Renewable and Sustainable Energy Reviews 135 (2021) 110087

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

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

Worldwide performance assessment of 95 direct and diffuse clear-sky irradiance models using principal component analysis

Xixi Sun [a], Jamie M. Bright [b,*], Christian A. Gueymard [c], Xinyu Bai [d], Brendan Acord [e], Peng Wang [f,a,g,**]

[a] LMIB & School of Mathematical Sciences, Beihang University, Beijing, China
[b] Solar Energy Research Institute of Singapore (SERIS), National University of Singapore, Singapore, 117574, Singapore
[c] Solar Consulting Services, Colebrook, NH, USA
[d] School of Computer Science and Engineering, Beihang University, Beijing, China
[e] Syene Clean Energy, Hong Kong, China
[f] School of Microelectronics, Beihang University, Beijing, China
[g] Beijing Advanced Innovation Center for Big Data and Brain Computing, Beihang University, Beijing, China

ARTICLE INFO

Keywords:
Clear-sky irradiance
Clear-sky model
Validation
Principal component analysis
Direct normal irradiance
Diffuse horizontal irradiance
Köppen-Geiger

ABSTRACT

Accurate estimations of clear-sky direct normal irradiance (DNIcs) and diffuse horizontal irradiance (DIFcs) are crucial in solar resources assessment. This study examines 95 and 88 popular clear-sky irradiance models for their worldwide estimation capability of DNIcs and DIFcs, respectively. Atmospheric inputs are from MERRA-2 reanalysis and irradiance observations for validation are extracted from 100 ground stations across five major Köppen-Geiger climate zones. They consist of 24 million 1-min measurements of DINcs and 18.7 million 1-min measurements of DIFcs after quality control and clear-sky detection, during the 5-year period 2015–2019. Using principal component analysis, the performance of each clear-sky irradiance model is ranked separately among the five climate zones, as well as given a global rank. For the Equatorial, Arid, Temperate, Cold and Polar climates, it is found that Heliosat1-I, ESRA-I, REST2v9.1, REST2v9.1 and CLS, respectively, are the best DNIcs models, while Modified Iqbal-C, Heliosat1-R, Calinoiu, Modified Iqbal-C and PSI-REST are the best DIFcs models. On a global worldwide basis, the three top-ranking models are REST2v9.1, REST2v5 and MMAC-V2 for DNIcs, and Modified Iqbal-C, PSI-REST and MRMv5 for DIFcs. The results and rankings presented are strictly relative to MERRA-2 input data, and should not be extrapolated to results from alternate sources of atmospheric data. This detailed validation exercise revealed inconsistent performance of many models across different climates, possibly due to insufficient training and/or over-fitting of empirical relationships. Codes of all clear-sky irradiance models in the public domain are available online on the Github repository JamieMBright/clear-sky-models.

1. Introduction

Accurate estimations of direct normal clear-sky irradiance (DNIcs) and diffuse horizontal clear-sky irradiance (DIFcs) are essential in a wide range of solar energy applications. Together, they form the clear-sky global horizontal irradiance (GHIcs) and provide a basis for irradiance evaluations under all-sky conditions.

Since high-quality observations of irradiance are generally sparse worldwide [1], a number of competing clear-sky irradiance models have been developed over the years to estimate DNIcs, DIFcs and GHIcs. Precise knowledge of DNIcs is particularly important for solar applications based on focusing devices, such as concentrating solar power or concentrating photovoltaics. In contrast, diffuse irradiance is a required quantity to evaluate the cooling loads of buildings or the potential of their different facades for daylighting. Moreover, the separate determinations of DNIcs, DIFcs and GHIcs are actually interrelated because of the closure equation that links them physically. While GHIcs is a very important benchmark in many solar applications, as well as an input to various intermediate calculations, such as diffuse fraction models (see a review in Ref. [2]), it is important to emphasise that most clear-sky irradiance models evaluate DNIcs and DIFcs independently in a first step, and then obtain GHIcs from their combination. From that perspective, the separate validation of GHIcs and its two components

* Corresponding author.
** Corresponding author. School of Microelectronics, Beihang University, Beijing, China.
E-mail addresses: jamie.bright@nus.edu.sg, jamiebright1@gmail.com (J.M. Bright), wang.peng@buaa.edu.cn, wang.peng@outlook.com (P. Wang).

https://doi.org/10.1016/j.rser.2020.110087
Received 13 March 2020; Received in revised form 9 July 2020; Accepted 10 July 2020
Available online 14 August 2020
1364-0321/

Nomenculture

Acronyms

DIFcs	Diffuse horizontal clear-sky irradiance
DNIcs	Direct normal clear-sky irradiance
GHIcs	Global horizontal clear-sky irradiance
DIF	Diffuse horizontal irradiance
DNI	Direct normal irradiance
GHI	Global horizontal irradiance
PCA	Principal component analysis
AOD	Aerosol optical depth
BSRN	Baseline Surface Radiation Network
WMO	World Meteorological Organization
BoM	Bureau of Meteorology, Australia
SAURAN	The South African Universities Radiation Network
NREL	National Renewable Energy Laboratory, USA-
NOAA	National Oceanic and Atmospheric Administration, USA-
NASA	The National Aeronautics and Space Administration, USA-
MERRA-2	Modern-Era Retrospective Analysis for Research and Applications, version 2 by NASA
KISR	Kuwait Institute of Scientific Research
RRTMG	Rapid Radiative Transfer Model
QC	Quality control
CSD	Clear-sky detection algorithm
KG	Köppen-Geiger climate classification
P50	The 50th percentile-
G/A/B/C/D/E	Referring to KG climate: Global/Equatorial/Arid/Temperate/Cold/Polar
PCHIP	Piecewise Cubic Hermite Interpolating Polynomial

Mathematical symbols

E_{bnc}	Beam normal clear-sky irradianceW m^{-2}
E_{dhc}	Diffuse horizontal clear-sky irradianceW m^{-2}
E_{ghc}	Global horizontal clear-sky irradianceW m^{-2}
E_{sc}	Solar constantW m^{-2}
E_{ext}	Extraterrestrial irradianceW m^{-2}
θ_z	Solar zenith angle deg
h	Elevation above sea level m
p	Surface atmospheric pressuremb
T	Ambient temperatureK
R_G	Surface albedo
T_L	Linke Turbidity
u_{O_3}	Total column ozone atm-cm
u_{NO_2}	Total column nitrogen dioxide atm-cm
u_{H_2O}	Total column precipitable water vapour atm-cm
τ	Aerosol optical depth at 550 m, 700 m, or broadband-
α	Ångström exponent
β	Ångström turbidity coefficient
ϖ	Aerosol single-scattering albedo
R	Mean sun-earth distance relative to astronomical unit
M_m	Error metric, where m of 1 is MBD etc.matches error
c_i	The PCA coefficient weighting for the i^{th} principal component
e_i	The explained variance of the i^{th} principal component %
r_x	The ranking score from PCA where x refers to a KG climate

Error metrics

MBD	Mean bias difference %
RMSD	Root mean square difference W m^{-2}
MAD	Mean absolute difference W m^{-2}
SBF	Slope of best-fit line-
U_{95}	Uncertainty at 95%
TS	t-statistic
R_2	Coefficient of determination
WIA	Willmott's index of agreement
LCE	Legate's coefficient of efficiency
KSI	Kolmogoov-Smirnov test integral
OVER	Critical limit of overestimation
CPI	Combined performance index

(DNIcs and DIFcs) appears an important objective and has attracted considerable interest in the literature (for a review, see Ref. [3]). In that framework, the present contribution can be portrayed as an extension of a previous study [4] that was limited to GHIcs. In that paper, Sun et al. [4] conducted a rigorous analysis and exhaustive validation of 75 popular clear-sky irradiance models for GHIcs, using observations from 75 high-quality stations belonging to various world networks, across five major Köppen-Geiger climate zones. Regarding DNIcs, various validation or benchmarking studies have appeared over the years [5–14]. Most of these studies only tested a small number of clear-sky irradiance models at a limited number of sites. In the latest known exercise of such kind, Antonanzas-Torres et al. [15] examined a larger number (70) of clear-sky models but only at two sites and for one year. Given the current worldwide interest in solar applications and the demand for accurate irradiance modelling at any global location, validation exercises conducted at only a few sites are clearly insufficient. The case of DIFcs is somewhat different because of the relatively lower number of validation studies [9,15–17]. This makes a more thorough validation of the many existing DIFcs models all the more pressing.

Two important limitations in the process of validating clear-sky irradiance models have been noted in the literature (e.g. Ref. [3,6]): (i) clear-sky periods need to be determined after the fact through careful examination of the measurement time series, which, while constituting a difficulty and a source of error due to cloud contamination, can now be considered satisfactorily resolved [18,19]; and (ii) appropriate and reliable atmospheric data inputs are needed to operate these models at any location of interest, but data sources of worldwide range are rare and may be of low resolution or low accuracy [20]. The latter issue, most importantly, has not always been addressed with sufficient care in the literature. A few studies (e.g. Refs. [9,11,14,15]) have used sunphotometric observations from NASA's AERONET network [21] to validate clear-sky irradiance models. This method assumes that an AERONET station is co-located with a radiometric station where high-quality irradiance components are measured. This is a rare occurrence, which prevents this method from being truly universal in engineering applications. Another approach is to use atmospheric data produced by a reanalysis, as introduced in recent contributions [3,14,22,23]. This solution offers great potential because reanalyses provide long data time series with no data gaps and global coverage. Recently, the accuracy of atmospheric aerosol data from two recent reanalyses has been validated against ground-based observations from AERONET [24]. Based on those results, NASA's Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) reanalysis [25] is selected here to provide the necessary atmospheric inputs to all clear-sky irradiance models under consideration. The outputs from a record number of DNIcs and DIFcs models are then validated against the 1-min ground measurements at 100 stations worldwide. To be specific, 88 clear-sky irradiance models that evaluate both DNIcs and DIFcs components, as well as 7 DNIcs-only models, are put to test here. To that end, an innovative ranking method based on principal component analysis (PCA) [4] is employed to rank the DNIcs and DIFcs models for each climate as well as globally.

The remainder of this paper is organised as follows. In Section 2, the clear-sky irradiance models are briefly described, as well as their inputs

and the ground irradiance measurements. Section 3 introduces the performance assessment methodology. This is followed by results and discussion of DNIcs and DIFcs models in Section 4. An update on the performance of GHIcs models and a combined GHIcs model is briefly discussed in Section 5. The future of clear-sky irradiance modelling is discussed in Section 6. The final conclusions are summarised in Section 7. For brevity of the main body of this paper, details on model descriptions and a complete spreadsheet of results are provided in the supplementary material. Finally, to encourage truly comparable studies in the future, the codes of the aforementioned models are made freely available through GitHub [26], at https://jamiembright.github.io/clear-sky-models/. Additionally, the models and methods of downloading and extracting MERRA-2 reanalysis are fully detailed and freely available in the `irradpy` Python package [27].

2. Clear-sky irradiance models, atmospheric data and validation sites

This section presents a general and brief description of the clear-sky irradiance models used in the subsequent analysis, along with their atmospheric input data. The 100 ground radiometric stations serving as validation sites are subsequently introduced.

