Renewable and Sustainable Energy Reviews 189 (2024) 113992

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

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

Original research article

Regime-dependent 1-min irradiance separation model with climatology clustering

Dazhi Yang [a,*], Yizhan Gu [a,b], Martin János Mayer [c], Christian A. Gueymard [d], Wenting Wang [a], Jan Kleissl [b], Mengying Li [e], Yinghao Chu [f], Jamie M. Bright [g]

[a] School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin, Heilongjiang, China
[b] Center for Energy Research and Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA, USA
[c] Department of Energy Engineering, Faculty of Mechanical Engineering, Budapest University of Technology and Economics, Műegyetem rkp. 3, H-1111, Budapest, Hungary
[d] Solar Consulting Services, Colebrook, NH, USA
[e] Department of Mechanical Engineering & Research Institute for Smart Energy, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region
[f] Department of Advanced Design and Systems Engineering, City University of Hong Kong, Hong Kong Special Administrative Region
[g] UK Power Networks, London, UK

ARTICLE INFO

Dataset link: https://github.com/dazhiyang/Yang5-Separation

Keywords:
Solar radiation
Separation modeling
Worldwide validation
Regime-dependent model
Cluster analysis

ABSTRACT

Since directly measuring beam and diffuse irradiance is not feasible on many occasions, one often has to resort to estimating the beam and diffuse irradiance components from the global irradiance, which is known as separation modeling. Separation modeling is essentially a nonlinear regression problem, with the clearness index being the main input and the diffuse fraction being the output. Hundreds of separation models with various complexities have been proposed, among which the Yang4 model was recently validated using worldwide data as the quasi-universal choice for 1-min data. In this work, Yang4 is further improved by regime-dependent fitting, i.e., fitting a separate set of model coefficients for each climatological regime. Different regimes are determined through clustering of cloud cover frequency, aerosol optical depth, and surface albedo climatology maps. The new Yang5 model is able to outperform its predecessor at the 126 stations tested, covering a wide range of climate types. Overall, the normalized root mean square errors for beam normal irradiance (BNI) and diffuse horizontal irradiance (DHI) of Yang5 are 17.55% and 32.92% on average, as compared to 19.13% and 34.94% for the next best model, namely, Yang4. Furthermore, through conducting pairwise Diebold–Mariano tests, Yang5 is shown superior to Yang4 at 110/126 sites for BNI prediction and 93/126 for DHI.

1. Introduction

In solar energy meteorology, a fundamental relationship between solar radiation components is the closure equation, which states the fact that the global horizontal irradiance (GHI) is constituted of a beam component and a diffuse component:

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

where G_h, B_n, B_h, and D_h are GHI, beam normal irradiance (BNI), beam horizontal irradiance (BHI), and diffuse horizontal irradiance (DHI), respectively; and Z is the solar zenith angle, which can be computed deterministically via a solar position algorithm, such as the SG2 algorithm [1]. Notwithstanding, measuring or estimating all irradiance components is often seen as a tedious and costly task. Consequently, under the frequent situations where only GHI is available, either from radiometric measurement or retrieved from remote-sensing data, one often obtains the other unknown components through what is known as "separation" or "decomposition" modeling. Separation modeling is an important stage of the model chain, which serves to convert irradiance to photovoltaic power, and thus is profoundly useful for solar resource assessment [2–4] and forecasting [5–7].

Since surface radiation components exhibit both a diurnal cycle and a seasonal cycle as a result of the changing apparent sun–earth position, it is customary to perform solar modeling on normalized quantities. In separation modeling, specifically, the two most relevant normalized quantities are the diffuse fraction k which is the ratio between DHI and GHI (i.e., $k = D_h/G_h$), and the clearness index k_t which is the ratio between GHI and its extraterrestrial counterpart (i.e, $k_t = G_h/E_0$, where E_0 is the extraterrestrial GHI). Stated differently, separation models seek to estimate k using k_t and some other auxiliary variables. Early separation models are often simply as univariate functions of k_t,

* Corresponding author.
E-mail address: yangdazhi.nus@gmail.com (D. Yang).

https://doi.org/10.1016/j.rser.2023.113992
Received 21 August 2022; Received in revised form 16 September 2023; Accepted 20 October 2023

Nomenclature

Abbreviations

AOD	aerosol optical depth
AST	apparent solar time
BHI	beam horizontal irradiance
BNI	beam normal irradiance
DHI	diffuse horizontal irradiance
ECMWF	European Centre of Medium-Range Weather Forecast
ERA5	fifth-generation ECMWF Reanalysis
GHI	global horizontal irradiance
IEA	International Energy Agency
KGC	Köppen–Geiger climate classification
MERRA-2	Modern-Era Retrospective Analysis for Research and Applications, version 2
MODCF	MODIS-based cloud frequency
MODIS	Moderate Resolution Imaging Spectroradiometer
nMBE	normalized mean bias error
nRMSE	normalized root mean square error
PCA	principal component analysis
PVPS	Photovoltaic Power Systems Programme
RCC	radiation climate classification

Notations

α	$90° - Z$, elevation angle [degree]
Δk_{tc}	$k_{tc} - k_t$, difference between the clearness index of clear-sky global horizontal irradiance and the clearness index [dimensionless]
ψ	three-point moving average of 1-min k_t [dimensionless]
B_h	beam horizontal irradiance [W/m^2]
B_n	beam normal irradiance [W/m^2]
D_h	diffuse horizontal irradiance [W/m^2]
E_0	extraterrestrial global horizontal irradiance [W/m^2]
G_{csky}	clear-sky global horizontal irradiance, obtained from the McClear database [W/m^2]
G_h	global horizontal irradiance [W/m^2]
k	D_h/G_h, diffuse fraction [0–1]
$k^{(s)}$	hourly or half-hourly satellite-derived diffuse fraction [0–1]
k_{csi}	G_h/G_{csky}, Starke's quantifier for cloud enhancement [dimensionless]
$k_{\mathrm{hourly}}^{\mathrm{Engerer2}}$	hourly diffuse fraction estimate, obtained by applying Engerer2 to hourly data [0–1]
k_t	G_h/E_0, clearness index [dimensionless]
k_{de}	$\max\left(0, 1 - \frac{G_{\mathrm{csky}}}{G_h}\right)$, part of the diffuse fraction that is attributed to cloud enhancement [dimensionless]
$k_{t,\mathrm{daily}}$	daily average of k_t [dimensionless]
$k_{t,\mathrm{hourly}}$	hourly average of k_t [dimensionless]
k_{tc}	G_{csky}/E_0, clearness index of clear-sky global horizontal irradiance [dimensionless]
Z	zenith angle [degree]

such as the often-cited Erbs model [8], which consists of a piecewise linear function of k_t determined empirically using hourly irradiance data. (Following the convention, model names are written in Small Caps, and in cases where multiple versions of the same model are available, a number is appended to distinguish them.) Over time, the modeling philosophy of separation models has advanced substantially, and it has become a commonly accepted fact that selecting auxiliary variables and performing feature engineering are both absolutely vital when constructing a high-performance separation model.

Another notable evolution is with respect to time scale. Initially, separation modeling was mostly conducted on hourly, daily, or even monthly data. However, considering the need for high accuracy and detail in current solar applications, the value of irradiance data at such coarse temporal resolutions now appears excessively limited, to the point where such models have become increasingly outdated or less useful. In contrast, those separation models for solar applications that have appeared since 2015 almost always utilize 1-min data. Compared to the older hourly time scale norm, this 60-fold increase in temporal resolution offers a considerable advantage of resolving high-frequency features, such as those caused by cloud-enhancement and/or albedo-enhancement events [9], which are not observable on an hourly basis.

Due to the combinatorial flexibility of input parameters as well as the source data used for model fitting and diagnosis, a large number of models with different degrees of sophistication and generalization abilities have been proposed, which has led to a heated debate on "what constitutes a good separation model?" Most certainly, insofar as predictive models are concerned, accuracy often comes as the foremost criterion of judgment. In the past, however, the conditions reported in one work always differed from those in another, in terms of location, time period, and modeling philosophy; hence, a fair comparison of model performance was not possible. The year 2016 marked a turning point, when Gueymard and Ruiz-Arias [10] compared a total of 140 separation models available then, using 1-min data from 54 research-grade radiometric stations spread across all seven continents and on islands in all four oceans. The scale of that work was unprecedented in terms of both the number of stations and the number of models compared. The conclusion made therein was that the Engerer2 model [11] could be considered quasi-universal because it attained the best overall predictive performance. Logically speaking, for any model proposed since, surpassing the performance of Engerer2 becomes a key criterion, which is why most separation models proposed post 2016 use Engerer2 as a benchmark.

Separation modeling is a fast-advancing field, and many new models have emerged since 2016, rendering the former question on the best separation model again opaque. For instance, Bright and Engerer [12] performed a re-parameterization of Engerer2 using more data points from more stations, and on more time scales, namely, 5, 10, 15, 30 min, 1 h, and 1 day. Besides, Starke et al. [13,14] proposed regime-dependent separation models under the framework of the BRL model [15], which employs a logistic function to account for the correspondence between k and k_t. A third example is the machine-learning-based model combination proposed by Aler et al. [16], who combined the predictions from the 140 models reviewed in Gueymard and Ruiz-Arias [10] with extreme gradient boosting as a regression tool. All these and other latest developments in the field, as well as the pros and cons associated with each innovation, have been summarized by Yang and Gueymard [17]. That review attempted to perform an unbiased comparison of the latest separation models using a common database, but the scale of validation was relatively small (i.e., only data from 11 sites spanning 1 year were used).

Two years after the publication of Yang and Gueymard [17], another larger-scale validation work was performed by Yang [18]. In that work, more than 80 million valid 1-min data points from a total of 126 sites worldwide, covering a period of 2016–2020, were used. The data were prepared as a joint effort of the members of the International Energy Agency (IEA), Photovoltaic Power Systems programme (PVPS), Task 16, which is an international collaborative research and development initiative established within the IEA and with its member

states. The 126-station database, which benefited from a thorough and state-of-the-art quality-control process [19], appears as the most comprehensive to date, and therefore must be regarded as rightly authoritative. Consequently, it can be stated that whichever model obtains the best overall statistical performance on this database can be considered quasi-universal. A total of ten recent separation models entered the contest, and the one proposed by Yang [20], namely, the Yang4 model, attained the highest rank. The modeling philosophy of Yang4 is based on the so-called "temporal-resolution cascade," which uses the preliminary k estimate from a low-resolution (i.e., 1 h) separation model as an input to the high-resolution (i.e., 1 min) separation model.

