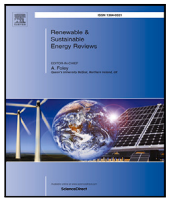
Renewable and Sustainable Energy Reviews 159 (2022) 112195

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser



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

Dazhi Yang

School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin, Heilongjiang, China

ARTICLE INFO

Keywords:
Solar radiation modeling
Separation modeling
Diffuse radiation
Worldwide validation
Benchmarking data
Review

ABSTRACT

Separation models, which are used to split beam and diffuse irradiance components from the global one, constitute the largest class of radiation models. Over the years, there have been more than 150 models proposed, and public views on the ranking of these models have been divergent, due to the climate-, weather-, and sky-condition dependency in model performance. In a study conducted in 2016, 140 separation models have been validated using high-quality radiometry data from 54 research-grade stations worldwide, which offer an objective and comprehensive assessment of the then available models. It was found that the Engerer2 model had the best overall performance. Since 2016, numerous other models have been proposed, and most of them are able to claim superiority over Engerer2, once again making the question "what is the best separation model to date" relevant. On this point, this article first reviews these latest advances in separation modeling. Next, as to promote fair comparison, an exceedingly comprehensive benchmarking dataset, which consists of 5 years (2016–2020) of 1-min data, from 126 stations located in all 7 continents and on islands in all 4 oceans, is considered in the empirical part of the article. With this dataset, which has more than 80 million valid 1-min data points, 10 latest separation models with diverse modeling philosophies are compared. It is found that the Yang4 model has the best overall performance, and thus is able to replace Engerer2 as the new quasi-universal separation model.

1. Introduction

In the field of solar energy meteorology, there are several well-known relationships in regards to shortwave radiation components. First and foremost, there is *closure equation*, which states the fact that the global irradiance is constituted of a beam component and a diffuse component:

$$G_h = B_n \cos Z + D_h, \tag{1}$$

where G_h, B_n, and D_h are global horizontal irradiance (GHI), beam normal irradiance (BNI), and diffuse horizontal irradiance (DHI), respectively, whereas Z is the solar zenith angle, which has the unit of degrees throughout this article. Since Z can be calculated to a high degree of accuracy via solar positioning, it is obvious from the closure equation that when two of three irradiance components are known, the third becomes deterministic.

The second relationship which is of absolute importance is the so-called *transposition equation*, which governs how irradiance components on a horizontal surface can be mapped to that on an inclined surface:

$$G_c = B_n \cos\theta + R_d D_h + \rho R_r G_h, \tag{2}$$

where G_c is the global tilted irradiance (GTI); θ is the incidence angle in degrees, which, just like Z, can be calculated with a solar positioning algorithm, for any inclined surface with known tilt and azimuth angles; R_d is the diffuse transposition factor, which can also be interpreted as the sky view factor, i.e., the effective part of sky "seen" by the inclined surface; R_r is the transposition factor due to ground reflection, and ρ is the foreground's albedo, which is a time-varying quantity describing the fraction of sunlight being reflected by the surface—it is often obtained through remote sensing or numerical weather prediction.

Besides the closure and transposition equations, there are k-indexes, which can be thought of as various normalized versions of irradiance:

$$\kappa = \frac{G_h}{G_{\text{csky}}}, \tag{3}$$

Abbreviations: BNI, beam normal irradiance; DHI, diffuse horizontal irradiance; GHI, global horizontal irradiance; GTI, global tilted irradiance; QC, quality control

E-mail address: yangdazhi.nus@gmail.com.

https://doi.org/10.1016/j.rser.2022.112195
Received 8 October 2021; Received in revised form 3 January 2022; Accepted 23 January 2022
Available online 23 February 2022

Nomenclature

α	$90° - Z$, elevation angle [degree]
Δk_{tc}	$k_{tc} - k_t$, difference between clearness index of clear-sky global horizontal irradiance and clearness index [dimensionless], as needed by Engerer2 and Yang4
κ	G_h/G_{csky}, clear-sky index for global horizontal irradiance [dimensionless], as needed by Starke1, Starke2, and Starke3
ψ	Three-point moving average of 1-min k_t [dimensionless], as needed by Starke1, Starke2, and Starke3
ρ	Foreground's albedo [0–1]
θ	Incidence angle [degree]
B_n	Beam normal irradiance [W/m^2]
D_h	Diffuse horizontal irradiance [W/m^2]
E_0	Horizontal extraterrestrial irradiance [W/m^2]
G_{csky}	Clear-sky global horizontal irradiance [W/m^2]
G_c	Global tilted irradiance [W/m^2]
G_h	Global horizontal irradiance [W/m^2]
k_d	D_h/G_h, diffuse fraction [0–1]
$k_d^{(s)}$	Hourly or half-hourly satellite-derived diffuse fraction [0–1]
k_t	G_h/G_{ext}, clearness index [dimensionless]
$k_{d,\text{hourly}}^{\text{Engerer2}}$	Hourly diffuse fraction estimate, obtained by applying Engerer2 to hourly data [0–1], as needed by Yang4
k_{de}	$\max\left(0, 1 - \frac{G_{\text{csky}}}{G_h}\right)$, part of diffuse fraction that is attributed to cloud enhancement [dimensionless], as needed by Engerer2 and Yang4
$k_{t,\text{daily}}$	Daily average of k_t [dimensionless], as needed by Starke1, Starke2, Starke3, and Paulescu
$k_{t,\text{hourly}}$	Hourly average of k_t [dimensionless], as needed by Starke3
k_{tc}	G_{csky}/E_0, clearness index of clear-sky global horizontal irradiance [dimensionless], as needed by Engerer2 and Yang4
R_d	Diffuse transposition factor [dimensionless]
R_r	Transposition factor due to ground's reflection [dimensionless]
Z	Zenith angle [degree]
AST	Apparent solar time [0–24]

$$k_t = \frac{G_h}{E_0}, \tag{4}$$

$$k_d = \frac{D_h}{G_h}. \tag{5}$$

The ratio between GHI and its clear-sky expectation (G_{csky}) is called the clear-sky index (κ). The ratio between GHI and the horizontal extraterrestrial irradiance (E_0) is called the clearness index (k_t). And the ratio between DHI and GHI is called the diffuse fraction (k_d). Although Eqs. (1)–(5) appear simple in form, a large fraction of energy meteorology research is in fact consisted in topics revolved around these equations—such works are jointly known as radiation modeling.

1.1. A bird's eye view on solar radiation modeling

The closure equation is useful during quality control (QC) of radiometry measurements, in that, if the three components do not conform to this relationship, at least one of them must be spurious, and thus the triplet needs to be rejected [1]. More specifically, for $Z \leq 75°$ and GHI > 50 W/m^2, if the difference between the left- and right-hand-side of Eq. (1) is more than 8%, the three-component closure test fails; for $Z > 75°$ and GHI > 50 W/m^2, the tolerance is 15% [1]. This basic QC routine has often been employed, as long as measurements of all three radiation components are involved [e.g., 2–4]. Other formalisms regarding QC of radiation data are described in a later section.

The transposition equation gives rise to transposition modeling, which, in large part, focuses on developing methods pertaining to the estimation of R_d in Eq. (2). Ranging from the simple isotropic model [5] to the more complex three-part geometrical framework as described by the Perez family of models [6–9], there have been over 30 transposition models proposed in the literature. Since the performance of these models is location-, weather-, climate-, and sky-condition-dependent, works and reviews which seek to compare transposition models were once popular. Nevertheless, very few conclusions from works of this kind can be deemed trustworthy, because they are either limited by the number of datasets or number of models compared, or contain one or more interpretation mistakes or coding errors. The reader is referred to the review by Yang [10] for a list of more than one hundred mistakes and errors found in the literature. Indeed, the review by Yang [10] is also the most extensive one by far, in terms of number of datasets and number of models considered. It was found that the Perez family of models has the best overall performance, and nothing much has changed since. Whereas the stature of the Perez model is difficult to surpass, recent innovations focus on other aspects of transposition modeling, such as inverse transposition modeling [11,12], uncertainty quantification [13], or inferring unknown orientations of distributed PV [14].

As suggested by Eq. (3), the calculation of clear-sky index requires G_{csky}, which is estimated via clear-sky models. Clear-sky models proposed in the literature are more numerous than transposition models, and they can be broadly categorized into those that consider atmospheric physics and those that do not. It should be noted that the word "clear" suggests a cloud-free sky, but not an "atmosphere-less" sky. Indeed, aerosol, water vapor, trace gases, surface reflectivity can all affect the amount of radiation reaching the earth's surface. A good clear-sky model could therefore account for the effect of atmospheric constituents on radiation through physical modeling—radiative transfer. On the other hand, empirical clear-sky models simply fit some data points to some bell-shaped function forms, and are generally inferior to the physical ones in terms of accuracy. Given the cardinal importance of clear-sky radiation modeling in solar resource assessment and forecasting, e.g., see [15–18], validation and comparison studies are too many to list. Nevertheless, the results presented in the pair of recent reviews by Sun et al. [19,20], in which 75 GHI clear-sky models and 95 BNI and DHI clear-sky models are assessed at worldwide locations, can be said to be truly conclusive. In that, a proprietary research version of the REST2 clear-sky model [21] was found to be the best one to date.

The clearness index in Eq. (4) is obtained as the ratio of G_h and E_0, where the latter describes the radiation just outside the earth's atmosphere. Whereas it is understood *a priori* that E_0 is function of the sun–earth distance, which varies throughout a year, it also depends on a quantity known as the solar constant. Because the sun's output is perpetually changing, solar constant is obtained through calculating the average of instantaneous total solar irradiance (TSI), over a long period of time, typically over a few decades. Since TSI is measured by multiple sets of spaceborne instruments, some modeling is needed for reconciliation of these measurements. On this point, Gueymard [22] performed an evaluation of the solar constant based on 42 years of

spaceborne observations of TSI. It is suggested that the value of 1361.1 W/m^2 is adequate, despite most software packages are still using the outdated value of 1366.1 W/m^2.

Aside from the above modeling topics, the field of solar energy meteorology is also interested in several other tasks. For instance, the irradiance components that have been discussed so far are broadband irradiance, as opposed to which there is spectral irradiance. Spectral irradiance has a wide range of applications not only in solar engineering, but also in photobiology and photochemistry, among others. The most representative spectral irradiance model is the SMARTS model, the detailed research issues of which has been fully elaborated by Gueymard [23]. Another major class of models is related to satellite-derived irradiance modeling, which seeks to derive irradiance from remote sensing imagery of top-of-atmosphere reflectance, and the reader is referred to Huang et al. [24] and Miller et al. [25] for reviews. Given the fact that satellite-derived irradiance is often found to be biased, in a spatially inhomogeneous fashion, post-processing of the raw irradiance estimates is of interest and is in fact necessary, which leads to another type of radiation models, namely, site-adaptation models, see [26–29] for reviews and latest advances. That said, since these topics are not central to this work, they are not discussed any further.

1.2. Separation modeling prior to 2016

At this stage, the only remaining type of radiation models which has not been discussed is separation modeling. In short, separation models aim at estimating the diffuse fraction, k_d. It is clear from Eq. (5) that once k_d is estimated, one is able to retrieve D_h from G_h. Subsequently, B_n can be obtained using Eq. (1). Stated differently, separation models use G_h (and other *computable* parameters, such as Z or k_t) as input, and output k_d. The reason that separation models are in high demand is three-fold, which is elaborated next.

The first source of demand for separation models is this: unlike GHI, which can be measured with a stationary pyranometer, DHI and BNI measurements both require a tracker, whose two-axis rotation follows the position of the sun as closely as possible. Such tracker contains high-precision mechanical moving parts, and therefore is very costly as compare to the pyranometer. Hence, very few ground-based radiometry stations measure all three components, instead, only GHI is frequently measured. (Even so, the number of radiometry stations falls far short of the number of stations for other basic meteorological variables, such as temperature, humidity, or pressure.) It is on that account that DHI and BNI need to be estimated. Following that, the second attraction of separation modeling lies in the fact that DHI and BNI are required by transposition modeling, which is essential insofar as flat-surface solar collectors, such as photovoltaic (PV), are concerned, because these collectors are installed with an inclination, as to maximize their annual power production. Indeed, during the conversion from irradiance to PV power output, splitting B_n and D_h from G_h is often the foremost step [30,31]. The third motivation for separation modeling is related to remote sensing. Because ground-based radiometry measurements are scarce, deriving irradiance from satellite imagery data is essential for resource assessment and forecasting purposes. However, almost all satellite-to-irradiance models at present only retrieve G_h. It is for that reason that B_n and D_h in satellite-derived databases, including the ones offered by the National Renewable Energy Laboratory (NREL) [15], SolarAnywhere,[1] Solargis,[2] Solcast,[3] are obtained through a separation modeling means.

The high demand for accurate separation models has spawned a rich literature. Philosophically, since the value of academic publications is materialized mainly through the notion of improvement, in that the new theory and methods should supersede the old ones, almost all publications on separation modeling are able to claim superiority in one form or another. Technically, unlike transposition models and clear-sky models, almost all separation models are empirical; this implies that the bar of entry in this line of research is fairly low, and anyone who owns a dataset can perform her (solar engineers are assumed to be feminine in this article) modeling and fitting. Combining both arguments, the views on the performance ranking of separation models had been divergent. It was not until the review by Gueymard and Ruiz-Arias [32] that the subject was fully reconciled.

In 2016, Gueymard and Ruiz-Arias [32] compared a total of 140 separation models, which represent all modeling techniques at that time, using 1-min data from 54 research-grade stations worldwide. Ranging from the single-predictor models, which use k_t as the sole input, to multi-predictor ones, which involve auxiliary parameters such as zenith angle, time of day, dry-bulb temperature, surface albedo, or variability index, the review shows an unprecedented level of effort on collecting, arranging, and assessing separation models, in an objective and rigorous fashion. Whereas none of the models was found to be able to outperform the others at all locations, the Engerer2 model[4] [33] was found to be quasi-universal, based on statistical results averaged over different climate types. The reader is referred to Section 2 for the modeling details of Engerer2. In short, the success of Engerer2 can be attributed to its explicit representation of cloud-enhancement events, which refer to the sky conditions with passing broken clouds, where the beam component is often not obstructed by clouds, but the value of diffuse irradiance may be boosted further by sunlight reflected off cloud edges [see 34,35, for reviews]. Also known as over-irradiance, cloud enhancement is able to bring G_h to a value much higher than G_{csky}, for a time period of a few seconds to a few minutes. Since 1-min G_h is highly likely to be influenced by cloud enhancement, Engerer2 uses an additive term to represent such boost in diffuse fraction.

1.3. Philosophy of conducting research in separation modeling

> "Compared with new methods of estimating solar radiation, sunshine data are of poor and inconsistent quality, and empirical correlations developed for one site rarely apply to conditions from other regions. There have been numerous studies of this type published in this journal or others, so this in not a scientifically original topic. Other types of empirical models, such as those used to separate the direct and diffuse components, are also of little interest if only applicable to specific sites. Only manuscripts offering a novel approach of 'universal' appeal in these areas, and providing some new insight into these problems, will be published."
>
> — Gueymard, Renné, and Vignola [36]

When we examine the literature of separation modeling, it would not be long before some peculiar articles can be identified [e.g., 37–42]. Albeit their publication date is often fairly recent, their modeling approaches and the time scale of interest seem to be detached from the rest of the literature. More specifically, these are works that leverage regression models on monthly or daily radiation and sunshine duration data to perform separation modeling. If we take a closer look, it is easy to conclude several characteristics from articles of this kind: (1) the literature review section contains excessive outdated references from the 1960s, 1970s and 1980s; (2) in each article, tens of models are developed; (3) the function form of the developed models is often

[1] https://www.cleanpower.com/.
[2] https://www.solargis.com/.
[3] https://solcast.com.

[4] It is convenient to express a model using the first author's name in Small Caps, and if there are multiple versions, a number is added.

limited to just simple polynomial regression; and (4) the developed model are validated on a very small number (a few or tens) of data points. It is no surprise that one should question the need for existence of these works. The rationale is very simple: if tens of models can be developed based on just a few data points from a few stations, how many of these models can we developed and how many do we really need? This question is a rhetorical one. Indeed, the limitation and inadequacy of models of this sort have long been argued by domain experts, who advocate abandoning such research works, formally, first in 1993 [43] and then in 2009 [36], see above quote. Ignorance is bliss, until something happens. The shadows on *Plato's Cave* wall are actually not reality, but only what the prisoner sees—in a world that is getting more and more closely interconnected, it is puzzling why there are still people conducting such research without knowing the field has leaped forward long ago.

If one can agree to the results and conclusions presented by Gueymard and Ruiz-Arias [32], it is then quite obvious that making comparison to Engerer2 is absolutely necessary whenever new separation models are being proposed. Otherwise, the so-claimed novelty and performance superiority cannot be truly justified at any rate. Why should one go to great lengths to demonstrate goodness of the proposed model, while she can simply compare it to Engerer2? The rule of thumb of separation modeling research is therefore this: Engerer2 ought to be included as a benchmark, until a new quasi-universal model is identified. Consequently, one can leverage this rule of thumb to screen the quality of manuscripts in the literature—those manuscripts that do not include Engerer2 as a benchmark [e.g., 44–46] either have conducted literature review in an incomplete way, by being ignorant about the existence of the quasi-universal model, or are written by intellectually dishonest wielders of *Occam's broom* who intentionally hide away the inconvenient fact that the proposed model is unable to outperform Engerer2. In both cases, dismissing the manuscript seems to be a wiser choice than using it.

