Reply to report of reviewer #2 "Evaluating parallax and shadow correction methods for global horizontal irradiance retrievals from Meteosat SEVIRI by Wiltink et. al."

bold italic font = reviewer's comment
regular font = authors' reply
red regular font = original text in the manuscript
blue regular font = newly added or updated text in the manuscript

---

General comments:

The manuscript "Evaluating parallax and shadow correction methods for global horizontal irradiance retrievals from Meteosat SEVIRI" by Wiltink et al. presents correction methods for satellite-based GHI retrievals. Two approaches are used here: purely geometric (parallax and shadow corrections) and empirical collocation shifts. The paper is overall well written and the scientific topic investigated is justified and matches well with AMT. The structure makes sense overall, although some main findings could be stressed a bit more clearly (e.g. the need to lower the cloud height to achieve the lowest RMSE and the counter-acting or opposing effects of two correction methods).

The authors thank the reviewer for taking the time to provide feedback on our preprint. The comments are appreciated and should further improve the quality of the manuscript. Please find our reply below.

We tried to present the main findings more clearly in the manuscript. To this end we have rewritten the abstract and made some changes to the results and conclusion sections. For the adjustments made to these sections, we refer to our reply to reviewer #1 who had similar concerns.

In the "conclusions and outlook" section I miss a bit to bring the findings of this study into a broader context. E.g. what are the implications/suggestions for the usage of satellite derived GHI in predictions for PV yields? Is the difference of the satellite retrievals of about 100 W/m2 w.r.t. the pyranometers considered critical in view of typical total solar irradiance values?

The current study stresses the need for a combined parallax and cloud shadow correction to achieve optimal GHI accuracy. This is of interest in particular for PV monitoring and forecasting. The article shows that the improvements due to the parallax and cloud shadow correction are largest for heterogeneous cloud regimes, which are usually the regimes where the largest PV forecast errors are made. Furthermore, we show that the applied corrections become increasingly relevant at higher resolution, which also is an important aspect considering the higher resolution retrievals of the current generation of geostationary satellites.

The error of ± 100 W m^{-2} is for 5 minute timescales where there is a large degree of spatio-temporal cloud variability. Thus, at these timescales larger errors are expected. Temporal averaging significantly reduces the observed errors. At hourly timescales the error is roughly halved (Wiltink et al., 2024). To a large extent, the achievable GHI accuracy depends on the time scale of interest. We would argue that any improvement in GHI accuracy is valuable for the PV community. Especially the combined parallax and cloud shadow geometric correction has potential, since it is relatively easy to implement, relies on (already available) SEVIRI observations, and requires little calculation time. The reduction in RMSE of about 10 % with respect to the uncorrected retrieval could be considered a 'free' improvement.

The relevance of this study in a broader context could indeed be mentioned more explicitly in the conclusion section, and therefore we have added the following paragraph to the conclusions:

Page 23, line 357: This study shows the relevance of accurately correcting GHI retrievals for parallax and shadow displacement at increasing resolutions. → This study shows the relevance of correcting GHI retrievals for both parallax and cloud shadow displacement to achieve a higher degree of accuracy. These findings are of special interest for users of satellite-based GHI products such as grid operators making PV energy forecasts. Especially the combined geometric parallax and cloud shadow correction has potential for broad adaptation. This correction relies only on the satellite and solar positions and satellite-observed radiances to retrieve H_c, and can thus be performed at any place and time as long as observations are available.
Another relevant aspect for users of satellite-based GHI products is that parallax and cloud shadow corrections become increasingly important at higher resolutions.

The presented correction methods seem to be able to reduce the errors between the satellite retrievals and the ground based retrievals of GHI by about 10-15 W/m2. This could be emphasized more clearly.

In the updated manuscript we now mention the improvement of the combined geometric correction w.r.t. the uncorrected retrieval (11.7 W m^{-2}) in the abstract, results and conclusions. The improvement for the timestep mean optimal shift (15.7 W m^{-2}) will be reported in the results and conclusions.