Despite its success, Yang4 has room for improvement. Firstly, its coefficients are fitted using data from only seven mid-latitude sites. Hence, they do not cover all the many other climatic or weather conditions outside the fitting data. As has been shown repeatedly, *conditioning* (also known as regime-dependent fitting) as a strategy is highly rewarding [14,21,22]. Conditioning means that different sets of model coefficients are empirically fitted using data corresponding to different regimes (e.g., climates), such that the fitted coefficients can best adapt to the condition-specific features embedded in the data. It is on this account that this work attempts to further improve on the previous Yang4 model by following this conditioning approach, which constitutes the main merit. In addition to that, the three-factor clustering method used to determine the radiation regimes is another novelty here. As discussed in Section 4.1, this regime-dependent modeling approach is indeed beneficial, making the new model outperform Yang4 by a significant margin, hence qualifying this development as a substantial contribution to the separation modeling literature.

2. Method

2.1. Input parameters

Before elaborating on the regime-dependent modification to Yang4, a short review of the previous development of the Yang family of models is first presented. To facilitate the discussion, some useful input parameters have been summarized in the nomenclature section. Among those parameters, the solar zenith angle (Z), solar elevation angle (α), and apparent solar time (AST) are calculated via a solar position algorithm; the clear-sky GHI (G_{csky}) can be obtained from the McClear database[1] or similar, and all remaining ones are calculable using GHI.

2.2. Engerer2 model

Engerer2 was historically the first model developed for high-resolution (1-min) irradiance data [11], taking the transient effects of cloud enhancement into consideration. More specifically, Engerer2 is a five-predictor model based on the logistic function, which attempts to reproduce the whole extent of the observed k–k_t space. The mathematical form of Engerer2 is given by:

$$k^{\mathrm{Engerer2}} = C + \frac{1-C}{1+e^{\beta_0+\beta_1 k_t+\beta_2 \mathrm{AST}+\beta_3 Z+\beta_4 \Delta k_{tc}}} + \beta_5 k_{de}, \tag{2}$$

where the model coefficients $C = 0.042336$, $\beta_0 = -3.7912$, $\beta_1 = 7.5479$, $\beta_2 = -0.010036$, $\beta_3 = 0.003148$, $\beta_4 = -5.3146$, and $\beta_5 = 1.7073$ were fitted with 1-min irradiance data from six radiometric stations in Australia. Variable k_{de} represents the fraction of k that is induced by cloud enhancement; this positive quantity is estimated as a simple function of k_{csi}, which is the ratio of GHI and clear-sky GHI. The other model inputs can be considered conventional because they have appeared in several previous models [e.g., 15,23].

[1] https://www.soda-pro.com/web-services/radiation/cams-mcclear

2.3. Evolution of the Yang family of models

Based on the construct of Engerer2, two separation models named Yang1 and Yang2 were proposed by Yang and Boland [24], who introduced the satellite-derived diffuse fraction, $k^{(s)}$, as an additional predictor. Yang1 includes $k^{(s)}$ as an additive trend component, whereas Yang2 adds $k^{(s)}$ as part of the main effect, i.e., inside the exponential term of the logistic function. Mathematically,

$$k^{\mathrm{Yang1}} = C + \frac{L}{1+e^{\beta_0+\beta_1 k_t+\beta_2 \mathrm{AST}+\beta_3 Z+\beta_4 \Delta k_{tc}}} + \beta_5 k_{de} + \beta_6 k^{(s)}, \tag{3}$$

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 $L = 0.6768$, and

$$k^{\mathrm{Yang2}} = C + \frac{1-C}{1+e^{\beta_0+\beta_1 k_t+\beta_2 \mathrm{AST}+\beta_3 Z+\beta_4 \Delta k_{tc}+\beta_6 k^{(s)}}} + \beta_5 k_{de}, \tag{4}$$

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 $\beta_6 = -2.8396$. The model coefficients of both Yang1 and Yang2 were fitted using data from seven stations in the United States.

Using half-hourly and hourly satellite-derived diffuse fraction was found to improve accuracy, for it can be regarded as a low-frequency version of the actual diffuse fraction. Conceptually, the strategy is highly similar to using a variability index in many former models, see [10] for a list. Furthermore, due to the worldwide availability of such data—see [25] for a review—satellite-augmented models can be applied at most locations, except for the high-latitude regions where high-resolution cloud information is missing because the images from geosynchronous satellites do not resolve. However, many satellite-derived irradiance databases are not updated in real-time (particularly those in the public domain). Hence, Yang1 and Yang2 are unsuitable for real-time applications, as is also the case with many other separation models that use time-averaged inputs.

To remedy the situation, [20] proposed replacing $k^{(s)}$ in Yang2 with a low-frequency estimate of diffuse fraction calculated from Engerer2, which led to Yang3 and Yang4. This modeling strategy is termed temporal-resolution cascade, for it uses sequentially two separation models at different temporal resolutions. Since Yang4 performs better than Yang3 and the only difference between them is the temporal resolution of the Engerer2-derived diffuse fraction, the latter one is not thoroughly discussed here. Denoting the hourly diffuse fraction estimate using Engerer2 as $k_{\mathrm{hourly}}^{\mathrm{Engerer2}}$, Yang4 reads:

$$k^{\mathrm{Yang4}} = C + \frac{1-C}{1+e^{\beta_0+\beta_1 k_t+\beta_2 \mathrm{AST}+\beta_3 Z+\beta_4 \Delta k_{tc}+\beta_6 k_{\mathrm{hourly}}^{\mathrm{Engerer2}}}} + \beta_5 k_{de}, \tag{5}$$

where all coefficients are inherited from Yang2 without modification. As mentioned in the introduction, Yang4 has been identified as the new quasi-universal model in the überreview and comparative validation conducted by Yang [18]. That said, a deficiency of Yang4 is that its model fitting is limited to stations in the contiguous United States, leaving room for improvement on the adequacy and universality of model coefficients [24]. To address this important concern, the regime-dependent version of Yang4, which is named Yang5, is introduced hence.

2.4. A new avenue: Regime-based model fitting

The basic idea of this work is to create a model with coefficients changing in accordance with different regimes. Whenever the phrase "climate regime" is mentioned in solar energy meteorology, it generally refers to the Köppen–Geiger climate classification (KGC); this is indeed the choice of Starke et al. [14], who fitted a separate set of model coefficients for each major class of KGC. However, defining regime with KGC may attract skepticism, since KGC is based primarily upon seasonal temperature and precipitation, which do not necessarily reflect the differences in the long-term statistical behavior of solar radiation, e.g., two locations with different KGC classes may share a radiation regime with

similar annual mean and variance. Since separation modeling deals with irradiance, rather than temperature or precipitation, it is herein argued that regime-dependent separation models based on KGC might be helpful to some extent, but overall insufficient if not inadequate. Consequently, alternative conditioning variables must be sought.

Whenever influencing factors to surface radiation are thought of, cloud properties must be regarded as the primary ones. Although high cirrus clouds are almost transparent to shortwave radiation, medium and low cumulus or stratus clouds attenuate or scatter most of the incoming radiation. That is why detailed information on cloud type, amount, and frequency is key to virtually all solar radiation modeling. Cloud processes are also intimately tied to the short-term variability of solar irradiance, which contributes most to modeling uncertainty. Indeed, various studies have documented the dependence on cloud cover of radiation modeling accuracy [e.g., 26–28]. For that reason, cloud cover frequency is selected as the first conditioning variable. It should be noted, however, that cloud cover frequency is only a very crude proxy of the actual cloud dynamics, which are exceedingly complex.

The second conditioning variable selected in this work is aerosol optical depth (AOD), which is the most relevant quantity influencing the surface radiation under a cloud-free atmosphere. Particularly for arid areas with low cloudiness but prevalent medium-AOD to high-AOD situations and frequent sand or dust storm episodes of extremely high AOD, the effect of aerosols is noteworthy. At present, the foremost representative application of aerosol data in solar energy meteorology ought to be clear-sky modeling—almost all physics-based high-performance clear-sky radiation models rely on various sorts of aerosol inputs [29,30]. In parallel, aerosols also act as cloud condensation nuclei and ice nuclei, which are key to the formation of clouds. That is, besides the primary effect of attenuation, aerosols also have a secondary (i.e., indirect) effect on radiation through affecting clouds.

The third conditioning variable considered here is surface albedo, which is the ratio of upwelling and downwelling global irradiance at the surface—a fraction between 0 and 1, see [31] for precise definitions. It is well known that a portion of the radiation reaching the Earth's surface is reflected back to the atmosphere, and through a process known as backscattering, a part of it returns to the surface. It is clear from simple observation and basic physics that the backscattering process is stronger when the ground surface is more reflective, which leads to the so-called "albedo enhancement" effect [31]. This phenomenon is particularly prominent under high cloudiness, which can lead to its combination with cloud enhancement effects, thus generating large variations in short-term irradiance. For instance, the observable impact on the diffuse fraction of situations of high surface albedo combined with high cloudiness was exemplified for a mountain site by Gueymard and Ruiz-Arias [10]. Considering the goal of regime-based separation modeling, the spatial variations in surface albedo should have a noticeable impact on modeling accuracy.

Besides cloud cover, aerosol, and surface albedo, there are many other meteorological variables that can be deemed useful in defining irradiance regimes. This is because ultimately, in meteorology, everything is related to everything. One can access a rich collection of meteorological variables from modern reanalysis products, so data availability is not of concern. However, as compared to cloud, aerosol, and surface albedo, the effects of other meteorological variables are mostly secondary or tertiary. For instance, although water vapor absorption can reduce clear-sky radiation by tens of W/m^2 [32], this reduction is smaller than that caused by clouds or aerosols. Additionally, the two most important indirect effects of water vapor on radiation, namely, cloud formation or modification of aerosol size [33], are also accounted for by clouds and aerosols themselves. Therefore, it is thought that the three selected variables are sufficient for the first attempt to create a radiation climate classification for separation modeling, although such a choice may not be optimal, and is worth investigating further in future works.

All three selected influencing factors vary in both space and time, and thus the corresponding data can be viewed as a time series of lattice processes. Theoretically, it is possible to not only segregate the meteorological regimes in terms of space, but also in terms of time, e.g., one set of model coefficients for each season, monsoon, or even month. This nonetheless greatly increases the dimensionality of the problem at hand. Recalling that previous regime-dependent separation models just use KGC, which is only distributed over spatial locations, the current work should follow that and consider only regimes defined over space but not time. In any case, a time-dependent classification would add much higher complexity to the model, as well as difficulty of operational implementation. For that reason, climatological values on an annual basis are deemed sufficient in this work—even though higher-frequency climatologies might also be worth exploring in the future. In summary, climatology maps of annual cloud cover frequency, aerosol optical depth, and surface albedo are used as raw data defining distinct radiation regimes, but they must first undergo some form of data fusion, such that the overall regimes are results of integrating and balancing of the information contained in all three factors. Clustering is chosen for this task. Clustering is a data-science technique that divides samples into groups in an unsupervised fashion such that the members within each group are more similar to one another than to members in other groups [34]. Once the model coefficients for a certain cluster are fitted, the coefficients are expected to perform well for other locations with that same regime.

