



HAL
open science

A Bayesian Inference Approach to Extract Circuit Model Parameters and Analyze Photovoltaic Degradation from Power Production Data

Joseph Chakar, Jean-Paul Calin, Marko Pavlov, Jordi Badosa, Jean-Baptiste Puel, Yvan Bonnassieux

► To cite this version:

Joseph Chakar, Jean-Paul Calin, Marko Pavlov, Jordi Badosa, Jean-Baptiste Puel, et al.. A Bayesian Inference Approach to Extract Circuit Model Parameters and Analyze Photovoltaic Degradation from Power Production Data. 2024. hal-04238791v2

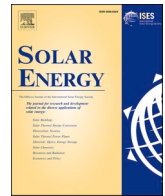
HAL Id: hal-04238791

<https://hal.science/hal-04238791v2>

Preprint submitted on 23 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



A Bayesian inference approach to extract circuit model parameters and analyze photovoltaic degradation from power production data

Joseph Chakar^{a,b,*}, Jean-Paul Calin^a, Marko Pavlov^c, Jordi Badosa^d, Jean-Baptiste Puel^{b,e}, Yvan Bonnassieux^{a,b}

^a Laboratoire de Physique des Interfaces et des Couches Minces (LPICM), Centre National de la Recherche Scientifique (CNRS), Ecole Polytechnique, Institut Polytechnique de Paris (IP Paris), Palaiseau 91120, France

^b Institut Photovoltaïque d'Île-de-France (IPVF), 18 Boulevard Thomas Gobert, 91120 Palaiseau, France

^c Feedgy, Paris 75009, France

^d Ecole Polytechnique, Institut Pierre-Simon Laplace (IPSL), IP Paris, Ecole Normale Supérieure (ENS), Sorbonne Université, CNRS, Laboratoire de Météorologie Dynamique (LMD), Palaiseau 91120, France

^e EDF, R&D, 18 Boulevard Thomas Gobert, 91120 Palaiseau, France

ARTICLE INFO

Keywords:

Solar energy
Photovoltaics
Bayesian inference
Performance analysis
Degradation

ABSTRACT

Power production data can be a valuable resource to analyze photovoltaic (PV) performance without the need for field surveys. Recent work has demonstrated the exciting possibility of leveraging this data to extract circuit model parameters and current-voltage properties of a PV system, but further development is needed to bolster these initial findings. Here, instead of using a classical optimization approach, we switch to the Bayesian framework to solve this complex multi-solution problem. This allows us to construct probability distributions over the model parameters, get a comprehensive picture of the solution space, and quantify prediction uncertainty. As a result, we can define confidence intervals for the system's electrical properties and consistently track their daily evolution. Our results are validated with laboratory measurements for five silicon and thin-film modules, and our scalable approach works with on-site as well as online weather data, which opens new prospects for remote PV monitoring, modeling, and degradation analysis for real-life applications.

1. Introduction

As the world photovoltaic (PV) fleet continues to grow and age, greater attention is being accorded to operations and maintenance (O&M) practices to ensure the technical and economic viability of solar energy systems over their lifetime. In order to meet production thresholds over more than 20 years, it is becoming increasingly critical to effectively monitor these systems to identify and address performance issues quickly. Today, thanks to new data analytics techniques, PV stakeholders can perform insightful remote analyses using limited but easily-accessible data, instead of relying on information-rich but expensive, expert-dependent, and interruptive field surveys.

One valuable online strategy is the modeling of deployed PV systems using operational data, as it allows us to understand deviations from the expected behavior and assess repowering opportunities. In this context, a simple yet effective tool that is often used to represent solar cells as

well as PV modules is the single-diode model (SDM) [13]. Even though the values of its parameters are normally not provided by solar panel manufacturers, several methods exist to estimate them from typical current-voltage (IV) properties.

On the one hand, analytical techniques [32,6,7,8,26,2,33] can fit IV data using mathematical techniques and physical models that describe the behavior of a solar cell. They are generally simple and easy to implement but rely on specific assumptions that may not always hold. On the other hand, numerical techniques [29,17,25] use iterative methods like the Newton-Raphson and Lambert W-function methods to extract model parameters from more complex IV information instead of handpicked data points. They can handle more realistic scenarios that cannot be solved analytically, but they require more computational power and suffer from convergence issues due to their sensitivity to initial conditions.

Alternatively, (meta)heuristic algorithms [44,35,27,31,1,43,36,

* Corresponding author at: Laboratoire de Physique des Interfaces et des Couches Minces (LPICM), Centre National de la Recherche Scientifique (CNRS), Ecole Polytechnique, Institut Polytechnique de Paris (IP Paris), Palaiseau 91120, France.

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

<https://doi.org/10.1016/j.solener.2024.112595>

Received 5 March 2024; Received in revised form 22 April 2024; Accepted 30 April 2024

Available online 8 May 2024

0038-092X/© 2024 The Authors. Published by Elsevier Ltd on behalf of International Solar Energy Society. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

37,42,15,41] – such as the particle swarm, flower pollination, and teaching–learning-based optimization algorithms – offer several advantages over analytical and iterative techniques. They are significantly less sensitive to the initialization process, more robust to noise, and able to handle larger quantities of multi-dimensional IV data (e.g. with temperature and irradiance dependence). Their speed, flexibility, and scalability have made them a popular choice among researchers who adopted them to extract circuit model parameters [40,3,19,22,21] and predict PV performance [28,5,45] from partial or full IV curves measured outdoors. However, despite being a valuable tool for evaluating PV performance, IV curves are rarely measured in actual PV installations, which makes it difficult to implement these studies in real life.

Instead, we recently demonstrated how such techniques can be used to extract the SDM parameters and infer IV curves from typical power production data [4] (i.e. time-series that list the operating current and voltage of a PV system at varying irradiance and temperature) measured by commercial solar inverters and data loggers. Specifically, we used the teaching–learning-based optimization (TLBO) algorithm to find an optimal combination of SDM parameters that can model a monocrystalline silicon (mono c-Si) module using a representative sample of its production data. Similarly, other researchers have proposed different optimization algorithms [38,24] to achieve this using data from the National Renewable Energy Laboratory (NREL) and National Institute of Standards and Technology (NIST).

Albeit promising, these initial findings have several common shortcomings. First, they all rely on classical optimization techniques, which are limited to finding only one “best” fit for the problem at hand. However, based on the desired level of accuracy, complex tasks like fitting the SDM can admit multiple solutions, especially when only maximum power point (MPP) data is available. Therefore, under the same conditions, these algorithms can yield different and often stochastic results depending on their starting point, convergence properties, and tolerance to noise. Moreover, these studies have so far only covered crystalline silicon, which is the predominant but not exclusive

PV module technology. Last but not least, results reported in the literature are often validated using outdoor data, which may not be as reliable as laboratory data obtained under controlled conditions.

Here, we reframe the task of extracting circuit model parameters from production data as a Bayesian inference problem. Using this probabilistic approach, we can explore the parameter search space more intelligently by considering a range of possible outcomes. We can then update these beliefs as more data become available, which enables us to gradually improve the accuracy of our predictions and reduce their uncertainty over time. We also extend our study on monocrystalline silicon to four other silicon and thin-film module technologies, test our new method using both on-site and satellite-based weather data, and validate our results using laboratory measurements.

