Summary Comments

The manuscript provides a detailed description of a series of airborne campaigns deploying a suite of instruments assembled to assess and validate all instruments on the EarthCARE satellite. The paper provides a concise overview of the EarthCARE objectives and clearly articulates the validation measurement needs of all four instruments on the EarthCARE satellite over a range of atmospheric conditions highlighting the uniqueness of the instrument suite deployed. The introduction was well written, and provided a clear description, application, and impact of the using airborne instrument suite for validation. In contrast, the abstract should be changed to be clear on the overall objective or purpose of this manuscript and provide a summary statement of the conclusions. For example, the manuscript describes the rationale, comparison approach for each instrument, and shows initial comparisons thus providing a substantial contribution to the evaluation and validation of EarthCARE's measurements. These comparisons start to build confidence in the EarthCARE measurements and data products for use by the science community and are already providing feedback to the algorithm developers to adjust processing methods and identifying additional calibration needs.

The aircraft remote sensing instruments deployed are clearly described in the manuscript with appropriate references used for calibration, demonstration of their heritage, and are of high quality. The measurements provide a capable and unique suite of measurements that was deployed in multiple geographic regions with varying atmospheric conditions of clouds and aerosols and spanned both daytime and nighttime as presented in the examples for each instrument. The paper clearly states, although not in the abstract, that the intent is not to provide an exhaustive validation analsysis but rather illustrate how the data has been and will continue to support assessments of EarthCARE's measurements and retrieval algorithms for key geophysical variables.

There are two appendices that came across as after thoughts in the manuscript. The first appendix provides specific flights coordinated underflights of NASA's PACE satellite during the deployments discussed in the manuscript (see specific comments on the flight track figures below) and is only discussed in the conclusions section at the very end. The second appendix is a summary of the previous coordinated flights with NASA's CALISPO and CloudSat satellites and similarly has only a few sentences at the end of the conclusions. In this case, the authors postulate that airborne flight campaigns can link the two separate datasets without providing specific details on how that might be done which is likely not trivial. It is recommended to remove this from the conclusions and compile all the tables (PERCUSSION, PACE, CALIPSO/CloudSat) of the under-flight information in a single supplement rather than include these as appendices. These tables can then be used as a reference for future more comprehensive validation efforts.

Lastly, these data will only be useful the larger science community if the datasets are available and open access. A location and description of how to download the CALIPSO/CloudSat flights was not clearly listed in the appendix or in the text. In addition, the data should be processed data (i.e. geophysical data products rather than raw signals) if they are to be fully exploited for validation by the larger community. Overall, the presentation quality is generally good. The manuscript has value to the larger community to better under this unique dataset with specific examples presented to address all four EarthCARE instruments. Specific detailed comments below can enhance clarity, particularly in sections discussing lidar measurements.

We thank you very much for your time you dedicated to review this manuscript, for your careful reading and valuable comments that helped us improve the quality of this

manuscript. Based on the reviews we revised the manuscript more substantially to better bring out the main points of this study.

We have considered all your comments and will give a point-by-point response. Answers are given in italics and bold. Text changes are marked in italics, bold, and blue.

Specific Comments/Suggestions:

- Provide clear statement of the overall objective of the manuscript and summarize conclusions in the abstract. It is not clear from the abstract what is the main objective paper and there is no conclusions/im pact statement provided.

We agree that the abstract did not well capture the overall objective of the manuscript. We revised the abstract:

'In May 2024, the Earth Clouds, Aerosols and Radiation Explorer (EarthCARE) satellite was launched. For the first time a satellite combines two active instruments, i.e., the Atmospheric Lidar and the Cloud Profiling Radar, together with two passive instruments, a multi-spectral imager and a broad-band radiometer, on one single spacecraft platform. EarthCARE is thus the most complex satellite mission to date to for collocated aerosol, cloud, radiation and precipitation measurements. To utilize the data collected by the EarthCARE mission to its full extent and to support and quantify the data quality and measurement uncertainty, careful and holistic validation activities are needed. For this purpose, we set up an airborne instrument payload on the German High Altitude and LOng-range research aircraft (HALO), which is similar to the EarthCARE instrumentation. We used this payload during an extensive measurement campaign in summer and fall 2024 in the tropic and mid- to high-latitudes targeting to validate the EarthCARE measurements and data products early in its commissioning phase. With this manuscript we aim to give a detailed overview of the PERCUSION mission, and to advertise the use of its data in future more detailed validation studies. We give examples of how to use PERCUSION data to approach the validation of all for instruments of EarthCARE as well as of higher level (i.e. multi-sensor) products, and give first confidence in the quality of EarthCARE data.'

