Global and Regional Evaluation of the CERES Edition-4A Surface Solar Radiation and Its Uncertainty Quantification

Ke Zhang, Long Zhao, Wenjun Tang, Kun Yang, and Jing Wang

***Abstract*—This article presents a comprehensive evaluation of the 2000–2018 Clouds and Earth's Radiant Energy System Synoptic 1° Ed4A (CERES SYN1deg Edition 4A) surface solar radiation (SSR) product. In particular, the global assessment is conducted over different temporal scales (i.e., hourly, daily, and monthly-average) with special attention given to the impact of clouds, and a regional evaluation is further implemented over the Mainland of China (MC) using SSR measurements from a denser observational network provided by the China Meteorological Administration. Evaluation across all valid station-grid pairs yields mixed performance with $|\text{MBE}| \leq 2.8$ (6.2) W m^{-2}, RMSE$\leq$89.5 (31.6) W m^{-2}, and $R \geq 0.95$ (0.93) over the globe (MC) for different temporal scales, and the monthly CERES SSR, with RMSE$\leq$20 W m^{-2}, is found to hold promise for global numerical weather prediction and climate monitoring. In addition, CERES is found to generally underestimate and overestimate SSR over land and ocean, respectively. Comparison between year-round and cloudy-season suggests that the presence of clouds may potentially impact the SSR retrievals, especially at the hourly temporal scales, with an increase in RMSE values larger than 10 W m^{-2} for most stations. Further investigation of subgrid heterogeneity suggests that most *in situ* SSR measurements can reasonably represent the 1° grid average except for some stations with specific geographic deployments, which may raise significant spatial representativeness issues and, therefore, need to be used with great caution.**

***Index Terms*—Clouds and Earth's radiant energy system synoptic (CERES), cloud, spatial representativeness, surface solar radiation, uncertainty quantification.**

Manuscript received December 2, 2021; revised March 1, 2022; accepted March 30, 2022. Date of publication April 4, 2022; date of current version April 27, 2022. This work was supported by the National Natural Science Foundation of China under Grant 41805133 and Grant 42171360. *(Corresponding author: Long Zhao.)*

Ke Zhang, Long Zhao, and Jing Wang are with the Chongqing Jinfo Mountain Karst Ecosystem National Observation and Research Station, School of Geographical Sciences, Southwest University, Chongqing 400715, China, and also with the Chongqing Engineering Research Center for Remote Sensing Big Data Application, School of Geographical Sciences, Southwest University, Chongqing 400715, China (e-mail: zk1011@email.swu.edu.cn; zhaol04@swu.edu.cn; wangjing929@email.swu.edu.cn).

Wenjun Tang is with the Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100101, China (e-mail: tangwj@itpcas.ac.cn).

Kun Yang is with the Department of Earth System Science, Ministry of Education Key Laboratory for Earth System Modeling, Institute for Global Change Studies, Tsinghua University, Beijing 100084, China (e-mail: yangk@tsinghua.edu.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/JSTARS.2022.3164471, provided by the authors.

Digital Object Identifier 10.1109/JSTARS.2022.3164471

I. Introduction

SURFACE solar radiation (SSR) is an essential component in the energy balance of the earth's surface nexus and has a profound impact on land-atmospheric interactions. In addition, SSR data provide a basis for the design of buildings and estimation of crop yields and are a major input in solar energy-related projects [1]–[6]. In this regard, reliable and high-quality SSR data are, therefore, urgently required for both hydrometeorological and climatic research as well as social-industrial applications.

Aided by meteorological stations, some studies [7]–[14] have attempted to estimate SSR through observed meteorological variables, including cloud fraction, air temperature, and sunshine duration, yet there are still limitations mainly due to inadequate spatiotemporal coverage and point-scale representativeness of meteorological stations. Tang *et al.* [15] proposed estimating global high-resolution SSR by jointly using meteorological and cloud information from various reanalysis and satellite retrievals. However, the reconstructed SSR highly depends on the quality and temporal coverage of the satellite products. Satellite remote sensing, on the other hand, holds better spatial and temporal coverage and is expected to provide more straightforward SSR retrievals at regional or global scales [16]. Recent decades have seen a series of satellite-based SSR products with different temporal coverages/resolutions, including those from the energy radiation budget experiment (ERBE: 1984–1999/monthly; [17]), the global energy and water cycle experiment-surface radiation budget (GEWEX-SRB: 1983–2007/3-h; [18]), the University of Maryland-Surface Radiation Budget (UMD-SRB: 1983–2007/3-h; [19]), the International Satellite Cloud Climatology Project-Flux Data (ISCCP-FD: 3-h/1983–2009; [20]), the Global Land Surface Satellite Products System (GLASS: 2000–2019/daily; [21]) dataset, and the Clouds and Earth's Radiant Energy Systems (CERES: 2000–present/1-h; [22]). Nevertheless, satellite-retrieved solar radiation products may be subject to unknown uncertainties that originate from many factors, such as inaccurate sensor measurements, defect retrieval algorithms, and biased retrieving parameters, and therefore, require comprehensive evaluations prior to practical applications.

Among the abovementioned satellite-based SSR products, the CERES data are characterized by a relatively longer timespan and higher temporal resolution compared with the others. CERES has recently released the fourth edition of the

Synoptic 1° product (Ed4A). In fact, the past three years have already seen independent evaluations regarding this specific dataset. Most of these studies have been conducted over a specific area at certain temporal scales (e.g., the Southern Ocean: hourly and annual mean [23]; Arctic: 3-h [24]; the three poles: hourly, 3-h, daily, and monthly [25]; and global land: daily [26]). Nevertheless, with regard to the needs for different applications and researchs [16], it is of great concern to know its accuracy for the whole globe (including land and ocean) across different temporal scales. Thus, this article first attempted to evaluate the Ed4A product over the whole globe (including land and ocean) using 85 global hourly *in situ* SSR observations collected by CERES by simultaneously considering different temporal scales (i.e., from hourly to monthly).

Owing to different geoclimatic conditions, the CERES SSR retrieval may perform differently in different spatial extents, especially over areas that are prone to heavy aerosol loads, and thus makes regional evaluations necessary. However, it is often hampered by the lack of reliable regional SSR *in situ* observations. Taking the Mainland of China (MC), a typical aerosol-affected region [27]–[30], as an example, it is necessary to know the performance of CERES Ed4A over MC. Unfortunately, there is only one station in MC for the *in situ* observations used in the global evaluation, which is not sufficient to support the in-depth evaluation over MC. Nevertheless, the China Meteorological Administration (CMA) initiated a dense SSR observatory over MC decades ago. Building upon these ground observations, this article further investigated the performance of CERES SSR over MC and compared it with the global evaluation results. Furthermore, it is well known that the presence of clouds can potentially influence most satellite SSR retrievals (e.g., Li *et al.* [26]), yet its impact on the CERES Ed4A SSR product remains to be better quantified. In this article, the cloudy-season evaluation was further conducted over the globe and MC and then compared with the year-round results to more deeply investigate the effect of clouds on CERES SSR performance. It is worth noting that other factors, such as the sunlight effect, water vapor, and trace gases (e.g., CO_2, O_2, and O_3) in the atmosphere, may also influence the accuracy of SSR retrievals but are beyond the scope of this study and are, therefore, not covered in this particular work.

In addition, accurate and spatially representative station SSR measurements are the basis for regional and global gridded SSR evaluation. For the *in situ* observations, it is particularly important to clarify, as it is commonly used for satellite and model evaluation. Schwarz *et al.* [31] studied this by examining to what extent they represent the spatial average over the monthly scale [31]. However, the representativeness (e.g., daily) remains unclear. This was further investigated by using the daily SSR from the Satellite Application Facility on Climate Monitoring (CM SAF) with a high-spatial resolution of 0.05°.

In summary, this article intends to provide a comprehensive spatiotemporal evaluation of the CERES Ed4A SSR product. It builds upon the previous work by 1) using both the ocean- and ground-available *in situ* SSR stations at the global scale as well as a dense regional SSR ground observational network over MC, 2) exploring the performance of CERES SSR over different temporal scales (i.e., hourly, daily, and monthly), 3) investigating the impacts of clouds on SSR retrieval accuracy through comparison of year-round and cloudy-season products, and 4) exploring the representativeness of *in situ* observations for the 1° CERES grid.

The rest of this article is organized as follows: Section II presents the CERES SSR, *in situ* observations, auxiliary data, and evaluation methods; Section III presents the evaluation results of CERES SSR over the globe and MC with regard to different temporal scales, comparison between year-round, and cloudy seasons with a focus on the impact of clouds, and subgrid heterogeneity and spatial representativeness of *in situ* observations; Section IV discusses the seasonal variation in SSR, difference between CERES performance over land and ocean, and comparison against previous studies, followed by concluding remarks in Section V.

II. Data and Methodology

The evaluation is conducted mainly by comparing the CERES SSR product against two collections of *in situ* observations, with additional investigation on evaluation uncertainties using a series of independent satellite and reanalysis datasets. To be more specific, *in situ* observational data consist of hourly SSR at 85 stations across the globe, as well as CMA daily SSR from 90 stations for MC. A finer-resolution radiation product from the CM SAF is used to explore the subgrid spatial variability in the SSR. In addition, total cloud cover (TCC) from the CERES and SSR from the European Centre for Medium Range Weather Forecasts Reanalysis 5 (ERA5) are further adopted to support studies on the cloud impact on satellite-retrieved SSR and representativeness of individual stations, respectively. The following briefly describes the satellite data, *in situ* observations, and reanalysis data used in this article as well as evaluation methods.

A. CERES SSR Product

The National Aeronautics and Space Administration (NASA) CERES sensors are onboard the Terra and Aqua satellites, which were successively launched in 1999 and 2002 and deployed on a sun-synchronous, near-polar orbit. The CERES project has delivered products since March 2000 in UTC time format at four levels, namely, Level 0 is the raw digitized instrument data, Level 1 is the raw engineering and status data, Level 2 is the instantaneous geophysical variables at the CERES footprint resolution, and Level 3 is the spatially averaged and temporally interpolated grid-averaged radiative fluxes and cloud properties. It adopts the Langley Fu–Liou radiative transfer model, which takes Moderate-resolution Imaging Spectroradiometer (MODIS) and geostationary satellite (GEO) cloud, Global Modeling and Assimilation Office atmospheric profile, MODIS aerosol, and CERES top of atmosphere fluxes as input to retrieve SSR [32], [33]. The CERES SSR is reserved at a $1° \times 1°$ resolution with each grid value obtained from the spatial averaging of radiation retrievals within the grid. Ed4A, as the Level 3 and latest release of CERES SSR products, succeeds its previous release (Ed3A) by refining temporal resolution (from 3-h to hourly) and with significant improvement in anomalous time series. Further details about the CERES SSR products can be found in Rutan *et al.* [34]. In this article, the hourly, daily, and monthly average Ed4A SSR products, with larger time spans from March 2000 to March

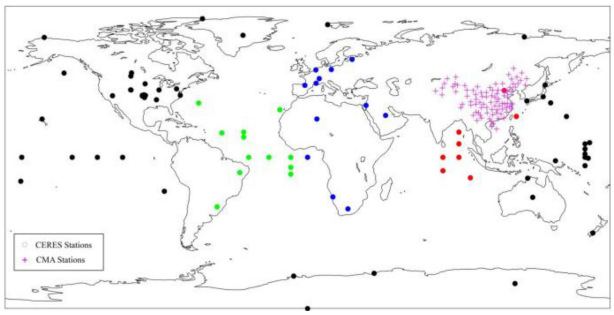


Fig. 1. Spatial distribution of *in situ* SSR stations used for CERES evaluation. Purple crosses indicate the CMA collection of meteorological monitoring stations, and dots indicate the CERES collection, of which red, green, and blue dots each denote those within CM SAF SARAH-E 1.1 and SARAH 2.1 and their shared spatial coverages, respectively.

