Dear Referee #2:

Thank you very much for your positive and constructive comments. We have studied the valuable comments carefully and have made corrections which we hope can meet with approval. We apologize that we did not present some important results and Figures in the original manuscript, limited by the length of the article. We will add these Figures and results in the revised manuscript or supplemental material, following the suggestions by Referee #2 and the editor. The point to point responds to the Referee #2's comments are listed as following:

Comment#1:

Using daily cumulative sunshine duration to derive hourly cloud transmittance and hourly solar irradiance is illogical and maybe an obvious mistake.

[Response to Comment#1]:

Thank you for your valuable advice. The actual sunshine duration during a given period is defined as the sum of that sub-period for which the direct solar irradiance exceeds 120 W m^{-2} (WMO). Sunshine duration is a potential good index that can reflect the influence of cloud on surface solar irradiance. The sunshine duration data are widely used to correct the cloud effect on global (direct and diffuse) radiation. The REST2 could efficiently estimate the GHI, DNI and DIF values, but should be corrected under all sky condition. However, getting hourly sunshine durations measurement throughout mainland China with long temporal range is impossible. Therefore, we tested the accuracy of correcting the solar irradiance values using daily sunshine duration measurements at Xianghe station. Figure 1 showed the line chart of the GHI estimations in clear sky, GHI estimations with cloud correction, GHI observation in random selected day of each month in 2007 at Xianghe station. Figure 2 showed the comparison results of the hourly estimated GHI/DNI/DIF values without cloud correction and the hourly measured GHI/DNI/DIF values at Xianghe station. The result showed that there are uncertainties in estimating GHI/DNI/DIF values without considering the cloud effect on solar irradiance. The accuracy of the estimated hourly GHI, DNI and DIF values have been significantly improved after correcting the cloud effect on GHI, DNI and DIF values.