2.1. Clear-sky irradiance models

A near-exhaustive review of the literature has identified a total of 95 clear-sky irradiance models for consideration in this paper. From this total, 88 are complete models that provide estimations of both the direct and diffuse components, and 7 are DNIcs-only models. Although, in principle, one can alternatively derive DNIcs and DIFcs from GHIcs using empirical estimates of the diffuse fraction K [2,28], various studies [29,30] suggest that it is actually more reliable to instead use a complete clear-sky irradiance model along with appropriate atmospheric information. A summary of all models and their original literature are listed in Table 1. The models are normally used as described by their original authors. A few models have been modified or combined to make them compatible with the present sources of data or to improve their universal validity, as described previously [4]. In particular, what was then referred to as the PR model is now renamed PSI-REST for clarity. It combines the non-aerosol transmittance functions described in Ref. [31] with the aerosol transmittance functions of the REST model [32].

The model input variables include: solar constant E_{sc} [W m^{-2}], zenith angle θ_z, site elevation above sea level h [m], local barometric pressure p [mb], ambient temperature T [K], surface albedo R_G, Linke turbidity factor (T_L), 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], aerosol optical depth at 550 nm, 700 nm, or broadband (τ), Ångström exponent α, Ångström turbidity coefficient β, and aerosol single-scattering albedo ϖ. All these variables are relevant to both DNIcs and DIFcs modelling, except R_G and ϖ, which are of interest only for the determination of DIFcs.

To facilitate the subsequent analyses, all models are categorised into five classes (with subscripts *a*, *b*, *c*, *d* and *e*) depending on the type of aerosol-related input variables they require.

- Class a: models without atmospheric inputs.
- Class b: models without aerosol inputs.
- Class c: models with input of Linke Turbidity (T_L).
- Class d: models with inputs of aerosol optical depth (AOD) and Ångström exponent.
- Class e: models with inputs of AOD, Ångström exponent, and ground albedo.

For detailed information on all models, as well as the present interpretations and formulations, the reader can refer to a previous study [4] or to the supplementary material. Note that all the clear-sky irradiance models in Table 1—except the web-based McClear model and the proprietary REST2v9.1 model—are coded in R and Matlab, and can be accessed from a dedicated repository on GitHub [26]. A release of Python code is also anticipated, as well as hosting it on the `irradpy` Python package [27].

In the present study, DNIcs, DIFcs and GHIcs are denoted with the symbols E_{bnc}, E_{dhc} and E_{ghc}, respectively. The extraterrestrial irradiance (E_{ext}) at the time of measurement is calculated as:

$$E_{ext} = E_{sc}/R^2, \quad (1)$$

where E_{sc} refers to the solar constant that is normally specified by each model under test. In case that value is not specified, a default value is assumed, E_{sc} = 1361.1 W m^{-2}, based on the most recent data [33]. R is the sun-earth distance relative to the astronomical unit [AU]. Its value, as well as the solar zenith angle (θ_z), are calculated via the SG2 sun position algorithm [34], which is selected here for its efficiency and accuracy. Longitude, latitude and elevation above sea level for each testing sites are obtained from their ground station metadata.

Two aerosol-related input variables of note — Linke turbidity (T_L) and broadband AOD $\tau_{broadband}$ — can both be derived from more fundamental aerosol data using a variety of empirical functions that have been proposed in the literature, which might introduce confusion. This is because imprecise input data can yield wildly different error propagation effects on the end results, though this strongly depends on individual model performance. It is important to evaluate the combined accuracy of two chained models: one that provides the input in proper form (T_L or $\tau_{broadband}$), and one that uses that input and evaluates the clear-sky irradiance. The five clear-sky irradiance models of class *c* with Linke turbidity as input, namely Heliosat-1, ESRA,Heliosat-2, Ineichen-Perez and Dogniaux, are thus paired with each of six existing formulations providing T_L, namely, T_{L2R} [35], T_{L2D} [36], T_{L2I} [37], T_{L2Gu} [38], T_{L2M} [39], and T_{L2Gr} [40], eventually leading to a total of 30 model combinations. In parallel, the two clear-sky irradiance models, MMAC and METSTAT, that use broadband AOD as input are paired with three existing formulations providing $\tau_{broadband}$, namely *V*1 [9], *V*2 [9] and *V*3 [38]. In what follows, the aforementioned combinations are labeled as "model name - formulation name". Details on the exact calculations of T_L and $\tau_{broadband}$ can be found in the supplementary material.

2.2. Atmospheric data

As mentioned previously, the atmospheric inputs required by the clear-sky irradiance models are taken from the gridded reanalysis dataset MERRA-2 [25]. MERRA-2 provides high-quality global coverage on a grid of 0.5° × 0.625° with 1-h temporal resolution from 1980 to the preceding 2 months. Whilst there is no "nearly-perfect" input dataset [92] and alternative global datasets for obtaining irradiance data do exist [93], such as the Copernicus Atmosphere Monitoring Service (CAMS) from the European Center for Medium-range Weather Forecasts (ECMWF), earlier studies [24,94,95] found that MERRA-2 had good balance between data availability and its spatio-temporal resolution and extent, and thus reduces the impact of data uncertainty onto solar irradiance estimations. Nevertheless, a recent validation of the MERRA-2 aerosol data revealed that they could be affected by significant bias and dispersion, depending on climate region or continent [24]. This bias becomes all the more important for DNIcs and DIFcs, as they are substantially more sensitive to aerosols than GHIcs [96]. Hence, by using the MERRA-2 reanalysis data as inputs, one must accept that error propagation is likely to occur. This means that the performance assessment herein is *strictly* tied to the MERRA-2 input data, and that different results or model rankings could well be obtained with different input data. However, this does not detract from the impact of this research because MERRA-2 is widely used in many solar energy studies and can be easily accessed and customised using the `irradpy` Python package [27].

Table 1
Summary of clear-sky DNIcs and DIFcs models along with their input variables: specified solar constant (E_{sc}), zenith angle θ_z, site elevation above sea level h [m], local barometric pressure p [mb], ambient temperature T [K], ground albedo R_G, Linke turbidity factor (T_L), 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], aerosol optical depth at 550 nm, 700 nm or broadband (τ), Ångström exponent α, Ångström turbidity coefficient β, and aerosol single-scattering albedo ϖ. The model number is assigned in descending order by the model class (a, b, c, d, e), number of input variables, publication year and alphabet. Seven models names are clustered with multiple model numbers, each corresponding to the various formulations of T_L or broadband-AOD τ. Models in boldfacae are those with only DNIcs output.

#	Class	Clear-sky irradiance model	Citation	E_{sc}	θ_z	h	p	T	R_G	T_L	u_{O_3}	u_{NO_2}	u_{H_2O}	τ	α	β	ϖ	Total
1	a	TJ	[41]		·													1
2	a	HS	[42]		·													1
3	a	Paltridge	[43]		·													1
4	a	Schulze	[44]		·													1
5	a	DPP	[45]		·													1
6	a	ABCGS	[42]		·													1
7	a	Biga	[46]		·													1
8	a	ASHRAE1985	[47]		·													1
9	a	ASHRAE2005	[48]		·													1
10	a	El Mghouchi	[49]	·	·													2
11	a	Sharma	[50]	·	·													2
12	a	**Meinel**	[51]	·	·	·												3
13	a	HLJ	[52]	·	·	·												3
14	a	**Samimi**	[53]	·	·	·												3
15	a	Fu & Rich	[54]	·	·	·												3
16	b	BCLSM	[55]	·	·								·					3
17	b	Kumar	[56]	·	·		·											3
18	b	Campbell	[57]	·	·		·											3
19	b	**Majumdar**	[58]	·	·		·						·					4
20	b	KASM	[59]	·	·		·						·					4
21	b	Capderou	[60]	·	·	·	·											4
22	b	**Allen**	[61]	·	·		·						·					4
23	b	MAC1	[62]	·	·		·		·				·					5
24	b	Josefsson	[42]	·	·		·		·				·					5
25	b	Badescu	[63]	·	·		·				·		·					5
26	b	CLS	[64]	·	·		·		·				·					5
27–32	c	Heliosat1	[65]	·	·		·			·								4
33–38	c	ESRA	[66]	·	·	·				·								4
39–44	c	Heliosat2	[67]	·	·	·				·								4
45–50	c	Ineichen-Perez	[68]	·	·	·				·								4
51–56	c	Dongiaux	[69]	·	·		·			·								4
57	d	Ineichen2008	[70]	·	·		·						·	·				5
58	d	Ineichen2018	[71]	·	·		·						·	·				5
59	d	Perrin	[72]	·	·		·				·		·			·		6
60	d	Iderіah	[73]	·	·		·						·		·	·		6
61	d	Bashahu	[74]	·	·		·						·		·	·		6
62	d	Yang	[75]	·	·		·				·		·			·		6
63	d	Dai	[76]	·	·		·						·		·	·		6
64	d	Calinoiu	[77]	·	·						·	·	·			·		6
65	d	**MLWT1**	[38]	·	·		·				·	·	·			·		7
66	d	**MLWT2**	[32]	·	·		·				·	·	·			·		7
67	d	Paulescu & Schlett	[78]	·	·	·	·				·		·			·		7
68	d	**REST**	[32]	·	·		·				·	·	·			·		7
69	d	Janjai	[79]	·	·	·					·		·		·	·		7
70	e	CEM	[80]	·	·		·		·				·	·				6
71	e	King	[81]	·	·		·		·				·			·		6
72	e	Atwater & Ball-2	[82]	·	·		·		·				·	·				6
73	e	RSC	[83]	·	·		·		·				·			·		6
74	e	PSIM	[84]	·	·		·		·				·			·		6
75–77	e	MMAC	[32]	·	·		·		·				·	·				6
78	e	MAC2	[62]	·	·		·		·				·		·	·		7
79	e	Iqbal-B	[85]	·	·		·		·		·		·			·		7
80–82	e	METSTAT	[86]	·	·		·		·		·		·	·				7
83	e	PSI-REST	[31,32]	·	·		·		·		·		·			·		7
84	e	MRMv5	[87]	·	·		·		·		·		·			·		7
85	e	CSR	[9]	·	·		·		·		·		·	·				7

(*continued on next page*)

Table 1 (*continued*)

#	Class	Clear-sky irradiance model	Citation	E_{sc}	θ_z	h	p	T	R_G	T_L	u_{O_3}	u_{NO_2}	u_{H_2O}	τ	α	β	ϖ	Total
86	e	MRMv6.1	[88]	·	·		·		·		·		·	·				7
87	e	Watt	[89]	·	·		·		·		·		·		·	·		8
88	e	Bird	[82]	·	·		·		·		·		·		·	·		8
89	e	Iqbal-A	[85]	·	·		·		·		·		·		·	·		8
90	e	Iqbal-C	[85]	·	·		·		·		·		·		·	·		8
91	e	CPCR2	[90]	·	·		·		·		·		·		·	·	·	9
92	e	Modified Iqbal-C	[9]	·	·	·			·		·		·		·	·		8
93	e	REST2v5	[91]	·	·		·		·		·	·	·		·	·	·	10
94	e	REST2v9.1		·	·		·		·		·	·	·	·	·		·	10
95	e	McClear	[22]	·	·	·	·	·	·		·		·	·	·			10

Table 2 summarises the input data from the MERRA-2 reanalysis database. Note that McClear (model of class *e*) is the only model that needs temperature T as an input. However, this model has a unique status here because its code is not user-accessible; its irradiance predictions can be obtained publicly, using an API web-service (http://www.soda-pro.com/web-services/radiation/cams-mcclear) that only requires spatio-temporal information to generate irradiance time series. Thus, temperature T is actually not required by any user-operable model in this study. Due to the patchy and discontinuous measurements of the nitrogen dioxide columnar abundance worldwide, and its relatively low impact on irradiance outside of heavily polluted areas, u_{NO_2} is set to a default value $u_{NO_2} = 0.0002$ [91]. The AOD at 700 nm, τ_{700}, and the Ångström turbidity coefficient, β, are derived from MERRA-2's τ_{550} and α following Ångström's law:

$$\tau_{700} = \tau_{550}\left(\frac{700}{550}\right)^{-\alpha}, \quad \beta = \frac{\tau_{550}}{0.550^{-\alpha}}. \tag{2}$$

The aerosol single-scattering albedo, ϖ, which is required by a few models of class *e*, is obtained as the ratio between the scattering aerosol optical depth and the total aerosol optical depth at 550 nm, i.e., TOTSCATAU/TOTEXTTAU in Table 2. The reader can find extensive details on MERRA-2 variables in Ref. [97].

