

Response to Editor

We thank the editor for their feedback. We have taken care to respond directly to each comment. The text from your comments are shown in black, and our responses are shown in blue. The responses include the manuscript text that was changed, removed, or added.

- 5 Thank you for your changes to the first draft of the manuscript. Two of the reviewers request major revisions, and I agree with their assessment. There is a lot of good information here, but some of the discussion is unclear and there appear to have been some decisions made that could be changed to improve the analysis. In addition to the comments of the reviewers, I have the following comments/questions:

10

1. Why do you apply the gamma parameterization (eq. 15) to extrapolate extinction to RH80 and ambient RH? The gamma parameterization is an arbitrary functional form (see, e.g., <https://doi.org/10.5194/acp-16-4987-2016>). Since you already have a kappa value, why not just apply the Koehler equation to calculate the growth to these other RH values? It seems like an unnecessary step to go to the gamma parameterization, and it adds uncertainty, especially at high RH values.

15 Thank you for the comment. Both κ and γ are somewhat arbitrary forms. At this point in the retrieval step, we do not have a CRI or κ . The reason that we do not use $C_{\text{RH, meas}(85\%)}$ directly is that the RH associated with that measurement is not reported. To reduce the uncertainty with what exact RH to prescribe, we opted to use the γ parameterization to fix the RH to 80% for the retrievals. We feel no changes are needed.

- 20 2. I want to confirm that you converted the values in the ICARTT-format files for size distributions and scattering and absorption values from the reported STP conditions to ambient T and P when calculating extinction for comparison with the lidar profiles. I strongly suspect you did (given the good agreement on the “unicorn” profile), but I’d like it clearly stated in the text somewhere.

25 Thank you for the comment. As you suggested, we performed the adjustments to compensate for temperature and pressure. To clarify this we added text to Sect. 3.2.2 and 3.3. We added the following text to Sect. 3.2.2 “These data are provided with a standard temperature and pressure conversion factor that translates the standard number concentrations from standard temperature and pressure to ambient temperature and pressure conditions (See Sect. 3.3)”. We also added the following text to Sect. 3.3: “The first step of calculating the ambient aerosol properties is to convert the number concentrations from standard temperature and pressure to ambient temperature and pressure using. After temperature and pressure adjustment, the hygroscopic adjustments are performed.”

30

3. You state that you use an extinction-weighted average across HSRL-2 altitude bins when comparing the ISARA data products with the HSRL-2 profiles. I understand why you do this for the comparison with RSP, whose signal will be weighted by extinction, but are you also doing this to compare with HSRL-2? It seems a simple unweighted mean would be appropriate for independent samples being averaged into HSRL-2 altitude bins for comparison with extinction within that bin as measured by the lidar. Please explain.

35

Thank you for the comment. We are only using extinction weighted averaging for the column-averaged data. We are using simple unweighted means for the vertically resolved data. We clarified the text as follows:

40

Once the desired data from the vertical profiles are identified, we select the remote sensing profile that is the nearest (horizontally) to a Falcon profile and is within 6 min of the start or stop time of the Falcon’s profile. Once a viable profile collocation is identified, ISARA products are averaged to the 225 m HSRL-2 bins for the HSRL-2 comparisons or weighted by extinction and averaged through the entire profile for RSP comparisons. For the case of the HSRL-2-ISARA comparisons, this averaging results in altitude-resolved properties for a given column of air. For the RSP-ISARA comparisons, this results in a column-average property. Note that for AOD, the ISARA-derived ambient extinction is vertically integrated with sample altitude and does not account for the altitude bins above or below the sample altitudes. This allows us to estimate the amount of the total aerosol column that was sampled by the Falcon.

45

4. In several of your figures you redefine parameters (e.g., in Fig. 5 you restate the equations for number, surface, and volume size distributions). This is not necessary.

Thank you for the comment. The redundant definitions from Fig. 5 have been removed.

5. I agree with the reviewer that additional synthetic size distributions should be evaluated. I suggest picking one that is characteristic of the marine boundary layer, one for the clean background free troposphere, one for smoke cases, and then the one you have for the mission-typical size distribution. It's helpful to choose a single size distribution observation for each of these cases, because averaging many size distributions tends to significantly broaden the standard deviation. It would also be useful to fit lognormal curves to each of these distributions and then use the lognormal fits to do the analysis—this aids in repeatability of your calculations by others if needed.

Thank you for the comment. We updated our synthetic analysis to include variability in the shape of the size distribution. We accomplished this by fitting the size-distributions that are observed in the ACTIVATE data and selecting the best fits from these to serve as the source of possible size-distributions that are available for the synthetic data generation. We updated all the sections related to the methods and results of the synthetic consistency analysis accordingly.

6. In your internal consistency checks you use a range of 1.51-1.55 for the RRI. The RSP retrieval gives a lower retrieved RRI. In many cases organic particles or sulfate particles might have a lower RRI. You may want to expand the range of allowed RRI to lower values.

Thank you for the comment. While it would be nice to investigate wider RRI ranges, it is not feasible due to the assumptions made with the particle sizing. We have used a wider range to account for some sizing uncertainties and for uncertainties in the scattering coefficient measurements. Finally, we did do a preliminary analysis to see if our measurements were sensitive to a wider range, but found that making the range wider does not improve our retrieval success rate. We feel there are no changes needed.

7. In Figure 7 you show excellent agreement between the ISARA-calculated and measured scattering coefficient values. It looks great! However, this is showing only the data that achieve a “match” and eliminates the ~40% of data for which ISARA did not provide a good match. Please plot all of the data (probably including the poor matches as grey points (you'd have to change the background color)) to show the performance of the ISARA approach when it fails as well as when it succeeds. Same comment for Fig. 8.

Thank you for the comment. We are not sure how to assign values as failed retrievals. The successful retrievals are an average of values within a given threshold. To do this we would need to rerun the ISARA several times with various thresholds on the successful retrieval. This method was proven successful in Sawamura et al. (2017), for example. This type of internal consistency is rarely published with any significant detail, and we feel it is outside of the scope of this work to analyze the impacts of various thresholds and optimization methods might be best. Additionally, RC4 has already indicated that this manuscript is lengthy and we would agree. Finally, the updated revision has many improvements that have increased the science value presented here. In our opinion, this optimization analysis deserves its own study. We feel there are no changes needed.

8. In Sect. 3.1.3 you don't really explain how the HSRL-2 and in situ data were compared. Are the profiles sloping or spiral? How many lidar shots are combined to produce the lidar profile for comparison? How are the times aligned (start to stop of the profile vs. start to stop of the lidar). This is a four-dimensional matching problem and there needs to be more detail here about how these comparisons were put together. That would help the reader understand some of the poorer agreement cases evident in Fig. 9, as opposed to the “unicorn” case.

Thank you for the comment. As we discuss in this section, Schlosser et al. (2024) was written to completely detail the collocation process. The following detail was added to explain the Falcon profiles better:

For this work, we limit the collocation periods to those where the Falcon performed a controlled ascent or decent. These vertical profiles are identified using the leg identifier available as part of the ACTIVATE data set (Sorooshian et al., 2023). The points of interest are classified as part of a Falcon spiral, ascent, or decent vertical profile. During the ascent and decent profiles, the Falcon would limit its airspeed to 180 m s^{-1} while performing a vertical profile while continuing to move in one direction. During the spiral profiles, the Falcon would spiral around a predetermined location while ascending or descending in addition to limiting its airspeed to 180 m s^{-1} .

9. In Fig. 9, please use more of the plotting space in panels c and i. Also, the highest altitude in the profile was $<4.5 \text{ km}$ (line 723), yet the color scale goes up to 5.5 so not all of the dynamic range is used.

Thank you for the comment. We increased the used space in panels c and i. Additionally, we have removed the altitude color bar and instead binned the points based on whether or not the corresponding in-situ profile had more than 3 points. The blue points are points where the profile had 3 or more points, and the red points correspond to the entire data set.

10. In the AOD comparison (Table 4), is the AOD for HSRL-2 calculated over the altitude range of the Falcon, or did you just use the archived AOD from the HSRL file? You need to describe how this comparison was calculated. What did you do with any missing in situ data from the profile (e.g., in Fig. 11 not every HSRL-2 altitude bin has a corresponding ISARA value)?