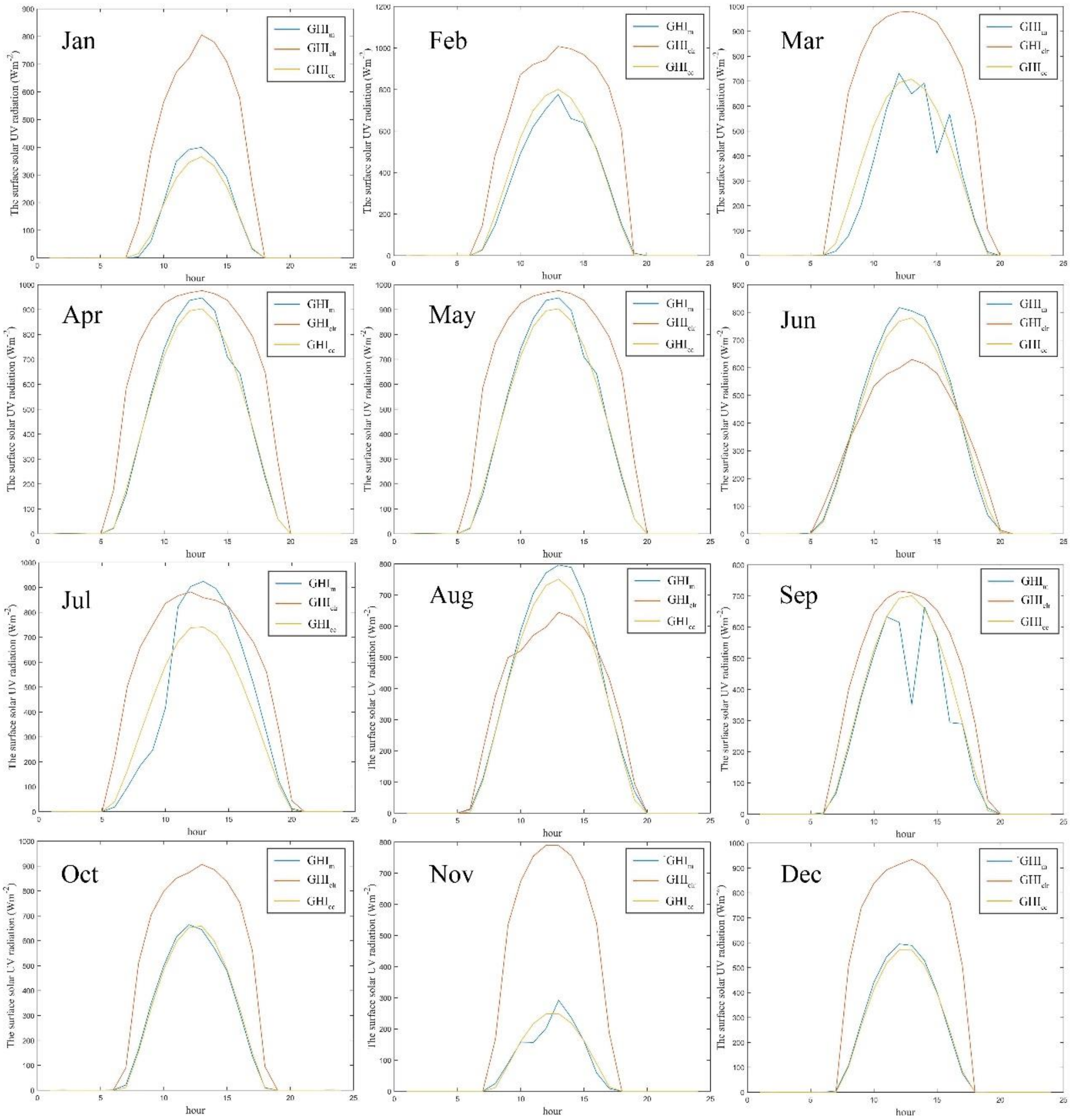

Figure 1. The hourly estimated GHIclr, GHIcc and GHIm in random selected day of each month in 2007 at Xianghe station. (GHIclr, GHIcc, and GHIm are the GHI estimations in clear sky, GHI estimations with cloud correction, GHI observation, respectively).

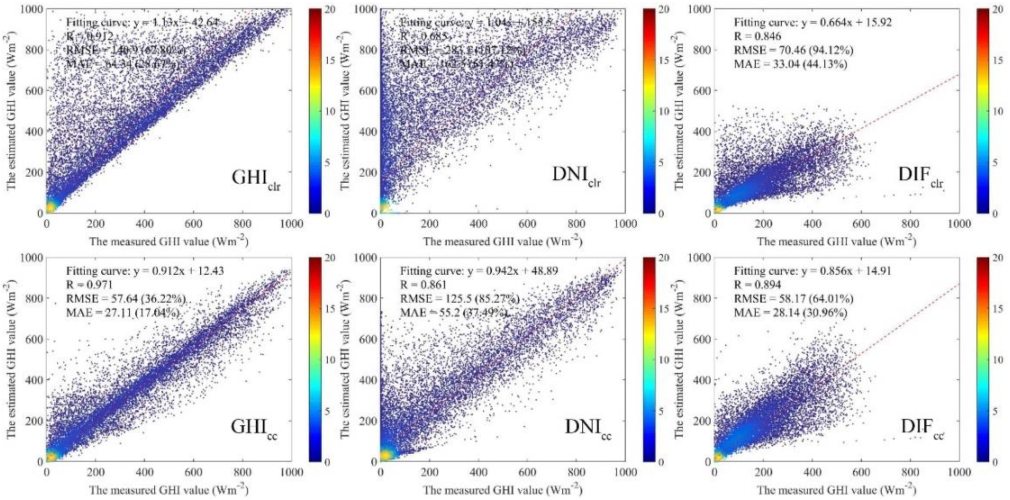

Figure 2. Validation of the hourly mean GHI, DNI and DIF values at Xianghe station.