If the authors could expand a bit more regarding these points mentioned above it would be appreciated, however, I will not insist on that. I recommend publication of the manuscript after the following minor revisons:

Specific comments

1) ***Line 33: I think what is meant here is "Geostationary Operational Environmental Satellites (GOES)" and not Global Earth Observing System (GEOSS)?***

   We indeed intended to refer to the "Geostationary Operational Environmental Satellites" and not the "Global Earth Observing System (GEOSS)". The reviewer's comment has been modified accordingly.

2) ***Line 189: What does "any direction" mean here? Fixed along either N, S, W, E direction or along the SEVIRI footprints or along any polar angle? Please specify.***

   With any direction we mean a fixed shift by multiples of 500 meter along the north-south and/or west-east axes. To prevent confusion, we have further clarified the directions of the shift of the SEVIRI grid:

   Page 8, line 189: For each day of the field campaign, the SEVIRI grid is shifted by multiples of 500 m in any direction. → For each day of the field campaign, the SEVIRI grid is shifted by multiples of 500 m along the north-south and/or west-east axes.

3) ***Figure 3 and Line 202: Is there a shading for the 5th to 95th percentile missing?***

   The 5^{th} to 95^{th} percentile does have a light shading, yet this might indeed easily be overlooked. We have reduced the transparency to make the shading appear less white in the figure.

4) ***Line 242: Does this mean that the retrieved Hc, that is used as input to calculate the GHI, is on average over-estimated? (since Hc has to be reduced to get the smallest RSME?)***

   On average, we have no indications that the H_c retrievals of SEVIRI are overestimated. In fact, this type of H_c retrievals, including the one used in this paper, tend to be lower than measured by space-based lidar instruments such as CALIPSO (Hamann et al., 2014). That the smallest errors are found for 70 to 90 % relative cloud top height, we mainly attribute to the neglect of the vertical structure of the cloud. Taking the vertical extent of a cloud into account means that radiation will be backscattered from lower altitudes within the cloud, which would reduce the required magnitude of the geometric corrections.

5) ***Line 248: Several cloud retrieval algorithms use a simplified Lambertian cloud model which is more representative of a "cloud mean height" than a "cloud top height". Have the authors considered to apply such cloud models instead of using Hc as input in order to mitigate this effect of ignoring the cloud's vertical extent?***

   We agree with the reviewer that applying a geometric correction based on cloud mean height would be an interesting study. Perhaps the computed parallax or cloud shadow displacement would become more accurate. However, to retrieve H_c, SEVIRI radiances in the thermal infrared are used, while the computation of the cloud mean height with a Lambertian cloud model would require additional observations, e.g. in the O_2 A-band, which are not available from SEVIRI. With the upcoming Sentinel-4 on MTG-S, O_2 A-band measurements from geostationary orbit will become available. Follow-up research could study the effect of cloud height retrievals from this instrument on geometric correction accuracy in more detail, but this is beyond the scope of the current paper.

6) ***Figure 5: If I understand correctly, a negative difference in panel c) indicates that the HR retrievals have a lower RMSE than the SR retrievals and that the HR-based daily or time-step corrections are better for the regimes CR2 and CR5-CR9. However, for CR1, CR3 and CR4 the SR-based corrections seem to be better. Do I understand that correctly? If so, then the term HR RMSE "improvement" in the Figure caption might be misleading since in some cases the improvement is with the SR and not the HR. Please clarify.***

   The reviewer is correct to point out that not for all cloud regimes higher resolutions lead to an improvement in RMSE. For CR1, CR3 and CR4, the optimal shift method indeed gives more accurate results at standard resolution. Moreover, in Figure 5c, we subtract the standard resolution RMSE from the high resolution RMSE which means that negative values correspond to HR improvements. This will be more clearly explained in the revised caption.

   Page 13, Figure 5 caption: (c) box-and-whisker plots indicating the HR RMSE improvement. → (c) box-and-whisker plots indicating the HR - SR RMSE difference (negative means that HR is better than SR).

