HAL
open science

Estimation of the Uncertainty due to Each Step of Simulating the Photovoltaic Conversion under Real Operating Conditions

Anne Migan-Dubois, Jordi Badosa, Vincent Bourdin, Moira Torres Aguilar, Yvan Bonnassieux

To cite this version:

Anne Migan-Dubois, Jordi Badosa, Vincent Bourdin, Moira Torres Aguilar, Yvan Bonnassieux. Estimation of the Uncertainty due to Each Step of Simulating the Photovoltaic Conversion under Real Operating Conditions. International Journal of Photoenergy, 2021, 2021, pp.4228658. 10.1155/2021/4228658 . hal-03381247

HAL Id: hal-03381247

https://hal.science/hal-03381247

Submitted on 21 Feb 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.

Hindawi
International Journal of Photoenergy
Volume 2021, Article ID 4228658, 14 pages
https://doi.org/10.1155/2021/4228658

Research Article

Estimation of the Uncertainty due to Each Step of Simulating the Photovoltaic Conversion under Real Operating Conditions

Anne Migan Dubois,[1] **Jordi Badosa**,[2] **Vincent Bourdin**,[3] **Moira I. Torres Aguilar**,[2] **and Yvan Bonnassieux**[4]

[1]*Université Paris-Saclay, CentraleSupélec, CNRS, Laboratoire Génie Électrique et Électronique de Paris (GeePs), 91192 Gif-Sur-Yvette, France*
[2]*École Polytechnique, IPSL, Université Paris-Saclay, ENS, Sorbonne Université, CNRS, Laboratoire de Météorologie Dynamique (LMD), 91128 Palaiseau, France*
[3]*CNRS, Laboratoire Interdisciplinaire Des Sciences du Numérique (LISN), 91405 Orsay Cedex, France*
[4]*CNRS, École Polytechnique, Laboratoire de Physique et Des Couches Minces (LPICM), Route de Saclay, 91128 Palaiseau, France*

Correspondence should be addressed to Anne Migan Dubois; anne.migan-dubois@geeps.centralesupelec.fr

Received 6 July 2021; Accepted 23 August 2021; Published 15 October 2021

Academic Editor: Alberto Álvarez-Gallegos

Copyright © 2021 Anne Migan Dubois et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The simulation of photovoltaic installations is a major issue for their sizing, their smart grid operation, and their fault detection and diagnosis. In this article, we study in detail every step of the simulation chain, either from the global horizontal irradiance and the ambient temperature (i.e., 4 steps of simulation) or considering the global in-plane irradiance and the module operating temperature (i.e., 1 step of simulation). The average quality estimation of the models is made through the calculations of average annual error between estimations and measurements, from 2016 to 2020. We have shown that the most uncertain step is the conversion of the global irradiance in its diffuse and direct components (17.2%, 2 models tested). If the model goes up to the in-plane irradiance, the average annual error decreases to 5.3% (6 models tested). The photovoltaic module temperature calculation induces an error of less than 2Â°C (4 models tested with 2 configurations). Meanwhile, the photoelectrical conversion shows a 3.5% error, similar to the measurement uncertainties, considering as input, the modules temperature, and the in-plane irradiance. If the simulation goes from the global irradiance and the ambient temperature measured locally, the estimation leads to a 6.7% average annual error. If the local measurements are not available, we can use the closest meteorological station's records (13 for our study), and the error becomes 12.1%. Finally, we can also use the satellite images that lead to a 15.2% error, for average per year. The impact of available input shows that modeling the DC photovoltaic production, using global horizontal irradiance and ambient temperature, gives rise to an error of 6.6% for local measurements, 12.1% for weather station measurements, and 15.2% for satellite images estimations. This article thus draws up a review of the existing models, allowing to calculate the DC production of a photovoltaic module, depending on the atmospheric conditions, and highlights the most precise or most critical steps, considering in situ and weather station ground-based measurements, and also estimation from satellite images.

1. Introduction

Photovoltaic (PV) production mainly depends on the solar radiation incident on PV modules. Solar resource variability and uncertainty associated with the modeling of PV energy production are one of the most important factors that influence the grid stability (with wind turbines), regardless of the size of the power grid [1]. The ability to precisely predict the energy produced by PV systems is of great importance and has been identified as one of the key challenges for massive

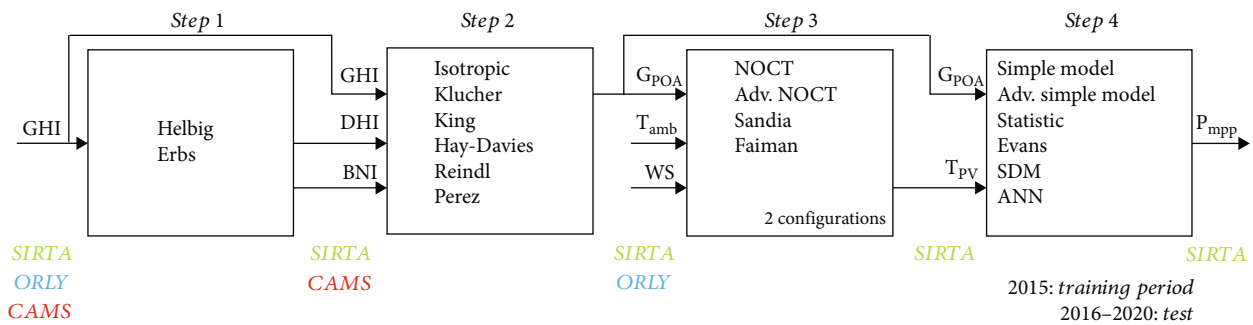

Figure 1: Schematic of the process for the simulation of PV production, studied models, and data sources available for each step.

PV integration [2, 3]. It is also a milestone in the sizing step of PV installations. This can also be used for PV fault detection and diagnosis.

This study focuses on evaluating the uncertainty on PV production estimation at each step of the modeling process. Each step is studied independently, and also its impact on the whole simulation chain is evaluated.

Few articles study the impact of the uncertainty of the modeling process. We can cite [4] which examines the uncertainty in long-term PV system yield predictions by statistical modeling (using Solar Advisor Model software) of a hypothetical 10 MW crystalline silicon PV system in Toronto (Canada). In this case study, uncertainties were estimated to be about 3.9% for year-to-year climate variability, 5% for long-term average horizontal irradiation, 3% for estimation of the in-plane radiation, 3% for power rating of modules, 2% for losses due to dirt and soiling, 1.5% for losses due to snow, and 5% for other sources of error. By performing statistical simulations, it was found that the combined uncertainty is approximately 8.7% for the first year of operation and 7.9% for the average yield over the PV system lifetime.

The study led by Sandia Laboratory [5] explains that the solar resource uncertainty (due to measurements, variability, spectrum...) is between 5 to 17%, the transposition of the horizontal irradiances to the in-plane irradiance is between 0.5 to 2%, the energy simulation and power plant losses induce uncertainty of 3 to 5%, and annual degradation uncertainty is about 0.5 to 1%. Using in situ ground-based measurements can reduce the uncertainty by up to 3.5%.

In reference [6], the authors compare yield predictions and monitoring data for 26 PV power plants located in southern Germany and Spain. If rather old radiation databases are used, the model systematically underestimates PV production by about 5% due to increased irradiance in recent years. Using recent satellite-derived irradiance avoids this underestimation. According to them, the main factor for the uncertainty of yield predictions is the aging, that is to say, the observed decrease of performance ratio. In this study, it decreases by 0.5% per year on average with a relatively high spread between systems. This decrease is attributed to nonreversible degradations and reversible effects, like soiling. The conclusion is that the uncertainty of the state-of-the-art yearly yield predictions using recent solar irradiance data is estimated to about 8%.

Only two publications compare the uncertainty of different photoelectric conversion models [7, 8]. Those two papers, written by the same authors, compare the annual PV yield prediction errors of four models: single-point efficiency, single-point efficiency with temperature correction, PVUSA (Photovoltaics for Utility Scale Applications), and single-diode model (SDM), against outdoor measurements for different grid-connected PV systems in Cyprus over a 4-year evaluation period. The best agreement between the modeled results and outdoor measurements for crystalline silicon PV technologies was obtained using SDM. The energy yield for thin-film technologies was more accurately predicted using the PVUSA model.

Our approach in this paper is similar to indirect forecasts: first, we calculate (or measure) the solar global in-plane irradiance (G_{POA}) and PV module operating temperature (T_{PV}), and then, using a PV performance model, we calculate the power produced at the maximum power point (P_{MPP}) [9]. Only the DC side of the photoelectric conversion is considered (no inverter). The different steps of the PV simulation are summarized in Figure 1, together with the considered models and the data sources.

This study focuses on the evaluation of the uncertainty on PV production estimation, step by step using different models, based on various data sources: in situ (SIRTA) and (ORLY) weather station ground-based measurements and geostationary satellite-based estimations (CAMS) during five years of operation.

For this purpose, we firstly present the data sources in Section 1. Section 2 describes the models that are used for each step of the modeling process with their accuracy. The results are presented in Section 3: it deals with modeling error estimation for the whole PV system simulation focusing on the uncertainty weight of each step and considering different data sources. These results are summarized in a conclusion and some perspectives are drawn.

2. Experimental Data Sources

In this study, three sources of data are explored. The highest level of precision corresponds to in situ ground-based measurements realized in our experimental research platform, where we do PV and atmospheric measurements, namely, SIRTA. Then, we consider the nearest weather station at ORLY airport where we use irradiances and ambient temperature measurements. Lastly, we use satellite-based estimations to proceed estimation of irradiance components

Figure 2: Outdoor characterization PV platform located at SIRTA.

using CAMS. All of these data sources are described in the following subsections.

The considered data are hourly averages and only with solar elevation angles higher than 10°.

2.1. Outdoor Photovoltaic Characterization Platform. Our experimental research platform is installed at the Atmospheric Research Observatory SIRTA [10] located in Palaiseau (France, 48.7 N, 2.2 E), 18.3 South-West of Notre-Dame de Paris.

A PV test bench was installed there in 2014 and comprises six commercial PV modules issued from different technologies (Figure 2). In this paper, we only consider the crystalline-silicon PV module, the second one from the left in Figure 2. All the equipment and sensor are listed in Table 4 of Appendix A, with their accuracy.