2018, are further evaluated over the whole globe and MC using much denser *in situ* observations. Given the available hourly CERES satellite SSR, its daily and monthly means are derived through a simple arithmetic average. Note that the nighttime hourly SSR was excluded from the averaging.

B. In Situ *SSR Observations*

The *in situ* data used to evaluate the CERES satellite radiation product include global hourly data and CMA daily data, as illustrated below.

The global hourly *in situ* observed SSR is obtained from the CERES atmospheric radiation measurement validation experiment (CAVE; [35]) database, which covers a total of 85 *in situ* stations, including 37 stations from the Baseline Surface Radiation Network (BSRN; [36]), 24 stations from NOAA Pacific Marine Environmental Laboratory (PMEL; [37]), and 24 stations from other worldwide agencies. This dataset is maintained at NASA Langley Research Center in UTC time format for use in the CERES project but is also open to the scientific community.[1] The global distribution of these hourly stations is indicated by dots in Fig. 1. It is worth noting that these stations are distributed unevenly across the globe, with more concentrated in North America and Europe and fewer over Africa, South America, Asia, and Oceania. There are 58 land stations and 27 ocean stations, and most of the offshore stations are located within 30°N–30°S. The geolocation information and a clearer spatial distribution of the 85 h stations and their data availability are provided in the supplemental materials (see Table S1, Figs. S1 and S3a).

The CMA daily SSR is obtained from 90 densely deployed meteorological stations in local time over MC (indicated by purple crosses in Fig. 1) and can be downloaded from the website[2]. These stations have employed the so-called DFY-4 thermopile pyranometer to measure total radiation since 1993. The radiometers are operated upon the thermoelectric effect, whereby a large thermoelectric charge is generated when a blackbody absorbs radiant energy and the output voltage is proportional to the intensity of the radiation. According to Shi *et al.* [38], the DFY-4 thermopile pyranometer produces reliable SSR observations with less than 5% errors. Here, prior to the evaluation, all CMA *in situ* SSR measurements underwent a rigorous quality control procedure through a four-step scheme as proposed by Tang *et al.* [39]. Detailed geolocation information and a clearer spatial distribution of the 90 CMA stations and their data availability are provided in the supplemental materials (see Table S2, Figs. S2 and S3b).

C. CM SAF SSR and Auxiliary Data

The CM SAF provides various surface radiation parameters and is mostly grouped into two product families, namely, Cloud, Albedo, and Radiation (CLARA; [40]), and surface solar radiation dataset–Heliosat (SARAH; [41]). The former is based on polar orbiting satellites that cover the whole globe but with a relatively coarse resolution of $0.25° \times 0.25°$, while the latter is based on geostationary METEOSAT satellites and mainly covers Europe, Africa, the Atlantic, and parts of South America. In addition, there are also surface solar radiation dataset-Heliosat-east data records (SARAH-E; [42]), which are generated through collaboration between CM SAF and the European Commission's Joint Research Center and cover the Indian Ocean, India, and parts of China. Both the SARAH and SARAH-E products feature a much finer spatial resolution of $0.05° \times 0.05°$ and show good agreement with *in situ* observations [43]. In this article, the recent releases of monthly average SARAH-E 1.1 (March 2000–December 2016) and SARAH 2.1 (March 2000–December 2017) SSR datasets are jointly adopted to study subgrid heterogeneities of the CERES SSR product. The hourly global SSR stations within CM SAF SARAH-E 1.1 and SARAH 2.1 and their shared spatial coverages are highlighted in Fig. 1.

The TCC data ($1° \times 1°$, monthly) used in this article is also provided by CERES SYN1deg Ed4A to minimize the impact of the cloud product itself on the evaluation results. It is derived based on the cloud properties measured by the MODIS sensor on board the Terra and Aqua satellites and GEO images with four cloud cover types (surface-700 hPa, 700–500 hPa, 500–300 hPa, and 300–50 hPa) [44]. TCC is the total cloud cover of the four layers, and it is applied to quantify the influence of clouds on CERES satellite retrievals (see Section III-C). ERA5 is the fifth generation of ECMWF succeeding ERA-Interim and consists of estimates of a large number of atmospheric, ocean wave, and land surface quantities through assimilation of high-altitude and near-surface observations [45]. In this article, we rely on third-party independent $0.1° \times 0.1°$ monthly ERA5-Land SSR estimates to support the diagnosis of certain stations that exhibit problematic SSR observations (see Section IV-A). Specifically, the ERA5 SSR is modeled through the Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land (H-TESSEL). More details on the physical process of ERA5 are available online.[3]

[1][Online]. Available: https://ceres-tool.larc.nasa.gov/ord-tool/jsp/CAVE41Selection.jsp

[2][Online]. Available: http://data.cma.cn

[3][Online]. Available: https://www.ecmwf.int/en/elibrary/18714-ifs-documentation-cy45r1-part-iv-physical-processes

TABLE I
FORMULAS RELATED TO THE FIVE METRICS

Metric	Formula
RMSE	$\sqrt{\frac{\sum_{i=1}^{N}(S_i - G_i)^2}{N}}$
rRMSE	$\lvert\frac{RMSE}{\bar{G}}\rvert \times 100\%$
MBE	$\frac{\sum_{i=1}^{N}(S_i - G_i)}{N}$
rMBE	$\lvert\frac{MBE}{\bar{G}}\rvert \times 100\%$
R	$\frac{1}{N-1}\sum_{i=1}^{N}(\frac{S_i - \bar{S}}{\sigma S})(\frac{G_i - \bar{G}}{\sigma G})$

Note that 1) RMSE and MBE are in W m^{-2}; 2) rRMSE and rMBE are in percentage. where N is the number of samples; S_i, $\bar{S}$, and σS are the satellite SSR data, its mean and standard deviation, respectively; and G_i, $\bar{G}$, and σG are the *in situ* observed SSR, its mean and standard deviation, respectively.

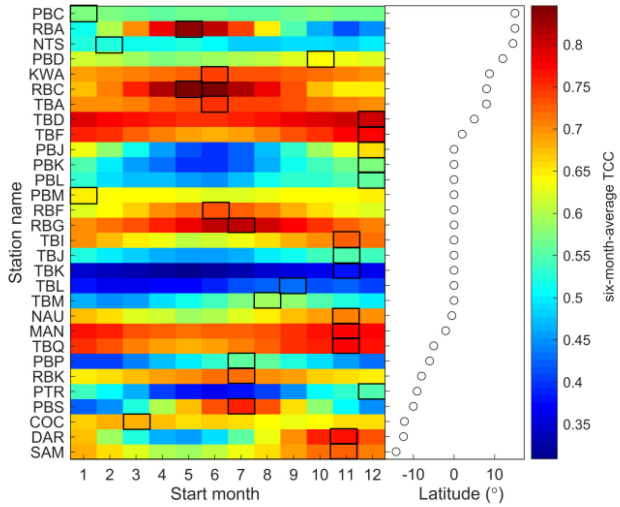


Fig. 2. Multiyear (2001–2017) mean of the six-month-average CERES TCC across *in situ* SSR stations between ±15° latitude. The horizontal axis in the left panel indicates the start month used for the six-month average, and the one corresponding to the maximum of the six-month-average TCC is marked with a black box. The right panel shows the latitudes of all stations within ±15° latitude.

D. Evaluation Methods

1) Preprocessing of in Situ *Observations:* The CERES-retrieved SSR is evaluated over the globe and MC by referring to global hourly and CMA daily *in situ* observations, respectively. The evaluation period spans from March 2000 to March 2018, for which the CERES satellite SSR products and the two *in situ* observations are available, and the global evaluation is conducted on an hourly, daily, and monthly basis, whereas the hourly evaluation is excluded over MC since *in situ* observations from CMA stations are available only at daily scales. The coarse temporal resolution *in situ* data are generated by simple arithmetic averaging of the original data since the hourly/daily data in a month are more than half for most months (see Fig. S3a and b). Note that the global hourly SSR data are the mean values for each hour, and the global hourly evaluation is only conducted toward the daytime records by excluding zero values of SSR at nighttime. Additionally, to improve the credibility of the evaluation results, all datasets were subjected to careful quality control, and the abnormal data and invalid data were removed prior to the evaluation.

2) Measurement of CERES Performance: During the evaluation, we followed Paech *et al.* [46] to compare the CERES gridded SSR to *in situ* observations, in which the pixels with the shortest distance between station locations and the center point of satellite pixels are selected for comparison with the *in situ* observations. In addition, five metrics, namely, root-mean-square error (RMSE), relative RMSE (rRMSE), mean bias error (MBE), relative MBE (rMBE), and correlation coefficient (R), are used to measure the performance of CERES SSR.

Fig. S3c depicts the number of valid observation-satellite SSR pairs. It can be found that the valid observation-satellite SSR pairs decrease as the temporal scale increases both in global evaluations and MC evaluations. In addition, for each independent evaluation, the valid observation-satellite SSR pairs are quite different. Despite this, the RMSE and MBE are valid for comparison because of the large sample numbers. Nevertheless, considering the effects caused by temporal scales as well as instrument differences in different independent evaluations, rRMSE and rMBE were introduced additionally to make those independent evaluations comparable. Table I lists the equations used to obtain the aforementioned five metrics. Furthermore, the grid-station comparison may result in subgrid heterogeneity-induced spatial representativeness issues at certain locations, which is further addressed in Sections III-C and IV-A.

3) Comparison Between Cloud-Season and Year-Round SSR: The presence of clouds may potentially impact the retrieval of CERES SSR [26], thus making the CERES product more questionable in cloudy seasons or areas. This issue is further investigated, in particular, through two approaches: 1) by illustrating the relationship between TCC and CERES SSR performance and 2) by comparing the year-round and cloudy-season evaluation results. For the first approach, the CERES TCC data (1°) are compared with the global monthly evaluation metric, i.e., rRMSE. For the second approach, different solutions are used to determine the cloudy seasons for areas beyond and between ±15° latitude. Specifically, for the former, summertime is commonly believed to be prone to a higher presence of clouds, and therefore June, July, and August for the Northern Hemisphere and December, January, and February for the Southern Hemisphere are considered cloudy seasons. The equatorial and intertropical zones, however, are dominated by the dry season and rainy season in which each last for approximately six months at most stations (see Fig. S4a). The presence of clouds is particularly an important sign of the rainy season (see Fig. S4b) with evidence that the six-month-average global precipitation measurement (GPM) precipitation is highly correlated with CERES TCC (see Fig. S5). Therefore, the six continuous months corresponding to the maximum of the multiyear average six-month-average monthly TCC are considered as cloudy seasons for stations within ±15° latitude (as black boxes in Fig. 2 show).

