I thank the authors for their efforts to respond to the original review and clarify my questions which was done appropriately. What I list below are changes of a technical nature mostly in respect of the additional text and changes. Line numbers refer to the revised version line numbers.

Line 72 Changes to this section are fine in general and address the original confusion but for complete correctness I suggest amending 'The thermal flux has the same dependencies as the solar flux except for θ_0,' to 'The thermal flux has the same form but without the dependence on the solar enith (θ_0) and with ϕ denoting the view azimuth rather than relative azimuth.'

The definition given in lines 114-117 for scene definition states angular geometry of the ADM is included in the scene definition and scene class is the part of the scene definition which determines the number of ADMs. This is fine but the rest of the discussion needs to be consistent with this definition but I think that scene definition is used several times in this section when you mean scene class. E.G. line 118 'scene definition' should be changed to 'scene classes' (with 'This classiiction' on line 121 changes to 'These classes' for consistency. Similarly table 1 should actually be titled Scene Classes rather than scene definition. And this should be reflected when referencing is (line 127).

Line 189 to 191.Change 'In clear-sky scenarios, the retrieval algorithm employs the two surface types with highest coverage within the BBR pixel, defining the observation as a mixed scene' to 'For clear sky cases which are a mix of surface types the observation is defined as mixed scene and the retrieval algorithm employs the two surface types with highest coverage within the BBR pixel'. Otherwise it sounds like you are defining all clear sky as mixed which presumably some might be a single surface type.

Lines 235 to 236. Given we have not yet discussed the LW this assertion, without any context or reference that the CTH is reliable for the thermal seems a bit odd. I would suggest removing 'The 90th percentile of the CTH derived from the MSI brightness temperature (BT), referred to as M-COP (Hünerbein et al., 2023a), is a reliable estimator for co-registering the BBR radiances in the thermal regime.' And just start the paragraph with the next sentence.

Lines 266 to 267. θ_{obl} needs to be defined. I think this needs to be rewritten to make the paragraph self-consistent, with some explanation as to why di starts at 1 and not for example zero. You are essentially saying that the minimum cloud height is 1/ tan (θ_{obl}), (so 0.7km assuming 55 is the angle), however you just stated that the cloud could be anywhere from surface (which I assume is zero) to the tropopause so the assertion that d starts at 1 is at odds with this. I think maybe you are saying that an image is only considered to be affected by parallax when there is some minimum difference between observed cloud and surface or nadir cloud height and oblique cloud height but this is not clear and in any case the latter would make d range from nadir CTH + something not 1. I think some correction is required here.

Lines 273 to 274. I think that selection of the excluded data reserved for testing as described in your reply to my question on this is fine. However I think calling it 'an entirely independent source' is incorrect and confusing making it sound like an entirely different dataset from CERES. I suggest replaincg 'Notably, this dataset for evaluation is an entirely independent source, as it was not used during the training of the ADM' with 'All the data in the evaluation dataset was excluded from the original training process'. Or 'The evaluation dataset represents a randomly selected subset of the originally identified potential training data that was excluded from the

training process'. I think it would be helpful to add one more sentence stating the RMS and bias errors found with this validation test.

Lines 292 to 294 Change: 'Even though in theory a multi-spectral model should be able to correctly handle all scene types, in practice it was demonstrated from previous GERB studies (Dewitte et al., 2008; Clerbaux et al., 2009) that a large bias was introduced in the case of semi-transparent clouds, fact that was indeed verified in the early stages of the selection and validation of the algorithm'. 'Previous studies for GERB (Dewitte et al., 2008; Clerbaux et al., 2009) has shown that using a single multi-spectral regression for all scenes can cause large biases for semi-transparent cloud. This problem was also highlighted in the early stages of the selection on and validation of the FMA-FLX processor algorithms'

Line 326 to 326 'To capitalize on this, the LW merging algorithm assigns a greater weight to the fore and aft views for plane-parallel scenes.' But you don't seem to identify plane parallel scenes and later you don't assign greater weight to fore and aft views so I think you mean "Thus plane parallel assumptions would indicate that greater weight should be placed on the fore and after views.'

Lines 322 to 323, "The three BBR thermal unfiltered radiances are co-registered at a reference level defined by the percentile 90th of the MSI CTH as described in section 2.1.5" Either specify for cloudy scenes or make general to all scenes by changing to "The three BBR thermal unfiltered radiances use the default surface co-registration for clear sky and for cloudy scenes are co-registered at a reference level defined by the percentile 90th of the MSI CTH as described in section 2.1.5" or "For cloudy scenes the three BBR thermal unfiltered radiances are co-registered at a reference level defined by the percentile 90th of the MSI CTH as described in section 2.1.5 instead of the default co-registration used for clear sky."

