

Worldwide performance assessment of 75 global clear-sky irradiance models using Principal Component Analysis

Xixi Sun^a, Jamie M. Bright^{b,*,1}, Christian A. Gueymard^c, Brendan Acord^e, Peng Wang^{d,a,**}, Nicholas A. Engerer^b

^a School of Mathematics and System Sciences, Beihang University, Beijing, China

^b Fenner School of Environment and Society, The Australian National University, 2601, Canberra, Australia

^c Solar Consulting Services, Colebrook, NH, USA

^d Beijing Advanced Innovation Center for Big Data and Brain Computing, Beihang University, Beijing, China

^e Synene Clean Energy, Hong Kong, China

ARTICLE INFO

Keywords:

Clear-sky irradiance

Clear-sky model

Validation

Linke turbidity

Principal component analysis

Global horizontal irradiance

ABSTRACT

This study evaluates the performance of 75 clear-sky global irradiance models against 75 ground stations worldwide covering five major Köppen-Geiger climate classifications and overall global performance. After quality control, clear-sky detection and data-availability criteria, there are 4.36 million 1-min valid global horizontal irradiance (GHI) data points for evaluation. This study represents the most encompassing evaluation of its kind in terms of number of models assessed, number of ground stations used, most consistent selection of input variables in terms of temporal and spatial resolution, also using the most rigorous and fair performance assessment criteria and suitable ranking system. A statistically rigorous Principal Component Analysis (PCA) ranking procedure is proposed to replace the conventional ordering method based on single and simple statistics. In particular, it is demonstrated that all 13 error metrics contribute to the variance of GHI estimation and thus must all be accounted for in the ranking procedure. The best performing models in each climate zone of equatorial, arid, temperate, cold and polar are REST2v9.1, REST2v9.1, MAC2, REST2v5 and CLS, respectively. Globally, the top three performing models are MAC2, REST2v5, and REST2v9.1. Many models appear to suffer from over-fitting empirical relationships to training data, resulting in inconsistent worldwide performance. Furthermore, six different formulations of the Linke Turbidity factor (T_L) are evaluated through its application to five clear-sky models that require it as an input. The best globally performing T_L formulations are either by Ineichen or Gueymard when combined with four of the five Linke-dependent clear-sky models, whilst the Grenier formulation performs best with the Ineichen & Perez model. Significant performance variation is observed depending on the T_L formulation selected, hence, careful selection is required for an optimal application. The model codes are available in R [1].

1. Introduction

The efficient utilisation of solar energy at the Earth's surface is crucially dependent on the correct estimation of beam, diffuse and global irradiance under all-sky conditions. Modelling all-sky irradiance conditions, which are impacted by various meteorological processes, is normally done by superimposing the cloud effects to the estimated clear-sky irradiance component. The latter thus serves as the basis for the former. Hence, the overall result can be biased if one or the other is inaccurate. Although the clear-sky irradiance can be measured at

ground level when there is no cloud interference, its accurate measurement is often elusive because ground observation stations are historically too sparse to provide data with high spatial resolution. In the vast majority of cases, solar analysts are forced to estimate the clear-sky irradiance by using mathematical models - the so called clear-sky irradiance models. Those models normally evaluate the radiative effects of various atmospheric variables, such as aerosols, water vapour, ozone, or nitrogen dioxide, while also incorporating (directly or indirectly) the impacts from geo-variables such as latitude, longitude, Earth-Sun distance, solar zenith angle, elevation above sea level, or surface albedo.

* Corresponding author.

** Corresponding author. School of Mathematics and System Sciences, Beihang University, Beijing, China.

E-mail addresses: jamie.bright@anu.edu.au, jamiebright1@gmail.com (J.M. Bright).

¹ Co-first author.

Clear-sky irradiance models are employed in a multitude of scientific or technical disciplines. Research applications of such models range from evaluations of the grid impact from distributed generation of solar panels [2,3], the determination of geographic areas where irradiance estimates are more uncertain [4], to the preparation of short-term forecasts with sky cameras [5]. Industrial applications often relate to solar forecasting, in the form of real-time estimates of the solar resource needed by stakeholders such as electricity distribution network service providers or solar power plant operators [6,7]. In urban environments, good estimates of clear-sky irradiance are necessary to evaluate cooling loads of commercial structures [8,9]. In radiometry, a clear-sky radiation model is often used as part of the quality control process of measured data [10–12].

The accuracy of the underlying clear-sky irradiance model is important because it has the potential to influence or derail scientific results, as well as industrial decisions and actions. Although many validation studies have been conducted in the past, each suffers from certain shortfalls including the fairness of the testing procedure, geographic extent, proper selection of clear-sky periods, selected performance metrics, etc. In this study, those earlier studies are reviewed first in subsection 1.1 to identify any substantial room for improvement. Based on that assessment, the objectives of this study are then listed in subsection 1.2 alongside a detailed description of the paper structure.

1.1. Previous validation studies

Numerous validation studies have been conducted to compare model predictions of clear-sky irradiance with ground data. A thorough literature review of the most prominent validation studies has recently been proposed [4], to which the reader is referred for details. Only key findings from that analysis are discussed below, wherever appropriate.

A survey of the current literature reveals that the largest number of clear-sky irradiance models that have been validated in a single study so far is 54 [13–15]. These models were assessed against hourly ground data from two radiometric stations in Romania. In parallel, the number of ground stations compared against in validation studies for clear-sky models is typically low. For instance, 16 stations were considered by Refs. [16,17], 22 by Ref. [18], 24 by Ref. [19], and 30 by Ref. [20]. Although [20] included a relatively large number of stations, all of them were located within the USA, thus lacking worldwide significance. Nevertheless, that study used 1-min resolution irradiance data to test 9 clear-sky models.

The most globally expansive study, defined as covering the most diverse climatic areas, was performed by Ineichen [18], who compared 7 different models at 22 ground data sources across Europe, Africa and Israel, albeit only at 1-h resolution. A more globally expansive study but with fewer ground sites was recently performed by Calinoui et al. [17], who used the Köppen-Geiger (KG) general climate classification and 16 sources of ground data. However, only 4 of the KG climate classes were explored, and class D (snow/cold) was represented by a single station only (see subsection 3.2 for more detail on climate classification).

The limitations of validation studies that compare model predictions to ground irradiance observations are also discussed in detail by Refs. [4,21]. To summarise, the following six limitations have been noted in the literature and directly motivate the current investigation.

1. It is stated that the validation studies are not tested at a sufficient number of locations to account for climatic effects. Whilst this climate disaggregation process was introduced by Ref. [17]—following similar work by Ref. [22]—the number of stations within each zone was inadequate. Whereas a clear-sky irradiance model usually considers the impact of solar attenuation from different variables, some attenuation processes may not be correctly modeled under all possible conditions. This was actually confirmed by Ref. [4].
2. The input data stream is subject to measurement errors that

propagate into the model output. Hence, the intrinsic model performance can never be assessed exactly due to such impact of input uncertainty on the prediction error. This is exemplified by Ref. [21] where it is stated that the ideal case with “nearly-perfect” input data is essentially never met for all practical solutions. In validation studies it is implicitly or explicitly accepted that there is always an inherent error with input data, even though all efforts are normally taken to use the best possible available data. There is no commonly accepted procedure to define what the best data could be, or how raw data would have to be interpolated, extrapolated or corrected to achieve the required level of quality at each instant. To facilitate a fair test, it is obvious that the exact same data should be used as input to all models under test. However, this may not be possible in all cases because models differ in their data requirement, particularly regarding aerosol-related inputs.

3. Ground irradiance observations are subject to cloud interference, and so methods to identify clear-sky irradiance periods are required. Some approaches to detect and exclude cloud-contaminated data have been proposed in the literature, but those methods have typically been empirically derived and validated under only a small number of climatic conditions. The validation of existing clear-sky identification methods is substantively outside the scope of this paper. Regardless, a cloud detection methodology must be used.
4. There are significant uncertainties within the publicly available observational datasets used for model validation due to which radiometers are used, how they are installed, calibrated and maintained, and how data streams are quality controlled. Assessing model performance against highly uncertain ground data is not overly useful, and can even lead to wrong conclusions. Therefore, care must be taken to consider only high-quality irradiance data for this kind of exercise.
5. The temporal resolution of irradiance data used to evaluate clear-sky irradiance models is very important. For instance, hourly radiation data are used extensively for energy simulations of solar systems. However, unless the actual sky conditions have been recorded the whole time, e.g. with a sky camera, it is virtually impossible to ascertain that any whole hour has been cloud free. This can lead to serious bias on the model validations results at many sites. Hence, finer temporal resolutions are necessary to evaluate clear-sky model performance with maximum certainty. For instance Ref. [23], recommended a time step of 5-min or better, but few validation studies have actually used a high (1-min) temporal resolution.
6. The method of model validation is often overly simplified and not always satisfactory [24]. Simply analysing the mean difference between time series is not enough to truly assess the longer temporal performance of a model. Rigorous testing is a must. Ref. [21] proposed the Combined Performance Index (CPI) that blends three prominent performance metrics. Still, CPI has not been statistically tested for general or “universal” significance yet. An opportunity is presented here to analyse the importance of many performance metrics so as to produce an absolute ranking between clear-sky models.

1.2. Research objectives

From the six limitations discussed above, it is concluded that a clear opportunity exists to perform a more globally comprehensive and rigorous clear-sky model performance assessment. The main goal is to perform the most rigorous and general validation study that addresses all of the six limitations previously described. The following bullet-pointed research objectives explain how this is attempted here:

- Test the most significant clear-sky models that are well documented in the literature.

The models used here have been selected from four comprehensive validation studies [4,13,15,25,26], and are all freely-available through literature or public-domain access. The only exception to the public-domain rule is the inclusion of the latest REST2v9.1, which is a proprietary model and as such cannot be released, and of McClear, which is also a proprietary model, but whose outputs are accessible as a free online service that just requires registration.

- Perform the analysis at a large number of locations to account for differing climate regions.

Ruiz-Arias and Gueymard [4] performed the most globally diverse analysis of clear-sky models, yet this was not a conventional validation based on ground observations as reference, but rather an extensive sensitivity analysis meant to evaluate the regional effects of input data on the dispersion between the predictions from such models. Apart from that investigation, the hourly study by Ref. [18] and the climate-disaggregated study by Ref. [17] had the most diversity in climate regions considered. One goal of the present study is to perform the most detailed analysis to date by increasing the diversity of climate regions represented. This climatic disaggregation actually represents an additional layer to compare model performance. The climate data and subsequent categorisation are detailed in [subsection 3.2](#).

- Use the best quality public-domain input data with global coverage.

Obtaining the appropriate input data is often difficult because modern clear-sky models often require up to 9 different atmospheric variables as inputs. Flat assumptions or monthly means are often used for particularly dynamic variables such as aerosols [26], which can detriment the optimal performance of some models. Whereas this has been previously considered appropriate due to the difficulty of obtaining certain variables, it does not facilitate a fair test for each model. To that end, much care is taken here to extract data from gridded sources (satellite and reanalysis) with the best possible temporal resolution. Sources of data and related quality-control methods are detailed in [subsection 3.1](#).

- Use high-quality and fine temporal resolution surface observation validation data.

The sites selected here for the validation study cover many areas of the world and come from a few different sources; details of the observation data are presented in [subsection 3.3](#). All stations provide observations at 1-min resolution.

- Perform a validation test that meets the best published requirements for clear-sky radiation model validation, and statistically demonstrate the suitability of the performance metrics.

This study uses the recommendations for detailed high-quality validation — or performance analysis — as published in Ref. [21], whose results are presented in [subsection 4.2](#). Additionally, a statistical Principal Component Analysis (PCA) is introduced to determine the significance of each error metric, and by doing so, enabling the production of a ranking score. The goal here is to single out which performance metrics are most important in this context. Consequently, the term “performance analysis” is used extensively here to recognise that there are also inherent errors in all sensors involved.

- Determine the most appropriate usage of the Linke turbidity factor.

Many clear-sky models require the Linke turbidity factor, T_L , which cannot be observed directly, and thus is a calculated intermediate variable with various possible formulations. To that end, six different derivations of T_L are identified and tested for each clear-sky model that

requires it as an input. Each unique combination of clear-sky model and T_L derivation is considered as a stand-alone model due to the unique irradiance outputs from each combination; a similar convention was adopted by Ref. [13]. The results of the validation are presented in [section 5](#).

1.3. Contribution

This study's ambition is to present the most comprehensive global performance assessment of the largest number of clear-sky irradiance models against the most sources of ground-measured validation data ever accomplished. It is also designed to be the fairest test performed to date, arising from its use of dynamic inputs for each variable so that every model is assessed over its full range of capability. This approach enables the first rigorous climate-disseminated assessment of each model considered herein. Furthermore, a novel and reproducible ranking methodology is introduced for the benefit of future validation studies of this type. Ultimately, the best clear-sky model is identified for each climate as well as globally.

2. Clear-sky models for validation

In this study, 75 clear-sky models have been selected based on three factors derived from the literature: popularity, past performance, and accessibility of the model formulation. The clear-sky model formulations are included as supplementary material, where the appropriate references are discussed, along with any necessary addition or correction. The clear-sky models are all coded in R [1] with the notable exception of McClear and REST2v9.1.

[Table 1](#) provides a summary of the clear-sky models and their required input variables, including the solar constant E_0 [W m^{-2}], the zenith angle θ_z , site elevation above sea level h [m], ground albedo R_G , local barometric pressure p [mb], ambient temperature T [K], Linke turbidity factor (T_L), aerosol optical depth at 550 nm, 700 nm or broadband (τ), Ångström exponent α , Ångström turbidity coefficient β , total ozone amount u_{O_3} [atm-cm], total nitrogen dioxide amount u_{NO_2} [atm-cm], total precipitable water vapour u_{H_2O} [cm], and aerosol single-scattering coefficient ϖ .

Information on longitude, latitude and site elevation above sea level is obtained from the ground observation data sets.

The direct normal, diffuse horizontal and global horizontal irradiance under clear-sky conditions (e.g. cloudless skies) are denoted as E_{bnc} , E_{dnc} and E_{ghc} , respectively. Without specification otherwise, the modeled global horizontal irradiance E_{ghc} is obtained from its direct and diffuse components (also modeled) as:

$$E_{ghc} = E_{bnc} \cos(\theta_z) + E_{dnc} \quad (1)$$

The extraterrestrial irradiance (E_{ext}) on the i -th day of the year is obtained using the Spencer equation [27].

$$E_{ext} = E_0 \left[1.00011 + 0.034221 \cos\left(\frac{360(i-1)}{365}\right) + 0.00128 \sin\left(\frac{360(i-1)}{365}\right) + 0.000719 \cos\left(\frac{720(i-1)}{365}\right) + 0.000077 \sin\left(\frac{720(i-1)}{365}\right) \right] \quad (2)$$

where E_0 is the solar constant [W m^{-2}].

Since most clear-sky models are derived empirically, their performance is sensitive to the assumed value of the solar constant. Hence, the value of E_0 assigned to each model is in accordance to the value specified in its nomenclature. If the solar constant value is unspecified, it is set as a default value of $E_0 = 1361.1$ [W m^{-2}], as befitting the latest solar constant derivation by Ref. [28]. All models require the solar zenith angle θ_z as an input. It is calculated following the well-known methodology defined by Ref. [29] as:

Table 1

Clear-sky models and their required inputs (*). The total number of input variables for each model is displayed in the last column. The variables displayed are zenith angle θ_z , site elevation above sea level h [m], surface albedo R_G , local barometric pressure p in [mb], ambient temperature T [K], Linke turbidity factor (T_L), aerosol optical depth at 550 nm, 700 nm or broadband (τ), Ångström exponent α , Ångström turbidity coefficient β , total ozone amount u_{O_3} [atm-cm], total nitrogen dioxide amount u_{NO_2} [atm-cm], total precipitable water vapour u_{H_2O} [cm], and aerosol single-scattering coefficient ϖ . The model number is consistent throughout the document. Note that those with multiple model numbers have different instances due to T_L variation.