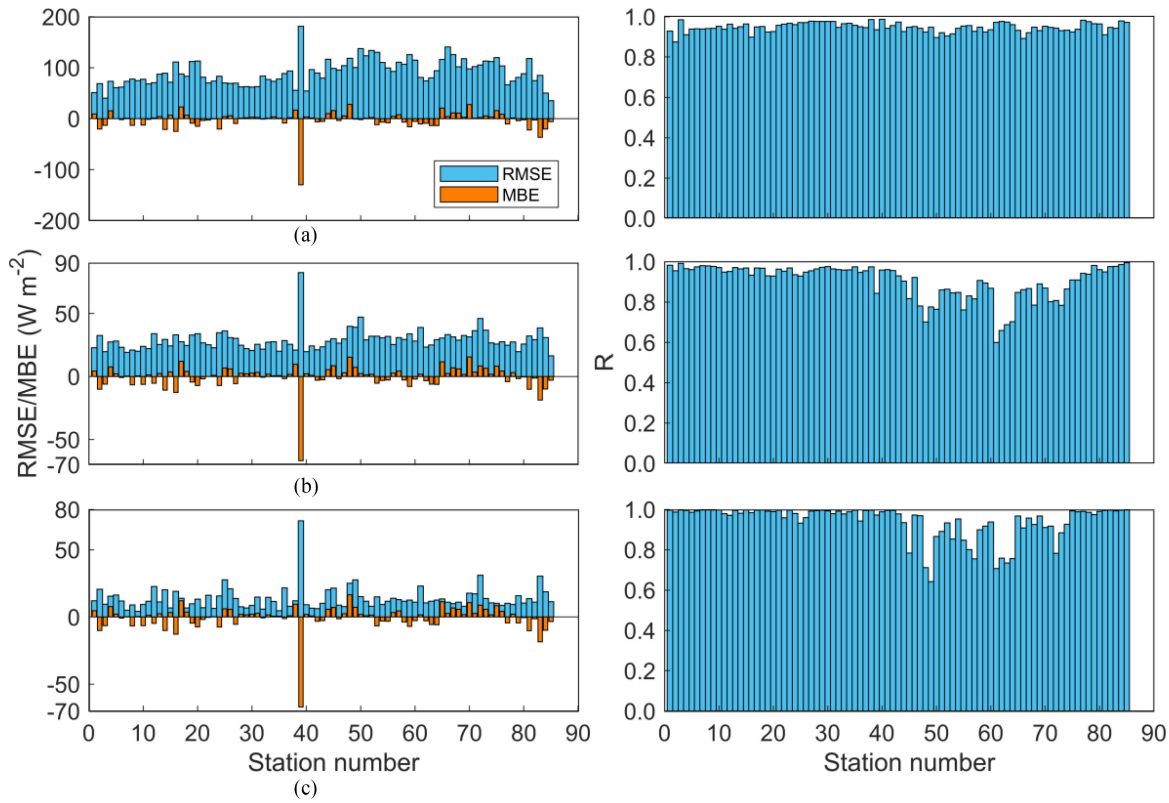


Fig. 3. Year-round evaluation of CERES satellite radiation products against CERES *in situ* observations from March 2000 to March 2018 over the globe. The left vertical axis denotes RMSE/MBE (W m^{-2}), while the right vertical axis denotes *R* values. The upper (a), Middle (b), and lower (c) Panels show the evaluation results for the global hourly, daily, and monthly temporal scales, respectively. For all subplots, the stations are ranked according to the latitude of the station (90°N–90°S) (same as their IDs listed in Table S1).

TABLE II
YEAR-ROUND COMPARISON OF CERES SATELLITE RADIATION PRODUCTS AND *IN SITU* OBSERVATIONS FROM MARCH 2000 TO MARCH 2018

Temporal scale	*(rRMSE) RMSE*	*(rMBE) MBE*	*R*	*N*
Global_hourly	(24.1)89.5	(0.8)-2.8	0.95	4532421
Global_daily	(15.6)29.6	(0.3)-0.6	0.96	370859
Global_monthly	(8.5)16.1	(0.4)-0.7	0.98	12449
MC_daily	(18.4)31.6	(3.6)6.2	0.93	378360
MC_monthly	(12.5)21.2	(2.8)4.8	0.94	17640

Note that 1) *N* indicates the total number of valid observation-satellite pairs across all stations; and 2) RMSE and MBE are in W m^{-2}, and rRMSE and rMBE are in percentage.

III. RESULTS

A. *SSR Evaluation With Global Hourly* in Situ *Observations*

The global evaluation of CERES-retrieved SSR is represented by statistical metrics of rRMSE, RMSE, rMBE, MBE, and *R* with respect to different temporal scales (hourly, daily, and monthly). The evaluation results are shown in Figs. 3 and 4. Overall, the (rRMSEs)RMSEs decrease as the temporal scales increase (see Table II; from (24.1%) 89.5 W m^{-2} to (8.5%) 16.1 W m^{-2}, Figs. 3 and 4), which is consistent with previous studies [26], [47]. The absolute quantity of (rMBEs)MBEs decrease slightly from hourly to daily (see Table II; from (0.8%) −2.8 W m^{-2} to (0.3%) −0.6 W m^{-2}) but there is no noticeable difference from daily to monthly scales (see Table II; from (0.3%) −0.6 W m^{-2} to (0.4%) −0.7 W m^{-2}). In addition, a lager MBE can be seen in coastal areas, on islands and in the polar region (see Fig. 4). In terms of *R* values, all temporal scales have a generally large value (≥0.95). For the independent evaluations at different temporal scales over the globe, some findings can be found, as follows.

1) For the hourly temporal scale, except for Station IZA (the 39th station in the hourly global *in situ* observations, see Table S1), which has generally larger RMSE (181 W m^{-2}) and MBE (−130 W m^{-2}), other stations yield RMSE varying between 35.1 W m^{-2} and 140.4 W m^{-2}, MBE −36.6 W m^{-2}–27.9 W m^{-2}, and *R* higher than 0.87 [see Fig. 3(a)]. Larger RMSE and MBE of IZA are also found at daily and monthly scales. This is further discussed in Section IV-A. In addition, as shown in Fig. 4(a), no apparent patterns are observed in terms of the spatial distribution of RMSE and *R* values. However, MBEs are mostly positive over oceans or coast lines and negative over land, which implies that CERES tends to underestimate SSR over oceans while overestimating over land. An overall performance based on comparisons across all stations with a total of 4532421 available observation-satellite SSR pairs at the hourly scale delivers RMSE = 89.5 W m^{-2},

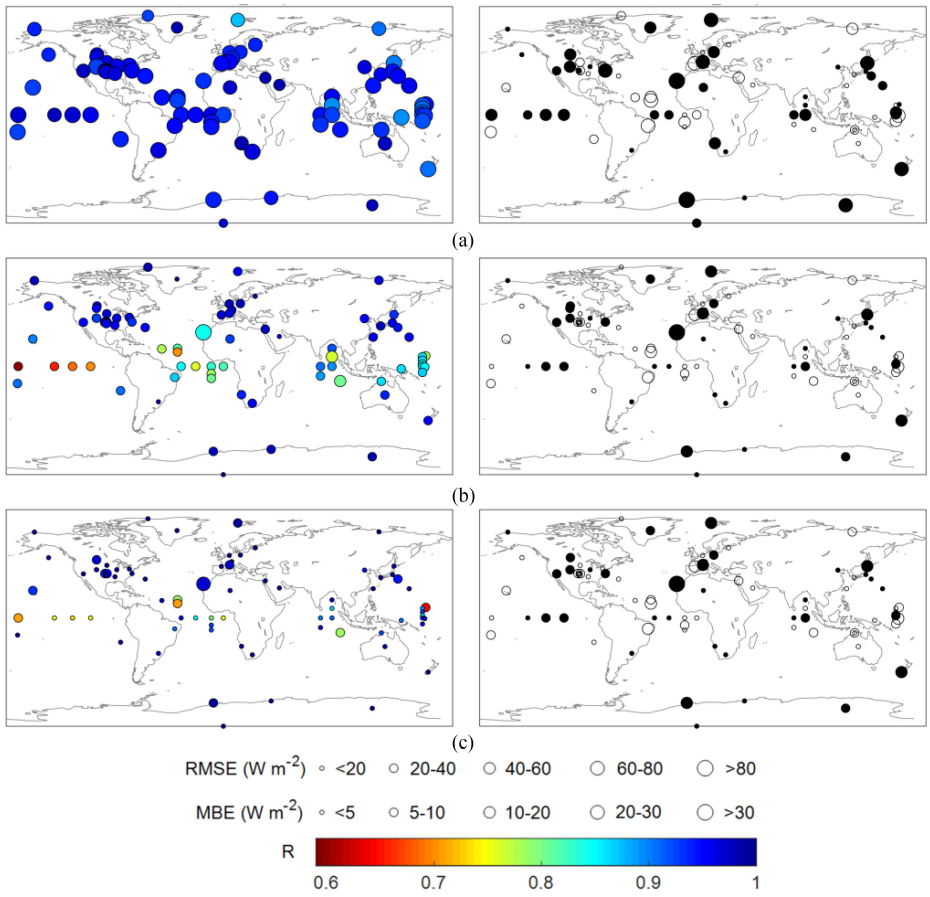


Fig. 4. Spatial distribution of evaluation results obtained from year-round global CERES SSR evaluation from March 2000 to March 2018. The left and right panels are RMSE/MBE (W m^{-2}) and R values. The upper (a), middle (b), and lower (c) panels show the evaluation results for the global hourly, daily, and monthly temporal scales, respectively. Note that empty and solid circles in the right panel each indicate positive and negative MBE values.

MBE $= -2.8$ W m^{-2}, and $R = 0.95$ (see Table II). In general, CERES Ed4A hourly data cannot satisfy the related scientific research requirements, as the RMSEs of all stations exceed 20 W m^{-2} [48].

2) The daily temporal scale (excluding Station IZA) yields a relatively lower RMSE (16.4 W m^{-2}–47.2 W m^{-2}) and MBE (-18.7 W m^{-2}–15.5 W m^{-2}) [see Fig. 3(b)] compared with the hourly scale. This can be attributed to the fact that the random errors that are frequently seen in the hourly data can be smoothed out through the daily average. In regard to R values, all stations, except for those between the 45th and 75th stations, exhibit higher R values (>0.9). This phenomenon is also observed in Fig. 4(b) and will be examined more in Section IV-B. Regarding the overall performance, the daily temporal scale with a total of 370 859 available observation–satellite SSR pairs delivers RMSE = 29.6 W m-2, MBE = -0.6 W m^{-2}, and $R = 0.96$ (see Table II).
3) Compared with the daily scale, the monthly scale delivers even more favorable consistency with a lower RMSE (4.2 W m^{-2}–31.0 W m^{-2}) and MBE (-18.5 W m^{-2}–16.5 W m^{-2}) [see Fig. 3(c)]. Only 14 stations have RMSE>20 W m^{-2} [see Fig. 3(c)], and most of these stations are near oceans or coast lines. The R values are slightly higher than those obtained from the daily scale but are still divergent across stations, with some approaching 1 and others being less than 0.90 [see Fig. 3(c)]. This can also be explained by the seasonal variations in SSR and smoother time series compared with the daily scale. Spatial patterns of RMSE, MBE, and R (see Fig. 4) echo these findings, and the overall performance with a total of 12449 available observation-satellite SSR pairs delivers RMSE = 16.1 W m^{-2}, MBE = -0.7 W m^{-2}, and $R = 0.98$ (see Table II). Generally, according to the Observing Systems Capability Analysis and Review Tool (see https://space.oscar.wmo.int/variables/view/downward_short_wave_irradiance_at_earth_surface) and Gupta *et al.* [49], the global monthly CERES SSR products almost satisfy the required criterion (RMSE$\leq$20 W m^{-2}) for global numerical weather prediction and climate monitoring.