7) ***Lines 340-341: The phrases "mainly optimal" and "less optimal" sound strange as "optimal" is an absolute term. Maybe the following sentence sounds more reasonable: "In other words, the optimal shift is mainly suitable for the more variable regimes with lower retrieved clouds and, therefore, slightly less suitable for the more homogeneous regimes with higher retrieved clouds." Please consider to re-phrase this part.***

   We agree with the reviewer and have replaced "optimal" by "suitable" for these sentences.

8) ***Line 362: While this is true for the two extreme cases (uncorrected and fully corrected), it seems that the SR RSME is smaller than HR RMSE at relative cloud top heights above 60% and larger than HR RMSE at lower relative cloud top heights. Although not being statistically significant it might still be worth mentioning if the authors consider this relevant.***

   That is correct. In Figure 7, for relative cloud top heights between 20 % and 80 % the resolution differences are statistically significant, which means that at 60-80 % relative cloud top height the SR retrieval performs significantly better than the HR retrieval. For a fair interpretation of the figure, we now mention this in the manuscript:

   Page 16, line 366: is also significantly better than the SR retrieval. For relative H_c's between 20 and 80 %, the resolution differences are all statistically significant meaning also that at 60-80 % relative H_c the SR retrieval performs significantly better than at HR.

9) ***Lines 365-366: I don't understand how the statement of this sentence can also be deduced from Figure 5b? Can this be backed up by giving the exact RMSE numbers of the 40% Hc from Fig 7 versus the daily shift RMSE from Figure 5b?***

   This statement cannot be made solely from Figure 5b or Figure 7 alone, but only by comparing both figures. The reasoning is as follows.

   From Figure 5b we read that for CR2:

- Mean RMSE daily optimal shift = 118.9 W m^{-2}
- (Mean RMSE timestep optimal shift = 113.0 W m^{-2})
- Mean RMSE of the combined geometrical shift at 100 % relative H_c = 133.1 W m^{-2}

From Figure 7, which shows RMSE for CR2 at multiple relative H_c fractions we read:
(Note: In Figure 7 the median RMSE is plotted but in this case it is nearly identical to the mean value.)

- At 40 % rel. H_c the mean RMSE = 114.6 W m^{-2}
- At 100 % rel. H_c the mean RMSE = 133.1 W m^{-2}
  Note that this value is (and should be) identical to the RMSE shown in Figure 5b for the combined geometrical shift.

Thus, for CR2 the combined geometrical shift at 100 % relative H_c is less accurate than the daily mean optimal shift. In contrast, at 40 % relative H_c the combined geometrical shift is minimal (which is shown in Figure 7), and it is more accurate than the daily mean optimal shift.

While we do believe this is an interesting observation, initially we decided not to include these values in the main text as we expect this might distract from our main findings. However, since the reference back to Figure 5 might not be completely clear we have now decided to remove this reference and include the values in the text instead.

Page 16, line 365: For the HR retrieval, the smallest RMSE is found using a relative H_c of 40 %, corresponding to around 4000 m. At this relative H_c, the combined geometric correction outperforms the daily mean optimal shift ~~(compare with Fig. 5b)~~ (the RMSEs are 114.6 W m^{-2} and 118.9 W m^{-2}, respectively) , and the HR retrieval is also significantly better than the SR retrieval.

10) ***Lines 457-460: Maybe it could be useful for the reader if a Table (or text) is added in Section 2 that specifies the wavelegth ranges of the pyranometers and those of the SEVIRI bands that have been used. Or even show a Figure comparing the spectra for one collocated example case?***

CPP retrieves cloud properties following the Nakajima-King bispectral reflectance method (Nakajima and King, 1990) based on the 0.6 and 1.6 μm channels of SEVIRI. To retrieve GHI, SICCS then uses LUTs that have been simulated with a broadband version of the Double Adding KNMI (DAK) code (Kuipers Munneke et al., 2008). The DAK simulations are performed for wavelengths between 0.240 and 4.606 μm, and the retrieved GHI represents the total solar irradiance in that wavelength range (Greuell et al., 2013).
The pyranometers in the HOPE network actually have a limited spectral response of 0.3 to 1.1 μm. See Madhavan et al. (2016) for a figure of the spectral response. To obtain GHI the pyranometer spectral response function was convolved with the (Gueymard, 2004) solar spectrum and scaled to total solar irradiance as outlined in Madhavan et al. (2016).
In our opinion a table or figure is not required. However, we will mention the spectral ranges of the HOPE pyranometers and the SICCS code in the text:

Page 4, line 97: [..] around Jülich (Madhavan et al., 2016). The spectral response of the HOPE pyranometers is limited between 0.3 and 1.1 μm, but for the calculation of GHI the spectral response function was convolved with the solar spectrum of (Gueymard, 2004) and scaled to total solar irradiance.
Page 5, line 125: [..] reflectance method of Nakajima and King (1990). The bispectral retrieval is based on the 0.6 and 1.6 μm channels of SEVIRI.
Page 5, line 126: [..] to determine GHI. This broadband version of DAK covers the wavelength range from 0.240 to 4.606 μm.

Related to the spectral sensitivities, in Section 5.2 about mismatch errors, a revision has been made. In this section, it is claimed that the spectral differences between SEVIRI and the pyranometer network mainly occur at wavelengths where limited energy is received and thus can be neglected. Indeed, GHI from both sources represents the total solar irradiance, but the measurements are not actually sensitive to the entire solar wavelength range. In particular, the HOPE pyranometers are sensitive to wavelengths between 0.3 and 1.1 μm, which means that they are insensitive to differential absorption by liquid and ice cloud particles and particles of different sizes which occurs at wavelengths in the shortwave infrared. In Madhavan et al. (2016)

a spectral error of 2-5 % is reported for the HOPE pyranometers. To illustrate: if GHI below a cloud is 400 W m^{-2}, this could result in a 20 W m^{-2} error. We have corrected the paragraph in the following way:

Page 21, line 457: ~~The spectral differences between SEVIRI and the pyranometer network are mainly at the wavelengths where limited solar radiation is received. Therefore, the magnitude of this error is negligible compared to the temporal and spatial mismatch (Urraca et al., 2024).~~ → GHI values retrieved with CPP-SICCS and measured by the HOPE pyranometers are representative of the total solar irradiance. However, the sensitivity of the HOPE pyranometers is limited to wavelengths between 0.3 and 1.1 μm. A considerable amount of energy is contained in the part of the solar spectrum that the pyranometers remain insensitive to. In particular, GHI variations due to differential absorption by liquid and ice cloud particles and particles of different sizes which occurs at wavelengths in the shortwave infrared are not accounted for. In Madhavan et al. (2016), the spectral errors of the pyranometers are reported to be 2 to 5 %, which means that, for example, cloudy pixels with a GHI of 400 W m^{-2} could have a 20 W m^{-2} error.

11) ***Lines 461-463: Could a rough estimation of the temporal mismatch error in a worst-case scenario be added here? E.g. by moving a high cloud for 5 minutes with a high wind speed into a fixed direction and then estimating the induced mismatch?***

The spatial and temporal mismatch of surface solar radiation between satellite and in-situ measurements is studied in detail by Urraca et al. (2024). Cloud variability is identified as the main driver of both spatial and temporal mismatch. Consequently, the exact magnitude of the spatial and temporal mismatch depends on the time of day, season and location. Moreover, since cloud variability introduces variations both in the temporal and spatial domains, the errors cannot be considered to be independent. Therefore, temporal aggregation can also be seen as a means to correct the spatial mismatch.

Concerning this study, the analysis performed by Urraca et al. (2024) on the impact of spatial and temporal mismatch on the satellite product validation is the most relevant. The authors investigate mismatch errors by averaging BSRN measurements to wider temporal intervals (temporal mismatch) and by aggregating GHI from the SARAH-2.1 dataset (Pfeifroth et al., 2019) to coarser grids (spatial mismatch).