Additionally, the reader is advised to perform altitude corrections to the MERRA-2 data to account for the natural decrease of AOD τ and precipitable water u_{H_2O} with altitude, and for the possible difference in elevation between the site under scrutiny and the corresponding MERRA-2 grid cell. This is done via a scale-height approximation [24,98]:

$$\tau(h) = \tau(h_0)\, e^{(h_0-h)/H_a}, \quad u_{H_2O}(h) = u_{H_2O}(h_0)\, e^{(h_0-h)/H_v}, \tag{3}$$

where h [amsl] is the elevation of the ground validation station, and h_0 [amsl] is the elevation above mean sea level of the corresponding MERRA-2 grid cell. This cell height is itself obtained by dividing the surface geopotential height (MERRA-2's PHIS variable) by the average standard gravity (9.80665 m s^{-2}) [99]. For simplification and lack of site-specific data, the scale heights for τ and u_{H_2O} are both considered constant, and taken as $H_a = H_v = 2100$ m [24,98]. Starting from the native MERRA-2 data at 1-h temporal resolution, interpolation using a Piecewise Cubic Hermite Interpolating Polynomial (PCHIP)—an option in Matlab's `interp2` interpolation function—is employed to derive the required 1-min estimates. PCHIP is preferred here over the more usual spline interpolation because of the latter's difficulty in handling two successive values that are very similar in magnitude and followed by a steep ramp, which can result in possible over- or under-estimation, and even sometimes in non-physical negative values. PCHIP is closer to linear interpolation with much more rounded transitions and fewer deviations from the estimated values.

Table 2
Hourly atmospheric input data extracted from the MERRA-2 database.

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

2.3. Validation sites

The validation database consists of 1-min irradiance time series from a pool of 100 ground radiometric stations. As illustrated in Fig. 1, those stations represent a substantial spread of climate and global distribution. For additional scrutiny, they are also collated using the popular Köppen-Geiger (KG) climate classification [100]. Only the five KG main classes are considered here for simplification and statistical significance: Equatorial, Arid, Temperate, Cold and Polar climates; they are denoted A, B, C, D and E, respectively, in what follows. The radiometric observations from all validation sites are obtained from the six sources succinctly described below. More details can be found in the supplementary material.

- *Baseline Surface Radiation Network (BSRN).* BSRN [101] is an affiliate of the World Meteorological Organization (WMO) and maintains the highest standards of quality control (QC) [92]. The solar irradiance observations are contained in BSRN's LR0100 product, and 53 of its stations are included in this study.
- *The Bureau of Meteorology, Australia (BoM).* BoM (`bom.gov.au/climate/data/oneminsolar`) provides 1-min solar data for Australia and the Cocos (Keeling) Islands. It employs Kipp & Zonen CM11 pyranometers and CH1 pyrheliometers to measure DIF and DNI, respectively. These observations follow the reporting and QC standards of WMO. Twelve of its validation sites are included in this study.
- *The South African Universities Radiation Network (SAURAN).* SAURAN (`sauran.ac.za`) employs Kipp & Zonen CMP11 pyranometers for DIF measurements and CHP1 pyrheliometers for DNI. Nineteen of its validation sites in southern Africa are included in this study.
- *The National Renewable Energy Laboratory, U.S.A. (NREL).* NREL (`midcdmz.nrel.gov`) contributes 13 validation sites in this study. Users can choose data from their desired sensors, whose brand and type vary depending on site.
- *National Oceanic and Atmospheric Administration, U.S.A. (NOAA).* The NOAA's Global Monitoring Division of the Earth System Research Laboratory (`esrl.noaa.gov/gmd/dv/data`) provides two

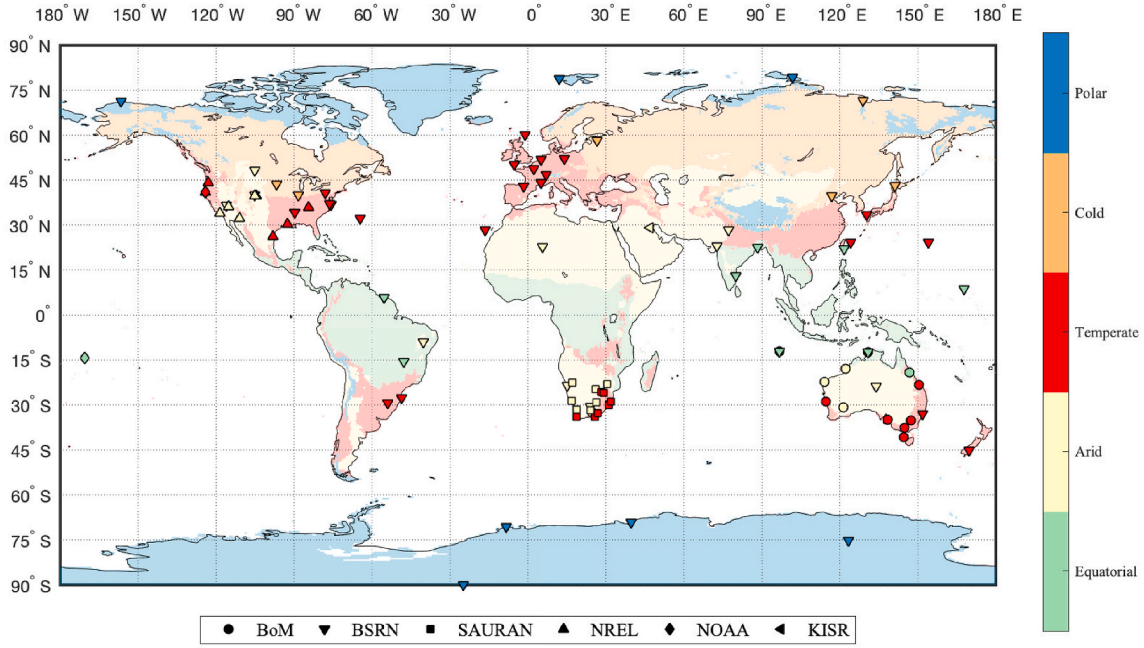

Fig. 1. World map marked with 100 ground radiometric stations. The colors and shapes of the marks denote the five Köppen-Geiger main classes and various data sources, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

validation sites for this study. Eppley PSP pyranometers and NIP pyrheliometers are used for DIF and DNI measurements, respectively.

- *The Kuwait Institute of Scientific Research (KISR).* KISR (`kisr.edu.kw`) [23,102,103] provides the Shagaya validation site with Kipp & Zonen CMP21 pyranometer for DIF and CHP1 pyrheliometer for DNI. It undergoes daily cleaning to minimise the soiling impact in arid environment.

3. Assessment methodology

The performance of the selected DNIcs and DIFcs models is assessed using the same methodology as in the earlier study that focused on GHIcs [4]. To summarize, the general framework contains the following steps:

1. *Data cleaning and clear-sky detection for all validation sites.* 181, 593, 072 raw 1-min pairs of DNI and DIF data points are obtained from the 100 validation sites for the 2015–2019 period. They are first subjected to QC screening for physical and extreme limits, and then to a process of extracting clear-sky periods. For that purpose, the Bright-Sun clear-sky detection (CSD) algorithm [19] is selected for its superior performance across different climates. The Bright-Sun CSD model is a freely available CSD method that is fully defined in Refs. [19]; in summary, an initial estimate of clear-sky periods is made before optimizing the clear-sky irradiance to the actual measured irradiance of the location. A second round of CSD is then performed using all three irradiance components (global, direct, diffuse) as indicators of clearness. All components must assert a clear-sky verdict for that period to be considered as clear. Finally, filters are passed over the data to remove periods of clear-sky that are close to a ramping event or that are too short a duration. If an irradiance measurement is detected as "not-clear" after clear-sky detection, or if the solar zenith angle exceeds 85°, the corresponding DNI and DIF data points are removed from consideration. Finally, 24, 009, 474 valid 1-min measurements (13.22%) remain for DNI, including 1, 508, 165 (6.28%) in Equatorial, 10, 154, 772 (42.29%) in Arid, 8, 988, 675 (37.44%) in Temperate, 2, 236, 187 (9.31%) in Cold, and 1, 121, 675 (4.67%) in Polar climates. For DIF, the corresponding numbers are 18, 709, 830 valid measurements overall (10.30%), including 999, 165 (5.34%) in Equatorial, 8, 484, 142 (45.35%) in Arid, 6, 733, 891 (35.99%) in Temperate, 1, 626, 389 (8.69%) in Cold, and 866, 243 (4.63%) in Polar climates.

2. *Calculate DNIcs and DIFcs for all models at all validation sites.* All models are operated at the center time of each 1-min measurement period, using the corresponding MERRA-2 inputs after they have been synchronised using the interpolation process explained earlier.

3. *Calculate 12 performance metrics* [92] *for all models at all validation sites.* The 12 performance metrics include mean bias difference (MBD), root mean square difference (RMSD), mean absolute difference (MAD), slope of best-fit line (SBF), uncertainty at 95% (U_{95}), t-statistic (TS), coefficient of determination (R^2), Willmott's index of agreement (WIA), Legates's coefficient of efficiency (LCE), Kolmogorov-Smirnov test integral (KSI), critical limit overestimation (OVER), and combined performance index (CPI). It is emphasized that identical validation data are used for all clear-sky irradiance models except McClear (#95), because it does not provide predictions for sites beyond ±89° of latitude, and thus is validated with fewer observations because it cannot use the 193, 913 valid DNIcs and 148, 309 valid DIFcs clear-sky measurements that exist at the BSRN South Pole station (SPO).

4. *Compute the ranking scores of each model r_x for both DNIcs and DIFcs.* This is done separately for six geographic areas: each of the five Köppen-Geiger (KG) climate classes, and the encompassing "global" world climate, i.e., $x = (A, B, C, D, E, G)$.

1. *Standardise the performance results from each metric used here, $M_m (m = 1, 2, \ldots, 12)$, for each model.* The absolute value of each error metric is used and then standardised with a normal score (or z-score). This usually leads to a standardised normal distribution for each error metric with zero mean and unit variance. As shown in Gueymard et al. [92], the subscript of M_m refers to the error metrics, e.g. $m = 1$ corresponds to the MBE, $m = 2$ is for RMSD, …, and $m = 12$ is for CPI.

