

Reply to Reviewer #1

Major Comments

The reason for the negative GHI biases in the EarthCare product is not well explained in this study. While the analysis shows that aerosol optical depth cannot explain the biases, it would be relevant to discuss other possible causes. Additionally, the reasons for the differences in GHI between EarthCare and CAMS are not explored, except for regions where positive differences appear (e.g., Sahara and Amazon). I do not expect the exact reasons to be fully resolved here, but the manuscript should clearly outline plausible contributing factors.

We thank the reviewer for highlighting the need for a clearer discussion regarding the drivers of the observed negative GHI biases and dataset discrepancies. We have expanded our discussion across several areas to provide a more comprehensive analysis of plausible contributing factors:

1. Solar zenith angle (SZA) specification error: We have added a technical subsection (Sect. 4.5) detailing an SZA specification error discovered in EarthCARE's ACM-RT product under baseline BA. This issue introduced an artificial underestimation in surface irradiance and accounts for a significant portion of the observed negative biases across both clear-sky (from -17.9 W m^{-2} to -5.4 W m^{-2}) and all-sky (from -9.8 W m^{-2} to 0.1 W m^{-2}) evaluations. In addition, a sentence has been added in the abstract: *"Most of this underestimation is traced to a solar zenith angle specification issue in the corresponding ACM-RT baseline of this study."*
2. BSRN validation factors: For the BSRN comparisons, we expanded the discussion to emphasize how the performance of the Scene Construction Algorithm (SCA) degrades with increasing off-track distance from the satellite ground track (Lines 505–509): *"Another reason could be the degradation of the SCA's performance with increasing distance from the satellite's ground track. Indeed, for distances below 25km, the 7 resulting EarthCARE overpasses are in close agreement with BSRN observations ($MBE = -4.5 \text{ W m}^{-2}$, $RMSE = 12.0 \text{ W m}^{-2}$), while the 12 overpasses that are further than 25km away from the ground station exhibit a significant negative bias ($MBE = -25.8 \text{ W m}^{-2}$) and much larger discrepancies ($RMSE = 50.7 \text{ W m}^{-2}$)."*
3. ACM-RT input field uncertainties: In the BSRN discussion, we added other possible variables which could contribute to the clear-sky bias (Lines 502–505): *"Beyond the SZA specification error, the remaining discrepancies under clear skies between ACM-RT and BSRN are likely related to uncertainties in the ACM-RT input fields. The residual negative bias cannot be explained by AOD alone (Sect. 3.3) and may also reflect errors in other aerosol optical properties (e.g. single scattering albedo), surface albedo and columnar water vapour."*
4. EarthCARE vs. CAMS Discrepancies: As already stated in the manuscript (Lines 736–737), our clear-sky irradiance and AOD comparisons with CAMS incorporate and discuss other influential atmospheric drivers beyond aerosol optical depth: *"Irradiance levels are also significantly influenced by additional aerosol optical properties (such as single scattering albedo), surface albedo, columnar water vapour, and varying concentrations of trace gases (e.g., O_3 , NO_2)."*

Specific Comments

* Section 2.2: I assume there are also uncertainties in the TOA fluxes derived from BBR measurements due to the unfiltering and angular correction processes. Please provide the uncertainty ranges of the

BBR fluxes or include relevant references. Are there studies comparing the accuracy of CERES and BBR TOA fluxes?

We have added BBR uncertainty ranges and a study discussing the comparison between CERES and BBR TOA fluxes to Section 2.2, Lines 192–199:

” The BBR serves as a key instrument for evaluating radiative closure and assessing uncertainties in the synergistic products derived from the other three EarthCARE sensors. However, its inherent uncertainties, originating from radiometric noise, spectral unfiltering, and multi-view flux inversion procedures, must be taken into account. In pre-launch synthetic evaluation studies, Velázquez Blázquez et al. (2024b) found that the RMSE of the SW combined flux in the BMA-FLX product ranged between 7 W m^{-2} and 13 W m^{-2} across three simulated test scenes using model-truth cloud fields. For early real EarthCARE observations, Clerboux et al. (2025) compared BBR SW combined flux estimates against the CERES Single Scanner Footprint product, yielding an RMSE of 47.8 W m^{-2} and an MBE of 2.6 W m^{-2} . ”

* Line 197: Figure 1 could potentially be moved to Appendix A, considering the large number of figures in the main manuscript already.