Certainly, there is still a lot of room for improving the accuracy of the cloud effect on solar irradiance. Hourly sunshine duration measurement are better reliable data to calculate the cloud effect on solar irradiance. Future work will be conducted to improve the model accuracy by incorporating high-quality cloud data or available hourly sunshine duration measurements.

Comment#2:

Also, I don't believe the accuracy of satellite retrievals would lower than that of reanalysis data corrected with ground observations (except for the station used to correct the reanalysis) because the reanalysis data is difficult to simulate the realistic clouds.

[Response to Comment#2]:

Thank you very much for your nice comments. We agree with you that the accuracy of satellite retrievals would be better than reanalysis data. However, the advantage of the datasets generated in this study is their high spatial and temporal continuity, rather than accuracy, because satellite signals are frequently affected by clouds. The spatial-temporal continuity and long temporal range could make up for the loss of accuracy. The dataset generated in this study could provide long-term solar irradiance data for the validation of the state-of-the-art climate model, such as the Coupled Model Intercomparison Project Phase 6 (CMIP6). We are looking forward to generate GHI, DNI and DIF data with higher accuracy, longer temporal range, higher spatial and temporal resolution, and spatial- temporal continuity in future. We will try to combine the advantages of satellite retrievals and reanalysis data in future work.

Comment#3:

I personally think the paper is outside the scope of regular articles because it is similar to the interpretation of data.

[Response to Comment#3]:

Thank you for your comment. There are four types of articles on ESSD including Data description paper, review article, brief communication and peer-reviewed comment. This article is a data discussion paper submitted on ESSD. We are sorry that we may not understand what “interpretation of data” means.

Comment#4:

The words “high-quality” in title is inappropriate because its spatial resolution is coarse.

[Response to Comment#4]:

Thank you for your good comments. The “high-quality” here means high spatial continuity, temporal continuity and temporal range (1980-2014). However, the title would be too long, if all these merits of this dataset are included in the title. The spatial resolution is the disadvantage of the generated datasets in this study. Further studies would be conducted to improve the spatial resolution of the datasets.

Comment#5:

Most of the error indicators for accuracy evaluation are not frequently used and also redundant**.**

[Response to Comment#5]:

Thank you for your nice comments. The normally used statistical indicators (RMSE vs RMSD, MAE vs MAD) are always in contradiction, for example, in Table S1, the RMSE, MAE, RMSD, MAD, and R for DOH are 129.13, 94.20, 50.89 and 0.91, respectively; while the RMSE, MAE, RMSD, MAD, and R for NMG are 146.53, 85.47, 49.72, and 0.89. The higher the R is, the better the estimated results are; the lower the

RMSE, MAE, RMSD and MAE are, the better the estimated results are. It was obvious that too few indicators may mislead the validation result. We could not tell the model accuracy from these statistical indicators. Therefore, 15 indicators are introduced in this study. Meanwhile, to evaluate the overall performance of the estimated GHI, DNI and DIF values, a global performance indicator (GPI) was used to reveal the overall accuracy of the estimate results in this article.

Comment#6:

Data description is unclear. For example, the observation instruments, error, frequency, length, quality and so on. Moreover. The quality control method should be introduced in detail because it is very significant for accuracy evaluation.

[Response to Comment#6]:

Thank you very much for your valuable suggestions. We are deeply sorry that we did not present detail data description in the original manuscript, limited by the length of the article. We will add data description in the revised manuscript or supplemental material, if the Referee #2 and the editor think they are necessary.

The sunshine duration measurements and solar irradiance measurements at CMA stations are provided by the China Meterological Administration. Figure 3 showed the structure diagram of sunshine duration sensor at CMA stations. The maximum permissible error and the annual stability for the sunshine duration sensor is ±10% per month and ±5%, respectively. The spectrum ranges, threshold of DNI values and maximum permissible error for the sunshine duration sensor are 400nm ~ 1100nm, $120W/m^2$, and $\pm 24\ W/m^2$, respectively. The observation frequency and length of the sunshine duration sensor are 10s and 24 hours, respectively. Figure 4 indicated the flowchart of the observation process for sunshine duration.

