

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 analysis 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 under flights 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 CALIPSO 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 is 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.

Specific Comments/Suggestions:

1. 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/impact statement provided .
2. Recommend moving Table 5 to a supplemental section along with other flight information tables for the PACE and CALIPSO/CloudSat.
3. 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 Rayleigh products are attenuated as shown.
 - b. Labels in upper panel have label of unpolarized - caption states this to be co-polarized and should be changed.
 - 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.
 - d. Why was 120km used for the comparison? Please provide a rationale.
 - e. No error bars provided for this comparison. Provide careful analysis to include error estimates for both datasets.
 - f. No statistical comparison is provided between the two different profiles (e.g. mean differences etc.).
 - 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.

- 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 with 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 of L1 data profiles. Again please clearly explain in the text.
 - i. Is the same molecular density profile used for the air density in WALES and ATLID data?
4. Figure 3 and discussion therein
- a. Why is a 100km averaged used? Please provide rationale.
 - b. What is the shading in the line plots (particle extinction and particle linear depolarization)
 - c. Provide error bars for data.
 - d. No statistical comparison is made between the measurements is presented.
 - 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 provide 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.
 - f. It looks like ATLID does not have backscatter coefficient measurements above 5km. Is this correct?
 - 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.
5. 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.
 - 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 differences. Please provide a discussion on the thresholds used for this comparison.
- c. Similarly, the averaging scales will also impact these comparisons.
 - 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)
6. 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 too small to read when printed.
 - c. Suggest adding quantitative analysis for these comparisons (mean bias, slope biases, etc) with a simple correlation plot?
 7. Figure 8 and discussion therein
 - a. Using white points on the bottom is difficult to see the points, which are of most interest 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 additional statistics such as mean bias for this segment of data.

Editorial Comments/Suggestions:

1. 1.line 53 – change “proofed” to “proved”
2. 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).
3. line 234 – remove “once”
4. line 245 – change “caught” to “flew under”
5. line 245 – change “...twice, one in northern direction and once in the southern direction of HALO.” to “in the northbound and southbound directions.”
6. Line 248 – remove “already”
7. Line 291-292 – rewrite this sentence as it is not well constructed and is confusing.
8. Line 324 – “After confirming good performance...” add “...for this day”.
9. Line 347 – Backscatter Ratio is now used and I do not think it has been defined.

10. Line 377 – not sure what is meant by “tendentially vertically thicker”
11. Line 385 – what is “bad layering”?
12. Line 401 – suggest to reference the figure in order (a) then (b).
13. Line 404 – define IQ offset
14. Line 432 – should this be 84 m or 84 km?
15. Line 543 – change ‘will be’ to ‘is’
16. Line 555 -suggest using ‘finer’ resolution rather than ‘larger’ resolution
17. Line 558 – remove ‘also’
18. 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.
19. Line 576 – suggest changing ...”are not affected from this issue anymore” to “... have been updated”.