Answers to Referee #1

The authors thank the referee for the evaluation of the manuscript. Please find our answers (blue text) to the comments (black text) below.

Review of "Retrieval of top-of-atmosphere fluxes from combined EarthCARE LiDAR, imager and broadband radiometer observations: the BMA-FLX product" by Velázquez Blázquez et al.

25 June 2024

General comments

This paper describes the BMA-FLX product that is used to retrieve radiative fluxes from the broadband radiometer (BBR) on the EarthCARE satellite mission. The product retrieves a SW and LW radiative flux for each of the three view angles of the BBR and then averages them to provide a best estimate of the fluxes. This best estimate is compared to the calculated fluxes for some example scenes, with an overall goal of obtaining radiative closure within 10 W/m2.

It is important to document the details of the BMA-FLX product so that future users of EarthCARE data can understand how the data is generated. The paper is thorough and mostly well written. However, I believe some substantial additional explanations are needed. I am also left concerned as to whether the performance of the product is sufficient to meet the EarthCARE goals based on the analysis presented. I outline these concerns in my specific comments below. After addressing these comments, I believe the study could be appropriate for publication in AMT.

Specific comments

Terminology: There is some debate in the ERB community as to whether "flux" is the correct term for the quantity considered in this paper that has units of W/m2. While "flux" is widely used, some argue that it is fundamentally wrong and that "flux density" is correct. Another alternative, "irradiance", is becoming more widely used because it is less ambiguous. I do not insist that the authors rename their product, but I do think that it is worth mentioning in the introduction that several different terms exist in the literature. These terms are often used interchangeably within the scientific community with some debate, but they all refer to the same quantity.

Thanks for your comment. This has been mentioned in the introduction as:
Various terms exist in the literature to refer to radiative fluxes, such as flux density and irradiance. Although these terms are often used interchangeably within the scientific community, with some debate, they all refer to the same quantity.

L57: It would be good to add a sentence or two to justify why the EarthCARE radiative closure goal is 10 W/m2, otherwise this seems rather arbitrary. What is the physical origin of this requirement? What does 10 W/m2 achieve that eg. 15 W/m2 does not? What limits a more ambitious goal of eg. 5 W/m2? Also, is this requirement defined based on the RMSE, which seems to be a focus of the results later in the paper and in the conclusions?

Thanks for your comment. Please let us refer to the official and public documentation for EarthCARE to answer your question. The EarthCARE radiative closure's goal of 10 Wm-2 is defined in the EarthCARE Mission Requirements Document (MRD) (Wehr T., 2006). The accuracy requirements placed under the

EarthCARE measurements are there described and justified. We added the corresponding reference to the manuscript in the text for better context.

T. Wehr, editor. 2006. EarthCARE Mission Requirements Document. Earth and Mission Science Division, European Space Agency, https://doi.org/10.5270/esa.earthcare-mrd.2006

L66-69: Following from the previous comment, a more general question is why is the EarthCARE radiative closure goal defined in terms of radiative flux? As noted in the paper, the conversion from radiance to flux inevitably introduces a substantial uncertainty. Unless I am missing something, it seems entirely possible to use a radiative transfer model to calculate *radiances* at the three view angles of the BBR using the scene properties derived from the other EarthCARE instruments, and compare those radiances to the directly observed radiances from the BBR. While flux is usually the quantity of most interest for atmospheric energetics, I think for the purpose of radiative closure, performing the closure in radiance-space provides a much tighter constraint and avoids the introduction of ADM uncertainty. Some mention of the justification to do radiative closure with fluxes in the introduction of this paper would be helpful to further motivate t*he BMA-FLX product.*

The radiative closure goal of EarthCARE is defined in the MRD and it stipulates that a retrieved (cloud–aerosol–precipitation) scene with a footprint size of 10 km×10 km shall be sufficiently accurate for its atmospheric vertical profile of short-wave (solar) and long-wave (thermal) flux to be reconstructed with an accuracy of 10 Wm−2 at the top of the atmosphere. (Wehr et al, 2023, https://doi.org/10.5194/amt-16-3581-2023)

This is a very pertinent and interesting question that has been properly addressed in an independent paper published on the radiative closure in this special issue. Please visit that paper for a complete picture.