In addition to the above, it is worth noting that: 1) all-station-based evaluation results for all temporal scales demonstrate slightly negative MBEs (-2.8, -0.6, and -0.7 W m^{-2} for hourly, daily, and monthly temporal scales, respectively), which indicates that CERES retrievals tend to slightly underestimate SSR over the globe. 2) A specific station named IZA shows poor consistency with CERES SSR across all temporal scales

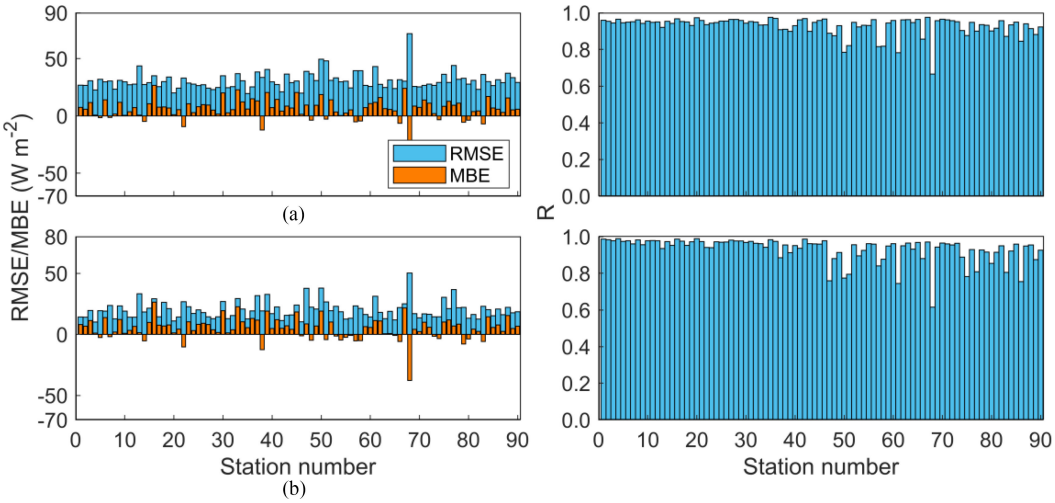


Fig. 5. Same as Fig. 3, but for evaluation results over the MC. The upper (a) and lower (b) panels show the evaluation results for the MC daily and monthly temporal scales, respectively.

(see Figs. 3 and 4). For the latter, a short discussion regarding the representativeness of single-station observations is presented in Section IV-A.

B. Regional SSR Evaluation With CMA in Situ *Observations*

Due to the availability of *in situ* SSR observations at the CMA stations, the evaluation of CERES SSR over MC is only conducted at daily and monthly temporal scales. In general, relative to global daily and monthly evaluation results, CERES SSR presents a slightly degraded performance with (rRMSE)RMSE, (rMBE)MBE, and R values of (18.4%) 31.6 W m^{-2}, (3.6%) 6.2 W m^{-2}, and 0.93 for the daily temporal scale and (12.5%) 21.2 W m^{-2}, (2.8%) 4.8 W m^{-2} and 0.94 for the monthly temporal scale, respectively (see Table II). Despite this, the overall RMSE (31.6 W m^{-2}) of the daily temporal scale falls in the normal range ($<$35 W m^{-2}) according to Huang *et al.* [16] and Wang *et al.* [48], and the overall performance at the monthly temporal scale (RMSE $=$ 21.2 W m^{-2}) is close to the minimum requirement of related research (RMSE$\leq$20 W m^{-2}). Contrary to the global evaluation results, CERES shows a significant overall overestimation in MC with MBE values of 6.2 W m^{-2} and 4.8 W m^{-2} for the daily and monthly temporal scales, respectively. This overestimation over MC is also found in other SSR products [50], [51]. The 68th station in Fig. 5, namely, #56385 in Table S2 and located in Southwest China, attracted our attention because it has a large RMSE and |MBE| and the smallest R value, indicating large discrepancies between *in situ* observations and satellite retrievals. A more in-depth discussion regarding this issue is presented in Section IV-B. Furthermore, a larger bias can be seen in Southwest China and Northwest China (see Fig. 6). This is probably because Southwest China is prone to prevalent clouds and complex terrain [52], and Northwest China is typically characterized by higher elevations and, hence, larger uncertainties in the retrieved CERES SSR. In terms of R value, they are mostly higher than 0.9 except for some low values observed mainly over the Southwest China (see Fig. 6), which may be attributed to the cloudy weather in this region. However, no significant law of decreasing R value with decreasing latitudes (see Fig. 6) is observed for daily and monthly temporal scales over MC, probably because of significant SSR seasonal variations beyond the Tropic of Cancer. This assumption is further confirmed in Section IV-B.

Specifically, for the daily temporal scale, among the 90 stations, 2 stations have RMSE below 20 W m^{-2}, while 69 stations have RMSE values between 20 and 35 W m^{-2} (see Fig. 5). In terms of MBE, it varies from $-$12 to 27 W m^{-2} (excluding station #56385). Out of all stations, there are 75 in total with positive MBEs, indicating remarkable overestimation in MC. For the monthly scales, the MBE varies between $-$12 W m^{-2} and 26 W m^{-2}, and more than half of the stations (52) possess RMSEs below 20 W m^{-2}, which suggests high consistency between the CERES satellite SSR and *in situ* SSR observations at the monthly temporal scale.

The abovementioned analysis reveals that special attention should be given to CERES satellite SSR in Northwest China and Southwest China.

C. Impacts of Clouds on CERES SSR

The impact of clouds on CERES SSR is investigated through two approaches. First, the SSR from cloudy seasons are evaluated and compared with the previously mentioned year-round evalutaions. The results are represented through barplots depicting the metric difference between year-round (yr) and cloudy-season (cs) evaluation (i.e., RMSE_d $=$ RMSE_yr-RMSE_cs; $R\_d = R\_yr - R\_cs$). In this case, the lower (larger) the RMSE_d (R_d), the greater the influence of clouds on the CERES SSR, and vice versa. As shown in Fig. 7, both the global and MC evaluation results suggest negative RMSE_d and positive R_d at most stations, and the magnitudes of RMSE_d and R_d decrease with increasing temporal scales. These results imply that clouds are indeed an important factor affecting the accuracy of CERES satellite SSR. The global results [see Fig. 7(a)–(c)] show that

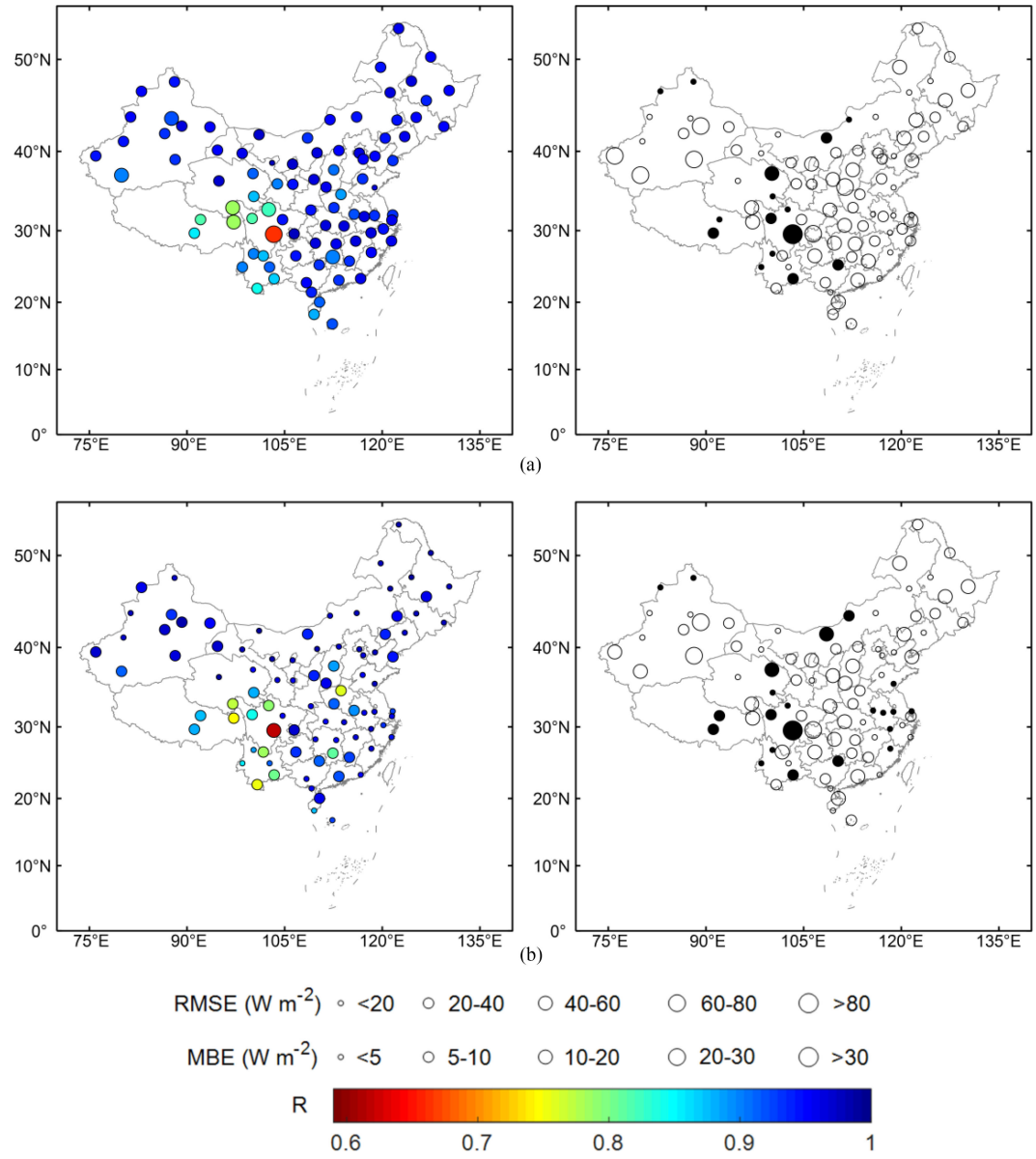


Fig. 6. Same as Fig. 4, but for evaluation results over the MC. The upper (a) and lower (b) panels show the evaluation results for the MC daily and monthly temporal scales, respectively.

the stations between the 40th and 70th stations have smaller magnitudes of RMSE_d and R_d than other stations (more obviously at daily and monthly temporal scales). This result is probably related to the fact that these locations are in the tropical zone (see Table S1) with less cloud variation in the time series, resulting in less difference between the year-round and cloudy season evaluations. Regarding the overall performance, the mean values of |RMSE_d| are 10.8, 5.8, and 3.4 W m^{-2} for global hourly, daily and monthly evaluations, respectively. The |RMSE_d|(|R_d|) averaged over all CMA stations is approximately 5.3 W m^{-2} (0.04) and 3.9 W m^{-2} (0.2) for daily and monthly scales, respectively. The significant negative daily R_d values over MC are observed at stations mainly located in Northwest and Southwest MC that are typically characterized by complex terrains and mountainous areas, which make SSR retrieval always a challenging issue throughout the year. Based on the abovementioned conclusions, it can be concluded that the effect of clouds on SSR is approximately 11, 6, and 4 W m^{-2} for hourly, daily, and monthly temporal scales, respectively.