With a temporal smoothing of the 1-minute BSRN data (i.e. using a window ± 1 minute), a Mean Absolute Deviation (MAD) in the order of 30 W m^{-2} is found. The temporal mismatch is minimized when the products are averaged to a ± 14 minute time interval. In that case, MAD is significantly reduced to about 23 W m^{-2} (see Figure 6 in Urraca et al. (2024) and supporting information S6 and S7). This seems to align with observations from our previous study, where a minimal RMSE between the HOPE pyranometer network and the satellite retrievals is observed with a 20-minute averaging period (Wiltink et al., 2024).

For the spatial mismatch, Urraca et al. (2024) finds an MAD of about 30 W m^{-2} at a 0.05° resolution. When spatially averaged to 0.25°, a small but significant reduction in MAD of 2-3 W m^{-2} is observed. This looks counterintuitive, as a smaller spatial mismatch may be expected at higher resolution. The reduction in MAD at coarser grids might be explained by 3D radiative effects, which are more prominent in higher resolution products. However, also the neglect of parallax and cloud shadow corrections could increase MAD in higher resolution products. This is not explicitly mentioned in Urraca et al. (2024). In the current, but also in our previous study (Wiltink et al., 2024), where we correct for parallax using the daily mean optimal shift method, in general, GHI retrieved at a 1×1 km^2 resolution is significantly more accurate than at 3×3 km^2.

The analysis by Urraca et al. (2024) shows that although the validation metrics might improve by spatial averaging because of the reduced impact of random errors (e.g. 3D effects), it does not necessarily mean that the spatial and temporal mismatch are reduced as well.

We did think of a worst-case scenario, as suggested by the reviewer. Imagine a 1 × 1 km^2 cloud, covering a SEVIRI HR pixel at the acquisition time. The satellite-measured GHI could be 200 W m^{-2}. If that cloud moves with a (not extreme) speed of 60 km h^{-1}, it means a movement of 5 km in 5 minutes. With some simplification, we can estimate the pyranometer to measure a clear-sky GHI of (say) 800 W m^{-2} for 4 minutes and a cloudy-sky GHI of 200 W m^{-2} for 1 minute, which would result in a 5-minute-average GHI of

680 W m^{-2} and a mismatch error of 480 W m^{-2}. This is an excessive error and is not very useful for giving a rough indication. For that reason, we decided not to add such a worst-case scenario estimate to the paper.

We will include the following paragraphs in the manuscript, providing an extended discussion of spatio-temporal mismatch effects.

Page 21, line 472: Thus, even [..] mismatch does remain.

In Urraca et al. (2024), the authors investigated mismatch errors by averaging BSRN measurements to wider temporal intervals (temporal mismatch) and by aggregating GHI from the SARAH-2.1 dataset (Pfeifroth et al., 2019) to coarser pixel grids (spatial mismatch). They found that the mean absolute deviation (MAD) for the temporal mismatch was minimized with a ± 14 minute temporal averaging window, while the MAD for the spatial mismatch was smallest if the retrievals were smoothed to $0.25 \times 0.25°$ (see their Figure 6). The width of the optimal temporal averaging window agrees well with our previous results in Wiltink et al. (2024), where the RMSE was found to be smallest at a 20-minute averaged temporal resolution. However, the results for the spatial mismatch in Urraca et al. (2024) appear counterintuitive, and might be related to the neglect of parallax and cloud shadow correction, which becomes more important at higher resolution. In the current study, as well as in Wiltink et al. (2024), higher resolution does lead to better correspondence with ground-based observations. Higher resolution retrievals should indeed be [..].

12) ***Lines 464-472: Same as above for the temporal mismatch errors, could a rough estimation for the order of magnitude of the spatial mismatch errors be added here?***

    Please see our reaction to the comment about the temporal mismatch (#11).

13) ***Lines 478-484: Some interesting studies on 3D cloud effects have been done here (although with an application to trace gas retrievals instead of GHI, but still it might be interesting for the authors): https://amt.copernicus.org/articles/15/1587/2022/, https://amt.copernicus.org/articles/15/5743/2022/, https://amt.copernicus.org/articles/15/3481/2022/***

    We thank the reviewer for bringing these articles to our attention. Currently, we are also studying GHI retrieval errors in Meteosat SEVIRI due to 3D radiative effects. These articles might include useful information for that project.