Thank you for the comment. We removed the altitude color mapping and instead categorize the points based on whether a profile had 3 or more coincident points and color them accordingly. We updated the figure caption and associated text to explain the details. Additionally we discussed this briefly in another comment, but we do not account for missing points in the column when integrating the ISARA-derived properties. This is because we do not expect to be consistent with something like AOD and we can estimate the amount of aerosol sampled by the Falcon by comparing the resulting AOD. We have added the following text to explain this “Note that for AOD, the ISARA-derived ambient extinction is vertically integrated with sample altitude and does not account for the altitude bins above or below the sample altitudes. This allows us to estimate the amount of the total aerosol column that was sampled by the Falcon.”.

11. There are several typos floating around. a) In Table 11 Altitude is in m, not km. b) Lines 269 and 276 the size distributions are n_0 , not N . c) Table 1 HSRL LDR is at only 355 and 532 wavelengths (1064 is extrapolated). e) Line 300, “size-dependent”. f) Line 760, “predominant”.

Thank you for the comment. Points a-c have been corrected. Point e is no longer relevant as this text was changed. Point f does not seem to be in error. Was there something about the use of predominant that is incorrect? We do not see any point d.

12. Line 291 “success rate” appears here before you’ve explained the ISARA algorithm, so it was puzzling.

Thank you for this comment. We have moved this discussion of “success rate” in to Sect 3.3.

13. “Stitching” is not explained (combining the LAS and SMPS size distributions).

Thank you for pointing this out. The sentence now reads “These calibrations resulted in good stitching between the SMPS and LAS distributions, where stitching refers to the process of combining aerosol size distributions from different instruments into a continuous data set. They also led to consistent integrated number concentrations measured when compared to ancillary CPC measurements (see Figure 7 of Sorooshian et al., 2023).”.

14. Table 2. Systematic uncertainties for size distributions are given as 10%. In diameter? In number concentration? Across all size ranges? Similar questions for the CAS and FCDP and CDP probes. I would think that counting statistics would be a source of random uncertainty in these instruments as well.

Thank you for the comment. These uncertainties are given in terms of number concentration. A random uncertainty is not provided with the data.

15. Line 372. Coarse mode particle may not contribute to concentrations, but can be absolutely important to surface area and volume.

Thank you for the comment. We agree that surface area and volume concentrations are important. We removed this sentence as it is not really accurate anyway given we are using the CAS for the measurements of coarse aerosol.

16. Fig. 1. Could you please use standard decades for x-axis labels? I don't know how to interpret the existing labels.

Thank you for the comment. We changed our x-axis labels to standard decades here and in the other figures showing size-distribution.

17. Line 421. Why do you use the smallest k values that produce scattering that matches to within one percent? For RRI and CRI you use the mean of all values that match. Why the difference here?

Thank you for the comment. This change would be a deviation from how this type of retrieval was done in Sawamura et al. (2017). We have already made several improvements over the previous and we would like to limit the number of changes made.

18. Line 427. Could you compare your kappa values with those derived from the AMS that I believe was on the Falcon for ACTIVATE? This is quite a low kappa for an aged, remote aerosol (e.g., <https://acp.copernicus.org/articles/21/15023/2021/acp-21-15023-2021.pdf>).

Thank you for the comment. RC4 had a very similar comment. Reiterating what was conveyed to them, there is AMS data available as part of the ACTIVATE dataset; however, that data would add even more length that they have stated is already a long paper. Additionally, there is currently a separate paper analyzing the differences between the different hygroscopicity parameterizations. The convenience of the $f(RH)$ derivation is that the retrieved hygroscopicity is internally consistent with the optical coefficient measurements.

19. when discussing biases (e.g., Fig. 6 and elsewhere), please clearly state how the bias is calculated. Is it simulated minus measured? In Fig. 6 the axis labels are unclear. Is "calculated" the synthetic data, or is that "measured"?

Thank you for the comment. We have clarified the x- and y-axis labels. For clarity, the bias is always calculated as $y - x$. We made statements in each of the consistency analysis sections where we explain which of the properties are used as the reference (y) and which are used for comparison (x).

20. In section 3.1.2, did you consider uncertainties in the gamma value for the extinction parameterization?

Thank you for the comment. We generate our dry and wet synthetic properties and add noise separately to each synthetic property. It is outside of the scope of this work to analyze the uncertainties in the γ and κ parameterizations as compared to real particle growth.

21. Table 3. Explain what "smoke counts" are. It appears elsewhere later, but is unexplained when it first appears.

Thank you for the comment. This is referring to the aerosol typing product derived from the HSRL-2. To clarify this, we added "This study also uses the advanced aerosol typing product provided by the HSRL-2 team (Burton et al., 2012, 2013). Specifically, we count the number of bins flagged as smoke in the HSRL-2 aerosol typing product within each $225 \text{ m} \times 60 \text{ s}$ bin. This HSRL-2-derived smoke count is used to analyze for the presence of elevated smoke layers." to Section 3.2. We also added "typing product" after HSRL-2 in this sentence.

- 185 22. Table 3. What is “platform separation”? Is this the average horizontal offset during the period of the profile? What is the range of values of separation (12 m is phenomenally close—shorter than fuselage length). How these profiles are compared needs some discussion, as noted above.
Thank you for the comment. We define platform separation as the horizontal separation between the center point of the Falcon’s profile and King Air. We added the following clarifying text to the caption of Table 3: “the horizontal separation between the center point of the Falcon’s profile and King Air (i.e., platform separation)”.
- 190 23. Table 5. “Optical N” is a pretty arbitrary comparison since the size distribution slopes so dramatically around these sizes. If you defined optical N as > 80 or as > 100 you would probably get a very different value. I’m not sure of the value of this product (I guess it goes into the calculation of σ_{ext} for comparison with RSP, but still seems arbitrary). Is 90 chosen because it is the LAS lower size limit?
195 Thank you for the comment. Because the particle size-distributions are in terms of ambient diameters, this is not necessarily the lowest LAS bin. It is possible, it is one of the SMPS bins. That being said, the primary justification is to match with the optically active particle size-distribution that is theoretically observed in the HSRL-2 RSP measurements. This decision is explained in more detail in Schlosser et al. (2022), but that work was limited to assuming the aerosol diameters are dry.
- 200 24. Table 6. I really don’t understand why, if C_{ext} and C_{bsc} are very well correlated with the calculated values, the ratio of these is not at all correlated with the calculated values. Would you please look into this?
205 Thank you for the comment. Note the comparisons in the revised manuscript are much improved, relative to the previous version. These improvements are mostly due to the consideration of a non-spherical coarse aerosol. That being said, LR consistency is still generally worse as compared to extinction and backscatter coefficients. One reason the LR is not as well correlated as the extinction and backscatter coefficients is because LR has fewer points of comparison. Additionally, there is more noise in the HSRL-2-derived LR.
- 210 25. Line 826, you state that under-sampling of coarse-mode particles is likely to blame for in situ/remote disagreements but never really explore this. Your size distributions go up to several microns behind the inlet, and the cloud probes are quite capable of measuring from a couple of μm and larger. Do you have evidence that the cloud probes are undersampling? What is the basis for this statement? It’s pretty important if our in situ techniques are not good enough to evaluate remote sensing methods. Why is the unicorn case so good if we can’t measure coarse particles well? Might it be that the other profiles are so short that it’s hard to match things up in less than ideal circumstances? This would argue that closure studies need to be performed extremely carefully only in ideal circumstances.
215 Thank you for the comment. We do agree that achieving spatiotemporal collocation is challenging. In our revisions, we were also able to improve our disagreement between the ISARA- and the remote sensing-derived properties. Still, we think that undersampling of the number concentration of aerosol with diameter greater than $5\ \mu\text{m}$ ($N_{>5\ \mu\text{m}}$) under background conditions ($< 1\ \text{cm}^{-3}$) is still a limitation of the in-situ instruments. For clarity, we added the following to the introduction: “The measurements of number concentration are less accurate for background concentrations ($N < 1\ \text{cm}^{-3}$), which is common for aerosol with diameters $> 5\ \mu\text{m}$ (Baumgardner et al., 2001).”
220
- 225 26. I agree that the case study “unicorn case” should be presented first, to show how the comparisons are done in an ideal situation. Then you can show the results for all the profiles together.
Thank you for the comment. We changed the order in which these sections are presented. We added a second case study that is presented in the case study section (Sect. 4.3)

Response to Reviewer 2

230 We thank the reviewer for their feedback. We have taken care to respond directly to each comment. The text from your comments are shown in black, and our responses are shown in blue. The responses include the manuscript text that was changed, removed, or added.