Numerous clustering algorithms exist in the literature and their standard implementations in popular statistical software tools are available in bulk. For instance, the `NbClust` package of R implements several different clustering methods, among which the most prevalent k-means clustering [35] is considered. Since the k-means algorithm is well known and much information is available elsewhere, this work does not reiterate its technical details in full. One should just be aware that k-means clustering seeks to partition n samples into k ($k \leq n$) clusters, denoted by $\mathcal{S} = \{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_k\}$, in such a way that the within-cluster variance is minimized. Mathematically, the optimization problem is

$$\underset{\mathcal{S}}{\operatorname{argmin}} \sum_{i=1}^{k} \sum_{\mathbf{x} \in \mathcal{S}_i} \left\| \mathbf{x} - \boldsymbol{\mu}_i \right\|^2 = \underset{\mathcal{S}}{\operatorname{argmin}} \sum_{i=1}^{k} |\mathcal{S}_i| \mathbb{V}\left(\mathcal{S}_i\right), \tag{6}$$

where $\boldsymbol{x}$ denotes a sample, $\boldsymbol{\mu}_i$ is the center of the ith cluster (i.e., the mean of points in $\mathcal{S}_i$), $|\mathcal{S}_i|$ is the cardinality of set $\mathcal{S}_i$, and $\mathbb{V}$ is the variance operator. As an unsupervised algorithm, the number of clusters, namely, k, needs to be specified *ex ante*. Whereas it is obvious that the minimum number of clusters must be greater than one, too large a number of clusters leads to redundant granularity as well, so in this work, the allowed number of clusters varies between two and ten.

The literature presents an eclectic mix of methods for determining the optimal number of clusters. Most of these methods rely on some index, and the optimal number of clusters is arrived at according to either maximized or minimized metrics of the index (i.e., maximum value of the index or minimum value of second differences between levels of the index). In particular, the `NbClust` package provides a total of 26 indexes of that sort, but not all have gained wide acceptance and their reliability has yet to be fully tested. In this work, a total of seven indexes, namely, "Cindex," "Hartigan," "Ratkowsky," "Scott," "Friedman," "McClain," and "Rubin" are jointly used as the basis of the optimal cluster number selection, which is thoroughly explained in Section 4.1. Once k is fixed, one may proceed to fit the regime-dependent model coefficients. It should be noted that for any new location, its corresponding cluster can be determined by the trained k-means model—the cluster and the separation model coefficients can thus be identified for any arbitrary location.

2.5. Benchmarks and evaluation metrics

After the new model is constructed, its accuracy needs to be tested against measured data and also against other competitive models. To ensure a fair comparison, four representative separation models are considered as benchmarks. Since Engerer2 [11] was the best model before 2016, it must be taken into account. Starke2 is a piecewise model differentiating cloud enhancement conditions, and Starke3 [14], as an extension of Starke2, is a high-performance piecewise regime-dependent model with KGC as a conditioning variable, so they are included as benchmarks—the model coefficients of Skarke3 are taken from Table 3 of [14]. The non-regime-dependent version of the proposed model, namely, Yang4, is also included to evaluate the potential progress. Arguably, more models such as those in [10,18] could be tested, but since they do not outperform Engerer2 or Yang4 in the respective reviews, they can be assumed to be inferior to these two models, at least in a general sense.

The process of verification of solar irradiance models is a delicate undertaking, for which the reader is referred to Yang et al. [36] for a fully expanded discussion. For the sake of conciseness, the performance of various separation models under comparison is here simply gauged on the basis of two widely used error metrics, namely, the normalized root mean square error (nRMSE) and normalized mean bias error (nMBE), which, when expressed in percent, are:

$$\mathrm{nRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{t=1}^{n}\left[\widehat{B}_n(t) - B_n(t)\right]^2}}{\frac{1}{n}\sum_{t=1}^{n} B_n(t)} \times 100, \tag{7}$$

$$\mathrm{nMBE} = \frac{\sum_{t=1}^{n}\left[\widehat{B}_n(t) - B_n(t)\right]}{\sum_{t=1}^{n} B_n(t)} \times 100, \tag{8}$$

where $\widehat{B}_n(t)$ and $B_n(t)$ are predicted and measured BNI at instance t, which indexes the testing samples. Similar metrics for DHI can be defined by replacing $\widehat{B}_n(t)$ and $B_n(t)$ with $\widehat{D}_h(t)$ and $D_h(t)$ respectively.

Historically, mean absolute error (MAE) is also often used to evaluate radiation forecasts and predictions. It has been shown recently that using both RMSE and MAE is inappropriate [5,37,38], for these two metrics are incompatible with each other under the framework of statistical consistency in verification. Stated simply, the evaluation metric should follow the choice of objective function that is used to optimize the model parameters. For instance, if a model is optimized by minimizing the squared loss, RMSE is consistent, whereas for an MAE-optimized model, MAE is consistent. Given the fact that the model coefficients for all separation models of concern are RMSE-optimized, using MAE as an evaluation metric is redundant and incorrect. The same argument extends to other popular error metrics for solar modeling. The reader is referred to Gneiting [39], who delivered the seminal paper that first discusses the consistency issue theoretically.

3. Data

Four different datasets are used in this work. The first one is a proprietary irradiance dataset from the IEA PVPS Task 16, and is identical to the one used by Yang [18]. It consists of 1-min ground-based irradiance measurements from 126 locations worldwide, as necessary for separation model development and validation. The remaining three datasets provide climatology maps for annual cloud cover frequency, aerosol optical depth, and surface albedo, which are used here to facilitate the cluster analysis, as well as to allow identification of local regimes for any unseen location of interest.

3.1. Ground-based irradiance measurements

The ground-based measurement dataset used in this work has been thoroughly described by Yang [18]. In short, the raw dataset collected by members of the IEA PVPS Task 16 is quality-controlled with a set of stringent filters (see [19] for details), after which more than 80 million valid 1-min data instances remain; each instance comprises three irradiance components in the closure equation, namely, G_h, B_n, and D_h, as well as auxiliary variables needed for separation modeling, such as Z, G_{csky}, or AST. Fig. 1 depicts the geographical distribution of these 126 sites, which are overlaid on the world map of Köppen–Geiger climate classification. Except for a few cold climates, the sites have good coverage for each climate zone.

Through clustering, the 126 stations are grouped into five clusters, of which the details are given in Section 4.1. For each cluster, all data instances from all stations within that cluster are gathered into a data table, which is subsequently divided into two halves, one for training and the other for validation. The splitting of data frames adopts random sampling without replacement. In usual radiation modeling and validation, a part of the data should contain "unseen" instances, such as to demonstrate and test for the generalization ability of the model of concern. Notwithstanding, considering that the current dataset already has the best possible inclusion of the currently available research-grade radiometric stations, the universality of any model fitted and successfully validated using the current dataset can be assumed with high confidence.

3.2. Cloud climatology

The cloud climatology data employed in this work comes from Wilson and Jetz [40], who integrated 15 years of remote-sensed cloud observations into climatology maps of 1-km-resolution cloud cover frequency. The remote-sensed cloud information was originally retrieved from twice-daily Moderate Resolution Imaging Spectroradiometer (MODIS) satellite imagery, which reveals cloud dynamics on a global scale through cloud frequencies (i.e., fraction of days during a month with a positive cloud flag). Wilson and Jetz [40] showed that this MODIS-based cloud frequency (MODCF) explains 78% of the variability in monthly mean cloud frequencies observed on the ground, with only a 7.99% RMSE between the satellite and ground station cloud data.

The dataset can be accessed for free from the EarthEnv project,[2] which is a collaborative effort intended to construct a database of standardized 1-km resolution data layers for environmental and climatological research. The MODCF dataset can be obtained either as monthly files or as a single file containing the annual mean, in `GeoTIFF` format, and the latter is used in this work. The original mean annual cloud frequency represents the percentage of cloudy days over the 15-year period, and the highest value is found to be 98.28%.

It is emphasized that all the three climatological datasets used in this work, namely, MODCF, the aerosol (Section 3.3), and albedo (Section 3.4) datasets, have differing spatial resolution. Hence, for consistency and ease of manipulation, it is of interest to standardize all of them onto the same grid. A common $0.5° \times 0.5°$ regular grid over latitude and longitude is selected here as a compromise between spatial resolution and data storage. The mapping from the original to the new standard grid uses the nearest neighbor interpolation technique with an ocean mask to remove regions with improbable solar applications at a $0.5°$ spatial resolution, and only excludes islands smaller than about 1 km^2, thus ensuring all 126 sites have climatology data. The mapping grid follows `kgc` package in R and the reader is referred to the University of Veterinary Medicine website[3] for detailed information.

[2] http://www.earthenv.org/cloud
[3] http://koeppen-geiger.vu-wien.ac.at/present.htm

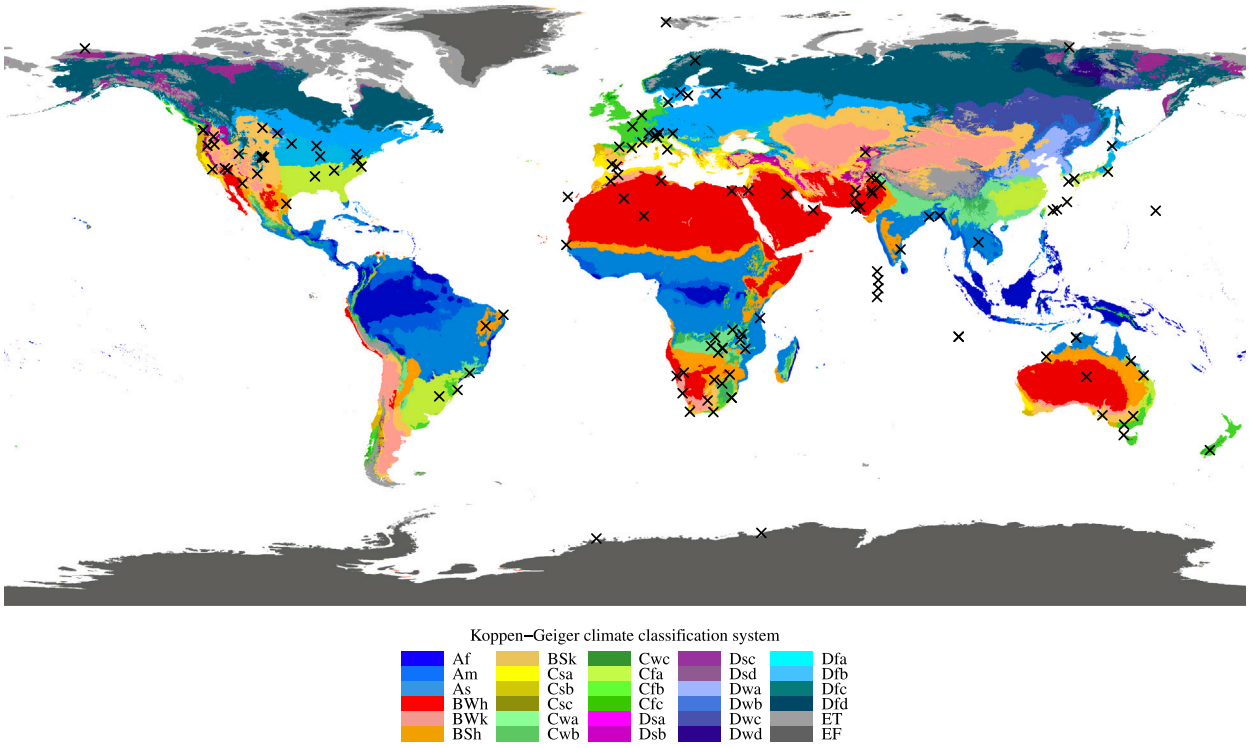

Fig. 1. Location of the 126 ground-based stations used in this work, superimposed on a world map of Köppen–Geiger climate classification.