#	Clear-sky Model	Citation	E_0	θ_z	h	R_G	p	T	T_L	τ	α	β	u_{O_3}	u_{NO_2}	u_{H_2O}	ϖ	Total
1	TJ	[31]		*													1
2	Schulze	[32]		*													1
3	DPP	[33]		*													1
4	Adnot	[34]		*													1
5	Biga	[35]		*													1
6	ASHRAE	[36]		*													1
7	Sharma	[37]	*	*													2
8	El Mghouchi	[38]	*	*													2
9	Yang & Walsh	[39]	*	*													2
10	HLJ	[40]	*	*	*												3
11	Kumar	[41]	*	*			*										3
12	Campbell	[42]	*	*			*										3
13	Fu & Rich	[43]	*	*	*												3
14	Atwater & Ball-1	[44]	*	*			*								*		4
15	KASM	[45]	*	*			*								*		4
16	Capderou	[46]	*	*	*		*										4
17–22	Kasten	[47]	*	*	*				*								4
23–28	Heliosat-1	[48]	*	*			*		*								4
29–34	ESRA	[49]	*	*	*				*								4
35–40	Heliosat-2	[50]	*	*	*				*								4
41–46	Ineichen & Perez	[51]	*	*	*				*								4
47	CLS	[52]	*	*		*	*								*		5
48	King	[53]	*	*		*	*					*			*		5
49	Josefsson	[14]	*	*		*	*								*		5
50	Badescu	[54]	*	*			*						*		*		5
51	Simplified Solis	[55]	*	*			*			*					*		5
52	Advanced Solis	[56]	*	*			*			*					*		5
53	Perrin	[57]	*	*			*					*			*		6
54	CEM	[58]	*	*		*	*			*					*		6
55	Atwater & Ball-2	[59]	*	*		*	*			*					*		6
56	RSC	[60]	*	*		*	*					*			*		6
57	PSIM	[61]	*	*		*	*					*			*		6
58	Bashahu	[62]	*	*			*				*	*			*		6
59	MMAC	[63]	*	*		*	*			*					*		6
60	Yang	[64]	*	*			*					*	*		*		6
61	Calinoiu	[17]	*	*			*					*	*	*	*		6
62	Hoyt	[65]	*	*		*	*					*	*		*		7
63	MAC2	[66]	*	*		*	*				*	*			*		7
64	METSTAT	[67]	*	*		*	*			*			*		*		7
65	PR	[68]	*	*		*	*					*	*		*		7
66	Paulescu & Schlett	[69]	*	*	*		*					*	*		*		7
67	MRMv5	[70]	*	*		*	*					*	*		*		7
68	MRMv6.1	[71]	*	*		*	*			*			*	*	*		7
69	Janjai	[72]	*	*	*		*				*	*	*		*		7
70	Bird	[59]	*	*		*	*				*	*	*		*		8
71	Iqbal-c	[29]	*	*		*	*				*	*	*		*		8
72	Modified Iqbal-c	[25]	*	*	*	*	*				*	*	*		*		8
73	REST2v5	[73]	*	*		*	*				*	*	*	*	*		9
74	REST2v9.1		*	*		*	*				*	*	*		*	*	9
75	McClear	[74]	*	*	*	*	*	*		*	*	*	*		*		10

$$\theta_z = \cos^{-1}[\sin(\delta)\sin(\phi) + \cos(\delta)\cos(\phi)\cos(\omega)], \quad (3)$$

where δ is the declination angle, ϕ is the geographic latitude with north positive and ω is the hour angle with noon zero and morning positive. δ is also defined by Ref. [29] with d being the day number with 1-min resolution:

$$\delta = 0.006918 - 0.399912\cos(B) + 0.070257\sin(B) - 0.006758\cos(2B) + 0.000907\sin(2B) - 0.002697\cos(3B) + 0.00148\sin(3B), \quad (4)$$

$$B = \frac{2\pi(d-1)}{365} \quad (5)$$

As one of the objectives of this paper is to determine the most appropriate T_L calculation methodology for clear-sky model utilisation, five radiation models that depend on T_L are combined with six different

formulations of T_L (as detailed in subsection 3.1). In turn, this produces 30 T_L -dependent clear-sky models.

3. Data description and sources

This section provides descriptions and sources for: 1) the raw input data for each model; 2) the climate region; and 3) the ground measurements for performance evaluations.

3.1. Input variables

Table 1 lists all the input variables for each clear-sky model used in this study. Since many of those variables exhibit rapid dynamic changes, high temporal resolution measurements are sought to capture this temporal variability. Certain variables also exhibit significant

variability over space. Hence, ideally, a very high spatio-temporal resolution input is desired. Because the present study intends to conduct a climatologically diverse validation, global coverage is needed for each input data. Therefore, a trade-off must be realised between data availability and its spatio-temporal resolution. Recourse to gridded datasets is the obvious choice, as also considered previously [4,75–77].

An appropriate gridded dataset that contains the required input variables (except those specifically derived or detailed here) is the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), by the NASA Global Modeling and Assimilation Office [78]. This global dataset provides hourly data since 1980 at a spatial resolution of $0.5 \times 0.625^\circ$.

It must be acknowledged that the source of input variables introduces a risk of bias, at least at some sites. The MERRA-2 Reanalysis data does not exactly represent the variables at the ground stations. As such, errors within the input data are likely to impact more upon those clear-sky irradiance models that depend on them more heavily. The appropriate input variables need to be selected from a public-domain and validated database; furthermore, the selected source must have global coverage such that users of these clear-sky models (as advised from this research) can expect similar performance. Such an approach was used previously in the literature [77,79]. Only a few databases—all based on reanalysis modelling—currently exist with all the detailed information required and at sufficiently high spatio-temporal resolution. MERRA-2 is such an ideal candidate and has been selected here. Individual variables from MERRA-2 have been validated around the world against AERONET, namely the aerosol and column water-vapour products [80,81]. The impact and accuracy of the MERRA-2 variables has been compared to ground measurements from AERONET stations [82] with particular concern for solar energy applications in a few studies. For instance Ref. [83], shows that using precipitable water data from MERRA (the predecessor of MERRA-2) did not result in significant changes in predicted clear-sky GHI, compared to using reference AERONET observations. Ref. [84] extensively compares many different sources of data inputs for solar clear-sky irradiance modelling and some of the idiosyncratic differences between certain datasets and opts to use MERRA-2 over alternatives. This reference shows results of tests undertaken at a specific BSRN site (Tamanrasset) where atmospheric conditions can vary dramatically over a wide range due to a normal clean background and occurrences of dust storm events. The resulting difference in modeled GHI is found small and insignificant (overall bias <1%) when alternatively using MERRA2 and AERONET as inputs to various clear-sky models. This suggests that the present approach is likely to have wide validity.

The accuracy of MERRA-2 variables against AERONET as well as sensors mounted on airborne and shipping measurement crafts is extensively documented in the literature [85]. Fundamentally, the MERRA-2 reanalysis dataset is not without inaccuracies or potential impacts. It does, however, present a high temporal and spatial resolution with a long historical database that makes it ideal for applications with solar clear-sky irradiance calculations. Naturally, the performance of certain models may benefit more or less depending on the reanalysis dataset selected. Hence, the final results presented here are somewhat specific to the use of the popular MERRA-2.

The variables that are extracted from the MERRA-2 reanalysis dataset are detailed in Table 2. There is need for altitude corrections to the MERRA-2 gridded data due to the natural decrease of τ and u_{H_2O} with increasing altitude. This is particularly significant when the reported reanalysis gridded cell elevation differs from the ground station elevation height. The proposed correction was defined by Ref. [8] and is based on a scale-height approximation:

$$k(h) = k(h_0) \exp[(h_0 - h)/H_a] \quad (6)$$

where k is the variable undergoing scale-height correction, h_0 is the reference site's elevation above mean sea level (amsl) as reported by MERRA-2, and h is the elevation of the current validation site. The scale

height for aerosols, H_a , is experimentally derived [86] and varies site-by-site. As there has been little global study on scale height [8], proposed an average value of $H_a = 2100$ m that is assumed satisfactory for general purpose, and for both τ_{550} and u_{H_2O} . The cell height for MERRA-2 is taken from the v5.12.4 2-dimensional constants dataset by extracting the surface geopotential height (PHIS variable) and dividing it by the average standard gravity (9.80665 m s^{-2}) [87].

Note that T is not extracted from MERRA-2 because the only model that requires it is McClear, which is an online model that requires only metadata specification.

The only variable that is required but not available from MERRA-2 is the total nitrogen dioxide amount u_{NO_2} . It is thus extracted from observations of NASA's Ozone Monitoring Instrument (OMI) aboard the Aura polar orbiting satellite. The OMNO2d product [88] is specifically used here, and is converted from OMI units [mol.cm^{-2}] to a columnar amount in atmosphere centimetres [atm-cm] by a division of 2.69×10^{16} . This OMI product provides daily resolution measurements and is quality controlled. However, since the OMI instrument is on-board a polar orbiting satellite and its data points have reasonably strict quality screening, there are frequent data gaps in-between flight paths. Those empty periods are not spatially or temporally filled and so are left blank. In the absence of OMI u_{NO_2} data, we default to a value of 0.0002 atm-cm [73].

Using the raw data extracted from MERRA-2, τ_{700} and the Ångström turbidity β can be derived from τ_{550} and α using Angstrom's law:

$$\tau_{700} = \tau_{550} \left(\frac{700}{550} \right)^{-\alpha} \quad (7)$$

$$\beta = \frac{\tau_{550}}{0.550^{-\alpha}} \quad (8)$$

The data processing scripts to extract the input data from the atmospheric databases are written in Matlab® and are available for inspection and adoption at [89].

A number of studies have found that the Linke turbidity factor at air mass 2 (T_{L2}) is correlated with u_{H_2O} and β , and have proposed various empirical conversion methods. Considering that the T_{L2} formulation is expected to affect the performance of Linke-based models [4], a specific study is desirable.

For a comprehensive analysis, six formulations of T_{L2} are considered here, and denoted as: T_{L2R} [90], T_{L2D} [91], T_{L2I} [92], T_{L2Gu} [93], T_{L2M} [94], and T_{L2Gr} [95]:

$$T_{L2R} = (1.8494 + 0.2425 u_{H_2O} - 0.0203 u_{H_2O}^2) + \beta(15.427 + 0.3153 u_{H_2O} - 0.0254 u_{H_2O}^2) \quad (9a)$$

$$T_{L2D} = \frac{1}{0.8662} \left[\beta(16 + 0.22 u_{H_2O}) + 0.1 + \frac{175 - \theta_z}{39.5 \exp(-u_{H_2O}) + 47.1} \right] \quad (9b)$$

$$T_{L2I} = 3.91 \tau_{550} \exp\left(0.689 \frac{1013.25}{p}\right) + 0.376 \ln(u_{H_2O}) + \left[2 + 0.54 \frac{1013.25}{p} - 0.5 \left(\frac{1013.25}{p} \right)^2 + 0.16 \left(\frac{1013.25}{p} \right)^3 \right] \quad (9c)$$

$$T_{L2Gu} = \frac{1}{0.8662} \left[1 + \frac{m_A \delta_w + \delta_{nt} + \delta_a}{m_R \delta_c} \right] \quad (9d)$$

$$T_{L2M} = 1.5 + 12.4\beta + 0.5 u_{H_2O}^{\frac{1}{3}} + 4(\beta - 0.1) \ln \left[\frac{\exp(-0.12h/1000)}{\cos(\theta_z) + 0.50572(96.07795 - \theta_z)^{-1.6364}} \right] \quad (9e)$$

Table 2

MERRA-2 variables alongside data code specification used within the time-averaged two-dimensional MERRA-2 collection. Any conversion away from the source data format is detailed. All data is provided at 1-h resolution. The aerosol single-scattering albedo, ω , is obtained as the ratio TOTSCATAU/TOTEXTTAU.

Variable	Symbol	Collection	Code	Raw data conversion	Final unit
AOD (550 nm)	τ_{550}	M2T1NXAER	TOTEXTTAU	–	–
Aerosol scattering	ω	M2T1NXAER	TOTSCATAU	–	–
Albedo	R_G	M2T1NXRAD	ALBEDO	–	frac.
Ångström exponent	α	M2T1NXAER	TOTANGSTR	–	–
Ozone	u_{O_3}	M2T1NXSLV	TO3	$\times 0.001$	atm-cm
Precipitable water	u_{H_2O}	M2T1NXSLV	TQV	$\times 0.1$	atm-cm
Pressure	p	M2T1NXSLV	PS	$\times 0.01$	mb

$$T_{L2Gr} = \frac{\beta - (-0.10545 - 0.02005 u_{H_2O} + 0.0050689 u_{H_2O}^2 - 0.0005202 u_{H_2O}^3)}{0.073554 - 0.0029011 u_{H_2O} + 0.00075553 u_{H_2O}^2 - 0.000078281 u_{H_2O}^3} \quad (9f)$$

where δ_x refer to the turbidity coefficients in Ref. [93].

Since the Linke turbidity at air mass 2 is a crucial input variable in five clear-sky models: namely, Kasten, Heliosat-1, Heliosat-2, ESRA and Ineichen-Perez, these are matched with the six formulations above, yielding 30 combinations of Linke-based models. For example, the ESRA model actually has 6 instances: ESRA-R, ESRA-D, ESRA-I, ESRA-Gu, ESRA-M and ESRA-Gr.

It is stated that T_{L2} is an input to all the aforementioned clear-sky models, although some internally convert it to T_L , based on $T_L = T_{L2} \times 0.8662$ [25].

3.2. Climate regions

Because the current study intends to evaluate various clear-sky models at a large number of locations to account for climate impacts, a method is needed to sub-categorise the different ground stations into distinct climatic regions. Although earlier studies have proposed clustering the regions by the mean annual irradiance [28], it is not a strong indication of climate in regions with distinct climate patterns. Instead, the popular KG climate classification system [17,96] is employed here to separate the ground stations into appropriate regions [17,97,98]. The KG classification includes 29 different climate regions. For simplicity, only the five major climate classes are considered: tropical/equatorial, arid/semi-arid, temperate, cold/continental and polar. They are denoted as class A, B, C, D and E, each containing 9, 23, 29, 7, and 7 stations, respectively. The different climate regions are displayed in Fig. 1.

In addition to the five KG climate regions, a “global” category is also defined as the ensemble of all climate classes. The formulation of the global model performance rank as derived from the ranking scores in each of the climates is detailed in subsection 4.2.

3.3. Ground measurements

To facilitate the worldwide performance evaluation of all models undertaken here, a total of 75 radiometric stations are considered for their data availability, quality and regular usage in the literature. As mapped and categorised in Fig. 1, many of those ground stations are located in the southern hemisphere, i.e., outside the usual areas considered in previous studies. This constitutes another novelty of the present study. Table 3 lists all the ground measurement sites with their names, coordinates, elevation, data range, and KG climate classification.

Based on the current literature (e.g., Refs. [99,100]), the reference GHI is obtained here as the sum of the horizontal projection of the direct normal irradiance (DNI) and the diffuse horizontal irradiance (DIF), because this provides a lower-uncertainty determination than

that of the conventional output of unshaded pyranometers. More specifically, the latter are affected by significant thermal offset and increased cosine response error at low sun angles. The latter source of error would be exacerbated at high-latitude sites, with the risk of inadvertently decreasing the reported performance of good models. Note, however, that the pyranometer-based GHI is still extremely important information in the quality-control process described later in this section.

The study is purposefully limited to the three-year period of 2015–2017, inclusive. Although there are substantially more GHI data available from the same sources, the number of years for the current analysis was restricted due to practical issues related to large data transfer, process and quality control among the authors collaborating in three continents. For instance, the entire global dataset of MERRA-2 must be downloaded per day because there is no option to extract site-specific or regional data time series [84]. Nevertheless, the three-year period maintains statistical significance of the research and such time span is consistent with studies on large-scale validation [4].

Seven different sources of irradiance data are identified to populate the whole 75-site database.

- The vast majority of the 1-min solar data used here are obtained from the LR0100 product of the Baseline Surface Radiation Network (BSRN) [101], an affiliate of the World Meteorological Organization (WMO), which maintains the highest standard of quality control (QC) of all institutions participating in solar radiation monitoring [21].
- The Bureau of Meteorology, Australia (BoM, <http://www.bom.gov.au/climate/data/oneminsolar/>), provides 1-min solar data for Australia and the Cocos (Keeling) Islands. For those BoM sites that actually double up as BSRN sites, a slight difference exists between the post-processing techniques applied. Wherever this overlap occurs, the BSRN data source is prioritised over that from BoM. All BoM sites use Kipp & Zonen CM11 pyranometers and CH1 pyrhemometers for DIF and DNI, and follow the WMO standards for reporting and QC. These radiometers are also recalibrated annually.
- NOAA's Global Monitoring Division of the Earth System Research Laboratory provides data for three sites (<https://www.esrl.noaa.gov/gmd/dv/data/>). At these sites, Eppley PSP pyranometers and NIP pyrhemometers are used for DIF and DNI, respectively.
- The National Renewable Energy Laboratory (NREL) contributes with three ground sites in the USA (<https://midcdmz.nrel.gov/>) where 1-min irradiance is measured with thermopile radiometers. At the OASIS station of Tucson, Arizona, DIF and DNI are measured with an Eppley 8–48 pyranometer and a Kipp & Zonen CHP1 pyrhemometer, respectively. The Las Vegas, Nevada station is an exception here, because DIF is not measured with a separate sensor. Hence, the component-sum GHI cannot be determined, and GHI is simply obtained from the global pyranometer measurements. The Golden, Colorado site employs intentional redundancy. For this study, three pyrhemometers and two pyranometers are used for DNI and DIF, respectively, allowing expanded quality control. In the vast majority of cases, the reading of these sensors is well within their

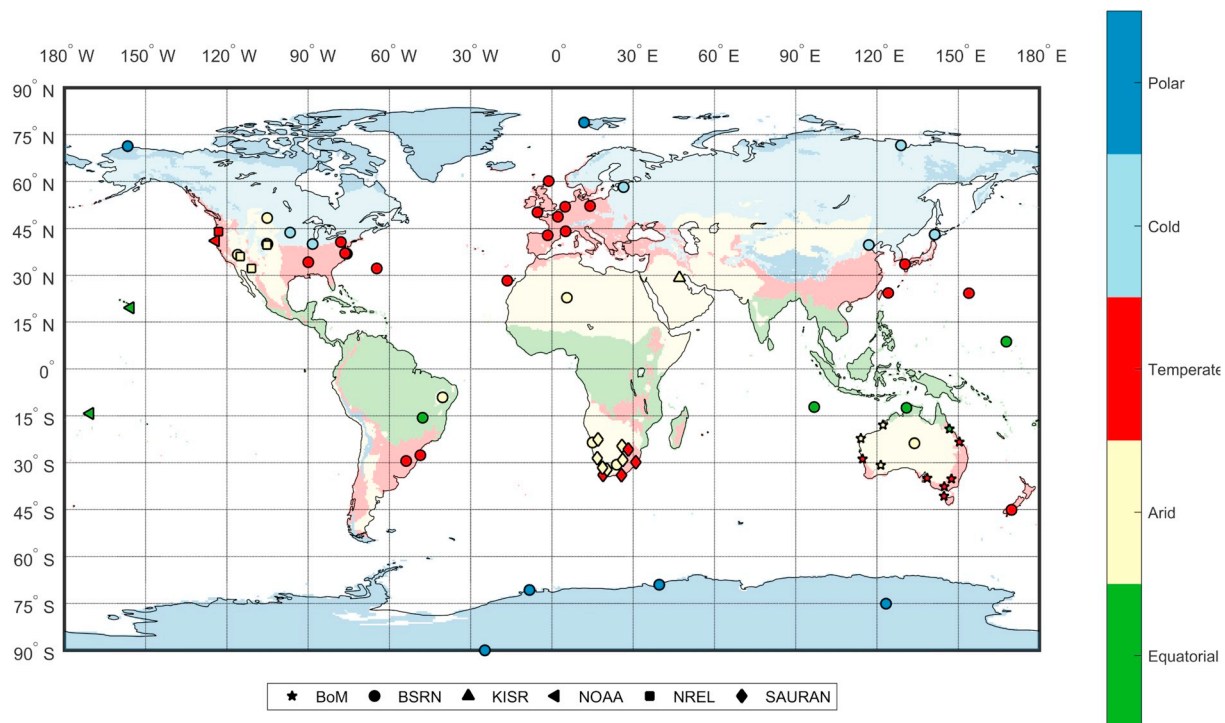


Fig. 1. World map indicating the Köppen-Geiger main classifications as well as the location and climate for each of the ground sites used in this study. For maximum legibility, each data source is indicated by a unique shape, combined with the colour of the appropriate climate class. An equidistant cylindrical projection is employed. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

combined uncertainties, and the average is used to define DNI and DIF. Slight disagreement typically exists when one instrument is being cleaned, resulting in a lower reading during one to 3 min. For DNI, the average of the two pyrheliometers providing the higher reading is then used. For DIF, the higher reading is used.