I consider the revised version of the manuscript ready for publication, pending a final check for any remaining typographical errors. I would also encourage the authors to reconsider the subheadings in the Methods section.

235 Thank you for the comment. We appreciate your effort in reviewing our revised manuscript. We made further improvements in this revision that were guided by the comments from the other reviewers and the editor. We feel the new version has even more scientific value.

Response to Reviewer 3

240 We thank the reviewer for their feedback. We taken care to respond directly to each comment. The text from your comments are shown in black, and our responses are shown in blue. The responses include the manuscript text that was changed, removed, or added.

I feel that the authors have been through a large effort towards focusing more on the science of their paper than the previous
245 version. They have taken an open mind and they have taken most comments into account. Congratulations because I can see your big efforts!

The starting point was however quite far from acceptable (if you remember my previous review). The authors have covered a great deal of the distance needed, but I feel that there are still some MAJOR points that need addressing, hence at this stage I
250 recommend a MAJOR REVISION.

Thank you for your continued efforts in reviewing this manuscript. Your comments have been integral to the marked improvements in the consistency results and further increased the scientific value of this work. In its newly revised version, we feel we have fully addressed each of your major and minor points.

255 The **MAJOR POINTS** of concern are highlighted below:

1. The synthetic consistency analysis is a bit limited because the authors have chosen to keep a fixed particle size-distribution, varying only the total number (see lines 476-477 and 578-580). This prevents a true synthetic consistency check, which should include exploring different PSDs, different mixes between larger and smaller particles, etc. I feel
260 that this could easily be addressed and that there are no justifiable reasons for this self-limitation. I therefore invite the authors into exploring different dominant particle sizes using modelled monomodal and bimodal PSDs.

Thank you for the comment. We updated our synthetic analysis to include variability in the shape of the size distribution. We accomplished this by fitting the size-distributions that are observed in the ACTIVATE data and selecting the best fits from these to serve as the source of possible size-distributions that are available for the synthetic data generation. We
265 updated all the sections related to the methods and results of the synthetic consistency analysis accordingly.

2. The external consistency presented in section 3.1.3 is not fully convincing. In particular, figure 9 shows a low bias and I feel that the authors are too simplistically trying to dismiss this bias (lines 676-677; 699-701) as not being something to discuss more in-depth. The fact is that the data in Figures 9a,b,d,e show that the data are not near the 1:1 line (as they
270 should be) and they follow a different dependency. This point had been raised already in the forst review (points 6 and 8).

Thank you for the comment. In the revised version of the manuscript, we show marked improvements to the results. The majority of improvements are related to our treatment of nonspherical coarse aerosol. We updated all of the methods, results, and conclusions to reflect these improvements and adjustments to our methods.

275 3. Similarly for the LR dependency, which seems not to be responding to the microphysical parameters (line 682) but this does not seem to be evaluated critically.

Thank you for the comment. In our revised manuscript we made significant improvements in the LR consistency. These improvements stem from adding using two distinct spheroid approximations for the coarse aerosol properties. These improvements are reflected throughout the methods, results and conclusions. We still agree that LR consistency has more
280 room for improvements even over our best efforts. There is a functional limitation in the available shape we can optically represent. Additionally, improved measurements of coarse aerosol composition and optical properties could help.

285 4. The case study (figures 11 and 12) shows a lot of promise, and it is much more convincing than the external consistency analysis in section 3.1.3. I suggest to present the case study before the full dataset so as to show what is working and

convincing before showing something a little more problematic.

Thank you for the comment. We reorganized the sections to accommodate this comment.

- 290 5. The authors have tried to avoid my question about VLDR and PLDR (first review, point 25, and line 524 of the revised manuscript). However it is quite an important point because the chosen threshold of 13% need to be understood by the reader and the reviewer. Until this is clarified, it is not possible for me to understand this choice fully. I am under the impression that the authors maybe do not know the distinction between VLDR and PLDR, and have therefore not fully understood the question: to advance on this point, I invite them to discuss it with their lidar experts.

295 Thank you for the comment. We reviewed our analysis presented in the previous work and found the we were indeed using the volume LDR instead of the aerosol LDR. This realization significantly changed our view on the HSRL-2 data as it made us realize that even our unicorn case had measurable presence of nonspherical aerosol as evident from the HSRL-2-derived 532 nm LDR which was as high as 0.18. This provided us with a unique opportunity to explore non-spherical aerosol in our analysis. While our fine aerosol are still constrained to spherical approximations, we use the HSRL-2-derived LDR and LR to inform our assumptions on the coarse aerosol. We have updated the entirety of the text to explain the added nonspherical information. This has led to a marked improvement in the C_{ext} , C_{bsc} , and LR consistency. Additionally it has opened the discussion to LDR, which has been added throughout the text. Additionally we now specify that the LDR in question is the aerosol LDR in Table 1.

- 305 6. The text at lines 752-753 raises an important question: how doe the authors account for layers that are not sampled in-situ? Perhaps, using the HSRL-2 data a rigorous way to account this is possible. In any case, this needs discussion.

As the goal of this work is to establish consistency rather than complete closure, we do not expect to retrieve AOD. We added the following text for clarity “Note that for AOD, the ISARA-derived ambient extinction is vertically integrated with sample altitude and does not account for the altitude bins above or below the sample altitudes. This allows us to estimate the amount of the total aerosol column that was sampled by the Falcon.”.

- 310 7. There are several minor points of revision that I highlight in the attached annotated manuscript. Although they are minor, their number is large, hence they overall constitute a major point of concern.

Thank you for the comment. Your minor points have been restated below along with our response to each.

- 315 – Line 1: ~~particle~~

Thank you for the comment. An aerosol is a suspension of solid particles or liquid droplets in air or another gas. While it is common to use aerosol to refer to just the particles, we feel this needs to be made clear within the context of our study. We initially refer to the particles as “aerosol particles” to provide context to our study. However, we removed “particle” and “particles” in the subsequent text.

- 320 – Line 2: ~~impacts~~ effects

Thank you for the comment. We changed “impacts” to “effects”.

- 325 – Line 5: ~~typically~~

Thank you for the comment. We removed the word “typically”.

- Line 7: ~~software framework~~ methodological framework

Thank you for the comment. We changed “software” to “methodological”.

- 330 – Line 11: ~~measurement~~ in-situ and remote sensing
Thank you for the comment. We replaced “measurement” with “in-situ and remote sensing”.
- Line 11: “...ISARA-calculated aerosol data...” I would suggest to be explicit and indicate the list of properties that ISARA calculates.
- 335 Thank you for the comment. We feel this level of detail would be difficult to capture succinctly in the abstract and would be distracting from the findings presented. We added the text “intensive and extensive” to add more generic detail.
- Line 12: ~~data~~ properties
- 340 Thank you for the comment. We replaced “data” with “properties”.
- Line 18: Move aerosol to be written before imaginary refractive index.
Thank you for the comment. We made this change.
- Line 18: “...probable under-sampling...” explain why.
- 345 Thank you for the comment. We attribute this to low aerosol concentrations ($< 1 \text{ cm}^{-3}$) for aerosol larger than $5 \mu\text{m}$ that are measured throughout the mission. We added the following explanation to the abstract: “low background concentrations ($N < 1 \text{ cm}^{-3}$) of coarse aerosol sizes greater than $5 \mu\text{m}$.” We also added the number concentration of aerosol with diameter greater than $5 \mu\text{m}$ ($N_{>5 \mu\text{m}}$) information for each of the cases to Table 4.
- 350 – Line 20: ~~overall~~
Thank you for the comment. We removed “overall” from the text.
- Line 20: ~~has~~ could have
- 355 Thank you for the comment. We replaced “has” with “could have”.
- Line 23: “central”, Line 25: “central”, Line 27: “key”, Line 27: “critical”, Line 35: “critical”. there is an inflation of emphasizing words (central, key, critical, etc.). I think that modesty dictates that you should have one emphasizing word at most (or even none?). Usually understatement raises interest in your work, whereas emphasis very seldom persuades readers.
- 360 Thank you for the comment. We removed the emphasizing words on lines 25, 27, and 35, but have kept “central” on line 23.
- Line 44: “Tsekeri et al., 2017”. In the response to my review (point 2 of response to reviewer 3) you have given an explanation of the differences between the work by Tsekeri and yours. I feel that the paper would benefit from having these comments in its introduction instead, as they put the reader into the wider context of different possible approaches and their pros and cons.
- 365 Thank you for the comment. We moved the following text from the conclusion to the introduction:
- 370 This work extends the methods established by Sawamura et al. (2017) by attempting to account for coarse aerosol in the bulk aerosol properties after retrieving the fine CRI and κ . Similar to the work of Tsekeri et al. (2017), we account for the contribution of coarse (ambient aerosol diameter $> 1.0 \mu\text{m}$) particles using wing probe measurements.
- Line 66: “these measurements”. I suppose you want to say “the humidification corrections” or “the humidification curves”? in any case “these measurements” is unclear and does not indicate to what you want to refer when mentioning these high errors.
- 375

Thank you for the comment. The text “these measurements” is referring to the ambient RH measurements. The text is now “these RH measurements” for clarity.

