Solar Energy 241 (2022) 525–537

Contents lists available at ScienceDirect

Solar Energy

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

Effects of spatial scale of atmospheric reanalysis data on clear-sky surface radiation modeling in tropical climates: A case study for Singapore

Xixi Sun [a,b], Dazhi Yang [c], Christian A. Gueymard [d], Jamie M. Bright [e,*], Peng Wang [f,*]

[a] School of Mathematical Sciences, Beihang University, Beijing, China
[b] Solar Energy Research Institute of Singapore (SERIS), National University of Singapore, Singapore
[c] School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin, Heilongjiang, China
[d] Solar Consulting Services, Colebrook, NH, USA
[e] UK Power Networks, London, United Kingdom
[f] School of Microelectronics, Beijing Advanced Innovation Center for Big Data and Brain Computing, Beihang University, Beijing, China

ARTICLE INFO

Keywords:
MERRA-2
Aerosol optical depth
Water vapor
Clear-sky radiation model
Inter-comparison
Spatial scale mismatch

ABSTRACT

Solar resource assessments most generally require atmospheric information, which is customarily acquired from gridded datasets. The spatial scale mismatch problem, i.e., the difference in spatial representativeness of gridded data and *in situ* measurements, therefore becomes relevant. This study examines how the gridded data used as inputs to clear-sky radiation models can affect their performance at urban scale. The tropical island of Singapore is selected for the case study. Aerosol optical depth at 550 nm (AOD550), Ångström exponent (AE), and precipitable water (PW) from both the MERRA-2 reanalysis and ground-based stations (AERONET and SuomiNet) are collected between 2013–2020. Firstly, it is found that, relatively to the AERONET ground truth, the bias in MERRA-2's AOD550 is more prominent than that in AE or PW. Next, the bias propagation from the gridded inputs (AOD550, AE, and PW) to clear-sky radiation predictions is explored using various models. The estimated clear-sky direct normal irradiance (DNIcs) is more sensitive to AOD550 variation than the clear-sky global horizontal irradiance (GHIcs). Six clear-sky radiation models, five of which accept MERRA-2 gridded inputs, are compared with each other, and with the *in situ* irradiance measurements recorded at 9 sites. The inter-model difference across Singapore is remarkably consistent because the whole island fits inside a single MERRA-2 grid cell. Under high-AOD550 situations, however, the inter-model deviation becomes large for both GHIcs and DNIcs. The conventional model-versus-measurement comparison shows that each model achieves very different site-to-site performance, largely because the spatially-averaged inputs cannot fully represent the micro-climatic variability. Relatively speaking, no clear-sky radiation model significantly outperforms its peers. The simple MAC2 model and the empirical (locally derived) Yang GHIcs-only model are recommended for Singapore.

1. Spatial scale mismatch between the gridded and *in situ* atmospheric variables

Information on certain atmospheric variables, such as the aerosol optical depth (AOD) or precipitable water (PW), is essential for a wide range of environmental sciences, energy, agriculture and forestry applications. The ideal way to obtain such information is through direct *in situ* measurements. However, limited by funding, laboratory support, maintenance personnel, or equipment, global or regional ground-observing networks are scarce. In the case of AOD, for instance, AERONET (AErosol RObotic NETwork) is the world's largest network providing long-term and globally distributed ground-truth spectral AOD based on specialized measurements made with multi-wavelength sunphotometers (Holben et al., 1998). Nevertheless, despite there being hundreds of AERONET stations worldwide, some regions are still very poorly represented. Moreover, many stations are only active for a short period (e.g., field campaigns), or suffer from extended periods of missing data. For all these reasons, it is customary in solar resource studies to instead acquire estimates of radiatively-active atmospheric variables from satellite-based or reanalysis databases (Gueymard, 2019).

In stark contrast to ground-based networks, satellite-based and reanalysis databases provide *gridded* atmospheric data. Hence, their spatial coverage is extensive and either nearly global (with satellites) or global (with reanalyses). Satellite-based AOD products are derived from

* Corresponding authors.
E-mail addresses: jamiebright1@gmail.com (J.M. Bright), wang.peng@buaa.edu.cn (P. Wang).

https://doi.org/10.1016/j.solener.2022.06.001
Received 3 January 2022; Received in revised form 4 May 2022; Accepted 1 June 2022
Available online 29 June 2022

data collected by instruments onboard polar orbiters, such as the Moderate Resolution Imaging Spectroradiometer (MODIS) onboard the Terra and Aqua satellites (Remer et al., 2005). Since polar orbiters revolve around the Earth in longitudinal direction, they only provide a complete scan of the Earth every several days, which implies low temporal resolution. Additionally, the remote-sensed records include many spatio-temporal gaps because of the incapacity of the retrieval algorithms over bright surfaces, under cloudy skies, or misclassification of heavy pollution as clouds or vice versa. These issues affect regions worldwide, and thus creates an insurmountable problem for many applications that require continuous time series. From that standpoint, an alternative way to arrive at gridded atmospheric variables with no missing data is by running dynamic weather models combined with chemical transport models. In particular, retrospective simulations of the atmospheric state, known as reanalyses, are commonly used. Two such reanalyses that provide detailed aerosol data are used in practice (Casagrande et al., 2021; Witthuhn et al., 2021; Ruiz-Arias et al., 2019; Boraiy et al., 2017), namely NASA's Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) (Gelaro et al., 2017) and ECMWF's Copernicus Atmosphere Monitoring Service (CAMS) (Benedetti et al., 2009).[1]

Since the accuracy of gridded products is generally lower than that of ground-based measurements, validation of gridded products is logically attractive and thus has become commonplace. For instance, Shi et al. (2019) compared AOD from MODIS, MERRAero (the precursor of MERRA-2), and MERRA-2 with ground-truth AERONET data from more than 400 world stations; Bright and Gueymard (2019) validated the MODIS Aqua and Terra AOD products at 452 AERONET stations; and Gueymard and Yang (2020) validated AOD and Ångström exponent (AE) from the CAMS and MERRA-2 against 793 AERONET stations over all continents and all climate classes. In short, these studies showed that the reanalysis MERRA-2 AOD product, although it tends to underestimate AOD values, generally outperforms CAMS, and is comparable to MODIS.[2] Whereas the above-mentioned studies aim at quantifying the bias and accuracy of gridded aerosol products, studies that focus on other variables, such as PW, are also available (e.g., Huang et al., 2022; Wang et al., 2020). It is typically found that the PW predictions from reanalysis products in general, and of MERRA-2 in particular, are of good accuracy overall.

A central goal of most validation studies is to estimate the level of confidence with which gridded atmospheric products can replace ground-based measurements in various applications. A fundamental challenge, however, is the spatial scale mismatch: Whereas each cell (or pixel) of a gridded product depicts the average value of the atmospheric variable within that cell, ground-based measurements are only representative of a small area within the neighborhood of the station. In particular, the spatial scale mismatch refers to the difference in the spatial representativeness (Wu et al., 2019; Molero et al., 2018). Spatial scale mismatch is related mainly to the fine-scale topographical features within a grid cell. The terrain across a spatial domain, including mountains, rivers, coastlines, forests, or urban areas, can affect the atmospheric composition aloft. For instance, the vertical water vapor column (most generally measured in terms of PW) can be denser over a lake than over the dry vegetation around it or over high-elevation terrain. Similarly, AOD normally decreases with elevation, but increases near sources of pollution. Hence, AOD can be expected to be larger over a dense urban area that includes heavy traffic and industries, compared to a low-density suburb. These effects can be entangled over a small island with variable urban density, such as Singapore, which is ringed by large-scale industrial complexes and a tropical rain forest. In particular, both PW and AOD are expected to present sharp spatial transitions over much shorter distances than the typical size of a reanalysis cell.

Even though the issue of spatial scale mismatch has been known for decades (Blöschl and Sivapalan, 1995), related studies are exceedingly rare, which can be attributed wholly to the lack of dense ground-based monitoring networks. Indeed, in order to quantify the spatial scale mismatch, one must start with a good collection of ground-based stations within a cell, such that the difference between the measurements from individual stations and their average value can be evaluated. Particularly true is the case of solar irradiance: High-quality radiometric stations, due to their cost and maintenance requirements, are usually much fewer than stations measuring other atmospheric variables. Thus, it is extremely rare that a network of densely distributed radiometers exists within the microscale or mesoscale area under scrutiny. This explains the extreme paucity of studies dedicated to the spatial scale mismatch of solar irradiance (Yang, 2020; Macke et al., 2017; Madhavan et al., 2017; Gao et al., 1998; Barnett et al., 1998). Further studies are obviously needed; in this context, the present investigation is meant as a preliminary contribution toward a better understanding of the spatio-temporal variability of the solar radiation field over Singapore.

2. Spatial scale mismatch in clear-sky radiation modeling

Clear-sky radiation models describe the solar irradiance reaching the Earth's surface under a cloudless atmosphere. They can be either physical or empirical, where the former type normally has a higher accuracy than the latter, with less geographic or climatic limitations. The principle of physical clear-sky radiation modeling is to solve the radiative transfer equation considering various atmospheric constituents. Since the clear-sky irradiance is required in virtually all solar energy applications, the quality of the model inputs and their impact on its output accuracy must be regarded as a key consideration.

To ensure the spatial availability and temporal representativeness of the clear-sky model's irradiance output, it is essential to rely on a source of input atmospheric data with global and decades-long availability. Reanalysis, therefore, becomes the obvious choice. For instance, Ruiz-Arias and Gueymard (2018) adopted the MACC reanalysis (precursor of CAMS) to evaluate 15 clear-sky radiation models by intercomparing their outputs. Previous investigations carried out by the present authors have also demonstrated the feasibility and efficiency of using global reanalyses databases, such as MERRA-2, to provide the necessary atmospheric inputs to clear-sky models in climate-specific or global validation studies (Sun et al., 2019, 2021). An important finding of those studies was that estimates of global and direct irradiance were satisfactory when using selected clear-sky models along with MERRA-2 inputs. The latter can be accessed with the `irradpy` package (Bright et al., 2020a) in Python, for instance.

One important difficulty, however, is that reanalysis products are gridded, and thus typically suffer from the spatial scale mismatch, whose effects propagate into the estimated clear-sky irradiance. In parallel, when those estimates are evaluated, ground-based irradiance observations are almost always used as the reference. Hence, the validation results depend on where the ground station is situated within the local reanalysis cell. In particular, if the geography and/or microclimate of the cell are diverse, i.e., the cell has strong spatial inhomogeneity, the representativeness of the results cannot be ensured, and the validity of the conclusion can thus be affected. In order to better

[1] The temporal resolution of the MERRA-2 and CAMS products are hourly and 3-hourly, respectively. Their spatial resolution—i.e., 0.625° × 0.5° for MERRA-2 and 0.5° × 0.5° for CAMS, in longitude and latitude, is substantially coarser than that of gridded satellite-based products, which is typically 0.05° × 0.05° or better under cloudy conditions. Under cloudless situations (such as those investigated here), however, the underlying resolution of the atmospheric product (MERRA-2 in the present case) used to predict the clear-sky irradiance is what matters.

[2] The various MODIS AOD products exhibit heterogeneity because of variance in sensor characteristics and performance, observation times, algorithms, or spatial resolution. For instance, the MODIS-Aqua product generally has higher accuracy than its MODIS-Terra counterpart.

understand the spatial representativeness of clear-sky modeled results powered by gridded products, one must proceed from the model's inputs.