2. Methodology

2.1. Data

We rely on the data collected at the SIRTA (Site Instrumental de Recherche par Télédétection Atmosphérique) observatory’s [16] PV test bench [9] located at Ecole Polytechnique in Palaiseau, France, and illustrated in Fig. 1. The test bench was installed in 2014 and hosts five free-standing South-oriented commercial solar panels of different technologies. The nameplate IV properties of these modules at Standard Test Conditions (STC) of 1000 W/m² and 25 °C are reported in Table 1.

A Class A CMP22 pyranometer (Kipp & Zonen) measures the Plane-of-Array (POA) irradiance, and four-wired class A platinum sensors (PT100) are used to measure the backsheet temperature of the modules. Full outdoor IV curve measurements are recorded by Chroma electronic loads roughly every 30 s, but we demonstrate Bayesian inference’s capability to extract circuit model parameters using only MPP data. By adopting this approach, we can quantify the uncertainty associated with this process, and as we accumulate more data over time, we can progressively narrow down this uncertainty, steering away from the stochastic outcomes typical of traditional optimization techniques [24].

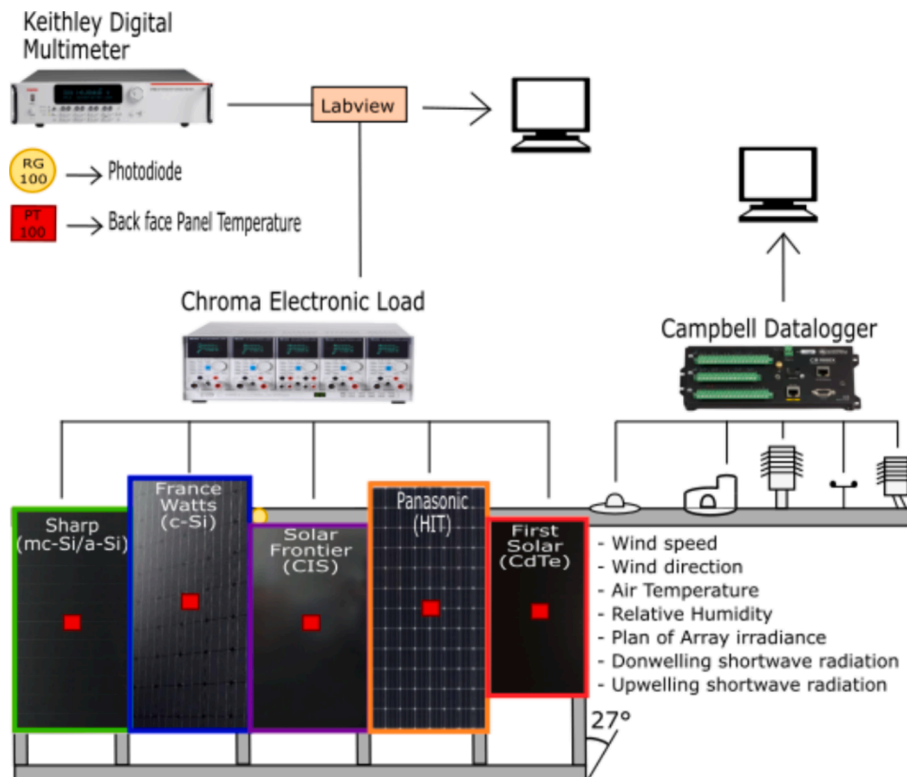


Fig. 1. Schematic of the SIRTA PV Test Bench.

Table 1

Current and voltage properties of the installed modules at Standard Testing Conditions (STC). From left to right: output power (P_{MPP}), voltage (V_{MPP}), and current (I_{MPP}) at the maximum power point; open-circuit voltage (V_{oc}); and short-circuit current (I_{sc}). From top to bottom: tandem amorphous/microcrystalline silicon (a-Si/ μ -Si), monocrystalline silicon (c-Si), Copper Indium Selenide (CIS), mono c-Si Heterojunction with amorphous silicon Intrinsic Thin layer (HIT), and Cadmium Telluride (CdTe).

	P_{MPP} [W]	V_{MPP} [V]	I_{MPP} [A]	V_{oc} [V]	I_{sc} [A]
a-Si/ μ -Si module	128 (+10 %/−5 %)	45.40	2.82	59.8	3.45
c-Si module	250 (± 3 %)	30.52	8.21	37.67	8.64
CIS module	150 (+10 %/−5 %)	81.5	1.85	108.0	2.20
HIT	240 (+10 %/−5 %)	43.7	5.51	52.4	5.85
CdTe	82.5 (± 10 %)	48.30	1.71	60.80	1.94

Since flash tests were performed for all the modules in early April 2022, we consider the three-month period preceding them (i.e. since the beginning of the year). We then use these lab measurements to validate our results. A full description of the measured data and installed equipment is available on the open-access GitLab repository of the test bench (see Data & Code Availability). More information on data preprocessing can also be found in the [Supplemental Information](#).

2.2. Equivalent circuit model

The single-diode model (SDM) is a simple physics-based model [13] that is commonly used to describe the behavior of a solar cell. Although solar cells are the primary power generation unit, PV modeling is usually done at the module level to reduce computational complexity and because solar panel specifications are more accessible. Assuming all the N_s series-connected cells of a PV module to be identical and under uniform and equal irradiance and temperature (i.e. they generate the same current), the governing equation [39] of the SDM can be written as:

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

Given a set of operating conditions (i.e. solar irradiance and cell temperature), the SDM depends on five parameters that relate the module's output current I_M to its operating voltage V_M : the light-induced current (I_L), saturation current (I_o), series resistance (R_s), shunt resistance (R_{sh}), and ideality factor (n). $V_T = \frac{k_B T_c}{q}$ is the thermal voltage of the cell, where k_B is Boltzmann's constant, T_c the cell temperature, and q the elementary charge.

Typically, manufacturers do not provide values for the SDM parameters of PV modules. However, reference values at STC can be estimated from classical datasheet properties using the California Energy Commission (CEC) method [8]. This method introduces a sixth parameter to the SDM, which is simply a correction factor for the temperature coefficient of the short-circuit current. It also relies on the conventional De Soto model [7] to adjust the five SDM parameters to the operating conditions (i.e. irradiance and temperature). As in [4], we leverage the *pvl* Python package [10] to compute these models and also factor in the series resistance's dependence on irradiance.

Once the SDM parameter values are calculated at predetermined conditions, the module's IV response can be simulated to determine its key properties – namely the DC output current (I_{MPP}), voltage (V_{MPP}), and power (P_{MPP}) at the MPP, as well as the open-circuit voltage (V_{oc}) and short-circuit current (I_{sc}). The simulated IV properties can then be compared with the measured data to evaluate how accurately the SDM parameters can model the PV module (or array).