- Recommend moving Table 5 to a supplemental section along with other flight information tables for the PACE and CALIPSO/CloudSat.

We agree that the information in the appendix appears somewhat unintroduced and unmotivated (with only a brief mention in the Conclusions). We include now the information of underflights under PACE mission during the second part of the PERCUSION campaign, and the information to CALIPSO/ClouSat already in the PERCUSION section to better introduce this addional objective.

'In addition to our primary goal of validating EarthCARE, our efforts also address the validation of measurements and algorithms from NASA's PACE (Plankton, Aerosol, Cloud, Ocean Ecosystem) mission. To this end, we conducted four dedicated flights in the PACE swath during the campaign phase over the tropical western Atlantic. Furthermore, in preparation for EarthCARE, we performed a series of underflights with the same payload under NASA's mission constellations CALIPSO and Cloudsat. Together with the measurements for EarthCARE validation, these underflights offer the opportunity to bridge some of the gaps arising from the two missions, such as those concerning wavelength

dependencies and sensitivities. Both topics will be addressed in the appendices of this manuscript.'

We do not want to put the information about the conducted underflights into a supplement, since we believe the detailed information about the underflights is an important part of this manuscript.

Figure 2 and discussion therein:

a) Care should be taken with the terms backscatter, backscatter ratio, and backscatter coefficient. For example, Rayleigh backscatter profile is used but not actually defined. Care should also be taken when comparing backscatter coefficient attenuated products with unattenuated products. For example, all of the Reyleight products are attenuated as shown.

   We agree that care should be taken here. Thus, we went through the text carefully. We introduced the term backscatter ratio where used the for the first time to clearly distinguish between backscatter coefficient and backscatter ratio. We also changed the phrasing from e.g. Rayleigh backscatter profile to Rayleigh backscatter coefficient to have no inconstancies between the figure and the description. However, we do not quantitatively compare the attenuated products to unattenuated products. The time-height cross-sections from WALES and ATLID are shown to give some context of the scene and confirm its stability.

b) Labels in upper panel have label of unpolarized – caption states this to be co-polarized and should be changed.

   We had a critical look again at figure label and figure caption and confirm, that the caption and the labels of the time-height cross-section is correct as given. The time-height cross-sections show the unpolarized (attenuated) backscatter coefficient for both instruments to give the context of the observed scene. In the profile comparisons we look at Mie, Rayleigh and cross-polar backscatter coefficient. Please be aware, that this is not the same variables as shown in the time-height cross-sections.

c) Why does the comparison not include the coincident location or be centered at the coincident location? Please explain/justify. Text refers to atmospheric conditions as being stable.

   Just around the overpass point the aerosol is very structured, which enhances the probability for representativeness errors. We extend the text to explain this better.

d) Why was 120 km used for the comparison? Please provide a rationale.

   The single profile data from ATLID is quite noisy, so a large averaging interval was chosen to reduce the statistical noise and make systematic effects stick out.

e) No error bars provided for this comparison. Provide careful analysis to include error estimates for both datasets.

The error bars for the statistical noise are so small that they were hardly visible. We added them now and you can see them if you zoom in.

f) No statistical comparison is provided between the two different profiles (e.g. mean differences etc.).

We agree, that doing a statistical comparison is very valuable. However, in this manuscript we only show one single example to advertise the use of PERCUSION for EarthCARE validation and give the main information. The statistical basis from one example is too low. We are preparing a manuscript targeting ATLID L1/L2 validation with the WALES measurements during PERCUSION where we include all 33 underpasses. We will add a statistical evaluation in this more detailed study.

g) The L1 ATLID data products are attenuated backscatter for the three components of scattering rather than unattenuated backscatter which is shown and therefore I suggest not stating this as a 'direct' L1 comparison. Showing and stating that a direct comparison of the attenuated ATLID product from the WALES data (unless a wavelength dependent attenuation correction is applied) is confusing. Although a cross reference (i.e. normalization to the ATLID data) is done for the WALES Rayleigh attenuated backscatter it is not clearly stated why this is required for comparison for the average reader. There are terms like 'extinction corrected' and 'aligned' which are not defined and discussed to the why they are used.