Fig. 2 (a) shows the annual mean cloud frequency, where the color of a pixel corresponds to how often weather is cloudy at that location in a year. Since the *k*-means algorithm requires its different input variables to be normalized, which constitutes a fundamental practice in data science, the cloud cover frequency is normalized to the $[0, 1]$ range, using the min–max normalization. The final cloud climatology map is shown in Fig. 2 (b), where darker pixels correspond to locations that are less cloudy.

3.3. Aerosol climatology

The aerosol climatology dataset used in this work results from a combination of two products, namely, the climatology developed by Yang and Gueymard [41] (hereafter, "Yang–Gueymard") and the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) [42]. Yang–Gueymard is a monthly AOD climatology obtained by merging five gridded AOD products, namely, MERRA-2, Multi-angle Imaging SpectroRadiometer, MODIS-Terra, MODIS-Aqua, and Visible Infrared Imaging Radiometer Suite; the merging procedure is complex, and the reader is referred to the original publication for technical details. It should be noted that data fusion is an important aspect of any modeling procedure that involves AOD. This is because each raw remote-sensed AOD dataset contains missing (or inaccurate) values due to difficulty in retrieval, but when multiple datasets are combined, missing (or inaccurate) data points from one product can be filled (or compensated) by those valid data points from another product, thereby leading to a final product with far fewer missing values and of higher quality than any single product.

The main AOD variable of interest here is the AOD at 550 nm (AOD550), which is the standard in solar energy meteorology and other disciplines. Yang–Gueymard contains monthly AOD550 climatology values over the years 2012–2020, at a $1° \times 1°$ spatial resolution. One limitation of the Yang–Gueymard climatology is that it only covers regions between $\pm 60°$ in latitudes. The year-long snow cover at high-latitude regions, such as Antarctica or Greenland, makes the AOD retrievals highly inaccurate there. Hence, one has to resort to using

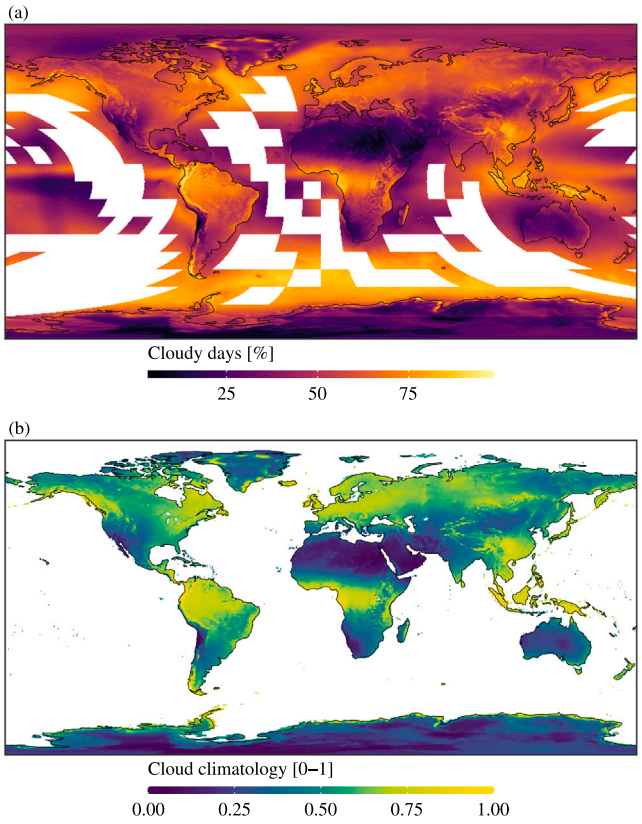

Fig. 2. (a) Map of the annual mean frequency of cloudy days from MODIS-based cloud frequency (MODCF), where "cloudy days [%]" means the percentage of cloudy days at a location, with 0 representing never cloudy and 1 meaning always cloudy. The MODCF database contains missing data, as represented by the white patches in the top figure. (b) A normalized version of the cloud climatology, with a range of $[0, 1]$ and excluding oceans.

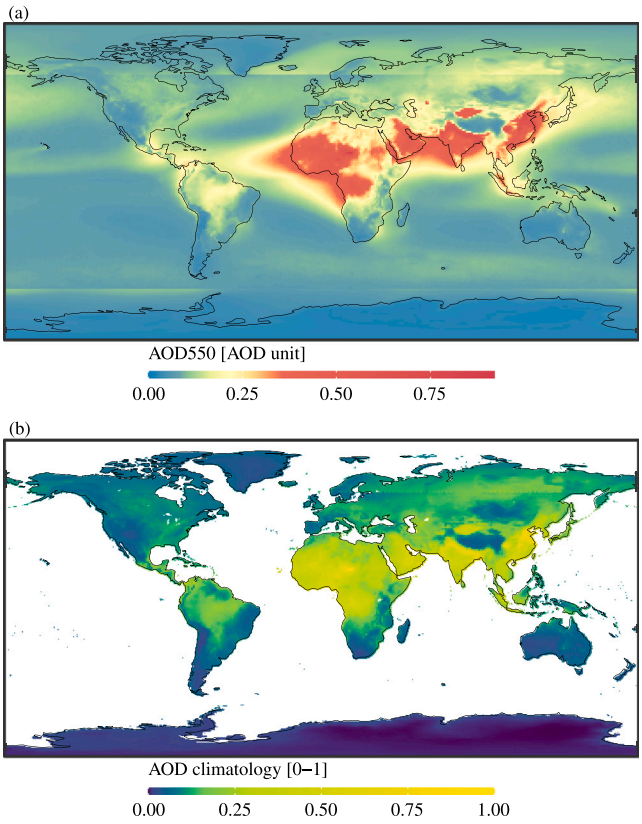

Fig. 3. (a) Map of the global climatological aerosol optical depth (AOD) by Yang–Gueymard (for locations within ±60° in latitudes) and MERRA-2 (for locations beyond ±60° in latitudes). (b) Regridded and normalized AOD climatology map with a range of $[0, 1]$.

modeled reanalysis data for those regions. MERRA-2 is selected here for that purpose, considering its relatively good accuracy globally [43]. Insofar as the desired AOD climatology is concerned, MERRA-2 has been validated to possess decent accuracy [42]. In this work, MERRA-2 monthly mean AOD550 data (also known as the "M2TMNXAER 5.12.4" product), over 2016–2020, with a $0.5° \times 0.625°$ spatial resolution, is obtained from the Goddard Earth Sciences Data & Information Services Center.[4] It is stressed that, by design, the AOD550 data obtained from the Yang–Gueymard climatology is more accurate than that from MERRA-2, which is why this combination of two products is more advantageous than solely using MERRA-2.

Yang–Gueymard and MERRA-2 are individually time-averaged into a single, long-term climatology map because of their difference in temporal coverage. Since their spatial resolutions also differ, both long-term climatology maps are again regridded onto a $0.5° \times 0.5°$ latitude–longitude grid, for consistency with the cloud and albedo climatology maps. Fig. 3 (a) shows the raw AOD550 climatology used in this work; the artifacts at ±60° are clearly visible due to the stitching of the two databases over an area where they are both uncertain. The normalized AOD climatology map is shown in Fig. 3 (b), in which pixels with darker colors correspond to low AOD locations.

3.4. Albedo climatology

The surface albedo climatology dataset is based on the fifth-generation ECMWF Reanalysis (ERA5) [44]—the latest-generation global reanalysis produced by the European Centre of Medium-Range Weather Forecasts (ECMWF). As compared to other albedo products based on remote sensing, such as those derived from MODIS, ERA5's albedo has lower spatial resolution ($0.25° \times 0.25°$ in latitude and longitude), and possible lower accuracy. Nevertheless, it inherits the advantage of reanalysis products, in that the data is spatio-temporally complete, which is a critical feature for the intended application here.

The data can be acquired from the ECMWF's Climate Data Store,[5] which is a centralized database containing various gridded products of ECMWF. The product family containing the monthly albedo climatology is named "ERA5 monthly averaged data on single levels from 1959 to present." The Climate Data Store offers an online user interface, from which the user can select the variable, location, time period, and file format. In this work, the variable called "forecast albedo" is selected, over the entire globe, covering a period of 2016–2020, to echo the span of the ground-based data, in `NetCDF` format. Since the dataset is spatio-temporally complete, the processing only requires regridding (onto the standardized $0.5° \times 0.5°$ grid adopted here) and normalization to $[0, 1]$. This procedure is consistent with those followed for clouds and aerosols. Fig. 4 shows the raw and processed albedo climatology used in this work, and darker colors denote smaller albedo.

At this stage, all three climatology variables have been introduced, regridded, normalized, and thus made ready for clustering. However, it is noted that the cloud climatology is the annual average of 15 years from 2001 to 2015, and the other two variables, namely aerosol and albedo, are averaged over the years 2016 to 2020. This temporal incompatibility would certainly have an effect on the results of clustering, and temporally aligned data should be preferred in general, i.e., using a cloud climatology over 2016–2020. However, the rule of thumb in selecting input data is to use better-quality data whenever they are available, and it is on this account that MODCF should be preferred. The same argument can be used to justify the choice of merging Yang–Gueymard and MERRA-2 AOD, instead of using just MERRA-2 AOD. Another drawback is the temporal coverage of data, because good climatology should be derived from data spanning multiple decades. That said, since the climatology variables are processed into a single "snapshot," which is similar in concept to the Köppen–Geiger climate classification map, the intra- or inter-annual variability only impacts clustering very marginally.

4. Results and discussion

4.1. Clustering results

The data preprocessing procedure results in regridded and normalized versions of the selected climatology variables, which is a $92,422 \times 3$ matrix, upon which the clustering depends. However, instead of using all these locations for clustering, the clustering is performed with just those pixels that collocate with ground-based stations. There are two compelling reasons for this choice. Firstly, if worldwide locations are used, such a data dimension is not conducive to the k-means algorithm, as it implies a $92,422 \times 92,422$ distance matrix, which is beyond the memory limit of regular computers. Although other clustering methods that can handle big data are available, there is a second reason preventing the use of worldwide locations for clustering. As the number of locations escalates, the optimal number of clusters is likely to become more numerous. As such, some clusters may not contain a sufficient number of stations or any station at all, and the fitting for separation model coefficients for those clusters is not possible. For these reasons, this work uses just the climatology variables at those pixels containing ground-based stations.