The current-voltage characteristics are measured with Agilent DC electronic loads (6060B), each minute from sunrise to sunset. The maximum power point (P_{MPP}) is derived from this characteristic. Under normal operating conditions, the voltage drop along the junction line is typically 0.5 at P_{MPP}, which results in a systematic error of less than 1.6% on the measured power. In addition, due to the transient mode of measurement, an uncertainty of ±5.6% of the measured power should be considered.

The PV module operating temperature (T_{PV}) is measured with 4-wired class A platinum sensors (Pt100) glued on the back-sheet. Their resolution is 0.01°C, and the accuracy is between ±0.15°C and ±0.30°C when the temperature is between -20°C and +100°C. In the transient mode of measurement with short pulses of 1 mA, there is no problem of self-heating of the probes. The overall uncertainty on the measured temperature is about ±0.40°C. The difference between the cell temperature and the probe glued on its back even with the highest irradiance is always less than 2°C, as we simulated it by common thermal laws [11]. These sensors are read by a digital multimeter Tektronix (DMM 3700A).

The global in-plane irradiance (G_{POA}) is measured with a second class solar radiometer (Hukseflux SR01) installed in the same plane as the PV modules. The SR01 is compliant with the ISO 9060 second class specifications. It was compared to a secondary standard for 2 weeks, and an overestimation of less than 0.3% on instantaneous measurements was observed. Thus, we can estimate that the SR01 is accurately calibrated.

2.2. In Situ Atmospheric Ground-Based Measurements SIRTA. Ground-based measurements are realized at SIRTA. It is a reference meteorological and climate observatory with more than 150 remote sensing and in situ instruments. In terms of radiometric measurements, the site is part of the Baseline Surface Radiation Network (BSRN) since 2003 [12]. Global horizontal irradiance (GHI), diffuse horizontal irradiance (DHI), and ground albedo (Albedo) measurements are realized following BSRN standards with Kipp & Zonen CMP22. The direct normal irradiance (BNI) is measured by a Kipp & Zonen CHP1 pyrheliometer. Both CMP22 and CHP1 are compliant with ISO 9060 secondary standard specifications, and they are expected to have a daily uncertainty lower than ±1%. The ambient temperature (T_{amb}) is measured at 1.5 m above the ground by a class A Guilcor platinum sensor (Pt100) with a precision of ± 0.15°C. A Vector A100R anemometer attached to the PV platform measures local wind speed (WS) with an accuracy of 0.1 ms^{-1} and a threshold of 0.3 ms^{-1}.

2.3. Weather Station Atmospheric Ground-Based Measurement ORLY. The weather station which is the nearest to our PV test bench is ORLY airport, located at Athis-Mons (France, 48.4 N, 2.2 E), about 13 km East from SIRTA, opened in 1921. T_{amb} is measured with class B sensors with a precision of ±0.5°C. WS measurements are made by class B anemometers whose accuracy is 10% or 0.5 ms^{-1}. GHI is measured with a pyranometer Kipp & Zonen CM6B compliant with ISO 9060 first-class specifications.

2.4. Satellite Irradiance Estimation CAMS. This study also considers estimations of irradiances from Meteosat geostationary satellite observations computed by Copernicus Atmosphere Monitoring Service (CAMS) [13, 14].

The temporal resolution is one minute, hourly averaged, and the spatial resolution results from interpolation to the point of interest. Meteosat pixels have an elliptic shape, and their average diameter ranges from 3 km to 7 km depending on the viewing geometry of the satellite. The estimations are GHI, DHI, and BNI in clear sky conditions; verbose mode with all atmospheric input parameters used for clouds, aerosols, ozone, water vapor, and the surface reflective properties. T_{amb} and WS are not estimated with satellite images. We use those that are measured in ORLY weather station, if needed.

Table 1: Available data versus data sources.

Data source	GHI	DHI	BNI	G_{POA}	WS and T_{amb}	T_{PV}
SIRTA	✓	✓	✓	✓	✓	✓
ORLY	✓				✓	
CAMS	✓	✓	✓			

Table 2: SDM validation in STC.

Parameter	STC	Model	Relative difference
P_{MPP}	245.7 W	246.5 W	0.32%
V_{MPP}	30.7 V	30.8 V	0.26%
I_{MPP}	8.0 A	8.0 A	0.58%

All the installed sensors and data-logging devices for the three data sources, with their accuracy, are listed in Table 4 of Appendix A.

The available data used for this study is detailed in Table 1:

3. Theoretical Modeling and Model Accuracy

G_{POA} and T_{PV} are needed to estimate P_{MPP}. Either they are available through in situ measurement or they should be estimated, adding up to three modeling steps (as presented in Figure 1).

All presented models are evaluated by yearly averages of data (2016, 2017, 2018, 2019, and 2020). Complete models can be found in the literature listed. All those theoretical models have been programmed using Python, an object-oriented, multiparadigm, and multiplatform programming language. All computed data are summarized in Figure 9 of Appendix B.

3.1. Modeling Performance Indicators. In order to compare estimations to measurements, we compute relative mean bias error (rMBE) and relative mean absolute error (rMAE), as defined in the following equations.

$$\mathrm{rMBE} = \frac{\sum_{i=1}^{N}\left[X_{\mathrm{calc}}(i) - X_{\mathrm{meas}}(i)\right]}{\sum_{i=1}^{N} X_{\mathrm{meas}}(i)} \times 100, \quad (1)$$

$$\mathrm{rMAE} = \frac{\sum_{i=1}^{N}\left|X_{\mathrm{calc}}(i) - X_{\mathrm{meas}}(i)\right|}{\sum_{i=1}^{N} X_{\mathrm{meas}}(i)} \times 100, \quad (2)$$

where X can be DHI, BNI, G_{POA}, or P_{MPP}, X_{calc} correspond to the result of X modeling, X_{meas} is the measured value of X, and N is the number of measured (or calculated) values of X, during one year. The indicators for T_{PV} are expressed in and not normalized which would have been nonsense.

Boxplot is drawn for each rMBE and rMAE calculated. The central red mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the "+" symbol. It should be noted that the use of a boxplot to represent 5 values (5 years from 2016 to 2020) is not necessarily very relevant, but it is visually very meaningful.

In this part, the performance of each step of Figure 1 is estimated individually. Input data are in situ ground measurements from SIRTA, for all presented results in this part, as the objective is to evaluate the uncertainty of each modeling step.

3.2. Step 1: Estimation of the Diffuse Fraction. In this study, we consider two models to compute DHI and BNI components of the irradiance from measured GHI. This input is taken from SIRTA in situ ground measurements as the objective is to evaluate only step 1.

For Helbig model, the fraction of the diffuse irradiance is calculated from GHI using an empirical relationship, giving an estimated value of DHI and BNI [15]. The empirical coefficients taken for this study are those found in [15] because modeling performance is not enhanced by using coefficients that have been computed with real data.

Assuming that the diffuse fraction of the hourly total radiation (DHI/GHI) is strongly correlated with the hourly clearness index (k_T), Erbs proposes an empirical relation using the combined data of four American locations [16]. k_T is the ratio of the hourly global radiation to the hourly extraterrestrial radiation.

rMAE and rMBE are calculated using SIRTA in situ ground measurements of DHI and BNI compared to modeled ones using GHI also measured at SIRTA.

As can be seen in Figure 3, the dispersion is wider for the estimation of BNI. DHI is more often underestimated, whereas BNI is overestimated. Helbig models seem to perform better than Erbs, with rMBE = −3.1 ± 1.1% and rMAE = 17 ± 0.9% for the estimation of DHI and rMBE = 5.4 ± 1.6% and rMAE = 16 ± 1.9% for BNI.

3.3. Step 2: Transposition of the Horizontal Irradiances in the Plane of Array. To calculate G_{POA}, needed input data are GHI, DHI, and BNI solar irradiances as well as the tilt angle and the ground albedo. These three measured input are taken from SIRTA in situ ground measurements as the objective is to evaluate only step 2.

G_{POA} is calculated by the sum of the beam, diffuse, and albedo irradiances in the plane of array. The beam irradiance is calculated with geometric relation between the BNI and the angle of incidence AOI (angle between the sun's rays and the PV array). AOI can easily be computed by knowing the sun's position (solar azimuth and zenith angles) and the array's geometry (tilt and azimuth angles). The irradiance due to the ground albedo is calculated thanks to the view factor in front of the array, the celestial vault is assumed to be uniform, and the ground is supposed to have a

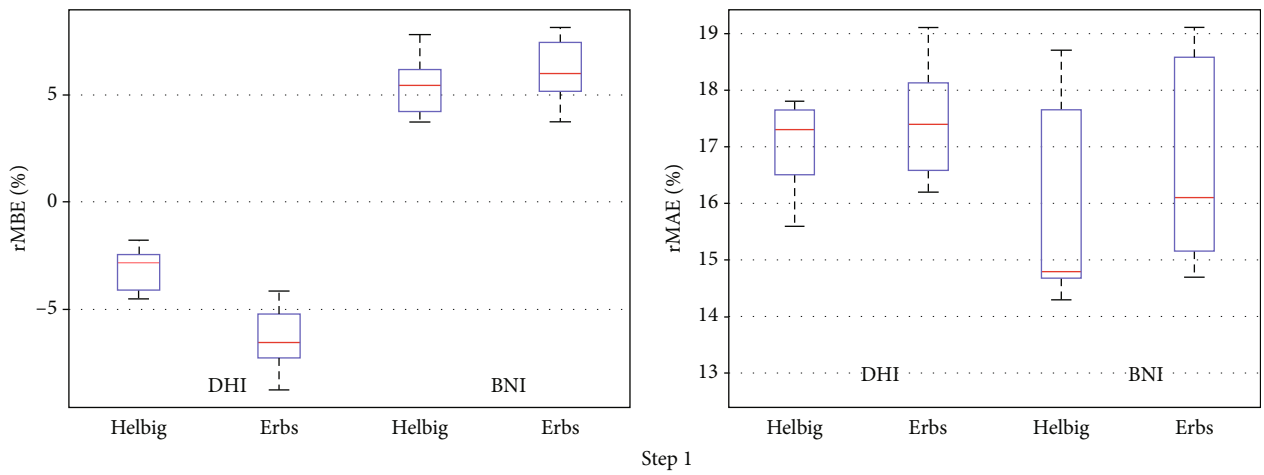

Figure 3: rMBE (left) and rMAE (right) of Helbig and Erbs models to calculate DHI and BNI from GHI measured at SIRTA. Computed data are summarized in Figure 9 of Appendix B.