As already mentioned for comment b), the time-heigh cross-sections are intended to give the scene context (geometry, homogeneity). The actual comparison is done on the co-polar, Rayleigh and cross-polar backscatter coefficient. For these quantities, of course, the different attenuation by molecular scattering for the different wavelengths has to be corrected.
This is done using the air density profiles given within the products. 'Direct' comparison here means that averaged profiles close to the raw L1 data are compared. As we see this might lead mis-interpretation we removed the word 'direct'. Furthermore, we already mention in the text, that we have to correct for different attenuation due to the different wavelengths; thus the data has to be extinction corrected for the comparison. ATLID data (co-polarized backscatter coefficient) is not given extinction corrected (ATLID raw) we use the information in the Rayleigh signal to correct the attenuation to make the profiles comparable to WALES co-polar backscatter coefficient, which is fully corrected for extinction. I that might not have been clear, we changed the text as following:

‚The backscatter coefficient for co-polar polarization for WALES and for ATLID is shown without extinction correction as contained in the L1 product and with correction for aerosol and molecular extinction by dividing the co-polar signal by the Rayleigh signal (HSRL method, purple line). The WALES profile shown is fully corrected for extinction.'

Since it was also not clear why we had to align the profiles for the Rayleigh backscatter coefficient, we added the following information:

'For evaluating the ATLID Rayleigh backscatter coefficient, we take into account the expected pure Rayleigh backscatter signal calculated from an atmospheric density profile contained within the L1 data and this pure Rayleigh signal at 355 nm with additional aerosol extinction derived from the WALES measurements.'

h) Suggest rewriting the discussion and figure caption being more precise on the specific terms and provide clear and concise description of the processing done for the comparisons. Explain why you are comparing unattenuated backscatter coefficients for the Mie components and why you are comparing attenuated backscatter coefficients for the Rayleigh scattering along when a scaling or normalization for above aircraft attenuation is being applied or not being applied. It might also help to define the ATLID L1 products as attenuated backscatter from top of the atmosphere. It looks like these comparisons are using the ratios o L1 data profiles. Again please clearly explain in the text.

   We have carefully looked at the description of the processing and added some additional explanations.

i) Is the same molecular density profile used for the air density in WALES and ATLID data?

   The ALTID and the WALES data use the air density profiles provided within their products which both are derived from ECMWF IFS analyses. The difference which comes from different interpolation schemes is below 0.5% for the case shown here and does not lead to deviations visible in the plotted data. This will be addressed in the detailed study and publication on L1/L2 validation.

Figure 3 and discussion therein:

a) Why is a 100km averaged used? Please provide rationale.

   Please see answers to Figure 2.

b) What is the shading in the line plots (particle extinction and particle linear depolarization)

   This is the error estimate given within the data products. We added text to the caption to explain this.

c) Provide error bars for data

   Please see answer above.

d) No statistical comparison is made between the measurements is presented.

   Please see the answer for point f) in the comments to figure 2.

e) It looks like there is a ~30% difference of the lidar ratio (mostly due to differences in backscatter it appears) and no comment is provided on these differences. In particular, it is stated that the lidar ratio and extinction are expected to be the same. Is this due to measurement error or differences in the wavelengths for this case? Please comment.

We agree that we have been short in the description of this comparison, especially as large differences occur, e.g. for the lidar ratio. We thus extended this paragraph:

'Our comparisons confirm the good quality of the A-PRO products in the current baseline version in general. But differences between the WALES and ATLID data of the different products are visible. While the particle extinction coefficient shows a good agreement within the uncertainty range, significant differences occur for the values of the lidar ratio in the height range between about 1-4 km. The lidar ratio retrieved from ATLID data is about 60 sr while the lidar ratio derived from WALES measurements is about 45 sr. This difference might result from the differences in the derived backscatter coefficient. The discrepancies might result from a wrong target classification and thus a wrong a priori information in the optimal estimate retrieval to derive A-EBD products, or a too strong weight on the a priori information. Investigations about this issue are ongoing, and the problem is hopefully solved in the new baseline version. In addition, for the example shown, significant noise on the depolarization values is observed. Furthermore, especially the depolarization ratio shows a strange behavior at the upper edge of the aerosol layer leading to high, unphysical values.'

f) It looks like ATLID does not have backscatter coefficient measurements above 5 km. Is this correct?