2. *Apply Principal Component Analysis (PCA) via the pca Matlab function.* The $pca(X)$ function takes an 95-by-12 (row-by-column; 88-by-12 for DIFcs) matrix where each column is an error metric and each

row is the value per model. The PCA returns a coefficient $c_{i,m}$ for each principal components ($i = 1, 2, \ldots, 12$) and for each error metric ($m = 1,2,\ldots,12$), and also returns the explained variance e_i. Here, only the first five principal components $i = (1, 2, 3, 4, 5)$ are considered because they can capture more than 99.2% of the explained variance e_i. For example, in the DNIcs evaluation under the global climate category, the first five $i = (1, 2, 3, 4, 5)$ components account for 79.6%, 13.2%, 5.0%, 1.1% and 0.7% of the explained variance, respectively.

3. *Compute the ranking score r_x of DNIcs and DIFcs for each model at each climate class $x = (A, B, C, D, E)$.* Performed using the metrics (M_m), their respective weighting per component ($c_{i,m}$) and the explained variance of each principal component (e_i).

$$r_x = \sum_{i=1}^{5} e_i \left(\sum_{m=1}^{12} M_m \times c_{i,m} \right), \tag{4}$$

4. *Repeat the same PCA procedure to obtain the global ranking score r_G by taking the mean of each error metric over the five climate classes.* Here, absolute values of MBE are used before being averaged across all climates.

5. *Rank the clear-sky irradiance models for DNIcs and DIFcs.* For each climate, this is done according to their ranking score, such that a larger value of r_x represents better performance, and vice versa.

4. Results and discussion

Following the assessment procedure above, the predictions from 95 clear-sky beam and 88 diffuse models (Table 1) are compared to 1-min irradiance observations at all 100 ground stations shown in Fig. 1 (see further details in the Supplementary Material). The results include a table of the 12 error metrics per model, a table of the PCA coefficients and explained variances, and a table of ranking score, for DNIcs, DIFcs and GHIcs, and for each of the six climate classes.

From those 54 tables (fully listed in the supplementary material), the significance of the ranking score r_x is discussed first in subsection 4.1. Further discussion is provided about the top 10 clear-sky irradiance models for DNIcs and DIFcs at each Köppen-Geiger climate (subsection 4.2). For solar applications that require universal coverage, the top-10 performers are discussed in subsection 4.3. Finally, the impact on model performance from various formulations of broadband AOD and Linke turbidity factor is addressed in subsection 4.4.

4.1. Ranking score significance

Ranking score, r_x, is essential to quantify the performance of clear-sky irradiance models. Fig. 2 illustrates the distributions of model ranking score r_x via six box plots (corresponding to the global coverage and five climate classes), and for both DNIcs and DIFcs. Compared to the DIFcs models, it is obvious that the DNIcs models exhibit a broader range of performance. Curiously, however, the interquartile range of the DNIcs results for the Polar climate is smaller than that of all other climate classes, whereas the situation is completely reversed in the case of DIFcs. It is important to note that the actual diffuse irradiance is normally very low at polar sites because of prevalent low-sun and low-AOD conditions. Frequent or systematic experimental issues, such as frost, snow, or thermal offset, might negatively bias the diffuse observations, causing a significant relative impact on the measurement time series that cannot be easily corrected. Additionally, the model predictions can be affected by the bias in the MERRA-2 AOD data, which is not negligible in the Polar climate class [24]. These sources of inaccuracy may either compensate or compound each other, depending on site, season, etc. One way of testing whether the source of disagreement noted above lies in the models' design or in the measurement accuracy would be to benchmark results with those from a reference physical model of proven universal accuracy, such as the Rapid Radiative Transfer Model (RRTMG) [11,13], and test the models at sites where a co-located AERONET station is available to avoid AOD-induced error propagation. In any case, this issue is beyond the scope of this paper, but would definitely require closer scrutiny in the future.

For a closer examination of multi-climate results, the statistics of r_x, $x = (G, A, B, C, D, E)$ for DNIcs and DIFcs are tabulated in Table 3. By taking the 50th percentile, $P50$, as the benchmark for moderate performance, its higher value in a certain climate class indicates good average performance of the whole pool of clear-sky irradiance models. For DNIcs prediction, the Temperate climate scores the lowest $P50$, whereas the Polar climate exhibits the lowest and only negative value of $P50$ for DIFcs. Relatively to DIFcs, and for each climate except Polar, larger values of $P50$ are also found for DNIcs. Such singularity may be due to the exceptional performance of model #76 PSI-REST, which yields the highest ranking score $r_x = 3.054$ among all climates. It is also noted here that the mean value of r_x in each climate class always approximately zero due to taking the z-score.

Overall, it is found that models #2 (HS), #14 (Samimi), and #15 (Fu & Rich) are the bottom performers of DNIcs in all climates. For DIFcs, models #14 (BCLSM) and #56 (Ideriah) provide the worst estimates.

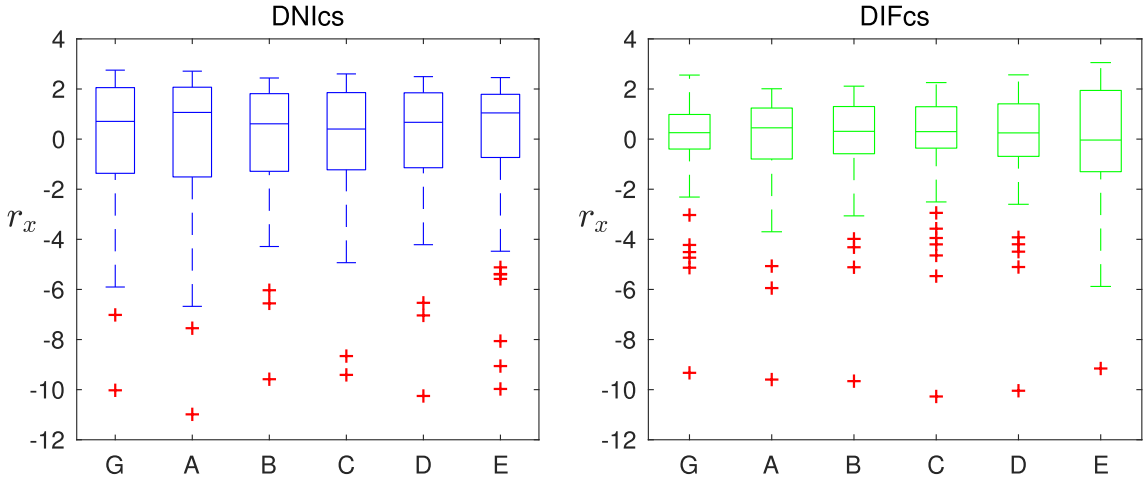

Fig. 2. Boxplots of the ranking score r_x for DNIcs and DIFcs within each climate x: Global (G), Equatorial (A), Arid (B), Temperate (C), Cold (D), and Polar (E). Outliers are determined if they are outside of $P75 \pm 1.5 \times (P75 - P25)$ where P is a percentile and 1.5 is the default whisker length in Matlab, which represents approximately 2.7σ if normally distributed.

Table 3
Statistical metrics of the ranking scores $r_x, x = (G, A, B, C, D, E)$ for DNIcs and DIFcs within each climate x. Here, $Pn(r_x)$ represents the n^{th} percentile.

	DNIcs						DIFcs					
Climate	mean	min.	max.	*P*25	*P*50	*P*75	mean	min.	max.	*P*25	*P*50	*P*75
G (Global)	0.0	−10.026	2.753	−1.365	0.708	2.055	0.0	−9.329	2.556	−0.396	0.256	0.984
A (Equatorial)	0.0	−10.985	2.712	−1.510	1.065	2.071	0.0	−9.597	2.009	−0.794	0.450	1.240
B (Arid)	0.0	−9.583	2.439	−1.285	0.611	1.814	0.0	−9.660	2.113	−0.584	0.310	1.300
C (Temperate)	0.0	−9.411	2.602	−1.225	0.402	1.856	0.0	−10.272	2.256	−0.360	0.298	1.293
D (Cold)	0.0	−10.252	2.494	−1.142	0.671	1.848	0.0	−10.044	2.565	−0.689	0.248	1.408
E (Polar)	0.0	−9.972	2.455	−0.733	1.045	1.788	0.0	−9.155	3.054	−1.300	−0.038	1.943

4.2. Climate-specific results

Tables 4–8 list the top-10 performing models for DNIcs and DIFcs under the A) Equatorial, B) Arid, C) Temperate, D) Cold and E) Polar climates, respectively. The REST2v9.1, REST2v5 and MMAC-V2 models show consistently good performance in the estimation of DNIcs for all climates, while only PSI-REST model exhibits such competence in estimating DIFcs in every climate class. The performance disparity may be caused by (i) the complex physical impacts from variant climate features; (ii) the empirical nature of most clear-sky irradiance models, which were developed with data from a specific climate only; and (iii) the possible lower quality of diffuse irradiance observations under harsh climates, such as at high-latitude or high-elevation stations.

For the Equatorial climate, the REST2v5 model is featured in the top lists for DNIcs and DIFcs. For the Arid climate, the Heliosat1-I and Ineichen2008 models show consistent performance for DNIcs and DIFcs. The Heliosat models exhibit good performance over the Temperate climate for both DIFcs and DIFcs. This result could be expected since the original Heliosat model was trained on data from mostly European sites. Reciprocally, these models do not perform as well in all other climate classes, owing to their empirical nature. For the Cold climate, REST2v9.1, REST2v5, Iqbal-C and Ineichen2008 show remarkable consistency in estimating both DNIcs and DIFcs. Although limited data, higher measurement uncertainty, and wildly fluctuating weather prevent high prediction accuracy in the Polar climate, REST2v5, Iqbal-C, MRMv5 and all three broadband-AOD versions of MMAC model show consistently good performance in estimating the irradiance components.

Among the five model classes, models with the largest number of inputs (class *e*) provide the best performance for all climates. This corroborates the results from previous studies [9,11], and could be expected because precise modelling tends to require a detailed description of the local atmospheric conditions. Models of class *e* collectively provide 76% and 48% of the top performing models in estimating DNIcs and DIFcs. The second best performing models are those of class *c*, which make use of the Linke turbidity coefficient. For DNIcs, models of this class capture the top spots in the Equatorial, Arid and Temperate climates. In particular, Heliosat1-I and ESRA-I are the best models in the Equatorial and Arid climates, respectively. For DIFcs estimation in the Arid climate, Heliosat1-R is the best performer. Also for DIFcs estimation, but in the Temperate climate, half of the 10 top models are from class *c*. In contrast, none of the class-*c* models are featured in the top performers list for the Cold and Polar climates, except ESRA-R that is ranked 6th for estimating DIFcs in the Cold climate. Three models from class *d*—namely, Calinoiu, Ineichen2008 and Yang—also demonstrate excellent performance in all climates except Polar. Models of class *b* only provide good estimates of DNIcs in the Polar climate, in which CLS is ranked #1. This surprising result might indicate an artifact caused by the larger measurement uncertainty in that climate. Regarding DIFcs, all class-*b* models are absent from the top ranks for any climate.