2.3. Bayesian inference approach

Instead of using a classical optimization algorithm to determine the

single “best” parameter values for the SDM, we opt for the Bayesian point of view. This allows us to get probability distributions over every parameter, find a set of parameter combinations that can model the PV system at hand, and quantify uncertainty over time. This approach is based on Bayes' theorem, which lets us update our prior beliefs about the parameters $P(\theta)$ as more data d becomes available.

$$P(\theta|d) = \frac{P(\theta)P(d|\theta)}{P(d)} \quad (2)$$

We first define the parameter search space based on the module's datasheet specifications and energy production during the analysis period. We assume all the possible parameter combinations (hypotheses) to be initially equiprobable then calculate the likelihood $P(d|\theta_i)$ of each parameter combination by comparing its corresponding MPP current and voltage with the module's actual output current and voltage at the same measured operating conditions. The normalization constant $P(d)$ being the sum of $P(\theta)P(d|\theta)$ over all the hypotheses, we can compute the posterior $P(\theta|d)$ using Bayes' theorem. The full procedure is detailed in the [Supplemental Information](#) and the code is publicly available on GitLab (see Data & Code Availability).

If we repeat this process on a regular basis while transferring the knowledge gained over time (i.e. by setting the posterior distribution obtained on the first day as the prior distribution of the following one and so on), we can progressively narrow down the 90 % confidence interval for our predictions and monitor their evolution over time. This temporal transmission of information serves as a safeguard against stochastic outcomes that do not reflect real physical daily or seasonal variability – a common pitfall of conventional optimization techniques found in the literature – thereby promoting stability and consistency in the results.

3. Results

3.1. Estimation of circuit model parameters

We first consider the mono c-Si module. Starting with the first day of the analysis period, we apply our Bayesian inference process on a daily basis (without any distinction between sunny and cloudy days) until the end of this 90-day time frame.

[Fig. 2](#) shows the posterior distributions of the module's SDM parameters on the last day of the analysis period (i.e. the probability mass functions on March 31st, 2022), which in this case reflect the state of the PV module after roughly eight years of operation. The orange dashed line represents the reference STC value estimated from the module datasheet properties using the CEC estimation method [8], and the shaded area accounts for the power tolerance specified by the manufacturer (± 3 % in this case). In other terms, this area shows how much each parameter would need to change to increase (or decrease) the module's rated power by 3 %, assuming the remaining parameters are kept constant. Note that this module's output power becomes insensitive to changes in the shunt resistance once the latter's value exceeds a few hundred ohms.

In the posterior distribution of I_L , the shift of the peak from the initial datasheet estimate indicates that the module's degradation is most likely induced by a drop in its light-induced and thus short-circuit current. This observation is consistent with field experience, which suggests that long-term current loss frequently impacts crystalline silicon modules [20,30]. Moreover, it appears that the dark saturation current I_o and series resistance R_s have not degraded over time, while the ideality factor n remains around 1. These observations reaffirm the ones we previously made [4] – even if the analysis periods are roughly one year apart – as this rather stable mono c-Si module has been well maintained since it was installed. In contrast, even though the expected value of R_{sh} is below its estimated reference value, there is more uncertainty around this parameter. Indeed, the shunt resistance's impact on the c-Si module's

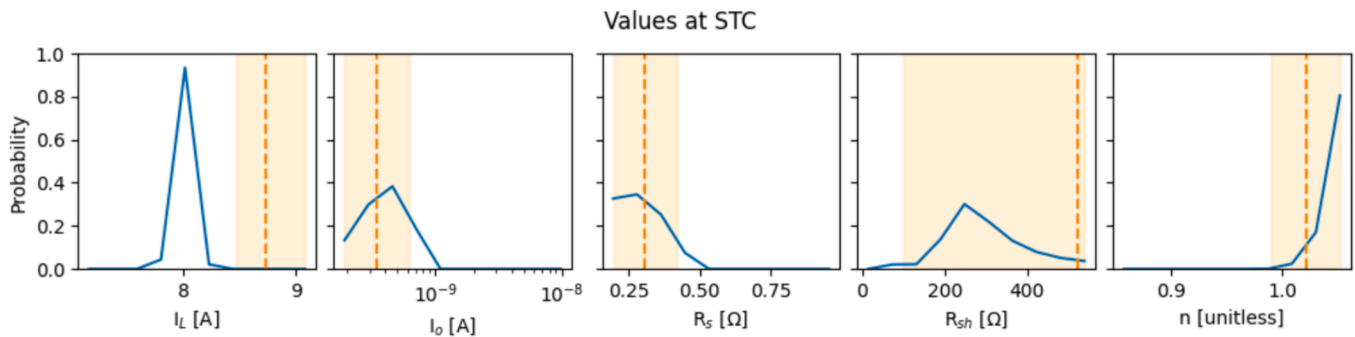


Fig. 2. Probability distributions of the single-diode model parameters at Standard Testing Conditions (STC) at the end of the analysis period. The dashed line represents the value estimated from the datasheet properties. The shaded area accounts for the datasheet power tolerance. From left to right: light-induced current (I_L), diode reverse saturation current (I_0), series resistance (R_s), shunt resistance (R_{sh}), and ideality factor (n).

electrical output is imperceptible once it crosses a certain threshold. This underscores the necessity of adopting a probabilistic point of view when dealing with complex models that display varying sensitivity to their correlated parameters and can yield identical outcomes across different parameter combinations – especially when there is limited data and sizable manufacturing variability.

Though the resulting drop in performance is in this instance well within the manufacturer’s performance warranty, these plots can provide us valuable insights into the module’s degradation modes. For example, a lower-than-expected value of I_{sc} can generally be an indication of uniform soiling or delamination, while shunt resistance degradation may be a symptom of Potential Induced Degradation (PID) [18]. One of the main advantages of switching to the Bayesian framework is that it enables us to identify a set of possible solutions by exploring the full parameter search space. In comparison, traditional heuristic algorithms used in prior work [44,35,27,31,1,36,43,37,38,42,15,41,4,24] are often designed to find the single best result using a fixed set of rules. These algorithms generally work well for deterministic problems where the outcome of a given input is always the same, which is not the case for extracting the parameters of a complex circuit model using only MPP data.

3.2. Estimation of current-voltage properties

Since we know the probability of each parameter combination included in our search space, we can monitor the evolution of the 90 % confidence interval for their associated module IV properties throughout the analysis period (see Fig. 3). This way, we can better understand how changes in the SDM parameters translate to changes in the module behavior (and vice versa) in case of a steady or sudden degradation.

Naturally, the size of the confidence interval starts large and then progressively shrinks as more production data becomes available. After analyzing several days of production data, the 90 % probability mass becomes confined to less than 5 % of the search space. Towards the end of the analysis period, it sits in less than 1 % of this space. Nevertheless, as we do not distinguish between cloudy and sunny days, occasional day-to-day fluctuations and outliers can occur.

Interestingly, the voltage converges to higher values, and the opposite is true for the current. This reaffirms our previous observation that this module’s power degradation is the result of a drop in the current rather than the voltage. Furthermore, there is more confidence around the values at the MPP than at the open-circuit voltage V_{oc} and short-circuit current I_{sc} . This comes as no surprise since we can pair the two latter points with different R_s and R_{sh} values to get IV curves that have different shapes but pass through the same MPP. Due to this inherent difficulty in determining unique values for the SDM parameters and IV properties from MPP data alone, confidence intervals should be considered to capture the estimation uncertainty. Given that no PV degradation occurs during our analysis period, the changes in the size

and trend lines of the confidence intervals seen during this short time-frame are not indicative of any physical changes and only reflect the evolution of the uncertainty surrounding the module’s predicted IV properties and SDM parameters (see Supplemental Information). In more complex cases involving long-term degradation, we would anticipate a gradual shift in trend lines, whereas an abrupt degradation event would likely prompt a sudden change in the size of the confidence interval followed by a gradual change in trend lines.