H. W. Barker, J. N. S. Cole, N. Villefranque, Z. Qu, A. Velázquez Blázquez, C. Domenech, S. L. Mason, and R. J. Hogan. "Radiative Closure Assessment of Retrieved Cloud and Aerosol Properties for the EarthCARE Mission: The ACMB-DF Product". Atmos. Meas. Tech., Special issue: EarthCARE Level 2 algorithms and data products, 2024. [Submitted]. Preprint to be available soon in the EarthCARE AMT special issue: https://amt.copernicus.org/articles/special_issue1156.html

L77-81: It is understandable that the algorithm needs to meet contingency requirements of the mission, but at the same time it does seem like a missed opportunity that the flux retrieval does not take advantage of the multi-angle views of BBR directly to constrain the radiance to flux conversion. Could this be mentioned as a potential future research activity?

Thanks for your suggestion. That approach was thoughtfully tested in the early stages of the BMA-FLX processor development. Several references included in the text ((Domenech and Wehr, 2011; Domenech et al., 2011a) describe methodologies that use the three BBR along-track radiances simultaneously for the retrieval of radiative fluxes. While the results were promising, the Agency ultimately decided to adopt a more conservative approach for the operational processor.

In any case, new approaches exploiting the multi-angular capabilities of the BBR will be explored though in later iterations of the processor. This has been mentioned and clarified in the text.

L81: I think there is an error here. The outgoing flux should NOT depend on the viewing geometry. Do the authors mean the solar geometry?

Yes, that's a typo. Thanks for spotting it. Changed to: As the outgoing flux is only dependent on the solar geometry and the radiometric properties of the atmospheric-surface domain, the three fluxes estimated by the ADMs should result in a similar flux assuming perfect instrument and retrieval responses.

L103-105: These two sentences do not make sense to me. I think I know what the authors mean, but I suggest revising the wording.

Thanks, for your comment. The paragraph has been rewritten as follows: "Scene definition refers to the classification of targets based on surface and cloud properties, as well as the angular geometry that defines the ADM model used for BBR observation. During the ADM development, the stratification of the scene definition, scene classes, determines the number of anisotropic models, which in turn dictates the number of datasets required for training the networks that construct the ADM."

L108: I do not understand why the scene type is defined separately for forward and backward scattering directions. Please clarify in the paper why this is needed.

Rewritten for clarification as: "The scene is defined separately for forward and backward scattering directions because the angular geometry significantly impacts the observed radiative properties of the target. This detailed classification is necessary because the scattering characteristics of a scene can vary significantly between forward and backward directions. By accounting for these variations, the ADM can more accurately model and predict the radiative properties of different scenes. This approach ensures that the ADM is trained on a comprehensive set of datasets, enhancing its reliability in various observational conditions."

L134-135: Using an AOD climatology might be an important source of uncertainty here. This is because the large spatio-temporal variability of aerosol means that the actual AOD for any given CERES measurement can be quite different from the climatology and therefore create a different anisotropy. Do the authors think this is important and have they considered using an aerosol retrieval product from MODIS to match instantaneously with the CERES measurements?

Thanks for your suggestion. That has been indeed considered for future developments. MODIS AOD cannot be used as part of the EarthCARE operational processing chain, however the AOD product from the MSI (M-AOT, Docter et al., 2023) is being currently being checked. To avoid introducing a new dependency the AOD climatology is still used but newer versions of the product will consider processing fluxes using NRT AOD retrievals. Discussed in the "Conclusions" section.

L143: I like the idea of using the MSI radiances directly as an input to the ANN because they contain information about cloud properties without introducing uncertainty from an intermediate cloud property retrieval. That said, I am a bit surprised by the choice of bands: 0.67, 0.865, 10.8, and 12.0 µm. The most relevant cloud properties for anisotropy are cloud fraction, optical depth, phase, and effective radius. Are the authors confident that information about these cloud properties is sufficiently represented by the 4 MSI bands that are chosen? Some supporting references would help.

Thank you for your comment. This is a very good point. The selection of MSI bands was largely influenced by the availability of bands with similar central wavelengths in the MODIS PSF-weighted radiances provided in the CERES SSF products. The authors prepared a CERES database using all daylight data from the following CERES SSF Ed4 products:

- CERES FM1, FM2 (Terra), FM3 and FM4 (Aqua) instruments in FAPS-mode between 2000-2005.
- CERES FM1 (Terra) and FM3 (Aqua) for 2007 in cross-track.

Both Terra and Aqua data was selected ensuring that the CERES instruments operated in cross-track scan mode most of the time. Cross-track scanning is optimal for developing the BBR angular models since it provides the best matching with MODIS. However, models trained using only cross-track measurements were not able to faithfully reproduce the angular geometry derived from along-track measurements. Therefore, observations from along-track scanning were also included in the database.