- Line 66: “15%”. I would not regard 15% as a very large error for the humidity-corrected properties. Are you sure that it is not much bigger?

Thank you for the comment. While it may seem trivial, a 15% error in ambient RH means that a measured RH of 80% can be as low as 68% or as high as 92%. We feel no changes are needed given this can be a significant source of error that propagates through to calculating the properties of humidified aerosol.

- Line 77: placed behind inlets

Thank you for the comment. This text has been added.

- Line: 78: “typical”. I would not regard 5 μm as “typical” unless (like here) there is an isokinetic inlet. Regular inlets could have a much smaller cutoff diameter.

Thank you for the comment. We removed “typical” from this sentence.

- Line 79: open-path.

Thank you for the comment. We added “open-path” to the text.

- Lines 79-80: “sizing can be highly uncertain”. please explain better why you think that wing-mounted probes are more uncertain than probes behind an inlet?

Thank you for the comment. We agree that they are at least similar to the errors derived from the measurements behind the inlet. We reworded this text as follows:

Wing-mounted open-path probes are commonly used to estimate coarse aerosol properties, but are designed in such a way that the sizing can be uncertain (e.g., Reid et al., 2003, 2006). The measurements of number concentration are less accurate for background concentrations ($N < 1 \text{ cm}^{-3}$) of aerosols with $D > 5 \mu\text{m}$ (Baumgardner et al., 2001). In addition to the coarse sampling limitations, particles are lost between the external inlet of the aircraft and the inlets of the instruments (Baron and Willeke, 2001; Kulkarni et al., 2011).

- Lines 80-81: “In addition to the coarse sampling limitations, particles are lost between the external inlet of the aircraft and the inlets of the instruments (Baron and Willeke, 2001; Kulkarni et al., 2011).”. Move this sentence to line 75.

Thank you for the comment. We moved this line to the start of the paragraph. We also reworded it as follows: “In addition to the measurement limitations in measuring relative humidity and hygroscopicity, particles are lost between the external inlet of the aircraft and the inlets of the instruments (Baron and Willeke, 2001; Kulkarni et al., 2011).”.

- Line 86: ~~Python-based~~. This is scientifically irrelevant. You may want to add the technical details (programming language, Mie scattering library used) at the end, otherwise they risk to shadow your scientific work and make it appear like a programming effort.

Thank you for the comment. We removed this “Python Based” and moved this information to the Code Availability section.

- Lines 88-89: ~~the Fortran-based Modeled Optical Properties of Ensembles of Aerosol Particles package (MOPSMAP; Gasteiger and Wiegner, 2018)~~ a Mie scattering library.

Thank you for the comment. We reworded the text to be as follows “Specifically, the algorithm uses a Mie scattering

library (e.g., Mishchenko et al., 2002; Bohren and Huffman, 2008) in conjunction with measured size distributions and optical coefficients to retrieve refractive indices and hygroscopicity from a "common" suite of in-situ instruments.". We move the information about MOPSMAP to the Code Availability Section.

- Line 97: “dry”. The same applies to wet sea salt.

Thank you for the comment. We removed “dry” from this sentence.

- Line 97: “are difficult”. yes they have a larger size, but this does not make these particles “difficult to consider”!

Thank you for the comment. We agree this was poorly constructed. We reworded these sentences for clarity:

Common coarse species that can have diameters $> 1.0 \mu\text{m}$ are sea salt and dust (Hussein et al., 2005). As discussed above, the larger ($D > 5 \mu\text{m}$) sizes of these coarse species are difficult to measure in background concentrations, however each species also poses a unique challenge from an optical perspective. Dry sea salt is non-spherical, non-absorbing, and very hygroscopic, which translates to larger values of κ (Sorribas et al., 2015; Moosmüller and Sorensen, 2018; Ferrare et al., 2023). Similar to sea salt, dust can also be non-spherical. In contrast to sea salt, dust can be moderately absorbing and has a complex refractive index (CRI) that is dependent on wavelength (Voshchinnikov and Farafonov, 1993; Veselovskii et al., 2010; Wagner et al., 2012; Sorribas et al., 2015).

- Line 101: ~~overcome these limitations by using~~ use

Thank you for the comment. We replaced “overcome these limitations by using” with “uses” as “use” would not have been grammatically correct.

- Line 102: open-path

Thank you for the comment. We added the text “open-path”.

- Line 102: ~~sample coarse-mode~~ permit

Thank you for the comment. We removed “sample coarse-mode” and reworded the sentence as follows: “ This work uses wing-mounted open-path probes to estimate the contribution of the coarse aerosol on the calculated ambient optical and microphysical data (Ryder et al., 2015; Tsekeri et al., 2017; Ryder et al., 2018)”.

- Line 103: ~~, which is common for studies looking to account for coarse-mode~~

Thank you for the comment. We removed the text “, which is common for studies looking to account for coarse-mode”.

- Line 124: “the coarse is assumed to be sea salt”. specify: dry or wet sea salt?

Thank you for the comment. We added the specification of “humidified sea salt”.

- Lines 131-132: “limited ability of in-situ instruments to sample coarse particles”. I do not agree with the stated "limited ability". Open-path instruments such as the ones that you use permit to sample the coarse mode.

Thank you for the comment. Please see our response to your comment regarding Line 18. We feel we have addressed this concern with the added information regarding low concentration of $N_{>5 \mu\text{m}}$.

- Line 152: (liquid)

Thank you for the comment. We feel this is an incorrect relation. Sulfate particles are well approximated as spheres in both solid and liquid states, for example. Also, dust is not always approximated as a sphere even when humidified. No changes are made as a result.

- Line 152: (solid)

Thank you for the comment. We addressed this in the previous comment. No changes are made as a result.

- Line 160: ~~composition~~ complex refractive index

Thank you for the comment. We replaced “composition” with “CRI”.

- Lines 169, 170, 173 and 187: ~~extrinsic~~ extensive

Thank you for the comment. We replaced the word “extrinsic” with “extensive” here and throughout the remaining text.

- Lines 170, 172, 173, and 188: ~~intrinsic~~ intensive

Thank you for the comment. We replaced the word “intrinsic” with “intensive” here and throughout the remaining text.

- Line 174: therefore using any Mie scattering library would produce equivalent results

Thank you for the comment. We added the text “; therefore, using any Mie scattering library would produce equivalent results.”.

- Line 186: where dN is the number of particles counted in a bin

Thank you for the comment. We added the text. “where dN is the number of particles counted in a bin.”.

- Lines 193-194: “For particles larger than the remote sensor’s observing wavelength, the remote sensor is most sensitive to particle cross-sectional area (i.e., surface area concentration). For particles smaller than the observing wavelength, the remote sensor is most sensitive to volume concentration.”. This statement is not correct. The cross-section is $Q \cdot A$ where Q = extinction efficiency and A is the area. Your statement is equivalent to saying that for small particles $Q \sim R$, but in fact typical Q curves are not like this. You can check it with your Mie code, or you can google it. For example: <https://www.researchgate.net/profile/Carynelisa-Haspel/publication/26454094/figure/fig3/AS:669075336359945@1536531399888/Theextinction-efficiency-Q-ext-as-a-function-of-size-parameter-x-of-polystyrene.png> This is the same as comment number 17 of my first review, not satisfactorily addressed.

We rewrote this sentence for clarity and relevance to the paper as follows: “For remote sensing of fine- and coarse aerosol particles, polarimeters and high-spectral-resolution lidar that operate in the visible to near-infrared part of the spectrum are sensitive to the aerosol effective radius, effective variance, real refractive index, and single-scattering albedo (Hansen and Travis, 1974; Burton et al., 2016; Stamnes et al., 2018).”.

