Combining physical models to estimate PV power: evaluation and optimal modeling in the solar resource-rich semi-arid Brazilian region

João Victor Furtado F. de Medeiros[a]*, Emerson Torres Aguiar Gomes[a], Rodrigo Alonso-Suárez[b], Olga de Castro Vilela[a]

Center for Renewable Energy (CER), Federal University of Pernambuco (UFPE), Recife, Brazil[a]
Laboratorio de Energía Solar (LES), CENUR Litoral Norte, Universidad de la República, Uruguay[b]

* Corresponding author
E-mail address: joao.furtado@ufpe.br (João Victor Furtado F. de Medeiros)

Abstract:

Accurate estimation of energy production in photovoltaic power plants is crucial for project feasibility assessment and O&M practices. This study evaluates and analyzes the impact of combining different physical models for PV power modeling, varying different techniques for global horizontal irradiance (GHI) separation, irradiance transposition, and optical, thermal and electrical modeling. High-resolution data collected at one-minute intervals from a 2.5 MWp PV plant located in the Brazilian semi-arid region are used. The PV generation is examined and modeled based on ground-measured GHI, considering a total of 11,340 possible combinations, through seven separation models, nine transposition models, four optical models, nine thermal models, and five electrical models. It is observed that the selection of physical models significantly impacts the estimation, when adopting inaccurate physical models relative differences of 49% in nMAE and 26% in nRMSE were evidenced. The models which achieved the best results among the top performers were Starke2 separation model, Perez's transposition model, Martin-Ruiz's optical model, Sandia or Mattei's thermal model and De Soto's electrical model. Additionally, selecting adequate models based on the literature proved to be a good choice for modeling, almost achieving the optimal performance of the best combinations.

Keywords: Photovoltaic modeling; Physical models; Grid-connected PV plants; 1-min data; GHI separation; Brazilian semi-arid;

1. Introduction

Photovoltaic (PV) energy has witnessed a remarkable surge in recent years, reaching 1411 GW in 2023 [1] with a projection of 5457 GW in 2030 under the 1.5°C scenario [2]. Its growth has been accelerated by implementing incentive policies in various countries, leading to technological advancements and cost reductions throughout the entire value chain of PV solar energy. The growing expansion of the photovoltaic market has led to a substantial increase in the number and size of photovoltaic systems, as well as in investment allocations and the associated risk-return dynamics [3]. Consequently, it becomes increasingly crucial for designers, investors, and financial institutions to identify and mitigate technical risks that could impact the feasibility and operation of these projects [4]. In grid-connected photovoltaic systems, the quantification of solar energy production is essential during the development and operational phases. In the development phase, it is required to assess the potential and profitability of PV power plants by estimating the expected site-specific performance. During operation, it enables the evaluation of overall system performance by comparing expected generation from modeling with actual measured generation [5].

The assessment of PV system performance and energy generation relies on various modeling tools, free or paid, designed to simulate electricity generation. These tools serve as valuable resources for solar PV designers and operators. They employ a series of chained models to predict energy generation based on the specifications of the PV system and meteorological data. While these simulations typically operate at an hourly resolution, recent studies have shown the importance of sub-hourly resolution, particularly minute-resolution, for accurately sizing inverters [6], implementing operational best practices [7], and conducting accurate simulations of photovoltaic systems [8, 9].

Several models exist in the literature, and their application is strongly dependent on the climatic conditions of the site [10]. Evaluating models that best suit the local climate and the photovoltaic system characteristics is critical. This ensures accurate estimation of the PV energy production, thereby minimizing uncertainties and contractual risks. Another aspect that may impact accurate estimation is the photovoltaic system losses, represented by the derating factor. An incorrect derating factor can result in significant errors in PV estimates [11]. An over-optimistic value may overestimate energy production, risking project viability. Conversely, a conservative factor can distort estimates, causing misguided investments. Thevenard and Pelland [12] discussed the uncertainty involved in the performance evaluation of large PV systems; the authors selected a derating factor of 3% with a 2% margin of uncertainty and observed that a deeper comprehension of some losses, such as dirt and soiling, could have improved estimates quality. The derating factor may exhibit variations in its magnitude depending on climatic conditions [13]. Different derating factors can affect the bias behavior in the PV modeling chain. Thus, assessing the performance of physical models against potential derating factors remains essential. Despite all the above, little attention has been paid in the literature to 1-minute PV simulations and how to optimally couple the physical sub-models for PV power estimation.

In Hofmann and Seckmeyer [8], the influence of various irradiance models and their combinations on the simulation of PV systems was evaluated for the cold semi-arid climate (BSk) and other climates with distinct characteristics. The authors examined solar irradiance direct-diffuse separation and tilted-plane transposition models at different temporal resolutions (1-hour and 1-minute), observing that the PV system simulations should ideally utilize 1-minute data either measured or derived from hourly values, the latter referred as synthetic 1-minute data [14]. The authors emphasize the importance of employing simulations with high temporal resolution data for accurate photovoltaic system sizing, enabling the estimation of losses due to inverter clipping, as observed by Burger and Rüther [15]. Additionally, Hofmann and Seckmeyer highlight in [8] the significant impact of transposition models on modeling PV systems, showing the need to carefully select these models, considering that radiation modeling is highly dependent on the location. Mayer and Gróf in [10] conducted a more extensive analysis in terms of physical models, incorporating different radiation models (separation and transposition) and optical, thermal and electrical models for estimating PV generation based on 15-minute mesoscale Numerical Weather Prediction (NWP) data. The analysis was performed for photovoltaic systems installed in Hungary, which has high latitudes and typical temperate and continental climates (Cfa, Cfb, Dfa and Dfb). The authors observed that from NWP data, the most critical steps were GHI separation and irradiance transposition to the Plane of Array (POA). In addition, they observed that model selection has a high effect on simulation accuracy, reaching differences of up to 13% in the mean absolute error and 12% in the nRMSE.

As researchers delve deeper into PV power modeling, there is a noticeable gap in the literature on several key issues that have not been addressed. These include questions about the impact of physical model selection on the accuracy of PV estimates in different climates, identifying critical modeling steps for accurate 1-minute simulation, and the effects of erroneous selection of the derating factor on the performance of physical models. These issues underscore the need to expand research efforts and experimental studies to clarify key aspects of PV modeling.

1.1. Article's contribution

The present work evaluates and analyzes the impact of combining different physical models for PV power modeling in a site located at the large solar resource-rich semi-arid region of Brazil, also known as Sertão, varying the different techniques for global horizontal irradiance (GHI) separation, irradiance transposition, and optical, thermal and electrical modeling of the PV systems. It assesses the results of more than 11,340 combinational simulations compared to 1-minute data from a 2.5 MWp PV plant located in Petrolina-PE (9.11°S, 40.44°W) located in the Brazilian semi-arid region. Comparative analysis of physical models allows the identification of models that perform better in high temporal resolution, highlighting critical steps for accurate estimation and potential implications that physical models may present under different derating factor scenarios. In this manner, the main contributions of this work can be summarized as follows:

- Compares 7 GHI separation models (designed at hourly or minutely resolution), 9 diffuse irradiance transposition models to the inclined plane, 4 optical, 9 thermal and 5 electrical models, identifying critical steps in PV generation modeling and techniques that tend to have better or worse performance.
- Quantifies the impact that choosing inaccurate physical models can have on the simulation of PV systems based on high temporal resolution data.
- Compare the results of the best-fit individual and coupled models with results from the literature, observing that the application of accurate models individually allows satisfactory statistical results to be achieved with a high level of reliability, but not necessarily the best combination.
- Provides a first representative evaluation of the performance of physical models under different loss scenarios, demonstrating that the average behavior of the models is not affected by the derating factor adopted.

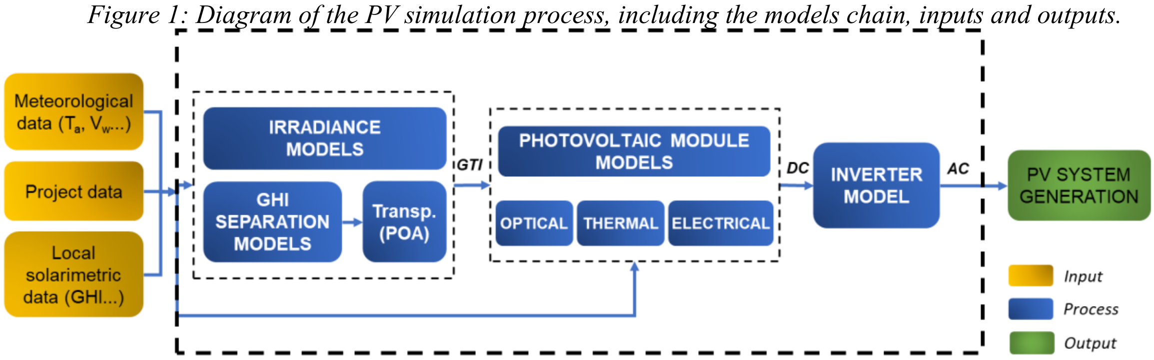
1.2. Article's outline

The paper is organized as follows. Section 2 presents the state of the art regarding the models that are typically chained in the process of simulating PV systems, parting from GHI to PV power estimation. Section 3 covers the data treatment and qualification, the models combinatorial approach, and the statistical indicators selected for the assessment. Section 4 presents the analysis of estimating the PV power generation using the different model combinations, in which insights are provided from the 11,340 possible cases. Finally, the conclusions and future prospects are summarized in Section 5.

2. Photovoltaic systems modeling

Modeling PV systems is structured into two macro-steps: one involves quantifying the irradiance on the modules, and the other relates to the optical, thermal and electrical modeling of the panels and inverters. Figure 1 illustrates the diagram of the PV power estimation process, grouping the models into conceptual categories (irradiance models, PV module models, inverter model). The inputs are solarimetric, meteorological and project data, as well as the PV losses adopted by the designer (discussed in section 2.4). The solarimetric data is then processed by the irradiance models, subdivided into GHI separation models and transposition models to obtain the Global Tilted Irradiance (GTI), also referred in the literature as the plane of array (POA) irradiance. Once the modules' incident irradiance is estimated, the optical, thermal and electrical models are applied to obtain the direct current (DC) generation to be injected into the inverter. Finally, the inverter model performs the DC-AC conversion, outputting the power generation.

Figure 1: Diagram of the PV simulation process, including the models chain, inputs and outputs.



2.1. Irradiance Models

In order to use the transposition models to estimate GTI, knowledge of the direct and diffuse components of the GHI is required. In many places, the direct and diffuse components are not measured due to the costs and effort involved in using tracking devices [16]. Therefore, designers often use GHI data, which needs to be divided into Direct Normal Irradiance (DNI) and Diffuse Horizontal Irradiance (DHI). Various models exist for this purpose, but their accuracy in estimating radiation on the tilted plane can vary based on local climate, sky conditions and local characteristics [17].

2.1.1. GHI Separation Models

The diffuse and direct components can be derived from GHI by examining the relationship between the diffuse fraction (k_d, the ratio of DHI to GHI) and the clearness index (k_t, the ratio of GHI to extraterrestrial irradiance on the horizontal plane), or by analyzing the transmittance of direct normal irradiance (k_n, the ratio of DNI to incident extraterrestrial irradiance on the plane normal to radiation) as a function of k_t. Once the relationships between k_d *vs.* k_t or k_n *vs.* k_t are established, it becomes possible to determine the diffuse or direct component accordingly. Subsequently, by using the irradiance relationship, also known as the closure equation, presented in Equation 1, it is possible to obtain the missing radiation component.

$$GHI = DNI \cos \theta_z + DHI \qquad (1)$$

Where θ_z is the solar zenith angle, corresponding to the angle between the local zenith and the Sun's center direction.

A pioneering minute-based separation model was proposed by [18], referred to in this paper as ENGERER2, which considers the cloud enhancement events (over-irradiance). In [19], this model is reparametrized, giving rise to the ENGERER4 model, which was proposed for various temporal resolutions (1-min, 5-min, 10-min, 15-min, 30-min, 1-h, and 1-day). However, as observed in Manni *et al.* [20], and Yang [21], this model has less ability to describe cloud enhancement events, often yielding inferior results to the original ENGERER2 model [21] and also to well-established hourly models in the literature such as SKARTVEIT or PEREZ [20].