The data within the ALTLID L2 files are flagged as bad (too low SNR) and only fill_values are provided, so yes there is no ATLID data above.

g) The only plot on the bottom panel that has data above ~5km is the first one and there is no ATLID data for comparison. Suggest expanding the scale to show the range of measurements better.

The plots were chosen to cover the range where WALES data is present. That there is no WALES data above 5 km for the more SNR sensitive products at least shows that the aerosol load in this altitude region is so low that the lack of ATLID data can be attributed to this fact.

Figure 4 and discussion therein:

a) Overall the discussion can be simplified which should help highlight the key points when describing layers. For example, "vertically extended layer" is not needed and "indicative of a marked" are extra words that make this hard to follow. Suggest spending time make this discussion starting at line 340 more concise for the readers.

Thank you for this comment. Following the reviewer's suggestion, we revised the paragraph starting at line 340 to simplify the wording and make the description of the observed layers more concise and easier to follow. The paragraph now reads as follows:

'Besides the L2 optical properties we use the WALES measurements for a comparison with the ATLID Target Classification (EarthCARE ATLID TC Level 2A) and the ATLID cloud top height (EarthCARE ATLID CTH Level 2A) (Wandinger et al., 2023) (baseline BA). An example

comparison is shown for the HALO underflight on the 16 August 2024 southwest of Cabo Verde (orbit 01240E). This flight was selected as it captures a wide range of cloud and aerosol conditions typical of the tropical atmosphere. The flight objective aimed at a south-north transect across the ITCZ, including a flight leg of about 700 km (50 minutes) flown along the predicted EarthCARE ground track, extending from the ITCZ centre toward its northern edge. Figure 4a shows a time-altitude cross section of the backscatter ratio (BSR) measured with the WALES instrument with the EarthCARE underpass at 16:14 UTC. The BSR measurements reveals several cloud and aerosol features that are captured in this flight section. At the highest altitudes (10-13 km) frequently large BSR values indicate a widespread field of ice clouds. This ice cloud field has the largest vertical extent near the underpass and is more patchy and thinner toward the northern part of the transect. At lower altitudes (1–3 km), smaller-scale regions of large BSR values are detected which are associated with stratocumulus and cumulus clouds. Additionally in the lowermost 6 km, an extended layer characterized by moderate BSR values (BSR between 2-6) is observed which indicates an aerosol layer. Notably, this aerosol layer appears thicker and more continuous in the northern segment (distance > 400 km) of the flight leg. In the southern part, it appears as two separated layers: an elevated layer around 5 km altitude and another layer within the lowest 2 km of the atmosphere'

b) So all the layers are identified by thresholds it seems from the text which is a common approach for lidar layer detection algorithms. Are these thresholds the same for both ATLID and WALES? If not, one should expect diLerences. Please provide a discussion on the thresholds used for this comparison.
c) Similarly, the averaging scales will also impact these comparisons.

Addressing b) and c)

Thank you for this comment. We would like to emphasize here that the thresholds, methodologies, and averaging strategies used for the generation of the A-TC product and our WALES dataset differ fundamentally. However, our approach is not to mimic the ATLID measurements or the processor algorithms but to provide a fully independent validation. We generally want to highlight in this paper the suitability of the WALES data set for ATLID validation.

In our study, we intentionally follow a fully independent approach by using independent datasets and developing an independent algorithm for cloud detection. Several previous studies have demonstrated that the backscatter ratio at 532 nm can reliably be used to identify and separate clouds and aerosols across different altitude/latitude ranges (Groß et al., 2014; Urbanek et al., 2017; Gutleben et al., 2019; Dekoutsidis et al., 2024).

In addition, the comprehensive A-TC algorithm (see Irbah et al., 2023) relies on individual thresholds for the co- and cross-polarized Mie channels as well as the Rayleigh channel. The algorithm also uses the tropopause height as input and applies different averaging strategies guided by the A-FM product (see Donavan et al., 2024; van Zadelhoff et al., 2023).

Concerning your comment: If, however, the same thresholds were applied, for example to the backscatter ratio, a larger number of cloud pixels would be detected in the WALES dataset. This is due to the wavelength dependence of the BSR (ATLID: 355 nm vs. WALES: 532 nm), which generally results in lower values for ATLID compared to WALES. This effect would lead to a systematic underestimation of cloud pixels by the A-TC products, but the opposite seems to be the case, so the chosen backscatter threshold does not seem to be the critical factor here. We see (Figure 4b) that the A-TC product increases the cloud area

especially of ice clouds, e.g., by filling in gaps (which are detected in WALES). However, these cloud gaps are visible in the A-FM product as well as in Level 1 products, which suggests that the increase in cloud size (as well as a systematic overestimation of cloud pixels) is caused by the processing rather than by an incorrect choice of thresholds.