- Line 204: campaign

If we were to write the acronym out this sentence would read “The Aerosol Cloud Meteorology Interactions over the western Atlantic Experiment campaign...”. The extra word “campaign” is redundant with the word experiment in the ACTIVATE acronym. We followed the examples of previous studies published related to ACTIVATE (e.g., Sorooshian et al., 2023, 2019).

- Line 212: “During the first five and a half ACTIVATE deployments”. move the description of the two aircraft used to before this text, and indicate their respective role (in-situ / remote sensing)

Thank you for the comment. We moved the next paragraph to be before this paragraph. This paragraph does state explicitly what roles each aircraft is serving. Specifically, the text states “...lower-flying HU-25 Falcon aircraft collected in-situ data...” and “...higher-flying King Air at approximately 9 km would conduct remote sensing and launch dropsondes...”.

- 515 – Line 269: “ N ”. this should be dN and not N , given that N is the total number of particles regardless of their size bin (not size-resolved)
Thank you for the comment. We addressed this by replacing N with n° . The text is now as follows: “In-situ measurements of dry logarithmic size-resolved aerosol number concentration (n°)...”.
- 520 – Line 277: “ N ”. dN
Thank you for the comment. We changed N to be n° .
- Lines 284-285: “The impact on the absorption coefficient from particles above $1\ \mu\text{m}$ is assumed to be negligible in the calculation of extinction coefficients.”. You have the tools to easily calculate it (the Mie scattering library), therefore why this assumption?
525 We make this assumption to avoid having to apply two different size distributions during the ISARA CRI retrieval. We would need to use the entire distribution up to $3.49\ \mu\text{m}$ one for the absorption coefficient, which is not behind the cyclone, and the truncated size distribution for the scattering coefficient, which is behind the cyclone. In reality these are only particles in the 2.24 to $3.49\ \mu\text{m}$ range which have size parameters that fall in the maximally scattering range observed in Moosmüller and Sorensen (2018). We updated this sentence to read as follows: “The impact on the absorption coefficient from particles between 2.25 and $3.49\ \mu\text{m}$ is negligible in the calculation of absorption coefficients (Moosmüller et al., 2009; Moosmüller and Sorensen, 2018). This assumption may result in an overestimation in the IRI in some cases, which is important to consider.”.
530
- Lines 286-292:
535 It is important to note that the optical particle size can be greater than the aerodynamic size by a factor of 1.2 to 1.8 depending on the particle density and shape. Additionally, the cyclone has a 50% efficiency at $1\ \mu\text{m}$. Due to the difference between the aerodynamic and optical particle sizes and the imperfect nature of the cyclone, we use a maximum cutoff diameter of $2\ \mu\text{m}$ for the upper bound of the LAS size distribution.
540 To further motivate this decision, using lower thresholds of 1.5 and $1.8\ \mu\text{m}$ resulted in a success rate of 33% and 10% , respectively. There is also a decrease in the internal consistency as all measures of bias increase when a cutoff of less than $2\ \mu\text{m}$ is used.
These six lines are completely unclear to me.
We updated this to read as follows:
545 For the cyclone, it is important to note that the optical size can be greater than the aerodynamic size by a factor of 1.2 to 1.8 depending on the density and shape. Additionally, the cyclone has a 50% efficiency at $1\ \mu\text{m}$. Due to the difference between the aerodynamic and optical sizes and the imperfect nature of the cyclone, we select the LAS bin that has a D_{gm} of $2\ \mu\text{m}$ as the last bin of size distribution. This bin samples aerosol with diameters ranging from 1.79 to $2.25\ \mu\text{m}$. While this bin extends to aerosol sizes as high as $2.25\ \mu\text{m}$ this bin was chosen to limit the erroneous over-truncation of the size distribution. This decision is discussed in more detail in Sect. 3.3. The impact on the absorption coefficient from aerosol between 2.25 and $3.49\ \mu\text{m}$ is negligible in the calculation of absorption coefficients (Moosmüller et al., 2009; Moosmüller and Sorensen, 2018).
550
- Line 317: and $C_{\text{scat,measured}}$
555 Thank you for the comment. We added “and $C_{\text{scat,measured}}$ are”.
- Line 320: “ $0\ \text{Mm}^{-1}$ ”. can the signal be less than zero? please explain how this is possible
Thank you for the comment. The measurement uncertainty can lead to the in-situ measurements having signals less than zero, which happens more at lower signals. We added the following text for clarification: “Note that the measurement uncertainty can lead to the in-situ measurements having signals less than zero. This is most important
560

for low signal measurements and can impact some or all channels. ”.

- Lines 321-331:

Measurements of ambient liquid water content (LWC) and cloud drop number concentration (N_d) are used to classify in-situ data as cloud-free, ambiguous, or cloud. This classification becomes important because ISARA retrievals are performed for cloud-free cases. Ambient LWC and N_d are both derived from ambient particle size distribution measured by a Cloud and Aerosol Spectrometer (Droplet Measurement Technologies CAS; Baumgardner et al., 2001; Lance, 2012), a Fast Cloud Droplet Probe (SPEC FCDP; Kirschler et al., 2022), and a Cloud Droplet Probe (Droplet Measurement Technologies CDP; Sinclair et al., 2019). The CAS, CDP, and FCDP measure particles in the ambient D size ranges of 0.5–50 μm , 2–50 μm , and 3–50 μm , respectively. Measurements are considered cloud-free where LWC is less than 0.001 g m^{-3} , respectively (Schlosser et al., 2022). Because the CAS, CDP, and FCDP provide redundant measurements of LWC, this work relies on the CDP primarily and only uses FCDP for flights where the CDP was not being used. The CAS is only used as a backup for LWC in the case the CDP and the FCDP are unavailable.

What about clouds that have an IWC instead of LWC, and which have particles larger than 50 μm (as is usual for ice clouds)? how are you excluding those? perhaps you limit the measurements to $T > 0^\circ$? Please state in the paper how you avoid ice clouds.

Thank you for the comment. Ice is not a concern for two reasons. First, we limit the maximum aerosol diameter of the size distribution to 20 μm so these data will not influence our coarse aerosol measurement. Second, the ice particles would not survive isokinetic sampling. The following text has been added accordingly: “The sampling inlet flag also filters out ice clouds. Filtering out stray ice aerosol is not required as they will not survive the isokinetic sampling process.”.

- Line 345: “shortcoming”. why be so negative? you could say that it impacts the horizontal resolution without expressing a judgement

Thank you for the comment. This has been rephrased and the text is now as follows: “results in a spatial resolution of 8 km, which assumes a ground speed of 180 m s^{-1} ”.

- Line 347: “vertical profiles where the in-situ platform samples the column of air above an arbitrary ground point”. I doubt that the aircraft does “vertical profiles above an arbitrary ground point”. Typically they will be slant profiles or spirals. Please clarify

Thank you for the comment. While we feel this information would be distracting here, We added the following text to Sect. 3.5:

For these comparisons, the ISARA-derived products are acquired where the Falcon performed a controlled ascent or decent are identified using the leg identifier available as part of the ACTIVATE data set (Sorooshian et al., 2023). The points of interest are classified as part of a Falcon spiral, ascent, or decent. During the ascent and decent profiles, the Falcon would limit its airspeed to 180 m s^{-1} while performing a vertical profile while continuing to move in one direction. During the spiral profiles, the Falcon would spiral around a predetermined location while ascending or descending in addition to limiting its airspeed to 180 m s^{-1} .

Additionally We added “Further details on the Falcon’s vertical profile are discussed in Sect. 3.5” after the text associated with this comment.

- Line 368: “(8)”. this equation is not needed (it is the same as equation 3)

Thank you for the comment. We removed the redundant equations.

- Line 379: ~~channel~~ bin

Thank you for the comment. We replaced “channel” with “bin”.

- Line 380: “the average of all valid CRI values”. why not a weighted average, with the weight inversely proportional to the difference between C_{calc} and $C_{RH=40}$?

Thank you for the comment. Using a weighted average in this way assumes that the measured value has no measurement error. Additionally, we are limiting the number of changes made from Sawamura et al. (2017), which is the basis of this work. We added the following explanation to the text: “, which was shown to be effective for previous studies that this work expands upon (e.g., Sawamura et al., 2017)”.

- Line 390: “(9)”. equation not needed (like for equation 8)

Thank you for the comment. See the above response to the comment on line 368. No actions have been taken.

- Line 407: ~~RRI~~ CRI

Thank you for the comment. We replaced “RRI” with “CRI”.