Yang and Boland [22] proposed a modification to the ENGERER2 model, adding the hourly or 30-minute diffuse fraction obtained from satellite data. Due to the increased complexity of implementation, Yang [23] suggests modifying this model, replacing the satellite data diffuse fraction with the k_d obtained from the ENGERER2 model at hourly resolution ($k_{d,h}^{Engerer2}$). This parameter aims to describe low-frequency variations, similarly to a variability index. With this alteration, the model referred to as YANG4 is expressed as Equation 2, where *AST* is the apparent solar time and Δk_{tc} corresponds to the difference between the measured k_t and the clear sky index obtained from the clear sky model TJ (Threlkeld and Jordan, 1957), which presents consistent results according to the works of Sun *et al.* [24] and Bright and Engerer [19]. The k_{de} represents a portion of the diffuse fraction assigned to cloud enhancement events and is calculated as GHI minus GHI from TJ clear sky model (also known as clear sky irradiance, CSI) divided by GHI.

$$k_d^{YANG4} = C + \frac{1 - C}{1 + \exp(\beta_0 + \beta_1 k_t + \beta_2 AST + \beta_3 \theta_z + \beta_4 \Delta k_{tc} + \beta_6 k_d^{Engerer2})} + \beta_5 k_{de} \quad (2)$$

In [16], the Ridley, Boland, and Lauret [25] model, also known as BRL, is adapted by adding the parameter K_{CSI} to the equation. This parameter is defined as the ratio between *GHI* and the clear sky irradiance (*CSI*) obtained from the Solis model [26]. The authors define two models, the STARKE1 model for Australian data and the STARKE2 model parameterized for several sites in Brazil. Equation 3 presents the adopted model in this work.

$$k_d^{STARKE2} = \begin{cases} \dfrac{1}{\left[1 + \exp\left(\beta_0 + \beta_1 k_t + \beta_2 HSA + \beta_3 \alpha + \beta_4 K_t + \beta_5 \psi + \beta_6 \dfrac{CSI}{277{,}78}\right)\right]} & , K_{CSI} < 1{,}05 \\ \dfrac{1}{\left[1 + \exp\left(\beta_7 + \beta_8 k_t + \beta_9 HSA + \beta_{10} \alpha + \beta_{11} K_t + \beta_{12} \psi + \beta_{13} \dfrac{CSI}{277{,}78}\right)\right]} & , K_{CSI} \geq 1{,}05\ ;\ k_t > 0{,}65 \end{cases} \quad (3)$$

The K_{CSI} and k_t limits, presented by the STARKE2 model, are associated with over-irradiance events. In their subsequent work, Starke *et al.* [27] changed the limits of k_t, establishing the boundaries as follows: K_{CSI} must be equal to or greater than 1.05, and k_t should exceed 0.75. Additionally, the authors incorporate an hourly index into the model as Yang [23], which is the hourly clearness index ($k_{t,h}$) and adjust the model for different types of climates. The STARKE3 model can be observed in Equation 4.

$$k_d^{STARKE3} = \begin{cases} \dfrac{1}{\left[1 + \exp\left(\beta_8 + \beta_9 k_t + \beta_{10} HSA + \beta_{11} \alpha + \beta_{12} K_t + \beta_{13} \psi + \beta_{14} \dfrac{CSI}{277{,}78} + \beta_{15} k_{t,h}\right)\right]} & , K_{CSI} < 1{,}05 \\ \dfrac{1}{\left[1 + \exp\left(\beta_0 + \beta_1 k_t + \beta_2 HSA + \beta_3 \alpha + \beta_4 K_t + \beta_5 \psi + \beta_6 \dfrac{CSI}{277{,}78} + \beta_7 k_{t,h}\right)\right]} & , K_{CSI} \geq 1{,}05\ e\ k_t > 0{,}75 \end{cases} \quad (4)$$

The present study selected some hourly and sub-hourly models for GHI separation. The Erbs *et al.* [28] model was selected because it is easy to apply and widely used in softwares and in literature. This model was adopted as a baseline for comparison with more elaborate and complex models to assess whether the greater complexity in modeling diffuse and direct radiation can lead to gains in estimating PV generation. In addition, the hourly models of Skartveit *et al.* [29] and DIRINT [30] were adopted due to their good results in a location with a similar climate to the region of this work, BSh [31], as well as the ENGERER2, STARKE2, STARKE3 and YANG4 models designed at the 1-minute resolution.

2.1.2. **Transposition**

The information on global horizontal irradiance, direct normal irradiance and diffuse irradiance allows the estimation of global irradiance on the inclined plane. The global tilted irradiance (GTI) can be estimated from the sum of the direct and diffuse irradiance incident on the inclined plane, where the latter has two broad sources, namely, the diffuse irradiance from the sky dome and from the ground, as seen by the PV array. The GTI is thus calculated by the expression:

$$GTI = DNI * \cos(AOI) + DHI * SVF + RHI * GVF \quad (5)$$

The direct irradiance incident on the tilted plane is obtained from a simple geometric transformation, where the direct normal irradiance (DNI) component is multiplied by the cosine of the incidence angle (AOI). The sky-dome diffuse portion is obtained by the product of diffuse horizontal irradiance (DHI) and the sky view factor (SVF). The diffuse irradiance reflected by the ground that is seen by PV array can be obtained by the reflected horizontal irradiance (RHI) multiplied by the ground view factor (GVF). The RHI is the product of global horizontal irradiance (GHI) and the ground reflectance (ρ_g), also called albedo.

Among the three components in the calculation of GTI, excluding complex surroundings leading to ad-hoc reflections, the sky-dome diffuse irradiance is typically the most difficult to compute because it strongly depends on the cloudiness and clearness conditions of the atmosphere [32]. Several authors have studied this component from different approaches, from isotropic models that consider an homogeneous isotropic diffuse radiance distribution in the sky [33], to more complex and elaborated models that treat the circumsolar diffuse and/or the horizon brightness in more detail, called anisotropic models [34]. All these models evaluate the sky view factor (SVF) between the collecting surface and the visible part of the sky.

Historically, one of the pioneers and most widespread work in the literature is the isotropic model proposed by Liu and Jordan [35] with the sky diffuse irradiance incident on the sloping surface being given by DHI corrected by a sky view factor, represented by (1 + cos β)/2, where β is the inclination angle of the tilted surface. Koronakis [36] proposes a correction in Liu and Jordan's sky view factor, correcting the SVF to (2 + cos β)/3, in order to increase the DHI contribution to the tilted irradiance, increasing the estimation of the diffuse tilted irradiance (DTI) compared to LJ's isotropic model. Tian *et al.* [37] also proposed a change in SVF corresponding to (1 – β/180), with β in degrees. In Badescu [33], a 3D approach is performed and compared to the isotropic model of Liu and Jordan [35], showing that Badescu's model estimation with SVF of (3 + cos 2β)/4 was slightly more accurate for low slope surfaces at high latitudes.

The Hay and Davies [34] model integrates isotropic diffuse radiation with circumsolar radiation resulting from solar radiation scattering concentrated within the solar disk, incorporating the anisotropy index (F_{HD}). Temps and Coulson [38] introduced a correction factor for isotropic diffuse radiation to address horizon brightness, later modified by Klucher [39] into a modulating function (*F*) for a comprehensive "all sky" model. Reindl *et al.* [40] enhanced the Hay and Davies model by introducing a horizon brightening term with a different modulating factor approach, leading to the HDKR model, combining the previous insights from these contributions, this model presented great results in the comparison of 26 models performed in Nassar *et al*. [41].

A different approach with good results [42] for low latitude inclinations (lower than 16°) is the Muneer [43] model. Muneer [43] proposed a model distinguishing between overcast and non-overcast sky conditions, relating diffuse radiation at an inclined surface to DHI, with parameters adjusted based on location. Another relevant anisotropic model is the widely used Perez *et al.* [44] model, where the isotropic, circumsolar and horizon brightness diffuse parts are examined in more detail. In this model, the coefficients representing solid angles of the circumsolar region and empirical sky brightness functions describing circumsolar radiation and horizon brightness are used. Table 1 provides sky view factors for all transposition models examined in this paper.

Table 1: Transposition models used to estimate the diffuse irradiance on tilted plane.

CODE	TRANSPOSITION MODEL	COMMENTS
LJ	$SVF_{LJ} = \frac{1 + \cos \beta}{2}$	
Ko	$SVF_{Ko} = \frac{2 + \cos \beta}{3}$	
Ba	$SVF_{Ba} = \frac{3 + \cos (2\beta)}{4}$	
Ti	$SVF_{Ti} = 1 - \frac{\beta}{180}$	*β in degrees*
Klu	$SVF_{Klu} = \left(\frac{1 + \cos \beta}{2}\right)\left(1 + F\left(\frac{\beta}{2}\right)\right) * [1 + FAI\ (90 - \alpha)\ ]$	$F = 1 - \left(\frac{DHI}{GHI}\right)^2$
HD	$SVF_{HD} = \left[(1 - F_{HD})\left(\frac{1 + \cos \beta}{2}\right) + F_{HD}R_b\right]$	$F_{HD} = \frac{DNI}{DNI_{ext}}$ and $R_b = \frac{\cos AI}{\cos (\theta_z)}$
Mu	$SVF_{Mu} = TF(1 - F_{HD}) + F_{HD}R_b$	*The fitting coefficients of the TF equation were considered based on the parametrization for the globe.*
Re	$SVF_{Re} = \left[(1 - F_{HD})\left(\frac{1 + \cos \beta}{2}\right) * \left(1 + f\left(\frac{\beta}{2}\right)\right) + F_{HD}R_b\right]$	$f = \sqrt{\frac{DNI \cos \theta_z}{GHI}}$
Pe	$SVF_{Pe} = \left[\left(\frac{1 + \cos \beta}{2}\right)(1 - F_1) + F_1\frac{a_1}{a_2} + F_2 \sin \beta\right]$	*Coefficients for all sites parametrization*

The 9 different techniques highlighted in Table 1 were selected to estimate the amount of solar power incident on the photovoltaic modules.

2.2. **Photovoltaic Module Models**

Once the radiation incident on the plane of the photovoltaic array is known, the optical, PV cell thermal, and electrical modeling is calculated considering the meteorological and project design data. In this manner, the generation of DC power produced by the photovoltaic modules is estimated.

2.2.1. **Optical Models**

PV module characteristics are specified for the standard test condition (STC), which consists of 1000 W/m², spectral composition of light conforming to an air mass of 1.5 (AM1.5), a 25°C module temperature and a flash emission perpendicular to the module, thus the transmittance of the glass is only evaluated based on the normal incident irradiance (0°). Modules operate in varying climatic conditions and experience different angles of incidence throughout the year, and also during the same day, especially when installed in a fixed structure. The amount of light that passes through the module glass and reaches the cell depends on the angle of incidence (AOI). The greater the AOI, the lower the transmittance of the glass, thus impacting the incident irradiance on the cell and the photogenerated current. To estimate this variation, a factor commonly referred to as the incidence angle modifier (IAM) is introduced to model the PV generation [45].

Optical models are addressed in literature as IAM losses, or angular losses [46], or reflection losses [47]. Despite their different names, optical models aim to describe the reduction in irradiance on the cells' surface compared to the normal incidence. Mathematical equations based on Fresnel's laws describe the phenomenon of radiation interaction, considering the angle of refraction (AOR) and are commonly referred to in the literature as Physical model. Some other approaches consider dimensionless empirical parameterizations, such as the Ashrae model, which depends on an adjustment parameter in the form of b_o, and the Martin-Ruiz model, which considers the angular factor (a_r). These models differ in their associated mathematical formulations, which are shown in Table 2.

Table 2: Optical models used to describe reflection losses.

