Reply to report of reviewer #1 "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

---

In this paper, the authors validate two corrections to satellite-based ground radiation retrievals using ground measurements: parallax correction and shadow correction. This is more involved than it looks at first glance and the paper contains some results that many might overlook. It presents some evidence that the parallax error and shadow error partially cancel out each other for geostationary-based radiation measurements, and that the combined error may therefore increase if only parallax correction is applied. The paper also shows that results improve when using a value smaller than the retrieved CTH, at least in some cloud regimes. These result should be stressed more clearly, including in the abstract. The science in this paper is good, but some conclusions should be presented more clearly. I have just one significant worry on the validation of the empirical collocation shift correction. Below is a list of suggestions before this paper can be accepted with minor revisions.

We thank the reviewer for the constructive feedback. Valid points were raised, which we addressed to further improve the quality of the manuscript. Please find our response below.

We tried to present the main findings more clearly in the manuscript and have made the following adjustments, starting with the conclusions Section:

Updates in conclusions:

We made the following adjustments to the conclusions Section:

- A paragraph has been added where the improvements of the different corrections w.r.t. the uncorrected retrieval are mentioned. Here, it is also more clearly stated that reducing H_c can further improve the accuracy of retrievals:
  Page 22, line 499: In general, GHI is retrieved most accurately when the time step optimal shift is performed, followed by the combined geometric shift. Compared to the uncorrected retrieval, the RMSE is reduced by 15.6 W m^{-2} (14.5 %) and 11.7 W m^{-2} (10.8 %) respectively. With the parallax-only or daily optimal shift correction a smaller improvement in accuracy is obtained because these correction methods do not account for diurnal variations in the cloud shadow position. Performing only a cloud shadow correction will in most cases even lead to an increase in RMSE as the correction is applied to the incorrect non-parallax-corrected cloud position.
  Depending on cloud regime and resolution, applying a parallax-only correction can also increase the RMSE of the retrieved GHI compared to applying no correction at all. On average, for the parallax-only correction, the best results are obtained if it is performed based on 40-50% of the originally retrieved cloud top height H_c. The reason is that by only performing a parallax correction, the cloud shadow position is not explicitly considered. However, in our study domain the cloud shadow displacement (away from the equator) is typically opposite to the parallax correction (toward the equator), so they partly cancel each other out. Thus, implicitly

the cloud shadow displacement is partly accounted for when the parallax correction is applied with reduced H_c. Overall, this underlines the need for a combined parallax and cloud shadow correction to achieve accurate GHI retrievals.
In addition, for the combined geometric correction, using a partial H_c of 70-90 % can also reduce the RMSE in the retrieval. This might be due to radiation scattered towards the satellite from altitudes lower than H_c, which would result in a smaller parallax and shadow displacement. However, the differences in RMSE with the full combined correction are not statistically significant.

- For readability we have rephrased and moved the statement about pixel- and area-based corrections:
  Page 23, line 536: [..] effect on GHI at the surface. Furthermore, the fully combined geometric correction shows a significant reduction in RMSE when the correction is performed with a median H_c around the region of interest rather than relying on the H_c of each pixel separately. The reductions are 2.4 and 1.5 W m^{-2} at HR and SR, respectively, indicating that the area-based correction better handles uncertainties related to the H_c retrieval. This study shows [..].
- We have condensed the paragraph on the effect of spatial resolution on correction accuracy.
- A new paragraph has been added where the findings are briefly put in a broader context as suggested by reviewer #2 (See our reply to that review).
- Finally some textual adjustments have been made to improve overall readability.

Updates in results Section:

The largest adjustments for clarity have been made in the abstract and conclusions section. In the results section, some small textual adjustments were included that should improve readability. We did not adjust the overall structure of the sections. The larger adjustments included in the results section are as follows:

Page 11, line 225: ~~If only a parallax correction is performed~~ → For the separate parallax correction, starting from a H_c of 0 %, [..].

Page 12, line 270: At both resolutions, applying a time step mean optimal shift results in the smallest RMSEs of the evaluated corrections~~.~~, 15.6 W m^{-2} (14.5 %) and 8.1 W m^{-2} (7.6 %) W m^{-2} lower than the respective uncorrected HR and SR retrievals.