Finally, we note that a dedicated study is currently in preparation (Krüger et al. 2026, in prep.), which will investigate both the sensitivity of cloud pixel detection to the applied thresholds and the impact of different spatial resolutions (i.e., averaging scales of the A-TC product).

d) I did not see a reference to the CTH lines for the colors (i.e. which one is for ATLID and WALES)

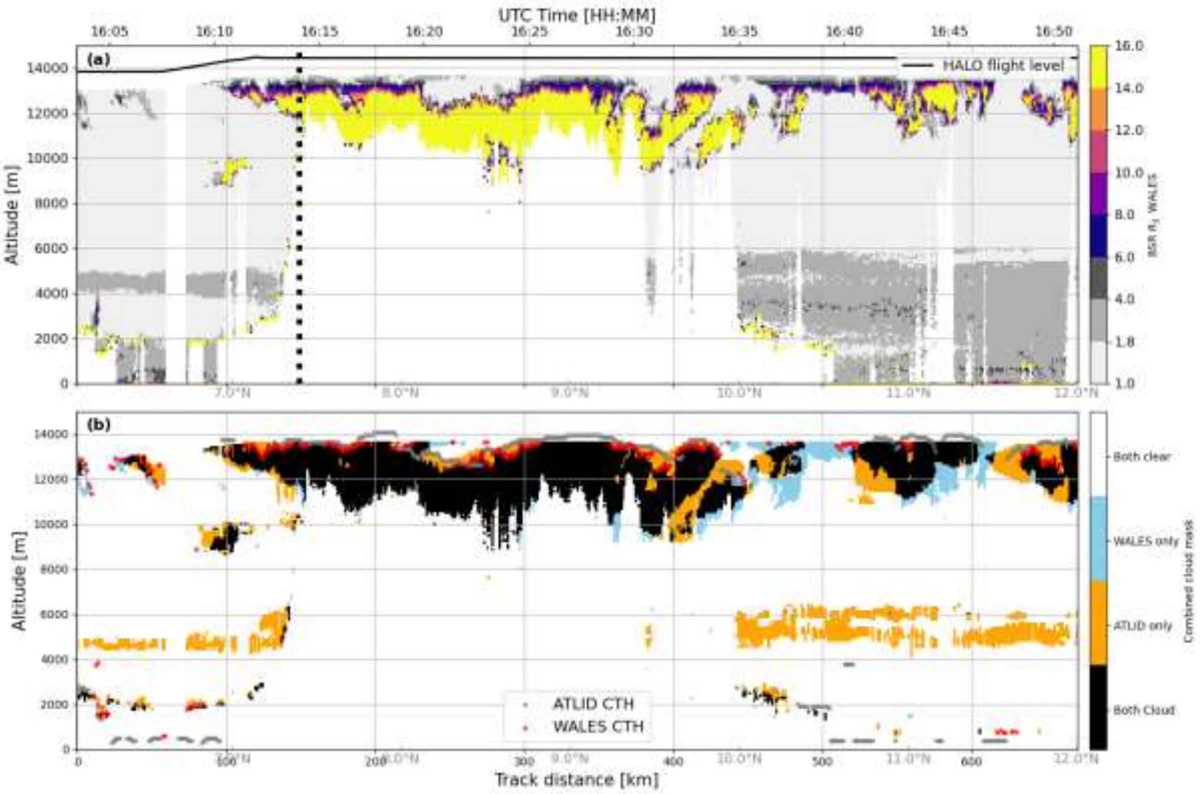
e) For WALES, does the cloud top height need to be below where the data starts? It seems like it is capped at the top of the highest altitude of the dataset (if WALES is the red color)

Comment regarding d) and e):

Thanks for the hints. We generally updated the figure and its caption to improve its clarity (also corresponding to requests from reviewer 1). The updates are:

- ***Adjusted the backscatter color bar by reducing the number of colors and setting new thresholds, corresponding to Krüger et al. (2026, in prep.).***
- ***Added a legend showing cloud tops from WALES (red) and the A-CTH product (gray).***
- ***Included a proper labelling of the lower x-axes in panel (a) and (b) (°N).***

We hope these changes also help the reader to better distinguish between cloud tops from WALES and A-CTH. To remove the "capped" cloud tops from WALES, we now exclusively display WALES cloud tops that could be unambiguously identified—i.e., where the distance between the cloud and the aircraft was sufficiently large—are displayed now.