CODE	OPTICAL MODEL	COMMENTS
Phys	$IAM_{Phys} = \frac{e^{-\left(\frac{KL}{cos(AOR)}\right)}\left[1-\frac{1}{2}\left(\frac{sin^2(AOR-AOI)}{sin^2(AOR+AOI)}+\frac{tan^2(AOR-AOI)}{tan^2(AOR+AOI)}\right)\right]}{e^{-(KL)}\left[1-\left(\frac{1-n}{1+n}\right)^2\right]}$	*Typical values for cSi PV modules* [48] *$K=4m^{-1}$, n=1.526, L=0.002m,* $AOR = arcsin\left(\frac{1}{n}sin(AOI)\right)$
Ashr	$IAM_{Ashr} = 1 - b_0\left(\frac{1}{cosAOI} - 1\right)$	*b_o = 0.05 for crystalline modules*
MR	$IAM_{MR} = 1 - \left[\frac{1-e^{-\frac{cos\,(AOI)}{a_r}}}{1-e^{\left(-\frac{1}{a_r}\right)}}\right]$	*a_r is the angular factor coefficient, a_r = 0.016*

2.2.2. **Thermal Models**

Thermal modeling of PV modules aims to determine the thermal behavior of the cells for different weather conditions. This step of the simulation chain is strongly dependent on solar irradiance, ambient temperature and wind speed [24]. Some models consider only the first two variables in estimating the temperature of the photovoltaic cell, disregarding convective heat exchanges, as is the case with the NOCT and Ross *et al.* [49] models. Other authors incorporate variables such as electrical efficiency (τ) and optical efficiency (τα), as well as information obtained from module datasheets, such as temperature coefficients, NOCT (Nominal Operating Cell Temperature), among other parameters [32].

More sophisticated models adopt a more intricate analysis of thermal exchanges, considering the influence of ventilation and convection on heat dissipation [50]. The influence of wind on the thermal behavior of cells is described by heat exchange coefficients (U_{PV} or U) [51, 52] or from the ratio of convective coefficients ($h_{w,NOCT}/h_w$) where h_w corresponds to the forced convection coefficient caused by wind action [53]. Because thermal losses substantially affect PV module's performance and power production [54], different thermal modeling impacts PV system simulation. Due to this, all models presented in Table 3 are evaluated in the present work.

Table 3: Thermal models used to describe PV cell temperature.

CODE	THERMAL MODELS	COMMENTS
NOCT	$T_{c,NOCT} = T_a + \frac{GTI}{G_{NOCT}}(T_{NOCT} - 20)$	
ross	$T_{c,ROSS} = T_a + k\ GTI$	*Ross thermal conductance coefficient for free standing PV systems:0.0208 Km²/W*
DB	$T_{c,DB} = \frac{T_a + (T_{NOCT} - T_{a,NOCT})\left(\frac{GTI}{G_{NOCT}}\right)\left[1 - \frac{\eta}{\tau\alpha}(1 + \gamma T_{ref})\right]}{1 - (T_{NOCT} - T_{a,NOCT})\left(\frac{GTI}{G_{NOCT}}\right)\frac{\gamma\,\eta}{(\tau\alpha)}}$	$\tau\alpha = 0.9$
King97	$T_{m,KING97} = T_a + \frac{GTI}{G_{STC}}[0.0712V_w^2 - 2.411V_w + 32.96]$	
Sandia	$T_{c,SANDIA} = T_m + \frac{GTI}{G_{STC}}\Delta T$	$T_m = GTI * exp\,(a + b * V_w) + T_a$
Mattei	$T_{c,MATTEI} = \frac{U_{PV}T_a + GTI(\tau\alpha - \eta - \beta\eta T_{ref})}{U_{PV} - \beta\,\eta\,GTI}$	$U_{PV} = 26.6 + 2.3\,V_w$
PVsyst	$T_{c,PVsyst} = T_a + \frac{\alpha GTI(1 + \eta)}{U_c + U_v V_w}$	$U_c = 29\,;\ U_v = 0$
Faiman	$T_{c,FAIMAN} = T_a + \frac{GTI}{U_c + U_v V_w}$	$U_c = 25\,;\ U_v = 6.84$
Skoplaki	$T_{c,SKOPLAKI} = \frac{T_a + \left(\frac{GTI}{G_{NOCT}}\right)\frac{h_{w,NOCT}}{h_w}(T_{NOCT} - T_{a,NOCT})\left[1 - \frac{\eta}{\tau\alpha}(1 + \beta T_{ref})\right]}{1 - \frac{\beta\,\eta}{(\tau\alpha)}\left(\frac{GTI}{G_{NOCT}}\right)\frac{h_{w,NOCT}}{h_w}(T_{NOCT} - T_{a,NOCT})}$	$h_w = 8.91 + 2.0\,V_w$ $h_{w,NOCT} = 10.91$

2.2.3. **Electrical Models**

Electrical models are divided into two large groups, one based on the equivalent electrical circuit of one or two diodes, and the second group which uses explicit equations to translate the maximum power point from STC to any operating condition.

The first group undertakes characteristic curve adjustment to determine a number of electrical parameters (4, 5, or 6) based on model-specific considerations and boundary conditions. De Soto *et al.* [48] introduced a modified diode ideality factor ('a') as one of the 5 parameters, commonly found to be less than 1, all extracted under Standard Test Conditions (STC). Equations derived from three known points alongside the temperature coefficient equation guide parameter determination, necessitating an iterative process. Once parameters are determined under STC, authors extrapolate the characteristic curve to any operational condition. In Dobos *et al.* [55], a sixth parameter, named Adjustment, is introduced in the electrical generation modeling of PV modules. This parameter adjusts the temperature coefficients of short-circuit current (α_{sc}) and open-circuit voltage (β_{oc}) provided by manufacturers. This model is also known as the *CEC* model (California Energy Commission). Wang *et al.* [56] evaluated the 6-parameter *CEC* model and two models from the second group. They observed that the *CEC* model showed better accuracy, but both the PVWatts and *CEC* models adequately described the generation of crystalline silicon modules. Roberts *et al.* [11] assessed the performance of three models based on the equivalent electrical circuit and two models that translate the maximum power point. They found that models in the latter group tended to overestimate the power output of the photovoltaic system, while those based on the equivalent electrical circuit tended to underestimate it. The De Soto *et al.* [48] model demonstrated the highest accuracy in simulating the photovoltaic system.

While most single-diode models [48, 55] require complex codes that require several iterations to converge, models that translate the maximum power point require few inputs and have fast computational processing. One of the models for translating the maximum power point is the PVWatts model used in the software developed by the National Renewable Energy Laboratory (NREL). The model estimates the output power of the PV array for different operating conditions, correcting the output power by the temperature. Huld *et al.* [57] proposed a model that estimates the output power as a function of operating temperature and irradiance, with some coefficients determined from indoor and outdoor measurements. Another simple model that performed well for monocrystalline modules [58] is the ideal diode model proposed by Saloux *et al.* [59]. This model simplifies by disregarding the effects of series resistance (Rs) and parallel resistance (Rp). In this paper, this model, which uses several explicit equations to estimate the output power of the PV array, will be referred to as the Saloux model. In addition, the De Soto, *CEC*, Huld and PVWatts models will be tested.

2.3. **Inverter Model**

Direct current (DC) from the PV generator is converted into alternate current (AC) by the inverter. The PVWatts inverter model, proposed by the National Renewable Energy Laboratory (NREL) [60], is used in the pvlib library. The California Energy Commission (*CEC*) conducted analyses that served as the basis for this model, which describes the inverter efficiency curve as a function of the loading condition, i.e., the ratio between the actual output power and the nominal power. When the simulated DC power exceeds the inverter's power limit, the model clips the output power to the nominal value. Those clipping events tend to be more evident in scenarios of high variability observed in 1-minute time series and in cases of overload, when inverter power is lower than the peak power capacity of the modules. In addition to the reduction in generation through inverter losses, there are other losses associated with the operation of photovoltaic systems.

2.4. **Photovoltaic System Losses**

Since models cannot predict certain losses, such as soiling, connections losses, Light Induced Degradation (LID), mismatch, among many others, the degradation factor or derating factor is typically derived from estimations or field measurements. They represent the negative impacts on the performance of PV systems [11], corresponding to a multiplier that reduces the output power.

The losses due to LID were considered by the commonly found value in the literature for similar polycrystalline modules, consisting of 2%. Cabling was calculated based on technical cable data, corresponding to 0.71% ohmic losses and soiling was adopted as 1.8% based on technical reports from the PV plant considered in this work. Although soiling shows seasonal variations, the average value adopted is considered representative for the plant evaluated [61], and similar annual values of 1.95% were found in the literature for semi-arid climates [62]. Mismatch losses were estimated following Lorente *et al.* [63], assuming that fewer than 25% of module strings operate at peak tolerance. This resulted in a 0.5% reduction in total power. Moreover, instead of considering the maximum tolerance value indicated in the datasheet, which would correspond to 1.5%, a more conservative value of ¼ of this tolerance was considered in the module quality reduction factor, equivalent to -0.38%. Overall, total system losses correspond to 4.6%, and the equivalent reduction factor is 0.954. As some losses may present high uncertainty depending on the considerations or methodologies adopted to determine their value, a certain degree of subjectivity is involved, hence a more detailed analysis of their impact on the physical models will be carried out in section 4.4.

3. Methodology

A total of 11,340 simulations were carried out to evaluate energy generation from observational GHI data using all the combinations of the previous models and data from the first year of operation of a 2.5 MWp PV plant. Seven distinct GHI separation models, nine transposition models, four reflective losses models derived from the incidence angle modifier (IAM) models, nine thermal module models, and five electrical models were considered in the analysis. Figure 2 shows all the models considered. The analysis works with the observational irradiance data (GHI) as input, and estimates the AC power generation (P_{AC}) of the PV plant. The intermediate variables are the GTI and the DC power (P_{DC}) injected into the inverter. The model chains used to derive P_{AC} are evaluated against their corresponding power generation data, without any type of local adjustment or adaptation, using their originally proposed coefficients. In the case of the segregation models, the coefficients proposed for the BSh climate or the closest are used. For the ground albedo, a value of 27.72 is used for all transposition models. The work was done in python language with some models available in the open library *pvlib* [64].

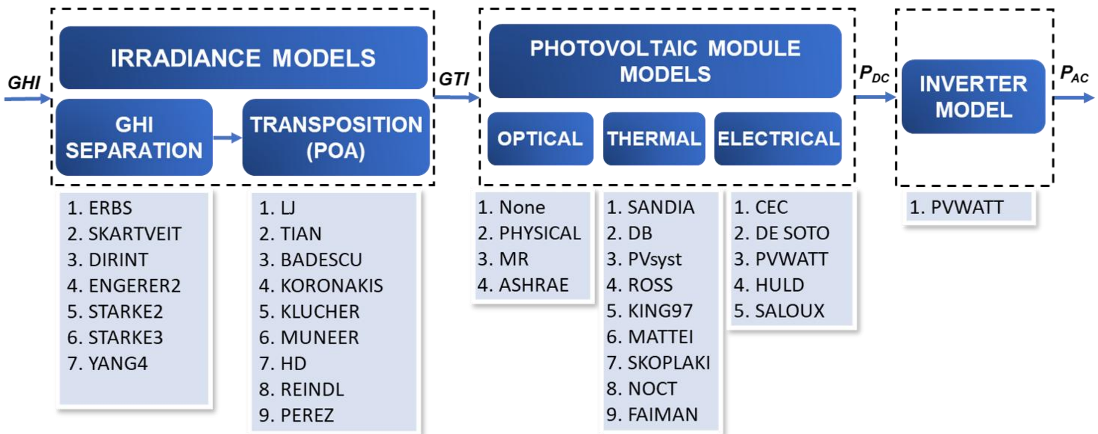


Figure 2: Process of estimating PV energy generation from horizontal global irradiance with the selected models highlighted in each step.

3.1. **Database**

The grid-connected PV system and solar station are located in Petrolina, Pernambuco, in the north-eastern part of Brazil, within a region known as Sertão. This is the region with the richest solar resource in Brazil, with average annual GHI resources of the order of 6.5 kWh/m2 day, the region is of great interest for large-scale solar energy projects [65]. The climate is hot and semi-arid, classified as BSh according to the updated Köppen-Geiger climate classification [66].