- Line 421: “the smallest κ values are taken”. why the smallest? we need (1) to know the reason for taking the smallest, and (2) to be wary of the fact that the minimum and the maximum can be prone to outliers

Thank you for the comment. As discussed in the comment regarding CRI averaging, we are trying to limit the number of changes made. Currently we are focusing on the consideration of coarse-aerosol and providing the foundation for future work. In response to your second point, there are several levels of averaging due to the retrieval step needing to be performed at the SMPS resolution. As such, we feel the impacts of outliers have already been limited. No actions have been taken.

- Line 425: fine-mode

Thank you for the comment. We added the text “fine-mode”.

- Line 427: “low hygroscopicity”. how does this low hygroscopy match the assumption that they are sulphates?

Thank you for the comment. We agree this is not very representative of pure sulfate aerosol species but common for mixtures of organic aerosol species with sulfate aerosol species. We changed the text as follows to capture this information:

The commonly observed low absorption is expected given the frequency of sulfate and secondary organic aerosol Nakayama et al. (2015). That being said, the observed IRI is between 0.01 and 0.08 in 32% of the data, which indicates the presence of moderately absorbing aerosol species such as aged smoke and dust. The observed low hygroscopicity in many of the retrievals also indicate organic aerosol species are present (Petters and Kreidenweis, 2007). Only 20% of the data had $\kappa > 0.2$, which is the upper limit of the range of κ for organic aerosol species (Massoli et al., 2010).

- Lines 429-432:

The ISARA-derived IRI and κ are combined with the measured ambient RH and dry size distribution data and are used to calculate ambient scattering and absorption coefficients ($C_{scat,amb}$ and $C_{abs,amb}$, respectively) for the total (e.g., bulk) particle size-distribution ($0.003 \leq D \leq 20 \mu m$), the fine particle size-range ($0.1 \leq D \leq 1 \mu m$), the coarse particle size-range ($1 \leq D \leq 20 \mu m$), and the optically active particle size-range ($0.1 \leq D \leq 20 \mu m$).

Move the text at lines 455-465 to here so that we can understand how you account for the coarse mode

Thank you for the comment. We left this paragraph where it is, but it has shifted up to just after lines 429-432 as a result of moving the ambient aerosol equations to the appendix.

- 655 – Line 434: “(16)”. equation not needed like for eq. 8
Thank you for the comment. See the above response to the comment on line 368. No actions have been taken.
- 660 – Line 437: “the following equations”. put the equations on this page in a table
Thank you for the comment. We feel that a table is not ideal for displaying equations, but we can understand how this series of equations is distracting here. We added these to the appendix (Section A).
- 665 – Lines 476-477: “the shape of the size distribution is fixed and only the total number concentration is allowed to vary.”. this is very limited. I suggest to take a different approach, using e.g. a lognormal where you can vary the number of particles and their geometric mean radius, or even bimodals resulting from the sum of two lognormals. I feel that the approach taken here is too simple and therefore not fully persuading.
Thank you for the comment. We updated our synthetic analysis to include variability in the shape of the size distribution. We accomplished this by fitting the size-distributions that are observed in the ACTIVATE data and selecting the best fits from these to serve as the source of possible size-distributions that are available for the synthetic data generation. We updated all the sections related to the methods and results of the synthetic consistency analysis accordingly. We have extended the synthetic data analysis to include variability in the aerosol size distribution. We have extensively updated Sections 3.4 and 4.1 to reflect these improvements.
- 670 – Line 506: “note that 26.49% synthetically-generated measurements did not fall within appropriate delta thresholds”. explain why this happens
Thank you for the comment. This is directly as a result of measurement noise. We updated the subsequent sentence as follows for clarity: “To demonstrate the functionality of this analysis, the synthetic data generation and retrieval processes were repeated with zero measurement noise, which results in a rate of successful retrievals of 100%. Synthetic consistency analysis can be extended further to include nonspherical aerosols, aerosols without a constrained RRI, and increasing the number of successful retrievals under higher noise and lower signal conditions (e.g., lower aerosol concentrations, weakly scattering or weakly absorbing aerosol).”.
- 685 – Line 524: “LDR”. Specify if this is volume depolarisation ratio (a property of the atmospheric volume comprising molecules and particles) or particle depolarisation ratio (a property of the particles only). This is quite important to understand your quantitative threshold at 13%, because these quantities can be quite different. If the lidar data are simply labelled “LDR” then please contact the scientists that work with that instrument to answer the question, because we need to know to which of the two depolarisation ratios you refer. This is the same as comment number 25 of my first review, not correctly addressed.
Thank you for the comment. We addressed this topic in response to your major point 5 above. Most important to note again here is that we have clarified that we are using the aerosol LDR.
- 690 – Line 532: “the ISARA products data are weighted by extinction”. Please clarify if this is a vertical integration giving columnar quantities. In the present formulation, this is unclear.
Thank you for the comment. Your understanding is correct. We clarified this in the text as follows:
For the case of the HSRL-2-ISARA comparisons, this averaging results in altitude resolved properties through for a given column of air. For the RSP-ISARA comparisons, this results in a column-average property. Note that for AOD, the ISARA-derived ambient extinction is vertically integrated with sample altitude and does not account for the altitude bins above or below the sample altitudes.
- 700 – Line 545: “(30)”. Usually the standard deviation has $(n_p - 1)$ at the denominator, to denote the fact that the mean is not independent from the data. It is my impression that this should be the correct approach here too. Please see: https://en.wikipedia.org/wiki/Bessel%27s_correction
Thank you for the comment. We added the missing “- 1” from the standard deviation equations to use Bessel’s

correction. There have been some updates to the standard deviations presented, but no changes were significant enough to change the overall results.

- 705 – Line 562: “range-normalized root-mean square deviation”. Again you are using in this formula the extremes of a variables (max and min) which are prone to be influenced by outliers. I encourage you to explore a metric that is more robust, e.g. using percentiles. On using the word "range" please note that in remote sensing this indicates the distance from the instrument (e.g. the variable R in the lidar equation). On the other hand in statistics this indicates the difference between max and min. So in this case, to avoid ambiguity, perhaps you could omit the word "range" here.

710 We updated this text to “normalized root-mean square deviation”.

- 715 – Lines 579-580: “Future work could explore the impacts of adjusting the synthetic size distribution creation process to analyze the impact of low total concentration conditions.”. As already highlighted, I encourage the authors into doing a proper study of the synthetic consistency, making use of the full variability of inputs and using lognormals. Thank you for the comment. As discussed in two previous responses, we have extended the synthetic data analysis to include variability in the aerosol size distribution. We have extensively updated Sections 3.4 and 4.1 to reflect these improvements.

- 720 – Line 587: “-0.9”. -0.9 is a very large bias on the IRI which is considered in this paper < 0.08 . But if this is a relative bias then I’d like to see a % symbol immediately after the "0.9"

725 Thank you for the comment. To clarify, this was written in terms of percent. To ensure this is clear in the text, we reworked this notation throughout the text. We no longer use the $\pm$ notation except when discussing the standard error of the mean in the statistical analysis.

- 730 – Figure 6: “ISARA-retrieved versus synthetic”. use these terms for the graph axes instead of measured and calculated, which are less clear

Thank you for the comment. We replaced “calculated” with “ISARA-derived” and “measured” with “synthetic” in this figure.

- 735 – Line 609: ~~points that had the successful retrieval of CRI and 12319~~. given that there is the same number of points, 12319, there is no need for a repetition

Thank you for the comment. This was an error on our part. The values here have changed slightly as we adjusted how we adjust the dry scattering measurements. We updated the numbers in the text..

- 740 – Line 610 “successful retrieval”. It is unclear what is a successful retrieval. A retrieval can be successful in terms of identifying a solution (the algorithm converges) or because the solution is close enough to the observations (which is more stringent). Please clarify. This is the same as comment 26 of my first review, not satisfactorily addressed.

Thank you for the comment. In this work we define a successful retrieval as one that is close enough to the in-situ observations. We have gone through the text and clarified what a successful retrieval is defined as. Text has been added in the retrieval section and associated with this comment at line 610. The following text has been updated accordingly:

A final check is performed to establish a successful CRI retrieval. Here we ensure $\overline{\text{CRI}}$ results in scattering and absorption coefficients that meet the same thresholds of 20% and 1 Mm^{-1} , respectively.