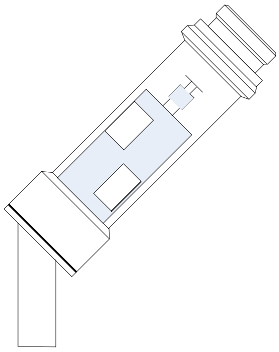

Figure 3. Structure diagram of the sunshine duration sensor at CMA stations.

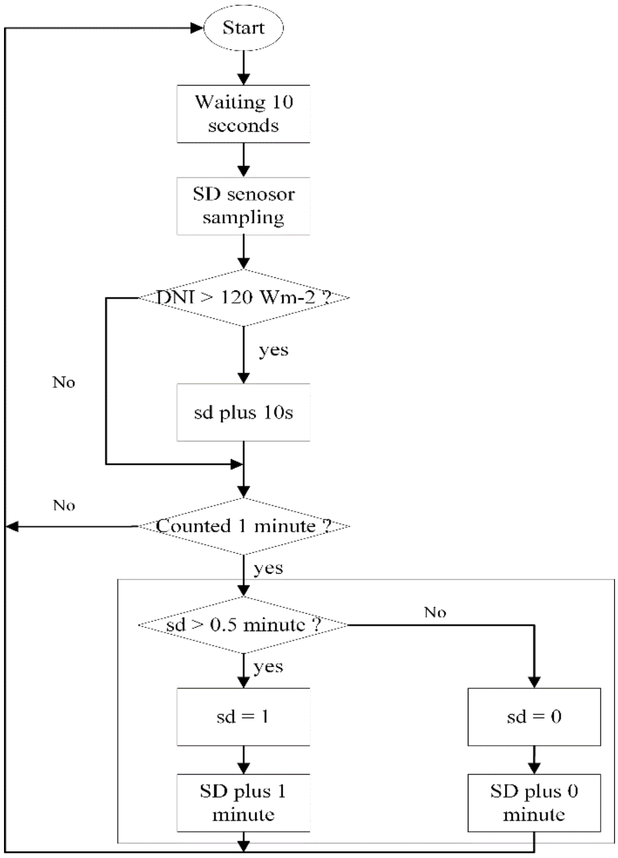

Figure 4. The flowchart of the observation process for sunshine duration.

The GHI, DNI and DIF values are measured with a pyrheliometer mounted on a sun tracker at 2474 CMA stations. Before 1993, the observation radiometers that were used at these CMA stations were similar to those used in the Soviet Union. However, the observation precision of these radiometers was impacted by imprecise instrument calibration, sensor aging and data quality control method (Tang et al., 2010). To combat these problems, the CMA updated the observation instruments to a thermopile pyrheliometer (DFY-4) with higher accuracy and robustness. The observation instrument for Wuhan station and Xianghe Station is Solys2 sun tracker. Figure 5 showed the Solys2 sun tracker. The detail information about Solys2 sun tracker could be obtained at 'https://www.kippzonen.com/Product/20/SOLYS2-Sun-Tracker'.

Figure 5. The physical photograph of Solys2 sun tracker.

The quality control processes have been done but not included in the original manuscript, considering the length of the article. Thus, we could only add a brief description of the quality control process in the revised manuscript. Detail description would be added the supplemental materials, according to the suggestions from Referee #2 and the editor. Five main quality check processes including the climate limit value or allowable value check, station extreme value check, internal consistency check among timing value, daily average value and daily extreme value, time consistency check and spatial consistency check have been done to ensure quality of the sunshine duration measurements. Then, the sunshine duration measurements were marked with 0, 1, 2 and 8, which represent correct data, suspicious data, wrong data and data missing, respectively. This work has been done by the China Meterological Administration. Meanwhile, the quality of solar irradiance measurements was also conducted following the rule that the GHI, DNI and DIF values should not exceed the extraterrestrial radiation at the same geographical location, otherwise the measurement will be directly deleted.

