References

[1] International Energy Agency (IEA). Renewables 2021 – Analysis and forecast to 2026; 2021.

[2] Chiantore PV, Hemetsberger W. Operation & Maintenance Best Practice Guidelines Version 4.0. Solar Power Europe; 2019.

[3] Batzelis EI, Papathanassiou SA. A Method for the Analytical Extraction of the Single-Diode PV Model Parameters. IEEE Trans Sustainable Energy 2016;7:504–12. https://doi.org/10.1109/TSTE.2015.2503435.

[4] De Soto W, Klein SA, Beckman WA. Improvement and validation of a model for photovoltaic array performance. Sol Energy 2006;80:78–88. https://doi.org/10.1016/j.solener.2005.06.010.

[5] Di Piazza MC, Luna M, Petrone G, Spagnuolo G. Translation of the single-diode pv model parameters identified by using explicit formulas. IEEE J Photovoltaics 2017;7:1009–16. https://doi.org/10.1109/JPHOTOV.2017.2699321.

[6] Dobos A. An Improved Coefficient Calculator for the California Energy Commission 6 Parameter Photovoltaic Module Model. J Sol Energy Eng 2012;134. https://doi.org/10.1115/1.4005759.

[7] Lo Brano V, Ciulla G. An efficient analytical approach for obtaining a five parameters model of photovoltaic modules using only reference data. Appl Energy 2013;111:894–903. https://doi.org/10.1016/j.apenergy.2013.06.046.

[8] Phang J, Chan D, Phillips J. Accurate analytical method for the extraction of solar cell model parameters. Electron Lett 1984;20:406–8. https://doi.org/10.1049/el:19840281.

[9] Ishaque K, Salam Z, Taheri H, Shamsudin A. A critical evaluation of EA computational methods for Photovoltaic cell parameter extraction based on two diode model. Sol Energy 2011;85:1768–79. https://doi.org/10.1016/j.solener.2011.04.015.

[10] Raj S, Sinha AK, Panchal AK. Solar cell parameters estimation from illuminated I-V characteristic using linear slope equations and Newton-Raphson technique. J Renewable Sustainable Energy 2013;5:401–4. https://doi.org/10.1063/1.4803748.

[11] Easwarakhanthan T, Bottin J, Bouhouch I, Boutrit C. Nonlinear Minimization Algorithm for Determining the Solar Cell Parameters with Microcomputers. International Journal of Solar Energy 1986;4:1–12. https://doi.org/10.1080/01425918608909835.

[12] Xiong G, Li L, Mohamed AW, Yuan X, Zhang J. A new method for parameter extraction of solar photovoltaic models using gaining–sharing knowledge based algorithm. Energy Rep 2021;7:3286–301. https://doi.org/10.1016/j.egyr.2021.05.030.

[13] Ye M, Wang X, Xu Y. Parameter extraction of solar cells using particle swarm optimization. J Appl Phys 2009;105(9):094502.

[14] Sheng H, Li C, Wang H, Yan Z, Xiong Y, Cao Z, et al. Parameters Extraction of Photovoltaic Models Using an Improved Moth-Flame Optimization. Energies 2019;12:3527. https://doi.org/10.3390/en12183527.

[15] Jervase JA, Bourdoucen H, Al-Lawati A. Solar cell parameter extraction using genetic algorithms. Meas Sci Technol 2001;12:1922–5. https://doi.org/10.1088/0957-0233/12/11/322.

[16] Xiong G, Zhang J, Shi D, Yuan X. Application of supply-demand-based optimization for parameter extraction of solar photovoltaic models. Complexity 2019;2019:1–21. https://doi.org/10.1155/2019/3923691.

[17] Rajasekar N, Kumar NK, Venugopalan R. Bacterial foraging algorithm based solar PV parameter estimation. Sol Energy 2013;97:255–65. https://doi.org/10.1016/j.solener.2013.08.019.

[18] Srihari G, Chandra JK. Parameter extraction of photovoltaic cell using an improved cuckoo search optimization. Sol Energy 2020;204:280–93. https://doi.org/10.1016/j.solener.2020.04.036.

[19] Alam DF, Yousri DA, Eteiba MB. Flower Pollination Algorithm based solar PV parameter estimation. Energy Convers Manage 2015;101:410–22. https://doi.org/10.1016/j.enconman.2015.05.074.