The next argument pertains to the generalization ability of separation models. As mentioned earlier, most separation models are empirical, which implies that they require fitting. And when the model coefficients are fitted using data from different locations, the performance of the model would naturally vary. In this regard, the generalization ability of a model refers to how well it is able to perform at "unseen" locations. Philosophically, the type of arguments as to claim the superiority of the proposed model is always inductive. Stated differently, through case studies, it is possible to show the proposed model is better than the benchmark at some locations, but the practical relevance of separation modeling is always in regard to those situations where the model has yet to be verified. It is on this account that another class of separation models—the one that leverages machine learning—attracts skepticism. For instance, some authors [e.g., 46,47] proposed mappings from k_t (and other inputs) to k_d via machine learning, whereas others [e.g., 48] used machine learning to post-process k_d estimates from existing separation models. Because machine learning is often sensitive to training data, the generalization ability of these models is unclear. When the training and testing datasets come from the same location, the situation is incestuous, and any conclusion made would be invalid; articles that commit such fallacy include [45–47] among others. Even if the machine-learning models are tested with extensive data, they are still faced with the transferability issue, in a sense that the trained models cannot be easily used by others, unless reproducibility is ensured in its entirety, which is rarely practiced anyway. It is perhaps for those reasons that the actual uptake of machine-learning separation models has hitherto been limited.

1.4. Contributions and article organization

Based on the above discussion, this article seeks to narrow the research gap by: (1) providing a review on recent developments on the subject of separation modeling, see Section 2; (2) introducing a universal benchmarking dataset that is able to test the generalization ability of separation models, see Section 3; and (3) conducting an extensive worldwide validation of latest separation models at 126 locations, see Section 4.

It is noted that the dataset adopted in this work is a collective effort of the International Energy Agency's (IEA's) PV Power Systems (PVPS) Task 16 Activity 1.4 members, and its comprehensiveness and quality can be regarded as the best-available ones. If any separation model can be shown to possess advantage over others on this dataset, it is most likely to have satisfactory performance elsewhere. In the present case, through the validation as shown in Section 4, it is found that there is no model that can consistently outperform its peers at all sites, which agrees with the findings of Gueymard and Ruiz-Arias [32]. Nevertheless, the Yang4 model [49] has the best overall performance, and therefore should replace Engerer2, as the new quasi-universal model. More importantly, as compared to some models, whose coefficients are climate-dependent, Yang4 contains just one set of coefficients for all scenarios, which is clearly more advantageous in terms of model simplicity. In fact, the benefits of using climate-dependent coefficients are not dominantly clear, as an all-climate version of a model is often found better than a climate-specific version of the same model.

2. Recent advances in separation modeling

Although cloud enhancement has been known at an earlier time, Engerer2 is perhaps the first model that explicitly considers its effects on diffuse radiation, and the importance of such consideration has been made fully evident through the validation by Gueymard and Ruiz-Arias [32]. Nevertheless, if we are to develop a truly universal model, other innovations that can elevate the performance of separation modeling must be taken jointly into consideration. Since many of the modeling issues had already been examined in 2016, this section should focus mainly on the state of affairs after that, and each innovation is arranged as a subsection below. To facilitate the present discussion, some useful input parameters are summarized in Table 1.

2.1. Model refitting

Engerer2 [33] is expressed as:

$$k_d^{\text{Engerer2}} = C + \frac{1-C}{1+e^{\beta_0+\beta_1 k_t+\beta_2 \text{AST}+\beta_3 Z+\beta_4 \Delta k_{tc}}} + \beta_5 k_{de}, \tag{6}$$

where $C = 0.042336$, $\beta_0 = -3.7912$, $\beta_1 = 7.5479$, $\beta_2 = -0.010036$, $\beta_3 = 0.003148$, $\beta_4 = -5.3146$, and $\beta_5 = 1.7073$. Given

$$k_{de} = \max\left(0, 1 - \frac{G_{\text{csky}}}{G_h}\right), \tag{7}$$

it can be seen that if and only if $G_h > G_{\text{csky}}$, which denotes a cloud-enhancement event, $k_{de} > 0$. The $\beta_5 k_{de}$ term in Eq. (6) therefore accounts for the additional amount of diffuse fraction that is due to cloud enhancement.

Since the original model coefficients were fitted using 1-min data from 6 sites in Australia, one logically attractive modification is to refit the model using data from other locations and from other temporal resolutions, such that the newly fitted models can be fine-tuned for situations other than that described by the 1-min Australia data. Such a work has been conducted by Bright and Engerer [50], who refitted a "world version" of Engerer2 using data from 70 sites worldwide, with different temporal resolutions, ranging from a minute to a day. Because the original article by Engerer [33] presented three models, the newly fitted 1-min model is herein named Engerer4. Engerer4 has the same expression as Engerer2, but with $C = 0.10562$, $\beta_0 = -4.1332$, $\beta_1 = 8.2578$, $\beta_2 = 0.010087$, $\beta_3 = 0.00088801$, $\beta_4 = -4.9302$, and $\beta_5 = 0.44378$. Despite the intention, the performance of Engerer4 has been shown to be inferior to Engerer2, in several later studies [49,51]. A possible explanation is under-fitting, which is potentially caused by the overwhelming diversity of the training samples [51].

$$k_d^{\text{STARKE1}} = \begin{cases} \dfrac{1}{1+e^{\beta_7+\beta_8 k_t+\beta_9 \text{AST}+\beta_{10}\alpha+\beta_{11}k_{t,\text{daily}}+\beta_{12}\psi+\beta_{13}G_{\text{csky}}/277.78}}, & \kappa \geq 1.05 \text{ and } k_t > 0.65; \\ \dfrac{1}{1+e^{\beta_0+\beta_1 k_t+\beta_2 \text{AST}+\beta_3\alpha+\beta_4 k_{t,\text{daily}}+\beta_5\psi+\beta_6 G_{\text{csky}}/277.78}}, & \text{otherwise,} \end{cases} \tag{8}$$

Table 1
Various input parameters as required by different separation models.

Parameter	Calculation method	Interpretation
G_{csky}	McClear clear-sky model	Clear-sky GHI
Z	Compute via solar positioning	Solar zenith angle in degrees
α	$90° - Z$	Solar elevation angle in degrees
AST	Compute via solar positioning	apparent solar time
k_t	G_h/E_0	Clearness index
$k_{t,\text{daily}}$	Average k_t over a day	Low-frequency k_t signal, a form of variability index
$k_{t,\text{hourly}}$	Average k_t over an hour	Low-frequency k_t signal, a form of variability index
ψ	Three-point moving average of k_t	Low-frequency k_t signal, a form of variability index
k_{tc}	G_{csky}/E_0	Clearness index of clear-sky GHI
Δk_{tc}	$k_{tc} - k_t$	Difference between clearness index of clear-sky GHI and clearness index
k_{de}	$\max(0, 1 - G_{\text{csky}}/G_h)$	Portion of the diffuse fraction that is attributable to cloud enhancement events
κ	G_h/G_{csky}	Clear-sky index for GHI
$k_d^{(s)}$	Retrieve from satellite-derived irradiance database	Half-hourly or hourly satellite-derived diffuse fraction
$k_{d,\text{hourly}}^{\text{ENGERER2}}$	Apply ENGERER2 on hourly G_h	Hourly diffuse fraction estimate from ENGERER2, which is a form of variability index

2.2. Piecewise modeling

Piecewise modeling of diffuse fraction has a long history, and many early models in fact had already employed this strategy, such as ERBS [52] or ORGILL [53]. Now that the benefits of including the effects of cloud enhancement are known, Starke et al. [54] proposed a piecewise model, which differentiates conditions with and without cloud enhancement. Furthermore, two sets of model coefficients are fitted using data from Australia and Brazil, respectively. In this article, the model fitted using Australia data is referred to as STARKE1, and that using Brazil data as STARKE2. The model is given as: see Eq. (8), where $\beta_0 = -6.70407$, $\beta_1 = 6.99137$, $\beta_2 = -0.00048$, $\beta_3 = 0.03839$, $\beta_4 = 3.36003$, $\beta_5 = 1.97891$, $\beta_6 = -0.96758$, $\beta_7 = 0.15623$, $\beta_8 = -4.21938$, $\beta_9 = -0.00207$, $\beta_{10} = -0.06604$, $\beta_{11} = 2.12613$, $\beta_{12} = 2.56515$, and $\beta_{13} = 1.62075$. And STARKE2 is the same as STARKE1 in form, except that $\beta_0 = -6.37505$, $\beta_1 = 6.68399$, $\beta_2 = 0.01667$, $\beta_3 = 0.02552$, $\beta_4 = 3.32837$, $\beta_5 = 1.97935$, $\beta_6 = -0.74116$, $\beta_7 = 0.19486$, $\beta_8 = -3.52376$, $\beta_9 = -0.00325$, $\beta_{10} = -0.03737$, $\beta_{11} = 2.68761$, $\beta_{12} = 1.60666$, and $\beta_{13} = 1.07129$. In Eq. (8), the condition $\kappa \geq 1.05$ and $k_t > 0.65$ marks the boundary of what constitutes a cloud-enhancement event.

It should be noted that the modeling philosophy of STARKE1 and STARKE2 follows that of RIDLEY2 [55], or the BRL model, as also commonly known, which utilizes the logistic function, whose shape resembles the k_t–k_d relationship. Furthermore, when multivariate input is used, the output becomes a surface instead of a line, which provides a better coverage of the k_t–k_d scatter, see Fig. 1 for visualization.

Table 2
Coefficients of ABREU. Abbreviation for climate zone: Arid (AR), High Albedo (HA), Temperate (TM) and Tropical (TR).

Coefficient	Climate			
	AR	HA	TM	TR
A	11.39	7.83	10.79	11.59
B	−6.25	−4.59	−5.87	−6.14
n	1.86	3.25	2.24	1.87

2.3. New mathematical function

Speaking of function forms, there are many other mathematical functions that could be deemed suitable for separation modeling, as far as the shape of function is able to sufficiently trace out the correspondence between k_t and k_d. Noticeable is the choice of Abreu et al. [56], which takes the form: $y = (1 + x^{-n})^{(-1/n)}$, where x itself is a second-degree polynomial function of k_t. That is,

$$k_d^{\text{ABREU}} = \left\{1 + \left[A \times (k_t - 0.5)^2 + B \times (k_t - 0.5) + 1\right]^{-n}\right\}^{-\frac{1}{n}}, \tag{9}$$

where A, B, and n take different values for different climates, as shown in Table 2. (We should circle back to climate-specific models shortly after.) This particular choice presents a function form that curls up for high k_t values, which leads to a better representation of data points under cloud enhancement. This model is noted as ABREU, which is a one-parameter model.

Besides ABREU, another choice is presented by Paulescu and Blaga [57], who combined multiple regression with indicator functions, such that the model can trace out a segmented surface, depending on whether the indicator functions are activated or not. More specifically,

$$\begin{aligned} k_d^{\text{PAULESCU}} &= \beta_0 + \beta_1 k_t + \beta_2 k_{t,\text{daily}} + \beta_3 \left(k_t - \beta_4\right) \mathcal{I}_{k_t \geq \beta_4} \\ &+ \beta_5 \left(k_t - \beta_6\right) \mathcal{I}_{k_t \geq \beta_6} + \beta_7 \left(k_{t,\text{daily}} - \beta_8\right) \mathcal{I}_{k_{t,\text{daily}} \geq \beta_8}, \end{aligned} \tag{10}$$

where $\mathcal{I}_{\text{condition}}$ is an indicator function whose value equals to 1 if the condition in the subscript is satisfied, 0 otherwise; $\beta_0 = 1.0119$, $\beta_1 = -0.0316$, $\beta_2 = -0.0294$, $\beta_3 = -1.6567$, $\beta_4 = 0.367$, $\beta_5 = 1.8982$, $\beta_6 = 0.734$, $\beta_7 = -0.8548$, and $\beta_8 = 0.462$.

2.4. Satellite-augmented modeling

One essential trick for separation modeling is to identify auxiliary input parameters that can enhance the explanatory power of the model. In the usual case, the input parameters need to be *computable*, which means they can be calculated for any arbitrary time instance, for any location, and most importantly, do not depend on diffuse or beam component. Commonly used auxiliary parameters include zenith angle, clear-sky irradiance, time of day, among others. That said, there are also auxiliary parameters that can only be obtained, for arbitrary locations, through remote-sensing or reanalysis means, such as surface albedo. At this stage, a new thought emerges: if remotely sensed surface albedo is useful, what other forms of satellite augmentation could we leverage? The most straightforward choice would no doubt be satellite-derived diffuse fraction, or $k_d^{(s)}$. This indeed has been considered by Yang and Boland [58], whose model takes half-hourly or hourly satellite-derived diffuse fraction as an additional parameter. This

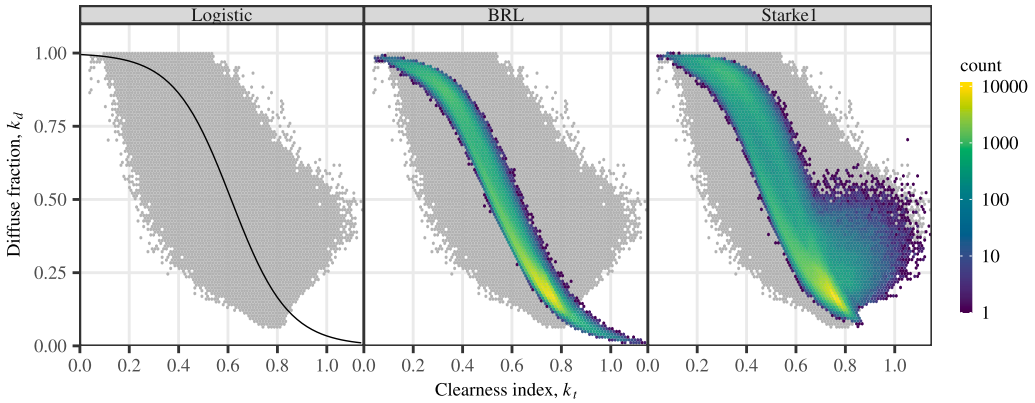


Fig. 1. One-minute diffuse fraction prediction using the logistic function, BRL, and Starke1 models, using data from Carpentras (44.083°N, 5.059°E), France, over 2015–2018. Measurements are shown as the gray background, and predictions are shown in viridis color scale.

innovation gives rise to two models, namely, Yang1 and Yang2. Both models are modified from Engerer2:

$$k_d^{\text{Yang1}} = C + \frac{L}{1 + e^{\beta_0+\beta_1 k_t+\beta_2 \text{AST}+\beta_3 Z+\beta_4 \Delta k_{tc}}} + \beta_5 k_{de} + \beta_6 k_d^{(s)}, \tag{11}$$

where $C = 0.0369$, $\beta_0 = -3.4986$, $\beta_1 = 7.9735$, $\beta_2 = -0.0030$, $\beta_3 = 0.0031$, $\beta_4 = -7.6836$, $\beta_5 = 1.0179$, $\beta_6 = 0.3505$ and $L = 0.6768$, whereas

$$k_d^{\text{Yang2}} = C + \frac{1 - C}{1 + e^{\beta_0+\beta_1 k_t+\beta_2 \text{AST}+\beta_3 Z+\beta_4 \Delta k_{tc}+\beta_6 k_d^{(s)}}} + \beta_5 k_{de}, \tag{12}$$

where $C = 0.0361$, $\beta_0 = -0.5744$, $\beta_1 = 4.3184$, $\beta_2 = -0.0011$, $\beta_3 = 0.0004$, $\beta_4 = -4.7952$, $\beta_5 = 1.4414$, and $\beta_6 = -2.8396$.

2.5. Temporal-resolution cascade

Using half-hourly or hourly satellite-derived diffuse fraction as an additional input parameter is able to bring substantial accuracy improvement [58]. However, the need for acquiring $k_d^{(s)}$ introduces additional workload, since each satellite-derived irradiance database uses a somewhat different convention for data dissemination, see [59] for a review on the popular databases. More importantly, most, if not all, satellite-derived irradiance databases are not updated in real time, instead, there is a delay ranging from a few days to a couple of years before the data is released. This fact hence limits the real-time application of Yang1 and Yang2. To remedy the situation, Yang [49] proposed replacing $k_d^{(s)}$ with a low-frequency estimate of diffuse fraction, calculated using, for instance, Engerer2. Denoting the hourly diffuse fraction estimate using Engerer2 as $k_{d,\text{hourly}}^{\text{Engerer2}}$, the model becomes:

$$k_d^{\text{Yang4}} = C + \frac{1 - C}{1 + e^{\beta_0+\beta_1 k_t+\beta_2 \text{AST}+\beta_3 Z+\beta_4 \Delta k_{tc}+\beta_6 k_{d,\text{hourly}}^{\text{Engerer2}}}} + \beta_5 k_{de}, \tag{13}$$

where all coefficients follow Yang2.

The reason that the model in Eq. (13) is named Yang4 rather than Yang3 is because, in the original publication, there is also a version which uses 15-min diffuse fraction estimate, i.e., $k_{d,15\text{ min}}^{\text{Engerer2}}$, which was named Yang3. Since both Yang3 and Yang4 can be thought of as cascading two separation models with different temporal resolutions, this modeling strategy has been coined the name of temporal-resolution cascade by Yang [49]. The strategy is general, as one can replace the low-frequency Engerer2-based diffuse fraction estimate with that computed using any other separation model. Nevertheless, such possibilities are not investigated further in this article, for brevity.