Page 14, line 297: First, [..] for all cloud regimes. Moreover, for CR1 to CR4, the mean HR RMSE of the parallax corrected retrieval is larger than for the uncorrected retrieval. For all cloud regimes besides CR7 and CR9 the shadow correction increases the HR RMSE with respect to the uncorrected retrieval. Second, [..].)

Find below our reply to the specific comments of the reviewer:

Abstract

1) ***This abstract undersells the paper. It should prominently mention that applying parallax correction alone can worsen validation results (as shown in Figure 4) and that using a height smaller than the retrieved CTH may be better than the full height.***

   The following adjustments were made to the abstract:

   - The first sentence of the abstract has been removed and replaced by the following sentence:
     Page 1 line 1: Satellite-derived GHI is an excellent data source for nowcasting solar power generation and validating weather and climate models.
   - The effect of cloud shadow displacement has been clarified (Also see our reply to point #2):
     Page 1, line 7: The geolocation of satellite retrievals is affected by parallax, a displacement between the actual and apparent position of a cloud, as well as by a displacement between the actual position of a shadow and the ~~cloud casting that shadow [..]~~ retrieved position of the shadow, which due to the 1-dimensional radiative transfer assumption, is directly below the cloud.

- We have reformulated the results mentioned in the abstract. For instance, it is now mentioned that the parallax and shadow-only corrections can worsen the accuracy while a reduced H_c might further improve the correction accuracy:

  Page 1, line 12: The time step averaged collocation shift correction generally yields the most accurate results, but a major drawback of this method is its reliance on ground measurements. The geometric correction, which does not have this disadvantage, achieves the most accurate results if a combined parallax and shadow correction is performed. It reduces the GHI root mean square error (RMSE) by 11.7 W m^{-2} (10.8%) compared to the uncorrected retrieval. Separate parallax or shadow corrections do not reach this level of accuracy. In fact, depending on the cloud regime, they may even increase the error compared to the uncorrected retrieval. In some cases, in particular when multilevel clouds are present, the retrieval accuracy improves if the geometric correction is based on a reduced H_c. Finally, it is demonstrated that GHI becomes increasingly sensitive to the applied correction at higher spatial resolutions, especially for variable cloud regimes. This has important implications for the retrieval accuracy of the current generation of geostationary satellites with spatial resolutions down to 500 m.

2) ***Page 1, line 7: This line is a bit confusing. The geolocation of the shadow (on the ground), as visible from the satellite, should not be affected by parallax. The 1-D assumption of "shadow is straight below the cloud" should be mentioned here, an assumption that is of course erroneous in most cases.***

We thank the reviewer for pointing out the unclarity when it comes to the cloud shadow position correction.

The CPP-SICCS retrieval for GHI assumes 1D radiative transfer. As a result, the retrieval places cloud shadows directly below the cloud. The radiances measured by SEVIRI and the pyranometer network are obviously governed by 3D radiative transfer and thus cloud shadows are normally displaced with respect to the below-cloud position.

Since the TOA radiance from cloud shaded pixels is low, the CPP-SICCS retrieval will (incorrectly) flag cloud-shaded pixels as clear sky. With the cloud shadow correction, we use the cloud top height and solar position to compute the distance between the position directly below the cloud (i.e. the assumed location of the cloud shadow in our 1D retrieval) and the actual location on the surface of this cloud's shadow.

In the end, with the cloud shadow correction, we can still take into account the effects of possible cloud shadow displacements, despite the 1D assumption in the GHI retrieval (See Figure 2 in the manuscript).

We agree with the reviewer that much of the confusion can be prevented if the 1D assumption in our retrieval is pointed out earlier, since this assumption is a crucial part in the correction.

The 1D assumption is now mentioned in the abstract. We also mention the 1D assumption in the introduction:

Page 3, line 63: [..] in addition. GHI retrievals almost exclusively assume 1D radiative transfer and as a result, in most cases, cloud shadows are incorrectly projected directly below the cloud. To correct the retrieved cloud shadow location [..] 2024-Preprint). With this correction the retrieved cloud shadow position can be shifted to the actual position of the cloud shadow at the surface.

Introduction

3) ***Page 3, line 64: same here, the shadow location would be retrieved accurately by just measuring the radiance from the shadowy pixel? But that's not what is done?***

Please see our response to comment #2.