Technical corrections:

14) ***Line 356: Remove the comma after Figure 7: "findings, Figure 7, shows"***

    The reviewer's comment has been modified accordingly.

15) ***Line 358: Remove the dot after Hc: "The median Hc. of the cirrostratus".***

    The reviewer's comment has been modified accordingly.

16) ***Line 391: Add a blank between "the" and "asymmetric".***

    The reviewer's comment has been modified accordingly.

17) ***Line 425: For instance, subtropical land regions → For instance, in subtropical land regions***

    The reviewer's comment has been modified accordingly.

References

Greuell, W., Meirink, J. F., and Wang, P.: Retrieval and validation of global, direct, and diffuse irradiance derived from SEVIRI satellite observations, Journal of Geophysical Research: Atmospheres, 118, 2340–2361, https://doi.org/10.1002/jgrd.50194, 2013.

Gueymard, C. A.: The sun's total and spectral irradiance for solar energy applications and solar radiation models, Solar Energy, 76, 423–453, https://doi.org/https://doi.org/10.1016/j.solener.2003.08.039, 2004.

Hamann, U., Walther, A., Baum, B., Bennartz, R., Bugliaro, L., Derrien, M., Francis, P. N., Heidinger, A., Joro, S., Kniffka, A., Le Gléau, H., Lockhoff, M., Lutz, H.-J., Meirink, J. F., Minnis, P., Palikonda, R., Roebeling, R., Thoss, A., Platnick, S., Watts, P., and Wind, G.: Remote sensing of cloud top pressure/height from SEVIRI: analysis of ten current retrieval algorithms, Atmospheric Measurement Techniques, 7, 2839–2867, https://doi.org/10.5194/amt-7-2839-2014, 2014.

Kuipers Munneke, P., Reijmer, C. H., van den Broeke, M. R., König-Langlo, G., Stammes, P., and Knap, W. H.: Analysis of clear-sky Antarctic snow albedo using observations and radiative transfer modeling, Journal of Geophysical Research: Atmospheres, 113, 17 118, https://doi.org/10.1029/2007JD009653, 2008.

Madhavan, B. L., Kalisch, J., and Macke, A.: Shortwave surface radiation network for observing small-scale cloud inhomogeneity fields, Atmospheric Measurement Techniques, 9, 1153–1166, https://doi.org/10.5194/amt-9-1153-2016, 2016.

Nakajima, T. and King, M. D.: Determination of the Optical Thickness and Effective Particle Radius of Clouds from Reflected Solar Radiation Measurements. Part I: Theory, Journal of Atmospheric Sciences, 47, 1878 – 1893, https://doi.org/10.1175/1520-0469(1990)047⟨1878:DOTOTA⟩2.0.CO;2, 1990.

Pfeifroth, U., Kothe, S., Trentmann, J., Hollmann, R., Fuchs, P., Kaiser, J., and Werscheck, M.: Surface Radiation Data Set - Heliosat (SARAH) - Edition 2.1, https://doi.org/10.5676/EUM_SAF_CM/SARAH/V002_01, 2019.

Urraca, R., Lanconelli, C., and Gobron, N.: Impact of the Spatio-Temporal Mismatch Between Satellite and In Situ Measurements on Validations of Surface Solar Radiation, Journal of Geophysical Research: Atmospheres, 129, https://doi.org/10.1029/2024JD041007, 2024.

Wiltink, J. I., Deneke, H., Saint-Drenan, Y.-M., Heerwaarden, C. C. V., and Meirink, J. F.: Validating global horizontal irradiance retrievals from Meteosat SEVIRI at increased spatial resolution against a dense network of ground-based observations, Atmos. Meas. Tech., 17, 6003–6024, https://doi.org/https://doi.org/10.5194/amt-17-6003-2024, 2024.