2.6. Variability index

The Yang family of models has been shown to have superior performance over all other above-mentioned models [49,58]. This is partly due to its inheritance of the cloud-enhancement modeling of Engerer2. However, since the Yang models are able to outperform Engerer2, by significant margins, the additional predictive power must be attributed to the inclusion of $k_d^{(s)}$ or $k_{d,\text{hourly}}^{\text{Engerer2}}$, which essentially represents the low-frequency trend in k_d. Stated differently, $k_d^{(s)}$ and $k_{d,\text{hourly}}^{\text{Engerer2}}$ can be viewed as a form of variability index, describing the low-frequency variations in k_t–k_d correspondence.

The original idea of including a variability index came from Perez et al. [60]. Subsequently, this idea has attracted attention over the years, and different forms of variability index have been proposed, among which the daily mean k_t, or $k_{t,\text{daily}}$, is one of the most popular choices. The parameter $k_{t,\text{daily}}$ has been employed in many derivative works of the BRL model, such as Starke1 and Starke2. In a recent work [61], an expansion of the Starke1 and Starke2 further includes the hourly mean k_t, or $k_{t,\text{hourly}}$, into modeling, which leads to Starke3: see Eq. (14), where climate-dependent coefficients are displayed in Table 3.

Four subtle differences between Starke3 and its earlier versions are to be highlighted. One of those is that the coefficient of G_{csky} in Starke3 is fitted using G_{csky} data with a unit of W/m^2, but the coefficients of G_{csky} in Starke1 and Starke2 are fitted using G_{csky} data with a unit of MJ/(h m^2). Since 1 MJ/(h m^2) equals to 277.78 W/m^2, one can see that Eq. (8) has already factored in the issue of inconsistent units, such that G_{csky} in W/m^2 can be used across all models. Secondly, the conditioning of Starke1 and Starke2, as described in the original work, was incomplete, because it did not include $\kappa \geq 1.05$ and $k_t \leq 0.65$ scenarios, see Eq. (2) of Starke et al. [54]. This oversight has also been corrected in this article, as written in Eq. (8). Thirdly, the conditioning

Table 3
Coefficients of Starke3. Climate abbreviation follows the first letter of Köppen–Geiger climate classification.

Coefficient	Climate				
	A	B	C	D	E
β_0	0.29566	−1.7463	−0.083	0.67867	0.51643
β_1	−3.64571	−2.20055	−3.14711	−3.79515	−5.32887
β_2	−0.00353	0.01182	0.00176	−0.00176	−0.00196
β_3	−0.01721	−0.03489	−0.03354	−0.03487	−0.07346
β_4	1.7119	2.46116	1.40264	1.33611	1.6064
β_5	0.79448	0.70287	0.81353	0.76322	0.74681
β_6	0.00271	0.00329	0.00343	0.00353	0.00543
β_7	1.38097	2.30316	1.95109	1.82346	3.53205
β_8	−7.00586	−6.53133	−7.28853	−7.90856	−11.70755
β_9	6.35348	6.63995	7.15225	7.63779	10.8476
β_{10}	−0.00087	0.01318	0.00384	0.00145	0.00759
β_{11}	0.00308	−0.01043	0.02535	0.10784	0.53397
β_{12}	2.89595	1.73562	2.35926	2.00908	1.76082
β_{13}	1.13655	0.85521	0.83439	1.12723	0.41495
β_{14}	−0.0013	−0.0003	−0.00327	−0.00889	−0.03513
β_{15}	2.75815	2.63141	3.19723	3.72947	6.04835

$$k_d^{\text{Starke3}} = \begin{cases} \dfrac{1}{1+e^{\beta_0+\beta_1 k_t+\beta_2 \text{AST}+\beta_3\alpha+\beta_4 k_{t,\text{daily}}+\beta_5\psi+\beta_6 G_{\text{csky}}+\beta_7 k_{t,\text{hourly}}}}, & \kappa \geq 1.05 \text{ and } k_t > 0.75; \\ \dfrac{1}{1+e^{\beta_8+\beta_9 k_t+\beta_{10}\text{AST}+\beta_{11}\alpha+\beta_{12} k_{t,\text{daily}}+\beta_{13}\psi+\beta_{14} G_{\text{csky}}+\beta_{15} k_{t,\text{hourly}}}}, & \text{otherwise,} \end{cases} \tag{14}$$

Table 4
Coefficients of Every2. Climate abbreviation follows the Köppen–Geiger climate classification.

Coefficient	Climate									
	Am	Aw	BSh	BSk	BWh	Cfa	Cfb	Csa	Csb	Other
β_0	−6.433	−6.047	−6.734	−7.310	−7.097	−6.484	−6.764	−7.099	−7.080	−5.38
β_1	8.774	7.540	8.853	10.089	9.416	8.301	9.958	10.152	10.460	6.63
β_2	−0.00044	0.00624	0.02454	0.01852	0.01254	0.01577	0.01271	−0.00026	0.00964	0.006
β_3	−0.00578	−0.00299	−0.00495	−0.00693	−0.00416	−0.00338	−0.01249	−0.00744	−0.01420	−0.007
β_4	2.096	2.077	1.874	1.296	1.661	1.607	0.928	1.147	1.134	1.75
β_5	0.684	1.208	0.939	1.114	1.130	1.307	1.142	1.184	1.017	1.31

of Starke3 has also been modified. Instead of using $k_t > 0.65$, this version uses $k_t > 0.75$. Last but not least, the indexing for coefficients of Starke3, that is, $\beta_0, \ldots, \beta_{15}$ is different from that of Starke1 and Starke2—for instance, β_0 in Starke3 takes the position of β_7 in Starke1. Generally speaking, these ambiguities and standardization issues are likely to cause implementation errors, and thus must be warned.

2.7. Climate-specific fitting

At this stage, two climate-specific models have been presented, namely, Abreu and Starke3, despite that their climate classification systems are different. This idea of obtaining climate-specific model coefficients is straightforward, and models with such fitting may appear to have *a priori* advantages. In other words, when the fitting data is climate-specific, it is natural to expect the model to demonstrate better performance under similar climatic conditions. For this reason, aside from Abreu and Starke3, there are a few other works which adopt climate-specific fitting, such as [62]. Every et al. [62] first refitted the BRL model with worldwide data:

$$k_d^{\text{Every1}} = \frac{1}{1+e^{\beta_0+\beta_1 k_t+\beta_2 \text{AST}+\beta_3\alpha+\beta_4 k_{t,\text{daily}}+\beta_5\psi}}, \tag{15}$$

where $\beta_0 = -6.862$, $\beta_1 = 9.068$, $\beta_2 = 0.01468$, $\beta_3 = -0.00472$, $\beta_4 = 1.703$, and $\beta_5 = 1.084$. Then, climate-specific fitting was performed, and the resultant climate-dependent model coefficients are shown in Table 4. In this article, the "world version" is referred to as Every1, and the climate-specific version is referred to as Every2. It should be noted that the coefficients of Every2 do not cover all climate types, hence, in the validation below, the original BRL coefficients are used for climates that are not covered by Every2, see the last column of Table 4.

3. The benchmarking dataset

Validating separation models, or any other radiation model for that matter, demands high-quality, research-grade, ground-based radiometry data, which has hitherto been scarce. The most prominent and largest network of radiometry stations is the Baseline Surface Radiation Network[5] (BSRN) [63]. Whereas BSRN consists of about 60 active stations alone, there are other smaller networks owned by autonomous organizations such as the Bureau of Meteorology[6] (Australia), National Renewable Energy Laboratory (United States) [64], NOAA's SOLRAD Network[7] (United States) [65], University of Oregon Solar Radiation Monitoring Laboratory[8] (United States) [66], or Southern African Universities Radiometric Network,[9] Brooks et al. [67] of which the data are also invaluable to the research community. All above-mentioned datasets are in public domain, and can be accessed at the URLs indicated.

To that end, members from the IEA'a PVPS Task 16 Activity 1.4 have acquired and compiled several years of data from a total of 126 sites,[10] to facilitate solar energy meteorology research of various sorts [68]. Most of these stations use thermopile pyranometers for G_h and D_h measurement, and a tracker-mount thermopile pyrheliometer for B_n measurement, which comply to the recommended radiometry practice. In that, this article considers data from these 126 sites, over a 5-year period from 2016 to 2020. The geographical distribution of these sites, alongside their Köppen–Geiger climate classification, are depicted in Fig. 2.

To ensure that the best-possible data are used for comparison, a very stringent QC sequence was employed by the IEA members [68]. The entire QC sequence can be divided into four quantitative parts, each containing several filters (or tests). More specifically, if any of the below tests is *not* passed, the corresponding data point is flagged, and thus excluded from subsequent analysis. Otherwise, if a data point passes all tests, or when the tests could not be performed (i.e., when the condition of the test is not applicable), it is deemed "usable."

(1) Physically-possible limits tests

- $-4 \leq G_h \leq 1.5 E_{0n} \cos^{1.2} Z + 100$
- $-4 \leq D_h \leq 0.95 E_{0n} \cos^{1.2} Z + 50$
- $-4 \leq B_n \leq E_{0n}$

(2) Extremely-rare limits tests

- $-2 \leq G_h \leq 1.2 E_{0n} \cos^{1.2} Z + 50$
- $-2 \leq D_h \leq 0.75 E_{0n} \cos^{1.2} Z + 30$
- $-2 \leq B_n \leq 0.95 E_{0n} \cos^{0.2} Z + 10$

(3) Closure equation tests

- abs(closr) ≤ 8% for $Z \leq 75°$ and $G_h > 50$
- abs(closr) ≤ 15% for $93° > Z > 75°$ and $G_h > 50$

(4) k-index tests

- $k_n < k_t$ for $G_h > 50$, $k_t > 0$, and $k_n > 0$
- $k_n < (1100 + 0.03 \times h)/E_{0n}$ for $G_h > 50$ and $k_n > 0$
- $k_t < 1.35$ for $G_h > 50$ and $k_t > 0$

[5] https://bsrn.awi.de/.
[6] http://reg.bom.gov.au/climate/reg/oneminsolar/.
[7] https://gml.noaa.gov/grad/solrad/index.html.
[8] http://solardat.uoregon.edu/SelectArchivalUpdatedFormat.html.
[9] https://sauran.ac.za/.
[10] The number of sites considered initially was more numerous. However, after quality control, some of those initial sites have too few data points to be useful, and are therefore dropped.

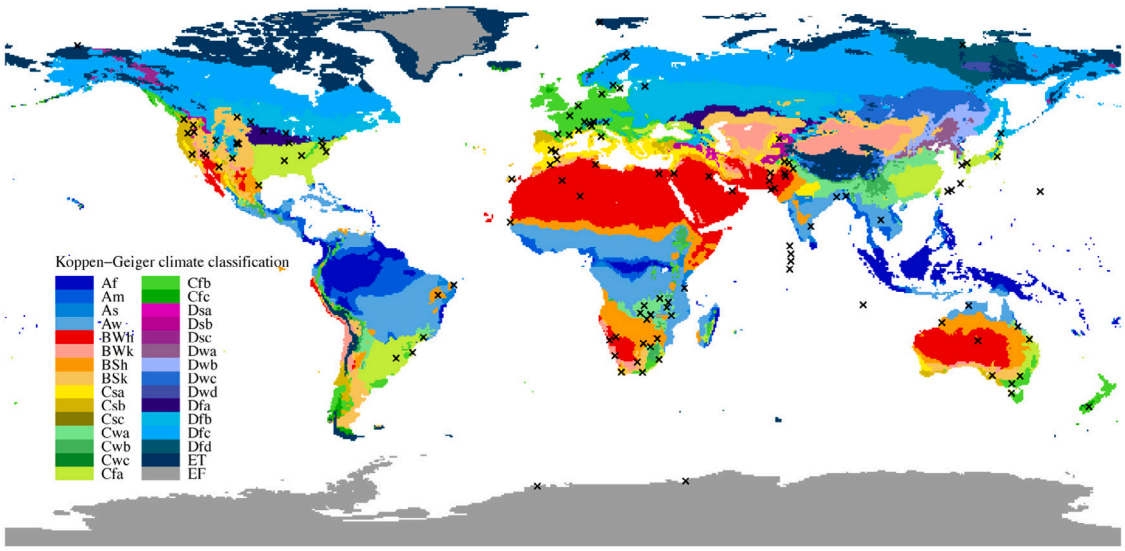


Fig. 2. The geographical distribution of 126 sites (black crosses) within the Köppen–Geiger climate classification system.

- $k_d < 1.05$ for $Z < 75°$ and $G_h > 50$
- $k_d < 1.10$ for $Z \geq 75°$ and $G_h > 50$
- $k_d < 0.96$ for $k_t > 0.6$, $Z < 85°$, $G_h > 150$, and $k_d > 0$

In the above QC sequence, all symbols have been introduced earlier, except for $E_{0n} = E_0/\cos Z$, which is the normal-incident extraterrestrial irradiance; closr $= G_h/\left(B_n \cos Z + D_h\right) - 1$, which is the decimal difference of the closure relationship; $k_n = B_n/E_{0n}$, which is the beam normal transmittance, a concept similar to clearness index, but for BNI; and h, which stands for the site's altitude in m.a.m.s.l. (meters above mean sea level). It should be noted that, for all numeric limits and additive constants in the above inequalities of radiation quantities, the unit is implied, in that it should follow that of the irradiance, which is W/m^2.

On top of the above QC routine, two more filters, namely, (1) $Z < 85°$, and (2) $G_h, B_n, D_h > 0$, are used to ensure the low-sun conditions are excluded from validation. This is because the errors in both the radiometry data and various separation models are large for low-sun conditions, but those instances are only of marginal importance in solar applications [10,32,69–71]. After pre-processing, there are more than 80 million valid data points remain at 126 sites. This scale is almost surely the largest collection of validation data points that the field has ever seen, surpassing the 25 million data points used by Gueymard and Ruiz-Arias [32].

4. Comparative assessment of 10 separation models at 126 sites worldwide

Based on the review in Sections 1 and 2, a total of 10 recent separation models are chosen for comparative assessment. They are: (1) Engerer2, (2) Engerer4, (3) Starke1, (4) Starke2, (5) Starke3, (6) Abreu, (7) Paulescu, (8) Every1, (9) Every2, and (10) Yang4. Except for Engerer2, all of these models are proposed after the year 2016, and are thought to be able to typify state-of-the-art separation models.[11] Each of the 10 models is applied to data from each site. Since separation modeling aims at predicting the beam and diffuse components, B_n and D_h are predictands (i.e., the quantities to be predicted), with which the performance evaluation is conducted.

When two or more prediction systems are present, the kind of assessment which one performs is known as comparative verification. The chief purpose of comparative verification is to find out which prediction system has the highest quality, and in what sense, such that the knowledge acquired through verification can be used to guide future decision-making. On this point, the Murphy–Winkler distribution-oriented framework for verifying deterministic (as opposed to probabilistic) predictions can be deemed most comprehensive, for its ability to assess different aspects of quality, such as association, calibration, resolution, or discrimination. This framework has been recently introduced to the solar community, and has been found useful for many applications [50,72–75].

That said, as one can see from those previous works, the Murphy–Winkler verification framework inspects the joint distribution of prediction and measurement, it is nevertheless often limited when the complexity and dimensionality of the verification task gets large [76]. The present comparative assessment involves 10 separation models, 2 predictands (B_n and D_h), 126 cases (i.e., datasets or sites), of which the data points can be further grouped according to climate, weather, and sky conditions. Clearly, reporting the results of Murphy–Winkler verification would require, most probably, hundreds of pages. To that end, more efficient alternative verification procedures need to be sought.

4.1. Assessment based on ranking statistics

When the subjects to be assessed are many and the samples are ample, one natural strategy is to examine the ranking data. In that, each subject is a model, and each sample refers to a case, that is, the performance ranking of different models at a particular site. By collecting such ranking data and performing statistical analysis, the overall performance of models can be summarized. Therefore, in the first part of the assessment, a linear ranking method [77] is considered. This method has been previously used to rank emerging technologies in the field of solar forecasting [78].

The linear ranking method is interested in knowing the *mean rank* of different models. Since there are a total of 10 models, the mean rank is expressed as $\boldsymbol{m} = (m_1, m_2, \ldots, m_{10})^{\mathrm{T}}$, and the rank for the ith model is expressed as:

$$m_i = \sum_{j=1}^{10!} \frac{n_j v_j(i)}{n}, \tag{16}$$

[11] The older models such as Erbs [52], Orgill [53], or Perez1 [60], are not included since they have been shown, using extensive and high-quality 1-min data worldwide, to be weaker than Engerer2 [32,51]. Yang1, Yang2, and Yang3 are not included because they are not as accessible or not as accurate as Yang4. Other models, such as those discussed in [37–42,44–48] are also ignored for their obvious deficiencies, as argued in the introduction.

Table 5
Ranking results of 10 separation models, based on the root mean square error of B_n estimates, at 126 sites. For each site, best model is ranked "1," and the worst model is ranked "10." Whereas the middle columns are omitted, the last column shows the mean rank of each model, the smaller the better.

Model	Site					Mean rank
	1	2	3	⋯	126	
Engerer2	4	1	6	⋯	6	4.32
Engerer4	7	6	7	⋯	8	6.68
Starke1	3	5	2	⋯	3	3.00
Starke2	5	8	3	⋯	2	5.04
Starke3	2	4	4	⋯	1	3.01
Abreu	8	7	8	⋯	7	7.87
Paulescu	6	2	5	⋯	5	4.67
Every1	10	9	9	⋯	9	9.10
Every2	9	10	10	⋯	10	8.84
Yang4	1	3	1	⋯	4	2.47

Table 6
Same as Table 5, but based on the RMSE of D_h estimates.