Comment#7:

How about the uncertainty about the interpretation of sunshine duration data because it is a main factor influencing the accuracy of your products?

[Response to Comment#7]:

Thank you very much for your comments. The cross-validation has been conducted to evaluate the uncertainty about the interpretation of sunshine duration data. We are deeply sorry that we did not present the cross-validation result in the original manuscript, considering the length of the article. However, if Referee #2 and the editor think this section should be included in the revised manuscript, we could put it in the final version of the manuscript or supplemental materials.

Firstly, 50 stations were randomly selected to conduct the cross-validation experiment. As shown in Figure 6, these stations covered most area of China. Then, sunshine duration measurements at remaining 2424 CMA were used as input data in Anusplin tool. Finally, the measured sunshine duration values are compared with the sunshine duration values after interpretation. Figure 7 and Figure 8 are the scatter plot showing the accuracy of the interpretation result of sunshine duration at 50 random selected CMA stations. The result indicated that the Anusplin Tools could be used to interpreted sunshine duration values at 2474 CMA stations with high accuracy.

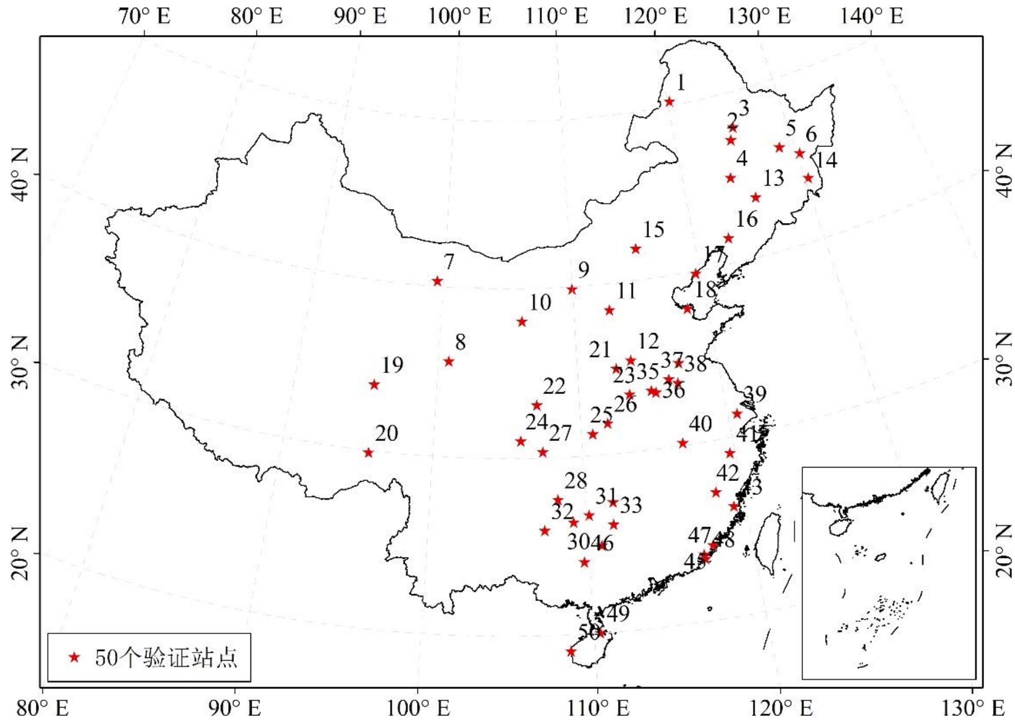

Figure 6. 50 random selected CMA stations for validating the accuracy of the sunshine duration values after interpretation.