We thank the reviewer for highlighting the need to optimize the figure count. To address this concern while preserving essential context for readers in the main manuscript, we decided on merging Figures 1, 2, and 5 into a single map (new Figure 1). This new Figure simultaneously presents the spatial coverage of the CAMS radiation service, the ground pyranometer stations, and the AERONET stations. This merging reduces the overall figure count in the main body by two figures while giving the reader an overview of all spatial domains and ground networks. Accordingly, the introduction of the AERONET stations used in the analysis has been moved prior to Figure 1, relocating from previous Section 2.9 to Section 2.5 (Lines 272–274).

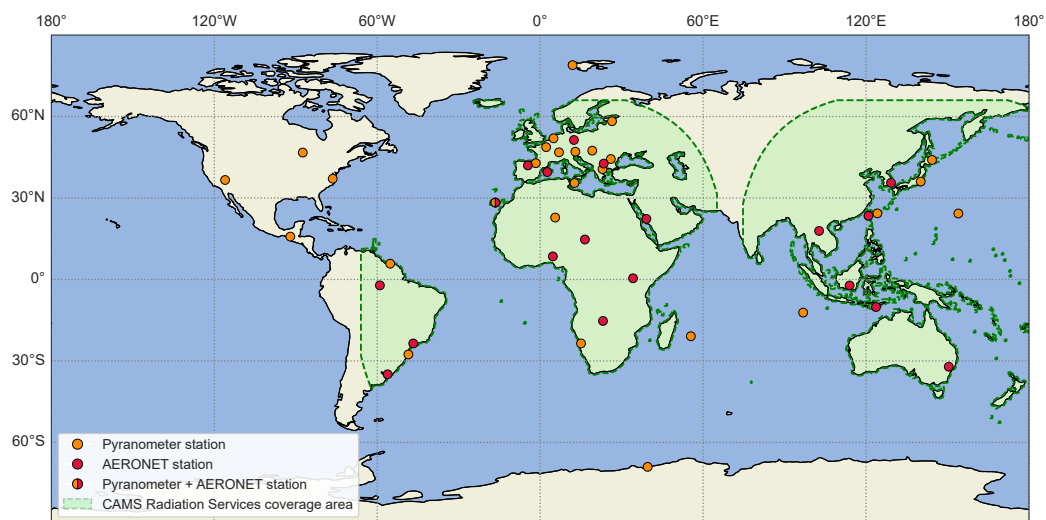


Figure 1: Spatial coverage of the surface GHI product estimated by CAMS Radiation Services (green dashed area) and geographical distribution of the ground stations. Shown are the 28 pyranometer stations (orange circles) and 19 AERONET stations (red circles) used in the analysis. Co-located BSRN and AERONET stations are indicated by split-color markers. The map is displayed as a Plate Carrée projection.

* Line 199: So only daytime EarthCare orbits were compared? This implies that daytime comparison was also considered for LW. Since daytime and nighttime LW biases often differ, it would be relevant to clarify that LW biases were examined only during daytime.

Yes, only daytime EarthCARE orbits were compared, including for LW fluxes, as the primary objective

of this work was to evaluate surface solar irradiance. To improve clarity, we have added the following sentence in Line 364:

"Once again, only daytime overpasses were included in the analysis."

* Line 210: Have you compared cloud fractions and cloud optical depths estimated from EarthCare CAPTIVATE and the McCloud model? These two parameters should be central to explaining differences in GHI for total skies (but not cloud-free regions).

We thank the reviewer for this insightful comment. We agree that cloud fraction and cloud optical depths are key drivers of total-sky GHI discrepancies. In this study, we focused our evaluation on the Cloud Modification Factor ($CMF = \frac{G_{all-sky}}{G_{clear-sky}}$) for both EarthCare CAPTIVATE and the McCloud model. The CMF integrates the combined radiative impact of the cloud fraction and the cloud optical depth into a single metric representing cloud-induced GHI attenuation. Consequently, we chose the CMF because it encompasses all cloud-related influences into a single metric, leaving the detailed exploration of individual cloud parameters for future work once at least one full year of data is available.