3.2 Shadow correction

4) ***Page 6, line 160: this 1-D assumption should be pointed out earlier (introduction and perhaps abstract) as otherwise the reader may be confused why there is an error in the shadow location, as a simple radiance retrieval without any assumptions will have the correct geolocation for the shadowed pixel (if the shadow is on the ground and not on a lower cloud).***

Please see our response to comment #2.

3.4 Empirical collocation shift correction

5) ***Page 8, line 195: variations in CTH are also not accounted for.***

The reviewer is correct in pointing out that another shortcoming for the daily mean optimal shift is that variations in CTH remain unaccounted for. Also the time step mean optimal shift does not account for the diurnal variation in CTH. This is because the shift computed for each time step is based on all corresponding times during the entire length of the field campaign. Only the geometric correction accounts for variations in cloud top height. The consequences of not accounting for cloud top height are also discussed in more detail in Section 4.2.3 about high clouds. We have decided to already provide a comment on this in the methodology:

Page 8, line 195: The mean optimal shift method does have some drawbacks. For instance, variations in H_c are not considered in the correction. ~~A~~ Another shortcoming of the daily mean optimal shift [..] unaccounted for.

6) ***Page 9, Figure 3, latitude shift: why is there no latitude shift (south/north) around 12 noon? Shouldn't there be a shadow in the north?***

In Figure 3, the diurnal variation in optimal longitude shift is more apparent than the latitudinal variation. For the longitudinal shift, the optimal shift location changes with shadow position from west to east and becomes zero around noon when the sun is directly south, meaning that there is no cloud shadow displacement in the east west direction at that moment.

We attribute the flatter diurnal curve for the latitude shift mainly to compensating effects of the cloud shadow position and parallax. In our domain (i.e. midlatitude northern hemisphere) around noon the cloud shadow will indeed be positioned north of the cloud as the sun is located in the south. Accounting for the cloud shadow position in our retrieval will result in a northward shift, while the parallax correction will result in a southward shift towards the subsatellite point at the equator. (We also discuss this in section 4.1.1 about the separate parallax and shadow corrections). Therefore, the effect of cloud shadow position and parallax partly cancel each other out. As this opposing effect is largest around noon, the curve for diurnal variations in optimal latitude shift is flattened. However, around noon, the mean optimal shift is not zero but ± 3.0 km south for both the time step and daily mean optimal shift.

7) ***Page 10/11, Figures 3 (both panels) and 4(right panel): please add a thin grid line at 0, like you have in figures 5, 7, 8.***

Grid lines at y=0 have been added.

4.1.3 Pixel-based and area-based corrections

8) ***Page 12, line 258: I think "uncertainties" here should be "errors", shouldn't it?***

We agree with the reviewer and have replaced "uncertainties" with "errors" in this sentence.

4.1.5 Resolution sensitivity

9) ***Page 12, line 283: You have derived the emipirical collocation shift correction using pyranometer measurements and now you are using comparisons with pyranometer measurements to validate them. It would seem your validation is not independent of your reference.***

The reviewer makes a valid point by noting that the HOPE pyranometer network is used both for the derivation of the empirical collocation shift as well as for the validation making the validation data not fully independent of the reference.

Normally one would use data from prior years to establish these optimal shifts but, these are not available here. Yet, we would argue that the influence on the results remains limited.

To compute the daily mean optimal shift, we use 96 days of pyranometer observations from the HOPE field campaign. Here, all timeslots between 06:15 and 16:45 UTC are used, which means that with the (averaged) 5 minute temporal resolution, each day 127 observations are made. In total 12192 observations are used to derive a single daily mean optimal shift.

Non-independence of the collocation shift from the reference data might be a larger concern for the time step optimal shift as the time step optimal shift is computed for each of the 127 timeslots separately. However, in this case each collocation shift is still based on 96 observations of the various days. These 96 days represent a large range of weather conditions and cloud types. Therefore, we expect the mean time step optimal shift to remain largely insensitive to the time-to-time variability in GHI measured by the pyranometers.