Based on the considerations above and on previous knowledge related to the sensitivity of surface irradiance to various atmospheric quantities (Gueymard, 2003, 2012b), it is desirable to evaluate the spatial variability of AOD and PW, and to evaluate their impact on clear-sky radiation modeling. Without loss of generality, the MERRA-2 reanalysis is selected here as the source of input. As mentioned earlier, the ideal situation for this kind of validation study when conducted over an inhomegeneous environment would be to have numerous high-quality radiometric stations within each MERRA-2 cell. In the present case, this requirement is fulfilled, at least in great part, by the sensor network installed by the Solar Energy Research Institute of Singapore (SERIS) over the tropical island of Singapore. This is a 9-station network equipped with Delta-T SPN1 pyranometers, which have the unique capability of simultaneously sensing both the global horizontal irradiance (GHI) and diffuse horizontal irradiance (DIF) through the help of a sophisticated computer-generated shading pattern within their glass dome. Knowledge of both GHI and DIF provides a way to derive the direct normal irradiance (DNI).

Singapore is small enough to fit inside a single MERRA-2 cell. Hence, any clear-sky radiation model under scrutiny adopts the same MERRA-2 inputs at all stations of the network. Consequently, the inputs to all models are identical at any point in time (except for minute differences caused by latitude variance, which is negligible), and any spatial variability in site-to-site modeled irradiance performance must be entirely attributed (within the uncertainty margin of measurements) to either the micro-scale spatial variance in the reference irradiance datasets or, less likely, the variance in intrinsic model performance solely caused by error propagation.

In short, the aim of the study is to determine the impact of using gridded products, i.e., cell-averaged input variables, in sub-cell clear-sky GHI and DNI modeling. If this impact ends up being small enough, MERRA-2 would be confirmed as an appropriate source of atmospheric data for local solar resource assessments. (Obviously, the availability of accurate AOD and PW at much higher spatio-temporal resolution would still be preferable.) Instead of relying on a single clear-sky radiation model, multiple models are employed here to increase confidence in the arrived conclusions. While a critical byproduct of this study is the qualitative and quantitative comparison of clear-sky models based on the data and information available in Singapore, another byproduct refers to the validation of three key quantities provided by MERRA-2: AOD550, AE, and PW, which is facilitated by high-quality ground-based measurements in Singapore.

The structure of the remaining part of the paper is as follows. Section 3 describes the ground-based measurements of irradiance and other atmospheric variables available in Singapore. Section 4 briefly introduces the selected clear-sky radiation models and their inputs. Section 5 evaluates the accuracy of MERRA-2's AOD and PW products, the spatial variation of aerosols and solar radiation, the impact of AOD's bias on clear-sky irradiance estimates, and the prediction discrepancies between these models. Finally, Section 6 summarizes the work.

3. Data

3.1. Ground irradiance data from SERIS

SERIS maintains 25 PV monitoring stations across the island, among which nine of them, as listed in Table 1 and indicated by purple circles in Fig. 1, collect both GHI and DIF measurements at 1-min resolution. Delta-T SPN1 pyranometers are used there, whereas other stations only collect GHI with silicon reference cells. Consequently, at the nine stations, DNI can be derived from GHI and DIF using the closure equation:

$$\mathrm{GHI} = \mathrm{DNI} \times \cos\theta_z + \mathrm{DIF}. \quad (1)$$

Table 1
Site information about the 9 ground stations in Singapore. East longitudes are positive. The station elevation is represented by h in meters above mean sea level.

#	Site name	Lat. [°]	Lon. [°]	h [m]
1	NE @ HDB Punggol	1.3993	103.9091	55
2	SW @ SERIS	1.3005	103.7710	63
3	N @ Republic Poly	1.4439	103.7843	40
4	S @ Quayside Condo	1.2491	103.8414	36
5	E @ AMAT	1.3527	103.9653	45
6	NW @ Lim Chu Kang	1.4458	103.7108	40
7	SE @ HDB Marine Parade	1.3025	103.9123	60
8	W @ CTO	1.3551	103.6925	57
9	C @ Bosch	1.3477	103.8413	52

where θ_z is the solar zenith angle. The period of record from May 2013 to December 2020 is adopted in this study. The quality-control procedure described in Long and Shi (2006) and Sun et al. (2019) is applied to all data. Besides, all low-sun situations when $\theta_z > 85°$ are removed from all subsequent analyses.

It should be acknowledged that the SERIS meteorological network, just like many other local (and often proprietary) radiometric networks, does not offer nearly the same high-quality irradiance measurements as BSRN (Driemel et al., 2018), for instance. Most critically, the SPN1 instrument is known for its limited performance (Vuilleumier et al., 2017; Badosa et al., 2014; Möllenkamp et al., 2020). However, the SERIS ground database used here is the best that one can get in Singapore for solar radiation studies. All radiation sensors are identical in terms of equipment model, maintenance routine, and calibration method. This source of irradiance data is deemed suitable for the objectives of this study, even though the model performance could change if alternatively undertaken against higher-quality reference measurements (which do not yet exist).

3.2. Atmospheric data

3.2.1. MERRA-2 product

Following previous work Sun et al. (2019, 2021), and also because of its generally superior performance (as compared to CAMS) with regard to AOD550 (Gueymard and Yang, 2020), the MERRA-2 reanalysis database is adopted in this study. Given MERRA-2's spatial resolution, the entire administrative area of Singapore is contained in just a single cell, which is centered on latitude 1.5°, longitude 103.75°, as shown in Fig. 1. Five of the six clear-sky radiation models evaluated in this study (see Section 4) use the *same* atmospheric data extracted from this MERRA-2 grid cell. The target cell is mostly land, but is also affected by the ocean. Note that the atmospheric data is cell-averaged, which means it is affected by the constitution of the grid cell. The hourly atmospheric information extracted from the MERRA-2 database over the period January 2013 to December 2020 includes AOD550, AE, PW, aerosol scattering optical depth, surface albedo, ozone, and pressure (see Sun et al., 2021, for method details). The aerosol scattering optical depth is divided by AOD550 to obtain the aerosol single-scattering albedo.

3.2.2. AERONET data

AERONET provides AOD at various wavelengths, PW, and Ångström exponent products with relatively high temporal resolution. Version 3 of the processing algorithm derives aerosol and water vapor data at various quality levels. At the highest quality level ("Level 2.0") used here exclusively, all data points are cloud-screened and quality-assured (Holben et al., 1998; Giles et al., 2019). Because of the heavy cloudiness in a tropical environment such as Singapore, there are relatively few valid data points compared to sunnier areas, and most are either in early morning or late afternoon, when GHI is relatively low. As shown in Fig. 1, two AERONET stations (separated by approximately 10 km) provide V3-L2 products in Singapore. Their information appears

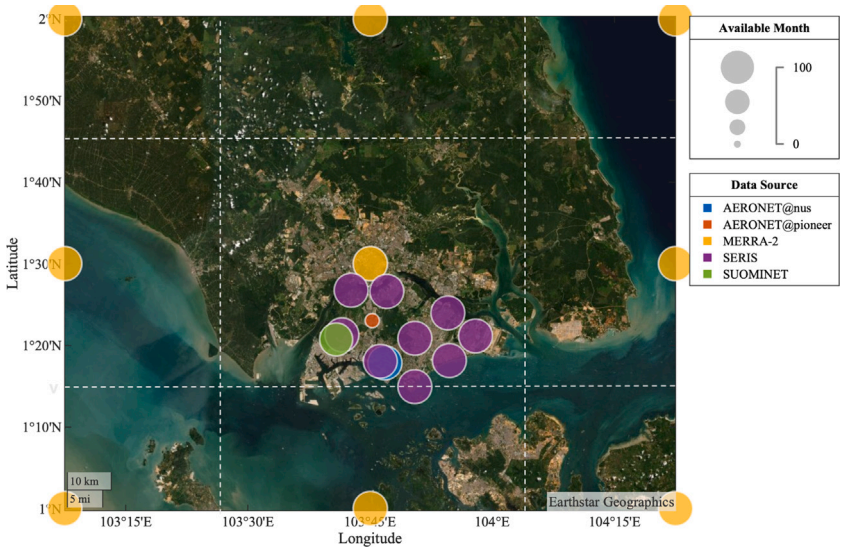

Fig. 1. Information on all the data sources used in this study, arranged by color code. The number of available months that have usable data between 2013 and 2020 is denoted by the circle size. Purple dots denote the 9 SERIS ground stations, with 93 months of data each. The blue dot denotes the AERONET station located at the National University of Singapore (NUS), with 94 months of data. The orange dot denotes the AERONET station at Pioneer with 11 data months. The green dot denotes a SuomiNet water vapor observation station at the Nanyang Technological University (NTU), with 88 data months. The yellow circles denote the 9 MERRA-2 nodes in and around Singapore, with 96 data months. The dashed white lines mark the MERRA-2 grid cell borders. Note that all ground stations fit inside the central MERRA-2 grid cell. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Site information on two AERONET stations located in Singapore. The original station names may cause misunderstanding, hence new simpler names (used throughout here) are provided within brackets. The station elevation is represented by h in meters above mean sea level. Start and end operational times are listed.

#	Site name	Lat. [°]	Lon. [°]	h [m]	Start	End
1	Singapore (NUS)	1.2977	103.7804	30	11/2006	–
2	Pioneer_JC (Pioneer)	1.3843	103.7545	25	06/2017	04/2018

in Table 2. Station NUS has a longer observational period (94 months) than station Pioneer (11 months). AOD550 is not observed directly by AERONET, so it is derived here from AOD observations at 440, 500, 675 and 870 nm, as described elsewhere (Gueymard and Yang, 2020). PW is the total columnar water vapor derived from the 935-nm sunphotometer channel. PW and AOD550 data points are aggregated into hourly resolution for intervals with more than one point per hour. In accord with the existing literature, AOD550 from AERONET is regarded as ground-truth in further comparisons with the MERRA-2 model estimates.

3.2.3. SuomiNet Data

SuomiNet is a university-based Global Positioning System meteorological (GPSmet) network to derive near real-time atmospheric PW measurements from high-precision surface meteorological observations (Ware et al., 2000). The accuracy of GPSmet-sensed PW is reported as better than 0.2 cm (Ware et al., 2000). To the difference of AERONET, which can only report valid PW measurements whenever the sun's disc is not obscured and is above a minimum elevation in the sky, the GPSmet retrieval algorithm is insensitive to sky or sun conditions. This means that, under a very cloudy tropical climate such as Singapore and during a specific period of, e.g., one year, the number of valid PW data points from GPSmet will be considerably larger than that from AERONET.

There is only one SuomiNet station (code: NTUS) in Singapore. Its latitude and longitude are 1.3458° and 103.6800°, respectively (see Fig. 1). SuomiNet reports PW twice an hour, aggregated here to hourly resolution. The available data is between May 2012 and April 2020, for a total of 88 valid data months.

4. Clear-sky radiation models

Since clear-sky models have been validated extensively and comprehensively in the literature, the present study does not attempt to repeat that sort of process for Singapore, because most of the knowledge gathered through the previous validations can be regarded as generally applicable for worldwide locations. Without loss of generality, six representative clear-sky radiation models are strategically selected based on their construct, popularity, and past performance. Table 3 provides a summary of the selected models and their required input variables. The model number is assigned in descending order by the number of input variables and publication year. (The MERRA-2 estimates of clear-sky GHI (GHIcs) are not considered here because they are known to be affected by an incorrect sun position algorithm (Gueymard, 2019), which introduces systematic errors.)

It should be noted that among these candidates, only the Yang model is empirical and thus does not rely on any atmospheric input. However, its model coefficients are localized, i.e., specifically obtained from measured irradiance in Singapore. Hence, it is not really independent from the reference data used to validate it here. Besides, the Yang model provides GHIcs predictions only, whereas all other models, operated here with appropriate inputs from MERRA-2, can provide predictions of both GHIcs and clear-sky DNI (DNIcs).