Besides quantifying uncertainty, another major advantage of using Bayesian inference is that it allows us to study short as well as long time frames and progressively update the output in a speedy yet consistent manner. As we previously noted [4], it would be difficult to obtain such results using classical optimization algorithms, since they are designed to find a single best and independent fit for each batch of data. This becomes even more evident once we look at Li et al.’s results [24], which exhibit seasonal variability and stochastic patterns that are atypical of rather stable silicon modules and can thus yield different and perhaps misleading conclusions depending on the length and size of the analysis window. Bayesian inference overcomes this issue because it continuously updates its (prior) beliefs about the system using the best consecutive data fits (i.e. the computed likelihoods).

As for the SDM parameters, we can inspect the distribution of the predicted module IV properties (see Fig. 4) to better understand the uncertainty surrounding their values at the end of the analysis period.

At the end of the analysis period, the posterior distributions of the predicted module IV properties are concentrated over a small portion of the initial SDM parameter search space. The 90 % confidence intervals for the module’s output power P_{MPP} , voltage V_{MPP} , and current I_{MPP} at STC are within ± 2.2 W, ± 0.5 V, and ± 0.13 A from their respective expected value, which corresponds to impressive precision levels around

± 1 %. However, the precision is expectedly lower for the open-circuit voltage (± 3.5 %) and short-circuit current (± 2.6 %). Indeed, the presence of multiple peaks shows how different values of V_{oc} and I_{sc} exist to get the same MPP predictions. These findings reaffirm the above observations and show that module degradation is in this case linked to a drop in current, as the voltage remains stable over time.

3.3. Flash test validation

Having performed flash tests shortly after the analysis finish date, we can evaluate the accuracy of the predicted module IV properties (at STC) by comparing their expected values (i.e. weighted averages) with the results of these indoor measurements.

The results of Table 2 show the degradation trends to be predicted correctly. However, it appears that the voltage tends to be overestimated while the current is more noticeably underestimated, especially at the MPP. This is in reality due to the outdoor equipment’s underestimation of the modules’ short-circuit and MPP currents (and thus their power

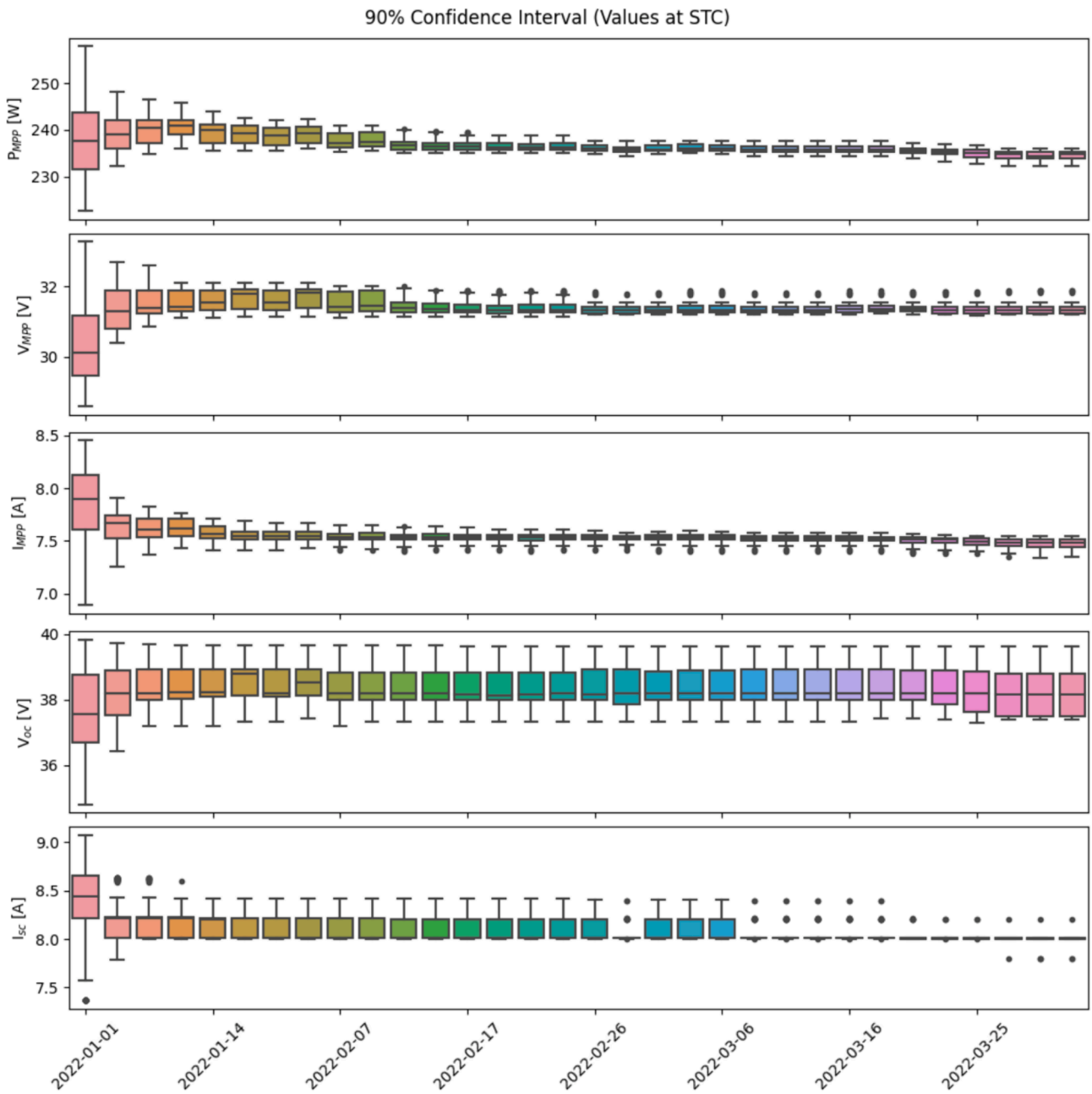


Fig. 3. Evolution of the 90 % confidence intervals of the module’s current–voltage properties at Standard Testing Conditions (STC) throughout the analysis period. Black dots represent less likely outlier values. The colors of the boxes and number of days shown carry no significance and are set solely for visualization purposes. From top to bottom: output power (P_{MPP}), voltage (V_{MPP}), and current (I_{MPP}) at the maximum power point; open-circuit voltage (V_{oc}); and short-circuit current (I_{sc}).