A surprising result is that McClear, despite its sophistication, does not perform better than many simpler models. An explanation is that its atmospheric inputs are *not* from MERRA-2, but from a different reanalysis database, namely the Copernicus Atmosphere Monitoring Service (CAMS) [104]. This suggests that the ranking of models found here (based on MERRA-2 inputs, except for McClear) is specific, and could be altered if a different database (e.g., CAMS) had been selected instead. This is an important finding that will require further scrutiny in the future.

Lastly, the models without atmospheric inputs (class *a*) demonstrate the poorest performance, as could be expected. They are barely featured in any top list except for the Polar climate, in which the Biga model is only ranked #10 for DIFcs with $r_E = 2.58$. All these findings suggest that leveraging more input variables tends to enable greater model accuracy and flexibility across climates.

4.3. Global performance

Table 9 shows the top 10 performers with respect to DNIcs and DIFcs, globally. For DNIcs, REST2v9.1 (class *e*) provides the best global performance ($r_G = 2.75$). One possible explanation is that the REST2 family of models incorporates a circumsolar correction to take the pyrheliometer aperture into account [105], which makes its predictions artificially closer to experimental measurements. Its former version, REST2v5 occupies the second best spot with only a slight difference in ranking score ($r_G = 2.67$). For DIFcs, the best performing model is Modified Iqbal-C of model class *e*. The second best DIFcs model is PSI-REST, also from model class *e*, which has excellent worldwide performance. Overall, models of class *e* are found to provide consistent global performance for DNIcs and

Table 4
Top 10 performing models for DNIcs and DIFcs in the Equatorial climate A.

	DNIcs				DIFcs			
	#		Model	r_A	#		Model	r_A
Climate A, Equatorial	1	c	Heliosat1-I	2.71	1	e	Modified Iqbal-C	2.01
	2	c	ESRA-I	2.69	2	e	PSI-REST	1.97
	3	e	MAC2	2.56	3	c	ESRA-GR	1.95
	4	e	McClear	2.48	4	c	Heliosat2-GR	1.92
	5	e	REST2v9.1	2.47	5	d	Calinoiu	1.87
	6	e	MMAC-V1	2.46	6	e	MRMv5	1.79
	7	e	REST2v5	2.45	7	d	Ineichen2008	1.76
	8	e	CPCR2	2.38	8	e	METSTAT-V3	1.76
	9	e	Bird	2.30	9	e	REST2v5	1.75
	10	e	MMAC-V2	2.30	10	d	Yang	1.74

Table 5
Top 10 performing models for DNIcs and DIFcs in the Arid climate B.

	DNIcs				DIFcs			
	#		Model	r_B	#		Model	r_B
Climate B, Arid	1	C	ESRA-I	2.44	1	c	Heliosat1-R	2.11
	2	C	Heliosat1-I	2.40	2	e	MRMv5	2.03
	3	E	METSTAT-V1	2.36	3	c	Heliosat2-R	1.94
	4	E	REST2v9.1	2.28	4	d	Yang	1.88
	5	E	MRMv6.1	2.23	5	d	Ineichen2008	1.83
	6	E	MMAC-V2	2.23	6	e	Modified Iqbal-C	1.81
	7	E	Iqbal-C	2.21	7	e	PSI-REST	1.74
	8	D	Ineichen2018	2.18	8	c	ESRA-GR	1.69
	9	E	REST2v5	2.14	9	c	Heliosat1-M	1.68
	10	D	Ineichen2008	2.13	10	c	Heliosat1-I	1.68

Table 6
Top 10 performing models for DNIcs and DIFcs in the Temperate climate C.

	DNIcs				DIFcs			
	#		Model	r_C	#		Model	r_C
Climate C, Temperate	1	e	REST2v9.1	2.60	1	d	Calinoiu	2.26
	2	e	REST2v5	2.55	2	d	Ineichen2008	2.21
	3	c	ESRA-I	2.49	3	e	Modified Iqbal-C	2.05
	4	c	Heliosat1-I	2.48	4	c	Heliosat1-R	1.95
	5	e	MMAC-V1	2.32	5	c	Heliosat1-GU	1.85
	6	e	MMAC-V2	2.30	6	c	Heliosat1-I	1.84
	7	e	Iqbal-C	2.30	7	c	Heliosat1-M	1.83
	8	e	MAC2	2.27	8	c	Heliosat2-R	1.77
	9	e	METSTAT-V1	2.22	9	e	PSI-REST	1.70
	10	e	Bird	2.19	10	d	Yang	1.66

Table 7
Top 10 performing models for DNIcs and DIFcs in the Cold climate D.

	DNIcs				DIFcs			
	#		Model	r_D	#		Model	r_D
Climate D, Cold	1	E	REST2v9.1	2.49	1	e	Modified Iqbal-C	2.56
	2	E	REST2v5	2.49	2	e	Iqbal-B	2.21
	3	E	METSTAT-V1	2.45	3	d	Calinoiu	2.20
	4	E	Iqbal-C	2.44	4	e	REST2v5	2.17
	5	E	PSIM	2.36	5	d	Ineichen2008	2.09
	6	E	CPCR2	2.34	6	c	ESRA-R	2.08
	7	E	MMAC-V2	2.31	7	e	REST2v9.1	2.04
	8	E	MMAC-V3	2.30	8	e	PSI-REST	2.03
	9	D	Ineichen2018	2.28	9	e	Iqbal-C	1.96
	10	D	Ineichen2008	2.25	10	d	Yang	1.86

Table 8
Top 10 performing models for DNIcs and DIFcs in the Polar climate E.

	DNIcs				DIFcs			
	#		Model	r_E	#		Model	r_E
Climate E, Polar	1	b	CLS	2.46	1	e	PSI-REST	3.05
	2	e	REST2v9.1	2.32	2	e	MMAC-V2	2.90
	3	e	MMAC-V1	2.30	3	e	MMAC-V3	2.79
	4	e	MMAC-V2	2.30	4	e	MMAC-V1	2.78
	5	e	REST2v5	2.28	5	e	MRMv6.1	2.78
	6	e	MMAC-V3	2.27	6	e	REST2v5	2.75
	7	d	REST	2.20	7	e	Bird	2.73
	8	e	Iqbal-C	2.19	8	e	MRMv5	2.69
	9	e	METSTAT-V1	2.17	9	e	Iqbal-C	2.59
	10	e	MRMv5	2.15	10	a	Biga	2.58

DIFcs, thus confirming expectations.

For a closer look in global estimates of DNIcs and DIFcs, heat maps are also provided in Fig. 3 for the top 30 models (minimum ($Rank_{DNIcs}$ + $Rank_{DIFcs}$)). Similar to scatter plots, the shade in these heat maps indicates a cluster of data points. From Fig. 3, REST2v5, Iqbal-C and MRMv5 model are found the three best models for the prediction of both

Table 9
Top 10 performing models for DNIcs and DIFcs on a Global climate basis.

	DNIcs				DIFcs			
	#		Model	r_G	#		Model	r_G
Global	1	e	REST2v9.1	2.75	1	e	Modified Iqbal-C	2.56
	2	e	REST2v5	2.67	2	e	PSI-REST	2.41
	3	e	MMAC-V2	2.61	3	e	MRMv5	2.36
	4	e	METSTAT-V1	2.61	4	e	REST2v5	2.08
	5	e	Iqbal-C	2.59	5	e	MRMv6.1	1.98
	6	e	MMAC-V1	2.50	6	e	Iqbal-C	1.92
	7	d	INEICHEN2018	2.46	7	d	CALINOIU	1.92
	8	d	REST	2.42	8	d	INEICHEN2008	1.83
	9	d	MLWT2	2.40	9	e	BIRD	1.83
	10	e	MRMv5	2.38	10	e	Iqbal-B	1.80

DNIcs and DIFcs. The second tier is populated by PSI-REST, REST2v9.1 and MRMv6.1, which are ranked in the top 13 models for both DNIcs and DIFcs. From the heatmaps, we observe that the DIFcs appears to have significant bias for all models with a significant dispersion. This demonstrates that precise modelling is difficult with the broadband radiative models because they tend to oversimplify the method of solution for the radiative transfer equation using empirical functions, whose validity is not necessarily guaranteed over all possible atmospheric situations, particularly regarding aerosol conditions. This can explain why some models tend to overestimate diffuse in all cases (e.g., REST2v5, MMACv2 or McClear), others tend to overestimate when DIFcs is low and underestimate when DIFcs is high (e.g., MRMv6.1), whereas others do the opposite (e.g., Ineichen2018). It is a mixed bag, but in any case the impact of this uncertainty is relatively minor compared to that in DNI (note the x- and y-axis limits in Fig. 3 are not equal for DNIcs and DIFcs).

Among the 7 DNIcs-only models, REST and MLWT2 are the top 2 performers. They also capture the 8*th* and 9*th* global ranking out of 95 models in estimating DNIcs. This was expected because both showed excellent results in earlier validation studies [3,6].

From the full ranking lists in the supplementary material, it is obvious that Fu & Rich (class *a*) consistently produces poor estimates of DNIcs and DIFcs in any climate class. Although various values (0.5 and 0.6) of the bulk atmospheric transmittance have been considered for "very clear" and "generally clear" conditions, the model performance hardly improves. This is consistent with previous analysis [9,54], and confirms that its (fixed) atmospheric transmittance constitutes an excessive oversimplification.

4.4. Impacts of inaccuracies in broadband AOD and Linke Turbidity

The broadband AOD ($\tau_{broadband}$) is meant to simplify the calculations related to aerosols by condensing their effective extinction effects into that single variable. It is an essential input to two clear-sky irradiance models: METSTAT and MMAC. It can be derived from other aerosol-related variables (α and β) with various empirical formulations. (All details can be found in the supplementary material.) To investigate the impacts of those formulations on the clear-sky irradiance model performance, each of the three known formulations is coupled with each of the two clear-sky irradiance models—MMAC and METSTAT—hence producing a total of six $\tau_{broadband}$-dependent models.

Fig. 4 plots the ranking score of those six models under the six possible climates for DNIcs and DIFcs. For DNIcs predictions, METSTAT performs best when associated with the *V*1 formulation [9]. For MMAC, however, things are less obvious because the *V*1 and *V*2 formulations yield relatively close results, with only a slight advantage for the latter on a Global basis. For DIFcs, the performance score fluctuates widely between −1 and 3, depending on climate class. This may be due to the generally poor performance of clear-sky irradiance models in estimating DIFcs, except in the Polar climate, where MMAC performs remarkably well, whereas METSTAT performs relatively poorly. This again illustrates the specificity of the Polar climate, which may call into question the reliability of diffuse data there. The *V*3 formulation does not fare well, so that neither METSTAT-*V*3 nor MMAC-*V*3 appear in the top 10 on a global climate basis. Nevertheless, results shown in the supplementary material indicate that METSTAT-*V*3 is ranked 14 globally for DIFcs. As before, things are different in the Polar climate: The different broadband AOD formulations do not affect the model performance much for DNIcs and DIFcs.