Lambertian behavior. The coefficient of reflection of the ground is 0.2 in our case, as representative of grass.

Six models have been considered representing different ways to estimate the diffuse irradiance arriving on PV modules, as shown in Figure 1.

The isotropic sky diffuse model assumes that the solar radiosity of the celestial vault is uniformly distributed over the complete skydome [17]. It is the simplest of the tilt surface models. Under completely cloudy skies, this model is quite a good calculation. The five other models do not consider diffuse radiation as isotropic.

Klucher found that the isotropic model gave good results for overcast skies but underestimates irradiance under clear and partly overcast conditions, when there is increased intensity near the horizon and in the circumsolar region of the sky [18].

King model provides the portion of the total radiation incident on the surface that will be used by the module for energy production. It accounts for variation in the solar spectrum as a function of absolute air mass through an empirical function as well as for optical losses incurred by the angle of incidence. These two empirical functions were proposed by the Sandia Laboratory [19] and developed through outdoor testing of relevant modules [20]. Klucher modified this clear sky model by imposing a modulating factor which forces the anisotropic correction factor to approach unity under cloudy sky conditions so that the model reduces to the isotropic sky model.

Hay and Davies developed a model to predict the tilted surface diffuse radiation accounting for both circumsolar and isotropic diffuse radiation [21]. Realizing that the isotropic hypothesis becomes less accurate under clear sky conditions, they defined an anisotropy index which corresponds to a portion of the diffuse radiation to be treated as circumsolar with the remaining portion considered isotropic. The circumsolar diffuse is projected onto the tilted surface, and the remaining diffuse radiation is treated as isotropic diffuse. Under clear skies, the anisotropy index is high, and the circumsolar diffuse is more critical than the isotropic diffuse. Under cloudy skies, the anisotropy index goes to zero, and all diffuse is treated as isotropic. This model does not account for horizon brightening diffuse radiation.

In the model of Perez et al. [22], the contribution of diffuse radiation from circumsolar, isotropic, and horizontal regions is determined by two empirically derived coefficients [23]. The empirical coefficients are based on two years of data from Carpentras and Trappes, both in France. The empirical coefficient for the diffuse radiation from circumsolar is completed by measurements realized in five U.S. locations [24]. Perez et al. approximated the horizon brightening effects by applying a correction factor to the isotropic diffuse radiation.

In the reference [25], Reindl et al. investigate the performance of five models which estimate the hourly tilted surface radiation by comparing with measured energy on six experimental sites. He proposes to apply the horizon brightening correction factor used by Perez to the isotropic term in the Hay-Davies model.

rMAE and rMBE are calculated using SIRTA in situ ground measurements of G_{POA} compared to modeled one, using GHI, DHI, and BNI also measured at SIRTA.

As can be seen in Figure 4, Klucher model best performs regarding error (rMSE = 4, 7 ± 0, 6%) and also bias (rMSE = − 0, 1 ± 0, 4%).

3.4. Step 3: Estimation of PV Module Temperature. In this part, we evaluate four models to calculate T_{PV} from T_{amb}, G_{POA}, and WS. These three inputs are taken from SIRTA in situ ground measurements as the objective is to evaluate only step 3.

The nominal operating cell temperature (NOCT) is defined as the temperature reached by open-circuited cells in a module under a G_{POA} of 800, a T_{amb} of 20 ($T_{amb,NOCT}$), and a WS of 1. An approximate expression for calculating the module operating temperature is given in [26]. In our first calculations, we have considered a typical value of NOCT (48). In a second step, we fit the value of NOCT to one year of measurements (2015) using the Levenberg-Marquardt method. The obtained NOCT is equal to 45.

This model does not consider the impact of wind speed on module temperature, and we can imagine a nonproportional relation between them. This effect is included in the following

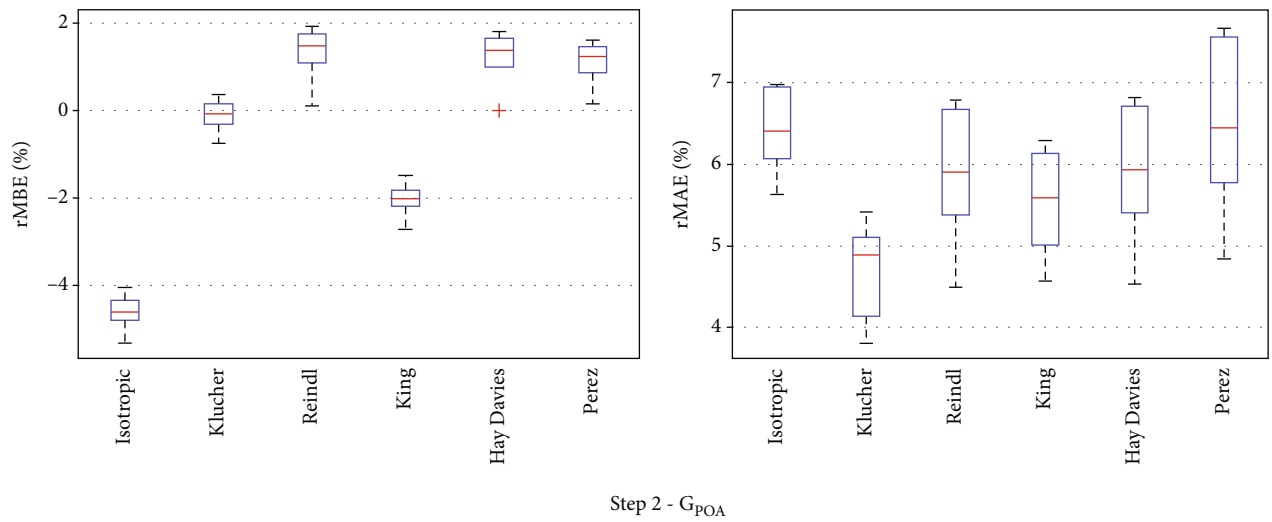

Figure 4: rMBE (left) and rMAE (right) of the presented models to calculate G_{POA} from GHI, DHI, and BNI measured at SIRTA. Computed data are summarized in Figure 9 of Appendix B.

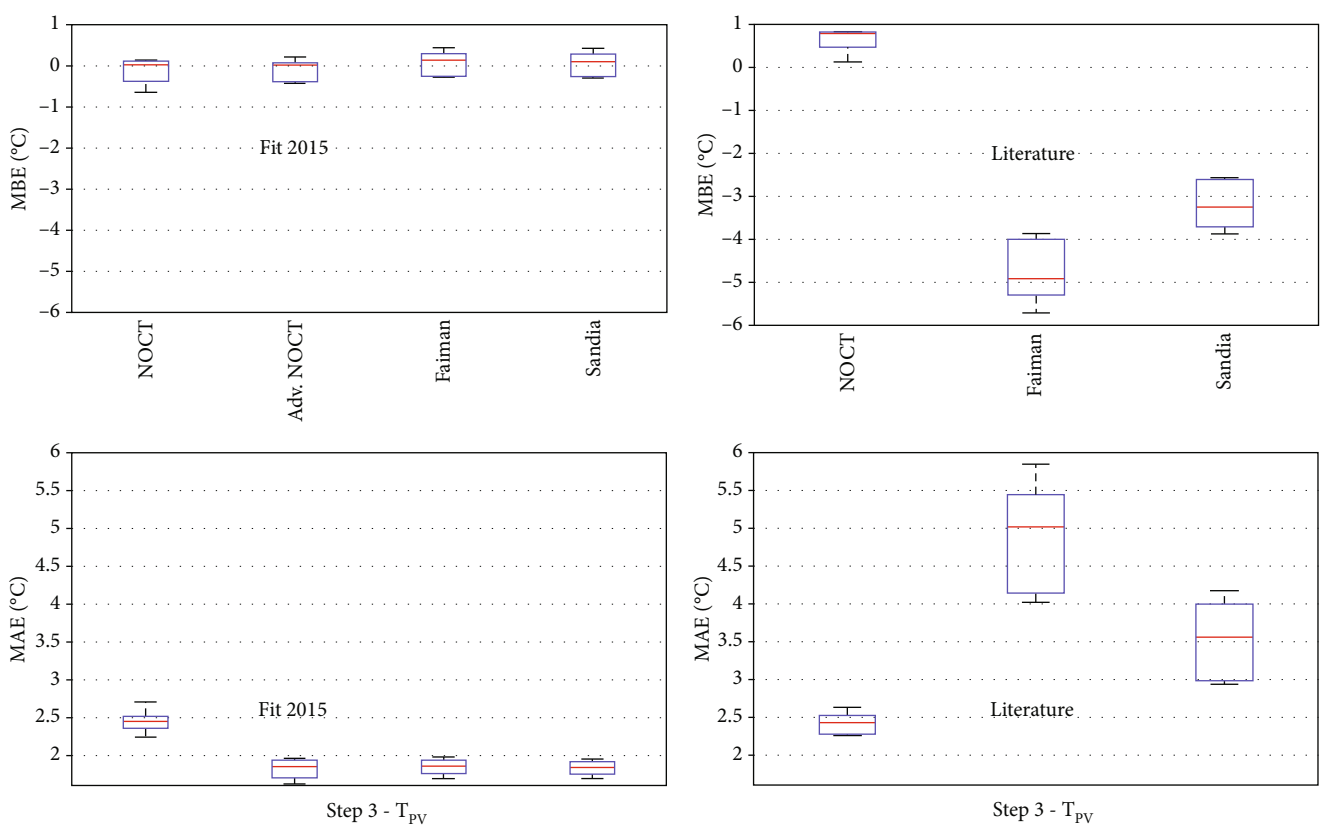

Figure 5: MBE (upper) and MAE (lower) of the presented models to calculate T_{PV} from GHI, T_{amb}, and WS, fitted coefficients (left) and literature coefficients (right). Computed data are summarized in Figure 9 of Appendix B.

improved NOCT formula that has been developed and named Advanced NOCT and is presented here [27]:

$$T_{PV} = T_{amb} + \frac{G_{POA,th}}{G_{POA,NOCT}} \times (a \cdot G_{POA,th} + b) \times \frac{\text{NOCT} - T_{amb,NOCT}}{1 + (\text{WS}/\text{WS}_{NOCT})^{c}}. \quad (3)$$