5. Results and discussion

With the data and models described in previous sections, the MERRA-2 cell that includes Singapore can be investigated thoroughly from various standpoints. An AOD product comparison between MERRA-2 and AERONET is presented in Section 5.1, whereas a PW comparison between MERRA-2, SuomiNet and AERONET appears in Section 5.2. Potential spatial variations in aerosols and surface solar radiation are discussed in Section 5.3. The error propagation from aerosols and water vapor inputs to clear-sky irradiance modeled predictions is discussed in Section 5.4. Finally, a clear-sky model inter-comparison, modeled-vs-measurement comparison, and the most suitable models for Singapore are discussed in Section 5.5.

Following Gueymard (2014a), five performance statistics are used in the following evaluations; they are: mean bias error (MBE), normalized mean bias error (nMBE), root mean-squared error (RMSE),

Table 3
Summary of six clear-sky radiation models along with their input variables: specified solar constant (E_{sc}), zenith angle θ_z, local barometric pressure p [mb], ground albedo R_G, total ozone amount u_{O_3} [atm-cm], total nitrogen dioxide amount u_{NO_2} [atm-cm], total precipitable water vapor u_{H_2O} [cm], aerosol optical depth at 550 nm (τ), Ångström exponent α, Ångström turbidity coefficient β, and aerosol single-scattering albedo ϖ. The model number is given in descending order by the number of input variables.

#	Clear-sky model	Citation	E_{sc}	θ_z	p	R_G	u_{O_3}	u_{NO_2}	u_{H_2O}	τ	α	β	ϖ	Total
1	YANG	Yang et al. (2014)	•	•										2
2	INEICHEN2018	Ineichen (2018)	•	•	•				•		•			5
3	MAC2	Sun et al. (2021)	•	•	•	•			•			•	•	7
4	MRMv6	Kambezidis et al. (2017)	•	•	•	•			•			•	•	7
5	BIRD	Bird and Hulstrom (1981)	•	•	•	•	•		•			•	•	8
6	REST2v5	Gueymard (2008)	•	•	•	•	•	•	•			•	•	10

normalized root mean-squared error (nRMSE), and mean absolute error (MAE). With N prediction–observation pairs, MBE is calculated as $N^{-1}\sum_{i=1}^{N}(P_i - O_i)$, and RMSE and MAE can be defined analogously. In the cases of nMBE and nRMSE, the mean observation is used for normalization, i.e., nMBE = MBE/$\mathbb{E}(O)$ and nRMSE = RMSE/$\mathbb{E}(O)$.

5.1. Validation of MERRA-2's aerosol products

Two MERRA-2 aerosol quantities, AOD550 and AE, are compared to AERONET ground-truth at two stations. The Pioneer station has 894 valid hours of AOD550 and AE, whereas the NUS station, which was established at a much earlier date, has a total of 8929 h of valid data. A further extraction is applied to make sure all aerosol products are compared using the same hours of valid data. The number of common valid hours is 628. This avoids the detrimental impact of temporal variations in disparate time periods, but considerably lessens the usable sample size.

Results of the four statistical metrics described above are provided in Table 4. It is found that the two closely located AERONET stations exhibit large differences in metric values, which is somehow unexpected. MERRA-2's AOD550 performs better at Pioneer than NUS with regard to all four error metrics. NUS-observed AOD550 not only has a high mean value of 0.452, but also a high standard variation of 0.407. The nRMSE at NUS is also two times higher than that of Pioneer. These results are confirmed in Fig. 2, showing a time series of hourly AOD550 for each site. Extremely high AOD550 values and discrepancies occurred in July 2017, when Singapore experienced a haze episode caused by smoke. Such episodes occur relatively frequently in Singapore (Hansen et al., 2019), and are well aligned with the Indonesia burning season. This usually peaks from July to October, when farmers in palm oil, pulp, or paper plantations use the slash-and-burn method to clear vegetation residues, which often spins out of control quickly. Moreover, these haze episodes are known to affect the performance of solar systems (Nobre et al., 2016). During such haze episodes, strong spatial variations in smoke density can exist. Furthermore, the AERONET cloud screening algorithm might be less discriminant during such high-AOD situations (Giles et al., 2019). These facts can explain the discrepancies noted between the two stations. A complementary explanation is that NUS is closer to the highway and factories along the west coast, and, therefore, is subject to more air pollution.

To take full advantage of all the available data, Fig. 3 is plotted with all data from 2013 and 2020. The first plot in Fig. 3 shows the histograms and kernel density estimates (KDEs) of AOD550 for Pioneer, NUS, and MERRA-2, which are both representations of the probability density function (PDF). It is found that the MERRA-2 AOD550 values are considerably underestimated and sparsely distributed. Hence, it can be anticipated that the GHIcs—and even more so, DNIcs—predictions made with models using MERRA-2 inputs will be too high.

MERRA-2's AE is relatively more stable spatially than AOD550. More specifically, the nRMSE for AE is only about half that of AOD550. The values of the four error metrics of MERRA-2's AE at the two

Table 4
Summary of performance metrics of AOD550 and AE for MERRA-2 compared with two ground stations of AERONET. The number, N, of data points for each station is controlled to be the same for absolute fairness. The mean and standard deviation (SD) for each AERONET station as well as the corresponding range of MERRA-2 are displayed. The third row specifies the unit of the variable, if any.

Site	N	AOD550					
		Mean±SD AERONET	Mean±SD MERRA-2	RMSE	nRMSE %	MBE	MAE
Pioneer	628	0.256 ± 0.122	0.212 ± 0.083	0.106	41.584	−0.044	0.076
NUS	628	0.452 ± 0.407	0.212 ± 0.083	0.484	107.16	−0.240	0.281
		AE					
		Mean±SD AERONET	Mean±SD MERRA-2	RMSE	nRMSE %	MBE	MAE
Pioneer	628	1.373 ± 0.281	1.223 ± 0.221	0.348	25.354	−0.151	0.285
NUS	628	1.415 ± 0.280	1.223 ± 0.221	0.382	26.970	−0.192	0.308

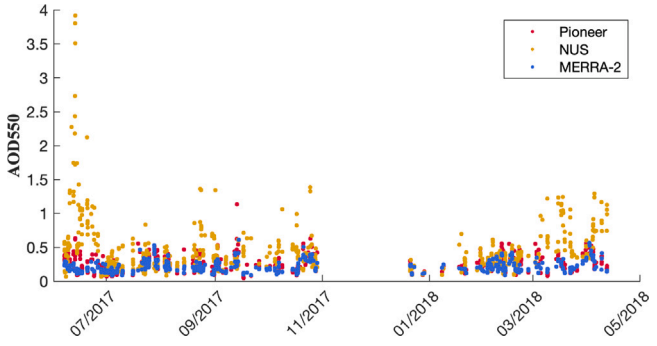

Fig. 2. Scatter plot of AOD550 from each data source. Only the 628 common hourly data points are shown. Red, yellow, and blue denote Pioneer, NUS and MERRA-2, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

AERONET stations are similar. As shown in the second plot of Fig. 3, the pdf plots of AE are adjacent. In comparison, MERRA-2 AE values are slightly underestimated.

In summary, the spatial mismatch between MERRA-2 and in-situ aerosol observations is not of the same magnitude for each variable. Whereas MERRA-2's AE can be a good replacement for actual observations over Singapore, a different picture emerges for AOD550 because of its strong underestimation, at least over some parts of Singapore.

5.2. Precipitable water comparison between MERRA-2, AERONET and SuomiNet

As with the aerosol data discussed above, precipitable water observations from AERONET (Pioneer and NUS stations) and SuomiNet (NTUS) are aggregated into hourly periods. The number of total available hourly data points for Pioneer, NUS and NTUS are 894, 8929, and 30,086, respectively. MERRA-2's PW is extracted separately to match

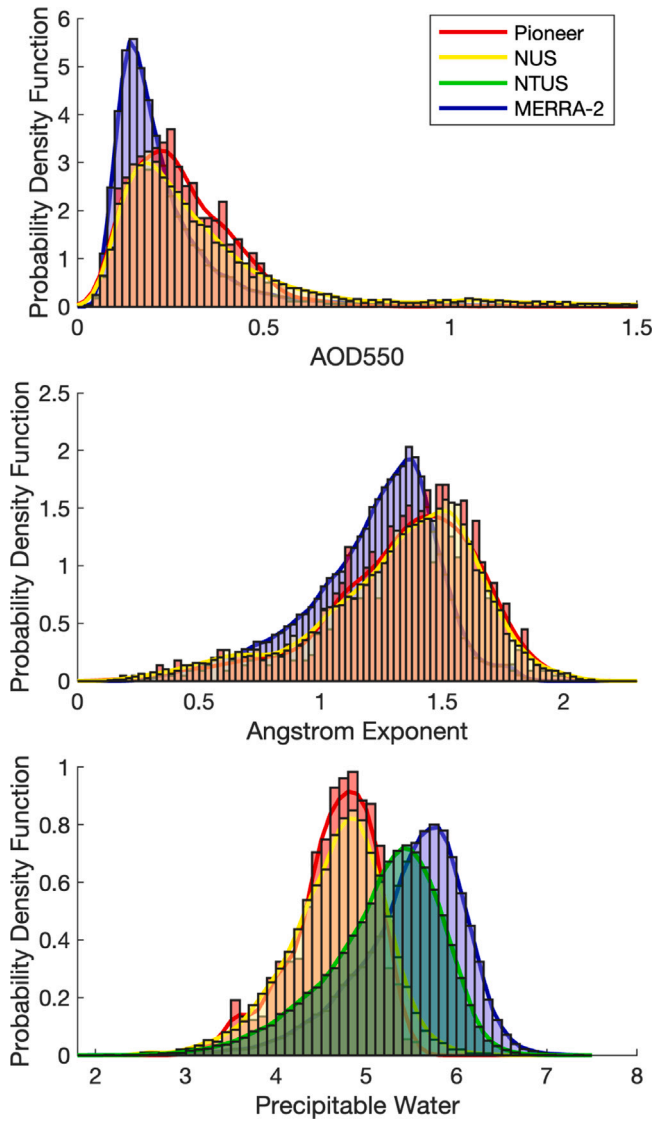

Fig. 3. Histograms and kernel density estimates (KDEs) of AOD550, AE and PW from each data source. The data used here are not extracted to exclude temporal variation, i.e., all available data between 2013 and 2020. Red, yellow, blue, and green denote histograms and KDEs for Pioneer, NUS, MERRA-2 and NTUS, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the corresponding ground observations, and is then compared with observations from Pioneer, NUS and NTUS. Note that the PW results are not from the same number of data points. Considering there are only 329 common hourly data points for the three ground stations, it is not representative of each data source and constitutes a huge reduction in the data sample size.

The performance metrics appear in Table 5 and show that MERRA-2's PW is somewhat overestimated as compared to both sources of ground observations. The differences being relatively small and PW being relatively high, however, the impact of using MERRA-2 PW data on the modeling of GHIcs and DNIcs should be small (Gueymard, 2014b). The lowest mean PW is recorded at Pioneer, possibly because of its interior location, although a sampling effect caused by the short observational period is likely too.

As shown in the bottom plot of Fig. 3, the PDF curve for PW obtained from MERRA-2 is closer to that from NTUS than to those from the two AERONET stations. Although the latter is in a dryer and hotter area than the former (see Fig. 4), a likely explanation is that either the AERONET retrieval algorithm underestimates PW in humid conditions, or conversely that the SuomiNet algorithm overestimates PW, which in turn would impact the MERRA-2 results since the reanalysis assimilates GPSmet data (but not AERONET-PW data).