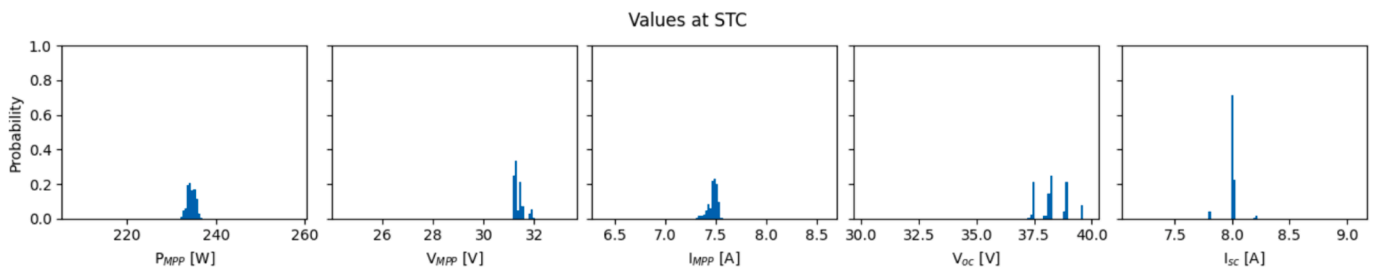


Fig. 4. Probability distributions of the predicted module current–voltage properties at Standard Testing Conditions (STC) at the end of the analysis period. From left to right: output power (P_{MPP}), voltage (V_{MPP}), and current (I_{MPP}) at the maximum power point; open-circuit voltage (V_{oc}); and short-circuit current (I_{sc}).

Table 2

Comparison between the predicted and lab-measured current–voltage properties. From top to bottom: output power (P_{MPP}), voltage (V_{MPP}), and current (I_{MPP}) at the maximum power point; open-circuit voltage (V_{oc}); and short-circuit current (I_{sc}).

	Datasheet	Measured (Flash test)	Predicted (Weighted average)	Predicted vs. Measured
P_{MPP} [W]	250 (± 3 %)	243.94	234.5	-3.87 %
V_{MPP} [V]	30.52	30.47	31.37	2.98 %
I_{MPP} [A]	8.21	8.01	7.47	-6.70 %
V_{oc} [V]	37.67	37.56	38.28	1.93 %
I_{sc} [A]	8.64	8.49	8.00	-5.77 %

outputs), as the values recorded outdoors at high irradiance (above STC) are consistently below those obtained from the flash tests performed in the lab at STC (see [Supplemental Information](#)). The measurement uncertainty is expected to be around ± 7.2 % for the site measurements [9], but more tests are needed to quantify this particular error and determine whether it caused by the equipment measuring the IV curves and/or the pyranometer overestimating the irradiance. Nonetheless, our analysis successfully detects the decrease in output current and overall performance, regardless of whether it is due to system faults and/or module degradation. It also provides a glimpse into the heightened complexity of remote performance and degradation analysis, as diverse and concurrent factors can contribute to the same observed phenomena.

3.4. Results for other module technologies

Since we have the flash test results for all the modules installed at the PV test bench, we can also check how accurately and precisely this approach can estimate the IV properties of the four remaining solar panel technologies, namely: tandem amorphous/microcrystalline silicon (a-Si/ μ -Si), Copper Indium Selenide (CIS), mono c-Si Heterojunction with amorphous silicon Intrinsic Thin layer (HIT), and Cadmium Telluride (CdTe).

Overall, the relative errors of the expected values (colored squares in [Fig. 5](#)) follow similar trend lines across the different PV technologies and are still within the measurement uncertainty of the power production data, which would increase if we also consider the meteorological data uncertainty. All the predicted MPP properties have a narrow 90 % confidence interval (black error bars), but the output power and current are consistently underestimated. Voltage is in contrast overestimated, which explains why P_{MPP} is slightly less underestimated than I_{MPP} is. This is mostly due to the outdoor equipment limiting the modules' output currents (i.e. our input data), as explained earlier.

The error bars are considerably but expectedly higher for the thin-

film modules. This is not only due to the higher modeling errors (see [Supplemental Information](#)) but also a difference in the manufacturer power tolerance (c.f. [Table 1](#)) and sensitivity to the different SDM parameters. While the flash tests were performed twice and have a low measurement uncertainty of ± 0.24 %, their setup is still mostly adapted for more prevalent silicon modules. The accuracy of the lab results obtained for the thin-film modules may thus be affected by a mismatch with the solar simulator's spectrum and/or mono c-Si reference cell used.

Despite these shortcomings, the average prediction error is below 5 % across all five technologies. Taking the power tolerance and measurement uncertainty into account, we can say that the SDM offers a good compromise between simplicity and accuracy when extracting module IV properties from power production data.

3.5. Sensitivity analysis

Knowing that weather data is seldom measured in the field, we report the results we get when we replace site measurements for irradiance and temperature with online satellite estimates using the same databases [23,34,12,14] and thermal model [11] described in our previous work [4] (see [Fig. 6](#)). This adds another layer of uncertainty to the input data but helps keep this research project grounded in real life.

At first sight, the error profiles closely resemble those obtained when onsite weather data is used. The main difference is that the power estimation error is lower here because the POA irradiance estimated using satellite data is lower than the one measured by the pyranometer (especially at high irradiance) [4], which results in higher values for the predicted output power and current.

Considering that the choice of weather database and physical models is not optimized for this PV test bench, these promising results attest to the robustness of this practical approach, which can be scaled to real PV systems of any size and technology.

4. Discussion

The ability to effectively extract circuit model parameters and IV properties from readily available power production and weather data is a powerful remote PV monitoring and modeling tool with a wide range of benefits for PV stakeholders. It allows system operators to continuously monitor important PV characteristics and detect potential issues early on, thus minimizing downtime and maximizing energy production. On the one hand, analyzing the evolution of the extracted circuit model parameters can provide valuable insights into the underlying physical degradation mechanisms, which helps system operators make more informed decisions remotely and reduce the need for expensive site inspections. On the other hand, DC modeling can be coupled with AC modeling to accurately simulate the future performance of grid-

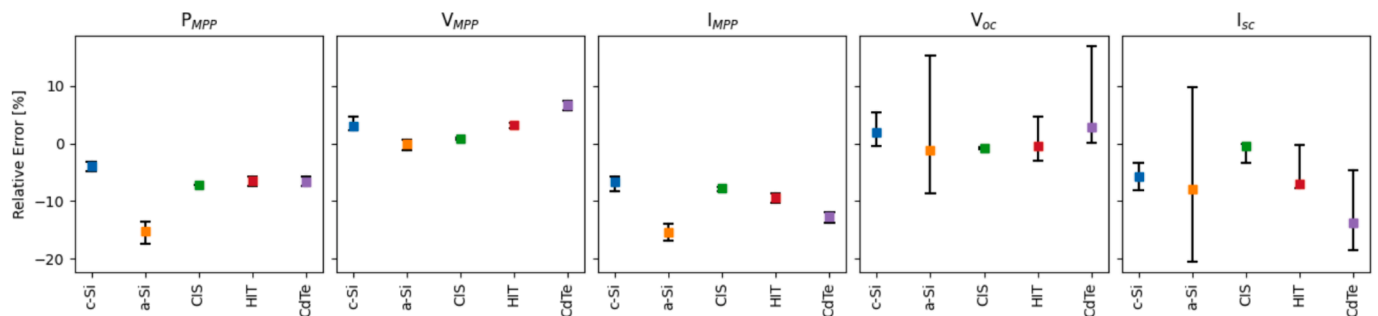


Fig. 5. Relative error between the predicted and measured current–voltage properties of all the installed modules at Standard Testing Conditions (STC). The colored squares represent the error for the expected value and the error bars correspond to the minimum and maximum errors of the 90 % confidence interval. From left to right on the title axis: output power (P_{MPP}), voltage (V_{MPP}), and current (I_{MPP}) at the maximum power point; open-circuit voltage (V_{oc}); and short-circuit current (I_{sc}). From left to right on the x-axis: monocrystalline silicon (c-Si), tandem amorphous/microcrystalline silicon (a-Si/ μ -Si), Copper Indium Selenide (CIS), mono c-Si Heterojunction with amorphous silicon Intrinsic Thin layer (HIT), and Cadmium Telluride (CdTe).