Model	Site					Mean rank
	1	2	3	⋯	126	
Engerer2	4	1	4	⋯	6	4.81
Engerer4	7	6	7	⋯	8	7.01
Starke1	3	5	2	⋯	4	3.47
Starke2	5	7	6	⋯	3	5.04
Starke3	2	2	3	⋯	1	2.55
Abreu	8	8	8	⋯	7	7.68
Paulescu	6	3	5	⋯	5	4.06
Every1	10	9	9	⋯	9	9.08
Every2	9	10	10	⋯	10	8.86
Yang4	1	4	1	⋯	2	2.45

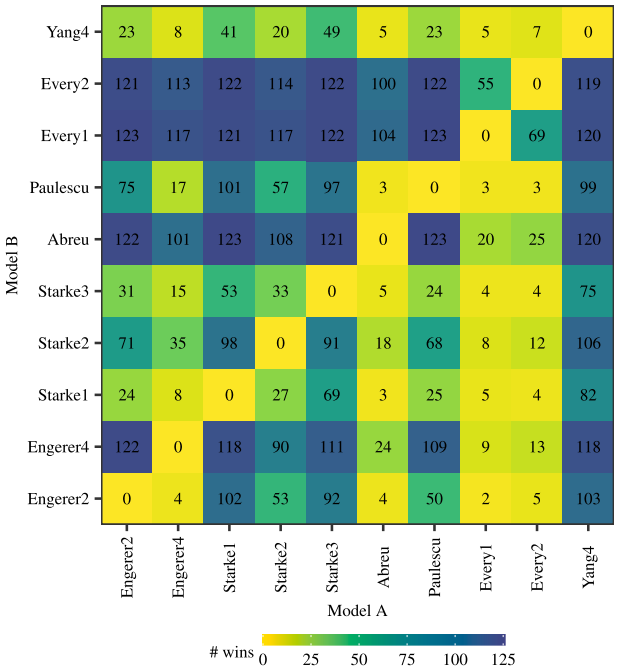


Fig. 3. Pairwise Diebold–Mariano (DM) tests for comparing the predictive accuracy of various separation models, in terms of B_n. The numbers show the number of instances the DM test statistics falls in the lower or upper 2.5% tail of a standard normal distribution. In other words, the entries denote the number of "Model A is better than Model B" instances. For example, in the lower-right corner, Yang4 performs significantly better than Engerer2 at 103 out of 126 sites.

where v_j with $j = 1, 2, \ldots, 10!$ represents all possible rankings (i.e., the permutation) of the 10 models; n_j is the frequency of occurrence of ranking j; $n = \sum_{j=1}^{10!} n_j$ is the number of samples; and $v_j(i)$ denotes the score of model i in ranking j. In this work, a negatively oriented ranking is used, which suggests that a better model would receive a smaller $v_j(i)$. Stated differently, if model i ranks the highest in ranking j, $v_j(i) = 1$; if it ranks the lowest, then $v_j(i) = 10$.

In order to perform the ranking, some criterion is needed. This criterion can be either based on objective metrics and measures, or from experts' judgment. In this article, the normalized root mean square error (nRMSE), in percent, is taken as the criterion—the smaller the nRMSE of a model is, as compared to its peers, the higher the ranking of that model is. For N prediction–measurement pairs, the nRMSE of B_n is given by:

$$\mathrm{nRMSE}\left(\hat{B}_n, B_n\right) = \frac{N}{\sum_{k=1}^{N} B_{n,k}} \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(\hat{B}_{n,k} - B_{n,k}\right)^2} \times 100, \tag{17}$$

where $\hat{B}_{n,k}$ and $B_{n,k}$ are the kth predicted and measured beam normal irradiance. The nRMSE of D_h can be evaluated in the same fashion, after replacing $\hat{B}_{n,k}$ and $B_{n,k}$ with $\hat{D}_{h,k}$ and $D_{h,k}$, respectively. (The full nRMSE tables of B_n and D_h are given in Appendix.)

Table 5 shows the ranking results of separation models in terms of the nRMSE of their B_n predictions. Each column in the table corresponds to the ranking at a particular site—due to space constraint, just a few columns are printed, with the rest omitted—and the mean rank is calculated through Eq. (16). For example, Yang4 ranks first at Site 1, and Every1 ranks last at that site. It can be seen from Table 5 that, for B_n prediction, Yang4 has the best mean rank of 2.47, followed by Starke1 (3.00), Starke3 (3.01), Engerer2 (4.32), so on and so forth. Similarly, the results for D_h predictions are presented in Table 6. And Yang4 has again attained the highest mean rank. Since the ranking statistics is but one approach for assessing the quality of these separation models, the interpretation of the results is postponed to the end of this section.

4.2. Assessment based on pairwise Diebold–Mariano test

The first approach of comparative assessment presents an overview of model performance through ranking statistics. In that regard, if Model $\mathcal{A}$ has a higher mean rank than Model $\mathcal{B}$, $\mathcal{A}$ is said to be better than $\mathcal{B}$. Nevertheless, for a pair of models $\mathcal{A}$ and $\mathcal{B}$, there may be situations where $\mathcal{A}$ outperforms $\mathcal{B}$ just 51% of cases, whereas $\mathcal{B}$ outperforms $\mathcal{A}$ in the remaining cases. It is, then, not quite right to adjudicate that $\mathcal{A}$ is better than $\mathcal{B}$—the judgment needs to be more "fuzzy." In such situations, one ought to take a closer look at the statistical significance of any claimed superiority, e.g., through hypothesis testing. Therefore, the second approach of assessment is based on the pairwise Diebold–Mariano (DM) test [79], which is a statistical test for comparing predictive accuracy of two models. Owing to the generality of its intended purpose, the DM test has been used, in this field, to compare forecasting models [80–82], post-processing models [83], and transposition models [10].

The null hypothesis of the DM test is that the expectation of loss differential is zero. The notion of "loss" is general, for one can opt from many scoring functions, such as absolute error or squared error. Defining the error for the kth B_n prediction from Model $\mathcal{A}$ as $e_k^{\mathcal{A}} = \hat{B}_{n,k}^{\mathcal{A}} - B_{n,k}^{\mathcal{A}}$, and that from Model $\mathcal{B}$ as $e_k^{\mathcal{B}} = \hat{B}_{n,k}^{\mathcal{B}} - B_{n,k}^{\mathcal{B}}$, the loss differential is:

$$d = g\left(e_k^{\mathcal{A}}\right) - g\left(e_k^{\mathcal{B}}\right), \tag{18}$$

where $g(\cdot)$ is the scoring function of choice, e.g., when absolute error is used, $g(x) = |x|$, or when squared error is used, $g(x) = x^2$. With such a preamble, the null hypothesis can be expressed as $H_0 : \mathbb{E}(d) = 0$. Since we are interested in knowing whether one model is significantly better than the other, or vise versa, a two-sided alternative $H_1 : \mathbb{E}(d) \neq 0$ is appropriate. Based on any two sets of predictions, the test statistic can be computed. This test statistic is then compared to the critical value—in this case, for two-sided alternative with a 95% confidence, the critical values are ± 1.96, which correspond to the 0.025 and 0.975

Table A.7
nRMSE [%] of beam normal irradiance, B_n, for sites in tropical climates (Köppen–Geiger classification: A). $\mathbb{E}(B_n)$ is the mean B_n [W/m^2] of the validation data. Row-wise best result is in bold.

stn	kgc	$\mathbb{E}(B_n)$	Engerer2	Engerer4	Starke1	Starke2	Starke3	Abreu	Paulescu	Every1	Every2	Yang4
A1	Aw	389.24	26.63	30.53	24.06	25.41	**23.15**	30.19	24.94	37.91	36.74	24.34
A2	Aw	513.92	17.14	19.08	16.16	**15.11**	15.20	20.38	18.37	22.61	21.21	15.72
A3	Aw	591.16	15.32	16.80	14.27	13.62	**13.54**	18.32	16.42	19.92	18.60	13.73
A4	Am	286.06	34.82	40.07	27.73	32.47	**26.18**	39.45	26.80	39.15	37.08	32.96
A5	Aw	253.99	33.43	36.44	31.30	**29.12**	30.38	36.97	33.28	36.07	33.19	31.86
A6	As	471.60	23.78	25.20	22.73	**20.99**	21.91	24.21	23.46	28.46	25.63	22.69
A7	Aw	355.06	45.55	31.92	28.86	30.23	29.06	45.37	**26.87**	47.56	46.93	39.39
A8	Aw	258.37	49.46	52.66	39.83	48.83	37.92	53.85	**36.02**	53.78	55.53	44.67
A9	Aw	585.18	15.12	17.84	14.78	15.40	13.95	18.87	16.90	21.35	20.98	**12.89**

Table A.8
Same as Table A.7, but for sites in dry climates (Köppen–Geiger classification: B).

stn	kgc	$\mathbb{E}(B_n)$	Engerer2	Engerer4	Starke1	Starke2	Starke3	Abreu	Paulescu	Every1	Every2	Yang4
B1	BWh	547.35	**25.72**	29.16	28.35	33.45	27.84	31.88	27.01	34.12	34.38	27.71
B2	BSk	677.83	15.56	16.73	14.14	15.09	15.40	18.30	15.49	20.91	21.47	**13.85**
B3	BWh	742.45	11.38	13.33	10.46	10.90	10.17	13.51	12.26	13.70	13.93	**8.97**
B4	BWh	732.63	12.53	14.47	10.89	10.95	10.37	14.74	12.53	14.82	15.08	**10.26**
B5	BSh	670.47	12.28	14.15	11.92	14.21	13.04	15.27	13.35	18.46	19.11	**10.99**
B6	BWh	400.04	24.64	32.82	26.92	33.99	22.07	33.47	**20.56**	29.51	29.89	29.22
B7	BSk	636.08	15.61	17.25	15.14	17.76	16.82	19.12	16.95	24.31	24.70	**14.67**
B8	BSh	635.70	14.17	15.84	12.87	11.67	13.16	16.13	14.75	16.18	15.80	**11.59**
B9	BSk	541.52	24.23	25.51	20.93	**20.54**	21.95	27.80	23.76	28.48	29.32	22.81
B10	BWh	517.82	20.29	22.00	19.67	24.39	19.98	25.36	**18.49**	27.45	27.88	20.85
B11	BSk	766.48	13.24	14.93	11.64	12.23	11.73	16.01	12.72	15.91	16.34	**11.45**
B12	BWk	738.25	12.29	12.67	**10.50**	12.55	11.42	13.81	11.82	16.09	14.28	10.87
B13	BSk	567.81	18.20	19.27	17.84	20.51	19.08	21.20	19.38	24.94	25.39	**17.28**
B14	BSh	629.87	14.16	16.06	12.34	12.18	**12.13**	16.96	13.92	18.02	17.92	12.62
B15	BWh	761.52	9.84	10.83	9.62	11.79	9.84	11.98	10.72	13.06	13.31	**8.98**
B16	BSk	531.56	17.54	19.26	**16.29**	17.83	17.55	21.28	19.67	25.27	25.88	16.60
B17	BSk	591.59	15.90	16.16	13.58	**13.51**	14.32	18.97	15.56	19.33	20.27	14.47
B18	BSk	538.69	23.45	24.39	20.88	**18.98**	21.45	26.52	24.08	26.41	27.51	22.56
B19	BWh	456.82	23.48	26.88	25.01	29.80	**18.82**	27.37	19.69	24.55	24.85	27.19
B20	BWh	679.52	**13.48**	14.67	14.61	18.10	15.48	16.88	14.15	19.82	20.14	14.18
B21	BWh	433.63	24.37	27.32	23.81	28.83	**18.72**	29.08	20.06	25.78	26.09	25.73
B22	BWh	645.82	**16.29**	18.13	18.29	22.52	18.24	20.02	17.35	22.56	22.73	17.31
B23	BSh	346.58	26.61	28.88	26.98	25.40	**25.17**	28.10	27.29	26.59	25.71	27.23
B24	BWh	711.76	13.07	13.89	12.79	14.99	13.75	14.95	13.73	17.93	18.15	**12.25**
B25	BWh	580.24	**18.21**	19.77	21.21	26.55	20.12	23.63	19.95	24.64	25.01	22.31
B26	BSh	320.51	35.32	43.79	36.22	43.72	**31.71**	44.81	31.84	42.47	45.40	36.17
B27	BWk	618.77	16.11	18.53	16.72	20.04	17.35	20.53	16.68	24.19	20.90	**15.93**
B28	BWh	370.02	23.77	28.04	24.55	26.85	**20.79**	26.95	22.32	24.30	24.92	26.31
B29	BSk	553.68	17.66	20.46	16.89	19.86	**16.54**	22.57	17.05	23.94	24.37	17.33
B30	BSh	384.52	25.96	29.08	24.67	23.68	23.92	30.76	25.60	33.16	33.21	**23.28**
B31	BSh	409.02	28.64	28.31	30.77	**24.74**	28.94	27.13	31.24	28.36	26.27	31.29
B32	BSh	499.55	20.20	22.38	18.94	18.19	18.40	23.45	20.36	27.04	27.01	**17.90**
B33	BSh	491.78	22.52	24.70	22.75	20.25	21.11	25.19	24.92	25.87	25.43	**19.43**
B34	BSk	634.59	**17.41**	19.73	20.19	24.85	20.16	21.23	19.10	24.19	23.89	19.00
B35	BSk	550.05	18.50	19.78	16.61	17.97	18.00	21.94	18.43	24.90	25.52	**16.53**
B36	BWh	646.45	**12.32**	12.93	15.07	19.53	14.47	15.52	12.68	17.15	17.49	15.53
B37	BSk	573.53	17.43	19.10	17.31	20.53	19.15	21.75	18.98	26.83	27.23	**16.26**
B38	BSk	524.54	17.51	19.42	**16.35**	17.92	17.77	21.45	19.85	25.48	26.10	16.72
B39	BWh	574.93	27.89	28.85	27.58	31.76	29.93	32.11	**26.09**	35.64	36.03	28.21
B40	BWh	558.76	17.47	20.81	17.53	21.20	17.29	22.48	**17.25**	24.22	24.51	17.60
B41	BWh	743.12	10.96	13.50	9.00	8.99	**8.84**	13.50	10.74	14.55	14.85	9.09

quantiles of a standard normal distribution. If the test statistic is smaller than −1.96 or greater than 1.96, the null hypothesis of zero loss differential is rejected, which indicates that one set of predictions is significantly better than the other.

Since there are 10 separation models and 126 cases, DM test is conducted for $C_2^{10} \times 126 = 5670$ times. The result for DM test on B_n prediction is depicted in Fig. 3. In this figure, the entries denote the number of times when Model $\mathcal{A}$ is better than Model $\mathcal{B}$. For example, the bottom-right entry says Yang4 is better than Engerer2 at 103 out of 126 sites; the top-left entry says Engerer2 is better than Yang4 at 23 out of 126 sites. Stated differently, the column sums suggest the total number of "winning" cases of each model, whereas the row sums suggest the total number of "losing" cases of each model. It can then be concluded that Yang4 has the best performance with a total of 942 wins, followed by Starke1 (879 wins), Starke3 (874 wins), Engerer2 (712 wins), so on and so forth. This result is consistent with earlier finding based on ranking statistics. Fig. 4, on the other hand, depicts the DM test results for D_h predictions, and Yang4 is again found to be the leading model with 945 wins.