A more straightforward way of illustrating the impact of clouds is to explore the relationship between cloud coverage and SSR evaluation metrics. In this regard, we rely on the multiyear average monthly CERES TCC (from March 2000 to March 2018) and the year-round monthly scale rRMSE to facilitate this investigation. As shown in Fig. 8, a larger rRMSE (>10%) is often associated with a much larger TCC (50%–100%). These findings, together with the above year-round and cloudy-season comparison, show that the presence of clouds impacts CERES SSR quality. The higher the cloud coverage is, the larger the uncertainties in the CERES retrievals. However, this uncertainty

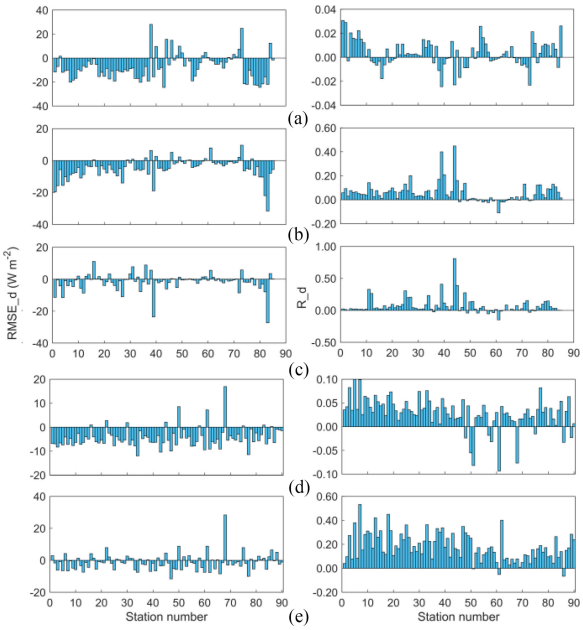


Fig. 7. Station-dependent evaluation metric differences (RMSE_d = RMSE_yr-RMSE_cs, and $R\_d = R\_yr\text{-}R\_cs$) between year-round (yr) and cloudy-season (cs) time periods from March 2000 to March 2018 over the globe and MC. The left vertical axis denotes RMSE/MBE (W m^{-2}), while the right vertical axis denotes R values. The upper three panels (a)–(c) show the results for the global hourly, daily, and monthly temporal scales, respectively. The lower two panels (d)–(e) show the results for the MC daily and monthly temporal scales, respectively. For all subplots, the stations are ranked according to the latitude of the station (90°N–90°S) (same as their IDs listed in Table S1). (a) Global_hourly. (b) Global_daily. (c) Global_monthly. (d) MC_daily. (e) MC_monthly.

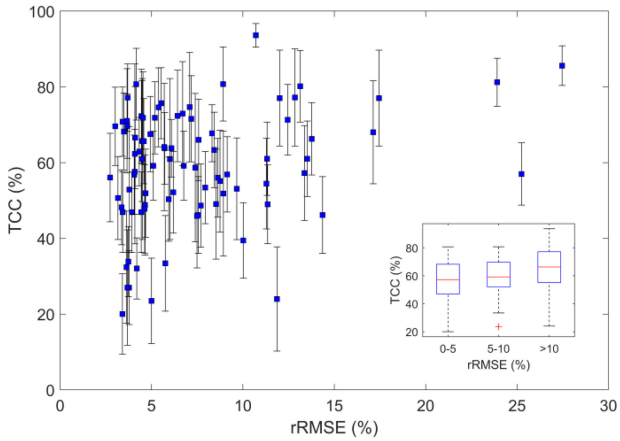


Fig. 8. Relationship between the multiyear average (March 2000 to March 2018) monthly CERES TCC and the year-round global rRMSE for the monthly temporal scale. The error bars denote the ±1 standard deviation of the CERES monthly TCC.

tends to decrease with increasing temporal scales. Therefore, more efforts should be made in retrieving satellite SSR in cloudy seasons.

D. Subgrid Heterogeneity and Spatial Representativeness of in Situ *Observations*

This article evaluates the CERES SSR by comparing the corresponding satellite gridded retrievals with *in situ* observations. However, for most hydrometeorological variables, spatial heterogeneity, and thus, the representativeness of *in situ* observations are of great concern across different atmospheric science studies [53], [55]. Here, the daily CM SAF SSR product (SARAH-E 1.1: March 2000–December 2016; SARAH 2.1: March 2000–December 2017) with a relatively smaller spatial extent but higher spatial resolution (0.05° × 0.05°) was applied to verify whether *in situ* station observations can represent the 1° CERES grid SSR through subgrid heterogeneity analysis. As CM SAF coverage does not include all 85 global stations (see Fig. 1), only the grids corresponding to the stations covered by CM SAF were analyzed.

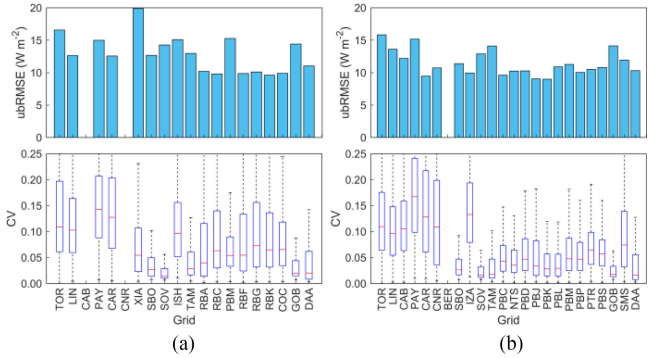


Fig. 9. Comparison between spatially upscaled CM SAF and CERES satellite SSR in terms of ubRMSD over a 1° × 1° CERES grid (shown as barplots; upper panel) and subgrid heterogeneity as measured by CV among all 20 × 20 CM SAF grids collocated in the 1° CERES grid (shown as boxplots; lower panel) for all stations located in the SARAH-E 1.1 (from March 2000 to December 2016; left panel) and SARAH 2.1 (March 2000 to December 2017; right panel) coverages. Note that some of the barplots and boxplots are left blank at certain stations where CM SAF SSR retrievals are missing.

Prior to the subgrid analysis, the rationality of using the CM SAF product is examined by comparing spatially average CM SAF data (20 × 20 CM SAF grids within the CERES 1° grid, hereafter upscaled CM SAF data) with the CERES satellite SSR. Since SSR from different sensors is not necessarily consistent with each other, the unbiased root-mean-square difference (ubRMSD, [56]) is used to measure the applicability regarding the random error only. It is expressed by the following equation:

$$\text{ubRMSD} = \sqrt{\text{RMSD}^2 - \text{MBD}^2} \tag{1}$$

where RMSD and mean Bias deviation (MBD) refer to the root mean square deviation and mean bias deviation, respectively, between the upscaled CMSAF data and the CERES satellite SSR. The RMSD and MBD are calculated in the same way as the RMSE and MBE listed in Table I. Here, "deviation (D)" is used instead of "error (E)" because there is no ground truth available.

As shown in the upper panel in Fig. 9, most collocated grids yield ubRMSDs less than 15 W m^{-2}, which are overall smaller than the global-daily SSR-evaluation-obtained RMSE values [varying from 20 W m^{-2} to 40 W m^{-2} at most stations; see Section III-A and Fig. 3(b)]. Hence, the use of a finer resolution of CM SAF for representing subgrid spatial heterogeneity is acceptable.

The spatial representativeness of *in situ* observations is indicated by the coefficient of variation (CV) of daily SSR among all 20 × 20 CM SAF grids collocated in the 1° CERES grid. The calculation formula for CV is as follows:

$$\mathrm{CV}_i = \frac{\sigma C_i}{\overline{C_i}} \tag{2}$$

where σC_i and $\overline{C}_i$ denote the standard deviation and mean values of 20 × 20 CM SAF grids collocated in the 1°CERES grid at the *i*th time series, respectively.

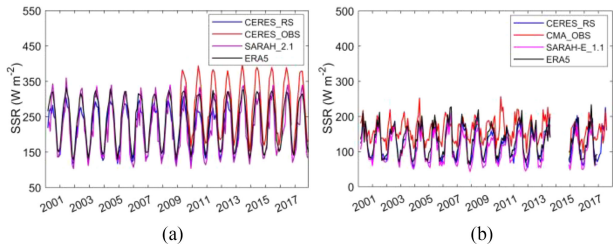
As seen from the lower panel in Fig. 9, most grids yield CVs less than 0.05, indicating relatively small spatial variability within the CERES grids. This agrees with the findings of Schwarz *et al.* [31], which suggest that the correlation between the grid-average SSR in the 1° grid and corresponding *in situ* observations can be up to 0.8. This further implies that the station-to-grid comparison should have induced limited uncertainties in the CERES SSR evaluation. Therefore, the aforementioned evaluation scheme is reasonably valid.

IV. Discussion

The abovementioned global and regional evaluation of CERES SSR indicates diverse consistencies compared with *in situ* observations, and larger RMSE and MBE values are seen at some stations. For example, larger MBEs can be seen in coastal areas, on islands and in polar regions (see Fig. 4). This may be attributed to large uncertainties in the inputs for CERES SSR retrieval, i.e., the great uncertainties for the MODIS aerosols over coastal or island stations and failure detection of clouds over the polar region [57], [58]. In addition, larger RMSE and MBE of IZA were found at all temporal scales (see Fig. 3). In most cases, these large inconsistencies with *in situ* observations may be due to 1) the satellite retrieval challenges related to the large spatial heterogeneity caused by clouds, surface conditions, etc. [59]–[62]; and 2) erroneous observations or flaws in *in situ* measurements. In general, the large inconsistency between CERES satellite SSR and *in situ* SSR observations may be due to 1) uncertainties in both *in situ* instruments and the remote sensing-retrieved SSR and 2) spatial mismatching between pixel-scale (1° × 1° in this article) satellite retrievals and point-scale station observations as induced by the latter's limited spatial representativeness, which was quantified in Section III-D. The following addresses this issue by focusing on the uncertainty at specific stations. In addition, *R* is found to decrease as the stations approach the equator at daily and monthly temporal scales over the globe. This latitudinal variation in *R* is also briefly discussed.

A. *Uncertainties of* in Situ *Measurements at Specific Stations*

As Figs. 3 and 5 show, stations named IZA in western Africa (see Fig. S1) and #56385 in Sichuan, China (see Fig. S2), each corresponding to the 39th and 68th stations in the CERES and CMA collection of *in situ* measurements (see Tables S1 and S2), stand out due to their extremely large RMSE and MBE values (see Figs. 3–6). By using additional independent SSR data from CM SAF retrieval and ERA5 reanalysis, the time series for Stations IZA and #56385 were conducted (see Fig. 10). For both IZA and #56385, it is found that the *in situ* observed SSR in general positively deviates from the other three products. According to Figs. S1–S2 and Tables S1–S2, Station IZA is located on a small island in the Atlantic with an elevation of 2737 m and is often surrounded by a stratocumulus sea below the station [63], whereas Station #56385 lies in the Emei Mountains with an even higher elevation of 3047 m. Under such geographic conditions, when solar radiation reaches the stations at high elevation and, thus, relatively shallower atmospheric optical depth, the sensors equipped on the *in situ* stations receive more radiance and, therefore, larger *in situ* SSR observations. In contrast, satellite retrieval is a measure of the pixel average, which exhibits a relatively lower elevation and, thus, less SSR. In short, geolocations of *in situ* SSR stations may remarkably impact spatial representativeness. In addition, the satellite retrieval challenges related to large spatial heterogeneity caused by clouds, surface conditions, etc., may be attributed to large inconsistencies with *in situ* observations. Moreover, inaccurate measurements can also potentially lead to large RMSE and MBE, and one should always be cautious when referring to station observations as "ground truth" for the evaluation of satellite retrievals and model outputs.