Meteorological variables and electrical data were measured every second and recorded every 1 minute. The meteorological variables were measured from the Meteorological Station of the Petrolina Solar Energy Reference Centre (CRESP), located at a latitude of 9.11 ºS, a longitude of 40.44 ºW, and an altitude of 385 m above sea level. The CRESP solarimetric station is regularly maintained by specialized technicians and is equipped with 3 EKO pyranometers, model MS-80 (class A according to ISO 9060:2018 standard), with a spectral flat range from 285 nm to 3000 nm and a response time of 0.5s. Two of them are used to measure GHI and one for DHI, the latter measured with a shadow ball attached to the tracker. The DNI pyrheliometer is also an EKO instrument, model MS-57 (class A), with the same response time and an extended spectral range from 20 nm to 4000 nm. In the scope of this work, DHI and DNI were used in the data quality control procedure. The electrical variables of the 2.5 MWp photovoltaic plant are recorded in the SCADA. Both databases operate with synchronized clocks. The grid-connected PV plant is composed of 7600 polycrystalline silicon PV modules of 330 Wp, model CS6U-330P, and 4 inverters of 600 kVA, model SIW700T600-33, 2 inverters operate at 8.6% overload and the other 2 are only 1% overloaded, more details in Appendix A. All modules are installed on a fixed structure with a 15° tilt. The database used in this research corresponds to the first year of operation, from November 2018 to October 2019, because over the following years the natural degradation of the modules could serve as a source of error.

3.2. **Data quality control**

The data quality control (QC) procedure was first applied to the irradiance magnitudes in order to remove anomalous data. Quality tests proposed by Baseline Surface Radiation Network (BSRN) in addition to physical and comparative filters discussed in Miranda *et al.* [67] were used. In the 1 year data period considered in this work, only 2.89% of all data were discarded. Samples close to sunrise and sunset are not considered for evaluation, and are removed by a solar elevation filter of 7°, as these samples tend to have more uncertainty due to the cosine error of the hemispherical instruments. For the CRESP site, these samples represent 0.21% of annual PV generation and can therefore be excluded due to their small impact on generation and the possibility of inducing relevant errors.

Apart from the solar radiation filters, the analysis only considered the moments when data was simultaneously available for all the solarimetric (GHI, DHI and DNI), meteorological (T_a and Vw) and electrical (DC and AC current and voltage) magnitudes. Days with more than 30% missing data were also discarded, as well as moments when generation is equal to 0, corresponding to inverter shutdowns. In the literature, these moments typically correspond to values between 1 and 3.4% of the year [10]. In the PV system evaluated in this work, these shutdown times correspond to 3.40%, 3.11%, 2.60% and 6.50% for the four inverters.

3.3. **Error and performance metrics**

The statistical indicators considered in this study are commonly employed in the literature [10, 11, 21] to evaluate and validate models. The selected statistics include Mean Bias Error (MBE), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and their normalized counterparts (nMBE, nMAE, and nRMSE). The first indicates how much the model underestimates or overestimates the measurement as stated at Equation 6 and 7, the second provides an average magnitude of the error based on the absolute differences (Equation 8 and 9), while the RMSE and nRMSE provide information on the error dispersions, where larger errors have greater significance due to the quadratic factor (Equation 10 and 11).

$$MBE = \frac{1}{N}\sum_{i=1}^{N}(x_{sim}^{i} - x_{meas}^{i}) \quad (6)$$

$$nMBE = \frac{1}{N\,\bar{x}_{meas}}\sum_{i=1}^{N}(x_{sim}^{i} - x_{meas}^{i}) \quad (7)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|x_{sim}^{i} - x_{meas}^{i}\right| \quad (8)$$

$$nMAE = \frac{1}{N\,\bar{x}_{meas}}\sum_{i=1}^{N}\left|x_{sim}^{i} - x_{meas}^{i}\right| \quad (9)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_{sim}^{i} - x_{meas}^{i})^2} \quad (10)$$

$$nRMSE = \frac{1}{\bar{x}_{meas}}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_{sim}^{i} - x_{meas}^{i})^2} \quad (11)$$

Where N is the amount of data, x_{sim}, x_{meas} and $\bar{x}_{meas}$ are respectively the simulated, measured and average measured values.

Additionally, a skill score known as Taylor's skill score (SS4) is incorporated, which involves the correlation (R) and the ratio of the estimated standard deviation (σ_{sim}) to the observed standard deviation (σ_{meas}) [68], as stated at Equation 12. These metrics play a crucial role in quantifying the performance and accuracy of models in various scientific domains, providing a comprehensive assessment of the agreement between modelled and observed values.

$$SS4 = \frac{(1+R)^4}{4[(\sigma_{sim}/\sigma_{meas}) + 1/(\sigma_{sim}/\sigma_{meas})]^2} \quad (12)$$

As discussed by Mermoud and Lejeune [69], for simulations of PV systems, parameters provided by manufacturers under Standard Test Conditions (STC) may only be partly representative of real operation, leading to errors in simulated production. However, these errors are linked to parameter uncertainties rather than the model's quality. In such cases, bias (MBE or nMBE) may be significant, but if the model is reliable, the error distribution represented by RMSE should remain low. Similarly, uncertainties associated with sensor measurements can impact model bias in radiation modeling. Nevertheless, models exhibiting lower nRMSE values demonstrate a greater capacity to describe the evaluated variable. In this context, nRMSE stands out as the preferred indicator for comparing the conducted simulations.

4. Results and Discussions

As stated before, a total of 11,340 simulations were carried out from 7 GHI separation models, 9 transposition models, 4 IAM scenarios, 9 PV thermal models and 5 electrical models. Each simulation aims to estimate the power generation of the 2.5 MWp plant. In this section, the error distribution and accuracy range of all the simulations is discussed, then the best sets of models are examined. Finally, the results regarding the impact that the radiation models have had on the PV modeling, as well as the effect of varying the derating factor has on the simulation of the PV plants, are discussed.

4.1. **Model accuracy range and distributions**

The variation in the accuracy of all simulations is shown in Figure 3 for the relative metrics and the skill score. The range and empirical distribution (quartiles) of each model's statistical indicators (grouped into the five categories of model's type) are obtained from all possible simulations using the given model. The first (Q1) and third (Q3) quartiles are shown in each box of the boxplot, with the median indicated by the red line. The whiskers extend from the box to the farthest data point lying within 1.5x the inter-quartile range (difference between Q1 and Q3) from the box. The furthest circles are called fliers and go beyond the end of the whiskers. For example, the ERBS model (separation model) has nRMSE values ranging from over 12.5% to almost 16.7%, which means that all the model chains that separate the global horizontal by the ERBS model have errors between those values and Q1 and Q3 concentrated below 14%.

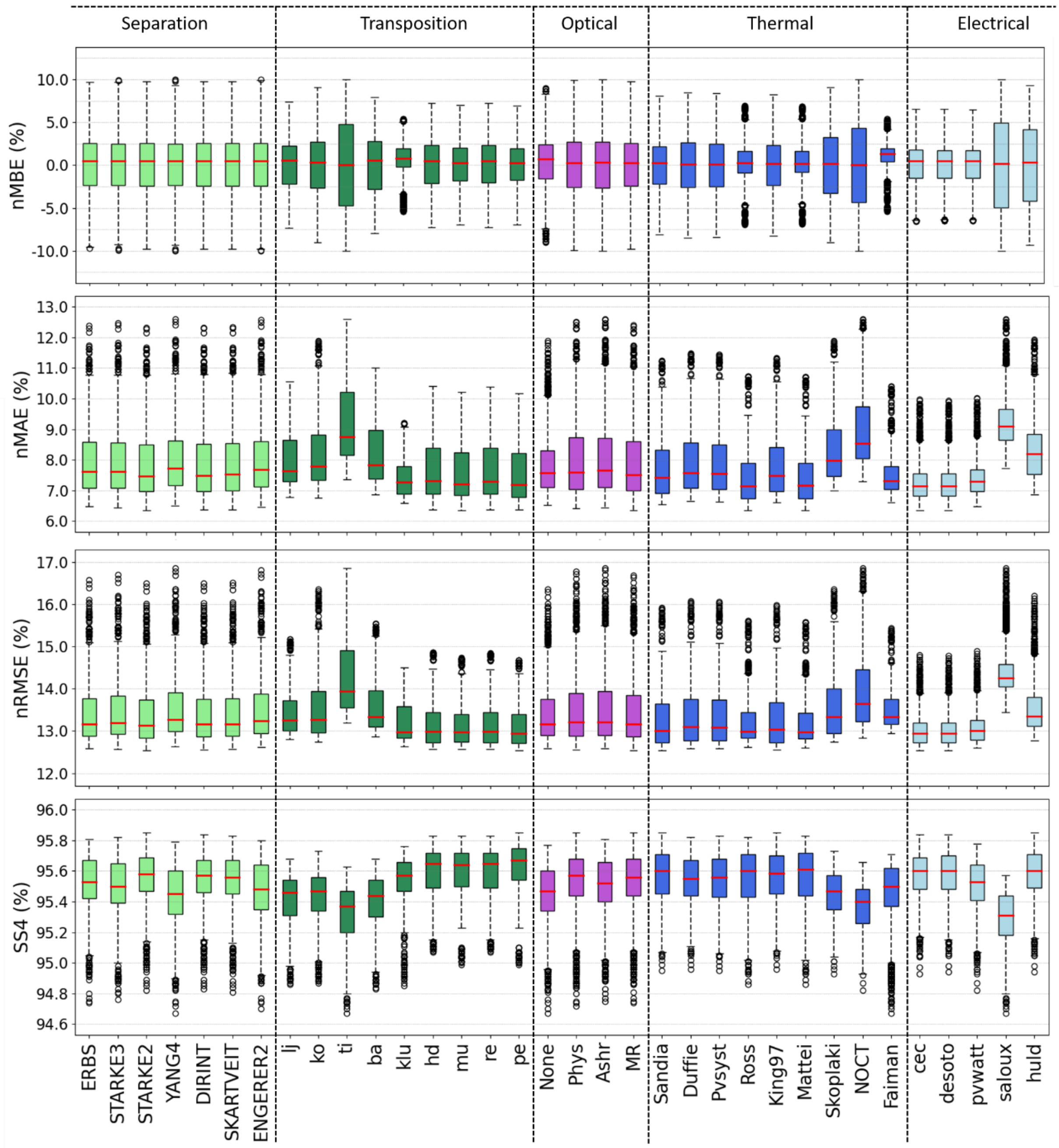


Figure 3: Accuracy distribution of each model for all possible model chains as a function of the statistical indicators nMBE, nMAE, nRMSE and SS4. Decomposition models are shown in light green, transposition models in dark green, IAM in purple, thermal models in dark blue and electrical models in light blue.

Based on Figure 3, it can be seen that the variation in nMBE of all the models is concentrated between -10% and 10%, with medians close to 0%. Klucher's anisotropic model, and Ross and Faiman's thermal models exhibited consistent overestimates, with positive values in the median and smaller variations between Q1 and Q3. The same behavior is observed for the Ross model in Deville *et al.* [9]. The most critical steps in 1-min modeling are given by the transposition models, thermal modeling and electrical models. Among all models, the Tian (ti) transposition model, the NOCT thermal model and the Saloux electrical model show the greatest variation in bias, corresponding to -10% to 10%, exhibiting greater dispersion in the distribution and consequently less consistency in modeling PV generation. The poor performance of these models is also confirmed for nMAE, nRMSE and SS4, with the highest errors associated with Q1 and Q3.

The minimum and maximum values found for the nMAE correspond to 6.3% and 12.6%, respectively, showing a high variation in the error associated with the selection of the models. In terms of nRMSE, the variation corresponds to 4.4%, with the maximum difference corresponding to 16.9% obtained by the YANG4 with tian model plus Ashrae, NOCT and Saloux models. The minimum value of nRMSE corresponds to 12.54% and was obtained by the sets of models that include the *CEC* and desoto electrical models, the King97 and Sandia thermal models, the Reindl (Re) and Perez (pe) transposition models, the Martin-Ruiz optical model and the STARKE2 GHI separation technique.

Furthermore, Figure 3 also shows that some models are generally more consistent depending on the statistical indicator, for example, the *CEC*, desoto and pvwatt electrical models were highly consistent with little variation between the Q1 and Q3 for nMBE, nMAE and nRMSE. In other words, regardless of the models used, these 3 models tend to be more consistent with more accurate simulations than the saloux and huld electrical models. Similarly, the ross and faiman thermal models showed smaller differences between Q1 and Q3 in terms of nMBE, nMAE and nRMSE. However, this result did not imply more accurate simulations, just less variation depending on the models in the simulation chain. It can be seen, for example, that other thermal models have a median and P25 below the nRMSE values of the Faiman model, indicating greater accuracy.

