Response to Referee #1

Thank you very much for your time and effort taken to review our manuscript submitted to AMT. We really appreciate the reviewer's constructive comments that are very useful to greatly improve the manuscript. We have revised the manuscript based on your comments as explained below. Please see below for our point-by-point responses to your comments, where the original comments are shown in *italics* and our responses are shown in normal text just below your corresponding comments.

General comments

1. *The evaluation of the algorithms using A-Train are useful but it would be helpful if the text could be restructured so that it is clearer what inputs will be used as input when EarthCARE data are available and what is being used from A-Train. For example, in Section 2.1 there is a mixture of EarthCARE products and A-Train data which is a bit challenging to parse.*

A. To clarify the distinction between the inputs used from A-Train data and those that will be used from EarthCARE products, we have added the following text at the beginning of Section 2.2:

'In the analysis of this paper, we utilize data from the A-Train constellation at the time of writing this paper before the EarthCARE data becomes available. While EarthCARE products will be used for future operational applications, A-Train data, including observations from CloudSat, CALIPSO, and MODIS, are currently employed to evaluate and refine the algorithm in preparation for application to the EarthCARE data. The A-Train data provides a valuable proxy for the type of information that will be available from EarthCARE, although there are some differences in instrument characteristics and data resolution. These differences are taken into account in our analysis to ensure that the results are relevant for the upcoming EarthCARE mission. For the EarthCARE mission, the algorithm will utilize data from the CPR_CLP (from CPR), the ATL_CLA (from ATLID), and the MSI_CLP (from MSI). These instruments provide vertical profiles of clouds and aerosols, which are critical inputs for calculating radiative fluxes and heating rates. The A-Train data, on the other hand, allows us to test and validate the algorithm using observations that are similar in nature to those expected from EarthCARE, ensuring that the algorithm is robust and ready for operational use once EarthCARE data becomes available.'

2. *In the manuscript evaluation of the RT products are compared with different datasets. It is easier to keep track of things if the products are introduced then consistently referred to after that point.*
3. *The methodology used for the TOA flux evaluation, Section 3.2, is difficult to understand in detail, especially the analysis of fluxes by cloud phase. Detailed comments have been provided below.*

A. The response to the general comments 2 and 3 will be addressed through the replies provided in each of the detailed comments below. We have added explanations, included citation of additional papers, and revised sentences to make the text clearer in accordance with the points raised in the comments.

Detailed comments

Line 47-48: This sentence should be rewritten to make it clearer that you are discussing space-based estimates of in atmosphere and surface radiative fluxes. The current text is a bit confusing, at least to me, since I wondered why RT calculations were used for surface radiation fluxes instead of surface-based radiometers.

A. The following statement has been added to clarify the explanation.

'Space-based RT calculations are commonly used to estimate radiative fluxes within the atmosphere and at the surface, complementing surface-based measurements where direct observations may be limited or unavailable on the global scale.'

Line 75: References for EarthCARE instruments or products?

A. We have cited Illingworth et al. (2015) and Wehr et al. (2023) as references for EarthCARE instruments and products in the revised manuscript.

Line 86: The +/- 10 W/m^2 has a particular spatial scale (averaged over 100 km^2) and temporal scale (instantaneous) which should be noted or referenced here.

A. References (ESA, 2001) and the following explanations have been added.

"The uncertainty of ±10 Wm^{-2} is associated with a spatial scale averaged over 100 km² and is based on instantaneous BBR measurements."

Line 126: Is this total cloud water content (liquid + ice) or liquid water content? If it is the former, how is it parsed into liquid and ice?

A. This intends to mean 'liquid water content'. We have revised the text as 'the effective particle radius and Cloud Water Content (CWC) for liquid clouds.

Line 140: Why is a constant sea surface albedo used?

A. A constant sea surface albedo is used in this study for simplicity and to reduce computational complexity. This approach assumes that the variability in sea surface albedo has a minimal impact on uncertainty of the overall radiative flux calculations, especially in comparison to other factors such as cloud cover and aerosol concentration. We have added a brief explanation of this setting in the revised manuscript as follows:

'In this study, a constant sea surface albedo is used to simplify the radiative transfer calculations and to minimize computational demands. This assumption is based on the understanding that variations in sea surface albedo have a relatively minor effect on uncertainty of the overall radiative flux compared to other variables such as cloud cover and aerosol properties.'