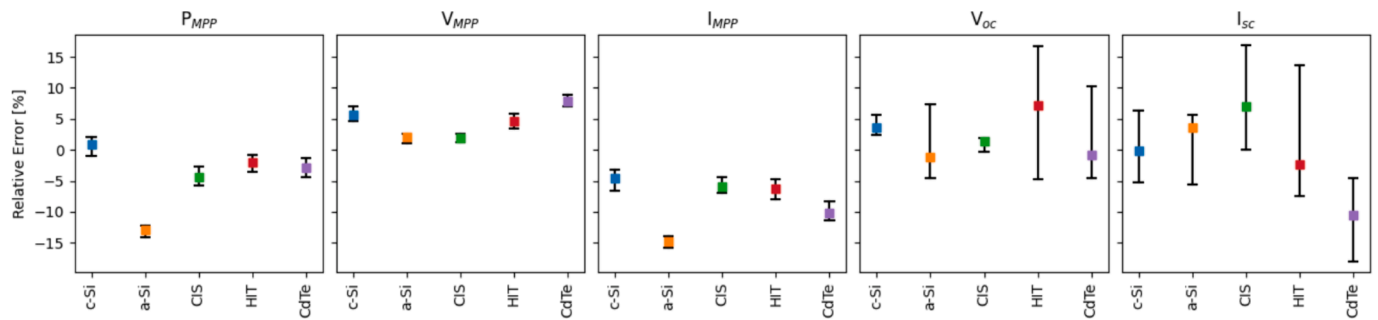


Fig. 6. Relative error between the predicted and measured current–voltage properties of all the installed modules at Standard Testing Conditions (STC) (using online weather data).

connected PV systems and assess the economic viability of replacing impacted modules and/or inverters for repowering applications.

The key advantage of using the probabilistic Bayesian framework lies in its ability to uncover all possible degradation scenarios and quantify prediction uncertainty. This ensures that no potential outcome is overlooked and provides a means to evaluate the variability surrounding expected returns on investment when making business decisions (e.g. for repowering applications). As previously stated, the existing literature predominantly relies on deterministic approaches that are ill-suited for this multi-solution problem. Such conventional optimization techniques ignore alternative possibilities and focus on finding a single optimal set of SDM parameter values based on predefined conditions, which likely translates to the oversight of critical degradation mechanisms in real-world scenarios.

Whether they are related to the modeling error or breadth of the search space, this approach's aforementioned limitations can be addressed to improve the quality of the results. The SDM is widely used to describe the behavior of silicon solar cells, which are the dominant PV technology, but one may consider alternative models to get improved accuracy and physical interpretation, particularly for thin-film technologies. In our proposed Bayesian inference method, we evaluate the same set of parameter combinations for every day of the analysis period, and the initial reference parameter values extracted from the solar panel datasheet strongly influence the combinations that constitute our search space. One way to enhance the final results is to refine the initial parameter search and rerun this analysis over and over until a predefined convergence criterion is met. Or, instead of having a limited number of hypotheses (i.e. a discrete parameter space), we can use an acquisition function to explore a continuous parameter space. Acquisition functions can balance global exploration (e.g. exploring solutions with larger shunt resistance values for the c-Si module) and local exploitation (e.g. refining the value of the light-induced current) of the search space by relying on a probabilistic model (typically a Gaussian process) that approximates the objective function (the SDM here). While these functions help reduce the number of evaluated hypotheses by exploring the search space more efficiently, they come with a heavy computational cost when compared with *pvliv*'s implementation of the SDM. Furthermore, this adaptive parameter search space strategy typical of Bayesian optimization would take away the ability to carry posterior distributions over time. In any case, there is only little room to improve the accuracy and precision of the V_{oc} and I_{sc} predictions since we are limited to MPP data. Ultimately, the choice of model and search space exploration strategy depends on the desired level of accuracy and computational efficiency for the application at hand.

Last but not least, while our approach effectively detected that our PV system suffers from current losses, it would not have been possible to determine which system component was responsible for the overall performance loss without performing flash tests for the modules. Indeed, the discrepancy between the predicted and lab-measured module IV properties does not disprove our predictions but instead confirms that

our system losses stem from equipment rather than module degradation. This simple case highlights the complexity of interpreting overall performance losses and identifying which system components are causing changes in the equivalent circuit model parameters extracted from MPP data. This said, additional efforts are needed to test this approach on larger PV systems, where partial shading, current mismatch, and more complex problems are also prevalent.

5. Conclusion

In this work, we present an innovative approach to extract the single-diode model parameters of a PV system from commoditized production and weather data. We shift from the conventional approaches of the literature to the more suitable Bayesian point of view, which allows us to identify a set of possible degradation mechanisms along with their associated probabilities. By automating the definition of the search space according to the system specifications and incorporating a transfer learning strategy, we steer clear of the stochastic outcomes of classical optimization techniques and strengthen the precision and interpretability of the predicted physical properties over short and long time frames. Using outdoor measurements and indoor flash tests for silicon and thin-film modules, we demonstrate our method's ability to correctly predict degradation patterns and highlight the challenges associated with deciphering the results of remote performance analyses. Moreover, our broadly-applicable Bayesian inference approach can be scaled to PV systems of any size and technology and works equally well with on-site weather data and satellite estimates, thus offering promising opportunities for real-life applications.

6. Experimental procedures

6.1. Data & code availability

The data and code that support the findings of this study are publicly available on GitLab <https://gitlab.in2p3.fr/energy4climate/public/sirta-pv1-data>.

6.2. Flash tests

The flash tests were performed at TotalEnergies laboratory facilities using a CetusPV-XF2-M Xenon flasher. Flash duration is 58 ms, and average results are reported at 1000 W/m² and 25 °C for four IV curve measurements (one forward and one reverse sweep per curve). The ± 0.24 % measurement uncertainty is based on a statistical process control (SPC) obtained by continuously measuring the MPP power output of a certified mono c-Si reference module.

CRediT authorship contribution statement

Joseph Chakar: Conceptualization, Data curation, Formal analysis,

Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Jean-Paul Calin**: Visualization, Writing – review & editing. **Marko Pavlov**: Conceptualization, Funding acquisition, Supervision, Writing – review & editing. **Jordi Badosa**: Data curation, Funding acquisition, Supervision, Writing – review & editing. **Jean-Baptiste Puel**: Supervision, Writing – review & editing. **Yvan Bonnassieux**: Funding acquisition, Supervision, Writing – review & editing.

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 Ph.D. candidate Moira Torres for helping access the data and flash test results 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 is also produced within the framework of the Energy4Climate Interdisciplinary Center (E4C) of IP Paris and Ecole des Ponts ParisTech. It is 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 Energie 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.solener.2024.112595>.