The MODIS bands available in the SSF Ed4 product include channels at 0.47, 0.64, 0.86, 11.03, and 12.02 microns for daylight CERES measurements. The MSI channels at 1.65 and 2.21 microns, which are important for cloud characterization, are present as additional footprint MODIS radiance statistics in the SSF Ed4 product. However, the PSF weighting is only done at clear and full CERES FOV, while the information used in the BBR ADMs includes imager statistics for both clear and cloudy areas over the BBR pixel.

The authors also analyzed data from the CERES SSF Ed3A Aqua and Terra products for the ADM development, which included the following bands:

Spectral channels	MSI	Aqua MODIS SSF Ed3A	Terra MODIS SSF Ed3A
VIS	0.67	0.65	0.65
NIR	0.86	-	-
SWIR	1.65	-	1.63
SWIR	2.21	2.11	-
MIR	-	3.79	3.79
TIR	8.80	-	-
TIR	10.80	11.03	11.03
TIR	12.00	12.02	12.02

All available MODIS radiances within the MSI bands were checked, and the results determined that the best combination of MSI-available radiances were the 0.65, 2.21, and 11.03 µm MODIS bands. The use of the MODIS band at 1.63 µm instead of the 2.11 µm resulted in similar ADM performance.

The 12.02 µm band was found relevant only if SWIR channels were not used. The 3.79 µm imager channel is essential for determining cloud phase. Introducing this band as input significantly improves performance, but this channel is not, unfortunately, available in the MSI design.

Paragraph rewritten as follows: The inputs for creating training sets for cloud scenes include cloud cover and radiances from the MODIS bands closest to those of MSI. The underlying assumption is that the non-linear combination of narrow-band radiances provides adequate information about the anisotropy of cloudy scenes, eliminating the need for using imager-retrieved cloud properties. Imager radiances are analyzed separately over cloud-free and cloudy parts of the observed scene. The optimal combination of narrow-band radiances includes the 0.67, 0.865, 10.8, and 12.0 µm MSI bands. This selection was primarily influenced by the availability of bands with similar central wavelengths in the MODIS PSF-weighted radiances provided in the CERES SSF products and by their importance in retrieving cloud properties using the MSI (Hunerbein et al. 2023a, Hunerbein et al. 2023b). While the short-wave infrared MSI bands significantly impact ADM performance, the CERES Ed4 product's PSF-weighted imager statistics for these MODIS bands did not include statistics for clear and cloudy areas over the CERES footprint, which are essential for constructing the BBR ADMs.

The improvement of the ADMs by using the MSI SWIR bands is foreseen and proposed for future updates. A discussion on this has been included in the “Conclusions” section.

L153-159: If I understand correctly, different datasets are used to determine the surface type for the scene identification (GLCC and X-MET) than what is used in the CERES training data (IGBP and NSIDC). Why not use the same datasets to minimize errors associated with different dataset definitions?

X-MET is the EarthCARE operational product for weather forecast data, therefore the BMA-FLX processor should use it. In the surface classification that misalignment does not really exist because the GLCC includes now the IGBP Land Cover Classification used by CERES. It is a change in the dataset name.

L232: Some more information about the ADM uncertainty is needed. Where does this uncertainty estimate come from?

Thank you for your comment. Agreed. The uncertainty estimate associated to the BBR ADMs has been now introduced in the text to better explain the merge of the flux estimates:

"The evaluation of the BBR ADMs was performed using CERES flux estimates from the database described in Section 2.1.3. BBR ADMs simulate the CERES radiative flux retrievals, which are considered the "truth" in this analysis. Notably, this dataset for evaluation is an entirely independent source, as it was not used during the training of the ADM. Following the strategy used in the training, the input parameters (NB radiances, BB radiances, cloud fraction, and surface parameters) are obtained from the CERES SSF and auxiliary products. For each ADM scene class, a flux estimate was obtained for every CERES radiance measurement, and the RMSE was computed for the entire bin. The RMSE for each scene class represents the performance of the BMA-FLX SW model compared to CERES, indicating the minimum uncertainty associated with the ADMs."

Fig. 3: The cloud optical depth would be easier to visualise with a much more limited colour bar, or maybe a non-linear colour bar. Otherwise, the vast majority of the data points are in a very limited range and it is difficult to see the variability.

Agreed, changed colour bar to logarithmic and discrete.

Figure 4: It is not very useful to plot the land cover codes and refer the reader elsewhere for the definition of the codes. Can the codes be replaced by the definitions directly in the plot? Or put in a nearby table?

Agreed, figure legend modified and added "To enhance the presentation of the results, the original land cover classes have been simplified" in the caption.