* Line 517: I assume the CAMS radiation dataset is also produced using the 1D RTM. If so, would the use of a 1D model for the EarthCare give better consistency between EarthCare and CAMS, assuming similar cloud properties? However, cloud fractions and optical depths from the two products may differ significantly, likely contributing to the GHI differences (as discussed in Fig. 10).

Yes, the reviewer raises a very valid point. The CAMS radiation dataset is generated using 1D RTMs, making EarthCare's 1D RTM methodologically more consistent with CAMS than the 3D product. This point is already addressed in the manuscript (Lines 592–593):

"As CAMS' McCloud and McClear models utilize 1D RTMs, EarthCare's 1D RTM is expected to give more comparable results than its 3D counterpart."

As the reviewer highlights, the comparison relies heavily on cloud fractions and optical depths. Nevertheless, this consistency is suggested by the lower RMSE obtained when evaluating EarthCare's 1D RTM against CAMS, compared to the 3D RTM over the same assessment domain (Fig. C1 in the manuscript). Therefore, to make the statement more robust, the following sentence was added in Lines 593–595:

"This is supported by EarthCare's 1D estimates yielding closer agreement with CAMS than the 3D estimates ($RMSE = 86.1 \text{ W m}^{-2}$ vs. 91.2 W m^{-2}) over the same assessment domain (Fig. C1)."

* Fig. 12b: Could the authors explain the large outliers when EarthCare GHI is 600–800 W/m^2 and CAMS GHI is 400–600 W/m^2 ? The majority of points lie near the one-to-one line, indicating strong correlation. However, after accounting for outliers, the mean negative difference suggests systematically smaller values in EarthCare GHI relative to CAMS.

We thank the reviewer for this observation. Further inspection reveals that these outliers correspond primarily to overpasses over South America. When comparing these specific cases (Fig. 2a) against independent low-AOD clear-sky irradiance calculations (accounting for the exact geolocation and timestamp of the overpass), the resulting irradiance aligns closely with the higher EarthCare values (600–800 W/m^2). While higher CAMS AOD estimates during South American biomass burning events could contribute to this discrepancy, only a few cases showed AOD differences large enough to fully explain it (Fig. 2b). This points to a localized underestimation of clear-sky irradiance within the CAMS dataset for these overpasses. The following three sentences were added in Lines 633–636:

"Some outliers can be observed which exhibit lower CAMS GHI (400–600 W m^{-2}) relative to EarthCare (600 – 800 W m^{-2}). These outliers correspond primarily to overpasses in South America. While higher CAMS AOD estimates contribute to some of these cases, independent clear-sky irradiance evaluation

using the SZA and standard atmospheric parameters indicates a localized underestimation of GHI within the CAMS product for these scenes.”

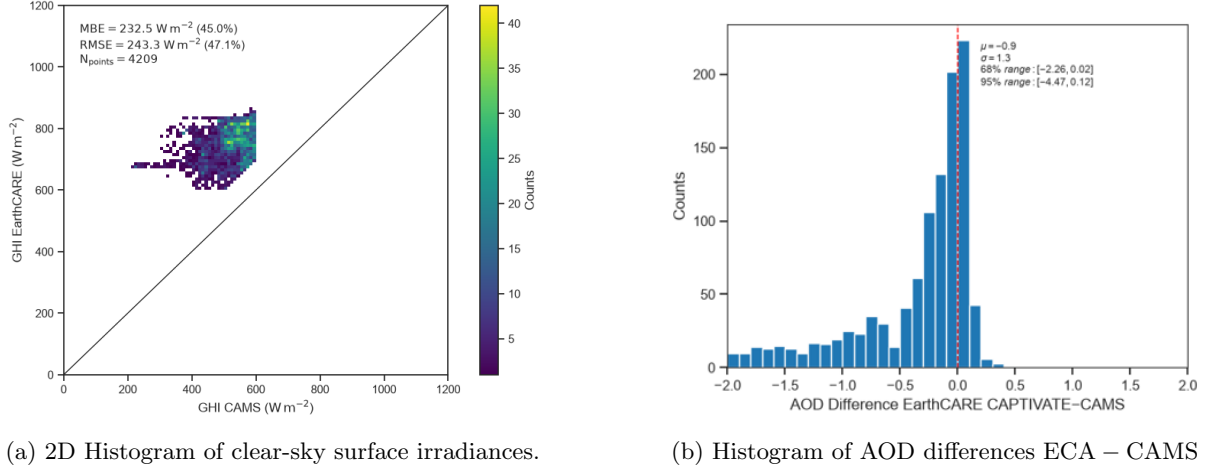


Figure 2: GHI and AOD comparison for outlier subset.