References

- [1] D.F. Alam, D.A. Younsi, M.B. Eteiba, Flower pollination algorithm based solar PV parameter estimation, *Energy Convers. Manag.* 101 (2015) 410–422, <https://doi.org/10.1016/j.enconman.2015.05.074>.
- [2] E.I. Batzelis, S.A. Papatthanassiou, A method for the analytical extraction of the single-diode PV model parameters, *IEEE Trans. Sustain. Energy* 7 (2016) 504–512, <https://doi.org/10.1109/TSTE.2015.2503435>.
- [3] A.A. Cardenas, M. Carrasco, F. Mancilla-David, A. Street, R. Cardenas, Experimental parameter extraction in the single-diode photovoltaic model via a reduced-space search, *IEEE Trans. Ind. Electron.* 64 (2017) 1468–1476, <https://doi.org/10.1109/TIE.2016.2615590>.
- [4] J. Chakar, M. Pavlov, Y. Bonnassieux, J. Badosa, Determining solar cell parameters and degradation rates from power production data, *Energy Convers. Manag.* X 15 (2022) 100270, <https://doi.org/10.1016/j.ecmx.2022.100270>.
- [5] Z. Chen, Y. Chen, L. Wu, S. Cheng, P. Lin, L. You, Accurate modeling of photovoltaic modules using a 1-D deep residual network based on I-V characteristics, *Energy Convers. Manag.* 186 (2019) 168–187, <https://doi.org/10.1016/j.enconman.2019.02.032>.
- [6] M.A. de Blas, J.L. Torres, E. Prieto, A. Garcia, Selecting a suitable model for characterizing photovoltaic devices, *Renew. Energy* 25 (2002) 371–380, [https://doi.org/10.1016/S0960-1481\(01\)00056-8](https://doi.org/10.1016/S0960-1481(01)00056-8).
- [7] W. De Soto, S.A. Klein, W.A. Beckman, Improvement and validation of a model for photovoltaic array performance, *Sol. Energy* 80 (2006) 78–88, <https://doi.org/10.1016/j.solener.2005.06.010>.
- [8] A.P. Dobos, An improved coefficient calculator for the California energy commission 6 parameter photovoltaic module model, *J. Sol. Energy Eng.* 134 (2012) 021011, <https://doi.org/10.1115/1.4005759>.
- [9] A.M. Dubois, J. Badosa, V. Bourdin, M.I. Torres Aguilar, Y. Bonnassieux, Estimation of the uncertainty due to each step of simulating the photovoltaic conversion under real operating conditions, *Int. J. Photoenergy* 2021 (2021) 1–14, <https://doi.org/10.1155/2021/4228658>.
- [10] F. Holmgren, W. W. Hansen, C. A. Mikofski, M., 2018. pvlby python: a python package for modeling solar energy systems. *J. Open Source Softw.* 3, 884. doi: 10.21105/joss.00884.
- [11] D. Faiman, Assessing the outdoor operating temperature of photovoltaic modules, *Prog. Photovolt. Res. Appl.* 16 (2008) 307–315, <https://doi.org/10.1002/pip.813>.
- [12] R. Gelaro, W. McCarty, M.J. Suárez, R. Todling, A. Molod, L. Takacs, C.A. Randles, A. Darmenov, M.G. Bosilovich, R. Reichle, K. Wargan, L. Coy, R. Cullather, C. Draper, S. Akella, V. Buchard, A. Conaty, A.M. da Silva, W. Gu, G.-K. Kim, R. Koster, R. Lucchesi, D. Merkova, J.E. Nielsen, G. Partyka, S. Pawson, W. Putman, M. Rienecker, S.D. Schubert, M. Sienkiewicz, B. Zhao, The modern-era retrospective analysis for research and applications, version 2 (MERRA-2), *J. Clim.* 30 (2017) 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- [13] Gray, J.L., 2005. The Physics of the Solar Cell, in: Luque, A., Hegedus, S. (Eds.), *Handbook of Photovoltaic Science and Engineering*. John Wiley & Sons, Ltd, Chichester, UK, pp. 61–112. doi: 10.1002/0470014008.ch3.
- [14] B. Gschwind, L. Wald, P. Blanc, M. Lefèvre, M. Schroedter-Homscheidt, A. Arola, Improving the McClear model estimating the downwelling solar radiation at ground level in cloud-free conditions – McClear-v3, *Meteorol. z.* 28 (2019) 147–163, <https://doi.org/10.1127/metz/2019/0946>.
- [15] S. Gude, K.C. Jana, Parameter extraction of photovoltaic cell using an improved cuckoo search optimization, *Sol. Energy* 204 (2020) 280–293, <https://doi.org/10.1016/j.solener.2020.04.036>.
- [16] M. Haefelin, L. Barthès, O. Bock, C. Boitel, S. Bony, D. Bouniol, H. Chepfer, M. Chiriaco, J. Cuesta, J. Delanoë, P. Drobinski, J.-L. Dufresne, C. Flamant, M. Grall, A. Hodzic, F. Hourdin, F. Lapouge, Y. Lemaître, A. Mathieu, Y. Morille, C. Naud, V. Noël, W. O'Hirok, J. Pelon, C. Pietras, A. Protat, B. Romand, G. Scialom, R. Vautard, SIRTA, a ground-based atmospheric observatory for cloud and aerosol research, *Ann. Geophys.* 23 (2005) 253–275, <https://doi.org/10.5194/angeo-23-253-2005>.
- [17] K. Ishibashi, Y. Kimura, M. Niwano, An extensively valid and stable method for derivation of all parameters of a solar cell from a single current-voltage characteristic, *J. Appl. Phys.* 103 (2008) 094507, <https://doi.org/10.1063/1.2895396>.
- [18] Itrec, F., 2014. Reprinted with permission from SolarPro.
- [19] Jones, C.B., Hansen, Clifford.W., 2019. Single Diode Parameter Extraction from In-Field Photovoltaic I-V Curves on a Single Board Computer, in: 2019 IEEE 46th Photovoltaic Specialists Conference (PVSC). Presented at the 2019 IEEE 46th Photovoltaic Specialists Conference (PVSC), pp. 0382–0387. doi: 10.1109/PVSC40753.2019.8981330.
- [20] D.C. Jordan, T.J. Silverman, J.H. Wohlgemuth, S.R. Kurtz, K.T. VanSant, Photovoltaic failure and degradation modes, *Prog. Photovolt. Res. Appl.* 25 (2017) 318–326, <https://doi.org/10.1002/pip.2866>.
- [21] H. Kalliojärvi, K. Lappalainen, S. Valkealahti, Feasibility of photovoltaic module single-diode model fitting to the current-voltage curves measured in the vicinity of the maximum power point for online condition monitoring purposes, *Energies* 15 (2022) 9079, <https://doi.org/10.3390/en15239079>.
- [22] K. Lappalainen, P. Manganiello, M. Piliouguine, G. Spagnuolo, S. Valkealahti, Virtual sensing of photovoltaic module operating parameters, *IEEE J. Photovolt.* 10 (2020) 852–862, <https://doi.org/10.1109/JPHOTOV.2020.2972688>.
- [23] M. Lefèvre, A. Oumbe, P. Blanc, B. Espinar, B. Gschwind, Z. Qu, L. Wald, M. Schroedter-Homscheidt, C. Hoyer-Klick, A. Arola, A. Benedetti, J.W. Kaiser, J.-J. Morcrette, McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions, *Atmospheric Meas. Tech.* 6 (2013) 2403–2418, <https://doi.org/10.5194/amt-6-2403-2013>.
- [24] B. Li, T. Karin, B.E. Meyers, X. Chen, D.C. Jordan, C.W. Hansen, B.H. King, M. G. Deceglie, A. Jain, Determining circuit model parameters from operation data for PV system degradation analysis: PVPRO, *Sol. Energy* 254 (2023) 168–181, <https://doi.org/10.1016/j.solener.2023.03.011>.
- [25] Y. Li, W. Huang, H. Huang, C. Hewitt, Y. Chen, G. Fang, D.L. Carroll, Evaluation of methods to extract parameters from current-voltage characteristics of solar cells, *Sol. Energy* 90 (2013) 51–57, <https://doi.org/10.1016/j.solener.2012.12.005>.
- [26] V. Lo Brano, G. Ciulla, An efficient analytical approach for obtaining a five parameters model of photovoltaic modules using only reference data, *Appl. Energy* 111 (2013) 894–903, <https://doi.org/10.1016/j.apenergy.2013.06.046>.
- [27] Q. Niu, L. Zhang, K. Li, A biogeography-based optimization algorithm with mutation strategies for model parameter estimation of solar and fuel cells, *Energy Convers. Manag.* 86 (2014) 1173–1185, <https://doi.org/10.1016/j.enconman.2014.06.026>.
- [28] H.G.G. Nunes, J.A.N. Pombo, S.J.P.S. Mariano, M.R.A. Calado, J.A.M. Felipe de Souza, A new high performance method for determining the parameters of PV cells and modules based on guaranteed convergence particle swarm optimization, *Appl. Energy* 211 (2018) 774–791, <https://doi.org/10.1016/j.apenergy.2017.11.078>.
- [29] A. Ortizconde, F. Garciasanchez, J. Muci, New method to extract the model parameters of solar cells from the explicit analytic solutions of their illuminated characteristics, *Sol. Energy Mater. Sol. Cells* 90 (2006) 352–361, <https://doi.org/10.1016/j.solmat.2005.04.023>.
- [30] J. Pascual, F. Martínez-Moreno, M. García, J. Marcos, L. Marroyo, E. Lorenzo, Long-term degradation rate of crystalline silicon PV modules at commercial PV plants: An 82-MWp assessment over 10 years, *Prog. Photovolt. Res. Appl.* 29 (2021) 1294–1302, <https://doi.org/10.1002/pip.3456>.
- [31] S.J. Patel, A.K. Panchal, V. Kheraj, Extraction of solar cell parameters from a single current-voltage characteristic using teaching learning based optimization algorithm, *Appl. Energy* 119 (2014) 384–393, <https://doi.org/10.1016/j.apenergy.2014.01.027>.
- [32] J.C.H. Phang, D.S.H. Chan, J.R. Phillips, Accurate analytical method for the extraction of solar cell model parameters, *Electron. Lett.* 20 (1984) 406, <https://doi.org/10.1049/el:19840281>.
- [33] M.C.D. Piazza, M. Luna, G. Petrone, G. Spagnuolo, Translation of the single-diode PV model parameters identified by using explicit formulas, *IEEE J. Photovolt.* 7 (2017) 1009–1016, <https://doi.org/10.1109/JPHOTOV.2017.2699321>.
- [34] Z. Qu, A. Oumbe, P. Blanc, B. Espinar, G. Gesell, B. Gschwind, L. Klüser, M. Lefèvre, L. Saboret, M. Schroedter-Homscheidt, L. Wald, Fast radiative transfer parameterisation for assessing the surface solar irradiance: The Heliosat-4 method, *Meteorol. z.* 26 (2017) 33–57, <https://doi.org/10.1127/metz/2016/0781>.
- [35] N. Rajasekar, N. Krishna Kumar, R. Venugopalan, Bacterial Foraging Algorithm based solar PV parameter estimation, *Sol. Energy* 97 (2013) 255–265, <https://doi.org/10.1016/j.solener.2013.08.019>.