PW is a direct function of temperature and relative humidity. For further reference, a map with the yearly average temperature and relative humidity measured at ten ground stations in Singapore is presented in Fig. 4. NUS, Pioneer, and NTUS are marked with black stars. The map indicates significant spatial variance in both temperature and relative humidity, which suggests an induced spatial variability in PW. In any case, for solar radiation modeling purposes, using MERRA-2 PW data as a spatial average for the whole Singapore appears acceptable.

5.3. Spatial variation of aerosols and radiation

Singapore is known as the *Garden City* with a 46.5% green cover. Nevertheless, the city also harbors chemical industries, as well as one of the world's three largest oil refineries, located southwest of the city. Many factories are concentrated all along the northern and western edges. Additionally, high-traffic ports and waterways exist along the south shore. One of the world's most active airports, Changi Airport, is also located to the east.

As mentioned above, Singapore is also likely to regularly suffer from haze caused by vegetation burning and forest fires from Indonesia or Malaysia during certain periods, which could cause extremely high AOD episodes. All these features make Singapore a suitable urban area to investigate the spatial variability in aerosols and solar radiation.

Through Fig. 4, it is found that the mean annual relative humidity (RH) varies 10% while the mean annual ambient temperature varies 1.6 °C among the ten ground stations. A similar figure, but for the number of clear-sky observations and average radiation between 2013 and 2020, is presented in Fig. 5 to evaluate the spatial consistency of the area's solar energy potential. The northwest and eastern regions have fewer clear-sky periods comparatively to other areas as observed with the SERIS ground sensors. Overall, despite the generally dense cloudiness, Singapore has a relatively good solar energy potential at around 1600 kWh m^{-2} per year in terms of GHI.[3] As could be expected, the DNI resource is much less, around 900 kWh m^{-2} per year.

The underlying cause for the observed spatial GHI variability of about ±5% is to be found in that in AOD, PW, or both. The AOD spatial variability can be asserted from the comparison results in Section 5.1. When comparing AOD550 and AE from MERRA-2 with AERONET, it is found that AOD550 suffers higher spatial inhomogeneity than AE. The mean AOD550 can vary up to nearly 70% over a 10-km distance, whereas the mean AE only varies 3% (Table 4). This suggests that AOD550 data would be needed at higher spatial resolution for more precise solar radiation modeling over Singapore.

5.4. Effects of input error propagation on clear-sky radiation model predictions

The sensitivity of clear-sky irradiance predictions on AOD550 is investigated here by applying a variable-controlling method. This approach focuses on both GHIcs and DNIcs, and differs from a previous effort (Gueymard, 2012b), which was focused on DNIcs only. The Ineichen2018 model in Table 3 is chosen here as the study case for its simplicity and its limited number of inputs (five). The default rural aerosol type is adopted here, per the author's recommendation (Ineichen, 2018), even though Singapore's atmosphere is likely to also contain maritime, urban, and smoke aerosols. The model derives GHIcs and DNIcs, noted G_{hc} and B_{nc} here, based on the Bouguer–Lambert–Beer attenuation relationship:

$$G_{hc} = E_{\text{ext}} \times \cos\theta_z \times T_g \tag{2}$$

[3] https://globalsolaratlas.info/map.

Table 5
Summary of performance metrics of MERRA-2's PW compared with ground stations (Pioneer and NUS) of AERONET and NTUS from SuomiNet. MERRA-2 data is available between January 2013 and December 2020. *N* indicates the number of available data points for each station. The mean and standard deviation (SD) for each ground observing station are displayed. The second row specifies the unit of the variable, if any.

Site	*N*	Period	Mean±SD cm	RMSE cm	nRMSE %	MBE cm	MAE cm
Pioneer	894	201706–201804	4.635 ± 0.457	0.609	13.136	0.528	0.541
NUS	8929	201302–202011	4.668 ± 0.540	0.662	14.162	0.539	0.579
NTUS	30086	201301–202004	5.180 ± 0.641	0.395	7.613	0.208	0.318

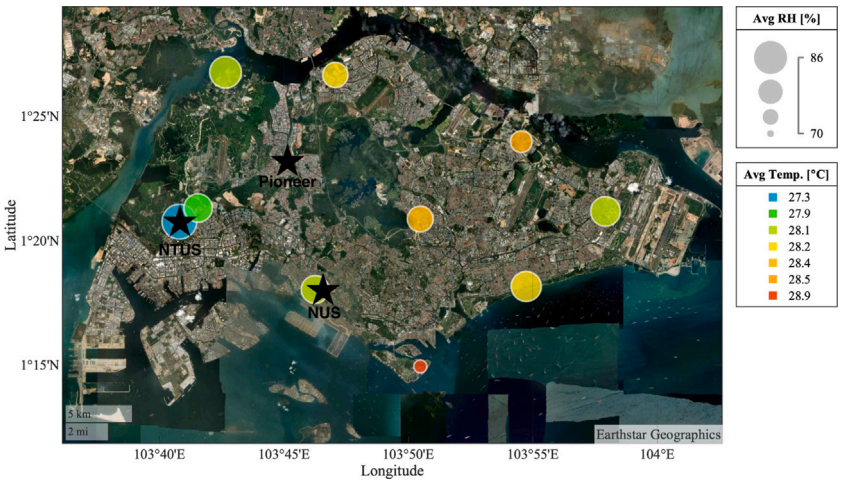

Fig. 4. Singapore map marked with nine SERIS ground meteorological stations and one SuomiNet ground station (in blue color). The sizes and colors of the solid circles denote the average relative humidity and ambient temperature between 2013 and 2020, respectively. The locations of AERONET Pioneer, NUS and SuomiNet NTUS are marked with dark stars. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

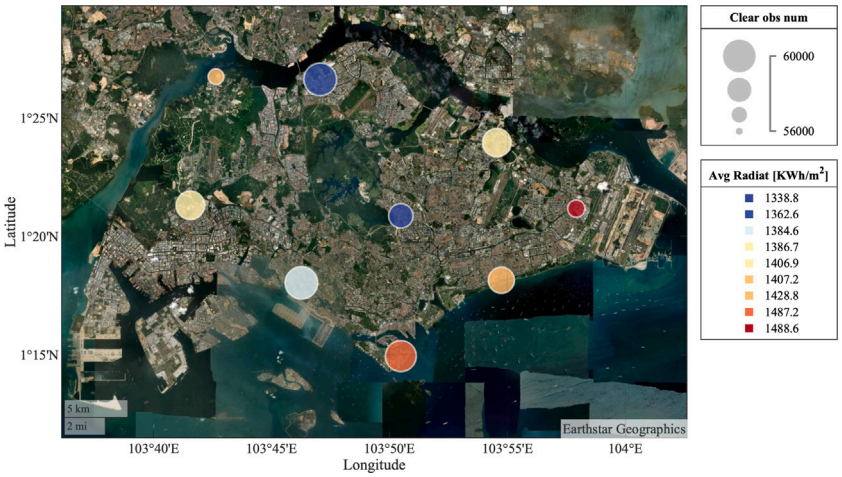

Fig. 5. Singapore map marked with nine ground radiometry stations. The sizes and colors of the solid circles denote the number of clear-sky observations and average radiation between 2013 and 2020, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$B_{nc} = E_{\text{ext}} \times T_b \tag{3}$$

where

$$E_{\text{ext}} = E_{\text{sc}}/R^2 \tag{4}$$

$$T_g = e^{-\tau_g/\cos^g \theta_z} \tag{5}$$

$$T_b = e^{-\tau_b/\cos^b \theta_z} \tag{6}$$

where T_g and T_b are the global and direct atmospheric transmittances, τ_g and τ_b are the extinction coefficients for global and direct irradiance, g and b are the corresponding exponent coefficients, E_{ext} is the normal-incident extraterrestrial irradiance, E_{sc} = 1361.1 W m^{-2} is the solar constant (Gueymard, 2018), R is the Sun–Earth distance relative to the astronomical unit, and θ_z is the solar zenith angle. R and θ_z are calculated via the SG2 algorithm (Blanc and Wald, 2012), wherever needed.

AOD550 and PW are selected as the control variables here, while the third atmospheric variable, local barometric pressure, is set to the annual mean value at Singapore, 1006.6 mb, as obtained from MERRA-2 between 2013 and 2020. The pressure is fixed here at a constant because of its negligible impact on irradiance predictions (Gueymard, 2003) and its small spatio-temporal variability over Singapore. The solar zenith angle θ_z is varied between 0–85° in steps of 0.1°. AOD550 being in a range of 0.05–2.0 over Singapore, intervals of 0.05 are used to capture the sensitivity of irradiance to AOD550. Similarly, PW is varied between 4–7 cm in steps of 1 cm. The main part of this

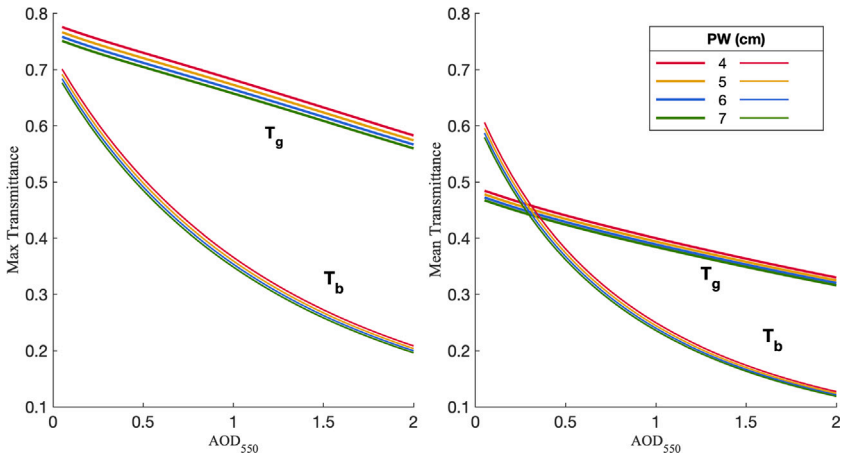

Fig. 6. Analysis of the sensitivity of the global transmittance T_g and direct transmittance T_b on AOD550 and PW. The left plot is for the maximum global and direct transmittances during an ideal clear day for a large range of AOD550 and PW values. The maximum transmittance is achieved at solar noon. The right plot displays the mean daily global and direct transmittances. Thicker lines are used to denote T_g.

analysis is to study the effect of varying AOD550 and PW on the global transmittance, T_g, and direct transmittance, T_b. This is summarized in Fig. 6, where the two subplots indicate that AOD550 has substantially more impact than PW on both T_g and T_b, and hence on both GHIcs and DNIcs. Similarly, T_b is more sensitive than T_g to AOD550, as revealed by the former's steeper curve. GHIcs is less sensitive to AOD550 partly because its direct and diffuse components swing contrary when AOD550 fluctuates (Gueymard, 2019). As could be expected, the lower the AOD or PW, the higher the transmittance. It is noteworthy that varying PW over the range 4–7 cm does not cause any significant change in GHIcs or DNIcs. This confirms the discussion in Section 5.2.