More specifically, those gridded values of the three climatology variables that collocate with the radiometric stations form a matrix

[4] https://disc.gsfc.nasa.gov/, one can search "MERRA-2 tavgM_2d_aer_Nx" from the website to find the exact product.

[5] https://cds.climate.copernicus.eu/cdsapp#!/home

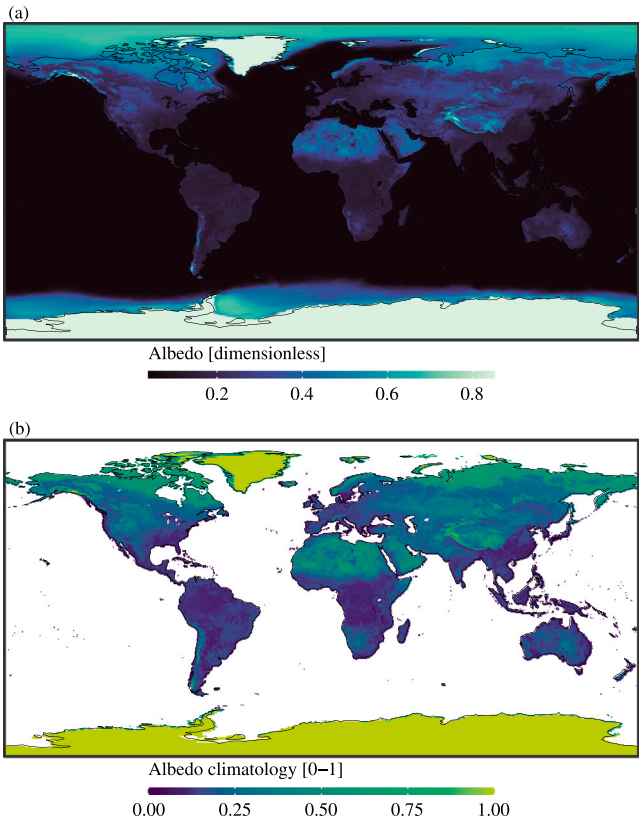

Fig. 4. (a) Climatology map of the ERA5 albedo, which is an aggregate of 5 years (2016–2020) of albedo, at 0.25°×0.25° resolution. (b) Regridded and normalized albedo climatology on a 0.5° × 0.5° grid, in the [0, 1] range.

of dimension 126 × 3, which serves as the input of the clustering algorithm. Its output is a set of integers representing the cluster to which every station belongs. One preliminary step, however, consists in specifying the cluster number such that the *k*-means algorithm can assign cluster centers according to that number. It is customary to choose the cluster number beforehand, taking into account that the optimal selection is often elicited from a mix of indexes provided by `NbClust` as mentioned before. Clearly then, the problem here is one of balance—choosing the "centroid" of them might be the most sensible solution.

As an illustration, Fig. 5 depicts the magnitude variation of seven indexes, when the number of clusters is increased from two to ten. It is clear that the divergence among the indexes is the smallest when the number of clusters is six, which would normally indicate the optimum number of clusters. Upon further verification, however, it is found that using six clusters actually leads to unbalanced clustering results: Some clusters contain tens of stations, whereas the smallest cluster contains just three stations. This can be problematic if out-of-sample prediction is to be carried out with new data. Therefore, considering this empirical pitfall, the number of clusters is decreased from six to five, which results in the desired more uniform clusters.

Fig. 6 (a) shows the *k*-means clustering results, visualized in the principal-component space. More specifically, the three-dimensional input data points (i.e., cloud, aerosol, and surface albedo) are projected onto the space spanned by the first two principal components. Fig. 6 (b) shows the same samples in the original three-dimensional space. In the left subfigure, Clusters 4 and 5 have a considerable amount of overlap, but this does not mean bad clustering—from the right subfigure, the two clusters appear to be well separated. The other three clusters have fewer samples than Clusters 4 and 5, and the within-cluster distances of the samples are also larger than those of Clusters 4 and 5,

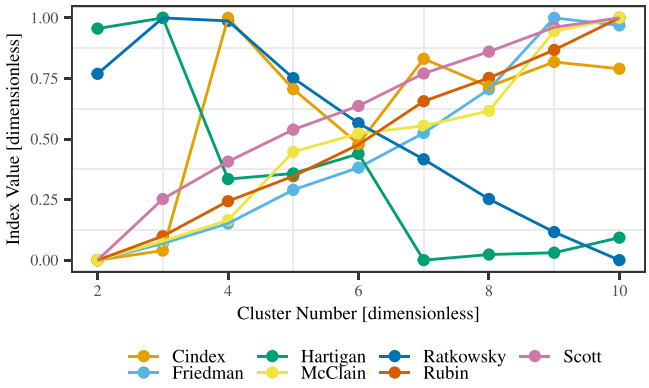

Fig. 5. Variation trend of normalized cluster indexes in NbClust.

which suggests the uneven spatial distribution of cloud–aerosol–albedo climatology.

As in the case with all machine-learning algorithms, the *k*-means clustering also allows prediction for new data: Once the model is fitted, one can obtain the cluster number of a new location by feeding the values of the three climatology variables at that place. Since the climatology variables are available globally, a new regime-classification map based on the three variables can be developed. After going through every point in the regridded lattice, a map depicting the worldwide cluster number can be drawn. Fig. 7 shows the so-called "radiation climate classification" (RCC) based on the three climatology variables. To be consistent with how KGC is developed, the oceanic pixels in the RCC map are omitted except for islands with significant land area.

To better understand the clustering result, one can compare the RCC map to the "Blue Marble" from NASA,[6] which provides an observation on the land surface, oceans, sea ice, and clouds distribution of the entire Earth from space. One can find in Fig. 7 that Cluster 2, in light blue, covers most of the Middle East and North Africa countries, which are characterized by relatively high surface albedo (desert), low cloud cover frequency, and medium to high AOD due to frequent dust episodes. Cluster 3, in green, occupies high-latitude regions, where the surface albedo is particularly high because of the extensive snow and ice coverage. It is also clear that the areas characterized by high AOD and cloud cover frequency—Cluster 1, in orange—tend to be geographically adjacent to Cluster 2 around the Sahara and the Taklamakan deserts. Lastly, in the mid-latitude areas, Clusters 4 (in yellow) and 5 (in dark blue) are often found adjacent to each other, where Cluster 5 is associated with large forests (e.g., the Amazon) and Cluster 4 with arid areas next to these forests (e.g., most of Australia). It is noted that clusters appear with irregular shapes in the middle of Asia, which appears to result from the local geomorphic complexity and diversity. Using Fig. 7, users are able to search for the corresponding RCC for any location of interest, and thus perform separation modeling using Yang5 for worldwide locations.

4.2. Performance evaluation

Ground-based irradiance data from each of the 126 sites are randomly split into two halves, as to be used for training and validation, respectively. For each cluster, data points from all stations within that cluster are lumped into overall training and validation sets. Since the new Yang5 model follows the same function form as Yang4, the least-squares fitting process is also inherited from its predecessor. The regime-dependent coefficients of the Yang5 are tabulated in Table 1. For comparison purposes, the coefficients of Yang4 are given in Table 2.

[6] https://visibleearth.nasa.gov/images/57752

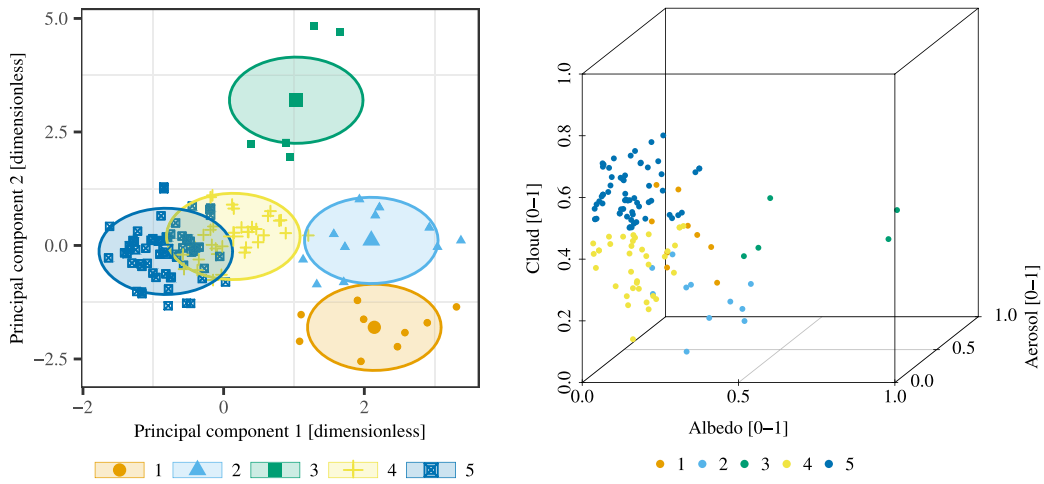

Fig. 6. Clustering of 126 stations by k-means. The left panel shows a 2D view of the clustering result, where stations from different clusters are marked with different colors and shapes. The right panel shows the 3D version of the clustering result.

Table 1
Model coefficients of the regime-switching Yang5.

Cluster	C	β_0	β_1	β_2	β_3	β_4	β_5	β_6
1	0.13105	−4.26740	7.68051	0.00540	0.01748	0.91590	0.52176	−1.68819
2	−0.01614	−3.33038	5.72307	0.01296	0.01230	−0.96483	0.94204	−1.68332
3	−0.27475	0.36085	0.39869	0.00479	0.00039	−10.20264	2.12475	−1.78455
4	−0.01095	−0.92129	3.65015	0.00767	0.00494	−3.76465	1.36482	−2.11867
5	0.04297	−1.64437	4.71808	0.01462	0.00745	−3.35223	1.25192	−2.36477

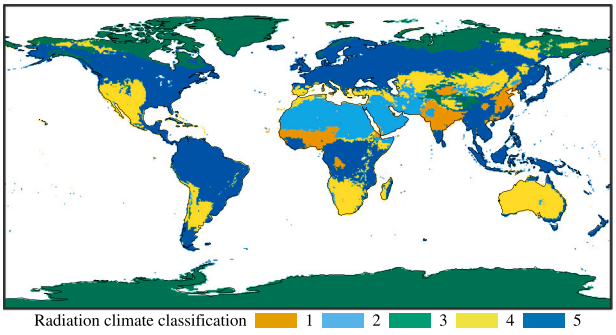

Fig. 7. Global radiation climate classification based on cloud, aerosol, and surface albedo climatology.