[20] Patel SJ, Panchal AK, Kheraj V. Extraction of solar cell parameters from a single current–voltage characteristic using teaching learning based optimization algorithm. Appl Energy 2014;119:384–93. https://doi.org/10.1016/j.apenergy.2014.01.027.

[21] Jiang LL, Maskell DL, Patra JC. Parameter estimation of solar cells and modules using an improved adaptive differential evolution algorithm. Appl Energy 2013;112:185–93. https://doi.org/10.1016/j.apenergy.2013.06.004.

[22] Niu Q, Zhang L, Li K. A biogeography-based optimization algorithm with mutation strategies for model parameter estimation of solar and fuel cells. Energy Convers Manage 2014;86:1173–85. https://doi.org/10.1016/j.enconman.2014.06.026.

[23] Chen X, Yu K, Du W, Zhao W, Liu G. Parameters identification of solar cell models using generalized oppositional teaching learning based optimization. Energy 2016;99:170–80. https://doi.org/10.1016/j.energy.2016.01.052.

[24] Ram JP, Babu TS, Dragicevic T, Rajasekar N. A new hybrid bee pollinator flower pollination algorithm for solar PV parameter estimation. Energy Convers Manage 2017;135:463–76.

[25] Xiong G, Zhang J, Yuan X, Shi D, He Y, Yao G. Parameter extraction of solar photovoltaic models by means of a hybrid differential evolution with whale optimization algorithm. Sol Energy 2018;176:742–61. https://doi.org/10.1016/j.solener.2018.10.050.

[26] Xu S, Wang Y. Parameter estimation of photovoltaic modules using a hybrid flower pollination algorithm. Energy Convers Manage 2017;144:53–68. https://doi.org/10.1016/j.enconman.2017.04.042.

[27] Rao RV, Savsani VJ, Vakharia DP. Teaching–learning-based optimization: A novel method for constrained mechanical design optimization problems. Comput Aided Des 2011;43:303–15. https://doi.org/10.1016/j.cad.2010.12.015.

[28] Gray JL. Handbook of Photovoltaic Science and Engineering (John Wiley & Sons); 2010. https://doi.org/10.1002/9780470974704.

[29] Tian H, Mancilla-David F, Ellis K, Muljadi E, Jenkins P. A cell-to-module-to-array detailed model for photovoltaic panels. Sol Energy 2012;86:2695–706. https://doi.org/10.1016/j.solener.2012.06.004.

[30] Holmgren WF, Hansen CW, Mikofski MA. pvlib python: A python package for modeling solar energy systems. J Open Source Software 2018;3:884. https://doi.org/10.21105/joss.00884.

[31] Haeffelin M, Barthès L, Bock O, Boitel C, Bony S, Bouniol D, et al. SIRTA, a ground-based atmospheric observatory for cloud and aerosol research. Ann Geofis 2005;23(2):253–75.

[32] Dubois AM, Badosa J, Bourdin V, Torres Aguilar MI, Bonnassieux Y, Álvarez-Gallegos A. Estimation of the Uncertainty due to Each Step of Simulating the Photovoltaic Conversion under Real Operating Conditions. Int J Photoenergy 2021;2021:1–14.

[33] Vázquez M, Stolle I. Photovoltaic module reliability model based on field degradation studies. Prog Photovoltaics Res Appl 2008;16:419–33. https://doi.org/10.1002/pip.825.

[34] Jordan DC, Wohlgemuth JH, Kurtz SR. Technology and Climate Trends in PV Module Degradation. 27th European Photovoltaic Solar Energy Conference and Exhibition; 2012. p. 3118–3124. https://doi.org/10.4229/27thEUPVSEC2012-4DO.5.1.

[35] Hernday P. Interpreting I-V Curve Deviations. SolarPro; 2014. Retrieved from http://resources.solmetric.com/get/SolarPro%20I-V%20feature%20article,%20Hernday.pdf.

[36] Florides M, Makrides G, Georghiou G.E. Characterisation of the Shunt Resistance due to Potential Induced Degradation (PID) in Crystalline Solar Cells. IEEE 7th World Conference on Photovoltaic Energy Conversion; 2018. p. 0695–0699. https://doi.org/10.1109/PVSC.2018.8547500.