* I see the point of including Fig. 13 to motivate the scale used in later figures. However, Fig. 13 could be replaced by simply referencing Line 573: “In contrast, the MBE remains scale-invariant...”

We agree that the figure can be omitted without loss of clarity. Figure 13 has been removed from the manuscript, and the key values across spatial scales have been integrated directly into the text (Lines 644–657).

* In Table 1, when the EarthCARE GHI was compared with BSRN, EarthCARE GHI was biased low for cloud-free cases, and EarthCARE GHI was biased high for overcast scenes. In Fig. 11, EarthCare GHI is smaller than CAMS for cloud-free scenes (CMF = 1), and EarthCARE GHI is larger than CAMS for overcast scenes (CMF = 0). Fig. 15 shows that BSRN is somewhere between EarthCARE and CAMS for all skies. How about cloud-free cases specifically—does BSRN lie between the two? What are the main factors deriving all these differences for cloud-free cases? Since AOT was ruled out, could the authors suggest other parameters driving differences under cloud-free conditions?

We thank the reviewer for this insightful question. Regarding cloud-free conditions, BSRN indeed lies between EarthCARE and CAMS (Fig. 3). However, because there were only four matching cloud-free data points available for this comparison, the sample size was too sparse to draw statistically robust conclusions and was therefore omitted from the manuscript text. Regarding the primary drivers behind these cloud-free differences, an explanation is provided in our response to the Major Comment. In short, a significant portion of the negative bias observed in EarthCARE’s ACM-RT baseline BA product was traced back to a solar zenith angle (SZA) specification error (now detailed in Sect. 4.5). Accounting for this SZA fix substantially reduces the cloud-free negative bias against BSRN (from -17.9 W m^{-2} down to -5.4 W m^{-2}).

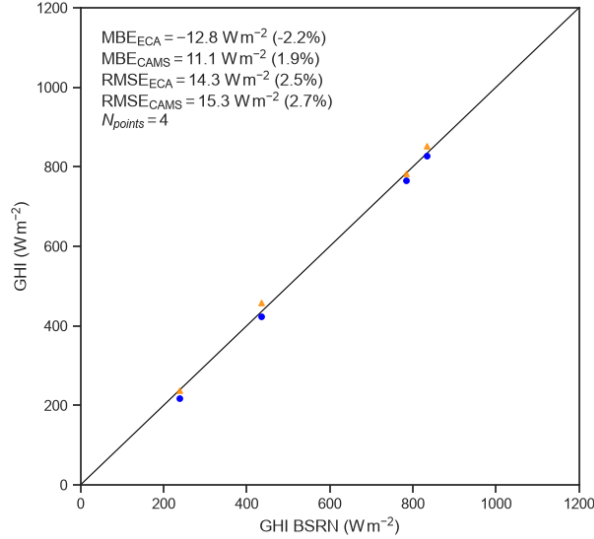


Figure 3: Comparison between EarthCARE (blue circles) and CAMS (orange triangles) GHI, validated by BSRN observations. Data points were recorded under cloud-free conditions.

* Considering CAMS uses multiple geostationary platforms, are there discrepancies in cloud properties depending on the satellite source?

Yes, there are discrepancies in cloud properties depending on the geostationary platform. The AHI (Advanced Himawari Imager) onboard the Himawari 8 satellite produces higher CMF bias (-0.036) and RMSE (0.138) relative to EarthCARE than the SEVIRI (Spinning Enhanced Visible and Infrared Imager) instrument onboard the MSG satellites (bias = -0.009 and RMSE = 0.125). We appreciate the reviewer bringing attention to this important point. In the manuscript, this platform dependency is discussed with a primary focus on GHI rather than CMF to align with CAMS' internal validation, which benchmarks performance using GHI (Lines 677–681):

"Finally, while CAMS GHI estimates using the MSG satellites (covering Europe, Africa, the Middle East) do not exhibit any significant bias according to their own validation, estimates based on the Himawari 8 satellite (covering eastern Asia and Oceania) exhibit a documented positive MBE of 25.7 W m^{-2} (Lefèvre, 2024). Consequently, the -31.4 W m^{-2} MBE observed in this region can be primarily attributed to the known overestimation inherent in the CAMS Himawari 8 retrievals rather than to an EarthCARE deficiency."