One may notice that there are some differences in magnitudes and signs of the model coefficients across regimes. Firstly, the magnitudes of model coefficients reflect the relative importance of each predictor as to predicting the diffuse fraction. As the predictor variables under different regimes exhibit different characteristics or profiles, it is reasonable for their relative importance *vis-à-vis* separation modeling to vary across regimes. Secondly, the occasional inconsistency in signs is thought to be purely mathematical, in that, the nonlinear least squares routine only minimizes the sum of squares but does not place any constraints on the sign or bounds of the coefficients. Since the routine converged for all five fitting exercises, the coefficients are guaranteed optimal with respect to the fitting data. More importantly, because the model is empirical in nature, making too many hypotheses, as to trying to understand why the coefficients behave in certain ways through physics, might not be meaningful. A more straightforward means to check the validity of these coefficients is through performance evaluation.

Yang5 needs to be evaluated against its peers. To ensure a fair comparison, the validation procedure of Yang [18] is followed precisely. In particular, the procedure is three-fold: (1) computing the error metrics, (2) model inter-comparison with linear ranking statistics, and (3) model inter-comparison through the Diebold–Mariano (DM) test. Given the large number of sites, tabulating all error metrics at each site would be inefficient. Hence, only the lumped statistics (i.e., nRMSE and nMBE), for all validation data points in each cluster, are displayed in Tables 3 and 4 for B_n and D_h, respectively. Interestingly, the overall nMBE of Yang5-estimated B_n in any cluster is below ±4%, which is unprecedented. This bias reduction is of substantial importance to solar resource assessment, which values bias more than accuracy. Moreover, as compared to Yang4 and other benchmark models, it is noteworthy that the nRMSEs of Yang5 for both B_n and D_h are significantly lowered in Clusters 1, 2, and 3, i.e., over regions with high cloudiness, high aerosol load, and high surface albedo, respectively. This empirically confirms that the proposed three-factor regime-dependent approach is sound.

The linear ranking method is applied to calculate the *mean rank* of the ith model, which is denoted as m_i:

$$m_i = \sum_{j=1}^{5!} \frac{n_j v_j(i)}{n}, \tag{9}$$

where v_j with $j = 1, 2, \ldots, 5!$ represents all possible rankings of the five models; n_j is the frequency of occurrence of the ranking j; $n = \sum_{j=1}^{5!} n_j$ 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 convention is used, which means that a better model receives 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 $v_j(i) = 5$. The ranking is based upon nRMSE, and the smaller the nRMSE of a model is, the higher the ranking of that model and the lower its v_j are. Table 5 shows the ranking results of the separation models for the D_h predictions, and Table 6 shows similar results for the B_n predictions. Each column in those tables corresponds to the ranking at a particular site; for conciseness, only a few columns (i.e., sites) are printed here with the rest omitted. The mean rank is calculated through Eq. (9). Among these models, Starke2 performs the worst, followed by Engerer2 and Starke3. This is somewhat expected, as both Starke2 and Engerer2 were fitted using local data, which can

Table 2
Model coefficients of Yang4.

C	β_0	β_1	β_2	β_3	β_4	β_5	β_6
0.03610	−0.57440	4.31842	−0.00112	0.00037	−4.79520	1.44143	−2.83961

Table 3
Cluster-wise normalized RMSE and MBE, both in percent, of five separation models. The column "mean" shows the mean B_n in W/m^2 for all data points in a cluster.

Cluster	Mean	Engerer2	Starke2	Starke3	Yang4	Yang5
		nRMSE [%]				
1	379.7	27.2	29.8	23.2	28.3	20.3
2	618.4	18.3	22.9	19.7	19.7	16.4
3	596.4	26.7	32.5	26.6	26.4	23.0
4	634.2	15.4	15.0	14.4	13.7	13.1
5	477.0	22.5	22.6	21.2	20.3	19.5
Overall	533.3	20.6	21.3	19.4	19.1	17.5
		nMBE [%]				
1	379.7	11.2	17.1	2.2	12.0	3.5
2	618.4	4.9	14.9	9.5	7.6	2.2
3	596.4	6.6	18.2	5.3	8.4	−3.4
4	634.2	−3.4	5.1	0.4	−1.7	0.1
5	477.0	1.0	9.3	−0.7	1.1	1.4
Overall	533.3	1.0	9.5	0.9	1.9	1.1

Table 4
Same as Table 3, but for D_h.

Cluster	Mean	Engerer2	Starke2	Starke3	Yang4	Yang5
		nRMSE [%]				
1	227.8	30.7	33.2	24.9	32.2	22.5
2	171.9	40.4	49.1	41.6	42.9	36.2
3	125.6	49.8	60.6	52.9	49.8	46.8
4	147.2	42.2	40.8	38.2	36.3	35.7
5	179.5	36.6	36.0	33.9	32.0	31.2
Overall	170.7	38.7	39.3	35.9	34.9	32.9
		nMBE [%]				
1	227.8	−10.9	−17.1	−1.1	−13.2	−1.0
2	171.9	−8.6	−29.7	−18.9	−16.2	−1.0
3	125.6	−8.8	−32.4	−6.6	−13.8	10.9
4	147.2	9.5	−12.5	−0.7	3.9	1.1
5	179.5	0.6	−12.6	3.8	−0.6	0.6
Overall	170.7	1.2	−15.2	−0.2	−2.0	0.9

Table 5
Ranking results of five separation models, based on the root mean square error of D_h estimates, at 126 sites. For each site, the best model is ranked "1," and the worst model is ranked "5." The middle columns for additional sites are omitted. The last column shows the mean rank of each model. A smaller rank indicates better performance.

Model	Station					
	1	2	3	⋯	126	Mean rank
Engerer2	4	2	4	⋯	5	4.16
Starke2	5	5	5	⋯	4	4.17
Starke3	3	3	3	⋯	1	2.58
Yang4	1	4	2	⋯	3	2.47
Yang5	2	1	1	⋯	2	1.62

impede their performance elsewhere. In contrast, Yang5 obtains the highest rank, beating its predecessor, Yang4, by a significant margin.

To give a visual representation of the predictive performance of the models, Fig. 8 shows the classic k–k_t plot at one of the stations, namely, the KWA station (−29.871°S, 30.977°E). Whereas the measurements (contained in the validation dataset at that site) are represented by the gray background, the viridis-colored scatter represents the predictions made by various models. A good separation model should cover the gray background as extensively as possible. Based on the relative k–k_t coverage, it can be concluded that Yang5 and Yang4 are able to "explain" more cases, followed by the two Starke models, and finally

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

Model	Station					
	1	2	3	⋯	126	Mean rank
Engerer2	4	2	4	⋯	5	3.98
Starke2	5	5	5	⋯	4	4.13
Starke3	3	4	3	⋯	2	2.90
Yang4	1	3	2	⋯	3	2.61
Yang5	2	1	1	⋯	1	1.37

Engerer2. Comparing the two Yang models more specifically, it can be observed that: (1) Yang5 extends more to the left as compared to Yang4, indicating the former is able to predict a smaller clearness index; and (2) the overall point cloud and the high-count-cloud is more dispersed for Yang5, which is a result of the regime-switching.

To better visualize the disparity between models, the DM test [45]—a statistical test for comparing the predictive accuracy of two models—is appropriate, as also demonstrated in previous works in which radiation models were compared [46–48]. The details of carrying out DM tests were thoroughly discussed by Yang [18], and thus are not reiterated. Figs. 9 (a) and (b) show the results in terms of B_n and D_h predictions, respectively. The number in each cell is the number of instances the DM test statistic falls in the lower or upper 2.5% tail of a standard normal distribution. Stated differently, it provides the number of "Model A is better than Model B" instances. For example, in the lower-right corner, Yang5 performs significantly better than Engerer2 at 122 out of 126 stations in terms of B_n. Since Yang5 has the largest number of "wins," it can again be considered the best model overall according to that criterion.

Given the ranking statistics and DM tests, it is still of interest to examine the distribution of error metrics, which is more intuitive. Tukey's boxplots of two error metrics, five clusters, and two variables (B_n and D_h) are displayed in Fig. 10. One can readily see that the median errors of Yang5, which are marked by the middle bar in each box, are often the lowest in terms of nRMSE (or closest to zero in terms of nMBE). Moreover, for Clusters 1 and 2, Yang5 presents a large advantage in predicting both B_n and D_h, which is consistent with the previous cluster distribution results. Therefore, this visual assessment confirms the conclusions from the ranking statistics and DM tests earlier, suggesting with high confidence that Yang5 performs better than Yang4, which completes the validation part of the work.

4.3. Methodological limitation

Yang5 performs best among all selected benchmarks, but, as with any model, there is potential for further improvement. When performing regime classification, this work uses a simple clustering method without weights, so the importance of the three climatology variables is assumed to be uniform, which might not be the optimal solution. On top of that, other means of enhancing the clustering, such as using an algorithm that can handle bigger data volume or including additional climatology variables,could be considered. That said, the RCC map depicted in Fig. 7 clearly resembles the geographical features of the "Blue Marble" of NASA. As such, the current clustering approach is admissible, as also evidenced by the superior performance of Yang5 than all other separation models to date.

Moving beyond the clustering technique, which is only of second-order significance to this work, the new model follows the form and modeling philosophy of Engerer2, which is mostly data-driven and marginally physical (i.e., by factoring the effects of cloud-enhancement

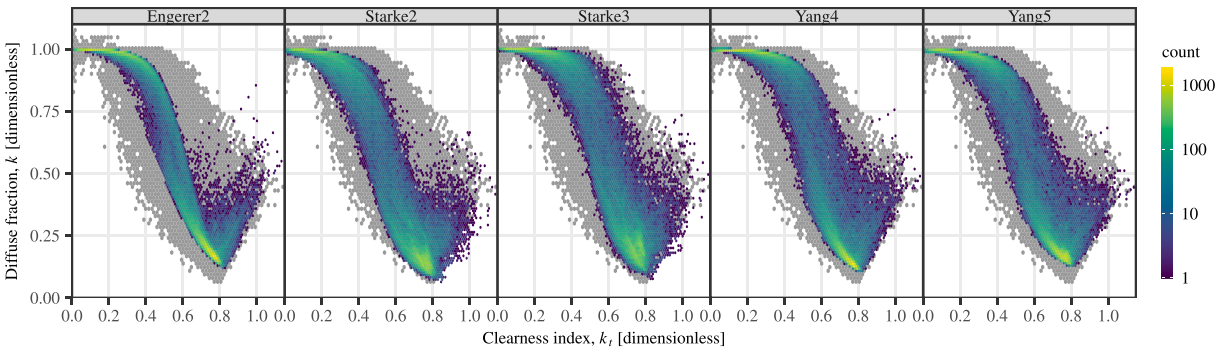

Fig. 8. Diffuse fraction measurements (gray background) at the KWA station (−29.871°S, 30.977°E), overlaid with the prediction results (viridis colors) of various separation models.