In terms of radiation modeling, the separation models do not vary the PV output so much from one to another, with very similar variations between P25 and P75. Among all the GHI separation models, STARKE2 with parameterization for Brazil, designed by [16], had the best quartile values for nMAE, nRMSE and SS4. Regarding the transposition models, the first 4 models (Liu and Jordan, Koronakis, Tian and Badescu) correspond to isotropic models and showed the worst results in terms of nMAE, nRMSE and SS4. These results are expected due to the greater simplicity of modeling diffuse radiation in the inclined plane [70], which impacts on the estimate of the solar resource available in the fixed plane of the PV array.

One of the radiation modeling steps that has the greatest impact on the simulation chain of photovoltaic systems is the irradiance transposition to the inclined plane [8]. For this reason, the analysis of the separation, optical, thermal and electrical models is examined in Figure 4 against the transposition models in order to address with higher detail the minimum results for the nMAE and nRMSE, and the maximum values of SS4 achieved for each of the combinations. The variation in the accuracy of the groups of models indicates that the best results are obtained by the anisotropic transposition models Hay and Davies (*hd*), Muneer (*mu*), Perez (*pe*) and Reindl (*re*), obtaining brighter colors, thus demonstrating a greater ability to estimate PV generation.

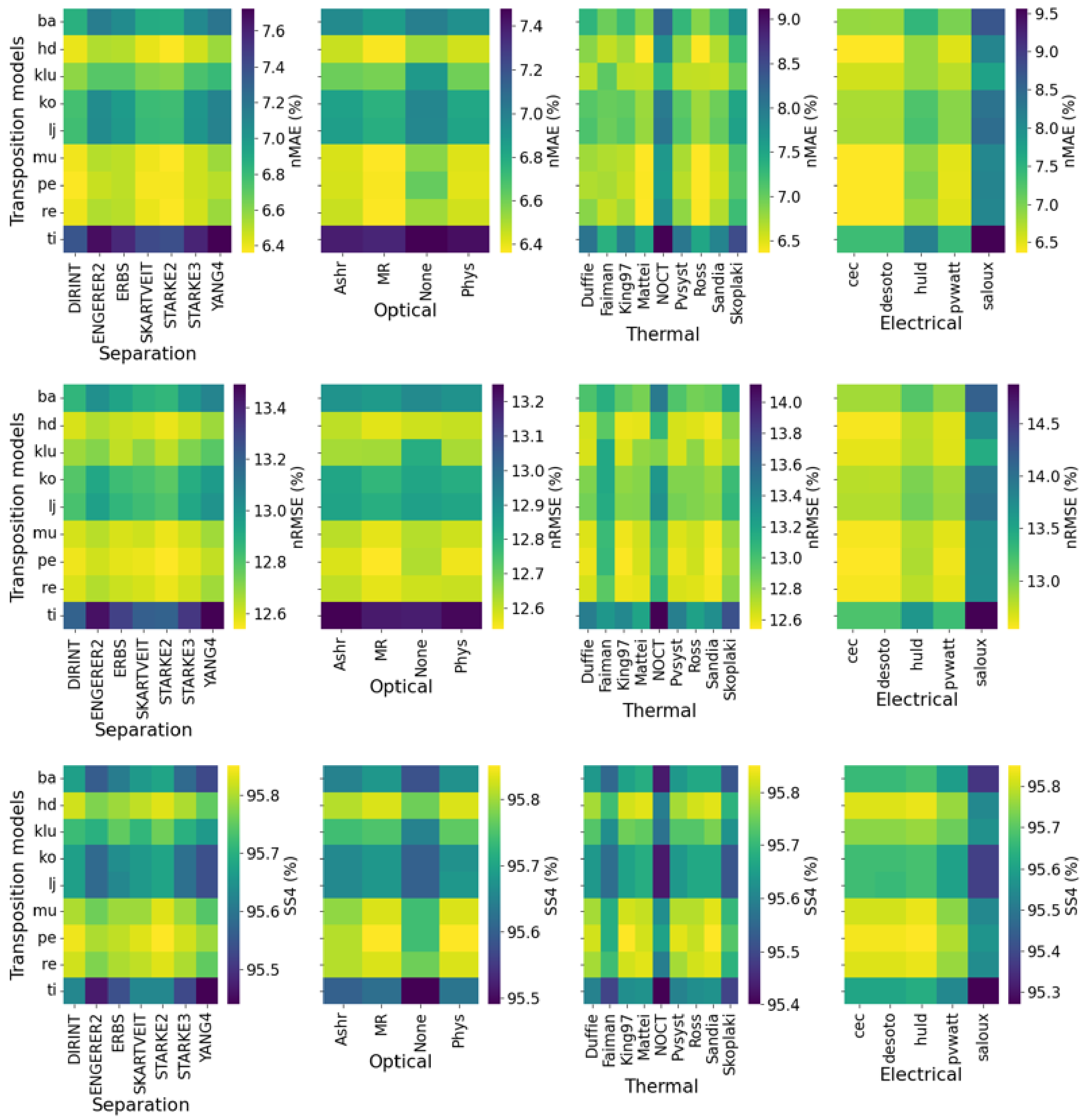


Figure 4: Distribution of the best results for nMAE, nRMSE and SS4 evaluated according to the combination of transposition models with the separation, optical, thermal and electrical models.

Among the anisotropic models, the Klucher model is the only one that underperforms, with lower statistics than the other anisotropic models. This is due to the fact that the Klucher modulation function does not exhibit a high ability to estimate the circumsolar irradiance and the brightness of the horizon in low inclinations at low latitude locations, tending to overestimate the incident irradiance on the inclined plane [71].

The results shown in Figure 4 indicate that the most significant variations between the best sets of models are found between the transposition model with the thermal and electrical modeling stage. The variations within the maximum and minimum of the thermal models and the transposition models range from 6.5% to over 9% in terms of nMAE, and from 12.6% to over 14% on nRMSE, indicating a variation of over 3.5% in nMAE and over 1.4% in nRMSE.

In general terms, the models that achieved good results when associated with the *hd, pe, re* and *mu* anisotropic models were the GHI *STARKE2* separation models, the *Martin-Ruiz* optical reflectance model, the thermal behavior of the *Skoplaki, King97, Mattei, Ross* and *Sandia* models, and the *CEC* and *desoto* electrical models. The best results were achieved by using the Perez transposition model.

4.2. **Best set of models**

Figure 5 illustrates the distribution of model performance across the top 1% (best-performing) and bottom 1% (worst-performing) simulation results based on four statistical metrics (nMBE, nMAE, nRMSE, and SS4). For each statistical metric, the 113 best and worst combinations have been identified from the 11,340 simulations considered. A combination is considered “best-performing” when it ranks within the top 1% of results for a given metric, representing the lowest error for nMAE, nMBE and nRMSE, or highest agreement, in the case of SS4. The heatmaps display the frequency of each model appearance within these best and worst combinations. Models highlighted in red denote poor performance as they feature among the worst combinations. In contrast, models that consistently appear in the top 1% are highlighted in green, reflecting a strong ability to accurately model PV generation.

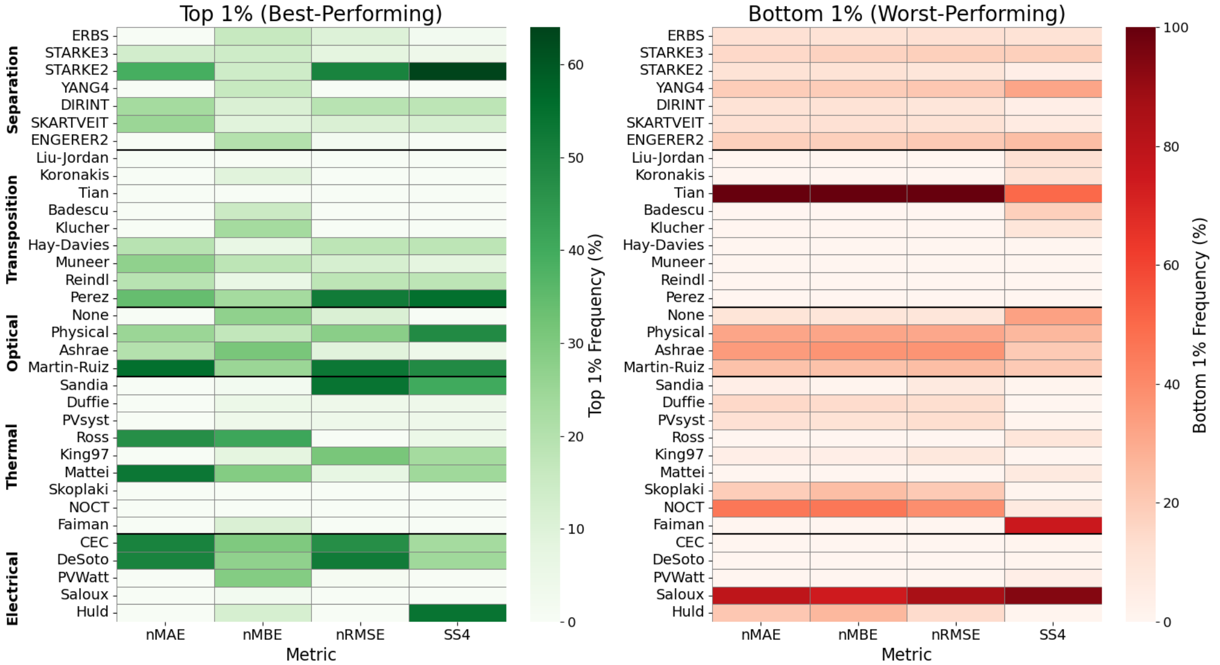


Figure 5: Frequency of individual models within the top and bottom 1% of 11,340 model combinations, evaluated by nMBE, nMAE, nRMSE, and SS4 metrics.

Figure 5 shows that the STARKE2 model is the GHI separation model that is most present among the best simulations, corresponding to 50% of the combinations with best nRMSE results, 64% for SS4 and 39% for nMAE, suggesting that its application with transposition, optical, thermal and electrical models tend to produce positive results in terms of nMAE, nRMSE and SS4. Subsequently, the DIRINT and SKARTVEIT models show high proportions in the best simulations. The three models stand out as the ones least present among the worst sets of models analyzed.

In terms of the transposition of radiation onto the inclined plane, as seen in Figure 4 and now described in terms of the percentage of the best and worst combinations, isotropic models tend to be absent from the best PV generation simulations in terms of the nMAE, nRMSE and SS4 metrics, demonstrating a low ability to estimate diffuse radiation on the inclined plane. In addition, it can be seen that the anisotropic models, with the exception of the Klucher model, tend to be among the best simulations conducted, this fact may be linked to the way the irradiance from the horizon brightness and circumsolar is estimated by Klucher model, the inferior performance of this anisotropic models is also recorded in Loutzenhiser *et al.* [72], Yang *et al.* [71] and Yang *et al.* [73]. Furthermore, the Tian model showed poor performance for the low slope evaluated (15°N), being the most frequent model present in the worst simulations. Arias-Rosales and LeDuc [74] compared the Tian, Badescu and Liu-Jordan sky view factors and observed that Tian tends to perform less than Liu-Jordan.

With regard to optical models, the model proposed by Martin and Ruiz [47] showed the best results, being present in more than half of the cases for the nMAE, nRMSE and SS4 statistics (55%, 53% and 48%, respectively). For thermal models, the Sandia model performed better in terms of nRMSE and SS4, being included in more than 2/5 of the combinations for those indicators, subsequently Mattei, Ross, and King97 also presented consistent results for the best simulations.