4.3. Visual assessment based on boxplot

Through ranking statistics and DM tests, the relative predictive performance of the 10 models has been investigated. Nevertheless, it is still of interest to examine how much difference there really is, using accuracy measures that are more familiar to solar engineers, namely, normalized mean bias error (nMBE), normalized mean absolute error (nMAE) and nRMSE. Since there are too many numbers to be reported, in the form of tables (see Appendix for instance), Tukey's boxplot is used for visualization, as shown in Fig. 5. Each boxplot in the figure

Table A.9
Same as Table A.7, but for sites in temperate climates (Köppen–Geiger classification: C).

stn	kgc	$\mathbb{E}(B_n)$	Engerer2	Engerer4	Starke1	Starke2	Starke3	Abreu	Paulescu	Every1	Every2	Yang4
C1	Csb	560.20	16.37	19.09	16.31	17.47	16.09	21.32	18.92	22.99	22.99	**13.74**
C2	Cfa	559.80	17.52	19.60	16.95	20.17	16.77	20.92	18.32	25.01	24.86	**15.97**
C3	Cfb	306.72	29.49	34.89	25.75	29.98	**25.53**	34.83	28.67	42.36	43.39	25.53
C4	Cfa	428.01	21.66	24.93	19.37	20.13	19.38	25.77	21.18	30.21	30.17	**18.58**
C5	Cfb	356.98	29.76	34.47	27.30	30.40	26.05	34.33	31.52	41.72	42.07	**25.19**
C6	Csa	610.18	15.06	17.24	13.50	13.86	14.09	20.30	16.22	19.37	20.22	**13.10**
C7	Csa	511.84	17.23	20.06	15.62	16.10	16.39	23.05	18.55	23.85	24.85	**15.29**
C8	Cfb	552.58	17.25	19.81	16.00	17.93	15.88	21.96	18.14	24.40	24.55	**15.14**
C9	Cfa	333.31	31.61	38.58	31.04	37.74	31.73	34.75	30.64	46.93	46.96	**29.69**
C10	Cwa	449.39	22.63	25.73	21.86	**21.06**	23.51	28.31	24.37	33.64	29.43	21.45
C11	Cwa	624.42	17.28	19.00	16.11	15.39	16.99	20.68	17.00	22.79	20.45	**15.09**
C12	Cwa	664.07	15.44	16.98	13.97	14.01	14.25	18.58	14.45	20.74	18.53	**13.26**
C13	Cfa	508.70	19.47	21.75	18.47	16.91	19.61	24.32	20.75	24.15	23.91	**16.45**
C14	Cfa	483.44	20.50	23.50	18.51	20.14	18.08	25.20	20.98	28.12	28.15	**17.45**
C15	Cfa	396.54	23.45	26.94	21.68	25.38	**20.46**	26.52	22.36	30.44	30.52	21.93
C16	Cfa	558.33	15.79	18.10	14.41	15.68	14.16	20.27	16.16	22.06	21.87	**13.85**
C17	Cwa	413.02	23.40	26.00	20.78	24.18	20.16	25.85	**20.07**	26.88	22.45	23.46
C18	Csb	881.75	11.37	11.52	11.26	13.25	11.23	22.35	11.93	13.90	13.54	**10.49**
C19	Csa	538.01	17.05	20.00	16.68	20.29	16.33	19.85	16.76	24.98	25.06	**16.30**
C20	Cwa	627.46	16.34	17.90	15.54	15.80	16.38	19.31	16.41	22.44	20.28	**14.15**
C21	Cwa	650.25	15.08	16.55	14.17	14.52	14.77	18.28	15.41	20.48	18.61	**12.94**
C22	Cwa	567.07	20.10	22.48	19.92	21.65	19.86	25.68	22.14	30.51	26.96	**18.58**
C23	Cfa	420.38	20.59	23.42	17.59	20.33	**16.81**	24.01	19.19	27.58	27.60	17.93
C24	Cfa	499.11	18.13	20.17	16.11	18.66	**15.76**	21.80	17.81	23.83	23.76	15.90
C25	Cfb	640.52	16.25	18.16	15.62	17.18	15.37	21.41	18.44	24.46	24.46	**14.28**
C26	Cfb	576.29	16.70	18.86	**15.36**	18.18	15.57	20.74	17.04	23.57	23.60	16.08
C27	Cwa	565.51	16.28	18.25	**14.68**	15.22	15.14	20.01	15.11	23.98	20.12	15.32
C28	Cfb	284.98	27.75	32.48	27.58	25.81	27.53	36.79	33.68	38.86	39.78	**24.55**
C29	Cfb	434.90	20.67	24.16	20.52	20.80	19.73	26.57	23.74	30.83	31.19	**17.46**
C30	Cfa	447.80	22.57	25.11	20.32	24.14	**20.08**	25.90	21.40	29.05	29.00	21.44
C31	Cwa	612.54	16.78	17.98	16.11	15.58	17.68	20.07	17.53	22.25	20.40	**14.72**
C32	Cwa	545.85	21.57	23.53	19.90	21.07	19.38	25.79	21.10	30.84	27.11	**19.13**
C33	Cfb	345.55	25.01	28.85	22.68	25.32	22.87	30.53	26.70	36.27	37.27	**22.28**
C34	Cfa	518.98	29.97	29.52	29.32	**23.47**	33.51	29.89	29.49	28.62	27.68	29.94
C35	Cfb	512.61	21.22	23.58	19.17	**18.23**	18.85	27.59	23.30	25.81	26.17	19.22
C36	Cfb	403.86	25.34	29.04	22.48	26.63	**21.98**	29.67	24.07	36.32	36.78	22.17
C37	Csa	541.27	18.29	19.82	**15.77**	17.67	15.93	21.92	17.17	22.98	23.63	16.53
C38	Cfb	630.96	15.10	17.12	15.48	18.91	15.69	18.50	16.81	22.54	22.39	**14.48**
C39	Cwa	616.27	14.34	16.22	**12.21**	12.39	12.44	19.14	13.84	19.17	16.65	12.50
C40	Cfa	567.67	16.05	18.22	15.23	14.63	15.03	20.70	17.46	22.06	21.82	**13.46**
C41	Cfb	519.48	18.93	21.44	17.67	20.60	17.24	22.70	19.31	27.58	27.51	**16.97**
C42	Cfa	453.02	20.40	23.31	18.07	20.50	**17.02**	25.14	20.06	29.71	29.60	17.26
C43	Csb	349.61	27.16	29.28	26.07	23.67	25.41	32.48	29.54	33.75	34.23	**23.52**
C44	Csb	356.50	31.66	32.73	30.40	**26.51**	29.49	38.18	35.53	37.49	37.91	29.63
C45	Csb	690.35	17.00	18.22	17.29	20.26	17.58	21.28	19.00	25.34	25.24	**15.37**
C46	Csb	671.29	14.96	16.97	13.42	12.39	14.14	20.25	15.62	18.59	19.12	**12.32**
C47	Cfa	387.69	24.58	26.66	22.35	23.13	**22.21**	31.01	25.79	33.69	33.36	22.45
C48	Csa	575.27	14.82	17.28	**13.62**	15.79	13.62	19.75	15.21	20.98	21.49	14.05
C49	Cfb	391.88	21.69	24.91	20.08	23.21	20.13	25.70	23.14	30.23	31.06	**19.38**
C50	Cwa	449.49	22.54	24.27	21.66	**19.17**	22.89	26.16	22.94	27.01	24.66	21.54
C51	Cfa	607.92	14.86	17.21	14.16	13.97	13.96	20.15	16.41	19.70	19.55	**12.03**

Table A.10
Same as Table A.7, but for sites in continental climates (Köppen–Geiger classification: D).

stn	kgc	$\mathbb{E}(B_n)$	Engerer2	Engerer4	Starke1	Starke2	Starke3	Abreu	Paulescu	Every1	Every2	Yang4
D1	Dfb	437.90	22.91	24.15	20.95	22.39	**20.69**	27.13	23.71	29.76	27.29	20.75
D2	Dfb	365.15	26.51	30.02	24.32	28.05	**23.64**	32.48	26.67	39.54	35.68	23.77
D3	Dfc	290.12	32.35	36.84	31.09	35.92	30.58	38.74	33.84	48.91	45.91	**28.99**
D4	Dfb	380.15	24.86	26.69	22.62	23.40	23.44	30.31	26.02	32.17	29.08	**22.04**
D5	Dfb	377.91	28.64	32.80	26.55	31.01	26.76	32.00	28.13	39.82	36.55	**26.25**
D6	Dfa	559.59	17.52	18.86	17.11	20.23	16.64	20.64	18.51	24.36	22.12	**16.01**
D7	Dsd	613.04	**16.16**	17.02	17.81	21.08	18.60	19.55	18.64	22.30	20.15	16.28
D8	Dfb	467.90	20.47	23.06	18.76	20.54	19.99	25.38	21.76	29.72	26.93	**18.65**

Table A.11
Same as Table A.7, but for sites in polar climates (Köppen–Geiger classification: E).

stn	kgc	$\mathbb{E}(B_n)$	Engerer2	Engerer4	Starke1	Starke2	Starke3	Abreu	Paulescu	Every1	Every2	Yang4
E1	ET	243.88	55.82	58.26	54.81	62.52	**45.24**	51.43	56.77	70.88	67.63	53.97
E2	ET	374.65	24.98	27.63	24.58	28.61	53.51	39.13	25.95	38.98	35.80	**23.27**
E3	ET	547.99	17.66	19.43	18.58	22.07	39.24	31.28	19.62	26.71	24.59	**16.59**
E4	ET	532.76	20.24	22.94	22.29	26.47	23.67	29.30	24.36	31.64	29.76	**19.50**
E5	EF	709.80	22.34	**22.08**	24.05	27.21	22.98	28.19	25.97	31.29	28.87	22.48

Table A.12
Same as Table A.7, but for sites on islands.

stn	kgc	$\mathbb{E}(B_n)$	Engerer2	Engerer4	Starke1	Starke2	Starke3	Abreu	Paulescu	Every1	Every2	Yang4
O1	Ocean	318.46	29.04	34.70	27.24	29.00	27.04	33.72	28.75	40.38	35.93	**25.99**
O2	Ocean	530.60	18.79	22.05	18.05	20.42	18.46	21.73	19.26	26.83	23.50	**16.76**
O3	Ocean	468.48	21.00	24.69	19.62	21.90	20.15	24.18	20.92	29.41	25.94	**18.51**
O4	Ocean	267.37	36.22	42.85	32.36	38.14	**31.17**	39.14	32.08	50.01	43.43	32.66
O5	Ocean	456.72	21.28	24.30	**21.03**	23.46	22.36	25.60	21.89	30.25	26.08	21.10
O6	Ocean	446.77	23.09	25.66	22.16	27.96	20.49	24.59	**20.02**	29.40	23.51	26.35
O7	Ocean	365.89	28.44	32.94	26.08	28.48	**24.85**	32.52	27.05	37.52	33.80	25.02
O8	Ocean	469.87	22.40	25.18	22.79	27.56	**21.83**	24.53	22.09	31.33	26.21	23.84
O9	Ocean	516.03	19.02	21.68	18.33	19.40	18.25	22.20	20.28	25.41	23.71	**16.64**
O10	Ocean	346.57	31.80	37.80	30.06	33.61	**26.91**	36.01	29.70	46.09	40.44	27.49
O11	Ocean	427.04	**21.62**	25.85	22.86	28.16	22.90	24.90	21.85	30.69	25.62	23.85
O12	Ocean	672.66	**26.45**	27.60	30.98	34.39	33.16	27.04	33.22	38.95	36.34	27.29

Table A.13
nRMSE [%] of diffuse horizontal irradiance, D_h, for sites in tropical climates (Köppen–Geiger classification: A). $\mathbb{E}(D_h)$ is the mean D_h [W/m^2] of the validation data. Row-wise best result is in bold.

stn	kgc	$\mathbb{E}(D_h)$	Engerer2	Engerer4	Starke1	Starke2	Starke3	Abreu	Paulescu	Every1	Every2	Yang4
A1	Aw	221.74	32.71	36.70	29.02	29.68	**26.66**	34.45	27.76	46.10	44.82	30.36
A2	Aw	173.08	35.50	39.29	33.19	30.49	**28.95**	37.99	34.25	45.98	44.39	31.47
A3	Aw	158.54	38.71	42.53	36.12	34.51	**31.59**	40.89	37.10	49.53	48.13	32.94
A4	Am	217.50	32.18	36.20	25.67	28.51	**22.07**	36.22	23.31	35.80	33.56	30.16
A5	Aw	227.60	25.84	27.63	25.47	**22.34**	24.33	27.35	25.12	26.62	24.66	24.25
A6	As	176.55	46.82	48.24	44.31	**39.98**	41.61	44.04	43.77	53.36	48.55	42.29
A7	Aw	235.78	45.13	35.40	30.60	31.83	27.62	50.62	**27.08**	51.72	51.61	37.98
A8	Aw	262.33	30.73	32.26	24.92	29.17	22.40	34.00	**21.74**	33.85	34.79	27.63
A9	Aw	161.13	38.87	45.07	38.01	38.90	34.37	45.10	41.14	53.40	53.60	**31.95**

Table A.14
Same as Table A.13, but for sites in dry climates (Köppen–Geiger classification: B).

stn	kgc	$\mathbb{E}(D_h)$	Engerer2	Engerer4	Starke1	Starke2	Starke3	Abreu	Paulescu	Every1	Every2	Yang4
B1	BWh	204.53	**48.04**	53.74	52.24	59.58	49.90	61.68	49.92	65.10	65.99	51.85
B2	BSk	132.90	42.06	47.93	39.46	44.46	41.16	55.62	43.49	60.03	60.56	**35.53**
B3	BWh	110.81	49.19	59.33	45.98	44.95	42.80	53.25	50.78	55.90	56.36	**37.02**
B4	BWh	115.60	54.42	63.99	49.22	46.78	44.12	61.17	53.93	60.34	60.95	**43.54**
B5	BSh	131.19	38.58	44.95	36.62	41.56	37.88	46.93	40.28	56.74	58.61	**33.59**
B6	BWh	216.73	29.06	39.01	32.52	38.50	23.59	42.90	**23.54**	36.32	37.05	34.43
B7	BSk	138.62	40.46	44.77	39.21	45.31	41.74	50.00	42.33	62.95	63.58	**36.98**
B8	BSh	135.95	44.57	50.58	41.57	36.25	41.47	45.13	44.01	47.95	47.92	**34.26**
B9	BSk	135.17	44.93	48.82	39.61	43.28	41.19	55.44	43.63	60.68	61.42	**39.55**
B10	BWh	174.90	36.92	39.30	34.74	41.73	34.07	45.11	**31.89**	49.46	50.40	37.60
B11	BSk	107.85	56.44	66.51	47.36	47.14	**42.92**	73.18	52.74	63.61	64.58	45.57
B12	BWk	111.93	44.13	49.49	39.46	46.00	41.41	52.30	43.37	61.16	53.40	**37.24**
B13	BSk	132.77	40.13	43.49	38.59	43.49	39.30	45.83	41.17	55.26	55.78	**36.21**
B14	BSh	146.45	39.92	45.58	36.23	36.32	**33.00**	44.69	37.73	52.20	53.31	33.65
B15	BWh	113.74	42.64	51.40	40.12	45.35	39.74	50.21	44.51	54.23	55.22	**36.04**
B16	BSk	140.81	37.98	41.93	**34.76**	37.22	36.45	46.65	42.17	54.69	55.76	35.42
B17	BSk	134.71	36.44	39.65	31.05	33.10	**29.67**	39.98	32.09	43.79	44.57	30.58
B18	BSk	133.54	45.67	48.81	41.11	42.36	**40.53**	51.22	45.41	55.62	56.46	41.25
B19	BWh	215.36	35.67	39.74	38.37	43.81	**26.01**	43.21	28.77	37.37	37.97	42.16
B20	BWh	133.09	41.34	46.74	42.69	51.53	44.19	52.49	**41.24**	59.92	61.10	42.52
B21	BWh	221.54	34.09	37.58	33.87	39.16	**23.69**	42.13	27.04	36.09	36.61	36.15
B22	BWh	166.21	**41.33**	44.25	45.24	54.03	43.17	52.60	42.30	58.75	59.71	45.01
B23	BSh	213.51	29.25	31.22	29.38	27.37	28.07	30.01	28.39	27.71	**27.24**	29.22
B24	BWh	121.21	43.62	48.34	43.77	50.58	46.09	49.29	45.77	62.07	62.98	**39.92**
B25	BWh	176.57	**31.67**	34.88	38.28	46.78	32.84	44.23	35.34	45.64	46.77	39.62
B26	BSh	249.29	33.34	39.26	34.28	39.13	**28.43**	43.24	29.13	40.38	42.94	34.51
B27	BWk	148.23	42.13	47.60	42.38	49.04	42.07	53.70	**41.21**	62.61	52.11	41.76
B28	BWh	217.96	27.66	32.20	29.38	31.00	**23.70**	31.40	24.54	27.10	27.64	30.92
B29	BSk	156.52	38.97	44.57	36.68	42.46	**33.56**	49.00	35.22	52.23	52.40	38.41
B30	BSh	183.25	37.93	41.36	35.61	**33.23**	33.48	42.19	33.61	46.81	47.24	33.70
B31	BSh	178.94	40.72	39.17	43.77	**35.06**	42.85	35.87	43.70	37.99	35.37	42.16
B32	BSh	193.77	37.06	40.40	33.38	31.84	31.35	39.85	33.21	47.93	48.42	**30.95**
B33	BSh	153.98	46.46	50.33	47.89	42.79	42.84	48.27	48.49	50.65	50.86	**38.10**
B34	BSk	169.05	**42.66**	46.12	46.96	56.42	45.80	53.84	45.18	60.42	59.63	47.53
B35	BSk	147.31	36.91	41.21	33.47	38.40	35.36	46.47	36.63	52.79	53.42	**31.81**
B36	BWh	157.67	**32.28**	33.63	41.46	51.05	39.00	45.11	35.85	50.95	52.36	41.95
B37	BSk	147.01	40.68	44.21	38.25	44.15	40.82	51.47	43.09	59.78	60.23	**36.09**
B38	BSk	142.23	37.52	41.43	**34.30**	36.68	36.24	46.21	41.99	54.01	55.11	35.35
B39	BWh	172.56	54.15	56.36	53.53	61.32	56.47	66.61	**49.40**	71.07	71.98	55.17
B40	BWh	162.43	38.39	45.70	38.06	45.13	**36.55**	49.76	36.58	53.81	54.61	39.01
B41	BWh	113.33	45.34	55.62	40.82	39.22	**36.07**	54.59	44.37	63.16	64.13	38.19

Table A.15
Same as Table A.13, but for sites in temperate climates (Köppen–Geiger classification: C).