This now mentioned in the main text as well:
Page 10, line 215: Note that the same pyranometer data is used for computation of the optimal shift and evaluation of the accuarcy of the emperical collocation shift method, which makes the data not fully independent. Ideally, data from previous years would be used to establish the optimal shifts, but these are not available for the field campaign. Yet, to derive each optimal shift, large volumes of data are used, representing a wide range of weather conditions and cloud types. Therefore, we expect the optimal shifts to remain largely insensitive to time-to-time variability in GHI measured by the pyranometers.

4.2 Separation into cloud regimes

10) ***Page 13, figure 13: bit surprised by seeing the clear sky regime here, as there will be no parallax or shadow error. But that actually provides a baseline for improved analysis of the rest of the results. If you have ca. 50 W/m² RMSE just from other sources, couldn't you subtract that error estimate from all the other figures so you get an estimate that covers only parallax and shadow effects? Or at least indicate this in the other figures for context.***

It is correct that in the clear sky regime no parallax or cloud shadow errors should be present, although, minor errors might arise from the imperfect classification of clear sky scenes. However, the current errors for CR9 are considerably larger than can be attributed to this effect. In this article, the error within the clear sky regime is not treated as a baseline error due to the following reasons:

1) Firstly, the RMSE of around 48 W m^{-2} for CR9 mainly represents a bias between what is measured by the pyranometers and the SEVIRI retrieval. The GHI measured with the pyranometer network underestimates what is retrieved by the CPP-SICCS retrieval as well as by the McClear model (Gschwind et al., 2019). This is what has also been observed in our previous publication: Wiltink et al. (2024), for instance in Figure 7a. In Section 4.3 of that article we mention imperfect calibration and tilt of sensor as possible reasons for this deviation.
2) Secondly, our retrieval for clear-sky pixels deviates from the cloudy pixel retrieval. In fact for clear-sky pixels, the CPP-SICCS retrieval does not rely on SEVIRI reflectance but computes GHI directly from the NWP input. For cloudy pixels, a Nakajima and King (Nakajima and King, 1990) bispectral retrieval is performed to derive cloud optical depth and effective radius which can then be used to compute (broadband) GHI. As a consequence, the error in cloudy pixels has entirely different characteristics than in clear-sky pixels.
3) Finally, the errors for cloudy pixels are largely due to uncertainties in the retrieval, for example related to cloud inhomogeneity. These uncertainties vary per cloud regime and explain the largest part of the additional error compared to clear sky. Parallax and shadow displacement errors form a relatively minor contribution, as can be infered from the difference between corrected and uncorrected retrievals.

The comment made us realize that it is useful to mention that the clear sky bias of 48 W m^{-2} is not due to parallax or shadow effects. Therefore, we have added the following sentence to the manuscript:

Page 14, line 295: The best agreement [..] clear-sky situations (CR9). The errors observed in this regime are not the result of parallax or cloud shadow displacement, but originate in a bias between the SEVIRI retrieval and the HOPE pyranometer network. Possible causes of this bias are imperfect calibration and sensor tilt as identified in Madhavan et al. (2016) and Wiltink et al. (2024).

11) ***Page 13, figure 13 caption: Note that the "natural color RGB" is known as the "day land cloud RGB" in the USA and perhaps some other communities. Adding this name (in addition to "natural colour" common in europe") 5may help some readers.***

We have adopted the term "natural color RGB" from the Satpy Python Library. The naming for this RGB-composite is in line with how it is used by EUMETSAT. For clarity, we have added the term "Day land cloud RGB" as well:

Page 13, Figure 5, caption: Satpy natural color RGBs of SEVIRI → Satpy natural color RGBs (also referred to as "Day land cloud RGB") of SEVIRI.

4.3 Diurnal Cycle

12) ***Page 19, line 398: "growing importance at increasing spatial resolutions", here one could mention FCI.***

This paragraph provides a good opportunity to mention the current generation of geostationary satellites that have spatial resolutions that go down to scales of 500 m. We do make a statement about this, however, this is only in the conclusions/ outlook. We have added the following sentence to the manuscript:

Page 19, line 398: Overall this illustrates [..] spatial resolutions. This also makes it increasingly relevant for the current generation of geostationary satellites like the GOES Advanced Baseline Imager (GOES ABI; Schmit et al., 2017) and the Meteosat Third Generation Flexible Combined Imager (MTG-FCI; Holmlund et al., 2021) which enable retrievals of GHI down to scales of 500 m.