"Figure 4: Distance–altitude cross sections for the HALO flight on 16 August 2024 showing (a) the backscatter ratio measured with the WALES HSRL, along with the flight altitude of the HALO aircraft (black solid line) and the EarthCARE underpass track (thick black dashed line). Panel (b) shows the combined cloud mask derived from the WALES BSR data and the collocated EarthCARE Level-2 target classification product (A-TC). Black shading indicates cloud pixels detected by both WALES and A-TC, while blue and orange colors represent cloud pixels identified exclusively by WALES or A-TC, respectively. White denotes regions with no cloud detection by either product. The red markers indicate the cloud top height derived from the WALES backscatter ratio, while the gray markers represent the cloud top height from the A-CTH product. Note that cloud tops from WALES are displayed only when they could be unambiguously identified (e.g., when the cloud top was sufficiently far from the HALO aircraft), while A-CTH reports a cloud top for these cases."

Please note the references that were mentioned above:

Dekoutsidis, G., Wirth, M., and Groß, S.: The effects of warm-air intrusions in the high Arctic on cirrus clouds, Atmos. Chem. Phys., 24, 5971–5987, https://doi.org/10.5194/acp-24-5971-2024, 2024.

Irbah, A., Delanoë, J., van Zadelhoff, G.-J., Donovan, D. P., Kollias, P., Puigdomènech Treserras, B., Mason, S., Hogan, R. J., and Tatarevic, A.: The classification of atmospheric hydrometeors and aerosols from the EarthCARE radar and lidar: the A-TC, C-TC and AC-TC products, Atmos. Meas. Tech., 16, 2795–2820, https://doi.org/10.5194/amt-16-2795-2023 , 2023.

Krüger, K., Groß, S., and Wirth, M.: Validation of cloud macrophysical properties from the ATLID L2a products A-TC, A-FM, A-CTH using airborne lidar observations during the HALO missions PERCUSION, ASCCI and NAWDIC, (to be submitted to AMT, 2026).

Figure 7 and discussion therein

a) The lines on the top panels are difficult to see. In fact, the coincident point was not noticed until referenced in the text and is difficult to see (especially when printed and not on a monitor). Suggest making lines more distinct.
b) The colors in lower panels are also difficult to see. The labels are two small to read when printed.
c) Suggest adding quantitative analysis for these comparisons (mean bias, slope biases, etc) with a simple correlation plot?

We used larger font and label sizes and also thicker lines. We did not add another plot in the manuscript, but added quantitative numbers (mean bias and correlation coefficient) in the discussion of the two lower panels:

'For the 670nm channel, we find a mean bias and standard deviation of (-1.28 +/- 6.93)%, and hence slightly smaller radiances of MSI compared to specMACS, and a correlation coefficient of 0.93 for the entire segment. For the 865nm band, we get (1.95 +/- 7.64)% for the mean bias and 0.93 for the correlation coefficient. For the SWIR channels, we find similar biases of (-2.20 +/- 8.51)% for 1650nm and (1.94 +/- 8.63)% for the 2210nm channel, however, they show slightly larger standard deviations. For the correlation coefficients we derive 0.92 for both channels underscoring the good agreement between the two instruments taking into account that we include time differences of around 10 minutes between the specMACS and MSI measurements.'

Figure 8 and discussion therein

a) Using white points on the bottom is difficult to see the points, which are of most interesting being more close in distance (i.e. suggest you highlight the better points to compare).
b) Figure caption, "close to the underpass" is redundant and can be removed.
c) Why are the histograms shown for the lower panel plot? No discussion is provided on the value to include them.
d) Suggest adding addition statistics such as mean bias for this segment of data.

Comment regarding a), c): We changed the color bar indicating the distance from the underpass accordingly and removed the histograms in the lower panel plot.

Comment on b) – We removed "close to the underpass".

Comment on d) – We calculated the mean bias of the swath averages of MSI and VELOX, which is about 0.15 K. We added this information in the text:

"A mean bias for the swath averages of MSI and VELOX brightness temperatures was found to be about 0.15 K."