stn	kgc	$\mathbb{E}(D_h)$	Engerer2	Engerer4	Starke1	Starke2	Starke3	Abreu	Paulescu	Every1	Every2	Yang4
C1	Csb	151.07	39.49	45.26	39.26	40.23	38.07	49.54	44.19	51.86	50.79	**31.86**
C2	Cfa	160.72	37.32	41.67	35.12	41.27	**32.05**	42.19	36.38	53.49	53.48	32.97
C3	Cfb	155.15	33.21	38.20	27.68	32.64	**25.39**	36.72	29.18	47.48	47.71	27.89
C4	Cfa	192.27	35.33	39.64	32.03	30.34	32.97	40.76	34.90	45.72	45.86	**28.94**
C5	Cfb	186.34	37.09	41.14	33.70	36.58	31.02	42.02	39.22	48.92	47.97	**31.00**
C6	Csa	124.67	40.84	46.78	36.84	38.28	36.59	48.75	40.16	50.37	50.87	**33.77**
C7	Csa	134.74	37.30	43.28	32.75	32.71	33.70	44.06	35.35	48.51	49.63	**31.23**
C8	Cfb	147.24	38.01	43.44	35.26	39.86	32.96	45.36	37.99	54.09	53.22	**32.47**
C9	Cfa	206.14	33.87	39.75	30.91	36.90	**29.02**	36.54	30.02	49.09	49.36	31.22
C10	Cwa	196.97	34.60	38.57	32.45	31.82	33.92	39.29	33.04	50.43	43.73	**31.71**
C11	Cwa	170.75	44.34	48.82	43.14	43.15	43.21	48.75	41.92	61.39	54.65	**37.35**
C12	Cwa	163.57	44.60	49.77	42.73	44.03	41.77	50.91	41.26	64.04	56.62	**37.25**
C13	Cfa	171.19	38.81	43.76	37.00	35.28	38.83	43.15	38.29	47.99	48.15	**30.92**
C14	Cfa	199.76	34.11	38.50	30.58	33.15	28.63	39.72	33.10	46.23	46.60	**28.11**
C15	Cfa	195.54	32.10	35.65	28.75	33.65	**25.21**	33.87	27.63	40.43	40.93	29.94
C16	Cfa	161.12	35.64	40.99	32.72	35.35	**29.71**	40.78	33.13	50.92	50.82	30.52
C17	Cwa	189.85	32.46	35.77	29.35	32.62	26.62	33.33	**25.70**	35.80	29.67	31.87
C18	Csb	98.38	69.40	71.50	62.09	72.14	**58.61**	153.56	64.86	81.60	76.46	64.87
C19	Csa	138.09	39.57	45.02	35.04	39.52	**33.27**	48.25	35.74	52.71	52.58	35.01
C20	Cwa	181.88	42.64	46.91	42.19	43.54	42.12	47.45	42.35	60.04	54.06	**35.93**
C21	Cwa	170.72	43.01	47.37	42.13	43.77	41.35	48.15	42.71	59.05	53.25	**36.54**
C22	Cwa	189.78	41.96	47.53	42.27	46.58	38.88	50.32	43.69	64.12	56.99	**36.76**
C23	Cfa	168.68	35.03	39.55	30.68	31.46	29.59	39.58	33.93	44.27	44.72	**28.64**
C24	Cfa	163.74	34.35	37.56	29.58	32.72	**27.91**	38.45	31.24	43.72	43.78	29.47
C25	Cfb	137.72	43.64	48.36	43.81	46.79	40.30	59.25	48.59	64.35	63.06	**37.02**
C26	Cfb	136.96	37.82	42.38	34.07	40.11	**32.42**	45.80	36.21	53.77	52.65	35.04
C27	Cwa	180.20	36.75	40.22	33.44	33.78	33.48	42.75	**32.21**	52.34	44.03	32.42
C28	Cfb	130.29	29.84	34.95	27.21	26.80	26.01	35.57	31.39	41.27	42.24	**24.78**
C29	Cfb	169.81	33.83	38.83	33.21	32.81	31.52	42.48	37.10	49.05	49.01	**27.09**
C30	Cfa	152.23	36.69	40.16	31.71	36.58	**29.82**	39.38	32.14	46.58	46.80	33.77
C31	Cwa	182.05	42.38	45.80	42.32	42.32	44.17	47.25	43.68	57.92	52.96	**36.30**
C32	Cwa	203.86	41.35	45.90	38.43	41.87	35.48	48.06	38.50	60.42	52.92	**34.44**
C33	Cfb	131.32	33.39	37.42	28.92	31.85	**27.50**	38.88	32.74	47.88	48.75	28.67
C34	Cfa	153.71	66.26	63.75	64.73	**52.58**	76.42	59.06	61.55	59.94	58.02	64.90
C35	Cfb	129.99	42.39	48.16	38.11	38.92	**36.70**	55.73	46.03	58.15	58.49	36.98
C36	Cfb	176.61	33.77	37.77	29.01	34.06	**26.42**	37.58	29.73	48.64	48.37	28.71
C37	Csa	138.39	40.08	45.26	34.72	38.52	**33.61**	46.38	36.50	50.99	51.33	34.28
C38	Cfb	139.41	39.26	44.08	40.45	48.58	38.27	45.38	43.09	59.64	57.61	**37.35**
C39	Cwa	150.05	38.65	44.17	33.83	34.84	32.72	47.04	34.65	53.42	46.16	**32.01**
C40	Cfa	159.92	39.69	44.64	37.58	36.41	35.93	45.86	39.67	55.08	54.81	**32.02**
C41	Cfb	168.73	36.70	41.57	33.80	39.25	**30.96**	41.91	35.25	54.36	53.05	32.18
C42	Cfa	164.55	37.77	41.70	31.93	32.67	30.31	46.82	34.69	48.65	48.79	**30.20**
C43	Csb	144.47	37.68	41.35	35.19	32.70	33.91	42.89	38.77	44.82	45.09	**30.41**
C44	Csb	138.71	39.77	42.66	36.02	34.45	**32.84**	46.26	41.99	49.67	49.21	34.17
C45	Csb	134.82	48.70	53.49	49.20	57.15	45.98	69.67	53.57	72.63	70.31	**43.30**
C46	Csb	120.10	49.82	57.51	46.74	42.33	49.27	63.16	47.82	58.74	60.31	**38.49**
C47	Cfa	148.76	33.10	37.23	29.12	31.51	**26.97**	38.90	32.19	45.39	45.47	27.96
C48	Csa	132.18	37.97	43.86	33.81	38.07	**31.97**	48.25	36.40	51.82	52.08	35.19
C49	Cfb	125.90	34.79	38.76	31.15	34.58	**28.81**	40.08	36.56	46.03	46.39	29.91
C50	Cwa	165.59	41.65	44.65	40.05	**36.28**	42.35	44.22	39.66	49.77	45.81	38.65
C51	Cfa	135.14	42.49	48.88	41.26	39.56	40.35	54.81	44.85	54.16	54.12	**33.03**

Table A.16
Same as Table A.13, but for sites in continental climates (Köppen–Geiger classification: D).

stn	kgc	$\mathbb{E}(D_h)$	Engerer2	Engerer4	Starke1	Starke2	Starke3	Abreu	Paulescu	Every1	Every2	Yang4
D1	Dfb	152.82	35.47	39.21	32.17	36.23	31.61	41.74	35.75	48.95	43.79	**31.06**
D2	Dfb	146.28	36.90	41.09	32.45	37.28	**31.45**	47.43	35.86	55.54	49.21	32.30
D3	Dfc	129.62	31.84	35.45	29.36	33.78	**27.16**	38.13	30.99	48.77	45.20	27.46
D4	Dfb	156.01	33.71	37.25	29.51	32.38	30.81	39.38	32.91	44.60	39.39	**28.92**
D5	Dfb	183.49	34.05	38.37	30.87	36.02	**30.78**	36.92	31.81	46.78	41.86	30.82
D6	Dfa	145.92	37.65	41.35	35.54	41.29	33.96	43.23	37.73	52.33	46.80	**32.96**
D7	Dsd	84.73	**42.91**	44.10	48.32	56.01	47.82	47.49	47.34	59.44	52.39	43.00
D8	Dfb	146.70	34.99	38.97	30.83	33.94	31.84	40.11	33.53	50.88	45.57	**30.65**

Table A.17
Same as Table A.13, but for sites in polar climates (Köppen–Geiger classification: E).

stn	kgc	$\mathbb{E}(D_h)$	Engerer2	Engerer4	Starke1	Starke2	Starke3	Abreu	Paulescu	Every1	Every2	Yang4
E1	ET	137.72	42.92	45.31	41.36	47.09	**33.41**	39.43	43.52	55.67	52.04	41.90
E2	ET	148.15	35.07	38.60	32.50	36.47	79.11	62.24	36.12	56.05	49.78	**31.81**
E3	ET	147.07	36.04	39.62	36.19	42.41	80.08	68.69	39.03	55.47	49.83	**32.71**
E4	ET	100.12	38.56	42.27	42.45	50.23	41.41	53.83	43.73	58.66	55.09	**36.47**
E5	EF	117.32	64.54	**63.96**	68.34	76.30	71.14	85.45	73.05	88.92	81.45	65.04

Table A.18
Same as Table A.13, but for sites on islands.

stn	kgc	$\mathbb{E}(D_h)$	Engerer2	Engerer4	Starke1	Starke2	Starke3	Abreu	Paulescu	Every1	Every2	Yang4
O1	Ocean	204.06	31.97	36.52	29.40	30.57	**27.10**	34.57	29.00	42.14	37.01	27.74
O2	Ocean	206.47	36.68	40.90	34.31	37.39	33.22	39.21	34.94	49.52	42.90	**31.73**
O3	Ocean	209.34	35.43	39.62	32.10	34.43	31.52	38.05	32.81	46.98	40.91	**30.36**
O4	Ocean	190.44	33.29	38.67	29.52	34.42	**26.85**	35.97	28.65	46.75	39.28	30.18
O5	Ocean	173.37	38.46	42.08	36.60	35.94	39.86	41.59	35.47	47.39	42.68	**33.89**
O6	Ocean	188.52	33.78	36.65	31.56	36.63	29.69	35.31	**27.06**	42.27	32.99	36.85
O7	Ocean	220.40	33.68	38.29	30.88	33.41	**27.18**	36.40	29.52	43.46	38.49	28.99
O8	Ocean	188.99	39.02	42.25	37.56	41.51	36.55	40.99	**34.71**	50.44	42.40	38.16
O9	Ocean	173.61	38.48	41.76	36.48	36.97	34.41	41.29	38.50	47.91	44.81	**32.39**
O10	Ocean	236.16	35.24	40.34	33.85	35.99	**27.44**	38.46	31.07	49.75	42.72	30.14
O11	Ocean	166.17	34.17	37.92	32.39	36.09	33.72	36.58	**30.62**	44.01	36.87	33.79
O12	Ocean	120.34	**66.00**	67.89	74.90	80.96	77.30	69.73	79.94	96.28	88.58	68.62

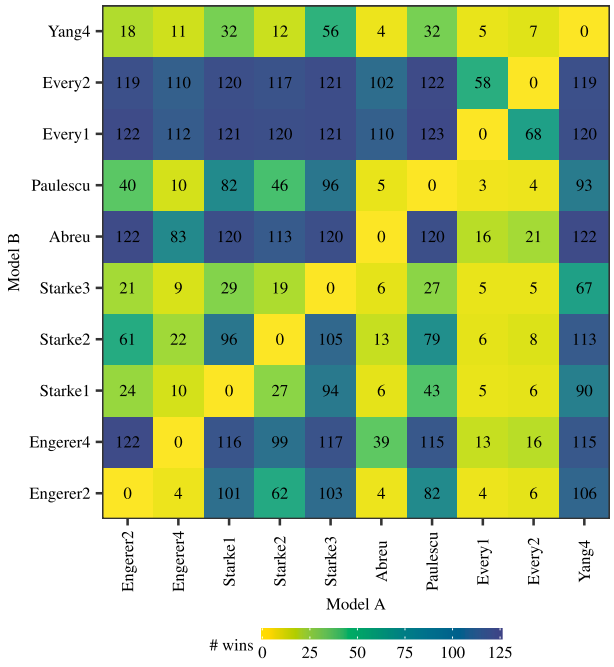


Fig. 4. Same as Fig. 3, but for the results of DM test on D_h predictions.

corresponds to the errors at all sites under one climate category, one accuracy measure, and one model. For example, there are 9 sites in the tropics, so each boxplot for "Tropical (A)" is plotted based on 9 samples. In order to avoid color overload, only four models, namely, Yang4, Starke3, Starke1, and Engerer2, which have been previously identified as the top ones, are displayed.

From Fig. 5, one can readily see that the median error of Yang4, which is marked by the middle bar in the box, is often the lowest (or closest to zero in terms of nMBE) among its peers. This is especially true for dry and temperate climates, which are most suitable for solar energy production. On this point, based on the visual assessment, as well as the ranking statistics and DM test results, one may confidently conclude that Yang4 *is able to replace* Engerer2 *as the new quasi-universal model.*

4.4. Observation and discussion

Before concluding the work, there are several important observations needing discussion. Firstly, although each of Abreu, Every2, and Starke3 is treated as one model, they are in fact composed of several models, whose fitted coefficients differ by climate. In this regard, anyone would logically take as true that climate-dependent coefficients are able to lead to better predictive performance. Nevertheless, this supposition has not really been confirmed by the results of the comprehensive assessment. For instance, for B_n prediction, Starke3 attained a slightly worse result than Starke1, whereas Abreu and Every2 are grossly dominated by other models. The reason for the poor performance of Abreu and Every2 is straightforward: Abreu has just a single predictor, and Every2 uses two predictors. One can thus claim that the importance placed on employing multiple relevant predictors should be higher than using climate-dependent coefficients. On the other hand, the reason that Starke3 is unable to outperform Starke1 by significant margins is more intricate. Since the function forms of the two models are basically the same (except that Starke3 has one more input parameter), it suggests that the climate-dependent features in k_t–k_d relationship are not distinctive. Recall that Köppen–Geiger classification divides climates based on seasonal precipitation and temperature patterns, which do not have a decisive effect on the variation in diffuse radiation. In that, whether radiation patterns follow climate classification is generally unclear. On this point, if we are to fit conditional separation models, the models may be better off if they are fitted based on different sky conditions, or different seasonal radiation patterns. This fact seems to have been largely overlooked by the literature thus far.

The second part of the discussion is related to the choice of input parameters. Variability index can come in different forms or combinations —e.g., Starke1, Starke2, Every1, and Every2 use $k_{t,\text{daily}}$; Starke3 uses both $k_{t,\text{daily}}$ and $k_{t,\text{hourly}}$; and Yang4 uses $k_{d,\text{hourly}}^{\text{Engerer2}}$. Evidently, these choices are not all essential, for the explanatory power for some of them are stronger than others, e.g., adding $k_{t,\text{hourly}}$ in Starke3 seems unnecessary. The usual way of identifying whether or not the effect of a parameter is significant is to use statistical testing, as done in [61]. Nevertheless, the validity of these tests is in doubt when the sample size is too large, as explained by Yang and Boland [58]—it is known as the "p-value problem"; this has been ignored by Starke et al. [61] when they conducted their hypothesis testing. Another questionable choice of input parameter is AST. There are sites where the mean atmospheric transparency is higher in AM than PM, but the reverse can be true elsewhere. AST was introduced by Boland et al. [84] to slightly improve the fit at their test sites in Australia, but this may lead to over-fitting, and likely worsens results at a significant fraction of sites with a different AM/PM pattern (pers. comm. Chris Gueymard). What should be included instead is albedo, which is the main factor affecting backscattering, and thus diffuse radiation, see [32,85] for more discussion.

The best way to improve accuracy, however, is always to inquire into the physics governing the radiative-transfer process, which has hitherto been lacking, with the notable exception of the pioneer works by Hollands and Crha [85] and Hollands [86]. Unfortunately, judging from the low citations of Hollands and Crha [85] and Hollands [86], most researchers do not seem to be even aware of these important early works on physical modeling. Empirical models are never too impressive, and the reluctance and incapacity of moving into the more difficult physics-based modeling is perhaps what slows down the progress of separation modeling research at the moment. Surely, we can refit the BRL model using new data, as done in [62], or add some

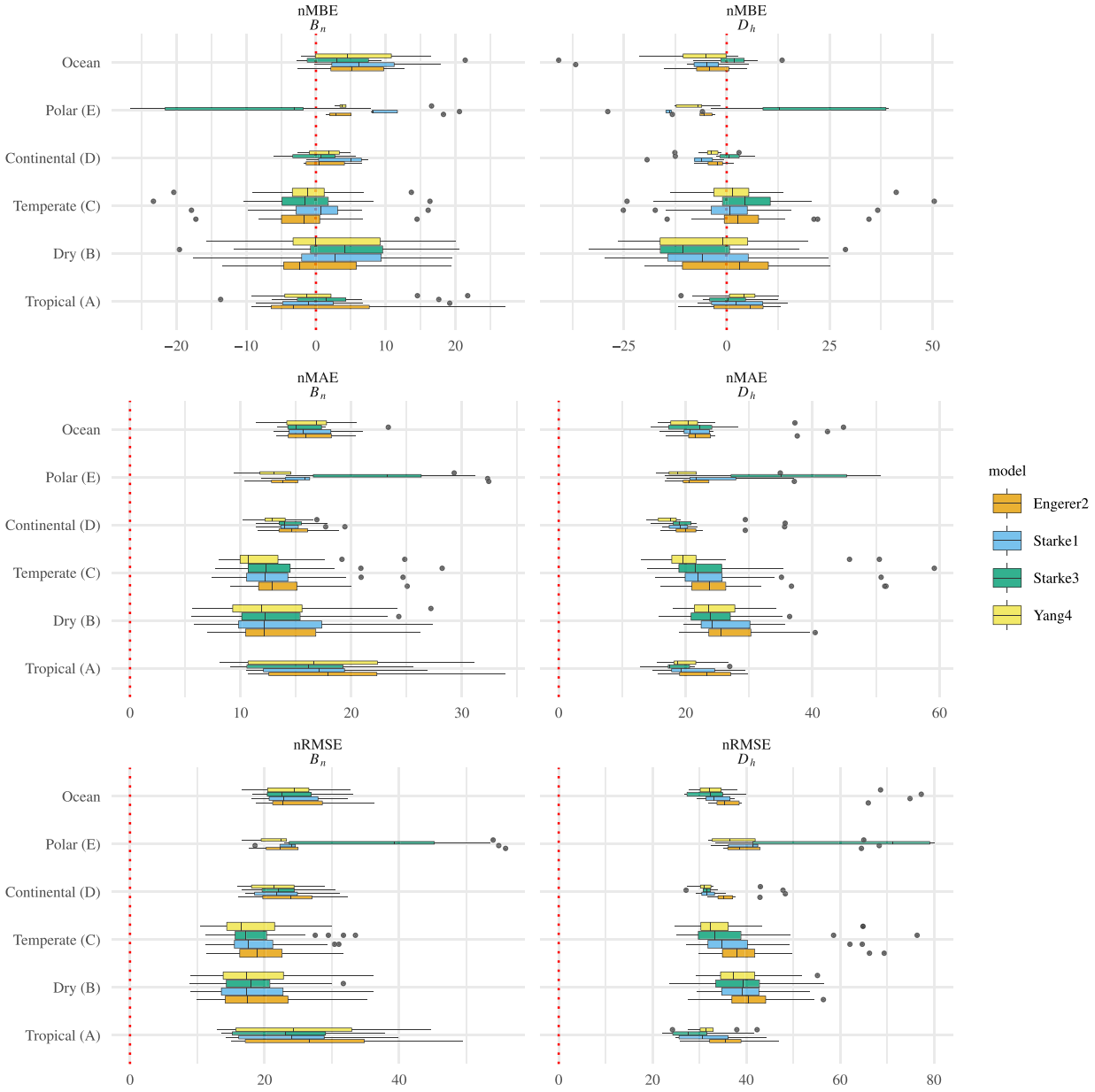


Fig. 5. nMBE, nMAE, and nRMSE (all in percent) of B_n and D_h estimates from three selected separation models, namely, Engerer2, Starke3, and Yang4. Tukey's boxplots are used for visualization. Dots beyond the ends of whiskers indicate outliers. The evaluation is grouped by Köppen–Geiger climate class. The height of the boxes is proportional to the number of sites.

new parameters here and there, as done in [61]—anyone with decent computer literacy can perform these tasks in a few days if not a few hours—but these so-claimed "innovations" or "novelties" have now been shown to be unable to bring about substantial improvement to the accuracy of separation models. "*Jootsing*," a word that is used by philosopher Daniel C. Dennett to describe processes in which one needs to understand the system in order to jump out of it [87], seems to be what we need, most desperately.