- At the Shagaya station of the Kuwait Institute of Scientific Research (KISR, <http://www.kisr.edu.kw>, [102–104]), irradiance is measured with individual Kipp & Zonen CMP21 pyranometers for GHI and DIF, and CHP1 pyrheliometer for DNI. Helpful redundancy is provided by a photodiode-based rotating shadowband irradiometer (RSI), whose output is comparable to that of the thermopile radiometers [105]. The stringent maintenance program includes daily cleaning of all instruments, which is conducive to high-quality data despite the rapid rate of soiling in this arid environment. The quality control procedure takes advantage of the instrumental redundancy, and selects the best of the two data streams with priority given to the thermopile data.
- For southern Africa (nine sites in South Africa, one in Namibia, and one in Botswana), 1-min solar datasets are provided by the South African Universities Radiation Network (SAURAN, <http://sauran.net/Data>). The SAURAN ground stations feature Kipp & Zonen CMP11 pyranometers for GHI and DIF measurements and CHP1 pyrheliometers for DNI measurements. Instruments are subject to regular maintenance and all glass domes and windows are cleaned multiple times each week.

All irradiance data collected for this study undergo a quality control (QC) routine as described by Ref. [106], whereby irradiance data points are tested for their physical acceptability and extreme limits. In particular, all situations of tracker-induced error are detected and eliminated from further analysis.

In addition to the [106] criteria, a GHI data point is removed if it does not pass one of the following conditions:

1. If any measurement is deemed “not-clear” during the clear-sky detection step, it is not considered for any model (detailed in subsection 4.1).
2. GHI data at solar zenith angles above 85° are removed. The reasons for this are that experimental errors are then typically largest (when irradiance is also very low), and that most clear-sky models are implicitly considering this limitation.

The complete database encompassing the 75 ground stations considered has 50,363,993 raw 1-min GHI data points. After the QC routine of [106] and additional filtering, only 4,360,012 valid clear-sky data points remain, or 8.66%. The total number of valid 1-min irradiance measurements at each climate region are: Equatorial, 107,990 (2.48%); Arid, 2,372,616 (54.42%); Temperate, 1,294,090 (29.68%); Cold, 258,111 (5.92%); and Polar, 327,205 (7.50%).

All the conditions just described help guarantee that all clear-sky models are validated against the same unique ground database. The only exception is McClear (#75), which cannot provide estimates at extreme latitudes poleward of $\pm 89^\circ$; there are 54,249 fewer valid measurements in its case.

Details on all ground stations used in this study are provided in Table 3.

4. Methodology

In this section, a general framework is proposed to evaluate the performance of global irradiance estimation by clear-sky models. More specifically, the measured irradiance time series are first filtered to extract cloudless periods (subsection 4.1). The methodology of assessing and evaluating model performance is detailed in subsection 4.2. Lastly, an objective ranking methodology is developed from various model performance indicators (subsection 4.3). A summary of the complete methodological approach is presented in Algorithm 1.

Table 3

Site information on 75 pre-selected solar measurement ground stations. The climate classes (last column) are based on the Köppen-Geiger standards [107]. East longitudes are positive and west longitudes are negative. The station elevation is represented by h in metres above mean sea level, as reported by the data source. The column labeled “Source” indicates where the data was obtained from.

#	Site Name (Country)	Lat. [°]	Lon. [°]	h [m]	Period	Source	Climate
1	Adelaide (Australia)	−34.9524	138.5196	2	06/16–12/17	BoM	Csb
2	Alice Springs (Australia)	−23.7951	133.8890	546	06/16–12/17	BoM	BWh
3	Alice Springs (Australia)	−23.7980	133.8880	547	01/15–06/16	BSRN	BWh
4	Barrow (USA)	71.3230	−156.6070	8	01/15–08/17	BSRN	ET
5	Bermuda (UKOT)	32.2670	−64.6670	8	01/15–08/17	BSRN	Csc
6	Bondville (USA)	40.0667	−88.3667	213	01/15–08/17	BSRN	Dfa
7	Boulder (USA)	40.1250	−105.2370	1689	01/15–08/17	BSRN	BSk
8	Brasilia (Brazil)	−15.6010	−47.7130	1023	01/15–12/15	BSRN	Aw
9	Broome (Australia)	−17.9475	122.2353	7.42	06/16–12/17	BoM	BSh
10	Cabauw (Netherlands)	51.9711	4.9267	0	01/15–12/17	BSRN	Cfb
11	Camborne (UK)	50.2167	−5.3167	88	01/15–07/17	BSRN	Cfb
12	Cape Grim (Australia)	−40.6817	144.6892	95	06/16–12/17	BoM	Cfb
13	Carpentras (France)	44.0830	5.0590	100	01/15–12/17	BSRN	Cfb
14	Chesapeake Light (USA)	36.9050	−75.7130	37	01/15–11/16	BSRN	Cfa
15	Cener (Spain)	42.8160	−1.6010	471	01/15–09/17	BSRN	Cfb
16	Cocos Islands (Australia)	−12.1930	96.8350	6	01/15–06/16	BSRN	Af
17	Cocos Islands (Australia)	−12.1892	96.8344	3	06/16–12/17	BoM	Af
18	Concordia Station (Antarctica)	−75.1000	123.3830	3233	01/15–08/15	BSRN	EF
19	Darwin (Australia)	−12.4239	130.8925	30.4	06/16–12/17	BoM	Aw
20	Darwin (Australia)	−12.4240	130.8925	32	01/15–06/16	BSRN	Aw
21	De Aar (SZA)	−30.6667	23.9930	1287	04/15–12/17	BSRN	BSk
22	Desert Rock (USA)	36.6260	−116.0180	1007	01/15–08/17	BSRN	BWk
23	Eskom Sutherland (SZA)	−32.2220	20.3478	1450	01/15–04/17	SAURAN	BSk
24	Florianopolis (Brazil)	−27.6047	−48.5227	11	01/15–08/17	BSRN	Cfa
25	Fort Peck (USA)	48.3167	−105.1000	634	01/15–08/17	BSRN	BSk
26	Fukuoka (Japan)	33.5822	130.3764	3	01/15–12/17	BSRN	Cfa
27	Georg von Neumayer (Antarc.)	−70.6500	−8.2500	42	01/15–01/17	BSRN	EF
28	Geraldton (Australia)	−28.8047	114.6989	29.7	01/17–09/17	BoM	Csa
29	GIZ Richtersveld (SZA)	−28.5608	16.7615	141	01/15–09/17	SAURAN	BWh
30	Gobabeb (Namibia)	−23.5614	15.0420	407	01/15–12/17	BSRN	BWh
31	Golden (USA)	39.7420	−105.1800	1829	01/15–12/17	NREL	BSk
32	Goodwin Creek (USA)	34.2547	−89.8729	98	01/15–08/17	BSRN	Cfa
33	Ishigakijima (Japan)	24.3367	124.1644	5.7	01/15–11/17	BSRN	Cfa
34	Izana (Spain)	28.3093	−16.4993	2372.9	01/15–10/16	BSRN	Csb
35	Kalgoorlie Boulder (Australia)	−30.7847	121.4533	365.3	01/17–12/17	BoM	BSh
36	Kwajalein (Marshall Islands)	8.7200	167.7310	10	01/15–09/17	BSRN	Af
37	Langley Research Center (USA)	37.1038	−76.3872	3	01/15–12/17	BSRN	Cfa
38	Lauder (New Zealand)	−45.0450	169.6890	350	01/15–01/17	BSRN	Cfb
39	Learmonth (Australia)	−22.2406	114.0967	5	01/17–09/17	BoM	BWh
40	Lerwick (UK)	60.1389	−1.1847	80	01/15–07/17	BSRN	Cfb
41	Lindenberg (Germany)	52.2100	14.1220	125	01/15–01/17	BSRN	Cfb
42	Mauna Loa (USA)	19.5362	−155.5763	3397	01/15–08/17	NOAA	ET
43	Melbourne (Australia)	−37.6655	144.8321	113.4	06/16–12/17	BoM	Cfb
44	Minami Torishima (Japan)	24.2883	153.9833	7.1	01/15–12/17	BSRN	Af
45	Nelson Mandela Univ. (SZA)	−34.0086	25.6653	35	01/15–11/17	SAURAN	BSh
46	NUST (Namibia)	−22.5650	17.0750	1683	07/16–09/17	SAURAN	BWh
47	Ny-Alesund (Norway)	78.9250	11.9300	11	01/15–10/16	BSRN	ET
48	Palaiseau (France)	48.7130	2.2080	156	01/15–07/17	BSRN	Cfb
49	Petrolina (Brazil)	−9.0680	−40.3190	387	01/15–01/16	BSRN	BSh
50	Rockhampton (Australia)	−23.3753	150.4775	10.4	06/16–12/17	BoM	Cfa
51	Rock Springs (USA)	40.7200	−77.9333	376	01/15–08/17	BSRN	Dfb
52	Sao Martinho da Serra (Brazil)	−29.4428	−53.8231	489	01/15–12/15	BSRN	Cfa
53	Sapporo (Japan)	43.0600	141.3286	17.2	01/15–12/17	BSRN	Dfb
54	Shagaya (Kuwait)	29.2099	47.0603	241.3	01/15–12/17	KISR	BWh
55	Sioux Falls (USA)	43.7300	−96.6200	473	01/15–08/17	BSRN	Dfa
56	South Pole (Antarctica)	−89.9830	−24.7990	2800	01/15–03/17	BSRN	EF
57	Stellenbosch Univ. (SZA)	−33.9281	18.8654	119	01/15–10/17	SAURAN	Csb
58	Syowa (Antarctica)	43.7300	−96.6200	473	01/15–07/15	BSRN	EF
59	Tamanrasset (Algeria)	22.7903	5.5292	1385	01/15–12/17	BSRN	BWh
60	Tiksi (Russia)	71.5862	128.9188	48	01/15–11/17	BSRN	Dfd
61	Toravere (Estonia)	58.2540	26.4620	70	01/15–12/17	BSRN	Dfb
62	Townsville (Australia)	−19.2483	146.7661	4.34	01/17–09/17	BoM	Aw
63	Trinidad Head (USA)	41.0541	−124.1510	107	01/15–04/17	NOAA	Csb
64	Tucson (USA)	32.2297	−110.9553	786	01/15–12/17	NREL	BSh
65	Tutuila (Samoa)	−14.2474	−170.5644	42	01/15–08/17	NOAA	Af
66	University of Nevada (USA)	36.0600	−115.0800	615	01/15–12/17	NREL	BWh
67	University of Oregon (USA)	44.0467	−123.0743	134	08/16–12/17	NREL	Csb
68	Uni. Free State (SZA)	−29.1107	26.1850	1491	01/15–08/17	SAURAN	BSk
69	Uni. Nata Howard (SZA)	−29.8710	30.9770	150	01/15–09/17	SAURAN	Cfa
70	Uni. Nata Westville (SZA)	−29.8170	30.9450	200	01/15–12/17	SAURAN	Cfa
71	Uni. Pretoria (SZA)	−25.7531	28.2286	1410	01/15–09/17	SAURAN	Cwa
72	USAid Gaborone (Botswana)	−24.6610	25.9340	1014	01/15–12/17	SAURAN	BSh

(continued on next page)

Table 3 (continued)

#	Site Name (Country)	Lat. [°]	Lon. [°]	<i>h</i> [m]	Period	Source	Climate
73	Vanrhynsdorp (SZA)	−31.6175	18.7383	130	01/15–10/17	SAURAN	BSh
74	Wagga (Australia)	−35.1583	147.4575	212	06/16–12/17	BoM	Cfa
75	Xianghe (China)	39.7540	116.9620	32	01/15–10/15	BSRN	Dwa

Algorithm 1

Evaluation framework.

Step 1: Collect as much high-quality 1-min irradiance data from around the world.
 Step 2: Extract corresponding dynamic input variables from MERRA-2 and OMI and apply appropriate scale-height corrections and temporal interpolations.
 Step 3: Perform quality control (QC) on data with the following steps:
 (3.1). QC as defined by Ref. [106].
 (3.2). Filter all time steps of $\theta_z > 85^\circ$.
 (3.3). Remove any time steps identified as not-clear following the procedure in Ref. [12], augmented with a new hour-duration criterion.
 Step 4: Categorise the ground data sources according to the five Köppen-Geiger climate classifications.
 Step 5: Calculate clear-sky GHI from all 75 models at all 75 sites that correspond exactly to the valid ground data.
 Step 6: For each climate and model estimate, calculate the 13 performance metrics as recommended by Ref. [21].
 Step 7: Perform the Principal Component Analysis (PCA) ranking procedure on results from each climate x to derive its ranking score r_x .
 Step 8: Perform PCA on the global error metrics taken as a mean from all KG climates in order to build the global rank r_G .
 Step 9: Draw conclusions on model performance in each climate, globally, and with regards to the best Linke turbidity combination.

4.1. Clear-sky period extraction

Since this study focuses on the performance of radiation models under clear-sky conditions, a preliminary task consists in selecting cloudless periods [12]. Under such clear conditions, the estimated GHI from a clear-sky model ($E_{\text{ghc}}^{\text{cs}}$) is directly comparable (and ideally equal) to its ground measurement ($E_{\text{ghc}}^{\text{gm}}$).

A recent review of clear-sky period extraction by Ref. [108] found 21 possible procedures in the literature. A recommendation was made to use multi-criteria methods and to apply a duration filter on the final result. Based on the discussion in Ref. [26], confirmed by the good overall performance noted in Ref. [108], the Reno method [12] has been selected. The method only requires the input of GHI ground measurements and corresponding clear-sky estimates, E_{ghc} , and thus is easy to implement. A moving window of 10-min is defined and is moved 1-min at a time. At each window, five criteria are evaluated from the GHI observations and corresponding clear-sky predictions to determine whether it is clear or not.

It is emphasised that the Reno method is assumed to be relatively insensitive to the choice of radiation model that it relies on to estimate the associated clear-sky irradiance. This stems from an important statement made by the authors of that study to the effect that “as long as the clear-sky model is reasonably accurate, we do not believe that choice of the specific model is critical.” [12]. Since some of the ground stations are located in areas with high concentrations of aerosols, such as the Shagaya site in Kuwait [103], the REST2v5 model is selected here to evaluate the clear-sky GHI as input to Reno's method because it has shown promising results over high-AOD areas [109].

As recommended by Ref. [108], a duration filter is added, so that a full consecutive hour must be identified as clear by the Reno method in order for those periods to be retained. This revised method is noted Reno_w.

The benefit of applying this filter can be immediately noticed from Fig. 2. The green lines indicate where the original Reno clear-sky detection (CSD) method diagnoses clear periods that are actually considered cloudy by the optimised Reno_w method. The latter method is more conservative because it identifies less clear periods than the

original version. This in turn is expected to improve the validity of the present performance analysis by decreasing the risk of cloud contamination.

4.2. Model performance evaluations

For high-quality model performance assessment, the suggestions from Ref. [21] are followed. Hence, thirteen performance metrics are considered and can be categorised into three classes:

- Class A: indicators of the dispersion (e.g. error) of individual points. These are the mean bias difference (MBD), root mean square difference (RMSD), mean absolute difference (MAD), standard deviation of the residual (SD), coefficient of determination (R^2), slope of best-fit line (SBF), uncertainty at 95% (U_{95}) and the t-statistic (TS). A value of 0 is perfect performance for class A metrics with exception of the SBF and R^2 where 1 rather indicates perfect performance.
- Class B: indicators of overall performance. These are Nash-Sutcliffe's efficiency (NSE), Willmott's index of agreement (WIA), and Legates's coefficient of efficiency (LCE). A maximum value of 1 indicates a perfect model.
- Class C: indicators of distribution similitude. These are the Kolmogorov-Smirnov test integral (KSI), the critical limit over-estimation (OVER) and the combined performance index (CPI). A value of 0 for all Class C metrics indicates excellent performance.

In Ref. [21], unfortunate printing errors occurred in the formulas describing the coefficient of determination (R^2) and the MBD. For the present case, the NSE in Class B becomes equivalent to R^2 and is thus ignored in what follows.

To acknowledge that the ground data used in this study inherently contain measurement errors (which are normally less than the modelling errors [109]), the word ‘error’ is replaced with ‘difference’, e.g. RMSD.