And have slightly adjusted the part in the conclusions/outlook where this is mentioned.

Page 23, line 538: The current generation of geostationary satellites like the GOES Advanced Baseline Imager (GOES ABI; Schmit et al., 2017) and the Meteosat Third Generation Flexible Combined Imager (MTG-FCI; Holmlund et al., 2021) enables retrievals of GHI down to scales of 500 m. → The current generation of geostationary satellites like the GOES ABI and the MTG-FCI enables retrievals of GHI down to scales of 500 m.

5.1 Generalizability of results

13) ***Page 19, line 414: this could be tested with IODC. Not suggesting you need to do it for this study, but you could mention it at least.***

Comparing retrievals from the prime and/or rapid scan service of Meteosat with the IODC service would indeed be a good way to test the relative importance of east-west versus north-south parallax. However, for the period of the HOPE campaign which we use in this study, direct comparison with IODC is not possible since the IODC service only became operational in 2016 and the campaign took place in 2013. We do agree with the reviewer that it is worth mentioning that there is a possibility for such a comparison at later periods, and have added the following sentence to the manuscript:

Page 12, line 283: For regions [..] more impact on the retrieval accuracy. The relative importance of the north-south and east-west parallax could be studied in more detail by comparing retrievals from the MSG Prime or RSS service, for which the satellite is positioned at 0/9.5 °E, to the MSG Indian Ocean Data Coverage (IODC) service, for which the satellite is positioned further east at 41.5/45.5 °E. However, this comparison is not possible for the dates of the HOPE field campaign as the MSG-IODC service became operational in 2016.

5.2 Remaining mismatch errors

14) ***Page 21, line 473: this could be tested with LEO, such as Sentinel data or even Landsat (if calibration is good enough).***

It would be interesting to test the magnitude of the spatial mismatch with LEO satellites. These higher resolution retrievals would become more sensitive to temporal misalignment and 3D radiative effects. Therefore, it would also be interesting to study the effect of these higher resolutions on the cloud shadow displacement and parallax corrections. This is beyond the scope of the current study.

15) ***Page 21, line 479: the comment on 3D is a bit oddly formulated.***

We have rephrased this sentence as follows:

Page 21, line 479: Furthermore, these 3D effects might limit the accuracy of geometric parallax and shadow corrections, for instance, when it comes to cloud-side illumination → For instance, 3D effects like cloud-side illumination might limit the accuracy of geometric parallax and shadow corrections.

Typos:

16) ***page 19, line 391, missing space in "theassymmetric".***

The reviewer's comment has been modified accordingly.

17) ***page 24, line 563: Github → GitHub***

The reviewer's comment has been modified accordingly.

References

Gschwind, B., Wald, L., Blanc, P., Lefèvre, M., Schroedter-Homscheidt, M., and Arola, A.: Improving the McClear model estimating the downwelling solar radiation at ground level in cloud-free conditions – McClear-v3, Meteorologische Zeitschrift, 28, 147–163, https://doi.org/10.1127/METZ/2019/0946, (data available at: `https://ads.atmosphere.copernicus.eu/datasets/cams-solar-radiation-timeseries`, last access: 7 March 2024)., 2019.

Holmlund, K., Grandell, J., Schmetz, J., Stuhlmann, R., Bojkov, B., Munro, R., Lekouara, M., Coppens, D., Viticchie, B., August, T., Theodore, B., Watts, P., Dobber, M., Fowler, G., Bojinski, S., Schmid, A., Salonen, K., Tjemkes, S., Aminou, D., and Blythe, P.: Meteosat Third Generation (MTG): Continuation and Innovation of Observations from Geostationary Orbit, Bulletin of the American Meteorological Society, 102, E990–E1015, https://doi.org/10.1175/BAMS-D-19-0304.1, 2021.

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.

Schmit, T. J., Griffith, P., Gunshor, M. M., Daniels, J. M., Goodman, S. J., and Lebair, W. J.: A Closer Look at the ABI on the GOES-R Series, Bulletin of the American Meteorological Society, 98, 681–698, https://doi.org/10.1175/BAMS-D-15-00230.1, 2017.

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.

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).~~ $\rightarrow$ 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 $\pm$ 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 $\pm$ 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 $\pm$ 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.