For module producing electricity, $G_{POA,th}$ is the fraction of energy converted into heat and is defined by the following equation:

$$G_{POA,th} = G_{POA} \times (1 - \text{CE}). \quad (4)$$

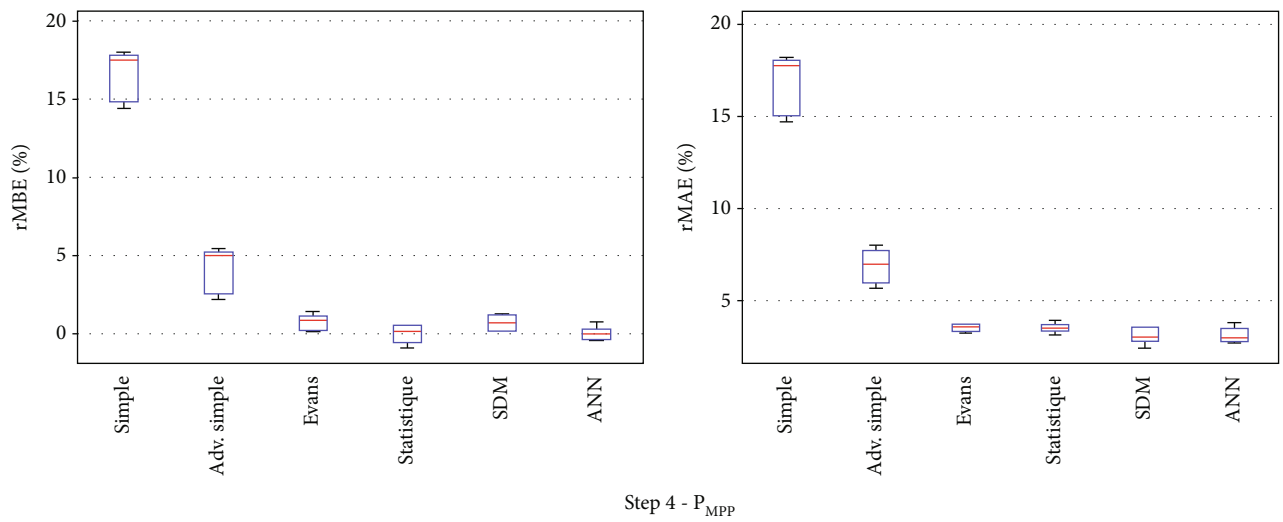

Figure 6: rMAE and rMBE of the presented models to calculate P_{MPP} from GHI and T_{PV}. Computed data are summarized in Figure 9 of Appendix B.

Table 3: Best results for each modeling step considering SIRTA in situ ground measurements. All computed data can be found in Figure 9 of Appendix B.

Step	Input	Model	Output	rMBE/MBE mean ± std	rMBE/MAE mean ± std
1	GHI	Helbig	DHI	−3.1 ± 1.1%	17.0 ± 0.9%
			BNI	5.4 ± 1.6%	16.0 ± 1.9%
2	GHI DHI BNI	Klucher	G_{POA}	−0.1 ± 0.4%	4.7 ± 0.6%
3	GHI T_{amb} WS	Adv. NOCT	T_{PV}	−0.1 ± 0.3	1.8 ± 0.1
4	G_{POA} T_{PV}	Evans	P_{MPP}	0.7 ± 0.5%	3.5 ± 0.2%

The coefficients a, b, and c are obtained by fitting with the Levenberg-Marquardt method, the value of T_{PV} measured during one year (2015), with the formula given in equation (3). We found that $a = 3.78e - 4$, $b = 2.29$, and $c = 0.474$. This study is in progress, in particular, to validate the location and technological universality of the proposed coefficients.

In the mid-1980s, a thermal model was developed at Sandia Laboratory for system engineering and performance modeling purposes [28]. Although rigorous, this early model has proven to be unnecessarily complex, not applicable to all PV technologies, and not easily adaptable to site-dependent influences. A simpler empirically-based thermal model was then developed at Sandia Laboratory [19]. This simple model has proven to be very adaptable and entirely adequate for system engineering and design purposes by providing the expected module operating temperature with an accuracy of about ±5. The empirical coefficients establish the upper limit for module temperature at low wind speeds and high solar irradiance, and the rate at which T_{PV} drops as WS increases, respectively. To evaluate this model, we firstly consider parameter values proposed by the authors ($a = -3.47$, $b = -0.0594$). In a second step, we fit these coefficients to one year of measurements (2015) using the Levenberg-Marquardt method. The obtained coefficients are $a = -3.1398$, $b = -0.305$.

Faiman presented a module temperature model [29] based on simple heat transfer concepts. To evaluate this model, we firstly consider empirical coefficient values proposed by the authors ($U_0 = 25.0$, $U_1 = 6.84$). In a second step, we fit these coefficients to one year of measurements (2015) using the Levenberg-Marquardt method. The obtained coefficients are $U_0 = 21.78$, $U_1 = 9.86$. Given the coefficients obtained for these last three models, Adv. NOCT is more sensitive to WS than the others.

MAE and MBE are calculated using SIRTA in situ ground measurements of T_{PV} compared to modeled one using G_{POA}, T_{amb}, and WS also measured at SIRTA.

As shown in Figure 5, the best performance is not obtained for coefficients taken directly from literature. Moreover, if you can do a data fitting during one year, all the four proposed methods have almost the same performance. For the rest of the study, we will choose the Adv

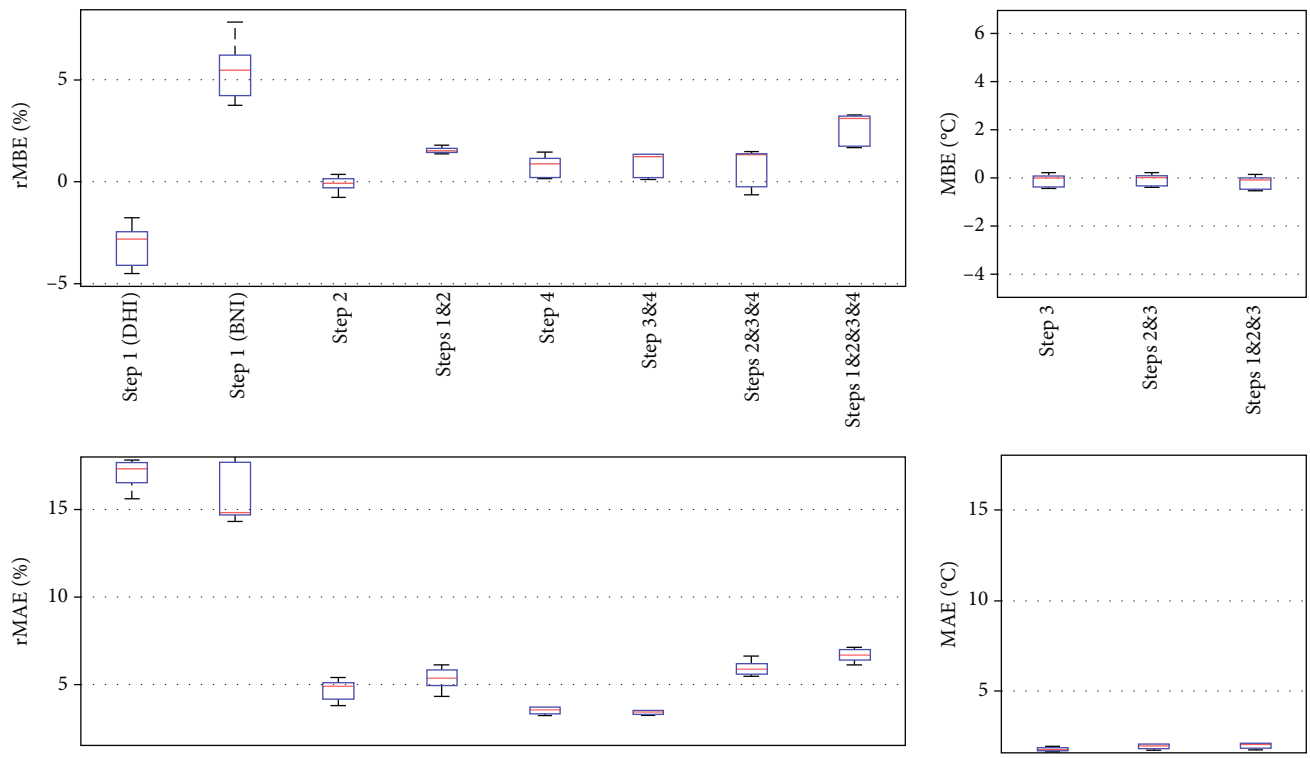

Figure 7: Modeling error estimation in each step of the simulation process for SIRTA measurements. The models are Helbig (step 1), Klucher (step 2), Adv. NOCT (step 3), and Evans (step 4). Computed data can be found in Figure 10 of Appendix C.

NOCT model because according to the coefficients found, it is a little more sensitive to WS. Its accuracy is MBE = −0.1 ± 0.3 and MAE = 1.8 ± 0.1.

3.5. Step 4: Photoelectric Conversion Models. In this study, we consider six different photoelectric conversion models [30] to calculate P_{MPP} from G_{POA} and T_{PV}. These two inputs are taken from SIRTA in situ measurements as the objective is to evaluate only step 4. Soiling effect and aging had not been taken into account.

The single point efficiency model, named here simple model, considers a constant conversion efficiency (CE) measured by the manufacturer during a flash test in standard test conditions (STC) [31]. For the studied PV module, $\mathrm{CE}_{\mathrm{STC}} = 15.28\%$.

We have improved this model by considering the conversion efficiency equal to the average measurements during 2015: $\mathrm{CE}_{2015} = 13.65\%$.

Evans model is a physics-based model that considers the linear variation of $\mathrm{CE}_{\mathrm{STC}}$ with module temperature and its exponential decreasing with the low light effect [32]. The temperature coefficient of the power is taken equal to -0.48% from the manufacturer datasheet and the low light effect coefficient is $\gamma = 0.12$, as advised by Evans in his article [32], for a crystalline silicon PV module.