745 “In this step, a successful κ retrieval is where a κ was found with a corresponding $\Delta C_{\text{scat,wet}} < 1\%$.” and “i.e., a solution was produced within the required retrieval thresholds”.

- Line 613: also
Thank you for the comment. We added “also”.

- 750 – Line 617: “of NRMSD”. add the reference to figure 7 here
Thank you for the comment. We added “ (see Figure 8)”.

- Line 621: “this bias increases with increasing wavelength,”. from the figure it looks like it increases with DE-
CREASING wavelength
755 Thank you for the comment. We agree this statement is incorrect and have removed it.

- Figure 7: “Figure”. use “ISARA-calculated” as x-axis label
Thank you for the comment. We replaced “calculated” with “ISARA-derived” in the figure x-axis label.

- 760 – Figure 8: “Figure”. 1) use “ISARA-calculated” as x-axis label and (2) pay attention to the location of the figure
titles. At the moment, “f(RH)” looks like the x-axis label for the top plot, whereas it is the title of the bottom plot.
Thank you for the comment. We replaced “calculated” with “ISARA-derived” in the figure x-axis label.

- Table 3: “Table”. add the location of each profile in the table. Could be a description such as “ocean, 1000 km East
765 of North Carolina”
Thank you for the comment. We have added the midpoint each of the Falcon’s profiles to this table. With the addi-
tion of Fig. 10, this should provide the information requested without adding more information to an already large
table.

- 770 – Table 3: “number of smoke counts”. explain what this is
Thank you for the comment. This is referring to the aerosol typing product derived from the HSRL-2. To clarify
this, we added “This study also uses the advanced aerosol typing product provided by the HSRL-2 team (Burton
et al., 2012, 2013). Specifically, we count the number of bins flagged as smoke in the HSRL-2 aerosol typing prod-
uct within each 225 m × 60 s bin. This HSRL-2-derived smoke count is used to analyze for the presence of elevated
775 smoke layers.” to Section 3.2. We also added “typing product” after HSRL-2 in this sentence.

- Table 3: “the RSP-derived total, fine-mode, and coarse AOD, the HSRL-2-derived AOD, in-situ-derived AOD,”
this description of Table 3 is incorrect
780 Thank you for catching this error. We replaced “the RSP-derived total, fine-mode, and coarse AOD, the HSRL-2-
derived AOD, in-situ-derived AOD,” with “HSRL-2-derived LDR”.

- Line 661: “Data from the 49 profiles”. I am getting confused now. It is unclear if you use 49 or 10 profiles. Please
make this clear in the text.
785 Thank you for the comment. The method here has changed from the previous revision. Our generalized study,
which has changed in title from the statistical analysis, focuses on the consistency within the subset of 10 profiles.
We also compare those consistency statistics to that observed in the entire set of 49 profiles. The text here is moved
to later in the text and changed to: “In this section we focus our analysis on the data from the 10 vertical profiles of
altitude-resolved ambient C_{ext} , C_{bse} , LDR, LR and N that had more than 3 points of comparison. For this section
we have applied the *a priori* assumption to the coarse aerosol as described in Sections 3.5 and 4.3. Additionally,
790 we compare the consistency results from this subset of 10 profile to the entire set of 49 profiles.”.

- Line 670: “These MRB indicate that the in-situ data data is biased low”. This low bias is very clearly seen in fig-
ures 9a,b,d,e and I don’t feel that simply using aggregate parameters such as MRB and r is sufficient to discuss this

evident low bias. I don't think that the authors can conclude that "the MB are fairly small" (line 676). Moreover, I'd expect the authors to play with the parameters that they can change (e.g. refractive index) to try and address this bias and correct the method. At this stage, from this plot I cannot agree that closure can be claimed between the remote sensing and the in-situ. This had been highlighted in my previous review (MAJOR points 6 and 8, not successfully addressed).

Thank you for the comment. We feel that the quantitative consistency statistics used in this study are useful for this type of analysis. While some trends seem obvious from scatterplots, the quantitative comparisons can often be surprising. That being said, we have rephrased our language throughout the manuscript to be more clear what these metrics are actually indicating. As was discussed in previous comments, our current revisions shows marked improvement in all consistency statistics used, which is largely due to adding in spheroidal approximation for the coarse aerosol shapes. We feel the improved correlations and reductions in biases do support the findings as they are stated in the current revision.

- Lines 676-677: “it is observed that the $MB \pm SB$ are fairly small and range from $1 \pm 11 \text{ Mm}^{-1}$ to $31 \pm 31 \text{ Mm}^{-1}$, which suggests that the MRB are partially inflated by low signal or noisy conditions”. ???

Thank you for the comment. We removed this text.

- Lines 682-683: “low correlation does not necessarily imply a lack of agreement.”. I am puzzled by this statement because figure 9h clearly shows that the LR calculation is not responding to the microphysical parameters (it is limited between 25 and 50). Perhaps the parameters used are not sufficient to model the LR. I would expect the authors to comment more critically on this point.

Thank you for the comment. We agree that the previous version of the manuscript showed very poor LR consistency. As discussed in several previous comments, our current analysis shows much better agreement in LR. We still offer a critical view on these comparisons and think there is room for improvement in LR and LDR; however, there are several limitations that include 1) the limited number of particle shapes available in our optical library, 2) the in-situ instrument suite does not sample coarse optical properties, and 3) sampling coincidence between can be difficult to get perfect and can always introduce uncertainty.

- Lines 699-701: “sensors is likely due to the loss of particles from the diameter cutoff of the inlet and through the in-situ sampling pathways as discussed in the Introduction and undersampling of the coarse aerosol particles by the CAS, CDP, and FCDP. Although in-situ values are lower than the HSRL-2 ones, reasonable agreement is evident by the $MB \pm SB$ ranges.”. as already discussed, I think that these statements are not supported by the scatter plots in figure 9

Thank you for the comment. We agree with this statement. In addition to improving these comparisons, we changed the text as follows to offer more explanations: “This discrepancy with the remote sensing retrievals is possibly due to difficulties in data coincidence, due to loss of aerosols from the diameter cutoff of the inlet and through the in-situ sampling pathways as discussed in the Introduction and undersampling of the coarse aerosol, or due morphologic and composition complexities.”.

- Line 722: over the Atlantic, XXX km East of the coast of North Carolina

Thank you for the comment. We added “On this day, a “unicorn” module was performed over the Atlantic that was ~450 km east of the coast of North Carolina.”.

- Line 733: likely to be

Thank you for the comment. We added “likely to be”.

- Figure 11: “Figure”. the case study figures 11 and 12 show a lot of promise and are more convincing than the overall external consistency statistics of section 3.1.3. I suggest to show the case study before section 3.1.3, so as

to persuade the reader about the method, and to be more critical in section 3.1.3 highlighting why the method is not universally as good as for the case study

Thank you for the comment. This change has been made. The sections have been reorganized to facilitate this change.

- Line 751: “the near surface aerosol are distinctly different”. we have no way to say that the aerosol type is different; only that the C_{ext}/C_{bsc} is larger.

Thank you for the comment. We do agree that the LR is larger. We changed the sentence as follows for clarity: “From the 1064 nm C_{ext} and the 355, 532, and 1064 nm C_{bsc} , it is possible that the near surface aerosol are likely different in size from the next highest data point (300 m).”.

- Lines 751-752: “is no available comparison with the ISARA-derived properties as the Falcon did not sample that low in the atmosphere.” This raises an important question: how do you account for layers not sampled < 238 m and > 4499 m, when doing the comparison with remote sensing columnar properties (e.g. AOD and variables in Table 7)? could you attempt using the lidar profile to quantify the correction for those layers that are not sampled in situ?

Thank you for the comment. As the goal of this work is to establish consistency rather than complete closure, we do not expect to retrieve AOD. We added the following text for clarity: “Note that for AOD, the ISARA-derived ambient extinction is vertically integrated with sample altitude and does not account for the altitude bins above or below the sample altitudes. This allows us to estimate the amount of the total aerosol column that was sampled by the Falcon.”

Response to Reviewer 4

We thank the reviewer for their feedback. We have taken care to respond directly to each comment. The text from your comments are shown in black, and our responses are shown in blue. The responses include the manuscript text that was changed, removed, or added.

The authors reported the ISARA retrieval framework in detail and attempted to validate it with ACTIVATE mission data. I appreciate the plethora of information re: the approach and the comparison datasets. However, I found the logic to be confusing at times despite the high level of details (see my specific comments and questions below). In addition, it would be helpful if the authors validated the retrieved aerosol properties (κ) with any available aerosol chemical composition from the same mission. Perhaps it is because the manuscript is rather long, there are quite a few typos and odd sentence structures. This manuscript would benefit from a major revision that I hope will address my questions and comments below.