Line 142: Significantly more detail is required about why the data was averaged, how the averaging was performed and the effect of the averaging on the resulting radiative transfer. Examples of some questions that should be addressed include:

1. *Why is the data not on the target 1 km along track grid for EarthCARE?*

   A. The 1 km grid calculation has not been implemented in order to achieve operational computational speed. Furthermore, the current grid resolution has been chosen to align with the footprint of instruments like BBR and CERES, which is around 10 km to 20 km, respectively. However, we are considering to perform the calculations on a 1 km grid as part of future research products. We have added these explanations in the revised manuscript as: "This averaging is primarily due to the computational cost of radiative transfer for meeting the latency requirement of data processing and is also for consistency with the footprint of BBR and CERES, which is around 10km and 20km, respectively."

2. *What is the original resolution of the individual datasets (cloud, aerosol, surface and meteorological fields)?*

A. The original resolution of all the individual datasets, including cloud, aerosol, surface, and meteorological fields, is 1 km × 240 m. MODIS global albedo product (MCD43C3) is gridded at a 0.05° by 0.05° spatial resolution. This information was added in the revised manuscript.

3. *If I assume that the retrieved cloud profiles are meant to represent ~ 1 km footprint (line 180), how were the cloud properties averaged in the horizontal? Are the resulting cloud profiles on the 5 km grid assumed to be overcast and horizontally homogeneous or instead partial cloud (cloud fraction < 100%) and inhomogeneous?*

A. If even a single grid within the 5 km grid contains clouds, the cloud profile for the entire 5 km grid is treated as uniformly cloudy, with values averaged horizontally. The original product is designed with a 1 km footprint resolution, but the 5 km grid assumes horizontal uniformity of cloud distribution within the grid, and values are averaged accordingly to account for any inhomogeneity. This explanation was added in the revised manuscript.

4. *How was the data averaged in the vertical?*

A. The vertical resolution of the radiation transfer model is 1 km from the Earth surface to 30 km altitude. The data was averaged onto the 1km resolution.

Line 151: Is the Voronoi ice particle shape consistent with the EarthCARE retrievals?

A. The particle shape of the Voronoi ice is consistent with assumptions in MSI cloud retrievals of EarthCARE. We have added the reference (Wang et al. 2023) to this and revised the text as follows.

'As an assumption of the ice cloud optical properties, Voronoi particles were used to account for the non-spherical shape of the ice particles in both the JAXA/A-Train product and the

EarthCARE mission (Wang et al., 2023). This assumption of ice particles in the RT simulation was consistent with that of the MODIS and MSI ice cloud retrievals.'

Line 154: The CERES product and its version that was used for evaluation should be specified.

A. We have added the specific product name, CER_ES8_Aqua-FM3_Edition3, to the manuscript to clarify the source of the CERES data used in our study.

Line 157: The method used to compute the diurnal fluxes should be explained. For example, is there a consistent method used for the CERES and the 2B-FLXHR-Lidar algorithms. Is the data in the product diurnal fluxes or instantaneous? Are diurnal fluxes computed for comparison with BSRN data? How was that done with the calculations and with the BSRN data?

A. All comparisons, including those with other products and BSRN data, were conducted using instantaneous data. Diurnal fluxes were not computed or used in this study. This comparison is consistent with the CERES and 2B-FLXHR-Lidar algorithms. The use of instantaneous values for comparisons is noted in the revised text as follows.

'All comparisons, including those with other products and BSRN data, were conducted using instantaneous data.'

Line 160: Is the analysis split to periods when MODIS was and was not available? This affects the availability of the COT constraint on the cloud properties.

A. All of the data were analyzed for the period over which MODIS was available. The data sampling for comparisons of the all-sky conditions includes the case with MODIS data not available, but the analysis for cloud phase type classification is based on the cases with MODIS data available. We hope the original text describing this now makes sense for you.

Line 165: When averaging the RT results, they are an average of 20 km along orbit? I assume the CERES footprint is not just along orbit but roughly a 20x20 km footprint.

A. The CERES flux data is 20 km x 20 km including both along-track and cross-track directions; however, the 1D radiative transfer calculation compares the flux calculated only in the along-track direction, so the comparison with CERES requires consideration of this point. This statement has been added to the text as follows in the revised manuscript.