The statistical model does not need internal information from the system to describe its performance. It is a data-driven approach that can extract relations on past historical data of couples (G_{POA}, T_{PV}) to predict the future behavior of the PV module (P_{MPP}) [33]. For a given hour h, in the year N, the statistical model looks for similar atmospheric conditions in 2015, that is to say, the ensemble of points $h_i \cap h_j$ respecting the following two conditions: $G_{\mathrm{POA}}(2015, h_i) = G_{\mathrm{POA}}(N, h) \pm I_G$ and $T_{\mathrm{PV}}(2015, h_j) = T_{\mathrm{PV}}(N, h) \pm I_T$. The algorithm starts with $I_G = 0$ and $I_T = 0$. If the ensemble of common points in illumination and temperature $(h_i \cap h_j)$ is empty, I_G and I_T intervals increase gradually by 5 compared to the target conditions, with a limit of 50. If several points in this ensemble are found, the given P_{MPP} at the output of the model is the average. Thus, the quality of the historical data is essential for an accurate simulation of this statistical model.

The single-diode electrical model (SDM) with 4 parameters is based on the Shockley diode equation [34], with a current source to model the photo-current, a single-diode junction, and a series resistance to model the contact losses. The photocurrent depends on T_{PV} and G_{POA}, and the diode saturation current and the series resistance are temperature dependent. The ideality factor of the diode is constant. All the parameters used in the above equation are determined by fitting the manufacturer flash test, and the STC modeling is compared in Table 2.

This 4-parameter SDM is more straightforward than the 5-parameter one and gives quite a good simulation accuracy.

The artificial neural network (ANN) was built using the feed-forward neural network structure with a weighted linear combination and sigmoid function [35, 36]. The architecture chosen is one output (P_{MPP}), two inputs (G_{POA} and T_{PV}), and one hidden layer. The training period was one year (2015), of which 2/3 of the data were used for the

Table 4: Equipment and sensors used for all the presented measurements.

Param.	Manufact.	Model	Std	Accuracy	Meas. uncert.
		PV module level measurements			
P_{MPP}	Agilent	6060B	IEC 60891	Electronic load: ±0.1% Voltmeter: 0.1% ± 300 mV Ammeter: 0.1% ± 350 mA	±7%
T_{PV}	TC	578-093 Class A	IEC 60751	[-20 °C; +100 C]: ±0.15°C 100.00 ± 0.06Ω at 0 C	±0.4 C $\lvert T_{cell} - T_{PV} \rvert \leq 1°C$
G_{POA}	Hukseflux Tektronix	SR01 2nd class 3706A	ISO 9060	±20% in hourly totals DC voltmeter: ±0.0025%	
		Atmospheric in situ measurements (SIRTA)			
GHI	Kipp & Zonen	CM22 2ry std	ISO 9060	Daily uncertainty ≤1%	
BNI	Kipp & Zonen	CHP1 2ry std	ISO 9060	Daily uncertainty ≤1%	
T_{amb}	Guilcor	Pt100 Class B	IEC 60751	[-100 °C; +100 C]: ±0.15 C 100.00 ± 0.06Ω at 0 C	±0.4 C
WS	Vector	A100R		[0.3 m/s; 10 m/s]: 0.1 m/s	
		Atmospheric measurements from nearby weather station (ORLY)			
GHI	Kipp & Zonen	CM6B 1 nd class	ISO 9060	Daily uncertainty ≤5%	
T_{amb}		Pt100 Class B	IEC 60751	[-100 °C; +100 C]: ±0.15°C	±0.4°C
WS		Class B		100.00 ± 0.06Ω at 0 °C 10%or 0.3 m/s	
		Geostationary satellite observations (CAMS)			
GHI, DHI, BNI	CAMS			Pixel diameter ≈4.5 km	

training and 1/3 for the validation. Dataset for the test is each year from 2016 to 2020.

rMBE and rMAE are calculated using SIRTA in situ measurements of P_{MPP} compared to modeled one using G_{POA} and T_{PV} also measured at SIRTA.

As can be seen in Figure 6, except for simple model, all of them have almost the same performance. SDM can be improved by either implement the 5-parameter SDM or better taking into account the dependence of the parameters with the temperature. The double-diode model could also be used to be more accurate at low irradiances. Moreover, SDM is tricky to tune because parameters are not given in the datasheet, and flash tests are not always available. Statistical and ANN needs a historical dataset to train the model. For these reasons, we decided to select the Evans model as the best for the following study, with rMBE = 0.8 ± 0.5% and rMAE = 3.5 ± 0.2%.

To summarize, this theoretical study shows that best performances are obtained, for each step, by using the Helbig model for step 1, the Klucher model for step 2, the Adv. NOCT model for step 3, and the Evans model for step 4 (see Figure 1). The modeling step with less accuracy is step 1. All those results are gathered in Table 3.

4. Results: Error in the Simulation of the P_{MPP} from Step to Step

In this section, we firstly study the uncertainty of each successive step, just for SIRTA in situ ground-based measurements, in order to evaluate the propagation of the error and the most uncertain simulation stem.

In the second part, we compare different ways to reach an estimation of P_{MPP}, with decreasing the difficulty of accessing data (that is to say, from SIRTA to ORLY and CAMS input data) and increasing the number of modeling steps (from only step 4 to all steps 1 to 4), as we can see in Figure 1.

4.1. Error Propagation in the Simulation of P_{MPP} from Step to Step. First, let us study the uncertainty of each step, just for SIRTA in situ ground-based measurements, in order to evaluate the propagation of the error and the most uncertain one. Figure 7 summarizes and completes what has been obtained in previous sections. All computed data can be found in Figure 10 of Appendix C.

First of all, this study confirms that the most uncertain step is the first one, to estimate the diffuse fraction (17.0%). Nevertheless, if step 2 (4.7%) is associated to step 1, having in mind

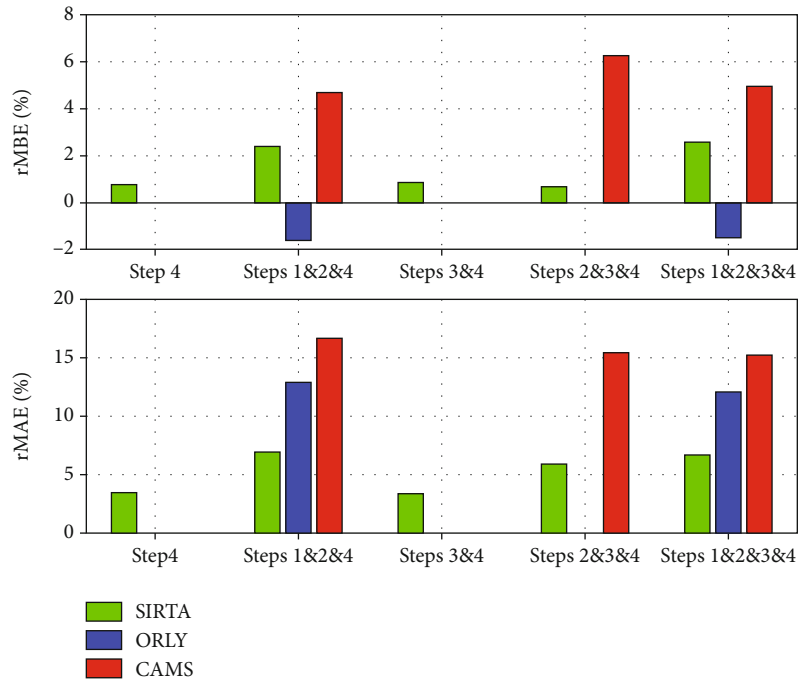

Figure 8: Modeling mean annual rMBE (top) and mean annual rMAE (bottom) of the whole simulation process as a function of the accessibility of the data sources (SIRTA, ORLY, and CAMS). The models are Helbig (step 1), Klucher (step 2), Adv. NOCT (step 3), and Evans (step 4).

that the critical variable is G_{POA} and not DHI, the error decreases down to 5.3%. This error is of the same order of magnitude as the precision of measuring instruments.

The estimation of T_{PV} is not a critical step. The associated uncertainty is lower than 2 (right part of Figure 7). This error must be separated from that of the other steps because they are not of the same type. Moreover, the impact of the T_{PV} modeling error does not linearly impact the modeling error of P_{MPP}.

In addition, the estimation of T_{PV} from environmental measurements is not critical in the simulation of P_{MPP} because the error whatever T_{PV} is measured (step 4 or steps 1 and 2 and 4) or estimated (steps 3 and 4 or steps 1 and 2 and 3 and 4) remains around 3.5% or 6.9%. On the other hand, the computation of G_{POA} has a very strong sensitivity on the simulation of P_{MPP}, confirming what has already been said before. The error goes from 3.5% with G_{POA} measured (step 4 or steps 3 and 4) to 6.9% when it is calculated (steps 1 and 2 and 4 or steps 1 and 2 and 3 and 4).

Moreover, we can see with this study that the errors almost accumulate, if steps 1 and 2 are considered together: the total error in P_{MPP} modeling (steps 1 to 4) is 6.7%, and the mean squared error of steps 1 and 2 and step 4 is 6.4%.

4.2. Error in the Simulation of P_{MPP} according to Data Source. In this part, different ways to reach an estimation of P_{MPP}, with decreasing the difficulty of accessing data (that is to say, from SIRTA to ORLY and CAMS input data) and increasing the number of modeling steps (from only step 4 to all steps 1 to 4) are studied, as can be seen in Figure 1.

Figure 8 presents the results of the global simulation, for all the data sources. The chosen models are the same as for the previous study. The differences between all the simulations are the data sources (in situ ground-based measurements SIRTA, ORLY weather station ground-based measurements, and CAMS satellite-based estimations).

This figure shows whatever the data sources, and the trends are the same: (i) modeling P_{MPP} with GHI is enough to have good performance compared to using GHI, DHI, and BNI; (ii) T_{PV} modeling hardly generates any error in the estimation of P_{MPP}.

Moreover, the less accessible the data are, the more accurate the simulation is. The best results are obtained for SIRTA in situ ground-based measurements with a rMAE of 6.7% for the whole simulation chain. Then, the nearest weather station ORLY (12.1%) and the worth are given by CAMS satellite images estimations (15.2%). This is mainly due to the fact that ground-based measurements are not biased but satellite-based estimations are. It can be explained by, on the one hand, the pixels size of the satellite images, and also, on the other hand, the uncertainties related to the estimation of GHI, DHI and BNI, mostly in overcast conditions.

5. Conclusion

In this paper, we have studied, step by step, the simulation process of the PV DC energy production, with a special focus on the calculation of the uncertainty.