* Fig. 14c: EarthCare GHI is larger than CAMS GHI over the Sahara and Amazon, likely due to smaller AODs in EarthCare. Other regions show opposite signs. Since AOD does not explain these differences, could the authors suggest additional relevant parameters?

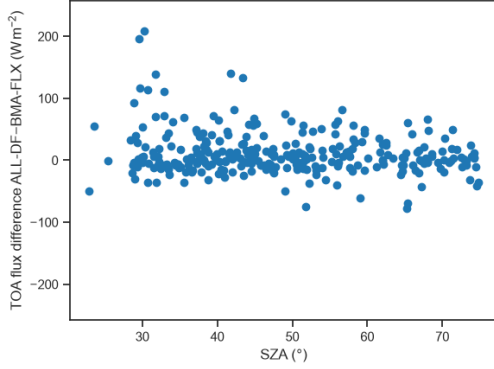
We thank the reviewer for this observation. While AOD is a key factor, we agree that other parameters play an important role in driving clear-sky GHI differences outside these specific regions. To address this, the manuscript already outlines additional atmospheric and surface drivers beyond AOD (Lines 736–737):

"Irradiance levels are also significantly influenced by additional aerosol optical properties (such as single scattering albedo), surface albedo, columnar water vapour, and varying concentrations of trace gases (e.g., O_3 , NO_2)."

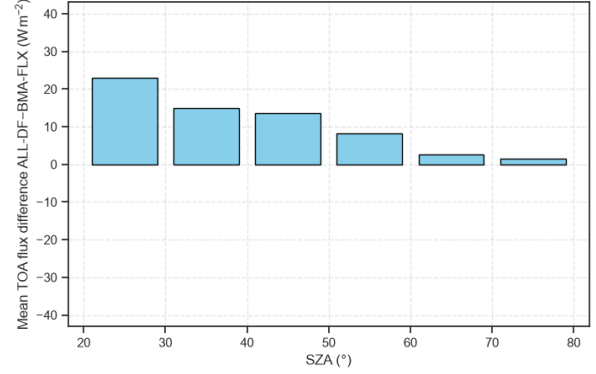
* When comparing RT simulations with TOA BBR or ground measurements, are the biases dependent on solar zenith angles?

Yes, a dependency on the solar zenith angle can be observed when comparing 1D RT simulations against validation datasets. For TOA BBR measurements (Fig. 4), lower solar zenith angles ($SZA < 40^\circ$) exhibit

a positive skew in flux differences exceeding $+100 \text{ W m}^{-2}$ with binned mean biases up to $+23 \text{ W m}^{-2}$ (8.6%), whereas higher SZAs yield mean differences near zero. A similar pattern appears in the surface GHI comparison against BSRN (Fig. 5), where mean biases are positive at low SZAs ($\text{SZA} < 40^\circ$) and shift to negative values at higher SZAs. However, these surface differences are strongly driven by high spatial and radiative variability under partly cloudy conditions. Similar trends can be observed for relative flux differences (Fig. 6). While evaluating SZA dependencies holds significant promise for future studies, the current ground dataset remains too small for a statistically robust SZA-dependent evaluation. Furthermore, TOA BBR measurements are utilized in this study primarily as auxiliary context to interpret surface flux discrepancies rather than as a standalone validation.

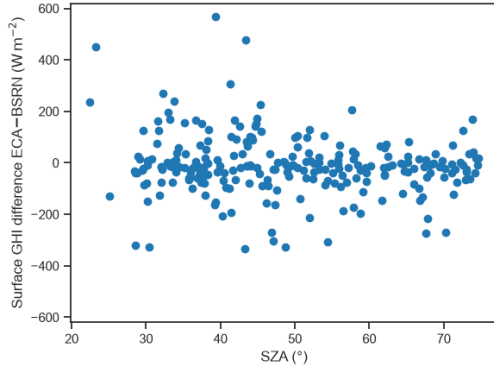


(a) Scatter plot of TOA flux differences.