'The CERES flux data is 20 km x 20 km including both along-track and cross-track directions; however, the 1D radiative transfer calculation compares the flux calculated only in the along-track direction, so the comparison with CERES requires consideration of this point.'

Line 176: Please indicate the value of heat content of air at constant pressure used in the calculation.

A. The specific heat content of air at constant pressure used in the calculation is cp=1005 J kg^{-1} K^{-1}. We have added this number in the revised manuscript.

Line 190: It would be clearer to refer to products used for comparison after they have been introduced earlier in the text. It is not clear what data is "the NASA CloudSat CALIPSO team".

A. We have specified the data product of 2B-FLXHR-Lidar for clarification and added the URL of the NASA team's website.

Line 195: Maybe more precise to call it "cloud top phase of MODIS"?

A. Corrected to "cloud top phase of MODIS".

Line 202: What is the latitude resolution of the data shown in Figure 1 b-e? Is it 5 km?

A. Yes, it is 5km. We have added this information in the revised manuscript as follows.

'The latitudinal resolution in panels (b) to (e) of Figure 1 is shown at 5 km.'

Line 217: The 24.4 W/m^2 bias is significantly larger than 2B-FLXHR-Lidar.

A. The reason for the large bias is due to the positive bias in the case of ice-phase cloud. The positive bias due to ice-phase clouds is discussed in the section where comparisons are performed for different cloud phases separately (Section 4.1).

Line 225: It would be good to indicate here the fraction of the full set of RT calculation that are used for the cloud type analysis. The text in this paragraph suggests that only data for which CloudSat, CALIPSO and MODIS are available will be used. As noted in 160, when the MODIS COT is not available that constraint is removed from the cloud properties used for the RT calculations.

A. Although the percentage of occurrence is different depending on cloud type, comparisons with CERES are made on a 5° monthly average and are therefore presented as a sample size of N.

Line 229-235: The categories are confusing, at least to me, and I suggest some restructuring and rewriting of the text to try and clarify them. Summarizing my understanding of the current text, cloud phase based on CloudSat/CALIPSO data is "Water" when all layers are liquid phase, "Ice" when all layers are ice phase and "Mixed" when both are present. However, only for single layer clouds is the combined CloudSat/CALIPSO and MODIS cloud phase categories defined. This results in the categories "Water/Water", "Water/Ice", "Ice/Water". Are the cloud phase categories unique? It is also not quite clear what is a single layer for the analysis. Is it a single Cloudsat/CALIPSO layer or it can be multiple adjacent layers?

A. The reviewer's understanding is correct for the cloud phase categories based on CloudSat/CALIPSO (CC), which generates "water", "ice" and "mixed". For the single-layer clouds, these CC-based cloud phase categories are further combined with MODIS-based cloud phase categories of “water” and “ice” to result in combined categories of “water/water”, “water/ice” , “ice/water” (in the order of CC/MODIS) and “mixed”, as described in the text. These four phase categories are determined uniquely for a given single-layer cloud. Additionally, because it is challenging for MODIS to capture multi-layer clouds, our analysis with the CC-MODIS combined cloud phase information focuses on single-layer clouds. The single-layer clouds are derived from CloudSat/CALIPSO, indicating cases where only one vertically continuous cloud layer was detected. To clarify this point, the following sentence has been added to the revised manuscript.

‘The single-layer clouds are derived from CloudSat/CALIPSO, indicating cases where only one vertically continuous cloud layer was detected.’

Summing the “N” values in the Figure 3a, 3b and 3c, does not result in a total “N” that matches “N” shown in Figure 2a so it is not clear if the categories are unique.

A. Figures 3(a)-(e) are derived from the classification and analysis of Figure 2(a), but since each cloud type is compared with CERES on a 5° monthly average basis, the sample sizes do not match exactly.

Line 234: No need to restate MODIS(MOD) since it is done in line 229.

A. MODIS (MOD) was corrected to MOD.

Line 237: How is the “Mixed” category a single layer cloud when it is defined as “a mixture of ice and water within the vertical profile”? This goes back to the comment about definition of a single cloud layer.

A. "Mixed" indicates cases where both liquid water and ice were detected within the vertical structure obtained from CloudSat/CALIPSO.