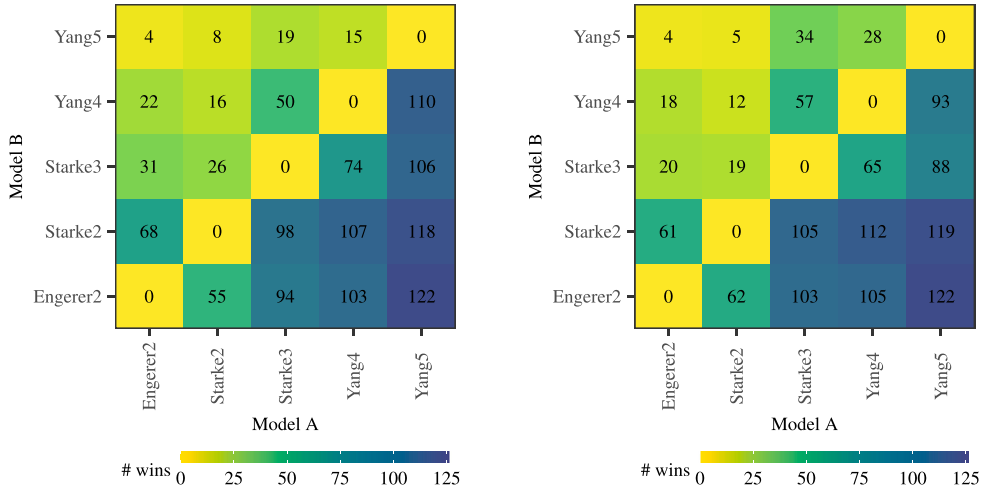

Fig. 9. Pairwise Diebold–Mariano tests for comparing the predictive accuracy of various separation models for B_n (left) and D_h (right).

events). It should be noted that separation models are often used in concert with satellite-derived GHI. However, deriving diffuse and beam radiation components from remote-sensing data can also be performed in a purely physical fashion through radiative transfer. How the empirical separation models compare to the physical derivation is largely unclear at the moment, which prompts future research in that direction.

When fitting the coefficients of Yang5, the sample sizes vary among stations because of the differing quality of the original data from the IEA PVPS database. After quality control, some sites have substantially more valid data points than others. Consequently, the sample size may exert an effect on the fitting, in that, stations with more samples contribute more towards the final coefficients. As a remedy, one may subset the same number of samples from each station, thus guaranteeing the equal contribution of each station, but that implies a smaller fitting dataset, which may or may not be beneficial in the end. Weighting serves as another alternative, but weighted nonlinear least squares is a method that goes beyond the scope of this work.

Finally, the model could be made even more complex by considering the seasonal variations in cloud, aerosol, and surface albedo, which are known to be large over most regions. This extension, however, would multiply the number of coefficient sets in Table 1, which might end up being overkill if the resulting gain in accuracy happens to be small.

5. Conclusion

Separation models are empirical functions that are needed in many applications to derive the diffuse fraction from the clearness index and other auxiliary variables, whenever directly measuring the direct or diffuse irradiance component is infeasible. Over the years, many such models were created with various parameters and variables. The Yang4 model was found to be of quasi-universal applicability and the most accurate model until this work. As an upgraded version of that predecessor, the proposed Yang5 model introduces regime-dependent coefficients through clustering the climatology maps of three variables, while preserving the temporal-resolution cascade characteristic from Yang4. Through error calculation, statistical linear-ranking analysis, Diebold–Mariano tests for comparing predictive accuracy, as well as visual inspection, Yang5 is found to outperform four other selected benchmarks on extensive 1-min irradiance data, which marks its superiority and general applicability. In particular, for the 126 radiometric stations used in this work, the overall nRMSEs of Yang5 for BNI and DHI are 17.5% and 32.9%, which are significantly lower than those of Yang4 and three other high-performance separation models of the recent literature. With the new model, BNI can be expected to be estimated with a small bias of less than 4% at any site, which is unprecedented. Compared to all previous models, the three radiation climate clusters with high cloudiness, high aerosol load, or high surface albedo are those for which the overall nRMSE is lowered the most, thus confirming the validity of the regime-dependent approach.

The development of Yang5 integrates several innovative separation modeling approaches in the literature. First, it resumes the logistic function shape, which gave rise to several high-performance models. Next, it considers, as per the Engerer2 model, the cloud enhancement as an important predictor, which improves the prediction at instances where k_t is high. Last but not least, Yang5 improves Yang4 by introducing three climatology quantities that have a significant impact on solar radiation, which are then subjected to dimensionality reduction through cluster analysis. Since Yang5 has a dominating performance over Yang4, which was deemed to be the quasi-universal model when it was first proposed, Yang5 should be the new quasi-universal model with

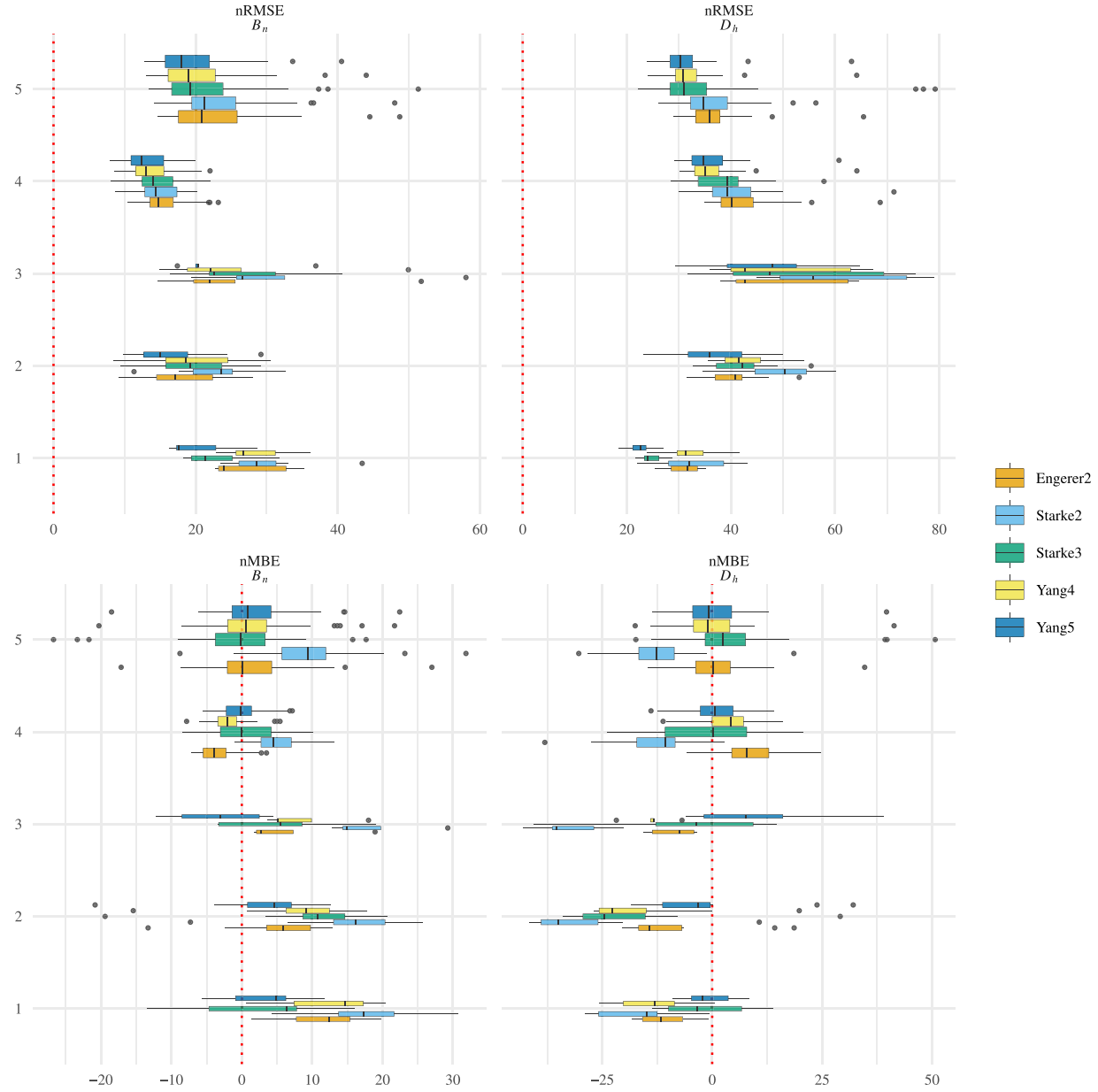

Fig. 10. nRMSE [%] and nMBE [%] of B_n and D_h estimates from five selected separation models, namely, Engerer2, Starke2, Starke3, Yang4, and Yang5. Tukey's boxplots are used for visualization. Dots beyond the ends of whiskers indicate outliers. The evaluation is grouped by clustering based on cloud, aerosol, and surface albedo climatology. The height of the boxes is proportional to the number of stations.

the overall best performance in the world. That said, even though Yang5 can efficiently predict diffuse fractions, empirical models have limited potential for improvement in comparison to physical models, which are thought to have *a priori* advantage. Besides physical modeling, another conspicuous alternative is to leverage machine learning, which has found success in many domains including energy meteorology. Finally, it should be reiterated that separation modeling has profound implications on the utilization of solar energy, as it is a part of the model chain, which converts irradiance to photovoltaic power. As such, continuous improvements in separation modeling are essential.

CRediT authorship contribution statement

Dazhi Yang: Conceptualization, Methodology, Software, Formal analysis, Validation, Investigation, Resources, Data curation, Writing – original draft, Visualization, Project administration, Funding acquisition. **Yizhan Gu:** Methodology, Software, Formal analysis, Investigation, Writing – original draft, Visualization. **Martin János Mayer:** Methodology, Validation, Writing – review & editing. **Christian A. Gueymard:** Methodology, Validation, Investigation, Data curation, Writing – original draft. **Wenting Wang:** Visualization, Writing – original draft. **Jan Kleissl:** Methodology, Writing – review & editing. **Mengying Li:** Writing – review & editing. **Yinghao Chu:** Writing – review & editing. **Jamie M. Bright:** Writing – review & editing.

Declaration of competing interest

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

Data availability

The R code and some additional results are uploaded to Github, at https://github.com/dazhiyang/Yang5-Separation. The three climatology variables and the cluster type of each point on the RCC map

are saved in `RCC.csv` to facilitate estimation for unseen sites. MBEs and RMSEs of the five separation models at 126 individual stations are listed in `Bn_error.csv` and `Dh_error.csv`, respectively.

Data: The raw data files as obtained from IEA members are not provided here for proprietary reasons. The raw climatology data, including cloud, aerosol, and surface albedo climatology, are also not provided due to their sizes. However, they can be downloaded from three sources as mentioned in Section 3. The cloud data is in `GeoTiff` format, with aerosol and albedo data provided in `NetCDF` format.