For electrical models, the *CEC* and De Soto single-diode models showed the best overall results, being green in all the indicators and scenarios (Best 1% and Worst 1%). Similar results of the high performance of these models were found in the comparative evaluation of several groups of models in the simulation of photovoltaic systems in Roberts *et al.* [11]. Also, in the comparative analysis of electrical models by Wang *et al.* [56], the authors found that for polycrystalline silicon modules the *CEC* model when compared to the PVWatt model in the temporal resolution of 1 min presents a greater capacity to describe the electrical output characteristics of the modules with better nMBE and nRMSE results. In general terms, the *CEC* and De Soto models are better at estimating the output power effectively, as they calculate the parameters of the equivalent circuit for each operating condition (GTI and T_{mod}) and consequently calculate the point of maximum power. Small variations are observed between the performance of the *CEC* model and the DeSoto model, because the difference between both models lies in the addition of the sixth parameter in the *CEC* model, the *Adjust* that corrects the temperature coefficients of the short-circuit current and the open-circuit voltage. This parameter can be obtained from the System Advisor Model library [75].

4.3. **Comparative evaluation of specific cases**

A wide range of possibilities for simulating PV system generation can be achieved from all the physical models selected. In Roberts *et al.* [11], one of the sets of models that showed the best estimate of PV generation was the DIRINT separation model with the Koronakis (ko) transposition model, no optical model, and the DeSoto electrical model coupled with the skoplaki thermal model. This combination achieved the best nRMSE results of around 15%. In Mayer and Gróf [10], the sets of models that showed excellent mean absolute error in simulating PV generation were composed of the Starke separation model, Mattei temperature, PVWatt electrical (referred by the authors as Evans, due to being one of the pioneers to approach output power following this method [76], Liu-Jordan or Perez transposition model and Martin-Ruiz or physical IAM model.

Table 4 shows the statistical metrics and the generation of each case. The best and worst results achieved in this work in terms of nMAE and nRMSE, as well as cases from selected papers, and also scenarios that consider the best models obtained individually according to the state of the art (Best individually 1 and 2) and a case of simplified models that are easy to apply are evaluated and presented in Table 4.

Table 4: Statistical metrics and total photovoltaic generation for specific modeling cases.

CASES	SET OF MODELS					nMAE (%)	nMBE (%)	nRMSE (%)	SS4 (%)	E_{PV} (GWh)
	Sep.	**Transp.**	**Optical**	**Thermal**	**Elect.**					
Measured	-	-	-	-	-	-	-	-	-	*2,136*
Best nMAE	*Starke2*	*Muneer*	*MR*	*Mattei*	*DeSoto*	*6,4*	*0,1*	*12,7*	*95,80*	*2,139*
Best nRMSE	*Starke2*	*Perez*	*MR*	*Sandia*	*DeSoto*	*6,5*	*-1,2*	*12,5*	*95,80*	*2,112*
Worst nMAE and nRMSE	*Yang4*	*Tian*	*Ashrae*	*NOCT*	*saloux*	*12,6*	*-10,0*	*16,9*	*94,99*	*1,923*
Roberts et al. (2017)	*DIRINT*	*Ko*	*None*	*Skoplaki*	*DeSoto*	*7,5*	*-1,5*	*13,1*	*95,37*	*2,104*
Mayer and Gróf (2021) 1	*Starke2*	*Perez*	*MR*	*Mattei*	*PVwatt*	*6,5*	*0,2*	*12,7*	*95,77*	*2,140*
Mayer and Gróf (2021) 2	*Starke2*	*LJ*	*Phys*	*Mattei*	*PVwatt*	*7,0*	*-0,5*	*12,9*	*95,60*	*2,126*
Best individually 1	*Yang4*	*Perez*	*Phys*	*Mattei*	*CEC*	*6,5*	*0,1*	*12,7*	*95,75*	*2,138*
Best individually 2	*Starke3*	*Perez*	*Phys*	*Mattei*	*CEC*	*6,5*	*0,1*	*12,7*	*95,78*	*2,139*
Simplified 1	*ERBS*	*LJ*	*None*	*NOCT*	*PVwatt*	*8,3*	*-2,8*	*13,5*	*95,12*	*2,076*
Simplified 2	*ERBS*	*LJ*	*None*	*Ross*	*Huld*	*9,9*	*-6,2*	*14,5*	*95,19*	*2,003*

Table 4 shows that the first two combinations with the best nMAE and nRMSE results, respectively, provide results well in line with the measured generation and PR, especially the combination with the best nMAE, which shows a 3MWh difference (corresponding to a nMBE overestimate of 0.1%) and an overall PR very similar to the measured one. It can also be seen that the best sets of models found by Roberts *et al.* [11] and Mayer and Gróf [10] achieve high performance results, also demonstrating good accuracy in estimating PV generation.

Furthermore, Table 4 shows that if the project designer decides to adopt models that stand out individually in the literature, i.e., assessing exclusively the separation models [21], or only the transposition models and the other stages of photovoltaic modeling, the errors found will be close to the best scenarios assessed here. The individually best scenarios correspond to models also adopted in Mayer and Gróf [10], with the exception of the separation models, corresponding to the Yang4 [23] and Starke3 [27] separation models for version 1 and version 2, the Perez model [73], the physical optical model, the Mattei thermal model (according to [77]) and the *CEC* electrical model due to its high performance, which is similar to DeSoto but with a slight improvement in nRMSE, as observed in Deville *et al.* [9].

Another important factor to mention is the additional errors that can be found if the simulation is conducted with simplified models that do not fit properly in the conditions evaluated, for example, simplified version 2 adopts the Badescu model, which was designed and evaluated in high-altitude locations, but for semi-arid climatic conditions at a low-latitude site it presents a low capacity to describe the diffuse radiation on the inclined plane, tending to strongly underestimate the irradiance that falls on the PV panels. When associated with the NOCT thermal model, which overestimates the thermal losses of the modules [78], the generation estimated based on this set of models will tend to strongly underestimate the measurement, as seen in Table 4 and also shown in Figure 6, underestimating in 6,2%.

Figure 6 shows the simulated versus measured graph for all the selected scenarios.

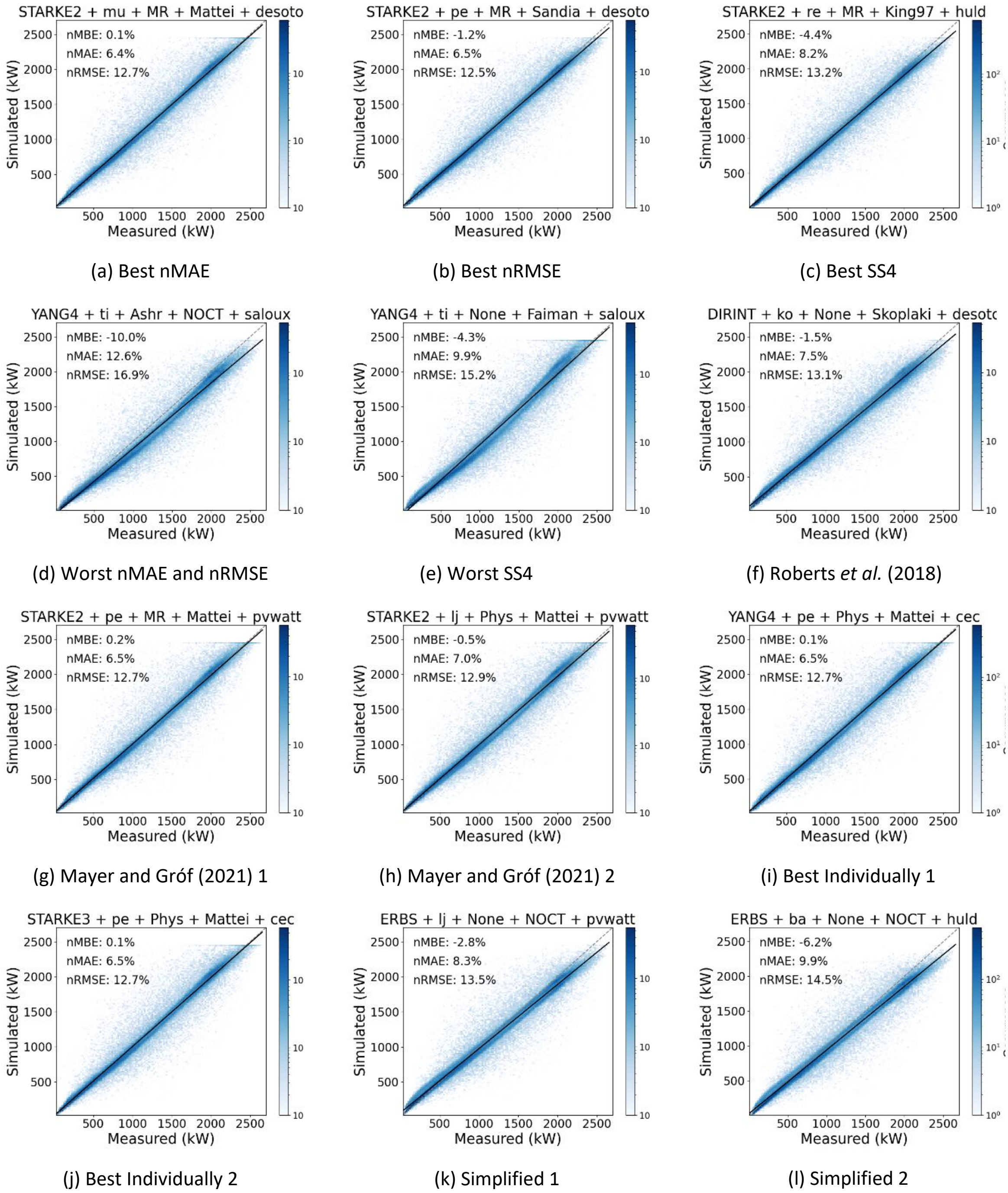


Figure 6: Simulated versus measured data for different scenarios.

Figure 6 shows that the adoption of accurate models according to the literature (Figures 6.i and Figure 6.j) results in accurate simulations with a high degree of reliability, featuring small differences in the statistics when compared to the simulations that best fit (Figure 6.a, 6.b and 6.c) the conditions evaluated. When comparing scenarios 1 and 2 of Mayer and Gróf [10] (Figures 6.g and 6.h) the use of an isotropic model tends to underestimate generation, adding error to the results.

It can also be seen that depending on the models selected, variations in nMAE of up to 6.2% (6.4% for the best nMAE and 12.6% for the worst nMAE) can be obtained, which means that the best model chains have 49% less error compared to the worst-performing ones. About nRMSE, the absolute variation found was 4.4% (12.5% for the best nRMSE and 16.9% for the worst nRMSE), corresponding to a relative difference of 26%.

Among the simplified cases (Figure 6.k and 6.l), it can be seen that adopting models that are easy to apply can lead to greater errors in estimating generation, especially when adopting isotropic models that tend to underestimate irradiance on the inclined plane (ba) associated with thermal models (NOCT) that overestimate the temperature of the modules, which increases losses and favors underestimating generation.

4.4. **Impact of the derating factor on model accuracy distribution**

The effect of the derating factor on the physical models was assessed by varying the loss coefficient from 0.914 to 0.994 and observing the average value of the statistic for all simulations using a selected model, i.e. if the nMAE of a model for the derating factor 0.994 corresponds to 8%, it indicates that the average nMAE of all simulations utilizing this model, when assembled, is 8%. Figure 7 shows all the average values for nMAE, nMBE, nRMSE and SS4. Yellowish colors indicate better average values, and greenish colors indicate worse results.