The diurnal T_g and T_b curves are plotted in Fig. 7(a) for a fixed PW. Considering ideally constant atmospheric conditions during the day, the highest value of T_g (and thus of GHIcs) occurs at solar noon. The evolution of T_b is somewhat different: T_b increases fast after sunrise and then remains high during mid-day. The impact of varying AOD550 on T_g and T_b is also explored. The subplots in Fig. 7(b) and (c) show that a 0.05 variation in AOD550 only causes a slight difference in T_g. The absolute deviation in T_g's mean and maximum values for a 0.05 variation in AOD550 is about 0.005. The mean and maximum of T_b vary more when AOD550 is between 0.05 and 1, as shown in subplots (d) and (e). The absolute mean and maximum deviations of T_b for a 0.05 variation in AOD550 is more than 0.01, i.e., two to five times higher than with T_g. This corroborates the findings in Gueymard (2012b).

In Sections 5.1 and 5.3, a concern was raised because of the apparently high spatial variability in AOD550. Based on the analysis of the sensitivity of GHIcs to AOD550 just described, this spatial variability is low enough to make its impact on GHIcs acceptably low too. For example, the MBE of MERRA-2's AOD550 compared with the Pioneer and NUS stations are −0.044 and −0.240, respectively (see Table 4), thus the induced difference in global transmittance can be expected to be approximately 0.005 and 0.025, respectively. A major concern, however, is the bias propagation from AOD550 to DNIcs. Using the same MBE data, much larger differences occur in the direct transmittance, reaching approximately 0.02 and 0.1, respectively. This strong dependence is also demonstrated in Gueymard (2012b).

For PW, MERRA-2 achieves relatively high accuracy (see Table 5). Moreover, the bias propagation from PW to GHIcs and DNIcs is only marginal as compared to that of AOD550 (see Fig. 6). For clear-sky radiation modeling, these results suggest that MERRA-2's PW is normally a good alternative to ground observations, at least for the evaluation of GHIcs in the Singapore area.

5.5. Clear-sky radiation model comparison

This section displays a quantitative analysis comparing *in situ* (ground) and modeled clear-sky irradiance, as well as a qualitative inter-comparison between the selected clear-sky radiation models. This setup is motivated by the fact that these two types of comparison target different properties of clear-sky models, which have been thoroughly discussed in earlier works (Sun et al., 2019, 2021). Multiple clear-sky radiation models are included to isolate the modeling-related effects from the spatial variability analysis itself and increase confidence in any conclusions made. All clear-sky radiation models, except Yang, are operated with MERRA-2 input rather than ground-measured atmospheric data because the latter has much fewer available points and is in conflict with the primary goal of investigating how gridded data influences clear-sky modeling.

5.5.1. Model predictions versus SERIS observations

This subsection elaborates a quantitative comparison of the GHIcs predictions from six clear-sky radiation models and DNIcs predictions from five[4] of them at nine Singapore radiometric stations, i.e., all irradiance predictions are compared with SERIS ground observations. Note that the following comparison is conditioned by the uncertainty in the observations, and by the error propagation from the atmospheric inputs, as discussed above. Based on the available literature (Vuilleumier et al., 2017; Badosa et al., 2014; Möllenkamp et al., 2020), the SPN1 uncertainty is estimated at 5% for GHI and 10% for DNI. Some systematic bias has also been noted in the latter case. Hence, different performance results would likely be obtained with better-quality instrumentation and/or with ground-based atmospheric data. After data quality control and clear-sky detection (Bright et al., 2020b), 47,522 1-min data points are usable for further analysis. Values of two error metrics, nRMSE and nMBE, for each model's GHIcs and DNIcs predictions are listed in Table 6 and Table 7, respectively. The best nRMSE and nMBE at each site are in bold. Note that the mean measured values of GHIcs and DNIcs of all sites are 258 and 405 W m^{-2}, respectively. The average values are relatively low because the clear-sky detection method identified most clear-sky periods in the morning and late afternoon as a result of the regular convective cloud systems throughout the day, which are typical in equatorial climates.

From a general perspective, the GHIcs predictions from all models and at each of the ground stations have a similar performance in terms of nRMSE, despite their modeling diversity. The nRMSE values of the

[4] Recall one of the models, namely, Yang, does not offer DNIcs prediction.

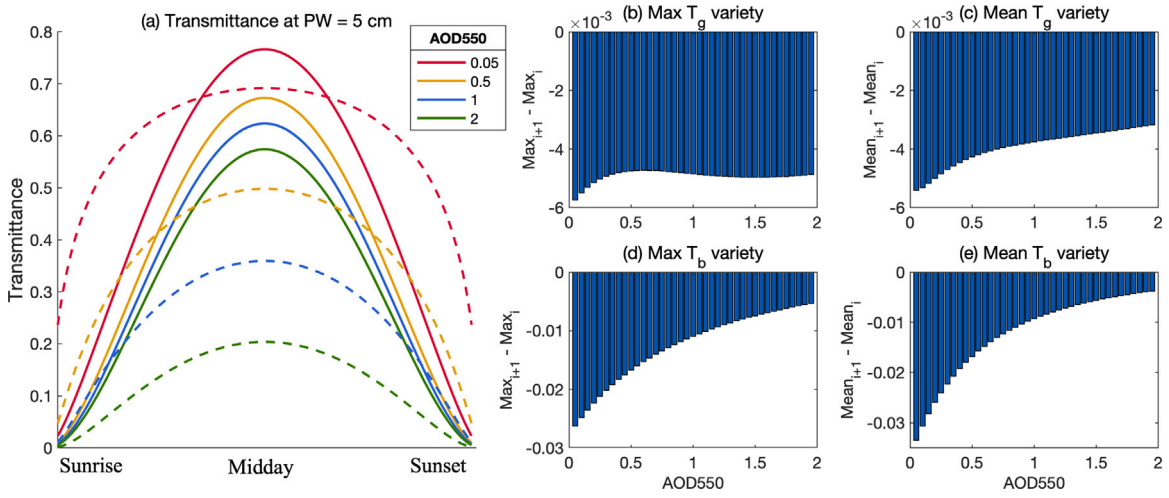

Fig. 7. Analysis of the AOD550 impact on the global transmittance T_g and direct transmittance T_b at PW = 5 cm. (a) Plots of T_g (solid lines) and T_b (dashed lines) during an ideal clear day. The transmittance increases when AOD550 decreases. (b) Difference between the maximum value of T_g for two adjacent AOD550 values, i.e., $\max(T_{g|AOD550=a+0.05}) - \max(T_{g|AOD550=a})$, where $a = 0, 0.05, \ldots, 1.95$. (c) Difference between the mean value of T_g at two adjacent AOD550 values. (d) Difference between the maximum value of T_b for two adjacent AOD550 values. (e) Difference between the mean value of T_b at two adjacent AOD550 values.

DNIcs predictions also exhibit the same features. Hence, it is fair to conclude that the model performance diversity is not significant, and thus echoes the adverse conditions noted above to a certain extent. The nMBE results cannot easily be summarized, except that a vast majority of the biases are positive, which may seem surprising. Barring any (highly unlikely) systematic faulty calibration of the instruments, the most reasonable explanation for this issue is the error propagation from the negatively biased MERRA-2 AOD550 data used as input, as suggested in Section 5.1, possibly combined with a systematic bias of the instrument in the case of DNIcs. The error propagation effect constitutes an important caveat in any validation study of radiation models. For that reason, the discussion and results below should only be considered as preliminary and subject to modification when better AOD information becomes available. With that caveat, and regarding the GHIcs results first, the nRMSE and nMBE results are found significantly larger than the assumed radiometer's uncertainty at most stations. Ineichen2018 is found to perform better at 5 ground stations in terms of nMBE, and also obtains the lowest mean nMBE overall. Another remarkable result is that of MAC2, which achieves the lowest average nRMSE over all stations. Yang unsurprisingly performs relatively well because it is empirically derived using global irradiance data from two radiometric stations in Singapore. Because that model is much simpler than all others and does not use any aerosol information, it avoids the error propagation issue altogether, which certainly contributes to its limited bias.

The DNIcs performance metrics in Table 7 indicate notably larger errors than for GHIcs and than the instrument's uncertainty at all stations. This is consistent with the findings in Section 5.4 that the aerosols bias has a greater influence on DNIcs than GHIcs predictions. This is also contrary to the expectation when model validations are based on AERONET input data and research-class instrumentation, e.g., Gueymard (2019, 2012a). All this tends to confirm the significance of the error propagation effect, in particular. With that caveat, MAC2 obtains the best scores at most stations, followed by the more sophisticated REST2v5.

Overall, it is found that no GHI or DNI model is able to outperform its peers by a significant margin, and no model is able to consistently be a top performer at all stations in Singapore. Moreover, the magnitude of the error metrics varies greatly from one site to another. Although the model's design and input quality directly impact its performance, the latter seems also to depend largely on the local conditions of the site itself, and possibly on each station's experimental error. With the currently available radiation data and MERRA-2 atmospheric inputs, MAC2 is found a good choice to evaluate both GHIcs and DNIcs in Singapore, especially considering its simplicity.

Table 6
Summary of nRMSE and nMBE error statistics for six clear-sky models' GHI predictions at nine ground stations. The "All" column is the absolute mean error value at the nine stations. 0 indicates a perfect value for nRMSE. The closer to 0, the better the nMBE value. The best nRMSE and nMBE at each site are in boldface. The mean measured values of GHIcs (GHIobs) of each site and the average value among all sites are displayed in the last row. The unit of GHIcs is W m^{-2}.

	nRMSE [%]									
	Site 1	2	3	4	5	6	7	8	9	All
Yang	8.3	29.5	29.7	18.9	23.5	26.5	19.5	31.7	29.7	24.1
INEI.2018	10.6	**27.8**	28.2	**17.8**	24.0	25.0	**17.4**	30.0	28.3	23.2
MAC2	8.5	28.1	28.4	18.0	**23.3**	**24.7**	19.4	29.4	28.2	**23.1**
MRMv6	11.0	28.0	**28.1**	18.7	24.3	24.9	20.3	**28.8**	**27.8**	23.5
Bird	9.9	30.6	30.7	20.6	24.2	26.8	23.4	31.0	30.1	25.3
REST2v5	**7.9**	28.6	28.9	18.3	**23.3**	25.1	19.9	29.9	28.6	23.4
	nMBE [%]									
	Site 1	2	3	4	5	6	7	8	9	All
Yang	−0.7	8.2	8.3	4.4	1.0	4.4	8.4	5.9	7.0	5.4
INEI.2018	−7.5	**0.8**	**0.9**	−2.7	−5.9	−2.8	**1.0**	**−1.4**	**−0.4**	**2.6**
MAC2	−2.0	6.7	6.8	3.0	**−0.4**	3.0	6.9	4.4	5.5	4.3
MRMv6	−3.2	5.5	5.6	**1.8**	−1.5	**1.8**	5.7	3.2	4.3	3.6
Bird	5.0	14.5	14.6	10.5	6.8	10.4	14.6	12.0	13.1	11.3
REST2v5	**−0.3**	8.7	8.8	4.9	1.4	4.8	8.9	6.4	7.4	5.7
Mean GHIobs	273	250	250	260	269	259	251	255	253	258

5.5.2. *Inter-comparison of clear-sky radiation models*

Considering the moderate quality of the SERIS radiometric observations, a qualitative inter-comparison between the clear-sky radiation models is conducted next. In this approach, which was previously applied in Suri et al. (2008) and Ruiz-Arias and Gueymard (2018), the analysis does not attempt to pick a "best" model, nor need high-quality ground measurements, but applies a relative cross-comparison between a number of models based on a common source of inputs to test their mutual consistency.