Editorial Comments/Suggestions:

line 53 – change "proofed" to "proved"

Done

line 194 – prefer to use "The specsMACS…" at the beginning of a sentence rather than starting with a lower case acronym (see other places in manuscript).

Done

line 234 – remove "once"

Done

line 245 – change "catched" to "flew under"

Done

line 245 – change "…twice, one in northern direction and once in the southern direction of HALO." to "in the northbound and southbound directions."

Done

Line 248 – remove "already"

Done

Line 291-292 – rewrite this sentence as it is not well constructed and is confusing.

We reformulated this sentence and the sentences be for and after as follows:

'Besides this aerosol layer, signatures of low clouds from about 0.5 km to about 2 km altitude are visible in the measurements of both systems. In addition, enhanced backscattering at around 15 km altitude can be seen in the ATLID measurements, this layer

is missed by the WALES measurements due to the lower flight altitude of HALO of about 11 km.'

Line 324 – "After confirming good performance…" add "…for this day".

Done

Line 347 – Backscatter Ratio is now used and I do not think it has been defined.

We added:

'which is the ratio of the total (aerosol + molecule) to the molecular backscatter signal,'

Line 377 – not sure what is meant by "tendentially vertically thicker"

We agree that this was not well-phrased. We changed it to

'This might lead to vertically thicker cirrus clouds in the A-TC, corresponding to the elevated cloud tops (red line) of the A-CTH product.'

Line 385 – what is "bad layering"?

By "bad layering," we refer to Section 2.1.1 of Irbah et al., 2023. This section describes that, for generating the target classification, the Feature Mask is used as input. The Feature Mask (A-FM, van Zadelhoff et al., 2023) serves as the starting point for ATLID Level 2 processing. Its purpose is to provide the best possible strategies for optimal averaging (e.g., determining boundaries between clear-sky and optically thin pixels), which are then used to generate further downstream Level 2 products. If transitions between different flags are less sharp, this can lead to suboptimal processing (for example causing optically thin clouds to be misclassified as aerosol). For further details, we refer to Irbah et al., 2023.

We have tried to clarify this sentence somewhat and have decided not to use the expression "bad layering" as follows:

'There is indication that cloud pixels in A-TC, especially ice clouds, appear too large horizontally and vertically, which could be due to suboptimal averaging in ATLID Level 2 processing (see Irbah et al., 2023).'

Line 401 – suggest to reference the figure in order (a) then (b).

Done

Line 404 – define IQ oLset

Since we want to provide an overview to a broader audience, we removed the technical term "IQ offset" and replaced it with the more descriptive explanation

'As this measurement was acquired during the commissioning phase, the redundant signal processing unit (SPU-A) of CPR was used which exhibited a drift in the Doppler velocity background signal (see clear-sky regions in Figure 5c), biasing low echo cloud regions which disappeared after switch to the operational signal processing unit (SPU-B) on 4 December 2024.'

Line 432 – should this be 84 m or 84 km?

HALO had a match distance of 84 m from the EarthCARE track during this underflight.

Line 543 – change ‘will be’ to ‘is’

Done

Line 555 -suggest using ‘finer’ resolution rather than ‘larger’ resolution

Done

Line 558 – remove ‘also’

Done

Line 575 – Suggest rewording, “comparison…can confirm the good quality if ATLID

> L1b”. This sentence combines that the airborne has the potential to confirm the quality but as stated it infers that this has been done already. Suggest better wording of this statement to reflect that initial measurements have shown qualitative agreement and has been used to identify know issues in the early processing of the ATLID data.

We have generally reworked the ‘Summary and Conclusion’ including this paragraph:

‘A definitive statement about the validation is difficult because the processing of the EarthCARE and HALO data is updated as we gain a better understanding of the measurements. Hence validation is a moving target. For instance, at the very beginning of the EarthCARE mission, PERCUSION measurements helped identify cross-talk effects, background suppression or calibration uncertainties that improved the first release of EarthCARE data products. of ATLID L1b data. It also helped to identify artefacts resulting from bugs in the layer assignment in the algorithm for the ATLID L2a optical products. Airborne radar data from PECUSION measurements helped to identify an offset of the CPR compared to MIRA measurements, and motivated a new calibration of the radar, which led to a much-improved data quality. With the help of the airborne data these problems could be solved quickly and the products publicly available have been updated.’

Line 576 – suggest changing …”are not affected from this issue anymore” to “… have been updated”.

Done