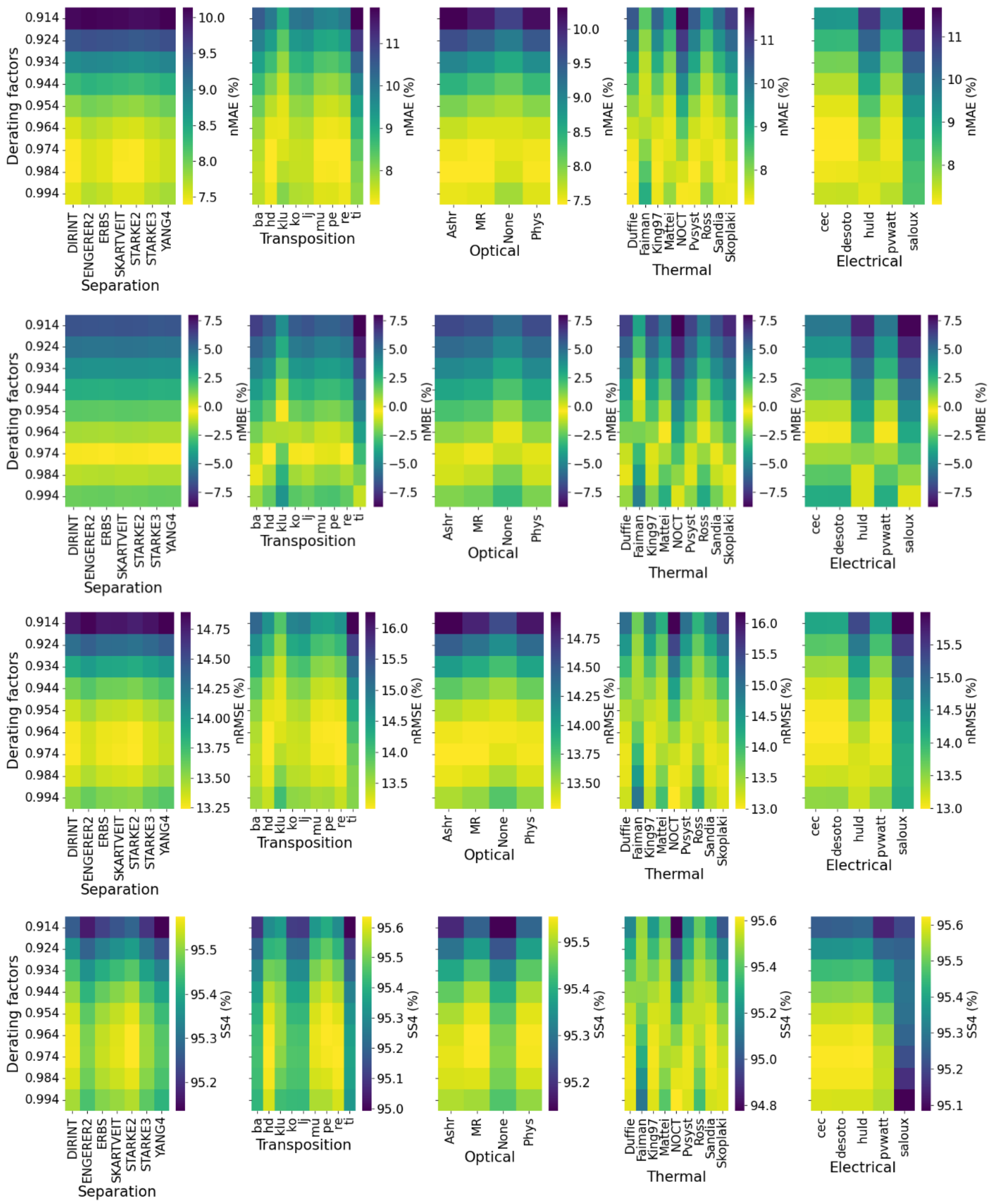


Figure 7: Impact of derating factors on average simulation values for each model.

Figure 7 shows that regardless of the loss values adopted, the *CEC*, desoto and pvwatt models tend to present more accurate results, with better accuracy for values close to 0.964 (approximately 3.6% losses). Furthermore, among the optical models, the best results in the scenario without IAM losses are offset by approximately 1%, demonstrating that the optical models produce losses of this magnitude.

Additionally, Figure 7 depicts that the top-performing models discussed in section 4.2 maintain a consistent performance profile regardless of the derating factor value, except for the thermal models, which display differing performances depending on the adopted derating factors.

5. Conclusions and limitations

This study analyzes different models and models chains used in the simulation of grid-connected PV systems, evaluating the PV generation from 11,340 model combinations consisting of seven GHI direct-diffuse separation models, nine transposition models to tilted plane, four optical models (IAM), nine thermal models and five electrical models using high-resolution (1 minute) data from a 2.5 MWp photovoltaic plant installed in the Brazilian northeast semi-arid region. The present study is the first detailed work for PV systems simulation in this relevant climate zone of Brazil (BSh) that evaluates all the groups of models required for PV power generation estimation under all-sky conditions.

The most critical steps observed correspond to the groups of transposition models, thermal models and electrical models. Among the best simulations, the anisotropic transposition models of Hay and Davies (HD), Reindl (Re), Muneer (Mu) and Perez (Pe) were the most prevalent models among the best combinations evaluated, with a special emphasis on Perez's diffuse transposition model. Among the thermal models, Sandia, King97, Mattei and Ross models were the most present in the 1% of the best combinations evaluated. Regarding the electrical models, the one-diode, *CEC* and De Soto models were the best models applied, with slight differences between the two. The PVwatt model also showed satisfactory results when compared to the other maximum power point translation models.

Significant increases in simulation error at high temporal resolution can be observed when erroneous or oversimplified physical models are used for PV generation modeling. Relative differences in nMAE of 49% and nRMSE of 26% were evidenced, indicating the importance of selecting appropriate models. In situations where data is unavailable to validate and identify the best models, it is advised to adopt the best models for the desired climate and system technology based on the literature, which is demonstrated in this work to be a good choice, almost achieving the optimal performance of the best combinations. Furthermore, apart from the thermal models, overestimating or underestimating the derating factor does not affect the performance profile of the models, showing that the best models tend to describe better the physical variables evaluated, demonstrating greater ability in modeling PV generation.

It is worth noting that the choice of separation models showed a low impact on the end-to-end modeling of PV generation in the present study. One possible hypothesis explaining this result is based on the characteristics of the evaluated PV plant that corresponds to a fixed system with low inclination. Since the system is fixed, the imbalance of these variables is reflected onto the inclined plane, resulting in minimal variation in POA irradiance. However, if the PV facility were a tracking solar plant, the separation models would tend to have a more significant impact, as POA irradiance is strongly influenced by DNI, necessitating accurate simulation of this variable. In this sense, in future studies, it’s essential to expand the analysis to solar PV tracker systems under hot semi-arid climate (BSh) conditions, and compare with the fixed-plane situation.

Acknowledgement

The authors acknowledge the financial support from CHESF/ELETROBRAS-ANEEL R&D Programs (CHESF - Petrolina PD-00048-0117 and CHESF - Casanova PD-00048-0217), PRH-ANP/FINEP (Nº 48610.201019/2019-38), and CAPES.

Appendix A

The PV plant is connected to Celpe's distribution grid with 2.5 MWp of DC capacity. The characteristics of the PV modules, inverters, PCS and other characteristics are shown in Table A.1.

Table A.1: Characteristics of the PV module, array and system.

GENERAL INFORMATION					
Latitude	9.11 °S	*PV Module*	CS6U-330P (CSI)	*Inverter Model*	SIW700-T600-33 (WEG)
Longitude	40.44 °W	*Rated power*	330 Wp	*Rated Power*	600 kW
Altitude	385 m	*Number of modules*	7600	*Number of inverters*	4
Structure / Tilted	15°N Fixed	*Efficiency at STC*	16.97%	*Modules per inv.*	1976 and 1824
DC Capacity	2508 kWp	*NOCT*	45°C	*Inverter overload*	8.6% and 0.3%
AC Capacity	2400 kW	*Temperature Coeff. ($P_{máx}$)*	-0.41%/°C	*PCS*	2
Area	45 ha	*PV technology*	poli-Si	*Inverters per PCS*	2

References

[1] IRENA, "Renewable energy statistics 2024," International Renewable Energy Agency, Abu Dhabi, 2024a.

[2] IRENA, "Tripling renewable power by 2030: The role of the G7 in turning targets into action," International Renewable Energy Agency, Abu Dhabi, 2024b.

[3] A. Bouich, I. G. Pradas, M. A. Khan and Y. H. Khattak, "Opportunities, challenges, and future prospects of the solar cell market. , 15(21), 15445," *Sustainability,* vol. 15, no. 21, p. 15445, 2023.

[4] U. Jahn, M. Herz, D. Moser, G. Belluardo and M. Richter, "Managing technical risks in PV investments: How to quantify the impact of risk mitigation measures for different PV project phases?," *Progress in Photovoltaics: Research and Applications,* vol. 26, no. 8, pp. 597-607, 2018.

[5] S. Kurtz, J. Newmiller, A. Kimber, R. Flottemesch, E. Riley, T. Dierauf, J. McKee and P. Krishnani, "Analysis of photovoltaic system energy performance evaluation method," National Renewable Energy Lab. (NREL). (No. NREL/TP-5200-60628), Golden, CO (United States), 2013.

[6] R. Ruther, L. R. do Nascimento and R. A. Campos, "Performance assessment issues in utility-scale photovoltaics in warm and sunny climates.," *Renewable Energy and Environmental Sustainability,* vol. 2, no. 35, 2017.

[7] J. Balfour, R. Hill, A. Walker, G. Robinson, T. Gunda and J. Desai, "Masking of photovoltaic system performance problems by inverter clipping and other design and operational practices.," *Renewable and Sustainable Energy Reviews,* vol. 145, no. 111067, 2021.

[8] M. Hofmann and G. Seckmeyer, "Influence of various irradiance models and their combination on simulation results of photovoltaic systems.," *Energies,* vol. 10, no. 10, p. 1495, 2017.

[9] L. Deville, M. Theristis, B. H. King, T. L. Chambers and J. S. Stein, "Open-source photovoltaic model pipeline validation against well-characterized system data.," *Progress in Photovoltaics: Research and Applications,* vol. 32, no. 5, pp. 291-303, 2024.

[10] M. J. Mayer and G. Gróf, "Extensive comparison of physical models for photovoltaic power forecasting.," *Applied Energy,* vol. 283, no. 116239., 2021.

[11] J. J. Roberts, A. A. M. Zevallos and A. M. Cassula, "Assessment of photovoltaic performance models for system simulation.," *Renewable and Sustainable Energy Review,* vol. 72, pp. 1104-1123, 2017.

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

[13] A. Younis and Y. Alhorr, "Modeling of dust soiling effects on solar photovoltaic performance: A review.," *Solar Energy,* vol. 220, p. 1074–1088, 2021.

[14] M. Hofmann, S. Riechelmann, C. Crisosto, R. Mubarak and G. Seckmeyer, "Improved Synthesis of Global Irradiance with One-Minute Resolution for PV System Simulations. 2014(1), .," *International Journal of Photoenergy,,* vol. 2014, no. 1, p. 808509, 2014.

[15] B. Burger and R. Rüther, "Inverter sizing of grid-connected photovoltaic systems in the light of local solar resource distribution characteristics and temperature. , 80(1), 32-45.," *Solar Energy,* vol. 80, no. 1, pp. 32-45, 2006.

[16] A. R. Starke, L. F. Lemos, J. Boland, J. M. Cardemil and S. Colle, "Resolution of the cloud enhancement problem for one-minute diffuse radiation prediction.," *Renewable Energy,* vol. 125, pp. 472-484, 2018.

[17] C. Reise, B. Müller, D. Moser, G. Belluardo, P. Ingenhoven, A. Driesse, G. Razongles and M. Richter, "Uncertainties in PV system yield predictions and assessments.," IEA PVPS Task 13, Report IEA-PVPS T13-12:2018., 2018.

[18] N. A. Engerer, "Minute resolution estimates of the diffuse fraction of global irradiance for southeastern Australia.," *Solar Energy,* vol. 116, pp. 215-237, 2015.

[19] J. M. Bright and N. A. Engerer, "Engerer2: Global re-parameterisation, update, and validation of an irradiance separation model at different temporal resolutions.," *Journal of Renewable and Sustainable Energy,* vol. 11, no. 3, 2019.

[20] M. Manni, A. Nocente, K. Skeie, M. Bellmann and G. Lobaccaro, "Validation of decomposition models for solar irradiance at high latitudes: A preliminary study.," *Journal of Physics: Conference Series,* vol. 2654, p. 012149, 2023.

[21] D. Yang, "Estimating 1-min beam and diffuse irradiance from the global irradiance: A review and an extensive worldwide comparison of latest separation models at 126 stations.," *Renewable and Sustainable Energy Reviews,* vol. 159, p. 112195, 2022.

[22] D. Yang and J. Boland, "Satellite-augmented diffuse solar radiation separation models.," *Journal of Renewable and Sustainable Energy,* vol. 11, no. 2, 2019.

[23] D. Yang, "Temporal-resolution cascade model for separation of 1-min beam and diffuse irradiance.," *Journal of Renewable and Sustainable Energy,* vol. 13, no. 5, 2021.