The impact of the Linke turbidity formulation is more severe than that of the $\tau_{broadband}$ formulation. Consistent with earlier results [4], the five T_L-based clear-sky irradiance models experience significant performance fluctuations when combined with the six possible T_L formulas. Globally, T_{L2Gu} [38] provides the best DNIcs performance for Heliosat1, ESRA and Heliosat2. For DIFcs, T_{L2R} [35] is the best formulation for the five clear-sky irradiance models in study, except Dogniaux. Overall, even the best possible T_L formulation does not necessarily make a T_L-based clear-sky irradiance model a top performer. This critical finding underlines the risks related to error propagation (from input to output), particularly in simple DNIcs and DIFcs models. Finding the ideal combination of a T_L-based clear-sky irradiance model and a T_L formulation that would guarantee accurate irradiance predictions worldwide still appears to be an elusive proposition. No T_L-based clear-sky irradiance model can thus be recommended for general usage, at least in terms of applications that involve both DNIcs and DIFcs.

5. Clear-sky global horizontal irradiance

This contribution advances the field of clear-sky irradiance model validation considerably through extended years of validation, an improved clear-sky detection algorithm that works in all climates, additional models, and additional validation sites, compared to previous studies. As such, it is possible to re-visit the findings from Sun et al. [4] and the GHIcs rankings therein. The supplementary material contains a complete re-evaluation of GHIcs models. For brevity, only the top five GHIcs models are discussed in this section. An additional topic is addressed, however, to explore whether combining the best possible DNIcs output with the best possible DIFcs output would improve the new hybrid model estimation of GHIcs.

Regarding the performance of existing GHIcs models, the present results are reasonably consistent with a previous assessment [4], despite the increased numbers of clear-sky irradiance models and validation sites. Table 10 lists the top 5 GHIcs models in each climate. On a global basis, the best two GHIcs models (MRMv6.1 and REST2v5) received almost the same highest ranking score of $r_G = 2.44$, with only a negligible difference of 0.0004. Three other models, namely PSI-REST, MAC2 and MRMv5, have a close ranking score between 2.20 and 2.30. Comparing such results to the previous analysis of GHIcs models [4], only 7 out of the original best 10 models are now found in the top global category, indicating some rank variations between the two studies.

In the present study, nearly all the top ranked GHIcs models of each climate are from class *e*, with the Temperate climate being the only exception. There, the best model is Ineichen2008 (from class *d*), and the

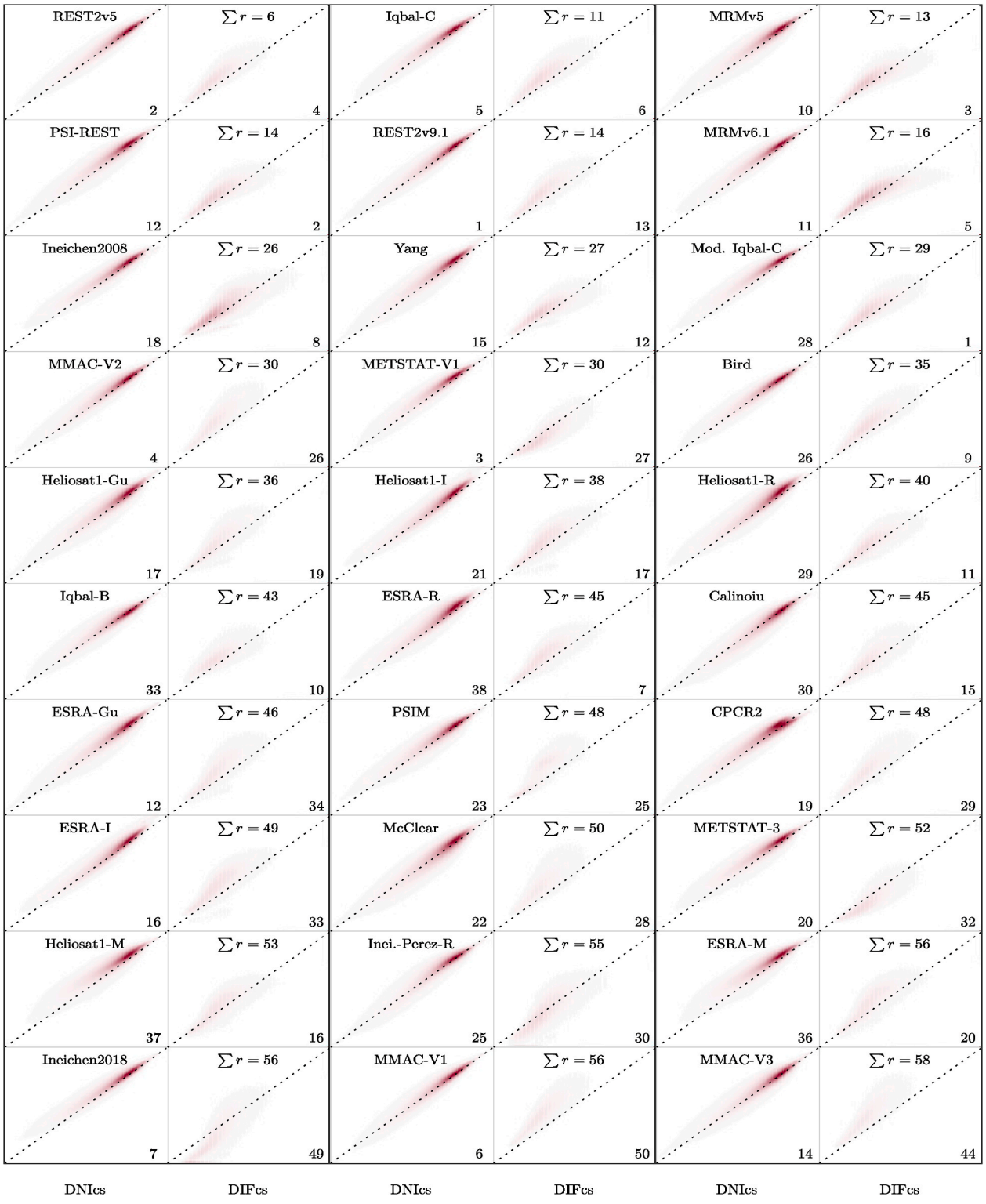

Fig. 3. Heat plots of the top 30 clear-sky irradiance models in global estimations of Minimum ($Rank_{DNIcs} + Rank_{DIFcs}$) against ground observations. The rank of the model in the global rank of DNIcs or DIFcs is marked at the bottom right corner of each subfigure and dark shade indicates a cluster of data points. Note the different scales in the DNIcs and DIFcs scatter plots. The X- and Y-axis limits are from 0 to 1200 W m^{-2} for DNIcs and only 0–200 W m^{-2} for DIFcs. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

5th best model is Josefsson (from class *b*).

In the present results, one important feature is that the DIFcs predictions still display significant scatter when plotted against ground observation, as shown in Fig. 3. This indicates that the fundamental drivers for DIFcs are not fully captured by any model. Other sources of inaccuracy include the presence of unresolved clouds and experimental errors (e.g., in the Polar climate). Note, however, that DIFcs is normally much lower than DNIcs, so that absolute RMS differences (in W m^{-2}) are of comparable magnitude. Considering the natural compensation of errors between DNIcs and DIFcs, their scatter is also much less than that

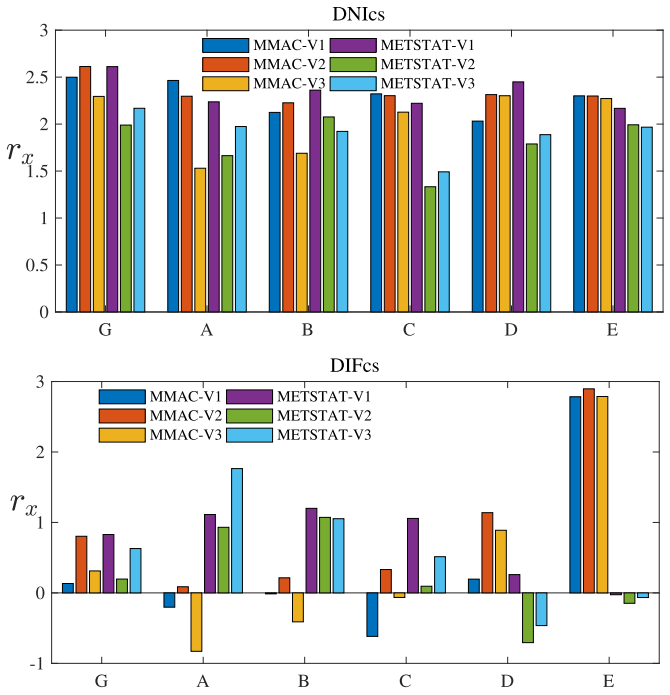

Fig. 4. Ranking score r_x of six clear-sky irradiance models with various broadband AOD formulations for their estimation of DNIcs and DIFcs within the climate x: Global (G), Equatorial (A), Arid (B), Temperate (C), Cold (D) and Polar (E).

in GHIcs predictions. Interestingly, the top performing global GHIcs model found here is MRMv6.1, even though it displays some non-linearity in its DIFcs heatmap (Fig. 3), and a negative bias in DNIcs. It is perhaps coincidental that these two patterns combine to form a very good GHIcs alignment. For these reasons, it is argued that, in the future, the focus of clear-sky irradiance modelling should move from GHIcs to its DNIcs and DIFcs components.

As a preliminary step in that direction, it seems logical that a hybrid model GHIcs model of potentially superior performance can be constructed by combining the best DNIcs and DIFcs models. This is attempted here using the best public-domain DNIcs model (REST2v5) and the best DIFcs model (Modified Iqbal-C). Although not thoroughly detailed here for conciseness, this new REST2v5-M.Iqbal-C construct does indeed perform well globally, with $r_G = 2.19$, and outperforms the M. Iqbal-C GHIcs model ($r_G = 1.7$), but surprisingly not the REST2v5 GHIcs model ($r_G = 2.44$). This mixed result is likely because there is no guarantee that corresponding DIFcs and DNIcs predictions have correspondingly good estimates, hence a poor DIFcs estimate may align with a bad DNIcs estimate, resulting in a mediocre GHIcs estimate. Conversely, in a good overall model, individual errors in DNIcs and DIFcs are perhaps more likely to compensate each other.

A deeper analysis by Sun et al. [106] presented all combinations of the top 10 DIFcs with the top 10 DNIcs models from this research to evaluate GHIcs performance against the top 10 GHIcs models from Refs. [4]. It was found that, with the exception of the REST2v5 GHIcs which came in 3rd place, a better GHIcs model can be produced through combining different DIFcs and DNIcs models using the closure equation.

6. The future of clear-sky irradiance modelling

The previous section indicated that the best MERRA-2-based GHIcs model is MRMv6.1, despite its DNIcs and DIFcs ranking only in 11th and 5th place, respectively. Moreover, the respective biases in each irradiance component were normally found to compensate each other so as to produce a good GHIcs estimate. Still, combining the predicted DNIcs and DIFcs from two different models in a hybrid way does not necessarily improve upon GHIcs, which was not expected. Hence, it is argued that the future of clear-sky irradiance modelling should focus more heavily on DNIcs and DIFcs rather than on just GHIcs. Training fundamental relationships between atmospheric constituents and the incident irradiance should also be explored in a large sample of climates to capture the whole gamut of relationships between meteorological components. Previous studies [11,13] have suggested that the best DNIcs models used in solar engineering could rival in accuracy with a physical (hence "universal") radiative transfer model such as RRTMG, which rather appeals to the atmospheric science community. The present results, however, question the validity of DIFcs modelling in the models tested here, which can be explained by their empirical approach. The potential gain of accuracy of RRTMG or similar physical model with regard to DIFcs modelling remains to be quantified. For solar applications, a physical model has the disadvantage of being more difficult to install and operate, along with executing much slower. Hence, the development of a fast and high-performance DIFcs model should be contemplated.