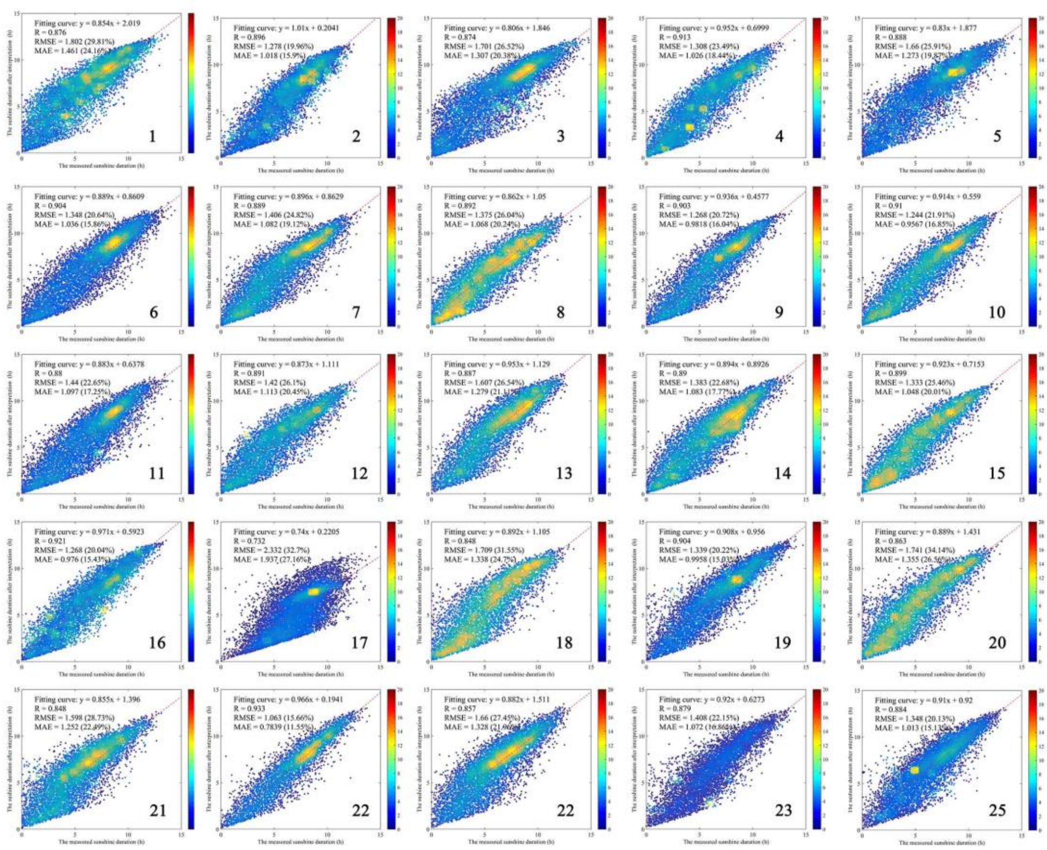

Figure 7. The validation result of cross-validation at 50 CMA stations

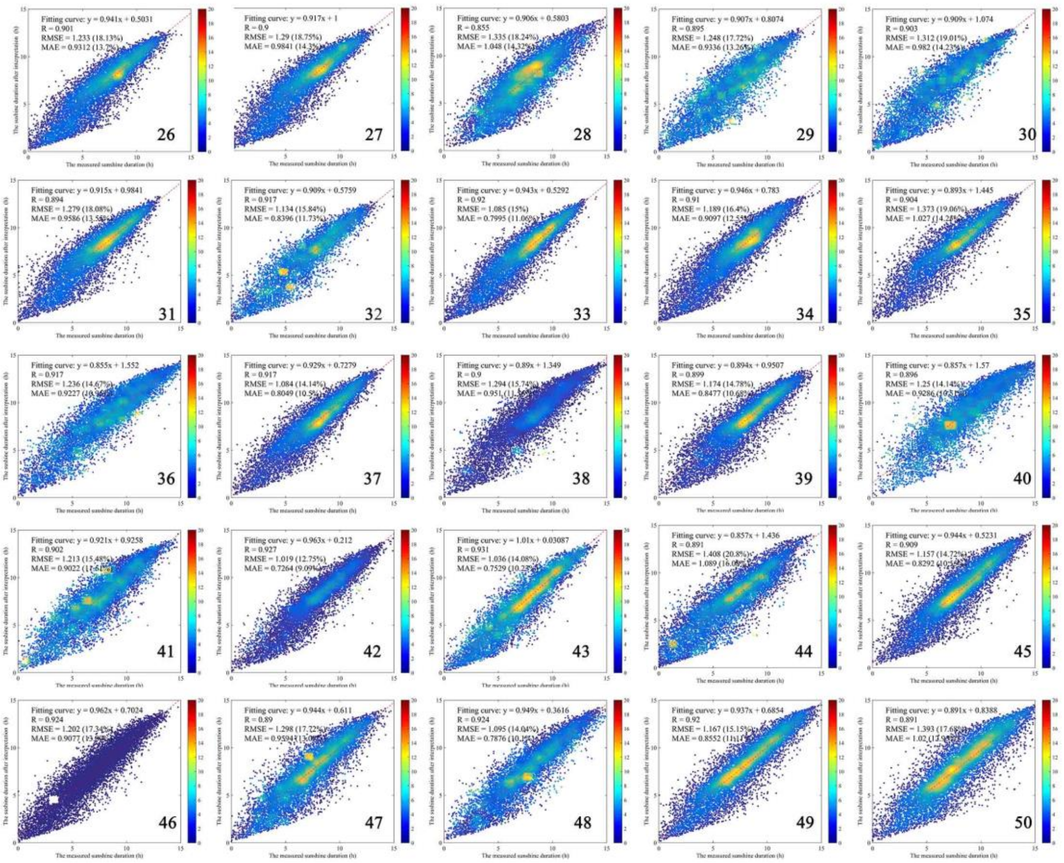

Figure 8. The validation result of cross-validation at 50 CMA stations.

Comment#8:

how to derive the cloud transmittances, and how about its uncertainty? how to correct the solar irradiance of MERRA-2 with the cloud transmittances? How about the uncertainty of the correction process?

[Response to Comment#8]:

Thank you very much for your nice comments. Following the example of the Angstrom-Prescott formula, Yang's definition (Yang et al., 2001) and Tang's definition of the cloud transmittances (τc) (Tang et al., 2018), τc was defined as the ratio between the solar irradiance in all-sky condition (the solar irradiance measurements) to the estimated clear sky solar irradiance. We parameterized