The mean inter-model deviation at location s and time t is defined following Ruiz-Arias and Gueymard (2018),

$$\delta_G(s,t) = \frac{1}{M}\sum_{i=1}^{M}\left|G_{hc}^{(i)}(s,t) - \mathrm{med}\left\{G_{hc}^{(j)}(s,t) : j = 1,\ldots,M\right\}\right|, \quad (7)$$

Table 7
Summary of nRMSE and nMBE error statistics for five clear-sky models' DNI predictions at nine ground stations. The "All" column is the mean error value at the nine stations. 0 indicate a perfect value for nRMSE. The closer to 0, the better the nMBE value. The best nRMSE and nMBE at each site are in boldface. The mean measured values of DNIcs (DNIobs) of each site and the average value among all sites are displayed in the last row. The unit of DNIcs is W m^{-2}.

	nRMSE [%]									
	Site 1	2	3	4	5	6	7	8	9	All
INEI.2018	19.2	42.2	44.8	40.1	35.0	43.3	44.4	45.0	37.2	39.0
MAC2	20.1	**38.2**	**39.4**	**36.0**	34.7	**40.2**	**39.1**	**40.5**	38.7	**36.3**
MRMv6	19.8	43.6	45.9	41.0	35.7	44.4	45.2	46.3	38.0	40.0
Bird	19.1	41.2	43.2	38.8	34.9	42.4	43.1	43.7	37.8	38.2
REST2v5	**17.0**	39.5	41.5	37.2	**33.6**	40.9	40.5	42.4	**36.4**	36.6
	nMBE [%]									
	Site 1	2	3	4	5	6	7	8	9	All
INEI.2018	7.8	21.5	24.2	19.2	10.2	19.2	22.5	23.0	7.7	17.3
MAC2	−7.7	**4.0**	**6.3**	**2.0**	−5.6	**1.9**	**4.9**	**5.2**	−7.8	**5.0**
MRMv6	9.3	23.3	26.0	20.9	11.8	20.9	24.2	24.7	9.3	18.9
Bird	3.1	16.2	18.8	14.0	5.4	14.0	17.2	17.6	3.0	12.1
REST2v5	**2.6**	15.6	18.2	13.4	**4.9**	13.4	16.6	17.0	**2.5**	11.6
Mean DNIobs	440	389	381	397	430	396	387	384	439	405

$$\delta_B(s,t) = \frac{1}{M}\sum_{i=1}^{M}\left|B_{nc}^{(i)}(s,t) - \text{med}\left\{B_{nc}^{(j)}(s,t) : j = 1,\ldots,M\right\}\right|, \quad (8)$$
$$\text{for } s \in \{1,\ldots,9\}, \text{ and } t \in \{1,\ldots,N\}.$$

where $G_{hc}^{(i)}(s,t)$ and $B_{nc}^{(i)}(s,t)$ are the GHIcs and DNIcs predictions for a particular *event*, from model i; med denotes the median operator; and $M = 6$ or 5 corresponds to the number of models under consideration (i.e., 6 GHIcs models and 5 DNIcs models). Here, an event denotes prediction at a particular location and time instance; given there are nine stations and 1,900,596 time instances, the total number of events is 17,105,364.

Two variants are then defined for further analysis, the time-averaged mean inter-model deviation ($\bar{\delta}_G^{(i)}(s)$ and $\bar{\delta}_B^{(i)}(s)$) and the site-dependent deviation ($\bar{\delta}_G(s)$ and $\bar{\delta}_B(s)$).

$$\bar{\delta}_G^{(i)}(s) = \frac{1}{N}\sum_{t=1}^{N}\left|G_{hc}^{(i)}(s,t) - \text{med}\left\{G_{hc}^{(j)}(s,t) : j = 1,\ldots,M\right\}\right|, \quad (9)$$

$$\bar{\delta}_B^{(i)}(s) = \frac{1}{N}\sum_{t=1}^{N}\left|B_{nc}^{(i)}(s,t) - \text{med}\left\{B_{nc}^{(j)}(s,t) : j = 1,\ldots,M\right\}\right|, \quad (10)$$

$$\bar{\delta}_G(s) = \frac{1}{M}\sum_{i=1}^{M}\bar{\delta}_G^{(i)}(s), \quad (11)$$

$$\bar{\delta}_B(s) = \frac{1}{M}\sum_{i=1}^{M}\bar{\delta}_B^{(i)}(s), \quad (12)$$
$$\text{for } s \in \{1,\ldots,9\}, \text{ and } i \in \{1,\ldots,M\}.$$

Fig. 8 shows the mean inter-model deviations δ_G and δ_B in AOD550–PW space for the 6 GHIcs models and 5 DNIcs models in Singapore. The deviations increase as the MERRA-2 AOD550 gets higher, most particularly for DNIcs. Conversely, such a trend is not observed with PW. A δ_G higher than 60 W m^{-2}, i.e., 9.2% of the median of all six models' GHIcs predictions for all locations and time, is observed for a few occasional AOD550 values that are greater than 1.5. This trend is more pronounced for DNI, for which such high deviations occur when AOD550 > 1. This confirms the difficulty of many simple radiation models to cope with high-AOD situations, as described in Gueymard and Ruiz-Arias (2015). In Singapore, high-AOD circumstances can be caused by intense pollution or smoke episodes. In the case of smoke, significant spatial variability in AOD can be expected because of the highly dynamic meteorological and chemical processes at play, as discussed in Section 5.1. Moreover, as discussed in Section 5.4, a 0.05 variation in AOD550 can cause a ≈0.5% difference in GHIcs and a 1%–2% difference in DNIcs. This is for one of the clear-sky radiation models; because these magnitudes are likely to be model-dependent, discrepancies can result when inter-comparing models of different construct. At most points where AOD550 is lower than 0.5, the discrepancy between models is relatively low.

Subplots (a) and (b) of Fig. 9 present the time-averaged mean inter-model deviation $\bar{\delta}_G^{(i)}(s)$ and $\bar{\delta}_B^{(i)}(s)$ of each model. It is found that all models perform consistently across all stations, as shown by the nearly horizontal lines. Thus the site-dependent deviation $\bar{\delta}_G(s)$ and $\bar{\delta}_B(s)$ values are very close at each location, as shown in subplots (c) and (d) of Fig. 9. The reason is that all models (except Yang) use the same spatial averaging inputs from MERRA-2. Interestingly, MAC2 and REST2v5 are almost indiscernible from the median of GHIcs. For DNIcs, the same is true for Ineichen2018 and REST2v5.

In contrast, Yang and MRMv6 consistently generate the highest values of the GHIcs deviation. In parallel, MAC2 and MRMv6 generate the highest values of the DNIcs deviation. These findings are compatible with those in Section 5.5.1.

However, since the models predict consistently across all ground stations (see Fig. 9), the variety in performance statistics based on different site observations in Tables 6 and 7 characterizes the spatial variance in atmospheric and environmental conditions in Singapore, which is not captured in MERRA-2's products because of their coarse resolution.

Another cross comparison of GHIcs and DNIcs predictions is shown as multiple scatter density plots in Fig. 10. The lower (left) triangle displays GHIcs and the upper (right) triangle displays DNIcs. Yang shows substantial discrepancy in comparison with the other models, which is normal since it does not use MERRA-2 data at all and thus cannot account for any hourly change in atmospheric conditions. In contrast, the GHIcs discrepancy among the other five models is not significant. One explanation is that these models share the same spatially averaged input variables from MERRA-2. Therefore, their outputs can be expected to vary in phase. Another reason is that GHIcs models are less sensitive to AOD than DNIcs models. Indeed, more model discrepancies are found for DNIcs in the upper triangular zone. The S shape of some of the DNIcs scatter plots are reminiscent of the findings in Gueymard (2019).

6. Conclusion

Gridded reanalysis databases such as MERRA-2 are widely used in solar resource assessments. The good spatio-temporal coverage of such reanalyses is an important asset to help improve clear-sky radiation modeling at global scale and over long time periods. Although some large-scale validation studies of MERRA-2 atmospheric predictions do exist, the impact of their use over urban-scale areas remains uncertain.

This study made use of MERRA-2's atmospheric data over Singapore, two AERONET ground stations with long records of aerosols and water vapor during clear periods, one SuomiNet GPSmet station providing continuous water vapor data, and nine SERIS radiometric stations that recorded 1-min global horizontal irradiance (GHI) and diffuse horizontal irradiance (DIF) between 2013 and 2020, from which direct normal irradiance (DNI) can be derived. These sources of data have enabled an in-depth exploration into the impact of spatial variability in atmospheric variables that relate to the prediction of clear-sky GHI (GHIcs) and clear-sky DNI (DNIcs).

Firstly, the study has validated the MERRA-2 aerosol and water vapor products with the corresponding ground truth from the AERONET and SuomiNet stations. The nRMSE of MERRA-2's aerosol optical depth at 550 nm (AOD550) at AERONET's Pioneer and NUS stations are 41.6% and 107.2%, respectively. The nRMSE of MERRA-2's Ångström exponent (AE) at the two stations is around 26%. It is clear that AOD550 suffers from high spatial variation at microscale and that its MERRA-2 predictions must be used with caution. The nRMSE of

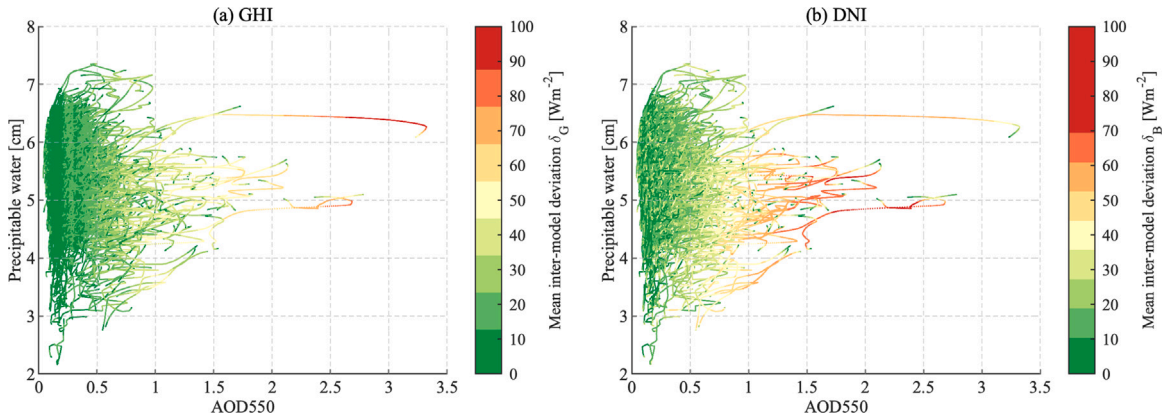

Fig. 8. Scatter density plot of the mean inter-model deviation δ in AOD550-PW space for GHIcs (left) and DNIcs (right) in Singapore, as calculated by Eqs. (7) and (8), respectively.