4.3. Ranking clear-sky models

Over the years, a number of studies have proposed various model rankings with different combinations of metrics. Ref. [26] employs the relative RMSD, relative MBD, and Pearson's correlation to rank models as “poor”, “average”, “good” and “excellent”. Alternatively, Ref. [13] uses a broad criterion of scores in both RMSD and MBD to distinguish clear-sky model performance. By using both the normalised RMSD and normalised MBD [17], separates models into relative performance instead of ranking them against each other. In a similar fashion, Ref. [18] assessed model performance with the help of MBD, SD and standard deviation of the bias. Ref. [21] developed the CPI, which consists of an empirical blend of KSI, OVER and RMSE, as a ranking index. Many comparative studies use qualitative measures such as the Taylor diagram [110] to visualise performance between different methodological approaches. A fundamental objective of our research is to provide definitive evaluations and ranking between models in order to ascertain which clear-sky irradiance model performs the best; qualitative measures are thus unsuited to this objective. As there is currently no consensus on how to rank clear-sky models, a novel and statistically rigorous approach is developed here. A thorough test is introduced to statistically combine all 13 performance metrics described above and eventually produce an overall ranking score, r_x , where x stands for one

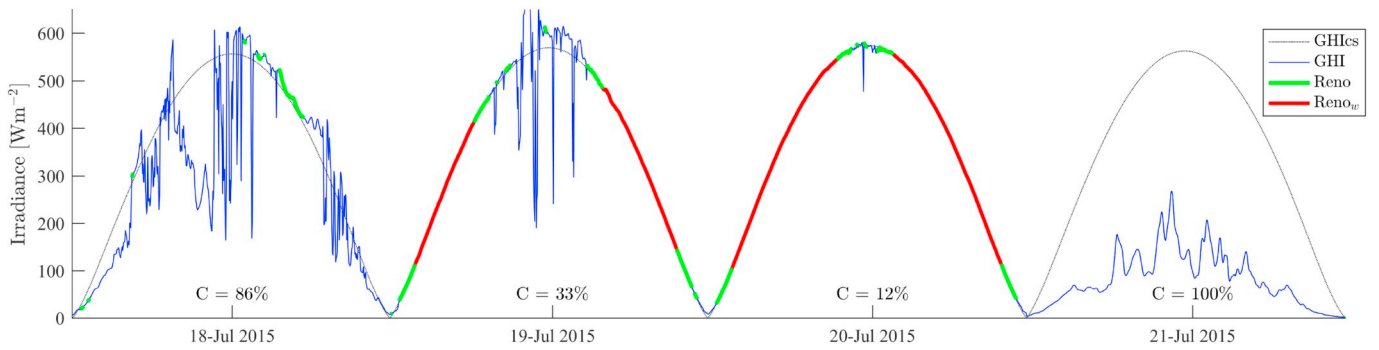


Fig. 2. Example performance of the Reno CSD methodology (green or red) and the optimised Reno CSD (Reno_w) with the new moving window criteria (red only). The measured GHI and modeled clear-sky GHI from the REST2v5 are displayed in blue and black, respectively. The daily-mean cloud amount C is indicated for each day. A sequence of four successive winter days at Adelaide Airport, Australia, is selected for illustrative purposes because it shows a mixture of cloud situations. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

of the five climate classes or the global climate. This is achieved using the Principal Component Analysis (PCA) method, which identifies the principal components contributing most to the overall variance. In most cases, a smaller sample of components account for the majority of the variance, thus reducing the number of dimensions belonging to a dataset. To be specific, PCA determines to what extent each variable p is the cause of variance within a set of potentially correlated variables. Exploring and presenting the cross-correlation between p number of variables would yield a number of $p(p-1)/2$ potential scatterplots to analyse. In the present case, there are $p = 13$ dimensions due to the 13 metrics, each with 75 samples (one per model). The degree of variance within each dimension corresponds to the typical dispersion of an error across the 75 different clear-sky models.

When variables are of different units or across different ranges as is the case here, it is important to first standardise the data before performing the PCA analysis, which is performed here with the “pca” function in Matlab. To standardise the error metrics, the two steps below are followed:

1. Transform all error metrics so that a larger value indicates the best performance. This is achieved as $M_t = \max|M| - |M_t - B|$, where M_t is the transformed error metric of M and B is the best possible score for that metric, such that $B = 0$ if 0 indicates best performance. Note that it is now assumed that underestimation and overestimation expressed by MBD are equally interpreted as poor performance.
2. Standardise all data by taking the z-score (also known as the standard score, or normal score) such that the mean is equal to 0 and variance is equal to 1. Larger values still represent better performance.

PCA produces a number (i) of principal components that is equal to the number of input variables. Each principal component (i) has a coefficient (c) for each error metric (m), and so $c_{i,m}$ refers to the coefficient of the i^{th} principal component of the m^{th} metric. In the present case, the 13 error metrics translate into 13 principal components, each with a set of 13 coefficients. This is best represented in Table 4 where the first five principal components' coefficients are displayed for all 13 metrics.

Also returned by PCA is the explained variance e attributed to the i^{th} principal component (e_i); the sum of all e_i is equal to 100%. For example, in the case of the 13 metrics used in this study for the global category, the percentage of variability that can be accounted for (e_i) from the first five principal components ($c_{1-5,m}$) are 90.46%, 5.32%, 2.82%, 0.86%, and 0.42%, respectively; the remaining eight components combined contribute <0.046%. The total amount of variance that should be accounted for is often subjective, hence the selected amount is what is deemed ‘sufficient’ by the user [111]. Ref. [112] summarises three methods of selecting the appropriate variance accounting limit: (i)

Table 4

First five PCA principal component coefficients (c_i), where i is the principal component that is ordered by e_i , i.e., how much variability can be explained by each principal component ($e_1 = 90.46\%$, $e_2 = 5.32\%$, $e_3 = 2.82\%$, $e_4 = 0.86\%$, and $e_5 = 0.42\%$). These coefficients are derived from the Global classification results.

Metric	c_1	c_2	c_3	c_4	c_5	$\sum_{i=1}^5 e_i c_i$
MBD	0.2900	−0.0663	0.0824	0.1152	−0.0411	0.2620
RMSD	0.2885	−0.0719	−0.1232	−0.2366	−0.3135	0.2503
MAD	0.2903	−0.0804	−0.0601	−0.0829	−0.1103	0.2555
SD	0.2885	−0.0719	−0.1232	−0.2366	−0.3135	0.2503
SBF	0.1777	0.9409	−0.1743	0.1847	−0.1342	0.2070
U ₉₅	0.2885	−0.0719	−0.1232	−0.2366	−0.3135	0.2503
TS	0.2565	−0.0584	0.7199	0.5078	−0.2977	0.2523
R ²	0.2709	−0.2086	−0.4692	0.4393	0.1876	0.2253
WIA	0.2827	−0.1259	−0.2487	0.4177	0.2837	0.2468
LCE	0.2902	−0.0816	−0.0686	−0.0404	−0.1601	0.2552
KSI	0.2871	0.0787	0.1870	−0.2176	0.3936	0.2689
OVER	0.2868	0.0814	0.1966	−0.2304	0.3818	0.2689
CPI	0.2871	0.0779	0.1874	−0.2244	0.3777	0.2688

accept only the components that account for 80–90% of the total variance; (ii) use the so-called Kaiser rule that discounts all components that have an eigenvalue less than 1 [113]; and (iii) perform a ‘Scree plot’ and determine at what component the slope of contribution plateaus.

As PCA is primarily intended as a dimensionality reduction technique, using more relaxed accounting is desirable to achieve the most significant reduction. Because the present analysis only considers 13 dimensions, this number is deemed suitable. Thus strict accounting criteria can be used, following a modified version of each of the three recommendations: account for >99.9% of the variance (as per a strict method i just above); rejection of any component contributing less than 0.05% (as per strict method ii); and subjective inspection of a noticeable flat line of contribution from the 6th onwards (as per method iii). Ultimately, the first five principal components are accepted as they satisfy these criteria. By only considering $i = 1$ to 5, the coefficients $c_{1-5,m}$ can be used to calculate a ranking score r_x per climate $x = (G, A, B, C, D, E)$.

The method of deriving a ranking score r_x from PCA requires the combination of the explained variance e_i , the standardised error metrics (M_m , where $m = 1$ is the MBE, $m = 2$ is the RMSD, ..., $m = 13$ is the CPI), and the first to fifth principal component coefficients of each metric m ($c_{1-5,m}$), such that:

$$r_x = \sum_{i=1}^5 e_i \left(\sum_{m=1}^{13} M_m \times c_{i,m} \right) \quad (10)$$

Note that a larger r_x represents better performance, and vice versa. This way, a ranking score r_x can be derived for each of the five climates

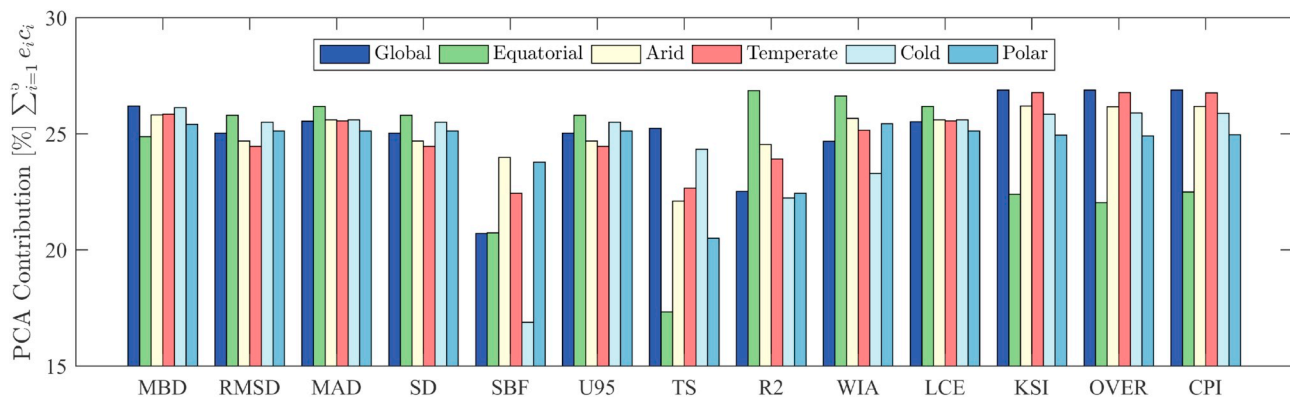


Fig. 3. Summary of the total principal component contribution of each metric within each climate category and globally, as expressed by Eq. (10).

(r_A , r_B , ...). The development of a meaningful global ranking score r_G is not as simple, however. There are many options available to develop r_G , such as (i) taking the mean of all r_x , (ii) performing a new PCA analysis on all 75 r_x values, (iii) calculating error metrics from all data available in the study, (iv) taking a representative sample size from each climate [114], or (v) taking the mean of each error metric across each climate. Options (i), (ii) and (iii) all result in biasing a particular climate: options (i) and (ii) bias the climate with largest variance and range in r_x , whilst (iii) biases the climate with the largest sample size. Option (iv) lowers the overall sample size, which is also not desirable. Hence, the best and simplest solution is option (v), which is followed below. The only downside is that taking a mean of a metric that can be either positive or negative with best performance at 0 (which applies here to MBE only) could potentially improve the error metric through averaging. A simple way around is to take the absolute value of MBE before averaging across all climates. Through this procedure, r_G can be treated with the same PCA procedure as with r_x for the other climates.

The PCA coefficients and overall effective weighting displayed in Table 4 for the global climate, as well as their equivalent per individual climate (see supplementary material), are displayed for each metric and all climates in Fig. 3. A value of 0% for a particular metric would indicate that it is not considered when ranking the models. Clearly, only subtle differences can be seen in the PCA contribution of most metrics. A fairly consistent weighting is observed between 22 and 26%. SBF and TS behave differently, depending on climate, and have somewhat lower contributions of 16–24%. As all metrics contribute with some significance, the suggestion by Ref. [21] to use a wide variety of metrics is corroborated here. In contrast, CPI is not found a systematically better metric, but interestingly its contribution is essentially equal to that of KSI or OVER. It is emphasised, however, that the results in Ref. [21] focused on DNI, whereas the present results focus on GHI, which is much less impacted by uncertainties in aerosol inputs than DNI.

5. Results and discussion

This section describes results of the worldwide validation of 75 clear-sky models at 75 ground stations scattered over five climate regions and the whole world. To fully display these results would require many repetitive and lengthy tables. For brevity, only a summary of the global rank alongside the global and climate-specific ranking scores r_x , as well as global error metrics, are presented in the main body of the article. The interested reader will find all error metric results for each individual climate and model in the supplementary material, where identical tables can be found. The supplementary material also provides a spreadsheet that contains all the raw error metrics and PCA coefficients.

For model performance analysis, four types of result are available: (i) all raw 13 error metrics; (ii) the PCA-derived ranking score r_x ; (iii) an ordered model overall rank in decreasing order of best performance

from 1 to 75; and (iv) box plots of 3 prominent error metrics for each model at all 75 sites. Result types (i), (ii) and (iii) are provided on a per model per climate.

The r_x statistics are discussed first. They are summarised to offer improved understanding of how they potentially differ between climates in subsection 5.1. Next, any climate specific observations of model performance is reported in subsection 5.2 before discussing the global performance and errors in subsection 5.3. Lastly, in subsection 5.4, an evaluation of the different Linke turbidity formulations is made to establish patterns and trends between the different options when applied to the Linke-based models. All the results are available as supplementary material.

5.1. Ranking score significance

For absolute clarity, two definitions exist when referring to rank: (i) the overall rank, and (ii) the ranking score r_x . The overall rank is the relative position of models (e.g. first, second, third etc.) when models are ordered by r_x . Note that r_x was thoroughly described in subsection 4.3. Whilst the overall rank is useful in distinguishing performance between models, it is actually the r_x score that truly matters in terms of performance. Each model is assigned an r_x score per climate region and globally (where $x = (G, A, B, C, D, E)$). Note that, in terms of r_x , the difference from first place to second place might be extremely small, but could be large between first and last rank. For this reason, a brief analysis of r_x results is offered below to assess their significance.

Box plots of r_x are presented in Fig. 4, where the extent of the four quartiles are indicated, as well as any outliers. The inputs to each box plot are the full 75 r_x scores per x , thus making each climate directly comparable to each other. Furthermore, statistics of the r_x scores are displayed in Table 5. As a side note, the mean r_x values are all approximately equal to 0, i.e. climate-independent.

A higher 50th percentile (median indicated by horizontal red line in box plot) P50 indicates that, collectively, models perform better in that climate and can be interpreted as a benchmark for reasonable performance. The Global and Polar climates have the smallest P50, thus indicating that success is generally harder under those conditions. The minimum r_x is always generated by the same model #13 Fu & Rich, which did not perform well (as discussed later).

P75 can be interpreted as a minimum threshold for good performance. A significantly higher P75 and a large interquartile range are found for the Polar climate. This means that there is a broader range of performance in this climate and that it will be highly competitive for top overall rank. The smallest maximum and P75 r_x values occurs in the Cold climate, making this climate also difficult to perform in. With a wider range of performance across all models, it is natural that errors offer larger variance. Recalling that PCA analysis depends on standardisation, having a larger range of performance results in more models being situated further from the mean, as such r_E can obtain a higher

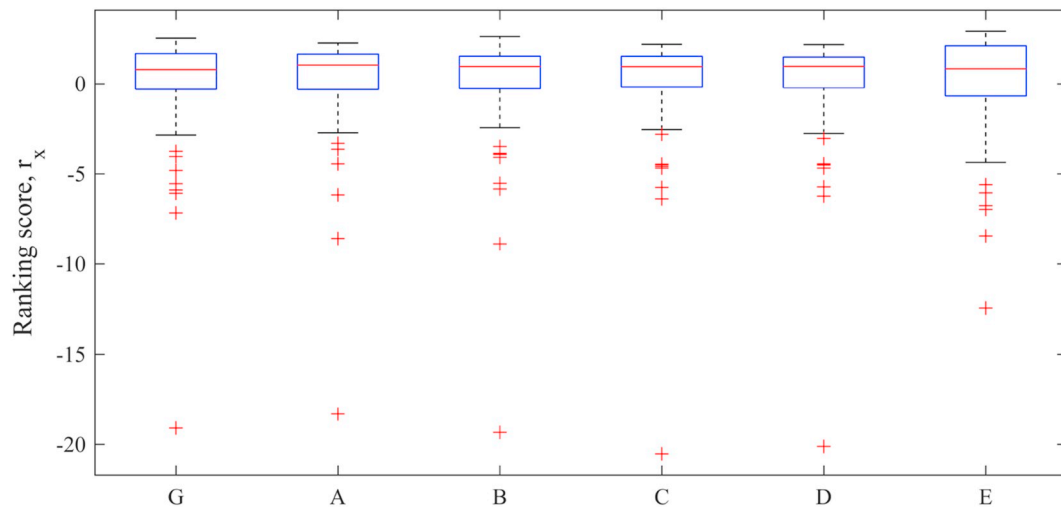


Fig. 4. Boxplots of the r_x scores within each climate $x = (G, A, B, C, D, E)$ indicating Global, Equatorial, Arid, Temperate, Cold and Polar, respectively.

Table 5

Statistical metrics of the ranking scores r_x within each climate x : Global, Equatorial, Arid, Temperate, Cold and Polar. P_n represents the n^{th} percentile. The mean (r_x) is 0 for all six climate categories.

Climate	min (r_x)	max (r_x)	P25 (r_x)	P50 (r_x)	P75 (r_x)
Global	−19.0970	2.5260	−0.3165	0.7900	1.6740
Equatorial	−18.3170	2.2560	−0.3272	1.0380	1.6420
Arid	−19.3350	2.6190	−0.2828	0.9580	1.5265
Temperate	−20.5260	2.1870	−0.1762	0.9490	1.5222
Cold	−20.1190	2.1710	−0.2342	0.9670	1.4828
Polar	−12.4450	2.9040	−0.6947	0.8360	2.1048

score. The supplementary material reveals that the worst MBE/RMSD [W m^{-2}] for climates G, A, B, C, D and E are 36.4/37.8, −29.8/30.7, −27.1/28.3, −31.8/33.0 and −41.5/42.7, respectively. By taking a z-score of error metrics whereby the mean must be scaled to 0, those climates with more extreme metric scores (notably the Polar climate) are able to ultimately achieve better r_x scores.

Remarkably, the r_x metric results are very similar over all climate regions and globally. A score of >2 is excellent, indicating performance more than 2σ . This means that the rank assigned to a model through ordering r_x between climate regions is meaningful, and that the KG climate categorisation is actually relevant, e.g. a rank of 15 in climate A can be assumed similar to a rank of 15 in climate B.