[24] Y. Sun, F. Wang, B. Wang, Q. Chen, N. A. Engerer and Z. Mi, "Correlation Feature Selection and Mutual Information Theory Based Quantitative Research on Meteorological Impact Factors of Module Temperature for Solar Photovoltaic Systems.," *Energies,* vol. 10, no. 1, p. 7, 2016.

[25] B. Ridley, J. Boland and P. Lauret, "Modeling of diffuse solar fraction with multiple predictors.," *Renewable Energy,* vol. 35, no. 2, pp. 478-483, 2010.

[26] P. Ineichen, "A broadband simplified version of the Solis clear sky model. , 82(8), .," *Solar Energy,* vol. 82, no. 8, pp. 758-762, 2008.

[27] A. R. Starke, L. F. Lemos, C. M. Barni, R. D. Machado, J. M. Cardemil, J. Boland and S. Colle, "Assessing one-minute diffuse fraction models based on worldwide climate features.," *Renewable Energy,* vol. 177, pp. 700-714, 2021.

[28] 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.

[29] A. Skartveit, J. A. Olseth and M. E. Tuft, "An hourly diffuse fraction model with correction for variability and surface albedo.," *Solar Energy,* vol. 63, no. 3, pp. 173-183, 1998.

[30] 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.

[31] J. V. F. de Medeiros, L. E. Barboza, D. R. Miranda, O. C. Vilela, E. T. Gomes and G. Salazar, "Modelos de separação de GHI: Validação para dados de 1-min e análise do impacto na geração de sistemas fotovoltaicos em baixa latitude. Avances en En," *Avances en Energías Renovables y Medio Ambiente-AVERMA,* vol. 26, pp. 291-302, 2022.

[32] J. Duffie and W. Beckman, Solar Engineering of Thermal Processes 2nd edition, Wiley, New York, NY, 1991., 1991.

[33] V. Badescu, "3D isotropic approximation for solar diffuse irradiance on tilted surfaces. , 26(2), 221-233.," *Renewable energy,* vol. 26, no. 2, pp. 221-233, 2002.

[34] J. E. Hay and J. A. Davies, "Calculation of the solar radiation incident on an inclined surface.," in *Proceedings of the 1980 Annual Meeting of the American Solar Energy Society*, 59-72, 1980.

[35] B. Y. H. Liu and R. C. Jordan, "Daily insolation on surfaces tilted toward the equator.," *ASHRAE Journal,* vol. 3, no. 10, pp. 53-59, 1961.

[36] P. S. Koronakis, "On the choice of the angle of tilt for south facing solar collectors in the Athens basin area. , 36(3), 217-225.," *Solar Energy,* vol. 36, no. 3, pp. 217-225, 1986.

[37] Y. Q. Tian, R. J. Davies-Colley, P. Gong and B. W. Thorrold, "Estimating solar radiation on slopes of arbitrary aspect. , 109(1), 67-74.," *Agricultural and Forest Meteorology,* vol. 109, no. 1, pp. 67-74, 2001.

[38] R. C. Temps and K. L. Coulson, "Solar radiation incident upon slopes of different orientations.," *Solar Energy,* vol. 19, no. 2, pp. 179-184, 1977.

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

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

[41] Y. F. Nassar, A. A. Hafez and S. Y. Alsadi, "Multi-factorial comparison for 24 distinct transposition models for inclined surface solar irradiance computation in the state of Palestine: A case study.," *Frontiers in Energy Research,* vol. 7, 2019.

[42] R. Wattan and S. Janjai, "An investigation of the performance of 14 models for estimating hourly diffuse irradiation on inclined surfaces at tropical sites.," *Renewable Energy,* vol. 93, pp. 667-674, 2016.

[43] T. Muneer, "Solar radiation model for Europe.," *Building Services Engineering Research and Technology,,* vol. 11, no. 4, pp. 153-163, 1990.

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

[45] M. H. Saw, H. L. Soh, A. Ng, K. E. Birgersson, S. E. R. Tay and M. Pravettoni, "A spot-area method to evaluate the incidence angle modifier of photovoltaic devices—Part 1: Cells," *IEEE Journal of Photovoltaics,* vol. 13, no. 2, pp. 267-274, 2023.

[46] A. Gonçalves, D. Rativa and L. A. Gómez-Malagón, "Model-based assessment of the incident angle modifier on the annual angular losses and gain of PV modules in tracking systems," *IEEE Journal of Photovoltaics,* vol. 14, no. 1, pp. 185-193, 2023.

[47] N. Martin and J. M. Ruiz, "Calculation of the PV modules angular losses under field conditions by means of an analytical model.," *Solar Energy Materials and Solar Cells,* vol. 70, no. 1, pp. 25-38, 2001.

[48] W. De Soto, S. A. Klein and W. A. Beckman, "Improvement and validation of a model for photovoltaic array performance.," *Solar energy,* vol. 80, no. 1, pp. 78-88, 2006.

[49] R. G. Ross, "Interface design considerations for terrestrial solar cell modules," in *12th Photovoltaic Specialists Conference*, 1976.

[50] D. L. King, W. E. Boyson and J. A. Kratochvil, "Photovoltaic array performance model.," in *Sandia National Laboratories (SNL)*, Albuquerque, NM, and Livermore, CA (United States), 2004.

[51] M. Mattei, G. Notton, C. Cristofari, M. Muselli and P. Poggi, "Calculation of the polycrystalline PV module temperature using a simple method of energy balance.," *Renewable Energy,* vol. 31, no. 4, pp. 553-567, 2006.

[52] D. Faiman, "Assessing the outdoor operating temperature of photovoltaic modules.," *Progress in Photovoltaics: Research and Applications,* vol. 16, no. 4, pp. 307-315, 2008.

[53] E. Skoplaki, A. G. Boudouvis and J. A. Palyvos, "A simple correlation for the operating temperature of photovoltaic modules of arbitrary mounting," *Solar Energy Materials and Solar Cells,* vol. 92, no. 11, pp. 1393-1402, 2008.

[54] R. Vaillon, O. Dupré, R. B. Cal and M. Calaf, "(2018). Pathways for mitigating thermal losses in solar photovoltaics. , 8(1), 13163.," *Scientific reports,* vol. 8, no. 1, p. 13163, 2018.

[55] A. P. Dobos, "An Improved Coefficient Calculator for the California Energy Commission 6 Parameter Photovoltaic Module Model.," *Journal of solar energy engineering,* vol. 134, no. 2, p. 021011, 2012.

[56] M. Wang, J. Peng, Y. Luo, Z. Shen and H. Yang, "(2021). Comparison of different simplistic prediction models for forecasting PV power output: Assessment with experimental measurements.," *Energy,* vol. 224, p. 120162, 2021.

[57] T. Huld, G. Friesen, A. K. R. P. Skoczek, T. Sample, M. Field and E. D. Dunlop, "A power-rating model for crystalline silicon PV modules.," *Solar Energy Materials and Solar Cells,* vol. 95, no. 12, pp. 3359-3369, 2011.

[58] R. Oliveira and C. L. T. Borges, "A comparison of photovoltaic models for estimating power generation: A case study of Brazilian data.," *Clean Technologies and Environmental Policy,* vol. 23, pp. 517-530, 2020.

[59] E. Saloux, A. Teyssedou and M. Sorin, "Explicit model of photovoltaic panels to determine voltages and currents at the maximum power point.," *Solar Energy,* vol. 85, no. 5, pp. 713-722, 2011.

[60] A. P. Dobos, "PVWatts version 5 manual," National Renewable Energy Lab. (NREL), Golden, CO (United States), 2014.

[61] M. R. Ramos, J. V. F. F. de Medeiros, D. Miranda, O. C. Vilela, A. C. Pereira, E. B. Jatoba and J. B. de Melo Filho, "Avaliação de desempenho da planta fotovoltaica de 2.5 MWp do CRESP com parametrização do software PVsyst.," in *Anais Congresso Brasileiro de Energia Solar-CBENS*, Florianópolis, SC (Brazil), 2022.

[62] M. Abraim, M. Salihi, O. El Alani, N. Hanrieder, H. Ghennioui, A. Ghennioui, M. Ydrissi and A. Azouzoute, "Techno-economic assessment of soiling losses in CSP and PV solar power plants: A case study for the semi-arid climate of Morocco. Energy," *Energy Conversion and Management,* vol. 270, p. 116285, 2022.

[63] D. G. Lorente, S. Pedrazzini, G. Zini, A. Dalla Rosa and P. Tartarini, "Mismatch losses in PV power plants.," *Solar Energy,* vol. 100, pp. 42-49, 2014.

[64] W. F. Holmgren, C. W. Hansen and M. A. Mikofski, "pvlib python: A python package for modeling solar energy systems.," *Journal of Open Source Software,* vol. 3, no. 29, p. 884, 2018.

[65] F. R. Martins, E. B. Pereira and S. L. Abreu, "Satellite-derived solar resource maps for Brazil under SWERA project.," *Solar Energy,* vol. 81, no. 4, pp. 517-528, 2007.

[66] M. C. Peel, B. L. Finlayson and T. A. McMahon, "Updated world map of the Köppen-Geiger climate classification.," *Hydrology and earth system sciences,* vol. 11, no. 5, pp. 1633-1644, 2007.

[67] D. R. Miranda, L. Petribu, J. V. F. F. de Medeiros, L. E. A. Barboza, O. C. Vilela, A. C. Pereira, E. B. Jatoba, A. Codeceira Neto and J. B. Melo Filho, "Quality control procedure for solar radiation at minute resolution.," in *EuroSun 2022*, Kassel (Germany), 2022.

[68] K. E. Taylor, "Summarizing multiple aspects of model performance in a single diagram," *Journal of Geophysical Research: Atmospheres,* vol. 106, no. D7, pp. 7183-7192, 2001.

[69] A. Mermoud and T. Lejeune, "Performance assessment of a simulation model for PV modules of any available technology.," in *Proceedings of the 25th European Photovoltaic Solar Energy Conference.*, Valencia (Spain), 2010.

[70] N. R. Kamphuis, C. A. Gueymard, M. T. Holtzapple, A. T. Duggleby and K. Annamalai, "Perspectives on the origin, derivation, meaning, and significance of the isotropic sky model.," *Solar Energy,* vol. 201, pp. 8-12, 2020.

[71] D. Yang, Z. Dong, A. Nobre, Y. S. Khoo, P. Jirutitijaroen and W. M. Walsh, "Evaluation of transposition and decomposition models for converting global solar irradiance from tilted surface to horizontal in tropical regions," *Solar Energy,* vol. 97, pp. 369-387, 2013.

[72] P. G. Loutzenhiser, H. Manz, C. Felsmann, P. A. Strachan, T. H. Frank and G. M. Maxwell, "Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation.," *Solar Energy,* vol. 81, no. 2, pp. 254-267, 2007.

[73] D. Yang, "Solar radiation on inclined surfaces: Corrections and benchmarks.," *Solar Energy,* vol. 136, pp. 288-302, 2016.

[74] A. Arias-Rosales and P. R. LeDuc, "Comparing view factor modeling frameworks for the estimation of incident solar energy.," *Applied Energy,* vol. 277, p. 115510, 2020.

[75] SAM, "System Advisor Model.," 2024. [Online].

[76] D. L. Evans and L. W. Florschuetz, "Cost studies on terrestrial photovoltaic power systems with sunlight concentration.," *Solar Energy,* vol. 19, no. 3, pp. 255-262, 1977.

[77] C. Schwingshackl, M. Petitta, J. E. Wagner, G. Belluardo, D. Moser, M. Castelli and A. Tetzlaff, "Wind Effect on PV Module Temperature: Analysis of Different Techniques for an Accurate Estimation.," *Energy Procedia,* vol. 40, pp. 77-86, Schwingshackl, C., Petitta, M., Wagner, J. E., Belluardo, G., Moser, D., Castelli, M., Tetzlaff, A..

[78] C. Li, S. V. Spataru, K. Zhang, Y. Yang and H. Wei, "A multi-state dynamic thermal model for accurate photovoltaic cell temperature estimation.," *IEEE Journal of Photovoltaics,* vol. 10, no. 5, pp. 1465-1473, 2020.