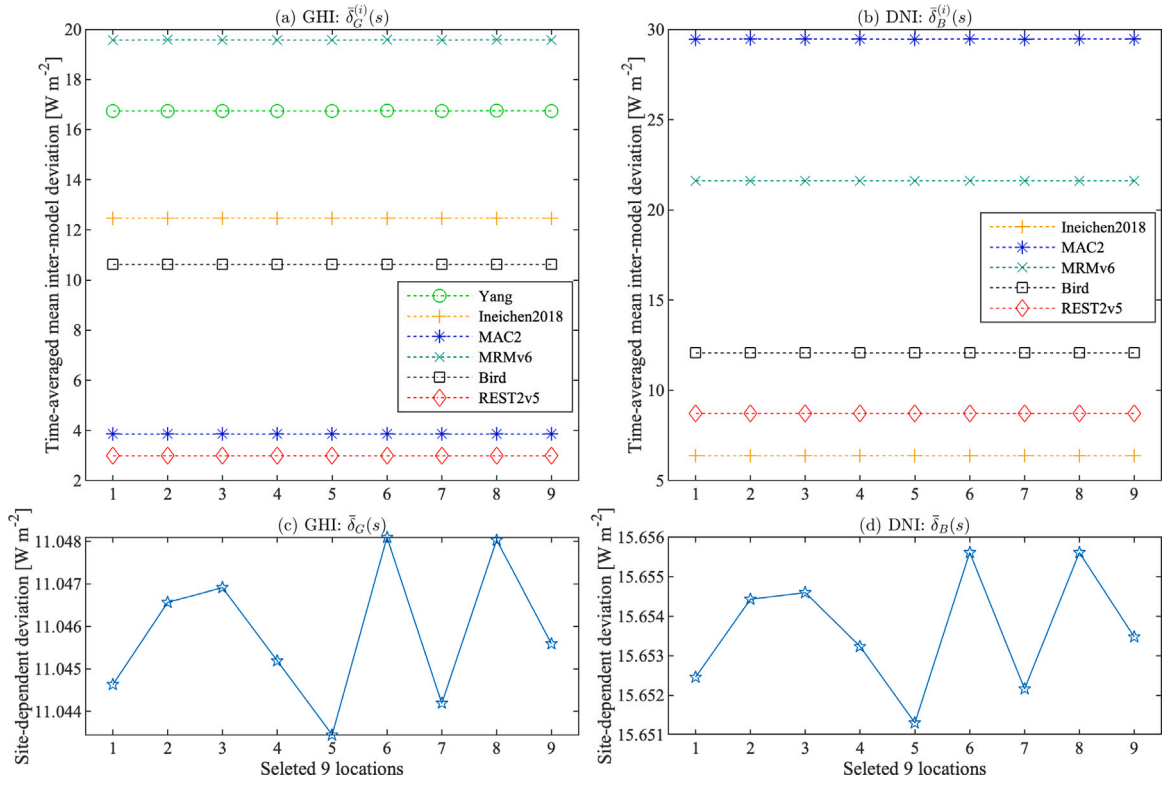

Fig. 9. Subplots (a) and (b): Time-averaged mean inter-model deviation $\bar{\delta}_G^{(i)}(s)$ for six GHIcs models and $\bar{\delta}_B^{(i)}(s)$ for five DNIcs models at nine stations (X-axis). Subplots (c) and (d): Site-dependent deviation $\bar{\delta}_G(s)$ for GHIcs or $\bar{\delta}_B(s)$ for DNIcs among the nine stations, as calculated by Eq. (9). The order of the nine stations is the same as in Table 1.

MERRA-2's precipitable water (PW) at Pioneer, NUS, and NTUS are reasonably low at 13.1%, 14.2%, and 7.6%, respectively.

The error propagation from aerosol data to clear-sky irradiance predictions was analyzed by applying a variable-controlling method. The Ineichen2018 model was chosen as a case study. It was found that an absolute variation (or error) of 0.05 in AOD550 yields a difference of 0.5% in GHIcs output, and a substantially larger difference of 1%–2% in DNIcs, thus underlining the critical role of error propagation from AOD to GHIcs or, even more so, DNIcs in the present analysis.

Finally, the MERRA-2 atmospheric variables were used as inputs to five selected clear-sky irradiance models to generate GHIcs and DNIcs predictions. A model inter-comparison has shown significant model discrepancies under high-AOD550 situations, confirming previous studies. The model validation based on measured irradiance data from the SERIS radiometric network has indicated that no clear-sky radiation model systematically outperforms any other in Singapore. The model-to-model differences are found essentially within the uncertainty limits of the radiometers. The model performance also varies with location, possibly because of the lack of accuracy and spatial resolution of the MERRA-2 data, in particular.

The simple Yang model achieves moderate performance in GHIcs predictions because it is empirically derived from local irradiance observations and is not affected by error propagation, contrarily to the five other models. Even though they are more elaborate than Yang, they do not outperform it significantly. In effect, they cannot really benefit from their more advanced modeling approach because of error propagation from their biased aerosol inputs. In contrast with the GHIcs-only Yang model, however, they can provide predictions of DNIcs, which constitutes an important advantage in practice. Based on various statistical results, the MAC2 model appears the proper choice for Singapore because it has a simple architecture, requires a limited number of atmospheric inputs, and generally obtains better error metrics in both GHIcs and DNIcs predictions. This conclusion is

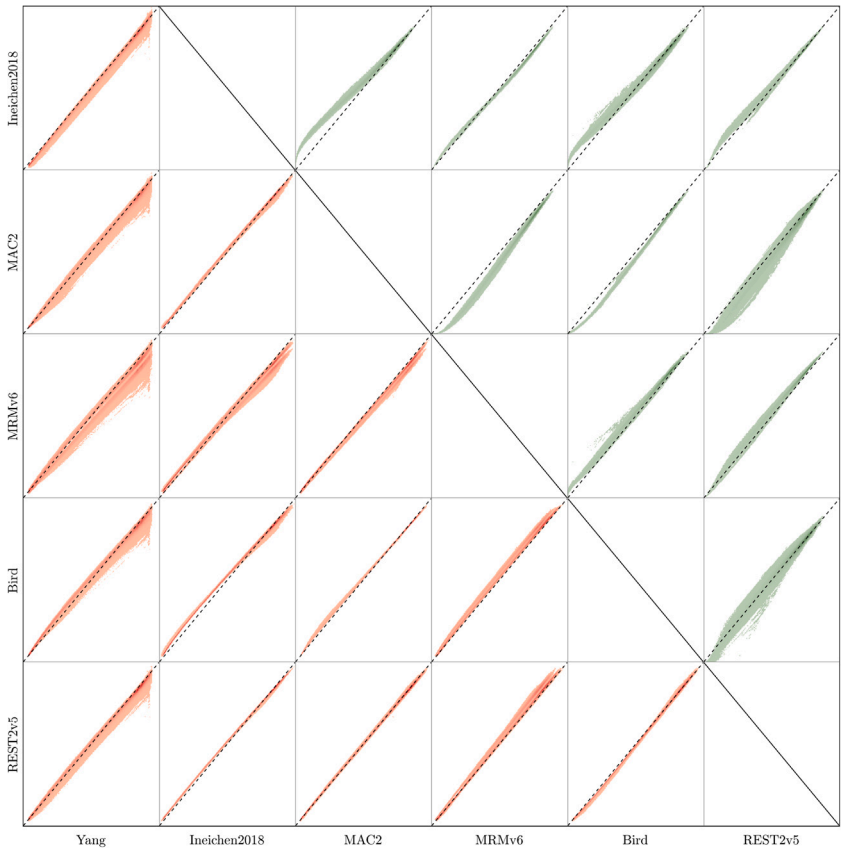

Fig. 10. Scatter density plots of GHIcs from six clear-sky models against each other in the lower (left, in orange color) triangle and DNIcs from five clear-sky models against each other in the upper (right, in green color) triangle. The shade in subplots indicates a cluster of data points. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

only preliminary, however, and would likely change if unbiased aerosol inputs and better DNI measurements became available.

Overall, in the absence of a dense network of high-quality ground-based measurements of aerosols and water vapor, the MERRA-2 reanalysis can be used to obtain relatively unbiased GHIcs and DNIcs predictions over local regions such as Singapore. The possibility of using aerosol databases with higher spatial resolution, even if at lower temporal resolution, would constitute a logical follow-up to mitigate the spatial scale mismatch.

Declaration of competing interest

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

Data availability

The MERRA-2 reanalysis database and the AERONET data are publicly available. The SERIS meteorological network is proprietary.

Acknowledgments

X. Sun is partially funded by the Energy Market Authority, Singapore (EMA), Energy Programme—Solar Forecasting Grant (NRF2017EWT-EP002-004). X. Sun and P. Wang were partially supported by the National Key Research and Development Program of China (Grant No. 2018YFB0703902). X. Sun is partly funded by the China Scholarship Council (No. 2020000000).

References

Badosa, J., Wood, J., Blanc, P., Long, C.N., Vuilleumier, L., Demengel, D., Haeffelin, M., 2014. Solar irradiances measured using SPN1 radiometers: uncertainties and clues for development. Atmos. Meas. Tech. 7 (12), 4267–4283. http://dx.doi.org/10.5194/amt-7-4267-2014.

Barnett, T.P., Ritchie, J., Foat, J., Stokes, G., 1998. On the space–time scales of the surface solar radiation field. J. Clim. 11 (1), 88–96. http://dx.doi.org/10.1175/1520-0442(1998)011<0088:OTSTSO>2.0.CO;2.

Benedetti, A., Morcrette, J.-J., Boucher, O., Dethof, A., Engelen, R.J., Fisher, M., Flentje, H., Huneeus, N., Jones, L., Kaiser, J.W., et al., 2009. Aerosol analysis and forecast in the European centre for medium-range weather forecasts integrated forecast system: 2. Data assimilation. J. Geophys. Res.: Atmos. 114 (D13), http://dx.doi.org/10.1029/2008JD011235.

Bird, R.E., Hulstrom, R.L., 1981. Simplified Clear Sky Model for Direct and Diffuse Insolation on Horizontal Surfaces. Technical Report, Solar Energy Research Inst., Golden, CO (USA), http://dx.doi.org/10.2172/6510849.

Blanc, P., Wald, L., 2012. The SG2 algorithm for a fast and accurate computation of the position of the sun for multi-decadal time period. Sol. Energy 86 (10), 3072–3083. http://dx.doi.org/10.1016/j.solener.2012.07.018.

Blöschl, G., Sivapalan, M., 1995. Scale issues in hydrological modelling: A review. Hydrol. Process. 9 (3–4), 251–290. http://dx.doi.org/10.1002/hyp.3360090305.

Boraiy, M., Korany, M., Aoun, Y., Alfaro, S., El-Metwally, M., Wahab, M., Blanc, P., Eissa, Y., Ghedira, H., Siour, G., et al., 2017. Improving direct normal irradiance retrieval in cloud-free, but high aerosol load conditions by using aerosol optical depth. Meteorol. Z. 26 (5), 475–483. http://dx.doi.org/10.1127/metz/2017/0844.

Bright, J.M., Bai, X., Zhang, Y., Sun, X., Acord, B., Wang, P., 2020a. Irradpy: Python package for MERRA-2 download, extraction and usage for clear-sky irradiance modelling. Sol. Energy 199, 685–693. http://dx.doi.org/10.1016/j.solener.2020.02.061.

Bright, J.M., Gueymard, C.A., 2019. Climate-specific and global validation of MODIS Aqua and Terra aerosol optical depth at 452 AERONET stations. Sol. Energy 183, 594–605. http://dx.doi.org/10.1016/j.solener.2019.03.043.

Bright, J.M., Sun, X., Gueymard, C.A., Acord, B., Wang, P., Engerer, N.A., 2020b. BrightSun: A globally applicable 1-min irradiance clear-sky detection model. Renew. Sustain. Energy Rev. 121, 109706. http://dx.doi.org/10.1016/j.rser.2020.109706.

Casagrande, M.S., Martins, F.R., Rosário, N.E., Lima, F.J., Gonçalves, A.R., Costa, R.S., Zarzur, M., Pes, M.P., Pereira, E.B., 2021. Numerical assessment of downward incoming solar irradiance in smoke influenced regions–A case study in Brazilian Amazon and Cerrado. Remote Sens. 13 (22), 4527. http://dx.doi.org/10.3390/rs13224527.