Four steps compose the simulation chain, and basic and well-known models were considered.

(1) Step 1 ⟶ from GHI to GHI, DHI, and BNI: 2 models

Variable	Source	Step 1	Step 2	Year	rMBE	rMAE	Mean rMBE	Std rMBE	Mean rMAE	Std rMAE
DHI	SIRTA	helbig		2016	−2,82	16,80	−3,148	1,089	17,020	0,879
DHI	SIRTA	helbig		2017	−3,96	15,60				
DHI	SIRTA	helbig		2018	−2,68	17,30				
DHI	SIRTA	helbig		2019	−4,51	17,80				
DHI	SIRTA	helbig		2020	−1,77	17,60				
DHI	SIRTA	erbs		2016	−5,59	16,70	−6,368	1,696	17,440	1,115
DHI	SIRTA	erbs		2017	−6,79	16,20				
DHI	SIRTA	erbs		2018	−6,56	17,40				
DHI	SIRTA	erbs		2019	−8,76	19,10				
DHI	SIRTA	erbs		2020	−4,14	17,80				
BNI	SIRTA	helbig		2016	5,66	18,70	5,408	1,557	15,980	1,920
BNI	SIRTA	helbig		2017	7,81	17,30				
BNI	SIRTA	helbig		2018	4,37	14,80				
BNI	SIRTA	helbig		2019	5,46	14,80				
BNI	SIRTA	helbig		2020	3,74	14,30				
BNI	SIRTA	erbs		2016	6,00	19,10	6,148	1,678	16,720	1,934
BNI	SIRTA	erbs		2017	8,14	18,40				
BNI	SIRTA	erbs		2018	5,64	15,30				
BNI	SIRTA	erbs		2019	7,23	16,10				
BNI	SIRTA	erbs		2020	3,73	14,70				
GPOA	SIRTA		Isotrophic	2016	−4,62	6,93	−4,606	0,460	6,428	0,555
GPOA	SIRTA		Isotrophic	2017	−5,32	6,97				
GPOA	SIRTA		Isotrophic	2018	−4,65	6,4				
GPOA	SIRTA		Isotrophic	2019	−4,44	5,63				
GPOA	SIRTA		Isotrophic	2020	−4,6	6,21				
GPOA	SIRTA		Klucher	2016	−0,164	5,41	−0,108	0,413	4,672	0,637
GPOA	SIRTA		Klucher	2017	−0,754	5				
GPOA	SIRTA		Klucher	2018	0,363	4,89				
GPOA	SIRTA		Klucher	2019	0,0933	3,81				
GPOA	SIRTA		Klucher	2020	−0,0774	4,25				
GPOA	SIRTA		Reindl	2016	1,7	6,78	1,325	0,709	5,894	0,915
GPOA	SIRTA		Reindl	2017	1,92	6,63				
GPOA	SIRTA		Reindl	2018	1,48	5,9				
GPOA	SIRTA		Reindl	2019	0,107	4,49				
GPOA	SIRTA		Reindl	2020	1,42	5,67				
GPOA	SIRTA		King	2016	−2,01	6,29	−2,034	0,441	5,538	0,697
GPOA	SIRTA		King	2017	−2,72	6,08				
GPOA	SIRTA		King	2018	−1,49	5,59				
GPOA	SIRTA		King	2019	−1,93	4,57				
GPOA	SIRTA		King	2020	−2,02	5,16				
GPOA	SIRTA		Haydavies	2016	1,59	6,81	1,221	0,711	5,926	0,914
GPOA	SIRTA		Haydavies	2017	1,81	6,67				
GPOA	SIRTA		Haydavies	2018	1,38	5,93				
GPOA	SIRTA		Haydavies	2019	−0,00338	4,53				
GPOA	SIRTA		Haydavies	2020	1,33	5,69				
GPOA	SIRTA		Perez	2016	1,6	7,52	1,102	0,564	6,508	1,153
GPOA	SIRTA		Perez	2017	1,42	7,66				
GPOA	SIRTA		Perez	2018	1,24	6,44				
GPOA	SIRTA		Perez	2019	0,15	4,84				
GPOA	SIRTA		Perez	2020	1,1	6,08				
GPOA	SIRTA	helbig	Isotropic	2016	−3,26	6,6	−3,132	0,189	5,924	0,590
GPOA	SIRTA	helbig	Isotropic	2017	−3,24	6,05				
GPOA	SIRTA	helbig	Isotropic	2018	−2,83	5,9				
GPOA	SIRTA	helbig	Isotropic	2019	−3,06	4,98				
GPOA	SIRTA	helbig	Isotropic	2020	−3,27	6,09				
GPOA	SIRTA	helbig	Klucher	2016	1,44	6,12	1,534	0,155	5,326	0,677
GPOA	SIRTA	helbig	Klucher	2017	1,39	5,71				
GPOA	SIRTA	helbig	Klucher	2018	1,79	5,36				
GPOA	SIRTA	helbig	Klucher	2019	1,55	4,32				
GPOA	SIRTA	helbig	Klucher	2020	1,5	5,12				
GPOA	SIRTA	helbig	Reindl	2016	3,57	7,9	3,222	0,925	6,756	1,242
GPOA	SIRTA	helbig	Reindl	2017	4,35	7,83				
GPOA	SIRTA	helbig	Reindl	2018	3,15	6,48				
GPOA	SIRTA	helbig	Reindl	2019	1,8	4,85				
GPOA	SIRTA	helbig	Reindl	2020	3,24	6,72				
GPOA	SIRTA	helbig	King	2016	−0,648	6,35	−0,558	0,168	5,490	0,732
GPOA	SIRTA	helbig	King	2017	−0,632	5,81				
GPOA	SIRTA	helbig	King	2018	−0,271	5,47				
GPOA	SIRTA	helbig	King	2019	−0,55	4,35				
GPOA	SIRTA	helbig	King	2020	−0,687	5,47				
GPOA	SIRTA	helbig	Haydavies	2016	3,47	7,9	3,128	0,928	6,756	1,242
GPOA	SIRTA	helbig	Haydavies	2017	4,26	7,83				
GPOA	SIRTA	helbig	Haydavies	2018	3,06	6,48				
GPOA	SIRTA	helbig	Haydavies	2019	1,7	4,85				
GPOA	SIRTA	helbig	Haydavies	2020	3,15	6,72				
GPOA	SIRTA	helbig	Perez	2016	2,19	7,8	1,996	0,573	6,858	1,110
GPOA	SIRTA	helbig	Perez	2017	2,76	7,97				
GPOA	SIRTA	helbig	Perez	2018	1,97	6,79				
GPOA	SIRTA	helbig	Perez	2019	1,17	5,22				
GPOA	SIRTA	helbig	Perez	2020	1,89	6,51				

Variable	Source	Coef source	Step 3	Year	MBE	MAE	Mean MBE	Std MBE	Mean MAE	Std MAE
TPV	SIRTA	Fit	noct	2016	0,10	2,24	−0,143	0,331	2,448	0,165
TPV	SIRTA	Fit	noct	2017	0,12	2,45				
TPV	SIRTA	Fit	noct	2018	−0,65	2,40				
TPV	SIRTA	Fit	noct	2019	−0,30	2,70				
TPV	SIRTA	Fit	noct	2020	0,02	2,45				
TPV	SIRTA	Fit	advnoct	2016	0,20	1,63	−0,118	0,279	1,820	0,139
TPV	SIRTA	Fit	advnoct	2017	0,02	1,73				
TPV	SIRTA	Fit	advnoct	2018	−0,44	1,93				
TPV	SIRTA	Fit	advnoct	2019	−0,39	1,96				
TPV	SIRTA	Fit	advnoct	2020	0,01	1,85				
TPV	SIRTA	Fit	Faiman	2016	0,43	1,70	0,054	0,310	1,848	0,111
TPV	SIRTA	Fit	Faiman	2017	0,25	1,78				
TPV	SIRTA	Fit	Faiman	2018	−0,26	1,92				
TPV	SIRTA	Fit	Faiman	2019	−0,27	1,98				
TPV	SIRTA	Fit	Faiman	2020	0,13	1,86				
TPV	SIRTA	Fit	Sandia	2016	0,41	1,70	0,035	0,310	1,832	0,100
TPV	SIRTA	Fit	Sandia	2017	0,23	1,77				
TPV	SIRTA	Fit	Sandia	2018	−0,26	1,90				
TPV	SIRTA	Fit	Sandia	2019	−0,30	1,95				
TPV	SIRTA	Fit	Sandia	2020	0,10	1,84				
TPV	SIRTA	Literature	noct	2016	0,79	2,29	0,626	0,296	2,420	0,151
TPV	SIRTA	Literature	noct	2017	0,81	2,49				
TPV	SIRTA	Literature	noct	2018	0,13	2,26				
TPV	SIRTA	Literature	noct	2019	0,57	2,63				
TPV	SIRTA	Literature	noct	2020	0,83	2,43				
TPV	SIRTA	Literature	Faiman	2016	−3,87	4,03	−4,734	0,769	4,874	0,763
TPV	SIRTA	Literature	Faiman	2017	−4,04	4,18				
TPV	SIRTA	Literature	Faiman	2018	−5,15	5,30				
TPV	SIRTA	Literature	Faiman	2019	−5,70	5,84				
TPV	SIRTA	Literature	Faiman	2020	−4,91	5,02				
TPV	SIRTA	Literature	Sandia	2016	−2,56	2,94	−3,188	0,591	3,522	0,549
TPV	SIRTA	Literature	Sandia	2017	−2,62	3,00				
TPV	SIRTA	Literature	Sandia	2018	−3,65	3,94				
TPV	SIRTA	Literature	Sandia	2019	−3,87	4,17				
TPV	SIRTA	Literature	Sandia	2020	−3,24	3,56				