(a) (b)

Fig. 10. Monthly average time series of SSR across Stations CERES_IZA (obs_IZA) and CMA_56385 (obs_56385) and their corresponding satellite pixels and/or model grids from March 2000 to December 2017. For each subplot, "CERES_RS", "SARAH_2.1"/"SARAH-E_1.1", and "ERA5" each denote the 1°CERES retrieval, the 0.05°CM SAF SARAH 2.1/SARAH-E 1.1 product, and the 0.1° ERA5 reanalysis, respectively.

B. *Seasonal Variations in SSR and its Impact on Global Evaluation*

Global daily and monthly SSR evaluations suggested lower *R* values around the equator and higher values beyond the tropics. Here, we better illustrate this finding by varying *R* values from the daily evaluation with increasing latitudes. Note that the absolute value of latitude is actually used to denote the deviation of stations from the equator.

In a statistical sense, as seen in Fig. 11(a), *R* values appear low around the equator by varying from 0.65 to 0.9, then quickly increase to ∼0.9 when approaching the edge of the tropics near 30°N and 30°S, and then remaining high at a constant value of beyond 0.9. Statistics from the monthly scale deliver similar results (not shown). The hourly scale, however, behaves totally differently, with *R* values maintained at a high value at all latitudes, as seen in Figs. 3(a) and 4(a). This is mainly because both the station-observed and satellite-retrieved SSR naturally

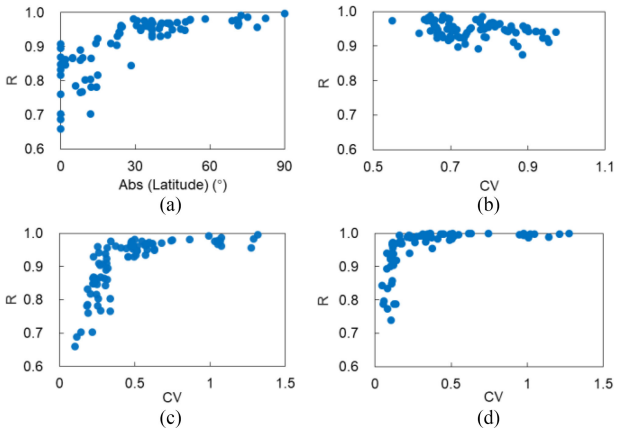


Fig. 11. Relationship between R values and (a) abs (latitude) for the daily temporal scale and CV of the SSR time series for the (b) hourly, (c) daily, and (d) monthly temporal scales over 85 global stations. (a) Global_daily. (b) Global_hourly. (c) Global_daily. (d) Global_monthly.

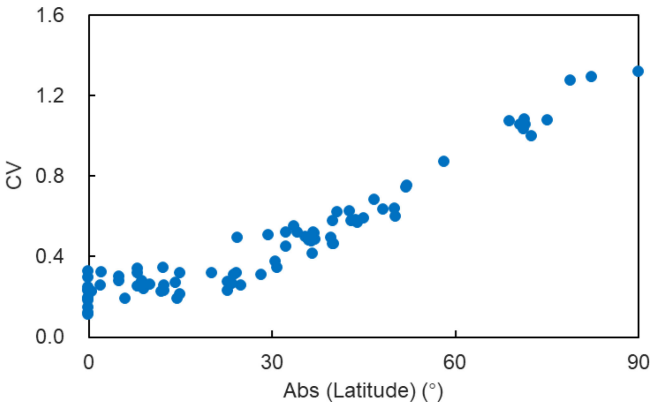


Fig. 12. Latitudinal evolution of the seasonal variation in SSR, as represented by the coefficient of variation (CV). Shown are the results from the global daily *in situ* observations.

exhibit significant diurnal variations, which dominate the correlation calculation at the hourly temporal scale and there by higher R values. However, at the daily and/or monthly temporal scale, the correlation is more determined by the seasonal variations in SSR. Taking the global daily *in situ* SSR as an example and expressing the seasonal variation in SSR with the CV of the time series, it is obvious that the CV of the SSR generally increases poleward and remains at a relatively low and stable value (<0.4) within $\pm 15°$ latitude (see Fig. 12). The latter may result in low R values [as seen in Figs. 3(b) and 4(b)] and further supports the use of $\pm 15°$ to characterize the equatorial zone, which has no significant seasonal variations in SSR. Note that CV is calculated by (2) listed in Section III-D, where i denotes the station number and σC_i and $\overline{C}_i$ denote the standard deviation and mean values of the SSR time series of the ith station, respectively. The monthly SSR behaves similarly and is not shown for conciseness.

Furthermore, the abovementioned interpretation is verified by evolving R values with CVs at all 85 global stations. As shown in Fig. 11(b)–(d), while R values exhibit no apparent variations at the hourly scale, they are found to vary significantly with CVs at daily and monthly scales–low R value with less CV (i.e., less seasonal variation) and high R value with higher CV (i.e., larger seasonal variations), implying that the correlation between station-observed SSR and CERES satellite retrievals is largely impacted by seasonal variations.

TABLE III
GLOBAL LAND AND OCEAN EVALUATION RESULTS
(MARCH 2000–MARCH 2018)

Temporal scale	Area	*(rRMSE)* *RMSE*	*(rMBE)* *MBE*	R	N
Global_hourly	Land	(23.9)82.7	(0.7)-2.6	0.96	3435031
	Ocean	(23.5)104.1	(0.3)1.3	0.95	1057387
Global_daily	Land	(15.9)28.1	(0.3)-0.5	0.96	281991
	Ocean	(13.4)30.5	(0.8)1.7	0.90	85633
Global_monthly	Land	(8.3)14.7	(0.4)-0.7	0.99	9410
	Ocean	(6.4)14.6	(0.7)1.6	0.96	2931

Note that 1) N indicates the total number of valid observation-satellite pairs across all stations; and 2) RMSE and MBE are in W m^{-2}, and rRMSE and rMBE are in percentage.

C. Comparison Between Global and Land Evaluations

Aided by global hourly *in situ* observations, land and ocean evaluations were further investigated, as shown in Table III . Note that Station IZA was excluded from global land evaluations because of its larger RMSE and MBE compared with other stations. Overall, for both land and ocean evaluations, rRMSE and RMSE decrease as the temporal scale increases, which is consistent with the findings of the global evaluation. When the MBE is calculated independently for the 58 global land stations and the 27 ocean stations, it is found to be negative on land and positive over the ocean at all temporal scales. Specifically, the MBEs are -2.6, -0.5, and -0.7 W m^{-2} for global land hourly, daily, and monthly evaluations and 1.3, 1.7, and 1.6 W m^{-2} for the global ocean hourly, daily, and monthly evaluations, respectively. This indicates that the CERES Ed4A SSR tends to underestimate the SSR over land while overestimating the SSR over the ocean. Fig. 4 further confirms this conclusion. Aided by ship cruise hourly data, Hinkelman *et al.* [23] also found an overestimation of CERES Ed4A with a bias of 21.5 W m^{-2} on Macquarie Island, Southern Ocean. In addition, the performance of CERES is better in the ocean than on land at all temporal scales, with rRMSE values of 23.9%, 15.9%, and 8.3% for global land hourly, daily, and monthly evaluations and 23.5%, 13.4%, and 6.4 for global ocean hourly, daily, and monthly evaluations, respectively.

D. Comparison With Previous Studies

Some studies have also evaluated the CERES SYN1deg Ed4A SSR product with regard to different spatiotemporal extents, as shown in Table IV and are briefly summarized below.

TABLE IV
LIST OF CERES SYN1DEG ED4A SSR PRODUCTS IN THE LITERATURE

Spatial coverage	Temporal coverage	Temporal resolution	Statistical Metrics	Reference
The Arctic	2007.1-2017.12	3-hourly	*R, RMSE, MBE, MAE and MBE*: 0.92, 53.9, 5.2 and 28.3	Sun et al., 2018
Macquarie Island, the Southern Ocean	2016.4-2018.3	hourly annual mean	hourly: *MBE* 21.5 annual mean: *MBE* 10	Hinkelman et al., 2020
Three poles	—	1-hourly 3-hourly daily monthly	3-hourly:(*R, (rRMSE)RMSE and (rMBE)MBE*) Ar: 0.89, (30.9)54.0 and (1.9)-3.4 An: 0.92, (23,0)59.4 and (1.7)-4.4 TP: 0.87, (29.5)113 and (7.6)-29.1 daily: Ar: 0.86, (20.5)35.1 and (1.6)2.8 An: 0.90, (5.1)39.5 and (1.8)-4.8 TP: 0.76, (17.8)39.7 and (7.6)-17.1	Wang et al., 2021
Global land	2008-2014	daily	*R, (rRMSE)RMSE and (rMBD)MBD:* 0.92, (17.0)27.6 and (1.3)2.2	Li et al., 2021

Note that 1) RMSE, MBE, MAE (mean absolute error) and mean bias deviation (MBD) are in W m^{-2}; 2) rRMSE and rMBE are in percentage; and 3) Ar, An, and TP denote arctic, antarctic and qinghai–tibet plateau, respectively.

Hinkelman *et al.* [23] conducted an evaluation of Macquarie Island, Southern Ocean, by using surface observations collected by the Macquarie Island Cloud and Radiation Experiment from March 2000 to March 2018. They found a positive bias of 21.5 W m^{-2} during daylight for the hourly temporal scale, which was also suggested by most stations over oceans or coastlines in the global hourly evaluation of our study [see Figs. 3(a) and 4(a)].

Sun *et al.* [24] evaluated the CERES Ed4A product over the Arctic region based on 23 *in situ* station observations from BSRN, CEOP (Coordinated Energy and Water Cycle Observations Project) and GC-Net (Greenland Climate Network) in 2007 at 3-h temporal scale. Wang *et al.* [25] further evaluated the CERES Ed4A by using *in situ* observations from BSRN, CEOP, TPDC-QTP (National Tibetan Plateau Data Center- Qinghai–Tibet Plateau), and NMC (Nam Co Watershed) at hourly, 3-h, daily, and monthly mean temporal scales over the three poles (Arctic, Antarctic, and Qinghai–Tibet Plateau). In the Arctic, both studies reported poor satellite SSR performance at the 3-h temporal scale, with RMSE of 53.9 W m^{-2} for the former and 54.0 W m^{-2} for the latter. This poor performance is also found in our study, with an RMSE of 71.1 W m^{-2} at the hourly temporal scale. In the Antarctic, Wang *et al.* [25] revealed that the *R*, (rRMSE)RMSE and (rMBE)MBE were 0.90, (5.1%) 39.5 W m^{-2}, and (1.8%) −4.8 W m^{-2} for the daily temporal scale. The BSRN Stations DOM, GVN, SPO, and SYO used in this article are all considered in this article, and the daily mean values of *R*, (rRMSE)RMSE and (rMBE)MBE derived in our work are 0.98, (21.53%) 28.7 W m^{-2}, and (6.1%) −8.2 W m^{-2}, respectively. CERES Ed4A SSR in both Arctic and Antarctic regions has a slightly larger uncertainty, which may be attributed to the difficulty in distinguishing between clouds and surface snow cover [15]. The Qinghai–Tibet Plateau was demonstrated to possess greater bias than polar regions with larger RMSE and MBE and smaller *R* values at different temporal scales [25]. This further supported the finding in Section III-B that there is a larger bias in Southwest China (see Fig. 6).