Line 239: It is stated that Figure 3 is the same as Figure 2 but broken down by cloud phase. This can be taken to mean that the data used to construct Figure 3 are derived from data averaged over 5 degrees and 1 month. If this assumption is correct, then it is unclear how to interpret the statement that the comparisons was limited to points when cloud in the CERES footprint were of the same type since that occurs on ~20 km and instantaneous data. When accumulated over space and time wouldn't there be heterogeneity arising from the CERES footprint level data, even if it was the same cloud type?

A. When classifying cloud types, we use 1 km grid data and analyze only cases where the entire approximately 20 km footprint of CERES along the track is covered by the same cloud type. It is true that the CERES footprint also has a 20 km observation width in the cross-track direction, meaning that other types of clouds could be mixed in. However, our approach does not include these cases for simplicity, and this is considered a limitation of the current analysis. The following explanatory text was added.

'When classifying cloud types, we use 1 km grid data and analyze only cases where the entire approximately 20 km footprint of CERES along the track is covered by the same cloud type.'

Line 243: Compared to what are the bias and RMSE are relatively small? While not necessary to include in the paper, it would be helpful to have the cloud phase analysis applied to the 2B-FLXHR-Lidar product to provide a point of comparison results using the EarthCARE algorithm.

A. The RMSE is smaller than that of ice-containing clouds. We have added the following text in the revised manuscript to clarify this point: 'When both CC and MOD indicate water clouds, the SW flux shows a slight negative bias, but both the bias (-11.7 Wm^{-2}) and RMSE (46.2 Wm^{-2}) are relatively small (Figure 5 (a)) compared to ice-containing clouds.' This paper focuses on the validation of the Japanese product, and therefore, classifying and analyzing the 2B-FLXHR-Lidar data by cloud type is beyond the scope of this validation. However, scientifically, it is very meaningful to validate the 2B-FLXHR-Lidar data by cloud type and to compare with our product, and we plan to do so in future EarthCARE validation studies.

Line 312: Could the biases also be compared with computed surface fluxes from CERES and 2B-FLXHR-Lidar? While not direct observations they would increase the amount of data that could be used for comparison with the EarthCARE RT algorithm.

A. As part of the EarthCARE validation plan, it has been decided to use observed flux data from BBR and BSRN, so we have conducted comparisons with these observations. In this paper, the validation was performed solely by comparing with observational data, in accordance with the validation plan. However, comparisons with surface fluxes from CERES and 2B-FLXHR-Lidar, suggested by the reviewer, would also be valuable to further validate our algorithm in future studies. Thank you very much for your suggestion.

Line 319: It would be good to explicitly document how the aerosol and cloud radiative forcing is computed in Section 2.1 since it is part of the product output.

A. The following description for computation of ARF and CRF have been added to section 3.

'Aerosol radiative forcing (ARF) and cloud radiative forcing (CRF) are calculated as the difference between the radiative fluxes with and without aerosols or clouds, respectively. Specifically, ARF is defined as the difference between the radiative flux calculated with all aerosol components included and the flux calculated without aerosols. Similarly, CRF is defined as the difference between the radiative flux with all cloud components included and the flux calculated in the absence of clouds. These calculations are performed for both the TOA and the SFC to assess the impact of aerosols and clouds on the Earth's energy budget.'

Figure 1: What is the wavelength for the extinction shown in panels "c"? Panel "e" is a bit hard to follow. Could it be split into a panel for SW and a panel for LW? For the current panel "e", the "obs" legend markers at the bottom of the plot are barely visible. It would also be good to have panel "e" aligned along the x-axis with the panels above it. Also, it is quite challenging to compare the markers for the computed and observed fluxes since they are fluctuating significantly, perhaps a line plot would be better.

A. The wavelength in panel (c) is 532nm, and it has been added to the figure. We have divided Figure 1 (e) into two separate panels: one for SW (panel (e)) and one for LW (panel (f)). This division makes it easier to see the value fluctuations and markers. Thank you for your suggestion.

Figure 8: It is difficult to see any structure to the cloud forcing on the plots. It would be helpful to consider modifying the plots so that some of the structure can be seen.

A . Widening the color bar range would make it difficult to capture the subtle effects in the LW radiation. Since the primary goal of this study is to demonstrate the overall cooling effect in both SW and Net radiation at the TOA and SFC, we would like to retain the current color bar settings. Also, the spatial pattern of heating and cooling in ATM for LW is clearly visible within the current color bar range.