Whereas, from Fig. 3, the scatter in the DIFcs predictions appears significant, it must be emphasized that part of it is actually caused by error propagation from inaccuracies in the MERRA-2 aerosol input data. Hence, an in-depth study would be needed to discriminate between intrinsic model errors and input-induced errors in DIFcs. This could be done using ideal setups combining high-accuracy point source simultaneous measurements of irradiance and atmospheric constituents, for example using AERONET ground stations co-located with irradiance measurement networks. However, the engineering solution of using worldwide gridded data is perhaps equally as important due to its

Table 10
Top five performing GHIcs models across all climates and globally. The model class (a to e) and ranking score (r_x where $x = (A,B,C,D,E,G)$) are indicated below each model's name.

	Global	Equatorial	Arid	Temperate	Cold	Polar
1st place class, r_x	MRMv6.1 e, 2.44	METST.-V1 e, 1.93	REST2v5 e, 2.33	Ineich.2008 d, 2.17	REST2v5 e, 2.24	MMAC-V1 e, 2.76
2nd place class, r_x	REST2v5 e, 2.44	MRMv6.1 e, 1.91	REST2v9.1 e, 2.15	METST.-V2 e, 2.12	MRMv5 e, 2.13	MMAC-V2 e, 2.75
3rd place class, r_x	PSI-REST e, 2.27	METST.-V2 e, 1.86	MRMv6.1 e, 2.12	REST2v5 e, 2.02	MRMv6.1 e, 2.08	MMAC-V3 e, 2.74
4th place class, r_x	MAC2 e, 2.24	REST2v5 e, 1.83	MAC2 e, 2.03	METST.-V3 e, 2.01	REST2v9.1 e, 2.06	PSI-REST e, 2.65
5th place class, r_x	MRMv5 e, 2.22	METST.-V3 e, 1.78	Iqbal-C e, 1.89	Josefsson b, 1.97	MAC2 e, 2.04	MRMv6.1 e, 2.63

uptake and application in solar resource assessment and forecasting tools. Variance can be expected in the performance of clear-sky irradiance models when supplied with either modeled gridded data or ground-based observations [93]. Hence, validating models using inputs from another reanalysis database (e.g., CAMS) or AERONET data would most likely change the ranking of models. It must be acknowledged that the best engineering solution is not necessarily the same as the ideal solution, because access to high accuracy input data is typically not practical at arbitrary locations around the world. The engineering solution is therefore dependent on the propagated errors from gridded datasets. For this reason, it would be worth training models using the MERRA-2 or other reanalysis input data in order to improve engineering-application performance. This would boost radiation model performance to the benefit of the whole solar industry. Considering the inadequacies in the MERRA-2 dataset, it is reiterated that the validation conducted here is not absolute and is strictly aimed at solar engineering applications.

7. Conclusion

This study has offered the most comprehensive and objective analysis of the performance of popular clear-sky irradiance models regarding their attempt at estimating the clear-sky direct normal irradiance (DNIcs) and diffuse horizontal irradiance (DIFcs) on a worldwide basis. A record number of 95 DNIcs models and 88 DIFcs models has been examined. Based on the large range of their inherent complexity, the models have been categorised into five classes (*a* to *e*), from simple inputs to multiple and detailed inputs. Those inputs have been extracted from the MERRA-2 hourly reanalysis, selected here for its high quality and global coverage. The validation has been conducted by comparing the modeled predictions to irradiance measurements conducted at 100 ground radiometric stations covering many climate regions of the world. After extracting 1-min clear-sky periods from the measured time series, all models have been validated against millions of quality-controlled observations of DINcs and DIFcs, during the period 2015–2019.

By using a PCA ranking procedure that incorporates 12 performance metrics, ranking scores have been obtained for each clear-sky irradiance model at five Köppen-Geiger climates (Equatorial, Arid, Temperate, Cold, Polar), as well as for the Global climate (whole world). For each climate class, the best models for DNIcs and DIFcs, as well as for their combination (GHIcs), have been found as follows:

	Best DNIcs	Best DIFcs	Best GHIcs
Global	REST2v9.1	Modified Iqbal-C	MRMv6.1
Equatorial	Heliosat1-I	Modified Iqbal-C	METSTAT-V1
Arid	ESRA-I	Heliosat1-R	REST2v5
Temperate	REST2v9.1	CALINOIU	INEICHEN2008
Cold	REST2v9.1	Modified Iqbal-C	REST2v5
Polar	CLS	PSI-REST	MMAC-V1

Interestingly, the global GHIcs performance obtained by combining the globally best DNIcs and DIFcs models (REST2v9.1 and Modified Iqbal-C, respectively), is ranked between its two components, and does not outperform the best existing GHIcs models of the literature. This suggests that more developments are necessary to improve the modelling of DNIcs and DIFcs separately, while preserving the best possible GHIcs predictions.

The modelling of DIFcs appears less satisfactory than that of DNIcs, probably because of the conceptual simplifications used in clear-sky radiation models compared to more physical radiative transfer models used in atmospheric sciences.

The impact of substituting between three possible broadband-AOD formulations on the clear-sky performance of two models using that input has been found to be relatively small. In contrast, the impact of selecting one of six possible formulations to derive the required Linke turbidity coefficient has appeared quite significant on the performance of six T_L-based clear-sky radiation models. Hence, the use of such models is not recommended for the prediction of DNIcs or DIFcs in worldwide applications.

The detailed results obtained here have shown that the PCA methodology is an efficient tool for the ranking of models, and that the Köppen-Geiger climate classification is useful in the process of evaluating the impact of local climate on model performance. In that respect, the results obtained for the Polar climate are significantly different than those obtained in any other climate. Considering the harsh conditions of irradiance measurements under a polar climate, both the actual uncertainty of these measurements and its ultimate impact on apparent model performance remain to be established.

One surprising finding has been the relatively disappointing performance of the most sophisticated specimen of the whole pool of 95 clear-sky radiation models—McClear. This is most likely related, at least in part, to the fact that McClear does not use MERRA-2 input data like all other models tested here, because it is hardwired by design to accept only data from another reanalysis. That model's performance is thus not exactly comparable to that of all other models. Hence, it is emphasized that the present model rankings are actually valid only when considering MERRA-2 inputs, and are not just an *intrinsic* measure of model performance. To address this important limitation, further studies would be needed to evaluate the sensitivity of the top models identified here to using a different source of input data. The present models, methodology, and findings specifically relate to applied solar engineering applications. In particular, this analysis is intrinsic to model pairing with MERRA-2 reanalysis data. It is emphasized that better results and a different model ranking could have been obtained with different (and hopefully more accurate) input data. It is suggested that future studies be conducted about the feasibility of appropriately correcting the bias in reanalysis input data, possibly on a regional climate basis.

All public-domain clear-sky radiation models, the Bright-Sun clear-sky detection method, and the assessment methodology have been coded in Matlab, R or Python. This code can be freely accessed online [26]. Access to MERRA-2 data and some python versions of clear-sky radiation models is found at the irradpy python package repository.

CRediT authorship contribution statement

Xixi Sun: Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing - original draft, Writing - review & editing, Visualization, Project administration. **Jamie M. Bright:** Conceptualization, Methodology, Software, Validation, Resources, Data curation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration. **Christian A. Gueymard:** Resources, Data curation, Writing - original draft, Writing - review & editing, Supervision. **Xinyu Bai:** Software, Resources, Data curation. **Brendan Acord:** Writing - review & editing, Supervision. **Peng Wang:** Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

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

Acknowledgements

X. Sun, X. Bai and P. Wang were partially supported by the National Key Research and Development Program of China (Grant No. 2017YFB0701700 and 2018YFB0703902) and the National Natural Science Foundation of China (Grant No. 61672080). J.M Bright is funded by the Energy Market Authority (EMA), Energy Programme—Solar Forecasting Grant (NRF2017EWT-EP002-004). SERIS is a research institute at the National University of Singapore (NUS). SERIS is

supported by the National University of Singapore (NUS), the National Research Foundation Singapore (NRF) and the Singapore Economic Development Board (EDB). The authors 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.2020.110087. The spreadsheet supplementary material contains every error metric from each model in terms of GHIcs, DIFcs and DNIcs for each climate (equatorial, arid, temperate, cold and polar) as well as the global category. Additionally, we provide the complete ranking lists and PCA outputs. The document supplementary material contains model descriptions that differ from our previous work.