Line 362 'Proper validation...' to 'Full validation..'

Line 374 'allows for the identification...' to 'allows the identification..'

Lines 377 to 380. You state discrepancy 'between the radiative transfer simulated used in the simulated geophysical data and the modelled EarthCARE products' and later 'discrepancies in surface definitions between RTC model and the BMA_FLX processor', but don't address the possible (likely if plane parallel) discrepancy between the simulated EarthCARE flux 'truth' and real non real world fluxes until much later. I suggest adding a sentence on this here as it is a major limitation you have with what you need to work with and is very relevant for considering what your results mean.

Line 430to 431 "The combined approach shows the greatest advantage in terms of error metrics, with lower values compared to the view-based flux estimations" to "The combined approach results in better error metrics than the those for the induvial views'

Line 432 "...the combined approach has significant lower error metrics than the individual fluxes, indicating better overall agreement." To "...the combined approach results in significantly better error metrics than those of the individual fluxes, indicating better overall agreement."

Line 435 " indicating the superior overall performance" to ".indicating superior overall performance"

Line 449 to 450 ' This comparison is shown as a "pre-launch" numerical assessment experiment for validating the performance and reliability of the BMA-FLX processor in diverse environmental

conditions.' Add a sentence along the lines or "Within the limitations of the accuracy of the simulated fluxes provided as truth.'

Line 456 to 457 "Consequently, SW fluxes, which are obtained from ADMs constructed using satellite measurements, retrieved in the regions with simulated broken clouds tend to be noisy and less reliable." This sentence seems to imply that your empirical ADMs are wrong, noisy and unreliable, rather than just not representative of the unrealistic plane parallel situation of the simulated truth I thin you need to rephrase this to make sense for example "Consequently, fluxes from BMA-FLX, which are obtained from ADMs constructed using satellite measurements of the real world non-plane parallel anisotropy will not follow the radiance to flux relationships found in plane parallel simulations used as the truth in this study." This applies to both the SW and LW I don't see a need to restrict it to the SW.

Lines 457 to 458 "In the LW estimates, this results in noisy flux retrievals that are flattened out when increasing the averaging region." Whilst the LW estimate c.f. to the 'truth' do appear noisy and the differences do average out I don't know if there is evidence that this is plane parallel in origin or even a model/real world cause. I suggest just changing these sentence to something more vague such as "LW flux retrievals appear to have a high frequency variations not observed in the simulations that are flattened out with increasing avering region". The issue with your longwave fluxes occurs in some clear sky cases and some cloud. It just appears that your fluxes are considerably more sensitive to some input that is varying on these high frequency scales that the simulated fluxes are not seeing. Either due to changing between regressions or I suppose the use of your z2 as a channel difference but you would need to look at your inputs to see where that is coming from. In any case I suggest just changing the sentence to not imply a known cause as suggested.

Line 477 'significantly impact in' to 'significantly impact'

Line 478 "The combined results for Baja do not differ significantly from the previous analysis." This statement does not match with what is shown in the tables or the plots, they seem very significantly different to me. I think you are not considering changes made to the values shown in table 3 presented here from those shown in the original submission. Please correct this statement or the results as appropriate.

Line 479 to 480 "This is primarily due to ..." "The average values are also influenced by ." I think it is far from clear is this is primary cause but it clearly influences comparison of the average values.

Line 481 to 482 "Overall, the algorithm for combining the view-based fluxes performs exceptionally well in mitigating the impact of incorrect retrievals from the nadir, aft, and fore models." If the 'incorrect retrievals' here are meant to mean the cloud retrievals you are discussing then this isn't very convincing for both Halifax and Baja the improvement for combination here seems very similar to what was obtained when you combined them without cloud retrieval errors. You do get a very significant reduction for Hawaii but given the similar strange inconsistency between views you presented for Baja in the original manuscript that seems to have gone away in this revision I wonder if maybe you should double check these Hawaii results and see if they have a a similar problem. I think this sentence needs to be rephrased to talk generally about the combination improvement rather than specifically for the effect of cloud retrieval errors although obviously helpful for those too.

518 to 520: "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" I would rephrase this and possibly move your discussion in lines 527 to 530 about the simulated environment here to add context. For example change to "Instances where the error metrics exceeded the 10 Wm−2 threshold may indicate that achieving the radiative goal will be challenging requiring both improvements to the cloud property retrieval and the BMA-FLX algorithm. However, this needs to be considered in the context of the simulated environment used for this study and the likely inaccuracy of the simulated fluxes used for 'truth' here particularly for broken cloud conditions. Further validation of the BMA-FLX will occur during the commissioning ......"

Line 638 – DOI is for preprint needs be updated to published version.