A model performing well in one climate and poorly in another denotes erratic performance, which is not desirable. This powerful detection process is an advantage and a novelty of the present approach. Fig. 5 is provided to show the error metric distribution at all 75 ground stations for each model; we observe that those models with the least variability are the same as ranked highly with r_x .

5.2. Climate specific results

The overall ranks and r_x scores are displayed in Table 6 for all climates and globally. For more climate-specific detail, the interested reader is directed to the supplementary material, where all results can be found.

The first important observation from Table 6 is that the erratic performance between climates as previously described is relatively frequent. It is observed that all models offer unusual performance in at least one climate, resulting in dramatically better or worse performance there than anywhere else. For the top four globally ranked models (MAC2, REST2v5, REST2v9.1 and Bashahu) the exception to their high scores occurs for the Polar climate. In contrast, it is the Arid climate for

the 5th ranked model #65 PR, etc. It is argued that erratic performance is indicative of overfitting a dependency to particular input variables. In the case of the Polar climate (and possibly Cold climate too), it is also possible that the harsh environmental conditions and low-sun conditions significantly impact the measurement uncertainty. These harsh climates additionally feature low aerosol loads and very high surface albedo. The latter circumstance can differently impact the performance of models depending on whether or not they take the surface albedo into consideration (most simple models do not). The other KG climates also have specific features and difficulties. In Arid climates, AOD can fluctuate rapidly over time, whereas water vapour remains typically low. Equatorial sites are dominated by significant levels of rainfall and water vapour that fluctuate in significant amounts on a sub-daily temporal resolution. They may also be impacted by smoke from seasonal forest fires. It is, therefore, easy to understand how overfitting may occur. Models are typically built with ground data from the Arid and/or Temperate climates, due to predominant availability of irradiance data. It is worth reminding the reader that this study focuses on GHI. It is likely that the difficulties mentioned are caused by significant discrepancies in the modelling of the diffuse and/or direct components, which should be explored in a future study. (All high-performance models, except Ineichen's Solis, evaluate the diffuse and direct components separately before adding them to obtain GHI.)

The impact of low sun angles may be important not just in Polar regions: (i) radiation models tend to perform worse when the air mass becomes large, thus exacerbating any inadequacy in attenuation modelling; and (ii) the performance of radiometers used to provide the reference measurements also tends to degrade. This is particularly the case for pyranometers that sense GHI, because of their imperfect cosine response, as documented elsewhere, e.g., Refs. [115,116,117], but other factors can impact DNI and DIF as well (e.g., shading). As preliminary tests have shown, the performance results described here would have been different (particularly over Polar regions) if the modeled GHI had been compared against the pyranometer-derived GHI rather than against the higher-accuracy component-sum GHI, as done here to follow the best practices. Another potential cause of variability in the specific case of the Cold climates could be that its ground stations are highly geographically dispersed (c.f. Fig. 1) with some sites having extreme latitudes, high altitude, island coastal, and all experiencing potential rapid albedo changes through snowfall at seasonal intervals.

One interesting feature in the model results is presented by the #47 CLS model which has impressively improved performance in the Polar climate. It takes first place with $r_E = 2.904$ - the best score in the entire study—yet ranking 56th, 61st, 66th, 61st and 63rd in G, A, B, C and D climates, respectively. Similar behaviour is presented by the #15 KASM

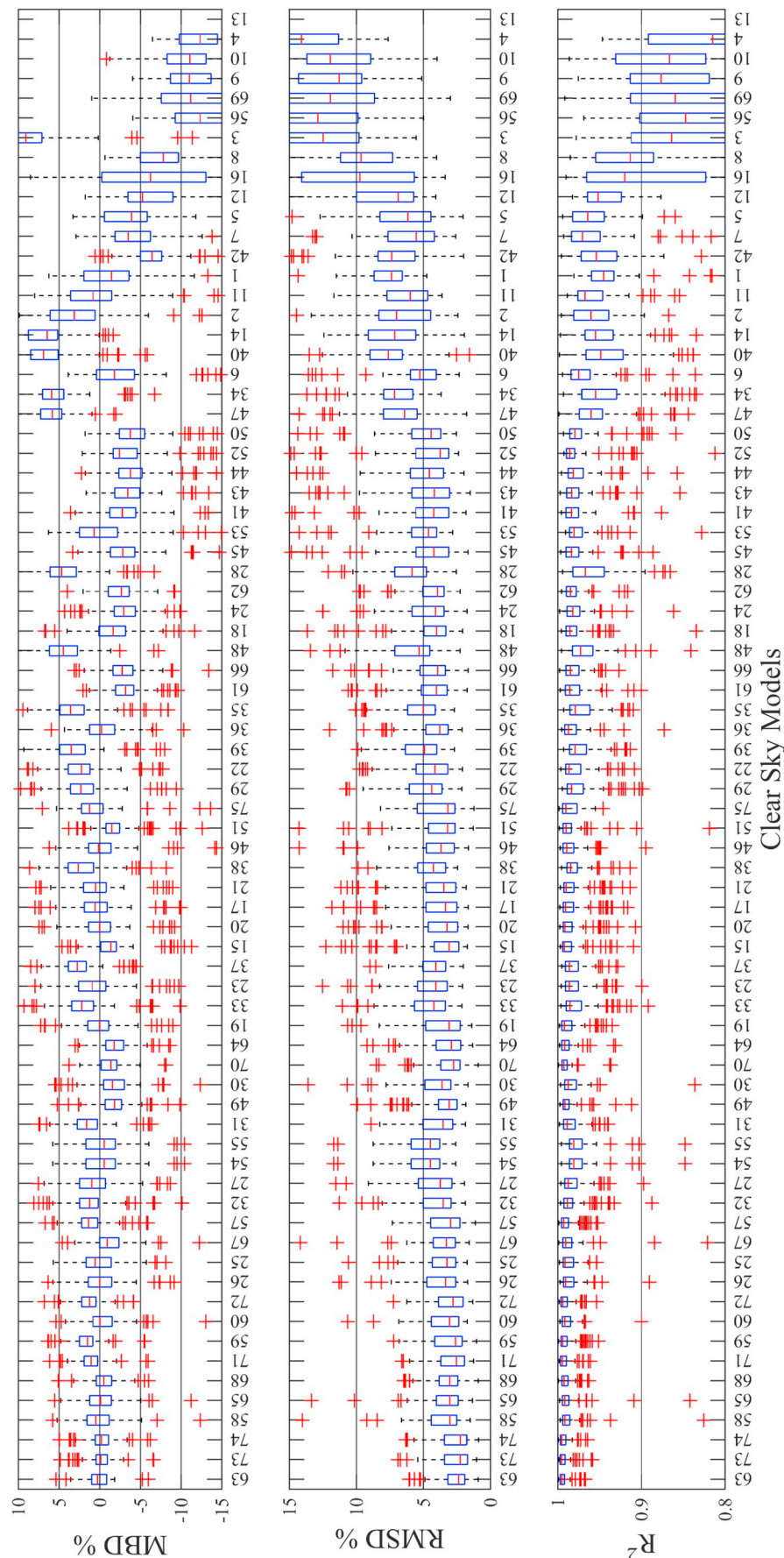


Fig. 5. Box plots of performance metrics for each of the 75 clear-sky models (Table 1) ordered by global performance ranking score r_G (Table 6), where each box-plot contains the result from each of the 75 ground sites (Table 3). The model number can be found in Table 1. The box edges indicate the interquartile range between Q2 and Q3. The median is represented by the horizontal red line, whereas the dashed line indicates Q1 and Q4, and plus signs (+) indicate outliers. The performance metrics are the mean bias difference (MBD), the root mean square deviation (RMSD), and the coefficient of determination (R^2). The y-axes are scaled to represent the majority of models. This leads to hiding some outliers and quartiles of poor performing models. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 6

Overall rank and ranking score r_x of each of the 75 clear-sky models in the six climate categories: G, A, B, C, D and E, representing global, equatorial, arid, temperate, cold and polar, respectively. The models are ordered according to the best global ranking score r_G , where a higher score is better.

#	Model	G	A	B	C	D	E	r_G	r_A	r_B	r_C	r_D	r_E
63	MAC2	1	4	3	1	2	14	2.526	2.174	2.494	2.187	2.076	2.308
73	REST2v5	2	5	2	4	1	15	2.510	2.153	2.595	2.033	2.171	2.295
74	REST2v9.1	3	1	1	2	3	26	2.388	2.256	2.619	2.154	2.019	1.792
58	Bashahu	4	16	10	3	10	18	2.279	1.676	1.864	2.046	1.677	2.139
65	PR	5	8	15	12	6	5	2.254	1.943	1.718	1.710	1.822	2.625
68	MRMv6.1	6	7	5	18	25	8	2.171	1.959	2.236	1.527	1.398	2.477
71	Iqbal-c	7	22	6	8	4	13	2.144	1.387	2.083	1.850	1.865	2.348
59	MMac	8	20	17	9	17	6	2.129	1.612	1.623	1.796	1.538	2.611
60	Yang	9	6	20	6	7	20	2.094	1.994	1.498	1.925	1.780	2.092
72	Iqbal-c	10	33	9	14	8	10	2.036	1.162	2.002	1.674	1.693	2.412
26	Heliosat1-Gu	11	15	8	7	15	30	2.007	1.701	2.022	1.897	1.604	1.459
25	Heliosat1-I	12	25	11	13	24	21	2.003	1.330	1.817	1.692	1.400	1.977
67	MRMv5	13	19	27	27	23	9	1.880	1.652	1.320	1.341	1.439	2.419
57	PSIM	14	21	13	15	22	25	1.856	1.391	1.758	1.573	1.451	1.847
32	Esra-Gu	15	38	19	19	9	19	1.824	1.038	1.536	1.527	1.688	2.109
27	Heliosat1-M	16	37	12	11	16	36	1.759	1.039	1.815	1.713	1.592	1.114
54	CEM	17	30	44	28	46	2	1.750	1.195	0.590	1.336	0.497	2.841
55	Atw.&Ball-2	18	31	45	29	47	3	1.750	1.195	0.590	1.336	0.497	2.841
31	Esra-I	19	45	30	32	11	7	1.725	0.433	1.280	1.274	1.645	2.578
49	Josefsson	20	13	7	30	18	41	1.521	1.793	2.080	1.291	1.511	0.790
30	Esra-D	21	26	37	22	34	22	1.519	1.329	0.959	1.397	1.124	1.916
70	Bird	22	10	16	35	28	31	1.494	1.855	1.704	1.070	1.295	1.316
64	Metstat	23	2	14	26	35	40	1.475	2.225	1.754	1.351	1.109	0.801
19	Kasten-I	24	24	18	10	12	43	1.436	1.342	1.572	1.760	1.637	0.730
33	Esra-M	25	51	31	34	20	24	1.417	0.075	1.194	1.077	1.476	1.883
23	Heliosat1-R	26	42	23	24	26	38	1.373	0.788	1.435	1.386	1.339	0.836
37	Heliosat2-I	27	57	22	38	27	23	1.358	-0.362	1.449	0.949	1.315	1.895
15	KASM	28	3	4	16	31	52	1.328	2.199	2.272	1.532	1.190	-0.361
20	Kasten-Gu	29	23	25	20	14	44	1.323	1.349	1.385	1.508	1.607	0.664
17	Kasten-R	30	34	24	17	13	47	1.270	1.143	1.429	1.530	1.622	0.544
21	Kasten-M	31	36	26	23	21	45	1.202	1.060	1.370	1.393	1.455	0.602
38	Heliosat2-Gu	32	49	29	42	36	28	1.195	0.174	1.289	0.744	1.063	1.513
46	Inei.&Per-Gr	33	27	21	5	19	50	1.139	1.308	1.451	1.946	1.485	-0.215
51	Simp. Solis	34	18	28	21	32	48	1.092	1.661	1.315	1.507	1.154	0.120
75	McClear	35	67	42	58	5	11	1.032	-2.483	0.755	-0.227	1.848	2.403
29	Esra-R	36	55	40	40	29	29	1.026	-0.220	0.843	0.808	1.268	1.471
22	Kasten-Gr	37	52	34	41	37	39	0.830	0.021	1.084	0.789	1.045	0.824
39	Heliosat2-M	38	58	32	50	42	33	0.790	-0.757	1.145	0.316	0.807	1.185
36	Heliosat2-D	39	40	53	46	33	34	0.785	0.816	0.028	0.478	1.127	1.166
35	Heliosat2-R	40	59	33	48	41	37	0.753	-0.856	1.131	0.435	0.814	0.913
61	Calinoiu	41	41	35	36	39	49	0.689	0.808	1.034	1.051	0.897	-0.116
66	Paulescu	42	39	52	45	48	32	0.681	0.830	0.034	0.564	0.492	1.226
48	King	43	54	43	56	56	27	0.664	-0.048	0.638	-0.096	-0.226	1.790
18	Kasten-D	44	44	46	44	38	46	0.657	0.636	0.556	0.586	0.967	0.571
24	Heliosat1-D	45	43	47	52	44	35	0.654	0.655	0.449	0.100	0.556	1.129
62	Hoyt	46	48	39	54	40	42	0.589	0.198	0.903	0.071	0.890	0.762
28	Heliosat1-Gr	47	64	57	57	45	16	0.485	-1.717	-0.323	-0.203	0.528	2.266
45	Inei.&Per-M	48	9	51	25	49	53	0.449	1.889	0.074	1.354	0.111	-0.444
53	Perrin	49	47	56	33	30	56	0.334	0.203	-0.162	1.225	1.216	-0.652
41	Inei.&Per-R	50	11	50	31	50	57	0.318	1.818	0.188	1.277	0.108	-0.709
43	Inei.&Per-I	51	14	49	39	59	51	0.243	1.715	0.416	0.827	-0.448	-0.344
44	Inei.&Per-Gu	52	28	58	37	54	54	0.132	1.284	-0.437	0.976	-0.168	-0.472
52	Adv. Solis	53	35	41	49	57	60	-0.054	1.104	0.801	0.334	-0.237	-1.371
50	Badescu	54	29	38	53	55	59	-0.092	1.207	0.958	0.080	-0.171	-1.362
47	CLS	55	61	66	61	63	1	-0.146	-1.286	-2.081	-0.747	-0.980	2.904
34	Esra-Gr	56	69	62	63	52	12	-0.207	-2.727	-1.098	-1.273	-0.065	2.399
6	ASHRAE	57	32	48	43	43	64	-0.353	1.192	0.420	0.718	0.596	-2.578
40	Heliosat2-Gr	58	70	59	65	62	17	-0.444	-3.312	-0.662	-1.596	-0.835	2.143
14	Atw.&Ball-1	59	68	63	64	64	4	-0.447	-2.580	-1.279	-1.283	-1.378	2.671
2	Schulze	60	66	36	51	51	61	-0.555	-2.460	0.982	0.146	-0.028	-1.691
11	Kumar	61	60	55	59	60	62	-0.837	-0.913	-0.106	-0.312	-0.526	-1.879
1	TJ	62	56	54	47	53	65	-0.896	-0.223	0.005	0.457	-0.134	-3.2840
42	Inei.&Per-D	63	62	67	62	65	55	-1.455	-1.358	-2.446	-1.091	-1.616	-0.510
7	Sharma	64	12	60	55	61	68	-1.502	1.814	-0.706	-0.072	-0.538	-4.364
5	Biga	65	17	61	60	58	69	-1.667	1.670	-1.094	-0.341	-0.336	-4.365
12	Campbell	66	50	64	66	68	67	-2.444	0.116	-1.307	-2.126	-2.772	-3.998
16	Capderou	67	63	65	68	67	66	-2.860	-1.404	-1.937	-2.809	-2.372	-3.659
8	El Mghouchi	68	53	71	67	66	70	-3.759	-0.010	-4.082	-2.557	-1.928	-5.596
3	DPP	69	74	70	70	69	58	-4.050	-8.583	-3.918	-4.504	-3.041	-0.956
56	RSC	70	73	69	73	73	63	-4.813	-6.161	-3.869	-5.749	-5.720	-2.141
69	Janjai	71	46	72	72	74	71	-5.542	0.215	-5.519	-4.677	-6.226	-6.052
9	YangWalsh	72	65	73	71	70	73	-5.888	-1.920	-5.836	-4.586	-4.449	-6.962
10	HLJ	73	71	68	69	71	74	-6.072	-3.638	-3.490	-4.461	-4.509	-8.433

(continued on next page)

Table 6 (continued)

#	Model	G	A	B	C	D	E	r_G	r_A	r_B	r_C	r_D	r_E
4	Adnot	74	72	74	74	72	72	−7.157	−4.451	−8.877	−6.386	−4.682	−6.757
13	Fu & Rich	75	75	75	75	75	75	−19.10	−18.32	−19.34	−20.53	−20.12	−12.45

in Equatorial and Arid, #75 McClear in Cold and Polar, and #14 Atw.& Ball-1 in Polar, yet delivering only very poor performance in all other climates. The Atw.&Ball-1 phenomenon just mentioned is perhaps due to the dynamic calculation of surface albedo with solar zenith angle that corrects the estimates very well for that climate. Despite the #63 MAC2 and #73 REST2v5 only scoring an overall rank of 14th and 15th, the r_E scores obtained are 2.308 and 2.295, respectively. These are both highly impressive scores that would have earned an overall rank of 1st in either the Equatorial, Temperate or Cold climates, thus their lower rank in the Polar climate is not indicative of poor performance. On the contrary, these two models consistently offer excellent performance. Hence, it is no surprise that they obtain global overall ranks of 1st and 2nd.

The fact that some overall high-ranking models (such as MAC2 or REST2v5) perform poorly in the Polar region, whereas overall low-accuracy models (such as CLS) perform well could be an indication of: systematic compensation of errors that benefit lower-accuracy models, or indication of significant bias in the measurements. Some reasons why the latter are possible have been mentioned previously. A current topic of interest in radiometry under the challenging arctic or polar conditions is related to the importance of proper ventilation, heating, and de-icing of radiometers [118]. Field campaigns, such as D-ICE [119], are being conducted to improve the quality of radiometric measurements under harsh icing conditions. It can thus be expected that stations operating under cold conditions will progressively receive instrumentation upgrades, which will decrease the risk of bias in future radiation measurement. This discussion suggests that the ranking of clear-sky models is less reliable in the Polar region than for all other KG climate classes, and that a follow-up of this study will be desirable when measurement accuracy has eventually improved over these regions.