References

[1] Bright JM. The impact of globally diverse ghi training data: evaluation through application of a simple Markov chain downscaling methodology. J Renew Sustain Energy 2019;11(2):023703. https://doi.org/10.1063/1.5085236.
[2] Gueymard CA, Ruiz-Arias JA. Extensive worldwide validation and climate sensitivity analysis of direct irradiance predictions from 1-min global irradiance. Sol Energy 2016;128:1–30.
[3] 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.
[4] Sun X, Bright JM, Gueymard CA, Acord B, Wang P, Engerer NA. Worldwide performance assessment of 75 global clear-sky irradiance models using principal component analysis. Renew Sustain Energy Rev 2019;111:550–70.
[5] Stafford J, Wendler G, Curtis J. Temperature and precipitation of Alaska: 50 year trend analysis. Theor Appl Climatol 2000;67(1–2):33–44.
[6] Gueymard C. Direct solar transmittance and irradiance predictions with broadband models . Part II : validation with high-quality measurements. Sol Energy 2003;74:381–95.
[7] Madkour M, El-Metwally M, Hamed A. Comparative study on different models for estimation of direct normal irradiance (DNI) over Egypt atmosphere. Renew Energy 2006;31(3):361–82.
[8] Ineichen P. Comparison of eight clear sky broadband models against 16 independent data banks. Sol Energy 2006;80(4):468–78.
[9] 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.
[10] Engerer N, Mills F. Kpv: a clear-sky index for photovoltaics. Sol Energy 2014;105:679–93.
[11] Gueymard CA, Ruiz-Arias JA. Validation of direct normal irradiance predictions under arid conditions: a review of radiative models and their turbidity-dependent performance. Renew Sustain Energy Rev 2015;45:379–96.
[12] Ineichen P. Validation of models that estimate the clear sky global and beam solar irradiance. Sol Energy 2016;132:332–44.
[13] 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.
[14] Polo J, Martín-Pomares L, Sanfilippo A. Solar resources mapping: fundamentals and applications. Springer; 2019.
[15] Antonanzas-Torres F, Urraca R, Polo J, Perpinan-Lamigueiro O, Escobar R. Clear sky solar irradiance models: a review of seventy models. Renew Sustain Energy Rev 2019;107:374–87.
[16] 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.
[17] Badescu V, Gueymard CA, Cheval S, Oprea C, Baciu M, Dumitrescu A, Iacobescu F, Milos I, Rada C. Accuracy and sensitivity analysis for 54 models of computing hourly diffuse solar irradiation on clear sky. Theor Appl Climatol 2013;111(3–4):379–99.
[18] Gueymard CA, Bright JM, Lingfors 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.
[19] Bright JM, Sun X, Gueymard CA, Acord B, Wang P, Engerer NA, Bright-Sun. A globally applicable 1-min irradiance clear-sky detection model. Renew Sustain Energy Rev 2020;121:109706.
[20] Bright JM, Gueymard CA. Climate-specific and global validation of modis aqua and terra aerosol optical depth at 452 aeronet stations. Sol Energy 2019;183:594–605. https://doi.org/10.1016/j.solener.2019.03.043. URL, http://www.sciencedirect.com/science/article/pii/S0038092X19302713.
[21] Giles DM, Sinyuk A, Sorokin MG, Schafer JS, Smirnov A, Slutsker I, Eck TF, Holben BN, Lewis JR, Campbell JR, et al. Advancements in the Aerosol Robotic Network (AERONET) Version 3 database–automated near-real-time quality control algorithm with improved cloud screening for Sun photometer aerosol optical depth (AOD) measurements. Atmospheric Measurement Techniques 2019;12(1).
[22] 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.
[23] 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, IEA SHC international conference on solar heating and cooling for buildings and industry and ISES solar world congress. 2017.
[24] Gueymard CA, Yang D. Worldwide validation of CAMS and MERRA-2 reanalysis aerosol optical depth products using 15 years of AERONET observations. Atmos Environ 2019:117216.
[25] Gelaro R, McCarty W, Suárez MJ, Todling R, Molod A, Takacs L, Randles CA, Darmenov A, Bosilovich MG, Reichle R, et al. The Modern-Era retrospective analysis for research and applications, version 2 (MERRA-2). J Clim 2017;30(14):5419–54.
[26] Bright JM, Sun X. A library of clear-sky irradiance models coded in R [online], https://jamiembright.github.io/clear-sky-models/; 2018.
[27] Bright JM, Bai X, Zhang Y, Sun X, Acord B, Wang P. irradpy: Python package for MERRA-2 download, extraction and usage for clear-sky irradiance modelling. Sol Energy 2020;199:685–93.
[28] Bright JM, Engerer NA. Engerer2: global re-parameterisation, update, and validation of an irradiance separation model at different temporal resolutions. J Renew Sustain Energy 2019;11(3):033701.
[29] Gueymard C. Importance of atmospheric turbidity and associated uncertainties in solar radiation and luminous efficacy modelling. Energy 2005;30:1603–21.
[30] Yang D, Boland J. Satellite-augmented diffuse solar radiation separation models. J Renew Sustain Energy 2019;11(2):023705.
[31] 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.
[32] Gueymard CA. Direct solar transmittance and irradiance predictions with broadband models. part i: detailed theoretical performance assessment. Sol Energy 2003;74(5):355–79.
[33] Gueymard CA. A reevaluation of the solar constant based on a 42-year total solar irradiance time series and a reconciliation of spaceborne observations. Sol Energy 2018;168:2–9.
[34] Blanc P, Wald L. The SG2 algorithm for a fast and accurate computation of the position of the Sun for multi-decadal time period. Sol Energy 2012;86(10):3072–83.
[35] Remund J, Wald L, Lefèvre M, Ranchin T, Page J. Worldwide Linke turbidity information. ISES Solar World Congress; 2003.
[36] Dogniaux R. Représentations analytiques des composantes du rayonnement lumineux solaire: conditions de ciel serein. Institut royal météorologique de Belgique; 1974.
[37] Ineichen P. Conversion function between the Linke turbidity and the atmospheric water vapor and aerosol content. Sol Energy 2008;82(11):1095–7.
[38] Gueymard CA. Turbidity determination from broadband irradiance measurements: a detailed multicoefficient approach. J Appl Meteorol 1998;37(4):414–35.
[39] Molineaux B, Ineichen P, Delaunay J-J. Direct luminous efficacy and atmospheric turbidity - improving model performance. Sol Energy 1995;55(2):125–37.
[40] 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.
[41] Threlkeld J, Jordan R. Direct solar radiation available on clear days, Heat. Piping Air Cond 1957;29(12). 1957.
[42] Badescu V, Gueymard CA, Cheval S, Oprea C, Baciu 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.
[43] Paltridge GW, Platt CMR. Radiative processes in meteorology and climatology. Develop. Atmosph. Sci. 1976;5:344.
[44] Schulze R. A physically based method of estimating solar radiation from suncards. Agric Meteorol 1976;16(1):85–101.
[45] Daneshyar M. Solar radiation statistics for Iran. Sol Energy(United States) 1978;21(4). 1978.
[46] Biga A, Rosa R. Contribution to the study of the solar radiation climate of Lisbon. Sol Energy 1979;23(1):61–7.
[47] of Heating AS, Refrigerating A-C. Engineers, ASHRAE handbook: refrigeration systems and applications. American Society of Heating, Refrigerating and Air Conditioning Engineers; 1985.
[48] of Heating AS, Refrigerating A-C. Engineers, ASHRAE handbook: refrigeration systems and applications. American Society of Heating, Refrigerating and Air Conditioning Engineers; 2005.
[49] El Mghouchi Y, El Bouardi A, Choulli Z, Ajzoul T. New model to estimate and evaluate the solar radiation. Int. J. Sustainable Built Environ. 2014;3(2):225–34.
[50] Sharma M, Pal R. Interrelationships between total, direct, and diffuse solar radiation in the tropics. Sol Energy 1965;9(4):183–92.
[51] Meinel AB, Meinel MP. Applied solar energy: an introduction. STIA 1977;77:33445.
[52] Hottel HC. A simple model for estimating the transmittance of direct solar radiation through clear atmospheres. Sol Energy 1976;18(2):129–34.
[53] Samimi J. Estimation of height-dependent solar irradiation and application to the solar climate of Iran. Sol Energy 1994;52(5):401–9.

[54] 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.
[55] Barbaro S, Coppolino S, Leone C, Sinagra E. An atmospheric model for computing direct and diffuse solar radiation. Sol Energy 1979;22(3):225–8.
[56] 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.
[57] Campbell GS, Norman J. An introduction to environmental biophysics. Springer Science & Business Media; 2012.
[58] Majumdar N, Mathur B, Kaushik S. Prediction of direct solar radiation for low atmospheric turbidity. Sol Energy 1972;13(4):383–94.
[59] Badescu V. Verification of some very simple clear and cloudy sky models to evaluate global solar irradiance. Sol Energy 1997;61(4):251–64.
[60] Capderou M. Theoretical and experimental models solar atlas of Algeria (in French) Tome 1 and 2. Algeria: University Publications Office; 1987.
[61] Allen RG. Assessing integrity of weather data for reference evapotranspiration estimation. J. Irrig. Drain. Div. Eng. 1996;122(2):97–106.
[62] Davies J, McKay D. Estimating solar irradiance and components. Sol Energy 1982; 29(1):55–64.
[63] 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.
[64] Suckling P, Hay JE. Modelling direct, diffuse, and total solar radiation for cloudless days. Atmosphere 1976;14(4):298–308.
[65] Page J. Algorithms for the satel-light programme. Technical report for the Satel-Light programme; 1996.
[66] 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.
[67] 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.
[68] Ineichen P, Perez R. A new airmass independent formulation for the Linke turbidity coefficient. Sol Energy 2002;73(3):151–7.
[69] Dogniaux R, Lemoine M. Programme de calcul des éclairements solaires énergétiques et lumineux des surfaces orientées et inclinées: ciel serein et ciel couvert. Institut royal météorologique de Belgique; 1976.
[70] Ineichen P. A broadband simplified version of the Solis clear sky model. Sol Energy 2008;82(8):758–62.
[71] Ineichen P. High turbidity solis clear sky model: development and validation. Rem Sens 2018;10(3):435.
[72] De Brichambaut CP. Estimation des ressources énergétiques solaires en France. Association francaise pour l'etude et le developpement des applications de l'energie solaire; 1975.
[73] Ideriah F. A model for calculating direct and diffuse solar radiation. Sol Energy 1981;26(5):447–52.
[74] Bashahu M, Laplaze D. An atmospheric model for computing solar radiation. Renew Energy 1994;4(4):455–8.
[75] Yang K, Koike T. A general model to estimate hourly and daily solar radiation for hydrological studies. Water Resour Res 2005;41(10).
[76] Dai Q, Fang X. A simple model to predict solar radiation under clear sky conditions. Adv Space Res 2014;53(8):1239–45.
[77] 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.
[78] Paulescu M, Schlett Z. A simplified but accurate spectral solar irradiance model. Theor Appl Climatol 2003;75(3–4):203–12.
[79] 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.
[80] Atwater MA, Ball J. A numerical solar radiation model based on standard meteorological observations. Sol Energy 1978;21(3):163–70.
[81] King R, Buckius R. Direct solar transmittance for a clear sky. Sol Energy 1979;22 (3):297–301.
[82] 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.
[83] 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.
[84] 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.
[85] Iqbal M. An introduction to solar radiation. New York: Academic Press Inc.; 1983.
[86] 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.
[87] Kambezidis HD. The meteorological radiation model (MRM): advancements and applications. Modeling solar radiation at the earth's surface. Springer; 2008. p. 357–92.
[88] 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.
[89] Watt D. On the nature and distribution of solar radiation. Report for US DOE; 1978.
[90] Gueymard C. A two-band model for the calculation of clear sky solar irradiance, illuminance, and photosynthetically active radiation at the earth's surface. Sol Energy 1989;43(5):253–65.
[91] 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.
[92] 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.
[93] Gueymard CA. Clear-sky radiation models and aerosol effects. Solar resources mapping. Springer; 2019. p. 137–82.
[94] Yang D, Bright JM. Worldwide validation of 8 satellite-derived and reanalysis solar radiation products: a preliminary evaluation and overall metrics for hourly data over 27 years. Sol Energy 2020. https://doi.org/10.1016/j.solener.2020.04.016. In press.
[95] 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.
[96] Gueymard C. Temporal variability in direct and global irradiance at various time scales as affected by aerosols. Sol Energy 2012:1–10.
[97] 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.
[98] 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.
[99] Gmao Ges-Disc. MERRA-2 const–2d–asm–Nx: 2d, constants V5.12.4. 2018. Greenbelt [online], https://disc.gsfc.nasa.gov/datasets/.
[100] Rubel F, Kottek M. Observed and projected climate shifts 1901–2100 depicted by world maps of the Köppen-Geiger climate classification. Meteorol Z 2010;19(2): 135–41.
[101] Wrmc-Bsrn. Data retrievel via pangaea [online], http://bsrn.awi.de/data/data-retrieval-via-pangaea/; 2018.
[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] 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.
[104] Gschwind B, Wald L, Blanc P, Lefèvre M, Schroedter-Homscheidt M, Arola A. Improving the McClear model estimating the downwelling solar radiation at ground level in cloud-free conditions–McClear-v3. 2019.
[105] Blanc P, Espinar B, Geuder N, Gueymard C, Meyer R, Pitz-Paal R, Reinhardt B, Renné D, Sengupta M, Wald L, Wilbert S. Direct normal irradiance related definitions and applications: the circumsolar issue. Sol Energy 2014;110:561–77.
[106] Sun X, Xiaoyi Y, Wang P. A study of models combination for global clear sky irradiance models. In: Proceedings of the 47th IEEE photovoltaic secialist conference (PVSC47), virtual meeting; 2020. p. 297. 15th June–21st August. 0608070718.