Moreover, the CERES Ed4A SSR product was examined at the global scale (land only) against 142 *in situ* observations at the daily temporal scale over the period 2008–2014 by Li *et al.* [26]. Their results showed that the *R* value and (rRMSE)RMSE were 0.92 and (17.0%) 27.6 W m^{-2} for the whole globe, respectively. The daily evaluation results over the globe (land only) in this research are similar to Li *et al.* [26], with *R* and (rRMSE)RMSE values of 0.96 and (15.9%) 28.1 W m^{-2}, respectively.

Overall, the findings of the global evaluation of the CERES Ed4A SSR product in this article are in line with previous studies. Readers are encouraged to refer to the abovementioned references (also listed in Table IV) for more details.

CONCLUDING REMARKS

In this article, we evaluated the year-round and cloudy-season CERES SYNIdeg Ed4A solar radiation products from March 2000 to March 2018 with regard to different temporal scales (hourly, daily, and monthly average) over the globe (land and ocean) and the MC using global hourly and CMA daily ground observations.

In general, CERES SSR yields $|\text{MBE}| \leq 2.3$ (6.2) W m^{-2}, $\text{RMSE} \leq 90.5$ (31.8) W m^{-2}, and $R \geq 0.95$ (0.93) compared with ground station observations for all temporal scales over the globe (MC). The discrepancies between satellite products and *in situ* observations decrease with increasing temporal scales, and the monthly SSR almost meets the minimum requirements of related scientific studies ($\text{RMSE} \leq 20$ W m^{-2}) across the globe. The former might be attributed to a decrease in the error of station representativeness as the time scale increases [64]. In addition, evaluations conducted over the land and ocean suggested that CERES Ed4A tends to underestimate the SSR over land while overestimating SSR over the ocean.

Degraded performance is found in the cloudy seasons of the CERES SSR product, which indicates that clouds may potentially impact the retrieval accuracy, implying the necessity of better representing clouds during the remote sensing of surface shortwave radiation. The subgrid spatial heterogeneity analysis suggested relatively small spatial variability with the CERES grids ($\text{CV} < 0.05$ for most stations). However, the subgrid cloud effects are a part of the spatial heterogeneity, and how to separate them still needs to be investigated in future studies. With

respect to the seasonal variation analysis, the results showed that the correlation coefficient between station observations and satellite SSR tends to be smaller near the equator due to the small seasonal variation in solar radiation. In addition, while the mismatching between point-scale station measurements and satellite pixel-scale retrievals is reasonably handled at most stations, some specific ones (i.e., Stations IZA and #56385) deployed at particular locations and mountainous areas may hold great uncertainties and, thus, need to be used with great caution in future evaluations.

Data Availability Statement

Availability of all the data used in this article are listed as follows.

The CERES SYN1deg Ed4A SSRs and TCC products, and the global hourly *in situ* observed SSR are obtained from the CERES project; This dataset is also from the CERES project, and therefore the online URL[4] CMA daily station observations are provided by CMA Meteorological Information Center.[5] CM SAF SARAH-E 1.1 and SARAH 2.1 SSR are available through CM SAF Surface radiation products of Climate Data Records;[6] ECMWF ERA5-Land SSR is accessible from the European Centre for Medium Range Weather Forecasts Datasets.[7]

GPM IMERG final run are accessible from the EarthData.[8]

References

[1] M. J. Al-Khawaja, "Determination and selecting the optimum thickness of insulation for buildings in hot countries by accounting for solar radiation," *Appl. Thermal Eng.*, vol. 24, no. 17/18, pp. 2601–2610, Dec. 2004.

[2] S. Chen *et al.*, "The potential of photovoltaics to power the belt and road initiative," *Joule*, vol. 3, no. 8, pp. 1895–1912, Aug. 2019.

[3] G. Hoogenboom, "Contribution of agrometeorology to the simulation of crop production and its applications," *Agricultural Forest Meteorol.*, vol. 103, no. 1/2, pp. 137–157, Jun. 2000.

[4] B. Sweerts, S. Pfenninger, S. Yang, D. Folini, B. van der Zwaan, and M. Wild, "Estimation of losses in solar energy production from air pollution in China since 1960 using surface radiation data," *Nat. Energy*, vol. 4, no. 8, pp. 657–663, Jul. 2019.

[5] H. Jiang, N. Lu, G. Huang, L. Yao, J. Qin, and H. Liu, "Spatial scale effects on retrieval accuracy of surface solar radiation using satellite data," *Appl. Energy*, vol. 270, no. 11, Jul. 2020, Art. no. 115178.

[6] T. Wang, G. Yan, X. Mu, Z. Jiao, L. Chen, and Q. Chu, "Toward operational shortwave radiation modeling and retrieval over rugged terrain," *Remote Sens. Environ.*, vol. 205, pp. 419–433, Feb. 2018.

[7] S. J. Reddy, "An empirical method for estimating sunshine from total cloud amount," *Sol. Energy*, vol. 15, no. 4, pp. 281–285, Apr. 1974.

[8] J. S. G. Ehnberg and M. H. J. Bollen, "Simulation of global solar radiation based on cloud observations," *Sol. Energy*, vol. 78, no. 2, pp. 157–162, Feb. 2005.

[4][Online]. Available: satellite SSR: https://ceres-tool.larc.nasa.gov/ord-tool/jsp/SYN1degEd41Selection.jsp; [Online]. Available: satellite TCC: https://ceres-tool.larc.nasa.gov/ord-tool/jsp/SYN1degEd41Selection.jsp; *in situ* observations: https://ceres-tool.larc.nasa.gov/ord-tool/jsp/CAVE41Selection.jsp *in situ* observations: https://ceres-tool.larc.nasa.gov/ord-tool/jsp/CAVE41Selection.jsp

[5][Online]. Available: http://data.cma.cn/data/cdcdetail/dataCode/RADI_MUL_CHN_DAY.html

[6][Online]. Available: https://wui.cmsaf.eu/safira/action/viewProduktList?dId=2&d-1342877-p=4

[7][Online]. Available: SSR: https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land-monthly-means?tab=overview

[8][Online]. Available: https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGM_06/summary?keywords=%22IMERG%20final%22

[9] K. L. Bristow and G. S. Campbell, "On the relationship between incoming solar radiation and daily maximum and minimum temperature," *Agricultural Forest Meteorol.*, vol. 31, no. 2, pp. 159–166, May 1984.

[10] K. Yang, G. W. Huang, and N. Tamai, "A hybrid model for estimating global solar radiation," *Sol. Energy*, vol. 70, no. 1, pp. 13–22, 2001.

[11] K. Yang and T. Koike, "A general model to estimate hourly and daily solar radiation for hydrological studies," *Water Resour. Res.*, vol. 41, no. 10, pp. 1–13, Oct. 2005.

[12] K. Yang, T. Koike, and B. Ye, "Improving estimation of hourly, daily, and monthly solar radiation by importing global data sets," *Agricultural Forest Meteorol.*, vol. 137, no. 1/2, pp. 43–55, Mar. 2006.

[13] W. Tang *et al.*, "Reconstruction of daily photosynthetically active radiation and its trends over China," *J. Geophys. Res. Atmos.*, vol. 118, no. 23, pp. 13292–13302, Nov. 2013.

[14] W. Tang, K. Yang, J. Qin, M. Min, and X. Niu, "First effort for constructing a direct solar radiation data set in China for solar energy applications," *J. Geophys. Res. Atmos.*, vol. 123, no. 3, pp. 1724–1734, Jan. 2018.

[15] W. Tang, K. Yang, J. Qin, X. Li, and X. Niu, "A 16-year dataset (2000-2015) of high-resolution (3 h, 10 km) global surface solar radiation," *Earth Syst. Sci. Data*, vol. 11, no. 4, pp. 1905–1915, Dec. 2019.

[16] G. Huang *et al.*, "Estimating surface solar irradiance from satellites: Past, present, and future perspectives," *Remote Sens. Environ.*, vol. 233, Nov. 2019, Art. no. 111371.

[17] B. R. Barkstrom, "The earth radiation budget experiment (ERBE)," *Bull. - Amer. Meteorol. Soc.*, vol. 65, no. 11, pp. 1170–1185, Nov. 1984.

[18] P. W. Stackhouse, Jr, S. K. Gupta, S. J. Cox, T. Zhang, J. C. Mikovitz, and L. M. Hinkelman, "The NASA/GEWEX surface radiation budget release 3.0: 24.5-year dataset," *GEWEX News*, vol. 21, no. 1, pp. 10–12, Feb. 2011.

[19] H. Jacobowitz and R. J. Tighe, "The Earth radiation budget derived from the nimbus 7 ERB experiment," *J. Geophys. Res.*, vol. 89, no. D4, pp. 4997–5010, Jun. 1984.

[20] Y. Zhang, C. N. Long, W. B. Rossow, and E. G. Dutton, "Exploiting diurnal variations to evaluate the ISCCP-FD flux calculations and radiative-flux-analysis-processed surface observations from BSRN, ARM, and SURFRAD," *J. Geophys. Res. Atmos.*, vol. 115, no. D15, pp. 1–21, Aug. 2010.

[21] X. Zhang *et al.*, "An operational approach for generating the global land surface downward shortwave radiation product from MODIS data," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4636–4650, Jul. 2019.

[22] B. A. Wielicki, B. R. Barkstrom, E. F. Harrison, R. B. Lee, G. L. Smith, and J. E. Cooper, "Clouds and the Earth's radiant energy system (CERES): An earth observing system experiment," *Bull. Amer. Meteorol. Soc.*, vol. 77, no. 5, pp. 853–868, May 1996.

[23] L. M. Hinkelman and R. Marchand, "Evaluation of CERES and cloudsat surface radiative fluxes over Macquarie Island, the Southern Ocean," *Earth Space Sci.*, vol. 7, no. 9, Sep. 2020, Art. no. e2020EA001224.

[24] D. Sun, C. Ji, W. Sun, Y. Yang, and H. Wang, "Accuracy assessment of three remote sensing shortwave radiation products in the Arctic," *Atmos. Res.*, vol. 212, pp. 296–308, Nov. 2018.

[25] G. Wang, T. Wang, and H. Xue, "Validation and comparison of surface shortwave and longwave radiation products over the three poles," *Int. J. Appl. Earth Observ.*, vol. 104, no. 15, Dec. 2021, Art. no. 102538.

[26] R. Li, D. Wang, and S. Liang, "Comprehensive assessment of five global daily downward shortwave radiation satellite products," *Sci. Remote Sens.*, vol. 4, Dec. 2021, Art. no. 100028.

[27] Z. Li *et al.*, "East Asian studies of tropospheric aerosols and their impact on regional climate (EAST-AIRC): An overview," *J. Geophys. Res.*, vol. 116, no. D7, Feb. 2011, Art. no. D00K34.