5. Conclusion

This article is concerned with three things. The first of those is to review the literature on separation modeling post 2016. As the bulkiest class of radiation models, separation models, which seek to split beam and diffuse components from the global one, are mostly empirical (as opposed to physical). For that reason, researchers have attempted to approach the task by opting different mathematical function forms, selecting different input parameters, and fitting different sets of coefficients with data from different climatic zones. Indeed, there have been more than 150 models proposed (the actual number could be way higher), and because each proposition is essentially trying to claim that the new model is able to outperform the old ones, there was not any consensus on what constitutes a good model until the review in 2016 by Gueymard and Ruiz-Arias [32], who identified Engerer2 [33] as the quasi-universal one. This article, therefore, reviews the literature since then. The merit of this review is that it argues what is the state-of-the-art and what are the possible innovations which one can make when it comes to separation modeling. At the same time, some outdated and non-representative separation modeling approaches have been identified, which should serve as an alert to prevent future repetitions on those low-value research ideas.

Next, a comprehensive benchmarking dataset is prepared, which can facilitate future research on separation modeling. As a collective effort of the IEA's PVPS Task 16 Activity 1.4 members, research-grade

1-min irradiance data from a total of 126 sites worldwide, spanning 5 years (2016–2020) are gathered, arranged, and quality controlled into a ready-to-use form. There is no doubt that this dataset is by far the most reliable and most comprehensive one. And any conclusion made based on this dataset would be most convincing, since it covers most, if not all, possible climatic, weather, and sky conditions on this planet. Since separation models, insofar as the current state of affairs is concerned, are often susceptible to over- or under-fitting due to their empirical nature, it is exceptionally important to test the universality of the model at unseen sites, with hold-out samples. Clearly, the present data is invaluable in that respect. It should be noted, however, that the data release is postponed to a later time, and at a separate venue, because there are some licensing issues that need to be sorted out first (see "data availability" below).

The third goal of this article is to investigate whether or not Engerer2 can be superseded by other more recent separation models. In order to do so, a comprehensive assessment of 10 recent models has been conducted. Through ranking statistics, Diebold–Mariano test for comparing predictive accuracy, as well as visual inspection, Yang4 [49] has been found to have the best overall performance, and therefore ought to replace Engerer2 as the new quasi-universal model. The modeling philosophy of Yang4 is two-fold. First is that it follows Engerer2 and explicitly models cloud enhancement in the form of an additive term, to quantify the amount of irradiance boost under cloud-enhancement events. Second is that it utilizes a so-called "temporal-resolution cascade" approach, which connects two separation models at different temporal resolutions, and thus accounts for variability in diffuse fraction at different frequencies. Although Yang4 is evidently the best model to date, it can still be improved in terms of: (1) including the effect of albedo enhancement, (2) introducing regime-switching coefficients (perhaps, based on the cloud amount data), and most importantly (3) improving the physical representation of the radiative-transfer process. These should be the subjects of future works.

Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

In the modern research world, reproducibility is being increasingly valued. On this point, it is clearly of interest to release the data and code alongside this article. Unfortunately, although a majority of stations used here provide publicly available data, there are several proprietary ones, of which the data cannot be shared in any form without the consent from the data owners. More importantly, since the cleaning and quality control of the dataset are performed as a collective effort of the IEA members, it is not appropriate to release it with this article. In a personal communication, Anne Forstinger, who leads the related work within the IEA's PVPS Task 16 Activity 1.4, noted that the dataset will be published on a separate server with a digital object identifier (DOI), but at a later time, after acquiring a special permission. The present author agrees with this decision. That said, once the original dataset is in the public domain, it is possible to supplement the present article with an addendum, in which the processed dataset and the computed code used to generate the results are to be included for total reproducibility.

Acknowledgments

The author would like to thank the following institutions and colleagues for the provision of data for this study:

- Baseline Surface Radiation Network
- Australian Government Bureau of Meteorology
- NOAA Global Monitoring Laboratory
- ESMAP program of the World Bank Group in particular Joana Zerbin, Clara Ivanescu, Branislav Schnierer, Roman Affolter, Geo-SUN Africa, Rachel Fox, Margot King
- SKYNET, in particular Hitoshi Irie and Tamio Takamura (CEReS/Chiba-U.), Chiba University, Tadahiro Hayasaka (Tohoku University), Chulalongkorn University
- Department of Civil Engineering at the Technical University of Denmark
- Swedish Meteorological and Hydrological Institute
- INPE National Institute of Space Research CCST Center for Earth System Sciences with FINEP Financier of Studies and Projects Ministry of Science and Technology and PETROBRAS Petróleo Brasileiro
- Cairo University in Egypt
- University of Oujda
- Institute Research Solar Energy Et Energies Nouvelles (IRESEN) in Morocco
- Research and Technology Centre of Energy (CRTEn) in Tunisia
- University of Jordan in Jordan and the Centre de Developpement des Energies Renouvelables (CDER) in Algeria.
- Grant Muller from NamPower in Namibia
- Yuldash Sobirov from the Institute of Material Science in Uzbekistan
- Frank Vignola from the University of Oregon
- David Pozo from the University Jaen
- Dietmar Baumgartner from the University of Graz
- Julian Gröbner at PMOD
- Nicolas Fernay from the University Lille
- Peter Armstrong at the Masdar Institute
- Laurent Vuilleumier at MeteoSwiss
- Irena Balog at ENEA
- Sophie Pelland at CanmetÉNERGIE Varennes
- Emmanuel Guillot at CNRS-PROMES, Odeillo

The author also thanks the IEA PVPS Task 16 partners CSP Services, MINES ParisTech/ARMINES, DLR, Solar Consulting Services, Harbin Institute of Technology, CENER, DTU, RSE, CIEMAT and Uni Malaga for providing their datasets, for combining the datasets, and for providing the results of the quality control.

Appendix. nRMSE Tables for individual sites

See Tables A.7–A.18.

References

[1] Long CN, Shi Y. An automated quality assessment and control algorithm for surface radiation measurements. Open Atmos Sci J 2008;2(1):23–37.
[2] Salazar G, Gueymard C, Galdino JB, de Castro Vilela O, Fraidenraich N. Solar irradiance time series derived from high-quality measurements, satellite-based models, and reanalyses at a near-equatorial site in Brazil. Renew Sustain Energy Rev 2020;117:109478. http://dx.doi.org/10.1016/j.rser.2019.109478, URL: https://www.sciencedirect.com/science/article/pii/S1364032119306860.
[3] Yang D, Yagli GM, Quan H. Quality control for solar irradiance data. In: 2018 IEEE innovative smart grid technologies - Asia (ISGT Asia). 2018, p. 208–13. http://dx.doi.org/10.1109/ISGT-Asia.2018.8467892.
[4] Yang D. SolarData: An R package for easy access of publicly available solar datasets. Sol Energy 2018;171:A3–12. http://dx.doi.org/10.1016/j.solener.2018.06.107, URL: https://www.sciencedirect.com/science/article/pii/S0038092X18306583.

[5] Kamphuis NR, Gueymard CA, Holtzapple MT, Duggleby AT, Annamalai K. Perspectives on the origin, derivation, meaning, and significance of the isotropic sky model. Sol Energy 2020;201:8–12. http://dx.doi.org/10.1016/j.solener.2020.02.067, URL: https://www.sciencedirect.com/science/article/pii/S0038092X20301948.

[6] Perez R, Stewart R, Arbogast C, Seals R, Scott J. An anisotropic hourly diffuse radiation model for sloping surfaces: Description, performance validation, site dependency evaluation. Sol Energy 1986;36(6):481–97. http://dx.doi.org/10.1016/0038-092X(86)90013-7, URL: https://www.sciencedirect.com/science/article/pii/0038092X86900137.

[7] Perez R, Seals R, Ineichen P, Stewart R, Menicucci D. A new simplified version of the Perez diffuse irradiance model for tilted surfaces. Sol Energy 1987;39(3):221–31. http://dx.doi.org/10.1016/S0038-092X(87)80031-2, URL: https://www.sciencedirect.com/science/article/pii/S0038092X87800312.

[8] Perez R, Stewart R, Seals R, Guertin T. The development and verification of the Perez diffuse radiation model. Technical Report SAND88- 7030, Albany, NY: Atmospheric Sciences Research Center, SUNY at Albany; 1988.

[9] Perez R, Ineichen P, Seals R, Michalsky J, Stewart R. Modeling daylight availability and irradiance components from direct and global irradiance. Sol Energy 1990;44(5):271–89. http://dx.doi.org/10.1016/0038-092X(90)90055-H, URL: https://www.sciencedirect.com/science/article/pii/0038092X9090055H.

[10] Yang D. Solar radiation on inclined surfaces: Corrections and benchmarks. Sol Energy 2016;136:288–302. http://dx.doi.org/10.1016/j.solener.2016.06.062, URL: https://www.sciencedirect.com/science/article/pii/S0038092X16302432.

[11] Yang D, Ye Z, Nobre AM, Du H, Walsh WM, Lim LI, Reindl T. Bidirectional irradiance transposition based on the Perez model. Sol Energy 2014;110:768–80. http://dx.doi.org/10.1016/j.solener.2014.10.006, URL: https://www.sciencedirect.com/science/article/pii/S0038092X14004927.

[12] Marion B. A model for deriving the direct normal and diffuse horizontal irradiance from the global tilted irradiance. Sol Energy 2015;122:1037–46. http://dx.doi.org/10.1016/j.solener.2015.10.024, URL: https://www.sciencedirect.com/science/article/pii/S0038092X15005757.

[13] Quan H, Yang D. Probabilistic solar irradiance transposition models. Renew Sustain Energy Rev 2020;125:109814. http://dx.doi.org/10.1016/j.rser.2020.109814, URL: https://www.sciencedirect.com/science/article/pii/S136403212030109X.

[14] Meng B, Loonen RCGM, Hensen JLM. Data-driven inference of unknown tilt and azimuth of distributed PV systems. Sol Energy 2020;211:418–32. http://dx.doi.org/10.1016/j.solener.2020.09.077, URL: https://www.sciencedirect.com/science/article/pii/S0038092X20310306.

[15] Sengupta M, Xie Y, Lopez A, Habte A, Maclaurin G, Shelby J. The National Solar Radiation Data Base (NSRDB). Renew Sustain Energy Rev 2018;89:51–60. http://dx.doi.org/10.1016/j.rser.2018.03.003, URL: https://www.sciencedirect.com/science/article/pii/S136403211830087X.

[16] Xie Y, Sengupta M, Dudhia J. A Fast All-sky Radiation Model for Solar applications (FARMS): Algorithm and performance evaluation. Sol Energy 2016;135:435–45. http://dx.doi.org/10.1016/j.solener.2016.06.003, URL: https://www.sciencedirect.com/science/article/pii/S0038092X16301827.

[17] Yang D. Choice of clear-sky model in solar forecasting. J Renew Sustain Energy 2020;12(2):026101. http://dx.doi.org/10.1063/5.0003495.

[18] Yang D. A guideline to solar forecasting research practice: Reproducible, operational, probabilistic or physically-based, ensemble, and skill (ROPES). J Renew Sustain Energy 2019;11(2):022701. http://dx.doi.org/10.1063/1.5087462.

[19] Sun X, Bright JM, Gueymard CA, Acord B, Wang P, Engerer NA. Worldwide performance assessment of 75 global clear-sky irradiance models using Principal Component Analysis. Renew Sustain Energy Rev 2019;111:550–70. http://dx.doi.org/10.1016/j.rser.2019.04.006, URL: https://www.sciencedirect.com/science/article/pii/S1364032119302187.

[20] Sun X, Bright JM, Gueymard CA, Bai X, Acord B, Wang P. Worldwide performance assessment of 95 direct and diffuse clear-sky irradiance models using principal component analysis. Renew Sustain Energy Rev 2021;135:110087. http://dx.doi.org/10.1016/j.rser.2020.110087, URL: https://www.sciencedirect.com/science/article/pii/S1364032120303786.

[21] Gueymard CA. REST2: High-performance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation – Validation with a benchmark dataset. Sol Energy 2008;82(3):272–85. http://dx.doi.org/10.1016/j.solener.2007.04.008, URL: http://www.sciencedirect.com/science/article/pii/S0038092X07000990.

[22] Gueymard CA. A reevaluation of the solar constant based on a 42-year total solar irradiance time series and a reconciliation of spaceborne observations. Sol Energy 2018;168:2–9. http://dx.doi.org/10.1016/j.solener.2018.04.001, URL: https://www.sciencedirect.com/science/article/pii/S0038092X18303463. Advances in Solar Resource Assessment and Forecasting.

[23] Gueymard CA. The SMARTS spectral irradiance model after 25 years: New developments and validation of reference spectra. Sol Energy 2019;187:233–53. http://dx.doi.org/10.1016/j.solener.2019.05.048, URL: https://www.sciencedirect.com/science/article/pii/S0038092X19305110.

[24] Huang G, Li Z, Li X, Liang S, Yang K, Wang D, Zhang Y. Estimating surface solar irradiance from satellites: Past, present, and future perspectives. Remote Sens Environ 2019;233:111371. http://dx.doi.org/10.1016/j.rse.2019.111371, URL: https://www.sciencedirect.com/science/article/pii/S0034425719303906.

[25] Miller SD, Rogers MA, Haynes JM, Sengupta M, Heidinger AK. Short-term solar irradiance forecasting via satellite/model coupling. Sol Energy 2018;168:102–17. http://dx.doi.org/10.1016/j.solener.2017.11.049, URL: https://www.sciencedirect.com/science/article/pii/S0038092X17310435. Advances in Solar Resource Assessment and Forecasting.

[26] Polo J, Fernández-Peruchena C, Salamalikis V, Mazorra-Aguiar L, Turpin M, Martín-Pomares L, Kazantzidis A, Blanc P, Remund J. Benchmarking on improvement and site-adaptation techniques for modeled solar radiation datasets. Sol Energy 2020;201:469–79. http://dx.doi.org/10.1016/j.solener.2020.03.040, URL: https://www.sciencedirect.com/science/article/pii/S0038092X20302784.

[27] Yang D. Ensemble model output statistics as a probabilistic site-adaptation tool for satellite-derived and reanalysis solar irradiance. J Renew Sustain Energy 2020;12(1):016102. http://dx.doi.org/10.1063/1.5134731.

[28] Yang D. Ensemble model output statistics as a probabilistic site-adaptation tool for solar irradiance: A revisit. J Renew Sustain Energy 2020;12(3):036101. http://dx.doi.org/10.1063/5.0010003.

[29] Yang D, Gueymard CA. Probabilistic post-processing of gridded atmospheric variables and its application to site adaptation of shortwave solar radiation. Sol Energy 2021;225:427–43. http://dx.doi.org/10.1016/j.solener.2021.05.050, URL: https://www.sciencedirect.com/science/article/pii/S0038092X21004205.

[30] Holmgren WF, Hansen CW, Mikofski MA. Pvlib python: a python package for modeling solar energy systems. J Open Source Softw 2018;3(29):884. http://dx.doi.org/10.21105/joss.00884.

[31] Mayer MJ, Gróf G. Extensive comparison of physical models for photovoltaic power forecasting. Appl Energy 2021;283:116239. http://dx.doi.org/10.1016/j.apenergy.2020.116239, URL: https://www.sciencedirect.com/science/article/pii/S0306261920316330.

[32] Gueymard CA, Ruiz-Arias JA. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol Energy 2016;128:1–30. http://dx.doi.org/10.1016/j.solener.2015.10.010, URL: http://www.sciencedirect.com/science/article/pii/S0038092X15005435.