Variable	Source	Step 4	Year	rMBE	rMAE	Mean rMBE	Std rMBE	Mean rMAE	Std rMAE
PMPP	SIRTA	ANN	2016	−0,02	2,77	0,031	0,473	3,127	0,456
PMPP	SIRTA	ANN	2017	−0,39	3,37				
PMPP	SIRTA	ANN	2018	−0,35	3,81				
PMPP	SIRTA	ANN	2019	0,14	2,74				
PMPP	SIRTA	ANN	2020	0,77	2,95				
PMPP	SIRTA	Evans	2016	0,24	3,34	0,751	0,545	3,509	0,214
PMPP	SIRTA	Evans	2017	0,16	3,71				
PMPP	SIRTA	Evans	2018	0,87	3,55				
PMPP	SIRTA	Evans	2019	1,04	3,23				
PMPP	SIRTA	Evans	2020	1,45	3,71				
PMPP	SIRTA	Simple	2016	15,00	15,15	16,565	1,718	16,757	1,687
PMPP	SIRTA	Simple	2017	14,42	14,70				
PMPP	SIRTA	Simple	2018	18,06	18,18				
PMPP	SIRTA	Simple	2019	17,57	17,75				
PMPP	SIRTA	Simple	2020	17,79	18,00				
PMPP	SIRTA	Adv. simple	2016	2,71	6,04	4,115	1,534	6,850	0,997
PMPP	SIRTA	Adv. simple	2017	2,20	5,65				
PMPP	SIRTA	Adv. simple	2018	5,45	7,61				
PMPP	SIRTA	Adv. simple	2019	5,01	6,96				
PMPP	SIRTA	Adv. simple	2020	5,21	7,99				
PMPP	SIRTA	Statistque	2016	−0,43	3,48	−0,024	0,634	3,513	0,284
PMPP	SIRTA	Statistque	2017	−0,91	3,63				
PMPP	SIRTA	Statistque	2018	0,54	3,90				
PMPP	SIRTA	Statistque	2019	0,15	3,13				
PMPP	SIRTA	Statistque	2020	0,54	2,43				
PMPP	SIRTA	SDM	2016	0,16	3,02	0,704	0,540	3,084	0,470
PMPP	SIRTA	SDM	2017	1,22	3,54				
PMPP	SIRTA	SDM	2018	1,27	3,53				
PMPP	SIRTA	SDM	2019	0,70	2,91				
PMPP	SIRTA	SDM	2020	0,17	2,42				

Figure 9: Computed rMBE and rMAE for SIRTA in situ ground-based measurements, for each modeling step. Colors are performance indicators (from green for the best to red for the worse).

(2) Step 2 ⟶ from GHI, DHI, and BNI to G_{POA}: 6 models

(3) Step 3 ⟶ from G_{POA}, T_{amb}, and WS to T_{PV}: 4 models in 2 configurations

(4) Step 4 ⟶ from G_{POA} and T_{PV} to P_{MPP}: 6 models

Step 1 is the one that leads to the greatest uncertainty (17%). As the output variables of this step are only intermediate variables, if we combine steps 1 and 2, then, the error drops to 5.3%. Moreover, it is the estimation of G_{POA}, the step which induces the greatest error (1.8 for step 3 and 3.5% for step 4).

This remark remains valid if we are interested in the estimation of P_{MPP}. Indeed, the modifications from measured G_{POA} (step 4) to estimated G_{POA} (steps 1 and 2 and 4) induces an increase of the error from 3.5% to 6.9%.

The temperature does not have a great influence in the estimation of P_{MPP}. Whether estimated (steps 3 and 4) or measured (step 4), the error remains stable around 3.5%.

For the global simulation, our study was interested in the performances obtained according to 3 sources of environmental data (GHI, T_{amb}, and WS). As might be expected, the closer this data is to the PV installation, the better the performance is. For example, for local measurements, the error is 6.6%. If data from the nearest weather station is used, the error increases to 12.1%, and for data estimated from satellite images, then, the error becomes 15.2%.

In conclusion, the proposed study makes it possible to evaluate the simulation quality of the P_{MPP} of a PV installation. Each step was evaluated independently and successively, according to 3 data sources. All the computed data are presented in Appendix B, and the raw data are available upon request.

The next steps will be to show the link between all the models and the error propagation, consider different meteorological forecast models, consider different time horizons, go through PV plants instead of a unique PV module, and compare with commercial tools, for instance.

Input	Source	Steps				Year	rMBE	rMAE	Mean rMBE	std rMBE	Mean rMAE	std rMAE
GHI, Tamb, WS	SIRTA	✓	✓	✓	✓	2016	1,74	6,49	2,602	0,804	6,663	0,394
GHI, Tamb, WS	SIRTA	✓	✓	✓	✓	2017	1,71	6,67				
GHI, Tamb, WS	SIRTA	✓	✓	✓	✓	2018	3,22	7,12				
GHI, Tamb, WS	SIRTA	✓	✓	✓	✓	2019	3,08	6,10				
GHI, Tamb, WS	SIRTA	✓	✓	✓	✓	2020	3,26	6,92				
GHI, Tamb, WS	SIRTA	✓	✓	✓	✓	2016	−1,71	16,32	−1,465	1,352	12,089	6,984
GHI, Tamb, WS	ORLY	✓	✓	✓	✓	2017	−2,87	15,80				
GHI, Tamb, WS	ORLY	✓	✓	✓	✓	2018	0,38	14,57				
GHI, Tamb, WS	ORLY	✓	✓	✓	✓	2019	−1,67	1,67				
GHI, Tamb, WS	ORLY	✓	✓	✓	✓	2020						
GHI, Tamb, WS	CAMS	✓	✓	✓	✓	2016	9,35	18,40	4,959	2,554	15,164	6,984
GHI, Tamb, WS	CAMS	✓	✓	✓	✓	2017	4,21	16,42				
GHI, Tamb, WS	CAMS	✓	✓	✓	✓	2018	4,51	14,78				
GHI, Tamb, WS	CAMS	✓	✓	✓	✓	2019	4,06	12,73				
GHI, Tamb, WS	CAMS	✓	✓	✓	✓	2020	2,67	13,49				
GHI, TPV	SIRTA	✓	✓		✓	2016	1,79	6,94	2,407	0,670	6,947	0,447
GHI, TPV	SIRTA	✓	✓		✓	2017	1,61	7,00				
GHI, TPV	SIRTA	✓	✓		✓	2018	2,81	7,37				
GHI, TPV	SIRTA	✓	✓		✓	2019	2,69	6,21				
GHI, TPV	SIRTA	✓	✓		✓	2020	3,13	7,22				
GHI, TPV	ORLY	✓	✓		✓	2016	−1,52	17,52	−1,588	1,117	12,894	7,538
GHI, TPV	ORLY	✓	✓		✓	2017	−2,95	16,84				
GHI, TPV	ORLY	✓	✓		✓	2018	−0,22	15,57				
GHI, TPV	ORLY	✓	✓		✓	2019	−1,65	1,65				
GHI, TPV	ORLY	✓	✓		✓	2020	NaN	NaN				
GHI, TPV	CAMS	✓	✓		✓	2016	9,76	19,88	4,655	2,937	16,558	7,538
GHI, TPV	CAMS	✓	✓		✓	2017	3,90	17,89				
GHI, TPV	CAMS	✓	✓		✓	2018	3,82	16,15				
GHI, TPV	CAMS	✓	✓		✓	2019	3,62	14,20				
GHI, TPV	CAMS	✓	✓		✓	2020	2,18	14,67				
GHI, DHI, BNI, Tamb, WS	SIRTA		✓	✓	✓	2016	−0,13	5,65	0,676	0,992	5,922	0,443
GHI, DHI, BNI, Tamb, WS	SIRTA		✓	✓	✓	2017	−0,65	6,01				
GHI, DHI, BNI, Tamb, WS	SIRTA		✓	✓	✓	2018	1,48	6,62				
GHI, DHI, BNI, Tamb, WS	SIRTA		✓	✓	✓	2019	1,34	5,46				
GHI, DHI, BNI, Tamb, WS	SIRTA		✓	✓	✓	2020	1,34	5,87				
GHI, DHI, BNI, Tamb, WS	ORLY		✓	✓	✓	2016						
GHI, DHI, BNI, Tamb, WS	ORLY		✓	✓	✓	2017						
GHI, DHI, BNI, Tamb, WS	ORLY		✓	✓	✓	2018						
GHI, DHI, BNI, Tamb, WS	ORLY		✓	✓	✓	2019						
GHI, DHI, BNI, Tamb, WS	ORLY		✓	✓	✓	2020						
GHI, DHI, BNI, Tamb, WS	CAMS		✓	✓	✓	2016	10,80	18,80	6,256	2,651	15,380	NaN
GHI, DHI, BNI, Tamb, WS	CAMS		✓	✓	✓	2017	5,90	16,80				
GHI, DHI, BNI, Tamb, WS	CAMS		✓	✓	✓	2018	5,72	15,00				
GHI, DHI, BNI, Tamb, WS	CAMS		✓	✓	✓	2019	4,87	12,90				
GHI, DHI, BNI, Tamb, WS	CAMS		✓	✓	✓	2020	3,99	13,40				
GPOA, Tamb, WS	SIRTA			✓	✓	2016	0,11	3,36	0,849	0,630	3,420	0,135
GPOA, Tamb, WS	SIRTA			✓	✓	2017	0,21	3,52				
GPOA, Tamb, WS	SIRTA			✓	✓	2018	1,23	3,53				
GPOA, Tamb, WS	SIRTA			✓	✓	2019	1,35	3,21				
GPOA, Tamb, WS	SIRTA			✓	✓	2020	1,34	3,48				
GPOA, Tamb, WS	ORLY			✓	✓	2016						
GPOA, Tamb, WS	ORLY			✓	✓	2017						
GPOA, Tamb, WS	ORLY			✓	✓	2018						
GPOA, Tamb, WS	ORLY			✓	✓	2019						
GPOA, Tamb, WS	ORLY			✓	✓	2020						
GPOA, Tamb, WS	CAMS			✓	✓	2016						
GPOA, Tamb, WS	CAMS			✓	✓	2017						
GPOA, Tamb, WS	CAMS			✓	✓	2018						
GPOA, Tamb, WS	CAMS			✓	✓	2019						
GPOA, Tamb, WS	CAMS			✓	✓	2020						
GPOA, Tpv	SIRTA				✓	2016	0,24	3,34	0,751	0,545	3,509	0,214
GPOA, Tpv	SIRTA				✓	2017	0,16	3,71				
GPOA, Tpv	SIRTA				✓	2018	0,87	3,55				
GPOA, Tpv	SIRTA				✓	2019	1,04	3,22				
GPOA, Tpv	SIRTA				✓	2020	1,34	3,71				
GPOA, Tpv	SIRTA				✓	2016						
GPOA, Tpv	ORLY				✓	2017						
GPOA, Tpv	ORLY				✓	2018						
GPOA, Tpv	ORLY				✓	2019						
GPOA, Tpv	ORLY				✓	2020						
GPOA, Tpv	CAMS				✓	2016						
GPOA, Tpv	CAMS				✓	2017						
GPOA, Tpv	CAMS				✓	2018						
GPOA, Tpv	CAMS				✓	2019						
GPOA, Tpv	CAMS				✓	2020						

Figure 10: Computed rMBE and rMAE for SIRTA in situ ground-based measurements, ORLY weather station ground-based measurements, and CAMS satellites images, for each modeling step. Colors are performance indicators (from green for the best to red for the worse).