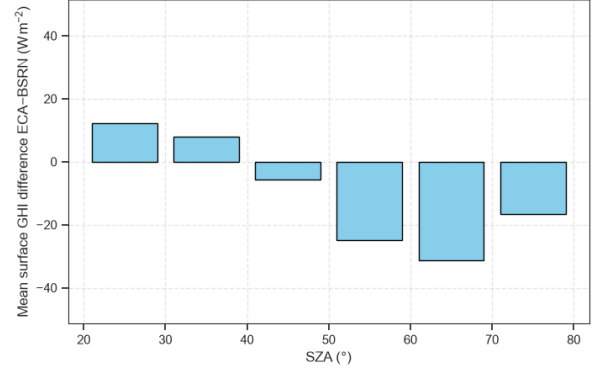


(b) Binned TOA flux differences across SZA intervals.

Figure 4: Dependence of TOA flux differences (ALL-DF – BMA-FLX) on solar zenith angles, showing (a) individual overpass values and (b) mean differences binned by SZA.

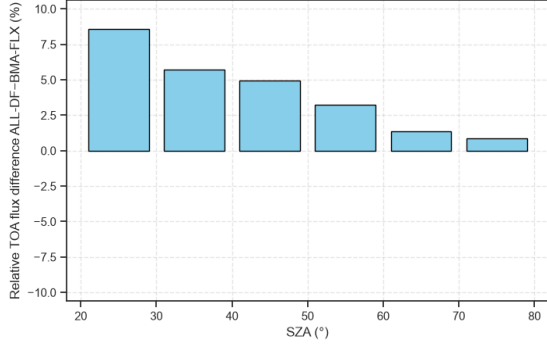


(a) Scatter plot of TOA flux differences.

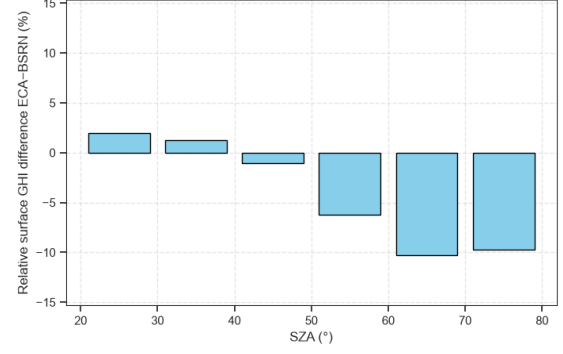


(b) Binned TOA flux differences across SZA intervals.

Figure 5: Dependence of surface GHI differences (EarthCARE – BSRN) on solar zenith angles, showing (a) individual overpass values and (b) mean differences binned by SZA.



(a) Relative TOA flux differences to BBR.



(b) Relative surface flux differences to BSRN.

Figure 6: Dependence of relative GHI differences (EarthCARE – BBR/BSRN) on solar zenith angles, showing (a) relative TOA flux differences to BBR (b) relative surface flux differences to BSRN.

* I understand that the comparison in Fig.18 was performed to ensure consistent sampling between the surface (SFC) and TOA evaluations. However, for the TOA comparison, additional sampling could potentially be included between the ACM-RT product and the BBR measurements. Have you, or has any previous study, carried out a direct TOA comparison between ACM-RT and BBR measurements? If so, are the findings of this study consistent with those earlier comparisons?

Yes, EarthCARE’s algorithm team regularly validates the 1D and 3D radiative transfer outputs against BBR measurements, such as those presented at the EarthCARE Science and Validation Workshop 2026 (Salas Molar et al. (2026)). Evaluating over a significantly larger sample size in the baseline BA, they found the opposite trend to ours: EarthCARE’s 3D algorithm performed better than the 1D algorithm. Consequently, our results reflect specific local conditions and sample size limitations, and should not be generalized as a broader trend. This point is already addressed in the manuscript (Lines 757–760):

” The 3D RTM exhibits greater deviation from the BBR observations, with an $rRMSE$ of 18.0 % compared to 14.9 % for the 1D RTM. However, Salas Molar et al. (2026) found the opposite trend over a significantly larger sample size in the baseline BA. Consequently, our results reflect specific local conditions and should not be generalized as a broader trend.”