Most models in the upper half of the rankings are of a hybrid nature, combining empirical and physical modelling approach, whereby the underlying physical mechanisms are parameterised based on available empirical data. One likely consequence of training a model on specific datasets is that it can end up suffering from the overfitting issue, and will not perform as well against datasets representing widely different climates. It is also possible that a model with a small number of atmospheric inputs and very simple—but yet physically sound—representation of radiation transfer throughout the atmosphere is impervious to the propagation of errors caused by imperfect inputs that might plague more sophisticated and detailed models. This might explain why MAC2 is one of the better performing models in all climates, despite its relative simplicity (Table 1).

There is potential to explore the relative methodological differences between different versions of the same model to understand which specific algorithm affect the performance, climate-wise. For instance, the newer REST2 model version (v9.1) improves in the Equatorial, Arid and Temperate climates, but degrades performance in Cold and Polar, ultimately resulting in a lower global overall rank. In the case of the Heliosat models, version 1 curiously performs better than version 2 in every climate. The Atw.&Ball model undergoes drastic improvement from version 1 to 2, although the latter is actually an adaptation by a different author [59]. The MMAC performs slightly worse than the alternative MAC2 formulation (the reader is referred to the supplementary material for distinct differences between these models). MRMv6.1 is a vast improvement over MRMv5, even though the latter performs marginally worse in the Cold climate. Perhaps most surprising is that the Advanced Solis obtains the 53rd rank only, whilst the older Simplified Solis comes in at 34th.

All models display highly varied performance across all KG classes, thus highlighting the importance of this type of climate-disaggregated analysis.

5.3. Global performance results and ranking

The ranks and ranking scores for global performance appear in Table 6 for G and r_G ; the error metrics for each model for G appear in Table 7.

The #63 MAC2 model offers such strong performance across all climates that it unsurprisingly earns the overall global rank of 1st place with $r_G = 2.526$. The #73 REST2v5 is a very close second with $r_G = 2.510$. In the third place, #74 REST2v9.1 has much more separation with $r_G = 2.388$, due to performance degradation over the Cold and Polar regions, which does not compensate for the improvement over the three other regions.

It may come a surprise that #63 MAC2 ranks so highly even though it has previously not validated well. One important reason is that various interpretations have been proposed for this model, particularly regarding its handling of aerosol attenuation. Originally [66], defined aerosol transmittance as $T_r(a) = (\exp(-\tau_a m_r)) = k^{m_r}$, stating that “ τ_a or k must be pre-assigned”. A subsequent study [120] provided estimates of k , in the range 0.87–1.0, for 15 world stations. This suggested a mean value of about 0.95, as used in subsequent validation studies, but leading to poor results [24,121]. A more physical definition of τ_a led to substantially improved results [77,122–124], typically similar to those of REST2, as also found here. More fine tuning would be desirable for this model, so that it can improve its performance under harsh conditions, for instance. More generally, the supplementary material lists all the present interpretations in coding the various models, as these interpretations can strongly influence model performance.

An important part of this study relies on a PCA analysis to identify the most significant error metrics. The use of RMSD, MBD and R^2 in previous validation studies is conventional, but also potentially simplistic or insufficient. Indeed, in many cases the ranking order would have changed if it had relied on just one or more of these three metrics, which confirms the added discriminating power of the PCA methodology. Interestingly, however, the top five models would have still performed much the same if using only the conventional metrics, as can be observed from Table 7.

A particularly strong example of the PCA impact can be seen with the #31 ESRA-I model. It has a very low global absolute MBD within the study with 1.68%, a good RMSD at 4.1%, and a very high R^2 at 0.987. These scores would have been enough to score very highly in previous validation studies (and did so in Ref. [26]). In contrast, the model's OVER and KSI metrics are significantly worse than other models surrounding it, hence, the PCA ranks it lower. Conversely, the overall globally ranked 11th #60 Yang model has a much higher RMSD than surrounding models. In this case, it is the significantly lower TS, MBD and CPI that promotes it. It can be observed that MBD, RMSD and R^2 often fluctuate up or down between ranks. Hence, developing a simple ranking based on any one of these metrics, as done before in, e.g. Ref. [115], would produce a different ranking order relative to the trustworthy and unbiased PCA ranking. The PCA metric weightings in Fig. 3 show that each error metric significantly contributes to the variance of clear-sky irradiance estimation. The list of 13 metrics detailed in subsection 4.2 is likely not exhaustive as other unexplored metrics may reveal significantly contributing variance. The PCA method enables a simple and unbiased process to consider any number of metrics

Table 7

Global performance metrics (G) for all clear-sky models ordered by rank r_G . The model name is not included, however, this table order corresponds to Table 6. Note that MBD is taken as absolute due to PCA rules on the global scale averaging across climates.

#	MBD	RMSD	MAD	SD	SBF	U ₉₅	TS	R ²	WIA	LCE	KSI	OVER	CPI
	%	%	%	%		%	%				%	%	%
63	0.904	3.03	2.22	3.03	0.991	8.41	194	0.993	1.000	0.928	90.3	63.3	68.0
73	0.764	3.15	2.22	3.15	0.984	8.72	159	0.993	1.000	0.927	79.4	45.5	55.2
74	1.10	3.20	2.47	3.20	0.987	8.88	201	0.992	0.999	0.920	106	78.5	78.2
58	1.14	3.94	2.95	3.94	1.00	10.9	182	0.988	0.999	0.903	118	94.8	85.2
65	0.544	3.82	2.74	3.82	0.974	10.6	129	0.989	1.000	0.909	93.1	66.0	62.9
68	0.948	3.63	2.53	3.63	0.968	10.1	200	0.990	1.000	0.917	111	81.3	80.5
71	1.31	3.33	2.47	3.33	0.972	9.24	333	0.992	0.999	0.919	120	90.1	100
59	1.61	3.52	2.56	3.52	0.994	9.77	494	0.991	0.999	0.915	138	109	129
60	1.24	4.01	2.96	4.01	0.978	11.1	241	0.987	0.999	0.903	111	82.9	85.3
72	1.66	3.31	2.57	3.31	0.974	9.18	431	0.992	0.999	0.915	138	114	124
26	1.50	4.38	3.41	4.38	0.996	12.1	228	0.986	0.999	0.889	150	131	111
25	1.56	4.01	3.06	4.01	1.01	11.1	256	0.988	0.999	0.900	170	148	126
67	1.56	4.17	2.93	4.17	0.969	11.6	325	0.986	0.999	0.902	138	112	112
57	2.05	3.81	3.07	3.81	0.997	10.5	466	0.989	0.999	0.900	209	192	174
32	1.65	4.39	3.44	4.39	1.01	12.2	302	0.985	0.999	0.887	172	153	133
27	1.81	4.83	3.80	4.83	0.996	13.4	259	0.983	0.998	0.876	192	170	139
54	1.29	5.99	4.04	5.99	0.996	16.6	160	0.969	0.999	0.865	131	113	92.6
55	1.29	5.99	4.04	5.99	0.996	16.6	160	0.969	0.999	0.865	131	113	92.6
31	1.68	4.10	3.14	4.10	1.03	11.4	394	0.987	0.999	0.896	188	168	152
49	2.40	4.50	3.60	4.50	0.976	12.5	469	0.985	0.997	0.885	202	186	170
30	1.95	4.55	3.47	4.55	0.968	12.6	359	0.984	0.999	0.886	216	198	166
70	2.04	3.95	3.07	3.95	0.952	11.0	484	0.989	0.998	0.902	211	189	175
64	2.75	4.30	3.52	4.30	0.988	11.9	638	0.986	0.997	0.888	233	218	207
19	1.73	4.41	3.40	4.41	0.960	12.2	227	0.986	0.998	0.891	272	249	186
33	2.46	4.93	3.93	4.93	1.02	13.7	445	0.981	0.998	0.870	228	211	183
23	2.23	5.10	4.11	5.10	1.01	14.1	349	0.981	0.998	0.867	255	237	190
37	2.87	4.38	3.71	4.38	0.993	12.1	640	0.985	0.997	0.876	276	262	234
15	2.60	4.94	4.08	4.94	0.993	13.7	461	0.981	0.996	0.872	277	258	214
20	1.50	4.63	3.56	4.63	0.955	12.8	165	0.984	0.998	0.885	296	276	195
17	1.94	4.81	3.74	4.81	0.961	13.3	248	0.983	0.998	0.879	290	270	200
21	1.85	4.82	3.74	4.82	0.955	13.4	225	0.983	0.998	0.879	305	284	206
38	2.77	4.81	4.03	4.81	0.980	13.3	481	0.983	0.997	0.867	297	278	229
46	2.90	5.57	4.50	5.57	1.02	15.4	339	0.975	0.996	0.857	281	254	201
51	2.96	4.93	3.85	4.93	0.975	13.7	552	0.982	0.997	0.877	306	286	242
75	1.51	8.86	4.37	8.86	0.994	24.6	165	0.906	0.999	0.852	169	151	117
29	2.79	5.32	4.30	5.32	1.03	14.8	478	0.979	0.998	0.859	281	266	220
22	3.17	5.17	4.21	5.17	0.963	14.3	543	0.980	0.997	0.862	341	322	263
39	3.56	5.51	4.70	5.51	0.982	15.3	595	0.977	0.996	0.845	344	330	272
36	1.62	4.79	3.73	4.79	0.931	13.3	287	0.983	0.999	0.878	399	388	275
35	3.81	5.66	4.83	5.66	0.996	15.7	653	0.976	0.995	0.840	369	354	293
61	4.07	5.41	4.57	5.41	0.988	15.0	867	0.978	0.995	0.852	355	334	305
66	3.30	5.22	4.06	5.22	0.946	14.5	659	0.980	0.997	0.867	323	312	267
48	4.11	5.78	5.03	5.78	0.976	16.0	744	0.975	0.995	0.834	299	287	260
18	2.30	5.24	4.04	5.24	0.933	14.5	358	0.980	0.998	0.869	381	367	270
24	3.49	5.46	4.30	5.46	0.947	15.1	637	0.978	0.996	0.859	311	297	256
62	3.05	5.14	4.12	5.14	0.929	14.2	605	0.981	0.997	0.866	333	321	267
28	4.14	5.98	4.99	5.98	1.03	16.6	752	0.970	0.995	0.831	323	308	275
45	4.46	6.49	5.33	6.49	1.000	18.0	707	0.967	0.993	0.831	395	381	315
53	3.36	6.83	5.51	6.83	0.958	18.9	401	0.964	0.994	0.821	373	354	268
41	4.60	6.64	5.53	6.64	1.01	18.4	702	0.965	0.993	0.825	392	377	312
43	4.81	6.26	5.42	6.26	1.01	17.3	921	0.970	0.993	0.828	415	403	351
44	5.17	6.85	5.77	6.85	1.000	19.0	868	0.964	0.992	0.815	432	415	354
52	4.84	6.16	5.51	6.16	1.03	17.1	984	0.971	0.992	0.827	472	453	391
50	5.38	6.69	5.98	6.69	0.991	18.5	1042	0.966	0.991	0.811	454	444	389
47	5.25	6.53	5.65	6.53	1.02	18.1	1293	0.963	0.991	0.808	395	383	378
34	4.91	6.98	5.94	6.98	1.05	19.4	921	0.959	0.992	0.799	384	373	332
6	5.38	8.16	6.96	8.16	0.972	22.6	714	0.947	0.987	0.781	411	391	324
40	5.87	7.32	6.50	7.32	1.02	20.3	1091	0.954	0.989	0.780	454	449	395
14	5.76	7.11	6.37	7.11	0.988	19.7	1463	0.956	0.988	0.784	428	419	417
2	5.97	9.00	7.69	9.00	0.983	24.9	598	0.939	0.987	0.752	438	425	330
11	4.35	8.33	7.03	8.33	1.05	23.1	460	0.948	0.990	0.777	594	589	416
1	5.74	10.6	9.01	10.6	0.980	29.5	461	0.914	0.984	0.714	414	403	301
42	7.34	8.90	7.65	8.90	0.961	24.7	1148	0.941	0.985	0.750	606	599	495
7	6.78	9.37	8.13	9.37	0.975	26.0	835	0.927	0.981	0.746	674	664	503
5	6.74	9.70	8.23	9.70	0.945	26.9	764	0.922	0.981	0.742	638	624	471
12	9.01	11.0	9.94	11.0	1.03	30.6	1144	0.906	0.975	0.688	708	695	555
16	9.57	13.6	11.3	13.6	0.985	37.8	699	0.861	0.975	0.638	749	742	536
8	10.5	12.8	11.2	12.8	0.920	35.5	1184	0.871	0.965	0.646	832	826	638
3	10.9	13.9	11.8	13.9	1.12	38.5	1046	0.850	0.970	0.606	736	728	563
56	13.3	14.1	13.3	14.1	1.03	39.1	2167	0.852	0.957	0.564	926	918	800
69	13.7	15.2	13.9	15.2	0.942	42.2	1705	0.824	0.949	0.559	1104	1097	862

(continued on next page)

Table 7 (continued)

#	MBD	RMSD	MAD	SD	SBF	U ₉₅	TS	R ²	WIA	LCE	KSI	OVER	CPI
	%	%	%	%		%	%				%	%	%
9	14.4	15.6	14.5	15.6	0.949	43.2	2004	0.816	0.945	0.536	1120	1115	904
10	14.7	16.1	14.9	16.1	0.914	44.7	1739	0.798	0.940	0.529	1058	1052	837
4	14.8	16.8	15.0	16.8	0.846	46.5	1591	0.790	0.937	0.514	1252	1250	943
13	36.4	37.8	36.4	37.8	1.10	105	2919	−0.024	0.789	−0.183	2214	2213	1676

and how they should be weighted.

Overall, one particular model, #13 Fu & Rich, is found to perform systematically poorly here, with a substantial MBD of −37% averaged across all climates, and an extreme MBD of −52% in the Polar climate. The source of this bias is the overly simplistic representation of atmospheric transmissivity stated as “typical values are 0.6 or 0.7 for very clear-sky conditions and 0.5 for a generally clear sky”. This description is static and subject to interpretation; a value of 0.5 was selected in Ref. [121] and used within this study. Since this a significant source of bias, a larger transmissivity value could be more appropriate—although ideally a site-specific value would need to be defined a priori in order to be considered globally. This is of concern because of the wide usage of this model in GIS applications [121].

It is instructive to revisit the findings in previous validation studies to see whether they are corroborated here. The top three performers here were also clearly identified in Ref. [4] and [77]. The REST2v5 ranks 2nd globally here, and was found to be a top performer in many previous validation studies [13,15,16,18,24,124–128]. In contrast, ESRA ranks 15th globally here, whereas it was found best by Engerer et al. [26]; whilst their study did include models that did rank higher here, limitations of climate diversity and input parameter simplicity may have biased their results. Surprisingly, many models claimed to be of high accuracy in previous validation studies rank very poorly in the present analysis. This can possibly be explained by a number of simplifications made in those studies relatively to insufficient quality control of observational data, high residual errors in input data, lack of climatic diversity, etc. This confirms the importance of the rigorous and unbiased approach followed here. The Bird model, at the 22nd global rank here, was surprisingly found to be a high performer by Refs. [16,124,124,127]. McClear appears only at the 35th global position here, whereas it was ranked highly by Refs. [18,129]. The Ineichen and Perez model only reaches the 46th global position here, yet was very highly rated by Refs. [126,130,131]. Similarly, Perrin is in 49th position globally but was highly rated by Ref. [132], Adv. Solis in 53rd position globally was highly rated by Refs. [15,18,125,127,131,133], and Adnot, which was highly rated by Refs. [45,134], is near the bottom here.

Finally, the results presented in the tables found in the supplementary material clearly demonstrate that the highest-ranked models are normally those that use detailed atmospheric information, most particularly about both aerosols and water vapour. This is consistent with previous findings, e.g. Ref. [121].

5.4. Linke Turbidity results

It is occasionally suggested that T_L is an outdated or overly simplified variable to describe the atmosphere. Remarkably, the present findings do not support this, at least not categorically. The Heliosat1-Gu and -I feature in 11th and 12th globally, the former 8th and 7th in climates B and C; the Inei.&Per.-Gr is 5th in climate C; the Kasten-I features in the top 10 of climate C; whilst the ESRA-I and -Gu feature in the top 10 of climate E and D, respectively.

Based on the performance results presented in Fig. 6, significant performance variation in a single clear-sky model is observed as a result of using alternative T_L formulas. This confirms the importance of testing

each Linke-based model with various Linke formulas. One likely cause for this variance is that each Linke formula has been derived empirically, and usually over temperate climates where neither AOD nor PW could reach large values. Eqs. (9a)–(9f) may not extrapolate correctly under extreme conditions or specific combinations of AOD, PW and air mass.

It would be unfair to compare T_L performance between different clear-sky models, e.g. ESRA-Gu performs better than Heliosat2-I, but this should not be interpreted to mean that Ineichen's T_L is better than Gueymard's, as the difference could be a result of the clear-sky model itself. In other words, it is only possible to evaluate T_L formula performance when comparing all 6 formulations from under the same clear-sky model; 5 repetitions of T_L evaluation are thus necessary to feed the 5 Linke-dependent models.

Overall, Ineichen's T_L provides the best global performance for the Kasten and Heliosat2 models. Gueymard's T_L narrowly beats the Ineichen combination with Heliosat1 and is best with ESRA, and is otherwise very competitive performance to Ineichen's with Kasten and Heliosat2; Grenier's T_L provides the best global performance for the Ineichen and Perez model, yet is the worst for both ESRA, Heliosat1 and Heliosat2; second worst with Kasten. This is perhaps related to the fact that the Heliosat-1, Heliosat2 and ESRA models have similar approaches and were all designed to be used with satellite data, whereas the other two are relatively simple clear-sky models. Dogniaux's, Molineaux's and Remund's T_L formulation yields similar average performance across clear-sky models. Notably, the Dogniaux T_L combined with Ineichen & Perez model performs very badly.