Driemel, A., Augustine, J., Behrens, K., Colle, S., Cox, C., Cuevas-Agulló, E., Denn, F.M., Duprat, T., Fukuda, M., Grobe, H., et al., 2018. Baseline Surface Radiation Network (BSRN): structure and data description (1992–2017). Earth Syst. Sci. Data 10 (3), 1491–1501. http://dx.doi.org/10.5194/essd-10-1491-2018.

Gao, W., Coulter, R.L., Lesht, B.M., Qiu, J., Wesely, M.L., 1998. Estimating clear-sky regional surface fluxes in the Southern Great Plains atmospheric radiation measurement site with ground measurements and satellite observations. J. Appl. Meteorol. 37 (1), 5–22. http://dx.doi.org/10.1175/1520-0450(1998)037<0005:ECSRSF>2.0.CO;2.

Gelaro, R., McCarty, W., Suárez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., et al., 2017. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). J. Clim. 30 (14), 5419–5454. http://dx.doi.org/10.1175/JCLI-D-16-0758.1.

Giles, D.M., Sinyuk, A., Sorokin, M.G., Schafer, J.S., Smirnov, A., Slutsker, I., Eck, T.F., Holben, B.N., Lewis, J.R., Campbell, J.R., et al., 2019. 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. Atmos. Meas. Tech. 12 (1), 169–209. http://dx.doi.org/10.5194/amt-2018-272.

Gueymard, C.A., 2003. Direct solar transmittance and irradiance predictions with broadband models. Part II: validation with high-quality measurements. Sol. Energy 74 (5), 381–395. http://dx.doi.org/10.1016/S0038-092X(03)00196-8.

Gueymard, C.A., 2008. REST2: HIgh-performance solar radiation model for cloudless-sky irradiance, illuminance, and photosynthetically active radiation: Validation with a benchmark dataset. Sol. Energy 82 (3), 272–285. http://dx.doi.org/10.1016/j.solener.2007.04.008.

Gueymard, C.A., 2012a. 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 86 (8), 2145–2169. http://dx.doi.org/10.1016/j.solener.2011.11.011.

Gueymard, C.A., 2012b. Temporal variability in direct and global irradiance at various time scales as affected by aerosols. Sol. Energy 86 (12), 3544–3553. http://dx.doi.org/10.1016/j.solener.2012.01.013.

Gueymard, C.A., 2014a. 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. 39, 1024–1034. http://dx.doi.org/10.1016/j.rser.2014.07.117.

Gueymard, C.A., 2014b. Impact of on-site atmospheric water vapor estimation methods on the accuracy of local solar irradiance predictions. Sol. Energy 101, 74–82. http://dx.doi.org/10.1016/j.solener.2013.12.027.

Gueymard, C.A., 2018. A reevaluation of the solar constant based on a 42-year total solar irradiance time series and a reconciliation of spaceborne observations. Sol. Energy 168, 2–9. http://dx.doi.org/10.1016/j.solener.2018.04.001.

Gueymard, C.A., 2019. Clear-sky radiation models and aerosol effects. In: Solar Resources Mapping. Springer, pp. 137–182. http://dx.doi.org/10.1007/978-3-319-97484-2_5.

Gueymard, C.A., Ruiz-Arias, J.A., 2015. Validation of direct normal irradiance predictions under arid conditions: A review of radiative models and their turbidity-dependent performance. Renew. Sustain. Energy Rev. 45, 379–396. http://dx.doi.org/10.1016/j.rser.2015.01.065.

Gueymard, C.A., Yang, D., 2020. Worldwide validation of CAMS and MERRA-2 reanalysis aerosol optical depth products using 15 years of AERONET observations. Atmos. Environ. 225, 117216. http://dx.doi.org/10.1016/j.atmosenv.2019.117216.

Hansen, A.B., Witham, C.S., Chong, W.M., Kendall, E., Chew, B.N., Gan, C., Hort, M.C., Lee, S.-Y., 2019. Haze in Singapore–source attribution of biomass burning PM 10 from Southeast Asia. Atmos. Chem. Phys. 19 (8), 5363–5385. http://dx.doi.org/10.5194/acp-19-5363-2019.

Holben, B.N., Eck, T.F., Slutsker, I., Tanre, D., Buis, J., Setzer, A., Vermote, E., Reagan, J.A., Kaufman, Y.J., Nakajima, T., 1998. AERONET—A federated instrument network and data archive for aerosol characterization. Remote Sens. Environ. 66 (1), 1–16. http://dx.doi.org/10.1016/S0034-4257(98)00031-5.

Huang, L., Wang, X., Xiong, S., Li, J., Liu, L., Mo, Z., Fu, B., He, H., 2022. High-precision GNSS PWV retrieval using dense GNSS sites and in-situ meteorological observations for the evaluation of MERRA-2 and ERA5 reanalysis products over China. Atmos. Res. 276, 106247. http://dx.doi.org/10.1016/j.atmosres.2022.106247.

Ineichen, P., 2018. High turbidity solis clear sky model: Development and validation. Remote Sens. 10 (3), 435. http://dx.doi.org/10.3390/rs10030435.

Kambezidis, H.D., Psiloglou, B.E., Karagiannis, D., Dumka, U.C., Kaskaoutis, D.G., 2017. Meteorological Radiation Model (MRM v6. 1): Improvements In diffuse radiation estimates and a new approach for implementation of cloud products. Renew. Sustain. Energy Rev. 74, 616–637. http://dx.doi.org/10.1016/j.rser.2017.02.058.

Long, C., Shi, Y., 2006. The QCRad Value Added Product: Surface Radiation Measurement Quality Control Testing, Including Climatology Configurable Limits. Technical Report DOE/SC-ARM/TR-074, Department of Energy, http://dx.doi.org/10.2172/1019540.

Macke, A., Seifert, P., Baars, H., Barthlott, C., Beekmans, C., Behrendt, A., Bohn, B., Brueck, M., Bühl, J., Crewell, S., et al., 2017. The HD(CP)^2 observational prototype experiment (HOPE)–an overview. Atmos. Chem. Phys. 17 (7), 4887–4914. http://dx.doi.org/10.5194/acp-17-4887-2017.

Madhavan, B.L., Deneke, H., Witthuhn, J., Macke, A., 2017. Multiresolution analysis of the spatiotemporal variability in global radiation observed by a dense network of 99 pyranometers. Atmos. Chem. Phys. 17 (5), 3317–3338. http://dx.doi.org/10.5194/acp-17-3317-2017.

Molero, B., Leroux, D.J., Richaume, P., Kerr, Y.H., Merlin, O., Cosh, M.H., Bindlish, R., 2018. Multi-timescale analysis of the spatial representativeness of in situ soil moisture data within satellite footprints. J. Geophys. Res.: Atmos. 123 (1), 3–21. http://dx.doi.org/10.1002/2017JD027478.

Möllenkamp, J., Beikircher, T., Häberle, A., 2020. Recalibration of SPN1 pyranometers against pyrheliometer and its relevance for the evaluation of concentrating solar process heat plants. Sol. Energy 197, 344–358. http://dx.doi.org/10.1016/j.solener.2019.12.055.

Nobre, A.M., Karthik, S., Liu, H., Yang, D., Martins, F.R., Pereira, E.B., Rüther, R., Reindl, T., Peters, I.M., 2016. On the impact of haze on the yield of photovoltaic systems in Singapore. Renew. Energy 89, 389–400. http://dx.doi.org/10.1016/j.renene.2015.11.079.

Remer, L.A., Kaufman, Y., Tanré, D., Mattoo, S., Chu, D., Martins, J.V., Li, R.-R., Ichoku, C., Levy, R., Kleidman, R., et al., 2005. The MODIS aerosol algorithm, products, and validation. J. Atmos. Sci. 62 (4), 947–973. http://dx.doi.org/10.1175/JAS3385.1.

Ruiz-Arias, J.A., Gueymard, C.A., 2018. 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 168, 10–29. http://dx.doi.org/10.1016/j.solener.2018.02.008.

Ruiz-Arias, J.A., Gueymard, C.A., Cebecauer, T., 2019. Direct normal irradiance modeling: Evaluating the impact on accuracy of worldwide gridded aerosol databases. AIP Conf. Proc. 2126 (1), 190013. http://dx.doi.org/10.1063/1.5117710.

Shi, H., Xiao, Z., Zhan, X., Ma, H., Tian, X., 2019. Evaluation of MODIS and two reanalysis aerosol optical depth products over AERONET sites. Atmos. Res. 220, 75–80. http://dx.doi.org/10.1016/j.atmosres.2019.01.009.

Sun, X., Bright, J.M., Gueymard, C.A., Acord, B., Wang, P., Engerer, N.A., 2019. Worldwide performance assessment of 75 global clear-sky irradiance models using principal component analysis. Renew. Sustain. Energy Rev. 111, 550–570. http://dx.doi.org/10.1016/j.rser.2019.04.006.

Sun, X., Bright, J.M., Gueymard, C.A., Bai, X., Acord, B., Wang, P., 2021. Worldwide performance assessment of 95 direct and diffuse clear-sky irradiance models using principal component analysis. Renew. Sustain. Energy Rev. 135, 110087. http://dx.doi.org/10.1016/j.rser.2020.110087.

Suri, M., Remund, J., Cebecauer, T., Dumortier, D., Wald, L., Huld, T., Blanc, P., 2008. First steps in the cross-comparison of solar resource spatial products in Europe. In: EUROSUN 2008, 1st International Conference on Solar Heating, Cooling and Buildings. ISES, Lisbonne, Portugal, p. CD.

Vuilleumier, L., Félix, C., Vignola, F., Blanc, P., Badosa, J., Kazantzidis, A., Calpini, B., 2017. Performance evaluation of radiation sensors for the solar energy sector. Meteorol. Z. http://dx.doi.org/10.1127/metz/2017/0836.

Wang, S., Xu, T., Nie, W., Jiang, C., Yang, Y., Fang, Z., Li, M., Zhang, Z., 2020. Evaluation of precipitable water vapor from five reanalysis products with ground-based GNSS observations. Remote Sens. 12 (11), 1817. http://dx.doi.org/10.3390/rs12111817.

Ware, R.H., Fulker, D.W., Stein, S.A., Anderson, D.N., Avery, S.K., Clark, R.D., Droegemeier, K.K., Kuettner, J.P., Minster, J.B., Sorooshian, S., 2000. SuomiNet: A real-time national GPS network for atmospheric research and education. Bull. Am. Meteorol. Soc. 81 (4), 677–694. http://dx.doi.org/10.1175/1520-0477(2000)081<0677:SARNGN>2.3.CO;2.

Witthuhn, J., Hünerbein, A., Filipitsch, F., Wacker, S., Meilinger, S., Deneke, H., 2021. Aerosol properties and aerosol–radiation interactions in clear-sky conditions over Germany. Atmos. Chem. Phys. 21 (19), 14591–14630. http://dx.doi.org/10.5194/acp-21-14591-2021.

Wu, X., Xiao, Q., Wen, J., You, D., Hueni, A., 2019. Advances in quantitative remote sensing product validation: Overview and current status. Earth-Sci. Rev. 196, 102875. http://dx.doi.org/10.1016/j.earscirev.2019.102875.

Yang, D., 2020. Quantifying the spatial scale mismatch between satellite-derived solar irradiance and in situ measurements: A case study using CERES synoptic surface shortwave flux and the Oklahoma Mesonet. J. Renew. Sustain. Energy 12 (5), 056104. http://dx.doi.org/10.1063/5.0025771.

Yang, D., Walsh, W.M., Jirutitijaroen, P., 2014. Estimation and applications of clear sky global horizontal irradiance at the equator. J. Sol. Energy Eng. 136 (3), 034505. http://dx.doi.org/10.1115/1.4027263.