[33] Engerer NA. Minute resolution estimates of the diffuse fraction of global irradiance for southeastern Australia. Sol Energy 2015;116:215–37. http://dx.doi.org/10.1016/j.solener.2015.04.012, URL: https://www.sciencedirect.com/science/article/pii/S0038092X15001905.

[34] Gueymard CA. Cloud and albedo enhancement impacts on solar irradiance using high-frequency measurements from thermopile and photodiode radiometers. Part 1: Impacts on global horizontal irradiance. Sol Energy 2017;153:755–65. http://dx.doi.org/10.1016/j.solener.2017.05.004, URL: https://www.sciencedirect.com/science/article/pii/S0038092X1730381X.

[35] Gueymard CA. Cloud and albedo enhancement impacts on solar irradiance using high-frequency measurements from thermopile and photodiode radiometers. Part 2: Performance of separation and transposition models for global tilted irradiance. Sol Energy 2017;153:766–79. http://dx.doi.org/10.1016/j.solener.2017.04.068, URL: https://www.sciencedirect.com/science/article/pii/S0038092X17303730.

[36] Gueymard CA, Renné D, Vignola FE. Editorial: Journal's performance and publication criteria. Sol Energy 2009;83(1):1. http://dx.doi.org/10.1016/j.solener.2008.07.007, URL: https://www.sciencedirect.com/science/article/pii/S0038092X08001801.

[37] Song Z, Ren Z, Deng Q, Kang X, Zhou M, Liu D, Chen X. General models for estimating daily and monthly mean daily diffuse solar radiation in China's subtropical monsoon climatic zone. Renew Energy 2020;145:318–32. http://dx.doi.org/10.1016/j.renene.2019.06.019, URL: https://www.sciencedirect.com/science/article/pii/S0960148119308407.

[38] Jamil B, Siddiqui AT. Generalized models for estimation of diffuse solar radiation based on clearness index and sunshine duration in India: Applicability under different climatic zones. J Atmos Sol-Terr Phys 2017;157–158:16–34. http://dx.doi.org/10.1016/j.jastp.2017.03.013, URL: https://www.sciencedirect.com/science/article/pii/S1364682616302735.

[39] Jamil B, Akhtar N. Comparative analysis of diffuse solar radiation models based on sky-clearness index and sunshine period for humid-subtropical climatic region of India: A case study. Renew Sustain Energy Rev 2017;78:329–55. http://dx.doi.org/10.1016/j.rser.2017.04.073, URL: https://www.sciencedirect.com/science/article/pii/S136403211730583X.

[40] Fan J, Wu L, Zhang F, Cai H, Ma X, Bai H. Evaluation and development of empirical models for estimating daily and monthly mean daily diffuse horizontal solar radiation for different climatic regions of China. Renew Sustain Energy Rev 2019;105:168–86. http://dx.doi.org/10.1016/j.rser.2019.01.040, URL: https://www.sciencedirect.com/science/article/pii/S1364032119300620.

[41] Xue X. Prediction of daily diffuse solar radiation using artificial neural networks. Int J Hydrogen Energy 2017;42(47):28214–21. http://dx.doi.org/10.1016/j.ijhydene.2017.09.150, URL: https://www.sciencedirect.com/science/article/pii/S0360319917338247.

[42] Fan J, Wang X, Zhang F, Ma X, Wu L. Predicting daily diffuse horizontal solar radiation in various climatic regions of China using support vector machine and tree-based soft computing models with local and extrinsic climatic data. J Cleaner Prod 2020;248:119264. http://dx.doi.org/10.1016/j.jclepro.2019.119264, URL: https://www.sciencedirect.com/science/article/pii/S0959652619341344.

[43] Kasten F, Duffie JA. Editorial. Sol Energy 1993;50(5):383. http://dx.doi.org/10.1016/0038-092X(93)90058-V, URL: https://www.sciencedirect.com/science/article/pii/0038092X9390058V.

[44] Zhou Y, Wang D, Liu Y, Liu J. Diffuse solar radiation models for different climate zones in China: Model evaluation and general model development. Energy Convers Manage 2019;185:518–36. http://dx.doi.org/10.1016/j.enconman.2019.02.013, URL: https://www.sciencedirect.com/science/article/pii/S0196890419301955.

[45] Hofmann M, Seckmeyer G. A new model for estimating the diffuse fraction of solar irradiance for photovoltaic system simulations. Energies 2017;10(2). http://dx.doi.org/10.3390/en10020248, URL: https://www.mdpi.com/1996-1073/10/2/248.

[46] Hassan MA, Khalil A, Kaseb S, Kassem MA. Exploring the potential of tree-based ensemble methods in solar radiation modeling. Appl Energy 2017;203:897–916. http://dx.doi.org/10.1016/j.apenergy.2017.06.104, URL: https://www.sciencedirect.com/science/article/pii/S0306261917308437.

[47] Claywell R, Nadai L, Felde I, Ardabili S, Mosavi A. Adaptive neuro-fuzzy inference system and a multilayer perceptron model trained with grey wolf optimizer for predicting solar diffuse fraction. Entropy 2020;22(11). http://dx.doi.org/10.3390/e22111192, URL: https://www.mdpi.com/1099-4300/22/11/1192.

[48] Aler R, Galván IM, Ruiz-Arias JA, Gueymard CA. Improving the separation of direct and diffuse solar radiation components using machine learning by gradient boosting. Sol Energy 2017;150:558–69. http://dx.doi.org/10.1016/j.solener.2017.05.018, URL: https://www.sciencedirect.com/science/article/pii/S0038092X17303870.

[49] Yang D. Temporal-resolution cascade model for separation of 1-min beam and diffuse irradiance. J Renew Sustain Energy 2021;13(5):056101. http://dx.doi.org/10.1063/5.0067997.

[50] Bright JM, Engerer NA. Engerer2: Global re-parameterisation, update, and validation of an irradiance separation model at different temporal resolutions. J Renew Sustain Energy 2019;11(3):033701. http://dx.doi.org/10.1063/1.5097014.

[51] Yang D, Gueymard CA. Ensemble model output statistics for the separation of direct and diffuse components from 1-min global irradiance. Sol Energy 2020;208:591–603. http://dx.doi.org/10.1016/j.solener.2020.05.082, URL: https://www.sciencedirect.com/science/article/pii/S0038092X2030582X.

[52] Erbs DG, Klein SA, Duffie JA. Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation. Sol Energy 1982;28(4):293–302. http://dx.doi.org/10.1016/0038-092X(82)90302-4, URL: https://www.sciencedirect.com/science/article/pii/0038092X82903024.

[53] Orgill JF, Hollands KGT. Correlation equation for hourly diffuse radiation on a horizontal surface. Sol Energy 1977;19(4):357–9. http://dx.doi.org/10.1016/0038-092X(77)90006-8, URL: https://www.sciencedirect.com/science/article/pii/0038092X77900068.

[54] Starke AR, Lemos LFL, Boland J, Cardemil JM, Colle S. Resolution of the cloud enhancement problem for one-minute diffuse radiation prediction. Renew Energy 2018;125:472–84. http://dx.doi.org/10.1016/j.renene.2018.02.107, URL: https://www.sciencedirect.com/science/article/pii/S0960148118302593.

[55] Ridley B, Boland J, Lauret P. Modelling of diffuse solar fraction with multiple predictors. Renew Energy 2010;35(2):478–83. http://dx.doi.org/10.1016/j.renene.2009.07.018, URL: https://www.sciencedirect.com/science/article/pii/S0960148109003012.

[56] Abreu EFM, Canhoto P, Costa MJ. Prediction of diffuse horizontal irradiance using a new climate zone model. Renew Sustain Energy Rev 2019;110:28–42. http://dx.doi.org/10.1016/j.rser.2019.04.055, URL: https://www.sciencedirect.com/science/article/pii/S1364032119302679.

[57] Paulescu E, Blaga R. A simple and reliable empirical model with two predictors for estimating 1-minute diffuse fraction. Sol Energy 2019;180:75–84. http://dx.doi.org/10.1016/j.solener.2019.01.029, URL: https://www.sciencedirect.com/science/article/pii/S0038092X19300386.

[58] Yang D, Boland J. Satellite-augmented diffuse solar radiation separation models. J Renew Sustain Energy 2019;11(2):023705. http://dx.doi.org/10.1063/1.5087463.

[59] Yang D, Bright JM. Worldwide validation of 8 satellite-derived and reanalysis solar radiation products: A preliminary evaluation and overall metrics for hourly data over 27 years. Sol Energy 2020;210:3–19. http://dx.doi.org/10.1016/j.solener.2020.04.016, URL: https://www.sciencedirect.com/science/article/pii/S0038092X20303893. Special Issue on Grid Integration.

[60] Perez R, Ineichen P, Maxwell E, Seals R, Zelenka A. Dynamic global-to-direct irradiance conversion models. ASHRAE Trans 1992;98(1):354–69.

[61] Starke AR, Lemos LFL, Barni CM, Machado RD, Cardemil JM, Boland J, Colle S. Assessing one-minute diffuse fraction models based on worldwide climate features. Renew Energy 2021;177:700–14. http://dx.doi.org/10.1016/j.renene.2021.05.108, URL: https://www.sciencedirect.com/science/article/pii/S0960148121007916.

[62] Every JP, Li L, Dorrell DG. Köppen-Geiger climate classification adjustment of the BRL diffuse irradiation model for Australian locations. Renew Energy 2020;147:2453–69. http://dx.doi.org/10.1016/j.renene.2019.09.114, URL: https://www.sciencedirect.com/science/article/pii/S0960148119314521.

[63] Driemel A, Augustine J, Behrens K, Colle S, Cox C, Cuevas-Agulló E, Denn FM, Duprat T, Fukuda M, Grobe H, Haeffelin M, Hodges G, Hyett N, Ijima O, Kallis A, Knap W, Kustov V, Long CN, Longenecker D, Lupi A, Maturilli M, Mimouni M, Ntsangwane L, Ogihara H, Olano X, Olefs M, Omori M, Passamani L, Pereira EB, Schmithüsen H, Schumacher S, Sieger R, Tamlyn J, Vogt R, Vuilleumier L, Xia X, Ohmura A, König-Langlo G. Baseline Surface Radiation Network (BSRN): structure and data description (1992–2017). Earth Syst Sci Data 2018;10(3):1491–501. http://dx.doi.org/10.5194/essd-10-1491-2018.

[64] Stoffel T, Andreas A. NREL Solar Radiation Research Laboratory (SRRL): Baseline measurement system (BMS); Golden, Colorado (data). Technical Report NREL/DA-5500-56488, Golden, CO: National Renewable Energy Laboratory; 1981, http://dx.doi.org/10.7799/1052221.

[65] Hicks BB, DeLuisi JJ, Matt DR. The NOAA Integrated Surface Irradiance Study (ISIS)—A new surface radiation monitoring program. Bull Am Meteorol Soc 1996;77(12):2857–64. http://dx.doi.org/10.1175/1520-0477(1996)077<2857:TNISIS>2.0.CO;2, URL: https://journals.ametsoc.org/view/journals/bams/77/12/1520-0477_1996_077_2857_tnisis_2_0_co_2.xml.

[66] Peterson J, Vignola F. Structure of a comprehensive solar radiation dataset. Sol Energy 2020;211:366–74. http://dx.doi.org/10.1016/j.solener.2020.08.092, URL: https://www.sciencedirect.com/science/article/pii/S0038092X20309312.

[67] Brooks MJ, Du Clou S, Van Niekerk WL, Gauché P, Leonard C, Mouzouris MJ, Meyer R, Van der Westhuizen N, Van Dyk EE, Vorster FJ. SAURAN: A new resource for solar radiometric data in Southern Africa. J Energy South Afr 2015;26(1):2–10.

[68] Forstinger A, Wilbert S, Kraas B, Peruchena CF, Gueymard CA, Collino E, Ruiz-Arias JA, Polo Martinez J, Saint-Drenan Y-M, Ronzio D, Hanrieder N, Jensen AR, Yang D. Expert quality control of solar radiation ground data sets. In: Solar world congress 2021. Virtual conference: International Solar Energy Society; 2021, (in press).

[69] Bright JM. Solcast: Validation of a satellite-derived solar irradiance dataset. Sol Energy 2019;189:435–49. http://dx.doi.org/10.1016/j.solener.2019.07.086, URL: https://www.sciencedirect.com/science/article/pii/S0038092X19307571.

[70] Killinger S, Engerer N, Müller B. QCPV: A quality control algorithm for distributed photovoltaic array power output. Sol Energy 2017;143:120–31. http://dx.doi.org/10.1016/j.solener.2016.12.053, URL: https://www.sciencedirect.com/science/article/pii/S0038092X16306600.

[71] Bright JM. The impact of globally diverse GHI training data: Evaluation through application of a simple Markov chain downscaling methodology. J Renew Sustain Energy 2019;11(2):023703. http://dx.doi.org/10.1063/1.5085236.

[72] Yang D, Alessandrini S, Antonanzas J, Antonanzas-Torres F, Badescu V, Beyer HG, Blaga R, Boland J, Bright JM, Coimbra CFM, David M, Frimane Â, Gueymard CA, Hong T, Kay MJ, Killinger S, Kleissl J, Lauret P, Lorenz E, van der Meer D, Paulescu M, Perez R, Perpiñán-Lamigueiro O, Peters IM, Reikard G, Renné D, Saint-Drenan Y-M, Shuai Y, Urraca R, Verbois H, Vignola F, Voyant C, Zhang J. Verification of deterministic solar forecasts. Sol Energy 2020;210:20–37. http://dx.doi.org/10.1016/j.solener.2020.04.019, URL: https://www.sciencedirect.com/science/article/pii/S0038092X20303947. Special Issue on Grid Integration.

[73] Yagli GM, Yang D, Gandhi O, Srinivasan D. Can we justify producing univariate machine-learning forecasts with satellite-derived solar irradiance? Appl Energy 2020;259:114122. http://dx.doi.org/10.1016/j.apenergy.2019.114122, URL: https://www.sciencedirect.com/science/article/pii/S0306261919318094.

[74] Yang D. Standard of reference in operational day-ahead deterministic solar forecasting. J Renew Sustain Energy 2019;11(5):053702. http://dx.doi.org/10.1063/1.5114985.

[75] Yang D, Perez R. Can we gauge forecasts using satellite-derived solar irradiance? J Renew Sustain Energy 2019;11(2):023704. http://dx.doi.org/10.1063/1.5087588.

[76] Murphy AH. Forecast verification: Its complexity and dimensionality. Mon Weather Rev 1991;119(7):1590–601. http://dx.doi.org/10.1175/1520-0493(1991)119<1590:FVICAD>2.0.CO;2, URL: https://journals.ametsoc.org/view/journals/mwre/119/7/1520-0493_1991_119_1590_fvicad_2_0_co_2.xml.

[77] Alvo M, Yu PLH. Exploratory analysis of ranking data. In: Statistical methods for ranking data. Frontiers in probability and the statistical sciences, Springer New York; 2014, p. 7–21.

[78] Yang D, Kleissl J, Gueymard CA, Pedro HTC, Coimbra CFM. History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. Sol Energy 2018;168:60–101. http://dx.doi.org/10.1016/j.solener.2017.11.023, URL: https://www.sciencedirect.com/science/article/pii/S0038092X17310022. Advances in Solar Resource Assessment and Forecasting.

[79] Diebold FX, Mariano RS. Comparing predictive accuracy. J Bus Econom Statist 1995;13(3):253–63. http://dx.doi.org/10.2307/1392185, URL: https://www.jstor.org/stable/1392185.

[80] Yang D. Reconciling solar forecasts: Probabilistic forecast reconciliation in a nonparametric framework. Sol Energy 2020;210:49–58. http://dx.doi.org/10.1016/j.solener.2020.03.095, URL: https://www.sciencedirect.com/science/article/pii/S0038092X20303418. Special Issue on Grid Integration.

[81] Yang D, Quan H, Disfani VR, Rodríguez-Gallegos CD. Reconciling solar forecasts: Temporal hierarchy. Sol Energy 2017;158:332–46. http://dx.doi.org/10.1016/j.solener.2017.09.055, URL: https://www.sciencedirect.com/science/article/pii/S0038092X17308423.

[82] Yang D, Quan H, Disfani VR, Liu L. Reconciling solar forecasts: Geographical hierarchy. Sol Energy 2017;146:276–86. http://dx.doi.org/10.1016/j.solener.2017.02.010, URL: https://www.sciencedirect.com/science/article/pii/S0038092X17301020.

[83] Yang D. Post-processing of NWP forecasts using ground or satellite-derived data through kernel conditional density estimation. J Renew Sustain Energy 2019;11(2):026101. http://dx.doi.org/10.1063/1.5088721.

[84] Boland J, Scott L, Luther M. Modelling the diffuse fraction of global solar radiation on a horizontal surface. Environmetrics 2001;12(2):103–16. http://dx.doi.org/10.1002/1099-095X(200103)12:2<103::AID-ENV447>3.0.CO;2-2.

[85] Hollands KGT, Crha SJ. An improved model for diffuse radiation: Correction for atmospheric back-scattering. Sol Energy 1987;38(4):233–6. http://dx.doi.org/10.1016/0038-092X(87)90044-2, URL: https://www.sciencedirect.com/science/article/pii/0038092X87900442.

[86] Hollands KGT. A derivation of the diffuse fraction's dependence on the clearness index. Sol Energy 1985;35(2):131–6. http://dx.doi.org/10.1016/0038-092X(85)90003-9, URL: https://www.sciencedirect.com/science/article/pii/0038092X85900039.

[87] Dennett DC. Intuition Pumps and Other Tools for Thinking. W. W. Norton & Company; 2013.