the cloud transmittance (τc) as a function of the relative sunshine duration (n/N), and the formula form was a quadratic polynomial formulation as follows:

$$\tau_c = \frac{R}{R_{clr}} = a + b\left(\frac{n}{N}\right) + c\left(\frac{n}{N}\right)^2 \qquad (22)$$

where R is the hourly and daily all-sky GHI, DNI and DIF; R_{clr} is the hourly and daily clear-sky GHI, DNI and DIF; and n and N are the sunshine duration and the maximum possible sunshine duration, respectively. Figure 9 showed the scatter plot between τc for daliy GHI/DNI/DIF values and the relative sunshine duration. It was clear that there is close relationship between τc and the relative sunshine duration.

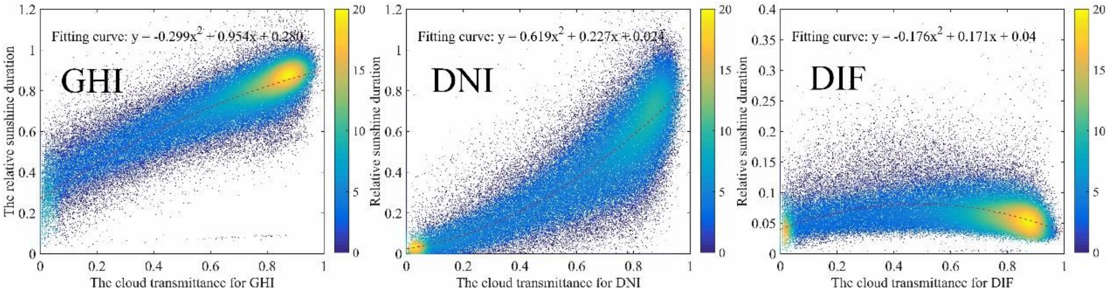

Figure 9. Empirical function between relative sunshine duration and cloud transmittance for GHI, DNI and DIF values at Xianghe station.