[28] H. Che *et al.*, "Analyses of aerosol optical properties and direct radiative forcing over urban and industrial regions in Northeast China," *Mrteorol. Atmos. Phys.*, vol. 127, no. 3, pp. 345–354, Jan. 2015.

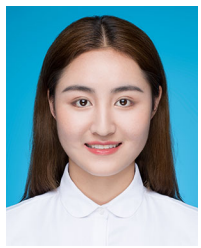
[29] B. Mai *et al.*, "Aerosol optical properties and radiative impacts in the pearl river delta region of China during the dry season," *Adv. Atmos. Sci.*, vol. 35, no. 2, pp. 195–208, Jan. 2018.

[30] L. Yu *et al.*, "Aerosol radiative effects from observations and modelling over the Yangtze River Basin, China from 2001 to 2015," *Int. J. Climatol.*, vol. 39, no. 8, pp. 3476–3491, Feb. 2019.

[31] M. Schwarz, D. Folini, M. Z. Hakuba, and M. Wild, "Spatial representativeness of surface-measured variations of downward solar radiation," *J. Geophys. Res. Atmos.*, vol. 122, no. 24, pp. 13–319, Nov. 2017.

[32] D. R. Doelling *et al.*, "Geostationary enhanced temporal interpolation for CERES flux products," *J. Atmos. Ocean Technol.*, vol. 30, no. 6, pp. 1072–1090, Jun. 2013.

[33] S. Kato and N. G. Loeb, "Twilight irradiance reflected by the Earth estimated from clouds and the Earth's radiant energy system (CERES) measurements," *J. Climate*, vol. 16, no. 15, pp. 2646–2650, Aug. 2003.
[34] D. A. Rutan *et al.*, "CERES synoptic product: Methodology and validation of surface radiant flux," *J. Atmos. Ocean Technol.*, vol. 32, no. 6, pp. 1121–1143, Jun. 2015.
[35] D. A. Rutan, F. G. Rose, N. M. Smith, and T. P. Charlock, "Validation data set for CERES surface and atmospheric radiation budget (SARB)," *WCRP/GEWEX News*, vol. 11, no. 1, pp. 11–12, Jan. 2001.
[36] A. Driemel *et al.*, "Baseline surface radiation network (BSRN): Structure and data description (1992–2017)," *Earth Syst. Sci. Data*, vol. 10, no. 3, pp. 1491–1501, Aug. 2018.
[37] M. F. Cronin, C. W. Fairall, and M. J. McPhaden, "An assessment of buoy-derived and numerical weather prediction surface heat fluxes in the tropical pacific," *J. Geophys. Res. Oceans*, vol. 111, no. C6, p. C06038, Jun. 2006.
[38] G. Y. Shi *et al.*, "Data quality assessment and the long-term trend of ground solar radiation in China," *J. Appl. Meteorol. Climatol.*, vol. 47, no. 4, pp. 1006–1016, Apr. 2008.
[39] W. Tang, K. Yang, J. He, and J. Qin, "Quality control and estimation of global solar radiation in China," *Sol. Energy*, vol. 84, no. 3, pp. 466–475, Mar. 2010.
[40] K. G. Karlsson *et al.*, "CLARA-A2: The second edition of the CM SAF cloud and radiation data record from 34 years of global AVHRR data," *Atmospheric Chem. Phys.*, vol. 17, no. 9, pp. 5809–5828, May 2017.
[41] S. Kothe, U. Pfeifroth, R. Cremer, J. Trentmann, and R. Hollmann, "A satellite-based sunshine duration climate data record for Europe and Africa," *Remote Sens.*, vol. 9, no. 5, May 2017, Art. no. 429.
[42] A. G. Amillo, T. Huld, and R. Müller, "A new database of global and direct solar radiation using the eastern meteosat satellite, models and validation," *Remote Sens*, vol. 6, no. 9, pp. 8165–8189, Aug. 2014.
[43] G. Alexandri *et al.*, "A high resolution satellite view of surface solar radiation over the climatically sensitive region of Eastern Mediterranean," *Atmos. Res.*, vol. 188, no. 15, pp. 107–121, May 2017.
[44] P. Minnis *et al.*, "CERES edition-2 cloud property retrievals using TRMM VIRS and terra and aqua MODIS data—Part I: Algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4374–4400, Nov. 2011.
[45] H. Hersbach *et al.*, "Global reanalysis: Goodbye ERA-Interim, hello ERA5," *ECMWF Newslett.*, vol. 159, pp. 17–24, Apr. 2019.
[46] S. J. Paech *et al.*, "A calibrated, high-resolution goes satellite solar insolation product for a climatology of Florida evapotranspiration," *J. Amer. Water Resour. Assoc.*, vol. 45, no. 6, pp. 1328–1342, Dec. 2009.
[47] L. Chen *et al.*, "Spatial scale consideration for estimating all-sky surface shortwave radiation with a modified 1-D radiative transfer model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 3, pp. 821–835, Mar. 2019.
[48] D. Wang *et al.*, "A new set of MODIS land products (MCD18): Downward shortwave radiation and photosynthetically active radiation," *Remote Sens,* vol. 12, no. 1, Dec. 2019, Art. no. 168.
[49] S. K. Gupta, D. P. Kratz, A. C. Wilber, and L. C. Nguyen, "Validation of parameterized algorithms used to derive TRMM–CERES surface radiative fluxes," *J. Atmos. Ocean Technol.*, vol. 21, no. 5, pp. 742–752, May 2004.
[50] F. Wu and C. Fu, "Assessment of GEWEX/SRB version 3.0 monthly global radiation dataset over China," *Meteorol. Atmos. Phys.*, vol. 112, no. 3, pp. 155–166, Apr. 2011.
[51] X. A. Xia, P. C. Wang, H. B. Chen, and F. Liang, "Analysis of downwelling surface solar radiation in China from national centers for environmental prediction reanalysis, satellite estimates, and surface observations," *J. Geophys. Res. - Atmos.*, vol. 111, no. D9, May 2006, Art. no. D09103.
[52] R. C. Yu, Y. Q. Yu, and M. H. Zhang, " Comparing cloud radiative properties between the eastern China and the Indian monsoon region," *Adv. Atmos. Sci.,* vol. 18, pp. 1090–1102, Dec. 2001.
[53] C. E. Bulgin, O. Embury, and C. J. Merchant, "Sampling uncertainty in gridded sea surface temperature products and advanced very high resolution radiometer (AVHRR) global area coverage (GAC) data," *Remote Sens. Environ.*, vol. 177, pp. 287–294, May 2016.
[54] N. A. Schutgens *et al.*, "Will a perfect model agree with perfect observations? The impact of spatial sampling," *Atmospheric Chem. Phys.*, vol. 16, no. 10, pp. 6335–6353, May 2016.
[55] N. A. Schutgens *et al.*, "On the spatio-temporal representativeness of observations," *Atmospheric Chem. Phys.*, vol. 17, no. 16, pp. 9761–9780, May 2017.
[56] X. Ling *et al.*, "Comprehensive evaluation of satellite-based and reanalysis soil moisture products using in situ observations over China," *Hydrol. Earth Syst. Sci.*, vol. 25, no. 7, pp. 4209–4229, Jul. 2021.
[57] J. C. Anderson *et al.*, "Long-term statistical assessment of Aqua-MODIS aerosol optical depth over coastal regions: Bias characteristics and uncertainty sources," *Tellus B Chem. Phys. Meteorol.*, vol. 65, no. 1, Sep. 2013, Art. no. 20805.
[58] T. Zhang *et al.*, "The validation of the GEWEX SRB surface shortwave flux data products using BSRN measurements: A systematic quality control, production and application approach," *J. Quantitative Spectrosc. Radiative Transfer*, vol. 122, pp. 127–140, Jun. 2013.
[59] H. Yan, J. Huang, P. Minnis, T. Wang, and J. Bi, "Comparison of CERES surface radiation fluxes with surface observations over Loess Plateau," *Remote Sens. Environ.*, vol. 115, no. 6, pp. 1489–1500, Jun. 2011.
[60] L. M. Hinkelman *et al.*, "Using CERES SYN surface irradiance data as forcing for snowmelt simulation in complex terrain," *J. Hydrol.*, vol. 16, no. 5, pp. 2133–2152, Oct. 2015.
[61] S. Kato *et al.*, "Surface irradiances of edition 4.0 clouds and the earth's radiant energy system (CERES) energy balanced and filled (EBAF) data product," *J. Climate*, vol. 31, no. 11, pp. 4501–4527, Jun. 2018.
[62] A. Riihelä *et al.*, "An intercomparison and validation of satellite-based surface radiative energy flux estimates over the Arctic," *J. Geophys. Res. Atmos.*, vol. 122, no. 9, pp. 4829–4848, Apr. 2017.
[63] W. Tang, K. Yang, Z. Sun, J. Qin, and X. Niu, "Global performance of a fast parameterization scheme for estimating surface solar radiation from MODIS data," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 6, pp. 3558–3571, Jun. 2017.
[64] Z. Li, M. C. Cribb, F. L. Chang, A. Trishchenko, and Y. Luo, "Natural variability and sampling errors in solar radiation measurements for model validation over the atmospheric radiation measurement Southern Great Plains region," *J. Geophys. Res. Atmos.*, vol. 110, no. D15, pp. D15S19, Apr. 2005.



Ke Zhang received the B.S. degree in geographic information science in 2020 from the Southwest University, Chongqing, China, where she is currently working toward the Postgraduate degree in geographic information science.

Her research interests include hydrometeorological monitoring and land surface modeling.

Long Zhao received the B.Eng. degree in hydraulics and hydropower engineering from Tsinghua University, Beijing, China, in 2008, and the Ph.D. degree in physical geography from the Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China, in 2013.

He was a Postdoctoral Fellow with the University of Texas at Austin from 2013 to 2017. He is currently an Associated Professor with Southwest University, Chongqing, China. His research interests include land data assimilation, soil moisture/temperature monitoring, and up-/down-scaling. His current research focuses on a NSFC project "Satellite Land Data Assimilation: Toward a Robust Model Parameter and Soil Moisture Estimation."

Wenjun Tang received the B.S. degree in applied meteorology from the Nanjing University of Information Science and Technology, Nanjing, China, in 2006, and the Ph.D. degree in physical geography from the Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China, in 2012.

He is currently an Associate Professor with the Institute of Tibetan Plateau Research, Chinese Academy of Sciences. His research interests include surface energy balance, climate change, and quantitative remote sensing.

Kun Yang received the B.E. degree in fluid mechanics and machinery and the M.S. degree in environmental hydraulics from the Department of Hydraulic Engineering, Tsinghua University, Beijing, China, in 1994 and 1997, respectively, and the D.Eng. degree from the Department of Civil Engineering, University of Tokyo, Tokyo, Japan, in 2000.

He was appointed as an Associate Professor with the University of Tokyo, in 2003 (the Coordinated Enhanced Observing Period Project – Water and Energy Budget Study) Chair in 2007, and Professor with the Institute of Tibetan Plateau Research, Chinese Academy of Sciences in 2007. Since 2016, he has been Full Professor with the Department of Earth System Science, Tsinghua University, Beijing, China. His research interests include land surface modeling, data assimilation, and plateau land-atmosphere interactions.

Jing Wang received the B.S. degree in geographic information science in 2021 from the Southwest University, Chongqing, China, where she is currently working toward the Postgraduate degree in geographic information science.

Her research interests include hydrometeorological monitoring and land surface modeling.