Appendix

A. Annex 1 Data acquisition equipment and sensors (Table 4).

B. Annex 2 Computed data used in Section 2 (Figure 9).

C. Annex 3 Computed data used in Section 3 (Figure 10).

Abbreviations

AOI: Angle of incidence
BNI: Hourly beam (direct) normal irradiance (Wm^{-2})
CE: Conversion efficiency
DHI: Hourly diffuse horizontal irradiance (Wm^{-2})
GHI: Hourly global horizontal irradiance (Wm^{-2})
G_{POA}: Hourly global plane of array irradiance (Wm^{-2})
MAE: Mean absolute error
MBE: Mean bias error
NOCT: Nominal operating cell temperature (°C)
P_{MPP}: Hourly power at the maximum power point (W)
T_{PV}: Hourly PV module temperature (°C)
T_{amb}: Hourly ambient temperature (°C)
WS: Hourly wind speed (ms^{-1})
k_T: Hourly clearness index
rMAE: Relative mean absolute error
rMBE: Relative mean bias error
ANN: Artificial neurons network
BSRN: Baseline surface radiation network
CAMS: Copernicus atmosphere monitoring service
ORLY: Ground-based measurements weather station
PV: Photovoltaic
PVUSA: Photovoltaics for utility scale applications
SDM: Single-diode electrical model
SIRTA: Site Instrumental de Recherche par Télédétection Atmosphérique (Atmospheric Research Observatory, in English)
STC: Standard test conditions.

Data Availability

Atmospherical data are available with free access: -from satellite images: http://www.soda-pro.com/web-services/radiation/cams-radiation-service-from meteorological weather stations: https://donneespubliques.meteofrance.fr/-from in situ measurements: https://bsrn.awi.de. Photovoltaic data used to support the results of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work benefited from the support of the Energy4Climate Interdisciplinary Center (E4C) of IP Paris and École des Ponts ParisTech. It was supported by 3rd Programme d'Investissements d'Avenir (ANR-18-EUR-0006-02) and by the Foundation of École polytechnique (Chaire "Defis Technologiques pour une Energie Responsable" financed by total).

References

[1] E. Ela, V. Diakov, E. Ibanez, and M. Heaney, *Impacts of variability and un certainty in solar photovoltaic generation at multiple timescales, tech. Rep. NREL/TP-5500-58274*, National Renewable Energy Laboratory, 2013.

[2] M. Rekinger, I.-T. Theologitis, G. Masson et al., "Connecting the sun," in *solar photovoltaic on the road to largescale grid integrationt Tech. rep*, EPIA, 2012.

[3] B. Barth, G. Concas, E. B. Zane et al., *PV Grid : Final Project Report*, BSW-Solar, 2014.

[4] D. Thevenard and S. Pelland, "Estimating the uncertainty in long-term photovoltaic yield predictions," *Solar Energy*, vol. 91, pp. 432–445, 2013.

[5] M. Schnitzer, *Reducing uncertainty in solar energy estimates: a case study*, Sandia PV Performance Modeling Workshop, Santa Clara, Canada, 2013.

[6] B. Müller, L. Hardt, A. Armbruster, K. Kiefer, and C. Reise, "Yield predictions for photovoltaic power plants: empirical validation, recent advances and remaining uncertainties," *Progress in Photovoltaics: Research and Applications*, vol. 24, no. 4, pp. 570–583, 2016.

[7] G. Makrides, B. Zinsser, M. Schubert, and G. E. Georghiou, "Energy yield prediction errors and uncertainties of different photovoltaic models," *Progress in Photovoltaics : Research and Applications*, vol. 21, no. 4, pp. 500–516, 2013.

[8] G. Makrides, B. Zinsser, A. Phinikarides et al., "Photovoltaic model uncertainties based on field measurements," in *2011 37th IEEE Photovoltaic Specialists Conference*, pp. 002386–002390, Seattle, WA, USA, 2011.

[9] A. Migan, F. Calderon Obaldia, J. Badosa, V. Bourdin, Y. Bonnassieux, and L. Dumbia, "Step-by-step evaluation of photovoltaic module performance related to outdoor parameters : evaluation of the uncertainty," in *2017 IEEE 44th Photovoltaic Specialist Conference (PVSC)*, pp. 626–631, Washington, USA, 2017.

[10] M. Haeffelin, L. Barthès, O. Bock et al., "SIRTA, a ground-based atmospheric observatory for cloud and aerosol research," *Annali di geofisica (Bologna, Italy : 1993)*, vol. 23, no. 2, pp. 253–275, 2005.

[11] D. Chigara, *Etudes de l'échange thermique des modules photovoltaïques: mise en place d'un protocole de calibration de sondes platine pour la mesure de température et thermographie infrarouge, Tech. rep.*, LGEP, 2014.

[12] "Baseline Surface Radiation Network (BSRN)," 2018, http://bsrn.awi.de.

[13] Z. Qu, A. Oumbe, P. Blanc et al., "Fast radiative transfer parameterisation for assessing the surface solar irradiance: the heliosat-4 method," *Energy Meteorology*, vol. 26, no. 1, pp. 33–57, 2017.

[14] M. Schroedter-Homscheidt, "The Copernicus atmosphere monitoring service (CAMS) radiation service in a nutshell," in *22nd SolarPACES Conference 2016*, UAE, Abu Dhabi, United Arab Emirates, 2016.

[15] N. Helbig, *Application of the radiosity approach to the radiation balance in complex terrain*, Thesis at University of Zurich, 2009.

[16] D. G. Erbs, S. A. Klein, and J. A. Duffie, "Estimation of the diffuse radiation fraction for hourly, daily and monthly- average global radiation," *Solar Energy*, vol. 28, no. 4, pp. 293–302, 1982.

[17] H. Hottel and B. B. Woertz, "Performance of Flat Plate Solar Heat Collectors," *Transactions of the American Society of Mechanical Engineers*, vol. 64, pp. 91–104, 1955.

[18] T. Klucher, "Evaluation of models to predict insolation on tilted surfaces," *Solar Energy*, vol. 23, no. 2, pp. 111–114, 1979.

[19] D. L. King, W. E. Boyson, and J. A. Kratochvil, *Photovoltaic array performance model, tech. Rep.*, Sandia National Laboratories, 2004.

[20] D. L. King, J. A. Kratochvil, and W. E. Boyson, "Measuring solar spectral and angle-of-incidence e_ects on pv modules and solar irradiance sensors," in *Proceedings of the 26th IEEE PV Specialists Conference*, pp. 1113–1116, Anaheim, CA, USA, 1997.

[21] J. Hay and J. Davies, "Calculation of the solar radiation incident on an inclined surface," in *Proceedings of the first Canadian solar radiation data workshop*, pp. 373–380, Toronto, Canada, 1980.

[22] R. Perez, P. Ineichen, R. Seals, J. Michalsky, and R. Stewart, "Modeling daylight availability and irradiance components from direct and global irradiance," *Solar Energy*, vol. 44, no. 5, pp. 271–289, 1990.

[23] R. Perez, R. Seals, P. Ineichen, R. Stewart, and D. Menicucci, "A new simplified version of the perez diffuse irradiance model for tilted surfaces," *solar energy*, vol. 39, no. 3, pp. 221–231, 1987.

[24] R. Perez, R. Stewart, R. Seals, and T. Guertin, *The development and veri_-cation of the perez di_use radiation model, Tech. rep.*, Sandia National Laboratories, 1988.

[25] D. Reindl, W. Beckman, and J. A. Duffie, "Evaluation of hourly tilted surface radiation models," *Solar Energy*, vol. 45, no. 1, pp. 9–17, 1990.

[26] R. Ross, "Flat-plate photovoltaic array design optimization," in *14th IEEE Photovoltaic Specialists Conference*, San Diego, CA, 1980.

[27] V. Bourdin, J. Badosa, F. Calderon-Obaldia, C. Abdel-Nous, A. Migan, and Y. Du Bois, *Bonnassieux, Comparaison de modèles de complexité croissante pour la simulation de la température de modules photovoltaïques*, Journ ées Nationales du PhotoVoltaïque, 2017.

[28] M. K. Fuentes, *A simpli_ed thermal model of photovoltaic modules, tech. Rep.*, Sandia National Laboratories, 1985.

[29] D. Faiman, "Assessing the outdoor operating temperature of photovoltaic modules," *Progress in Photovoltaics : Research and Applications*, vol. 16, p. 307, 2018.

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

[31] A. Labouret, P. Cumunel, J.-P. Braun, and B. Faraggi, *Cellules solaires : Lesbases de l'énergie photovoltaïque*, ETSF, Dunod, dunod Edition edition, 2010.

[32] D. L. Evans, "Simplified method for predicting photovoltaic array output," *Solar Energy*, vol. 27, no. 6, pp. 555–560, 1981.

[33] G. Graditi, S. Ferlito, and G. Adinolfi, "Comparison of photovoltaic plant power production prediction methods using a large measured dataset," *Renewable Energy*, vol. 90, pp. 513–519, 2016.

[34] G. R. Walker, "Evaluating mppt converter topologies using a matlab pv model," *Australian Journal of Electrical ans Electronics Engineering*, vol. 21, no. 1, 2001.

[35] A. Mellit and S. A. Kalogirou, "Artificial intelligence techniques for photovoltaic applications: A review," *Progress in Energy and Combustion Science*, vol. 34, no. 5, pp. 574–632, 2008.

[36] M. Karamirad, M. Omid, R. Alimardani, H. Mousazadeh, and S. N. Heidari, "ANN based simulation and experimental verification of analytical four- and five-parameters models of PV modules," *Simulation Modelling Practice and Theory*, vol. 34, pp. 86–98, 2013.