It should be stated that we did not correct the solar irradiance of MERRA-2 GHI record, because the data we download is the surface incoming shortwave flux in all-sky condition (SWGDN) from MERRA2 official website. The MERRA-2 products are produced with version 5.1.4 of the GEOS atmospheric data assimilation system. The key components of the system are the GEOS atmospheric model and the GSI analysis scheme. The Radiative transfer calculations necessary for the assimilation of satellite radiances in MERRA-2 are performed using the CRTM. Detail information for the official algorithm for calculating GHI could be found in an article named "The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2)". The high errors of the GHI values derived from MERRA-2 are consistent with a known tendency for the GEOS-5 systems to underestimate mid-latitude continental cloud cover (Draper et al. 2018).

We are deeply sorry that we did not present these results and Figures in the original manuscript, limited by the length of the article. We will add these Figures and result in the revised manuscript or supplemental material, if the Referee #2 and the editor think they are necessary.

Comment#9:

How to derive the hourly cloud transmittance because the sunshine duration is daily cumulative. This is the obvious mistake of this article.

[Response to Comment#9]:

Thank you very much for your valuable comment. We agree with Referee#2 that deriving hourly cloud transmittance is a shortcoming in this study. We are deeply sorry for this shortage of this experiment. Please kindly give us time to explain to you.

As shown in Figure 1, the REST2 always overestimated or underestimated the solar irradiance value, but it could preliminarily simulate the hourly solar irradiance value following the hourly change track of solar irradiance measurement. Thus, we try to corrected overestimated solar irradiance values using sunshine duration measurements. However, as described in “Response to Comment #6”, we could only conduct the experiment using daily mean sunshine duration measurement, limited by the data availability of hourly sunshine duration data.

It should be stated that the main goal of this article is correcting the uncertainty of the estimated clear-sky solar irradiance values by REST2 model, rather than simulating the hourly cloud transmittance. Certainly, the accuracy of the estimated solar irradiance values could be further improved using hourly cloud transmittance data. This limitation has been described in the revised manuscript, which would further investigate in future work.

Comment#10:

In Fig.3 for GHInew, we can not observe the overall overestimation, but the MAE indicates that the GHInew is significantly overestimated. It’s a contradiction and the results is unbelievable.

[Response to Comment#10]:

Thanks for your comment. We agree with you that the GHInew values did not show good agreement with GHI measurements. The GHI, DNI and DIF estimations in this study are subject to the input data quality, the interpolated method and the relatively coarse resolution of MERRA-2 products (AOD550, rog, p, w). However, the purpose of this study is to generated an hourly, daily and monthly GHI/DNI/DIF datasets with acceptable accuracy, high spatial and temporal continuity, long temporal range. Indeed, the data accuracy is still not high, but have been greatly improved than that of MERRA

GHI records. The spatial- temporal continuity and long temporal range could make up for the loss of accuracy. High quality input data is the premise for this goal. However, existing remote sensing products, reanalysis products and ground meteorological station measurements cloud not meet this requirement.

Comment#11:

Table 3 is meaningless and the comparison with other studies is extremely unfair because the spatiotemporal resolution, the input data, and the observation data (also number of observation stations) for these studies are completely different with you.

[Response to Comment#11]:

Thank you very much for your valuable suggestions. We intended to compare the accuracy of the GHI estimations in previous studies. Following your suggestions, we have deleted this section in the revised manuscript.