Figure 4 and 5: They are referenced in reverse order in the text. I suggest switching the order.

Corrected. Figure 4 referenced before Figure 5.

L392-393: Are the significantly worse results in Figure 6 concerning? I believe this is the most relevant uncertainty estimate because when the algorithm is applied to actual EarthCARE observations, BMA-FLX will need to use the M-CLD retrieved cloud properties, correct?

The BMA-FLX processor will indeed ingest results retrieved by the M-CLD processor. But please note that the real validation of the BMA-FLX products will take place during the Commissioning Phase, where the actual impact of MSI and ATLID products' uncertainties will be fully assessed.

L423, Table 1, and Table 2: There are several instances for both the assessment domain and the standard resolution where the combined RMSE in the SW exceeds 10 W/m2. If the RMSE is indeed the relevant quantity (see earlier question), I interpret this to mean that the radiative closure goal will regularly be exceeded, which then leads me to question whether BMA-FLX is sufficient to achieve the closure goal of the mission. I expect that this is not the impression that the authors would like to leave the reader. I think the conclusions need some additional discussion to relate these results back to the mission requirements and what this means for the adequacy of the BMA-FLX product.

Your observation indeed points to the potential challenge in meeting the ambitious radiative closure goal of EarthCARE. However, it's important to remember that these results are part of a simulating environment, and validation with actual BBR measurements will take place in a latter stage.

The current error metrics exceeding the threshold in several instances might indicate that the radiative closure goal will not always be met. This underscore the complexity of achieving the mission's goals, and emphasizes the accuracy requirement's ambition. However, the BMA-FLX product is a significant step towards achieving them and still provides significant scientific value. The mission's goals extend beyond achieving a specific RMSE threshold to include broader scientific objectives, such as understanding radiative processes. The data obtained, even with higher uncertainties in certain scenarios, contributes to these overarching goals by providing insights into the Earth's radiation budget. The ongoing improvements to the algorithms and the comprehensive data generated in the mission will contribute to better understanding how to achieve the mission's objectives and the potential for further refinement and optimization of the BMA-FLX product.

This discussion has been added to the conclusion as follows:

In the end-to-end uncertainty assessment of the standard resolution product, the LW fluxes demonstrated strong alignment with the model truth fluxes across all three scenes, benefiting from reduced anisotropy in the oblique BBR views. The RMSEs for these scenes were consistently below 6 Wm-2. In contrast, the SW fluxes showed greater deviations from the model truth, primarily due to the more complex anisotropy of solar radiation footprints and their dependence on cloud-retrieved fields. The RMSEs for the combined fluxes varied from 7 Wm-2 in the Halifax scene to 18 Wm-2 in the Baja scene. Instances where the error metrics exceeded the 10 Wm-2 threshold suggest that achieving the radiative closure goal might be challenging, highlighting the complexity of meeting the mission's objectives and underscoring the ambitious accuracy requirements.

Despite these challenges, the BMA-FLX product represents a significant advancement toward achieving the mission's goals and provides considerable scientific value. The mission's objectives extend beyond meeting specific RMSE thresholds to encompass broader scientific aims, such as understanding radiative processes. Even with higher uncertainties in certain scenarios, the data collected contribute valuable insights into the Earth's radiation budget. Ongoing improvements to algorithms and comprehensive data generated by the mission will enhance our understanding of how to meet the mission's objectives and offer potential for further refinement and optimization of the BMA-FLX product.

It is important to note that these results are based on a simulation environment, and the proper validation of the BMA-FLX products will occur during the Commissioning Phase. This phase will involve evaluating input ingestion with actual retrievals from X-MET, BBR, MSI, and ATLID data, and will include thorough testing of flux retrieval accuracy, the parallax algorithm, and the reference level methodology.

For future improvements in the ADMs, we plan to test the use of AOD and albedo climatologies in clear-sky conditions. The inclusion of AOD climatology might introduce significant uncertainty due to the large spatio-temporal variability of aerosols, potentially causing discrepancies between the actual AOD for any given BBR measurement and the climatology, thus affecting anisotropy. We will evaluate the operational use of the EarthCARE MSI's AOD product (M-AOT, Docter et al. 2023). Additionally, the SWIR imager channels (1.65 and 2.21 mic), crucial for determining cloud parameters, are not currently utilized in the SW ADM. Future updates will include these bands to enhance cloud field characterization. Furthermore, alternative algorithms that leverage the multi-angular capabilities of the BBR will be explored in future iterations of the BMA-FLX processor to complement the current method of integrating fluxes from each BBR telescope.