It is worth concluding that the selection of an optimal T_L for each model yields drastically different results. For instance, the r_G values of the Ineichen and Perez-Gr and Ineichen and Perez-D combinations are 1.139 and −1.455, respectively, inducing a strong overall global rank decline from 33rd to 63rd. This is highly significant variance for a difference between the two models that reduces to just one line of code.

Although it is concluded that T_L can still be a useful variable in clear-sky radiation modelling, the choice of the underlying T_L formula is critical. Hence, careful consideration is required to select the most appropriate formula to provide the main input to any T_L -based model. That selection process is climate-specific to some extent, as revealed in Table 6, which further complicates the issue. Furthermore, it is emphasised that, to achieve good or acceptable performance, the Linke-based models must be made dependent on both AOD and PW inputs at high temporal resolution, just as most high-ranking radiation models reviewed here. The apparent simplicity of the Linke-based models is somehow lost in that process. It is highly likely that using any Linke-based model with just a monthly climatology of Linke turbidity factor—as usually done in the literature—would result in severe loss of performance.

6. Conclusion

This detailed contribution offered the most comprehensive and fair worldwide validation study of 75 clear-sky irradiance models and was organised to address seven key research objectives (c.f. bullets in subsection 1.2) and (c.f. subsection 1.3).

These 75 models represent a large range, and quite comprehensive

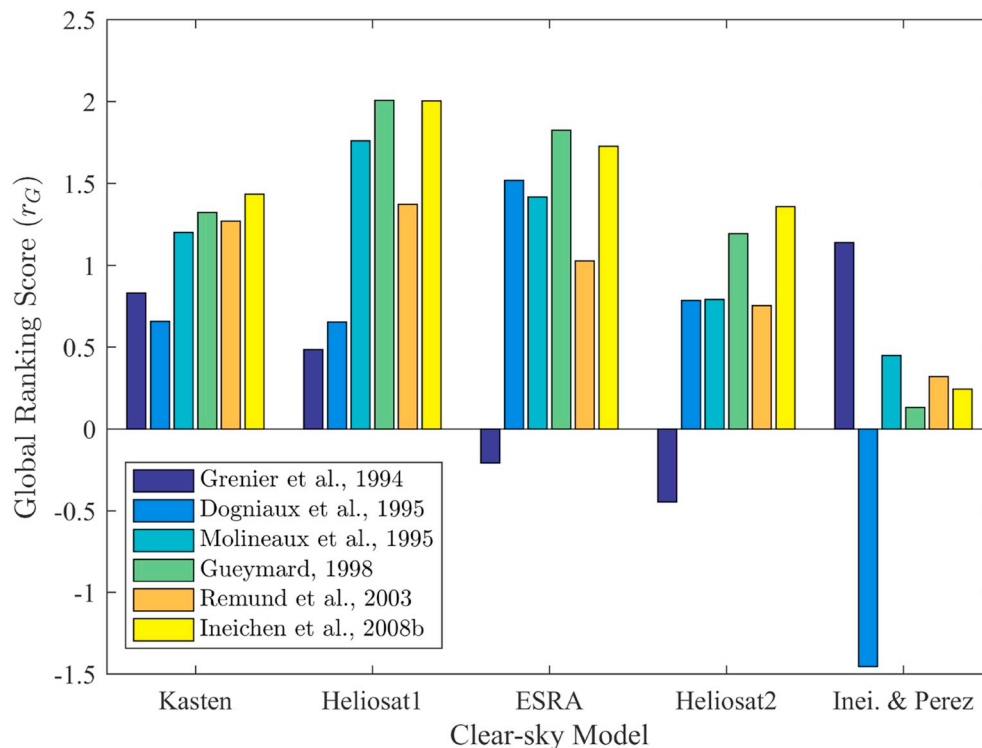


Fig. 6. Bars of the PCA-derived global ranking score r_G of each Linke-based T_L formulation with all Linke-based clear-sky irradiance models.

list, of what the solar literature has to offer in terms of possible calculation methods for the clear-sky broadband global irradiance. Only a few models of the not-recent literature or that were insufficiently described had to be excluded. Because clear-sky radiation models are typically used in a wide variety of applications, a thorough performance assessment methodology, coupled with an innovative ranking analysis, has been devised to offer significant improvements over existing validation studies. The 75 models were tested against high-quality measurements from 75 ground stations, representing the largest sampling of observations in the literature for this kind of validation.

The geographical dispersion of the selected pool of ground stations represents a large diversity of climate zones. Analysing model performance separately for each of the five main climate regions considered here constitutes a substantial improvement over previous studies, and also offers a much fairer test and demonstration of model capabilities when targeting “worldwide” applications. A particular novelty is that many ground stations are located at extreme latitudes or in the southern hemisphere. To operate the models, gridded input data were employed at the highest possible spatio-temporal resolution for each required input variable. After quality control and clear-sky suitability checks, each model was tested against 4,360,012 1-min clear-sky GHI data points, using observations with the lowest possible uncertainty, i.e. based on the component-sum method. This attention to detail constitutes a considerable step-up from previous validation studies, whereby certain variables were fixed or used inaccurately due to lack of availability/accessibility or mismatch in spatio-temporal resolution. The MERRA-2 reanalysis provided all input data at hourly resolution, with the exception of the NO_2 abundance, which was obtained from OMI satellite observations. The Reno clear-sky detection method was selected after considering various options. The method was further improved upon by considering an additional temporal filter for heightened scrutiny. It is argued that there is still a need for a thorough validation study of such clear-sky detection methodologies, particularly when considering longer time-averaging periods of irradiance

measurements, such as at 10-min or hourly intervals.

An important goal of this study was to challenge the conventional method of benchmarking model performance. To do so, 13 conventional or unconventional performance error metrics were calculated first. In a second step, the weighted contribution of each metric to the overall variance in modeled clear-sky irradiance was derived using the Principal Component Analysis (PCA). It was found that all 13 metrics contribute significantly to the ranking score r_x . Hence, it is strongly recommended that this type of multi-metric validation is used in future validation studies.

Six formulas to derive the Linke turbidity factor T_L from aerosol and water vapour data were identified in the literature and tested for performance. These were combined with five different clear-sky models that accept T_L as an input, thus enabling an assessment of the most appropriate T_L methodology. When considering all climate regions, the Ineichen or Gueymard formulas were found to offer the best global performance when combined with all but one T_L -clear-sky model. The Grenier formula performed better for the Ineichen & Perez model. Significant performance differences were observed with the same clear-sky model using different T_L formulation, hence, a careful climate-specific selection of T_L methodology is strongly advised. Overall, the best Linke-based model, Heliosat-1-Gu, only reached the 11th position. Considering that, as implemented here, the intermediate calculation of the Linke coefficient requires the same AOD and PW data as directly used by other models that perform much better, the apparent simplicity of Linke-based models does not compensate for their generally limited accuracy and relative lack of universality.

The performance results obtained here suggest that many hybrid empirical-physical models are built around overfit transmission functions, which introduces erratic performance across climates. This reinforces the importance of assessing the performance of radiation models over as many climatic conditions as possible.

Based on the PCA-derived ranking of all models, the models' performance is segregated below into six distinct categories, with the three

top-ranking models stated in each:

	1st	2nd	3rd
Global	MAC2	REST2v5	REST2v9.1
Equatorial	REST2v9.1	Metstat	KASM
Arid	REST2v9.1	REST2v5	MAC2
Temperate	MAC2	REST2v9.1	Bashahu
Cold	REST2v5	MAC2	REST2v9.1
Polar	CLS	CEM	Atw.& Ball-2

From a “worldwide” standpoint, the MAC2 model is the best overall because it performs excellently and evenly across all climates, despite being one of the simplest models evaluated here. The REST2v5 performs very similarly to the MAC2 with only $r_G = 0.016$ separating the two, and in fact outperforms the MAC2 on many of the metrics. The third placed REST2v9.1 has a much larger separation to second place due to poor performance in the Polar climate. The very harsh conditions of the latter are known to be a source of past experimental errors in radiometers. The ranking of models under the Polar climate is thus considered less certain than under the other climates.

Conflicts of interest

Dr. Gueymard has a potential conflict of interest as his model formulations are included in the evaluation. Conflict has been avoided by him having no possible influence over the results either in ranking score development, calculation of clear-sky irradiance or error metrics calculation, all of which were derived independently by X. Sun and J.M. Bright. All other authors declare no conflict of interest.

Acknowledgements

X. Sun and P. Wang were partially funded by the National Natural Science Foundation of China (Grant No. 11571028). J.M. Bright and N.A. Engerer are funded by the Australian Renewable Energy Agency (ARENA, Research and Development Programme Funding G00854). The authors also thank the BSRN, BoM, SAURAN, KISR, NOAA, and NREL for the high-resolution radiation data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.rser.2019.04.006>.

References

- [1] Bright JM, Sun X. A library of clear-sky irradiance models coded in R. 2018. [online]. <https://jamiembright.github.io/clear-sky-models/>.
- [2] Bright JM, Smith CJ, Taylor PG, Crook R. Stochastic generation of synthetic minutely irradiance time series derived from mean hourly weather observation data. *Sol Energy* 2015;115:229–42.
- [3] Bright JM, Babacan O, Kleissl J, Taylor PG, Crook R. A synthetic, spatially decoupled solar irradiance generator and application to a LV grid model with high PV penetration. *Sol Energy* 2017;147:83–98.
- [4] Ruiz-Arias JA, Gueymard CA. Worldwide inter-comparison of clear-sky solar radiation models: consensus-based review of direct and global irradiance components simulated at the earth surface. *Sol Energy* 2018;168:10–29.
- [5] Yang H, Kurtz B, Nguyen D, Urquhart B, Chow CW, Ghoniya M, Kleissl J. Solar irradiance forecasting using a ground-based sky imager developed at UC San Diego. *Sol Energy* 2014;103:502–24.
- [6] Engerer NA, Bright JM, Killinger S. Himawari-8 enabled real-time distributed PV simulations for distribution networks. Proceedings of the 44th IEEE photovoltaic specialist conference (PVSC44), Washington DC, USA. 25th–30th June. 2017. p. 1–6.
- [7] Bright JM, Killinger S, Lingfors D, Engerer NA. Improved satellite-derived PV power nowcasting using real-time power data from reference PV systems. *Sol Energy* 2018;168:118–39.
- [8] Gueymard CA, Thevenard D. Monthly average clear-sky broadband irradiance database for worldwide solar heat gain and building cooling load calculations. *Sol Energy* 2009;83(11):1998–2018.
- [9] Hosobuchi H, Yoshida H, Uetani Y. Calculation of the heating and cooling load of buildings using a sky radiance distribution model. 9th international IBPSA conference, no. 2 in 1, Montreal, Canada. 2005. p. 427–34.
- [10] Long CN, Ackerman TP. Identification of clear skies from broadband pyranometer measurements and calculation of downwelling shortwave cloud effects. *J Geophys Res* 2000;105(D12):15609. <https://doi.org/10.1029/2000JD900077>.
- [11] Ineichen P. Long term satellite global, beam and diffuse irradiance validation. *Energy Procedia* 2014;48:1586–96.
- [12] Reno MJ, Hansen CW. Identification of periods of clear sky irradiance in time series of GHI measurements. *Renew Energy* 2016;90:520–31.
- [13] Badescu V, Gueymard C, Cheval S, Oprea C, Baci M, Dumitrescu A, Iacobescu F, Milos I, Rada C. Computing global and diffuse solar hourly irradiation on clear sky. Review and testing of 54 models. *Renew Sustain Energy Rev* 2012;16(3):1636–56. <https://doi.org/10.1016/j.rser.2011.12.010>.
- [14] Badescu V, Gueymard CA, Cheval S, Oprea C, Baci M, Dumitrescu A, Iacobescu F, Milos I, Rada C. Accuracy analysis for fifty-four clear-sky solar radiation models using routine hourly global irradiance measurements in Romania. *Renew Energy* 2013;55:85–103.
- [15] Badescu V. Assessing the performance of solar radiation computing models and model selection procedures. *J Atmos Sol Terr Phys* 2013;105:119–34.
- [16] Ineichen P. Comparison of eight clear sky broadband models against 16 independent data banks. *Sol Energy* 2006;80(4):468–78.
- [17] Calinoiu D, Stefu N, Boata R, Blaga R, Pop N, Paulescu E, Sabadus A, Paulescu M. Parametric modeling: a simple and versatile route to solar irradiance. *Energy Convers Manag* 2018;164:175–87.
- [18] Ineichen P. Validation of models that estimate the clear sky global and beam solar irradiance. *Sol Energy* 2016;132:332–44.
- [19] Ineichen P. Long term irradiance clear sky and all-weather model validation. Proceedings of the 4th southern African solar energy conference (SASEC 2016), Stellenbosch, South Africa. 2016. p. 1–8.
- [20] Reno M, Hansen C, Stein J. Global horizontal irradiance clear sky models: implementation and analysis. Tech. rep., Sandia National Laboratories; 2012.
- [21] Gueymard CA. A review of validation methodologies and statistical performance indicators for modeled solar radiation data: towards a better bankability of solar projects. *Renew Sustain Energy Rev* 2014;39:1024–34.
- [22] Seo D, Krarti M. Hourly solar radiation model suitable for worldwide typical weather file generation. *J Sol Energy Eng* 2011;133(4):041002.
- [23] Long C, Gaustad K. The shortwave (SW) clear-sky detection and fitting algorithm: algorithm operational details and explanations. Pacific Northwest National Laboratory; 2004. Rep. ARM TR-004.1.
- [24] Gueymard CA, Myers DR. Validation and ranking methodologies for solar radiation models. In: Badescu V, editor. Modeling solar radiation at the earth's surface. Springer; 2008. p. 479–510.
- [25] Gueymard C. Clear-sky irradiance predictions for solar resource mapping and large-scale applications: improved validation methodology and detailed performance analysis of 18 broadband radiative models. *Sol Energy* 2012;86(8):2145–69. <https://doi.org/10.1016/j.solener.2011.11.011>.
- [26] Engerer N, Mills F. Validating nine clear sky radiation models in Australia. *Sol Energy* 2015;120:9–24.
- [27] Spencer J. Fourier series representation of the position of the sun. *Search* 1971;2(5). 172–172.
- [28] Gueymard CA. A reevaluation of the solar constant based on a 42-year total solar irradiance time series and a reconciliation of spaceborne observations. *Solar Energy* 1 July 2018;168:2–9.
- [29] Iqbal M. An introduction to solar radiation. New York: Academic Press Inc.; 1983.
- [30] Threlkeld J, Jordan R. Direct solar radiation available on clear days. *Heat, Piping Air Cond.* 1957;29(12).
- [31] Schulze R. A physically based method of estimating solar radiation from sun cards. *Agric Meteorol* 1976;16(1):85–101.
- [32] Daneshyar M. Solar radiation statistics for Iran. *Sol Energy*; (United States) 1978;21(4).
- [33] Adnot J, Bourges B, Campana D, Gicquel R. Utilisation de courbes de fréquences cumulées d'irradiation solaire globale pour le calcul des installations solaires. In: Lestienne R, editor. Analyse statistique des processus météorologiques appliquée à l'énergie solaire. Paris: CNRS; 1979.
- [34] Biga A, Rosa R. Contribution to the study of the solar radiation climate of Lisbon. *Sol Energy* 1979;23(1):61–7.
- [35] Handbook AF. American society of heating, refrigerating and air-conditioning engineers. Atlanta, GA, USA: Inc.; 2009.
- [36] Sharma M, Pal R. Interrelationships between total, direct, and diffuse solar radiation in the tropics. *Sol Energy* 1965;9(4):183–92.
- [37] El Mghouchi Y, El Bouardi A, Choulli Z, Ajzoul T. New model to estimate and evaluate the solar radiation. *Int J Sustain Built Environ* 2014;3(2):225–34.
- [38] Yang D, Walsh WM, Jirutitijaroen P. Estimation and applications of clear sky global horizontal irradiance at the equator. *J Sol Energy Eng* 2014;136(3):034505.
- [39] Hottel HC. A simple model for estimating the transmittance of direct solar radiation through clear atmospheres. *Sol Energy* 1976;18(2):129–34.
- [40] Kumar L, Skidmore AK, Knowles E. Modelling topographic variation in solar radiation in a GIS environment. *Int J Geogr Inf Sci* 1997;11(5):475–97.
- [41] Norman J, Campbell G. An introduction to environmental biophysics. New York: Springer; 1998.
- [42] Fu P, Rich PM. Design and implementation of the solar analyst: an ArcView extension for modeling solar radiation at landscape scales. Proceedings of the nineteenth annual ESRI user conference vol. 1. 1999. p. 1–31.
- [43] Atwater M, Ball J. Effects of clouds on insolation models. *Sol Energy* 1981;27(1):37–44.
- [44] Badescu V. Verification of some very simple clear and cloudy sky models to