- [36] J.P. Ram, T.S. Babu, T. Dragicevic, N. Rajasekar, A new hybrid bee pollinator flower pollination algorithm for solar PV parameter estimation, *Energy Convers. Manag.* 135 (2017) 463–476, <https://doi.org/10.1016/j.enconman.2016.12.082>.
- [37] H. Sheng, C. Li, H. Wang, Z. Yan, Y. Xiong, Z. Cao, Q. Kuang, Parameters extraction of photovoltaic models using an improved moth-flame optimization, *Energies* 12 (2019) 3527, <https://doi.org/10.3390/en12183527>.
- [38] X. Sun, R.V.K. Chavali, M.A. Alam, Real-time monitoring and diagnosis of photovoltaic system degradation only using maximum power point—the Suns-Vmp method, *Prog. Photovolt. Res. Appl.* 27 (2019) 55–66, <https://doi.org/10.1002/ppp.3043>.
- [39] H. Tian, F. Mancilla-David, K. Ellis, E. Muljadi, P. Jenkins, A cell-to-module-to-array detailed model for photovoltaic panels, *Sol. Energy* 86 (2012) 2695–2706, <https://doi.org/10.1016/j.solener.2012.06.004>.
- [40] F.J. Toledo, J.M. Blanes, Analytical and quasi-explicit four arbitrary point method for extraction of solar cell single-diode model parameters, *Renew. Energy* 92 (2016) 346–356, <https://doi.org/10.1016/j.renene.2016.02.012>.
- [41] G. Xiong, L. Li, A.W. Mohamed, X. Yuan, J. Zhang, A new method for parameter extraction of solar photovoltaic models using gaining-sharing knowledge based algorithm, *Energy Rep.* 7 (2021) 3286–3301, <https://doi.org/10.1016/j.egy.2021.05.030>.
- [42] G. Xiong, J. Zhang, D. Shi, X. Yuan, Application of supply-demand-based optimization for parameter extraction of solar photovoltaic models, *Complexity* 2019 (2019) 1–22, <https://doi.org/10.1155/2019/3923691>.
- [43] S. Xu, Y. Wang, Parameter estimation of photovoltaic modules using a hybrid flower pollination algorithm, *Energy Convers. Manag.* 144 (2017) 53–68, <https://doi.org/10.1016/j.enconman.2017.04.042>.
- [44] M. Ye, X. Wang, Y. Xu, Parameter extraction of solar cells using particle swarm optimization, *J. Appl. Phys.* 105 (2009) 094502, <https://doi.org/10.1063/1.3122082>.
- [45] M. Zaimi, H. El Achouby, O. Zegoudi, A. Ibral, E.M. Assaid, Numerical method and new analytical models for determining temporal changes of model-parameters to predict maximum power and efficiency of PV module operating outdoor under arbitrary conditions, *Energy Convers. Manag.* 220 (2020) 113071, <https://doi.org/10.1016/j.enconman.2020.113071>.