Code: Several R scripts are provided for the reader's information. Running these scripts is not possible, because that requires the original data files, which are either proprietary or too large to be uploaded to Github. However, readers who are interested in running the code can contact the corresponding author for more information.

- `ArrangeAOD.R`, `ArrangeCloud.R`, and `ArrangeAlbedo.R` are used to rearrange the raw climatology data and draw global maps of three variables.
- `dividing.R` melts the three climatology variables and gets the corresponding climatology value of 126 stations.
- `fitting.R` first performs clustering with the climatology variables at station locations, and then fits and obtains the regime-switching coefficients of YANG5.
- `validation.R` computes error metrics of the five selected models, visualizes the result of DM tests, and makes the k–k_t plot along with the bar plot.
- `Fig6_clusterPlot3D.R` depicts the 3D version of the clustering result of 126 stations.
- `Fig7_clusterMap.R` draws the global RCC map based on the clustering result.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (project no. 42375192). The authors 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 programme of the World Bank Group, in particular Joana Zerbin, Clara Ivanescu, Branislav Schnierer, Roman Affolter, GeoSUN Africa, Rachel Fox, and Margot King
- SKYNET, in particular Hitoshi Irie and Tamio Takamura (CEReS/Chiba-U.), Chiba University, Tadahiro Hayasaka (Tohoku University), and 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 in Morocco
- Institut de Recherche en Energie Solaire et Energies Nouvelles (IRESEN) in Morocco
- Research and Technology Centre of Energy (CRTEn) in Tunisia
- University of Jordan in Jordan and the Centre de Développement des Energies Renouvelables (CDER) in Algeria
- Majed Al Rasheedi from the Kuwait Institute for Scientific Research
- 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 of Jaen
- Dietmar Baumgartner from the University of Graz
- Julian Gröbner at PMOD
- Nicolas Fernay from the University of 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 authors also thank 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 procedure.

References

[1] Blanc P, Wald L. The SG2 algorithm for a fast and accurate computation of the position of the Sun for multi-decadal time period. Sol Energy 2012;86(10):3072–83. http://dx.doi.org/10.1016/j.solener.2012.07.018, URL: https://www.sciencedirect.com/science/article/pii/S0038092X12002800.

[2] Shi H, Yang D, Wang W, Fu D, Gao L, Zhang J, et al. First estimation of high-resolution solar photovoltaic resource maps over China with Fengyun-4A satellite and machine learning. Renew Sustain Energy Rev 2023;184:113549. http://dx.doi.org/10.1016/j.rser.2023.113549, URL: https://www.sciencedirect.com/science/article/pii/S1364032123004069.

[3] Wang W, Yang D, Huang N, Lyu C, Zhang G, Han X. Irradiance-to-power conversion based on physical model chain: An application on the optimal configuration of multi-energy microgrid in cold climate. Renew Sustain Energy Rev 2022;161:112356. http://dx.doi.org/10.1016/j.rser.2022.112356, URL: https://www.sciencedirect.com/science/article/pii/S1364032122002660.

[4] Yang G, Zhang H, Wang W, Liu B, Lyu C, Yang D. Capacity optimization and economic analysis of PV–hydrogen hybrid systems with physical solar power curve modeling. Energy Convers Manage 2023;288:117128. http://dx.doi.org/10.1016/j.enconman.2023.117128, URL: https://www.sciencedirect.com/science/article/pii/S0196890423004740.

[5] Yang D, Kleissl J. Summarizing ensemble NWP forecasts for grid operators: Consistency, elicitability, and economic value. Int J Forecast 2023;39(4):1640–54. http://dx.doi.org/10.1016/j.ijforecast.2022.08.002, URL: https://www.sciencedirect.com/science/article/pii/S016920702200111X.

[6] Mayer MJ, Yang D. Pairing ensemble numerical weather prediction with ensemble physical model chain for probabilistic photovoltaic power forecasting. Renew Sustain Energy Rev 2023;175:113171. http://dx.doi.org/10.1016/j.rser.2023.113171, URL: https://www.sciencedirect.com/science/article/pii/S1364032123000278.

[7] Mayer MJ, Yang D. Probabilistic photovoltaic power forecasting using a calibrated ensemble of model chains. Renew Sustain Energy Rev 2022;168:112821. http://dx.doi.org/10.1016/j.rser.2022.112821, URL: https://www.sciencedirect.com/science/article/pii/S1364032122007043.

[8] 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.

[9] 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.

[10] 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: https://www.sciencedirect.com/science/article/pii/S0038092X15005435.

[11] 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.

[12] 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.

[13] 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.

[14] Starke AR, Lemos LFL, Barni CM, Machado RD, Cardemil JM, Boland J, et al. 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.

[15] 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.

[16] 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.

[17] 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.

[18] Yang D. 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. Renew Sustain Energy Rev 2022;159:112195. http://dx.doi.org/10.1016/j.rser.2022.112195, URL: https://www.sciencedirect.com/science/article/pii/S1364032122001174.

[19] Forstinger A, Wilbert S, Jensen AR, Kraas B, Fernández-Peruchena C, Gueymard CA, et al. Expert quality control of solar radiation ground data sets. In: Solar World Congress 2021, International Solar Energy Society, Virtual Conference). 2021, p. 0104. http://dx.doi.org/10.18086/swc.2021.38.02.

[20] 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.

[21] 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.

[22] 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.

[23] Boland J, Huang J, Ridley B. Decomposing global solar radiation into its direct and diffuse components. Renew Sustain Energy Rev 2013;28:749–56. http://dx.doi.org/10.1016/j.rser.2013.08.023, URL: https://www.sciencedirect.com/science/article/pii/S1364032113005637.

[24] 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.

[25] 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.

[26] Yang D, Wang W, Bright JM, Voyant C, Notton G, Zhang G, et al. Verifying operational intra-day solar forecasts from ECMWF and NOAA. Sol Energy 2022;236:743–55. http://dx.doi.org/10.1016/j.solener.2022.03.004, URL: https://www.sciencedirect.com/science/article/pii/S0038092X22001645.

[27] Yang D, Sharma V, Ye Z, Lim LI, Zhao L, Aryaputera AW. Forecasting of global horizontal irradiance by exponential smoothing, using decompositions. Energy 2015;81:111–9. http://dx.doi.org/10.1016/j.energy.2014.11.082, URL: https://www.sciencedirect.com/science/article/pii/S0360544214013528.

[28] Fouquart Y, Buriez JC, Herman M, Kandel RS. The influence of clouds on radiation: A climate-modeling perspective. Rev Geophys 1990;28(2):145–66. http://dx.doi.org/10.1029/RG028i002p00145, URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/RG028i002p00145.

[29] 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.

[30] Gueymard CA. Clear-sky radiation models and aerosol effects. In: Polo J, Martín-Pomares L, Sanfilippo A, editors. Solar resources mapping: Fundamentals and applications. Cham: Springer International Publishing; 2019, p. 137–82. http://dx.doi.org/10.1007/978-3-319-97484-2_5.

[31] Gueymard CA, Lara-Fanego V, Sengupta M, Xie Y. Surface Albedo and reflectance: Review of definitions, angular and spectral effects, and intercomparison of major data sources in support of advanced solar irradiance modeling over the Americas. Sol Energy 2019;182:194–212. http://dx.doi.org/10.1016/j.solener.2019.02.040, URL: https://www.sciencedirect.com/science/article/pii/S0038092X19301653.

[32] Kiehl JT, Trenberth KE. Earth's annual global mean energy budget. Bull Am Meteorol Soc 1997;78(2):197–208. http://dx.doi.org/10.1175/1520-0477(1997)078<0197:EAGMEB>2.0.CO;2, URL: https://journals.ametsoc.org/view/journals/bams/78/2/1520-0477_1997_078_0197_eagmeb_2_0_co_2.xml.

[33] Salamalikis V, Vamvakas I, Gueymard CA, Kazantzidis A. Atmospheric water vapor radiative effects on shortwave radiation under clear skies: A global spatiotemporal analysis. Atmos Res 2021;251:105418. http://dx.doi.org/10.1016/j.atmosres.2020.105418, URL: https://www.sciencedirect.com/science/article/pii/S0169809520313557.

[34] Saxena A, Prasad M, Gupta A, Bharill N, Patel OP, Tiwari A, et al. A review of clustering techniques and developments. Neurocomputing 2017;267:664–81. http://dx.doi.org/10.1016/j.neucom.2017.06.053, URL: https://www.sciencedirect.com/science/article/pii/S0925231217311815.

[35] MacQueen J, et al. Some methods for classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, Vol. 14. Oakland, CA, USA; 1967, p. 281–97.

[36] Yang D, Alessandrini S, Antonanzas J, Antonanzas-Torres F, Badescu V, Beyer HG, et al. 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.

[37] Mayer MJ, Yang D. Calibration of deterministic NWP forecasts and its impact on verification. Int J Forecast 2023;39(2):981–91. http://dx.doi.org/10.1016/j.ijforecast.2022.03.008, URL: https://www.sciencedirect.com/science/article/pii/S0169207022000486.

[38] Kolassa S. Why the "best" point forecast depends on the error or accuracy measure. Int J Forecast 2020;36(1):208–11. http://dx.doi.org/10.1016/j.ijforecast.2019.02.017, URL: https://www.sciencedirect.com/science/article/pii/S0169207019301359.

[39] Gneiting T. Making and evaluating point forecasts. J Amer Statist Assoc 2011;106(494):746–62. http://dx.doi.org/10.1198/jasa.2011.r10138.

[40] Wilson AM, Jetz W. Remotely sensed high-resolution global cloud dynamics for predicting ecosystem and biodiversity distributions. PLOS Biol 2016;14(3):1–20. http://dx.doi.org/10.1371/journal.pbio.1002415.

[41] Yang D, Gueymard CA. Probabilistic merging and verification of monthly gridded aerosol products. Atmos Environ 2021;247:118146. http://dx.doi.org/10.1016/j.atmosenv.2020.118146, URL: https://www.sciencedirect.com/science/article/pii/S1352231020308761.

[42] Gelaro R, McCarty W, Suárez MJ, Todling R, Molod A, Takacs L, et al. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). J Clim 2017;30(14):5419–54. http://dx.doi.org/10.1175/JCLI-D-16-0758.1, URL: https://journals.ametsoc.org/view/journals/clim/30/14/jcli-d-16-0758.1.xml.

[43] Gueymard CA, Yang D. Worldwide validation of CAMS and MERRA-2 reanalysis aerosol optical depth products using 15 years of AERONET observations. Atmos Environ 2020;225:117216. http://dx.doi.org/10.1016/j.atmosenv.2019.117216, URL: https://www.sciencedirect.com/science/article/pii/S1352231019308556.

[44] Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, et al. The ERA5 global reanalysis. Q J R Meteorol Soc 2020;146(730):1999–2049. http://dx.doi.org/10.1002/qj.3803, URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803.

[45] 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.

[46] 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.

[47] 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.

[48] 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.