- evaluate global solar irradiance. *Sol Energy* 1997;61(4):251–64.
- [46] M. Capderou, Theoretical and experimental models solar atlas of Algeria (in French) tome 1 and 2, Algeria: University Publications Office.
- [47] Kasten F. Parametrisierung der Globalstrahlung durch Bedeckungsgrad und Trübungsfaktor. *Annalen der Meteorologie Neue Folge* 1984;20:49–50.
- [48] J. Page, Algorithms for the satel-light programme, technical report for the satel-light programme.
- [49] Rigollier C, Bauer O, Wald L. On the clear sky model of the ESRA - European Solar Radiation Atlas - with respect to the heliostat method. *Sol Energy* 2000;68(1):33–48.
- [50] Lefevre M, Albuisson M, Wald L. Joint report on interpolation scheme 'meteosat' and database 'climatology' i (meteosat), SoDa Deliverable D3-8 and D5-1-4. Internal document. 2002.
- [51] Ineichen P, Perez R. A new air mass independent formulation for the Linke turbidity coefficient. *Sol Energy* 2002;73(3):151–7. [https://doi.org/10.1016/S0038-092X\(02\)00045-2](https://doi.org/10.1016/S0038-092X(02)00045-2).
- [52] Suckling P, Hay JE. Modelling direct, diffuse, and total solar radiation for cloudless days. *Atmosphere* 1976;14(4):298–308.
- [53] King R, Buckius R. Direct solar transmittance for a clear sky. *Sol Energy* 1979;22(3):297–301.
- [54] Badescu V. Use of sunshine number for solar irradiance time series generation. Modeling solar radiation at the earth's surface. Springer; 2008. p. 327–55.
- [55] Ineichen P. A broadband simplified version of the Solis clear sky model. *Sol Energy* 2008;82(8):758–62.
- [56] Ineichen P. High turbidity Solis clear sky model: development and validation. *Rem Sens* 2018;10(3):435.
- [57] Perrin de Brichambaut C. Estimation des ressources énergétiques solaires en France. Association française pour l'étude et le développement des applications de l'énergie solaire en France; 1975.
- [58] Atwater MA, Ball J. A numerical solar radiation model based on standard meteorological observations. *Sol Energy* 1978;21(3):163–70.
- [59] Bird RE, Hulstrom RL. Simplified clear sky model for direct and diffuse insolation on horizontal surfaces. Tech. rep.. Golden, CO (USA): Solar Energy Research Inst.; 1981.
- [60] Carroll J. Global transmissivity and diffuse fraction of solar radiation for clear and cloudy skies as measured and as predicted by bulk transmissivity models. *Sol Energy* 1985;35(2):105–18.
- [61] Gueymard C. Mathematically integrable parameterization of clear-sky beam and global irradiances and its use in daily irradiation applications. *Sol Energy* 1993;50(5):385–97.
- [62] Bashahu M, Laplace D. An atmospheric model for computing solar radiation. *Renew Energy* 1994;4(4):455–8.
- [63] Gueymard CA. Direct solar transmittance and irradiance predictions with broadband models. part i: detailed theoretical performance assessment. *Sol Energy* 2003;74(5):355–79.
- [64] Yang K, Koike T. A general model to estimate hourly and daily solar radiation for hydrological studies. *Water Resour Res* 2005;41(10).
- [65] Hoyt DV. A model for the calculation of solar global insolation. *Sol Energy* 1978;21(1):27–35.
- [66] Davies J, McKay D. Estimating solar irradiance and components. *Sol Energy* 1982;29(1):55–64. [https://doi.org/10.1016/0038-092X\(82\)90280-8](https://doi.org/10.1016/0038-092X(82)90280-8).
- [67] Maxwell E. Metstat: the solar radiation model used in the production of the national solar radiation data base (NSRDB). *Sol Energy* 1998;62(4):263–79.
- [68] Psiloglou B, Santamouris M, Asimakopoulos D. Atmospheric broadband model for computation of solar radiation at the earth's surface. application to Mediterranean climate. *Pure Appl Geophys* 2000;157(5):829–60.
- [69] Paulescu M, Schlett Z. A simplified but accurate spectral solar irradiance model. *Theor Appl Climatol* 2003;75(3–4):203–12.
- [70] Psiloglou BE, Kambezidis HD. Performance of the meteorological radiation model during the solar eclipse of 29 March 2006. *Atmos. Chem. Phys.* 2007;7:6047–59.
- [71] Kambezidis H, Psiloglou B, Karagiannis D, Dumka U, Kaskaoutis D. Meteorological radiation model (MRM v6. 1): improvements in diffuse radiation estimates and a new approach for implementation of cloud products. *Renew Sustain Energy Rev* 2017;74:616–37.
- [72] Janjai S, Sricharoen K, Pattarapanitchai S. Semi-empirical models for the estimation of clear sky solar global and direct normal irradiances in the tropics. *Appl Energy* 2011;88(12):4749–55.
- [73] Gueymard C. 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. <https://doi.org/10.1016/j.solener.2007.04.008>.
- [74] Lefèvre M, Oumbe A, Blanc P, Espinar B, Gschwind B, Qu Z, Wald L, Homscheidt MS, Hoyer-Klick C, Arola A, et al. McClear: a new model estimating downwelling solar radiation at ground level in clear-sky conditions. *Atmos Meas Tech* 2013;6:2403–18.
- [75] Zhong X, Kleissl J. Clear sky irradiances using REST2 and MODIS. *Sol Energy* 2015;116:144–64.
- [76] Zhang Y, Li X, Bai Y. An integrated approach to estimate shortwave solar radiation on clear-sky days in rugged terrain using modis atmospheric products. *Sol Energy* 2015;113:347–57.
- [77] Ruiz-Arias JA, Gueymard CA. A multi-model benchmarking of direct and global clear-sky solar irradiance predictions at arid sites using a reference physical radiative transfer model. *Sol Energy* 2018;171:447–65.
- [78] Gelaro R, McCarty W, Suárez MJ, Toddling R, Molod A, Takacs L, Randles CA, Darmenov A, Bosilovich MG, Reichle R, et al. Modern-Era Retrospective Analysis for Research and Applications, Version2 (MERRA-2). *J Clim* 2017;30(14):5419–54. <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- [79] Lefèvre M, Wald L. Validation of the McClear clear-sky model in desert conditions with three stations in Israel. *Adv Sci Res* 2016;13:21–6.
- [80] Mikkavilli S, Prasad A, Taylor R, Huang J, Mitchell R, Troccoli A, Kay M. Assessment of atmospheric aerosols from two reanalysis products over Australia. *Atmos Res* 2019;215:149–64.
- [81] Shi H, Xiao Z, Zhan X, Ma H, Tian X. Evaluation of MODIS and two reanalysis aerosol optical depth products over aeronet sites. *Atmos Res* 2019;220:75–80.
- [82] Holben BN, Eck TF, Slutsker I, Tanre D, Buis J, Setzer A, Vermote E, Reagan JA, Kaufman YJ, Nakajima T. AERONET-A federated instrument network and data archive for aerosol characterization. *Rem Sens Environ* 1998;66(1):1–16.
- [83] Gueymard CA. Impact of on-site atmospheric water vapor estimation methods on the accuracy of local solar irradiance predictions. *Sol Energy* 2014;101:74–82.
- [84] Gueymard CA. Clear-sky radiation models and aerosol effects. In: Polo J, Martin-Pomares L, Sanfilippo A, editors. *Solar resources mapping*. Springer; 2019. p. 137–82.
- [85] Randles C, Da Silva A, Buchard V, Colarco P, Darmenov A, Govindaraju R, Smirnov A, Holben B, Ferrare R, Hair J, et al. The MERRA-2 aerosol reanalysis, 1980 onward. part I: system description and data assimilation evaluation. *J Clim* 2017;30(17):6823–50.
- [86] Hayasaka T, Satake S, Shimizu A, Sugimoto N, Matsui I, Aoki K, Muraji Y. Vertical distribution and optical properties of aerosols observed over Japan during the atmospheric brown clouds-East Asia regional experiment 2005. *J Geophys Res: Atmosphere* 2007;112(D22).
- [87] GMAO, GES-DISC. MERRA-2 const_2d_asm_nx: 2d, constants v5.12.4, greenbelt. 2018. [online]. <https://disc.gsfc.nasa.gov/datasets/>.
- [88] Krotkov N. Omi/Aura NO2 cloud-screened total and tropospheric column l3 global gridded 0.25 degree x 0.25 degree v3, NASA Goddard Space Flight Center. Goddard Earth Sciences Data and Information Services Center (GES DISC); 2013.
- [89] Bright JM, repository Github. Dynamic variable tool [online]. 2018 https://github.com/JamieMBright/dynamic_variable_tool_development.
- [90] Remund J, Wald L, Lefèvre M, Ranchin T, Page J. Worldwide Linke turbidity information. ISES Solar World Congress 2003 vol. 400. International Solar Energy Society (ISES); 2003. p. 13.
- [91] Dogniaux R. Représentations analytiques des composantes du rayonnement lumineux solaire: conditions de ciel serein. Institut royal météorologique de Belgique; 1974.
- [92] Ineichen P. Conversion function between the Linke turbidity and the atmospheric water vapor and aerosol content. *Sol Energy* 2008;82(11):1095–7.
- [93] Gueymard CA. Turbidity determination from broadband irradiance measurements: a detailed multicoefficient approach. *J Appl Meteorol* 1998;37(4):414–35.
- [94] Molineaux B, Ineichen P, Delaunay J-J. Direct luminous efficacy and atmospheric turbidity - improving model performance. *Sol Energy* 1995;55(2):125–37.
- [95] Grenier J, de La Casinière A, Cabot T. A spectral model of Linke's turbidity factor and its experimental implications. *Sol Energy* 1994;52(4):303–13.
- [96] Kottek M, Grieser J, Beck C, Rudolf B, Rubel F. World map of the Köppen-Geiger climate classification updated. *Meteorol Z* 2006;15(3):259–63. <https://doi.org/10.1127/0941-2948/2006/0130>.
- [97] Norton M, Amillo AM, Galleano R. Comparison of solar spectral irradiance measurements using the average photon energy parameter. *Sol Energy* 2015;120:337–44. <https://doi.org/10.1016/j.solener.2015.06.023>.
- [98] Nonnenmacher L, Kaur A, Coimbra CF. Verification of the SUNY direct normal irradiance model with ground measurements. *Sol Energy* 2014;99:246–58. <https://doi.org/10.1016/j.solener.2013.11.010>.
- [99] Michalsky J, Dutton E, Rubes M, Nelson D, Stoffel T, Wesley M, Splitt M, DeLuise J. Optimal measurement of surface shortwave irradiance using current instrumentation. *J Atmos Ocean Technol* 1999;16(1):55–69.
- [100] Gueymard CA, Myers DR. Evaluation of conventional and high-performance routine solar radiation measurements for improved solar resource, climatological trends, and radiative modeling. *Sol Energy* 2009;83(2):171–85.
- [101] WRMC-BSRN. Data retrieval via pangaea. 2018. [online]. <http://bsrn.awi.de/data/data-retrieval-via-pangaea/>.
- [102] Al-Rasheedi M, Gueymard C, Ismail A, Hajraf S. Solar resource assessment over Kuwait: validation of satellite-derived data and reanalysis modeling. *Proceedings of EuroSun ISES conference*. France: Aix-les-Bains; 2014. –.
- [103] Gueymard CA, Al-Rasheedi M, Ismail A, Hussain T. Long-term variability of aerosol optical depth, dust episodes, and direct normal irradiance over Kuwait for CSP applications. ISES solar world congress 2017. IEA SHC international conference on solar heating and cooling for buildings and industry. 2017. <https://doi.org/10.18086/swc.2017.04.04>.
- [104] Gueymard C. Validation of real-time solar irradiance simulations over Kuwait using WRF-Solar, EuroSun2018 12th international conference on solar energy for buildings and industry. Switzerland: Rapperswil; 2018.
- [105] Al-Rasheedi M, Gueymard CA, Ismail A, Hussain T. Comparison of two sensor technologies for solar irradiance measurement in a desert environment. *Sol Energy* 2018;161:194–206.
- [106] Long C, Shi Y. The QCRAD value added product: surface radiation measurement quality control testing, including climatology configurable limits, Atmospheric Radiation Measurement Program Technical Report. 2006.
- [107] Peel MC, Finlayson BL, McMahon TA. Updated world map of the Köppen-Geiger climate classification. *Hydrol Earth Syst Sci* 2007;11(5):1633–44. <https://doi.org/10.5194/hess-11-1633-2007>.
- [108] Gueymard CA, Bright JM, Lingfords D, Habte A, Sengupta M. A posteriori clear-sky identification methods in solar irradiance time series: review and preliminary validation using sky imagers. *Renew Sustain Energy Rev* 2019;109:412–27.
- [109] Eissa Y, Beegum SN, Gherboudj I, Chaouch N, Al Sudairi J, Jones RK, Al Dobayan

- N, Ghedira H. Prediction of the day-ahead clear-sky downwelling surface solar irradiances using the REST2 model and WRF-CHIMERE simulations over the Arabian Peninsula. *Sol Energy* 2018;162:36–44.
- [110] Taylor KE. Summarizing multiple aspects of model performance in a single diagram. *J Geophys Res: Atmosphere* 2001;106(D7):7183–92.
- [111] Yang D, Dong Z, Lim LHI, Liu L. Analyzing big time series data in solar engineering using features and PCA. *Sol Energy* 2017;153:317–28.
- [112] Davò F, Alessandrini S, Sperati S, Delle Monache L, Airolidi D, Vespucci MT. Post-processing techniques and principal component analysis for regional wind power and solar irradiance forecasting. *Sol Energy* 2016;134:327–38.
- [113] Kaiser HF. The application of electronic computers to factor analysis. *Educ Psychol Meas* 1960;20(1):141–51.
- [114] Killinger S, Lingfors D, Saint-Drenan Y-M, Moraitis P, van Sark W, Taylor J, Engerer NA, Bright JM. On the search for representative characteristics of pv systems: data collection and analysis of PV system azimuth, tilt, capacity, yield and shading. *Sol Energy* 2018;173:1087–106.
- [115] Gueymard C, Myers D. Solar radiation measurement: progress in radiometry for improved modeling. In: Badescu V, editor. *Modeling solar radiation at the earth surface*. Springer; 2008.
- [116] Habte A, Sengupta M, Andreas A, Wilcox S, Stoffel T. Intercomparison of 51 radiometers for determining global horizontal irradiance and direct normal irradiance measurements. *Sol Energy* 2013;133:372–93.
- [117] Sengupta M, Habte A, Gueymard C, Wilbert S, Renne D, Stoffel T. Best practices handbook for the collection and use of solar resource data for solar energy applications. second ed. National Renewable Energy Laboratory; 2017. Rep. NREL/TP-5D00-68886 <https://www.nrel.gov/docs/fy18osti/68886.pdf>.
- [118] Matsui N, Long CN, Augustine J, Halliwell D, Uttal T, Longenecker D, Niebergall O, Wendell J, Albee R. Evaluation of arctic broadband surface radiation measurements. *Atmos Meas Tech* 2012;5(2):429–38. <https://doi.org/10.5194/amt-5-429-2012> <https://www.atmos-meas-tech.net/5/429/2012/>.
- [119] Earth System Research Laboratory - Physical Sciences Division, D-ICE. De-icing comparison experiment. <https://www.esrl.noaa.gov/psd/arctic/d-ice/>; 2018.
- [120] Davies J, McKay D. Evaluation of selected models for estimating solar radiation on horizontal surfaces. *Sol Energy* 1989;43(3):153–68. [https://doi.org/10.1016/0038-092X\(89\)90027-3](https://doi.org/10.1016/0038-092X(89)90027-3).
- [121] Gueymard C. Temporal variability in direct and global irradiance at various time scales as affected by aerosols. *Sol Energy* 2012;1–10. <https://doi.org/10.1016/j.solener.2012.01.013>.
- [122] Gueymard C, Ruiz-Arias J. Validation of direct normal irradiance predictions under arid conditions : a review of radiative models and their turbidity-dependent performance. *Renew Sustain Energy Rev* 2015 2015:379–96.
- [123] M. Sengupta, A. Weekley, A. Habte, A. Lopez, C. Molling, Validation of the National Solar Radiation Database (NSRDB) (2005–2012). 31st European Photovoltaic Solar Energy Conference and Exhibition, Hamburg, Germany.
- [124] Sengupta M, Habte A, Gotseff P, Weekley A, Lopez A, Molling C, Heidinger A. A physics-based GOES product for use in NREL's National Solar Radiation Database. European Photovoltaic Solar Energy Conference and Exhibition, Amsterdam, Netherlands. National Renewable Energy Laboratory; 2014.
- [125] Antonanzas-Torres F, Antonanzas J, Urraca R, Alia-Martinez M, Martinez-de Pison F. Impact of atmospheric components on solar clear-sky models at different elevation: case study Canary Islands. *Energy Convers Manag* 2016;109:122–9.
- [126] M. J. Reno, C. W. Hansen, J. S. Stein, Global horizontal irradiance clear sky models: implementation and analysis, SANDIA report SAND2012-2389; 2012.
- [127] Gueymard CA. Clear-sky irradiance predictions for solar resource mapping and large-scale applications: improved validation methodology and detailed performance analysis of 18 broadband radiative models. *Sol Energy* 2012;86(8):2145–69.
- [128] Copper J, Sproul A. Comparative study of mathematical models in estimating solar irradiance for Australia. *Renew Energy* 2012;43:130–9. <https://doi.org/10.1016/j.renene.2011.11.050>.
- [129] Cros S, Liandrat O, Sébastien N, Schmutz N, Voyant C. Clear sky models assessment for an operational pv production forecasting solution. 28th European Photovoltaic Solar Energy Conference and Exhibition. 2013. 5BV–4.
- [130] Zhandire E. Predicting clear-sky global horizontal irradiance at eight locations in South Africa using four models. *J Energy South Afr* 2017;28(4):77–86.
- [131] Mikofski MM, Hansen CW, Holmgren WF, Kimball GM. Use of measured aerosol optical depth and precipitable water to model clear sky irradiance. Proc. 44th IEEE PV specialists conf., Washington, DC. IEEE; 2017.
- [132] El Mghouchi Y, Ajzoul T, Taoukil D, El Bouardi A. The most suitable prediction model of the solar intensity, on horizontal plane, at various weather conditions in a specified location in Morocco. *Renew Sustain Energy Rev* 2016;54:84–98.
- [133] Badescu V, Gueymard CA, Cheval S, Oprea C, Baciu M, Dumitrescu A, Iacobescu F, Milos I, Rada C. Computing global and diffuse solar hourly irradiation on clear sky. review and testing of 54 models. *Renew Sustain Energy Rev* 2012;16(3):1636–56.
- [134] Barbieri F, Riffart C, Vo B-T, Rajakaruna S, Ghosh A. A comparative study of clear-sky irradiance models for western Australia. Power and energy society general meeting (PESGM), 2016. IEEE; 2016. p. 1–5.