ESSD https://doi.org/10.5194/essd-2024-91

RC1: 'Comment on essd-2024-91', S. Kato, 1 June 2024 reply

The authors describe the surface solar radiation data product SARAH-3, which is a revised version of SARAH-2. The product includes global surface irradiance, direct irradiance, direct normal irradiance, photosynthetic active radiation, daylight, effective cloud albedo, and sunshine duration. Computations of these variables are based on parameterized model. Resulting global irradiances are evaluated against BSRN and GEBA data set. The manuscript is a thorough description of the data set. I only have minor comments/suggestions on this version.

A: Dear S. Kato, Thank You for taking note of our manuscript and for Your review and positive feedback.

Line 114: Could you include units of 160 and 112, or convert this to m s-1 if you know the size of the pixel.

A: Indeed, we have been unclear about the optical flow speed. We missed to add some essential explanations concerning the numbers. The unit of the optical flow speed generally is pixels per image sequence (30min in this case). For our application we wanted to have sufficient contrast at low speeds, so the resulting optical flow speed in truncated at 1 and stored as 8-bit greyscale image with a maximum value of 255. This means a value 255 corresponds to an optical flow speed of 1 pixel/30 min, and a threshold value of 160 for MVIRI corresponds to a speed of 160/255 ~ 0.63 pixel/30 min. As the native pixel size of the visible channels for SEVIRI is larger, the speed threshold is reduced accordingly to 112/255 ~ 0.44 pixel/30 min. In other words, the threshold used for the optical flow speed is quite low in order to assure clouds are excluded from the snow detection process. We will include/update the values in the revised manuscript.

Figure 3: Please include the definition of optical flow speed in the caption.

A: We will try to be clear concerning the optical flow speed in the revised manuscript.

Line 145-148. Are you saying that snow-coverage information is replaced by that of ERA5? Then why don't you use the ERA5 snow map and skip step 1 and 2?

A: The snow-coverage information is not replaced, but corrected by ERA-5 in the way that snow is removed if it is not included in ERA-5. A reason for this correction is that there is a chance of misclassifying fog as snow, which is considered in the algorithm but not fully solvable.

Line 154. Albedo is a hemispherical variable. It is awkward that albedo is derived by the ratio of radiance differences.

A: You are referring to the Effective Cloud Albedo here (CAL). Yes, the albedo is defined as integral of reflectances over all wavelengths and over the whole hemisphere. In our approach we assume the cloud reflectance to be isotropic. Different viewing geometries are considered by calculating the CAL based on the normalized reflection and by using the minimum reflectance per time slot (e.g. for the 13:00 UTC time slot) and per pixel.

CAL might also be expressed as the normalized difference of the all sky and clear sky reflection. Since several decade this quantity is called Effective Cloud Albedo.

Line 186 change radiation to irradiance.

A: Agreed. We will use the term irradiance here. This will be included in the updated manuscript.

Line 187 Need some explanations of how to get the relationship.

A: For the calculation of SID we would like to refer to Müller et al., 2015 and Skartveit, A., Olseth, J.A. and Tuft, M.A. (1998) An Hourly Diffuse Fraction Model with Correction for Variability and Surface Albedo. Solar Energy, 63, 173-183. We will point to these references for SID in the revised manuscript.

Line 300 Need brief description of Roesch et al's approach to derive monthly means.

A: We will add a brief description of the how the BSRN data is averaged to daily and monthly means. Roesch et al. proposes a monthly mean calculation method of 1-minute solar irradiance data that makes use of the diurnal cycle of solar irradiance. Hence the monthly mean is not disturbed from missing values.

BSRN and GEBA data: Are these stations listed in Tables 2 and 3 available for the entire period of 1983 to 2020? If they are not, please add column to indicate available time period for each station.

A: Thank you for the comment. GEBA data is available for the entire period of 1983 to 2020, but BSRN is not. We will add the temporal data coverage for the BSRN stations.

Figure 11: Why are absolute values used in these plots. I suggest using RMS differences.

A: The mean absolute difference (MAD) has been found to be very easy to understand and interpret by the users. Further the MAD is used in the CM SAF surface radiation data record validation as the main measure to define the quality of a data record. For these reasons we would like to stick to the absolute values. By using the boxplots the spread of the differences is nicely shown. However, we consider to add the RMS values for each station to the plots.

Figure 13: Are there any reasons for using sunshine duration, or not using global irradiances, in the figure. Figures 14 and 15 uses global irradiances. Please include global irradiance anomaly time series plot for the BSRN comparison.

A: **TO DO**: The reason for using sunshine duration in this plot is that there are much more reference measurements for sunshine durations and their temporal availability allows the analysis of the whole SARAH-3 climate data record. BSRN data starts in 1994 which is likely too late to see any effect of the Pinatubo eruption. Considering the small number of BSRN stations and their heterogeneous data coverage, an anomaly series over time would be difficult to interpret. Those issues diminish by using the monthly sunshine duration data from CLIMAT as reference in Figure 13. For the long-term stability analysis, the GEBA data is more appropriate due to its better spatial and longer temporal coverage, even though limited to Europe. Nevertheless, we will consider adding the BSRN bias time series to the revised manuscript.

Figures 13, 14, and 15. It is not clear whether a downward trend exists in parameterized irradiance, or observations, or both. Figure 15 suggests that trends for both computed and observed irradiances are mostly positive, but the trend of the difference is negative because observations have a large positive trend. Is this true for BSRN? Also, the negative trend is largely due to positive difference before 1995 for both BSRN and GEBA. Please investigate further why the difference is larger before 1995. In addition, if the number of surface sites changes over the course of the time period, it might introduce a trend.

A: You are right, Figure 13 and 14 show the bias time series of SARAH-3 vs station observations and do not show absolute trends. Both Figures 13 and 14 show that there is a small negative trend in the bias for sunshine duration (Figure 13) and for global irradiance. That means that a trend in the surface reference measurements would be underestimated by the SARAH-3 data record. Indeed, the overall negative trend is partly caused by the positive anomalies in the early years of the data record, which in turn might be caused by the underestimation of the direct aerosol effect during that period. The strong Pinatubo eruption seems to play a role here, too. For Figure 14 and 15 the number of stations available over time is constant to avoid artificial trends. Figure 15 shows the absolute trends of global irradiance. A constant number of stations over the whole time period has be used. The “Trendraster”-plots have been calculated and plotted for SARAH-3 and the GEBA surface measurements. Figure 15 reveals that both data sources see a positive trend in global irradiance of ~3.3 W/m2/decade for the GEBA statins and ~2.7 W/m2/decade for SARAH-3 global irradiance. Hence there is a negative trend in the bias (SARAH-3 minus GEBA) of ~-0.6 W/m2/decade. This value is also seen in Figure 14 (left) as number in lower part. The data basis for Figure 14 and 15 are similar.

Line 514 to 525. I assume that these discussions are for Figure 16. There are many lines in Figure 16 and hard to see. I suggest separating Figure 16 into two or more plots.

A: **TO DO** Yes, lines 514 to 525 discuss Figure 16. We will add the missing reference to Figure 16 here. Figure 16 is meant to qualitatively explain the relation between the SARAH-3 variables by showing zonal means. Splitting the Figure is an option but the relation between the variables would be less well visible. We suggest enlarging the figure and better streamline the discussion.