Thank you for the comment. We have attempted to clarify and improve our logic throughout the text. We reorganized the paper to include a significant primer on aerosol properties. Additionally, we have extended our synthetic data analysis to account for variations in the size distribution. We have also reorganized the results to present the synthetic closure before the internal closure and the case studies to be before the generalized external closure. We also improved our consistency metrics throughout, which is primarily as a result of the efforts made to account for nonspherical coarse aerosol. In our extensive revision, we feel we have successfully addressed the majority of your comments. We would like to directly address why we are not analyzing data from the ACTIVATE AMS. That data would add even more length that you have stated is already a long paper. Additionally, there is currently a separate paper that is under review analyzing the differences between the different hygroscopicity parameterizations. The convenience of the $f(RH)$ derivation is that the retrieved hygroscopicity is internally consistent with the optical coefficient measurements.

General comment:

– For particle diameters $> 1 \mu\text{m}$, the authors should consider reporting them in μm instead of nm (e.g., 3488 nm on L281).
Thank you for the comment. We followed your suggestion and reported μm instead of nm for diameters above $1 \mu\text{m}$ throughout the text.

– The authors should consider updating the symbol of surface area concentration from $dA/d\log D$ to the conventionally used name of $dS/d\log D$.
Thank you for the comment. We replaced “ $dA/d\log D$ ” with “ $dS/d\log D$ ” throughout the text.

– In a few places (e.g., L427), the authors used the term “aerosol particles” but all aerosols are particles.
Thank you for the comment. An aerosol is defined as a mixture of a solid or liquid particle in a gas (see <https://glossary.ametsoc.org/wiki/Aerosol>). While it is common to use aerosol to refer to just the particles, this needs to be made clear within the context of our study. We initially refer to the particles as “aerosol particles” to provide context to our study, however refer to them as just aerosol in sections where particle is an ancillary word.

– There are a lot of repeated definitions of acronyms (e.g, HSRL-2 on L233, Table 2 title and content) that should be cleaned up. There are also instances of using acronyms first then defining them later (e.g., CRI appeared on L156 then defined on L177).
Thank you for the comment. In the case of CRI, it was defined in full on Line 100. We have removed the redundant definition on Line 177.

– Was there relevant aerosol chemical composition available from the ACTIVATE mission (e.g., AMS) that the authors could have used to compare with derived hygroscopicity parameter κ and real refractive index?
Thank you for the comment. There is AMS data available as part of the ACTIVATE dataset; however, that data would

- 910 add even more length to what you have stated is already a long paper. Additionally, there is currently a separate paper analyzing the differences between the different hygroscopicity parameterizations. The convenience of the $f(RH)$ derivation is that the retrieved hygroscopicity is internally consistent with the optical coefficient measurements.
- Specific comments/questions:
- 915 – L120-122: the authors already defined the diameter range for fine mode and coarse mode particles on L95-97, so remove them here.
Thank you for the comment. These repeated definitions have been removed.
- 920 – L166: define what phase function represents and why it is introduced
Thank you for the comment. We added the following: “Within the angular scattering matrix, the first element is the phase function (a_1), which describes the angular distribution of scattered light intensity by a particle. It is normalized over all directions so that it represents the probability of scattering into a given direction. In the context of this study, it defines the angular scattering characteristics relevant to the viewing geometries of the lidar, polarimeter, and in-situ instruments. Note that only a_1 is considered here, as polarization effects are beyond the scope of this work.”.
- 925 – L205-211: is it necessary to list out the dates of the ACTIVATE deployments instead of referencing the campaign overview by Sorooshian et al. (2023)? This information does not seem to be used later.
Thank you for the comment. We agree, this information is extraneous. We have removed this information.
- 930 – L239-247: a lot of repeated information on HSRL-2 products. Please rewrite this part of the paragraph to be more concise.
Thank you for the comment. This repeated information has been removed and the paragraphs have been combined.
- 935 – L274: extra words “the 30”
Thank you for the comment. The sentence now reads: “The LAS sampled aerosols that were actively dried with a 6” Monotube dryer (Perma-Pure, Model 700) for all flights except those conducted between 14 May and 30 June 2021, which were only dried passively.”.
- 940 – L275: what is “ram heating” in the context of SMPS data?
Thank you for the comment. This passive drying occurs as a result of the isokinetic inlet and the air increasing in speed as it is sampled through that inlet. To clarify this information, we updated the text as follows: “All SMPS data rely exclusively on passive drying from ram heating (i.e., air speeding up as it enters the aircraft inlet) and a generally warmer cabin temperature than ambient air. In addition to passive drying, the aerosols sampled by the LAS are actively dried with a 6” Monotube dryer (Perma-Pure, Model 700) for all flights except those conducted between 14 May and 30 June 945 2021, which were only dried passively.”.
- 950 – Table 1: in HSRL-2 section, missing a comma between “532” and “1064 nm”. Also, repetitive definitions of HSRL-2 and RSP in both the table title and the table content. Consider adding acronym of extinction coefficient, backscatter coefficient, etc. in parameter description for clarity.
Thank you for the comment. We removed “1064” as 1064 nm LDR is not an HSRL-2 product. We removed the repeated HSRL-2 and RSP definitions from the title but did not remove them from the table as otherwise there would be excess

white space and we feel it is not detracting from the table's information or purpose.

- L291: “success rate” of...what exactly? Please be specific. Is it in retrieving coarse-mode particles?

Thank you for the comment. We are referring to the success rate of ISARA for a given set of data. We have clarified this as follows “The successful retrieval rate (i.e., the success rate) of ISARA is defined as the number of points where the required fitting thresholds are met divided by the number of data points with measurements required to perform the retrieval ($\text{successrate} = \frac{\text{number of successful retrievals}}{\text{number of attempts}}$).”.

- L319: extra word “are”

Thank you, we have removed “are” from this sentence.

- L339: water vapor “density” or mixing ratio?

Thank you for the comment. You are correct, this should be mixing ratio. We replaced “density” with “mixing ratio”.

- L346: travel “up to” 8 km?

Thank you for the comment. We clarified the statement as follows: “The 45-second resolution results in a spatial resolution of 8 km, which assumes a ground speed of 180 m s^{-1} .”.

- L369-370: repeated definitions of no, ao, vo (previously defined on L190-197). This repeat of definitions also happened to other acronym and symbols (e.g., Table 2 title). Please re-check the other sections of the manuscript as well.

Thank you for the comment. We removed redundant definitions throughout the manuscript.

- L383: why did the authors restrict calculated absorption coefficients to within 1 Mm^{-1} of the measured values? Restricting the calculated values within a constant 1 Mm^{-1} may bias the filtering process to higher absorption values, where 1 Mm^{-1} difference is a smaller fraction of the total value (e.g., 10% of 10 Mm^{-1}) vs. smaller absorption values (e.g., 50% of 2 Mm^{-1}). Should this restriction be a percentage difference like with scattering?

Thank you for the comment. This process was originally developed by Sawamura et al. (2017), which is where these thresholds were first used. In the context of the current work we have already made several improvements over the methods developed by Sawamura et al. (2017), and want to limit the changes made in this iteration. As such, we feel it is appropriate to continue using their proven thresholds rather than investigate using our own. In future iterations we may be able to use optimal estimation, but it is outside of the scope of this study to investigate how that would change the results of the algorithm. Additionally, this change would only impact the internal consistency and would not have impact on the external consistency. We have made no changes based on this justification.

- L427: the authors stated that the derived hygroscopicity for aerosol in the ACTIVATE region is low (≤ 0.1) but also concluded that the aerosol population is sulfate-dominant. This value of κ seems really low. Organic is often assumed to have κ value of 0.1. Sulfate (inorganic) are usually with $\kappa > 0.4$. An aerosol population dominated by sulfate and organic as stated by the authors should have $\kappa \gg 0.1$.

Thank you for the comment. Echoing our comment response to RC3 regarding this topic, we agree this is not very representative of pure sulfate aerosol species but common for mixtures of organic aerosol species with sulfate aerosol species. We changed the text as follows to capture this information:

The commonly observed low absorption is expected given the frequency of sulfate and secondary organic aerosol Nakayama et al. (2015). That being said, the observed IRI is between 0.01 and 0.08 in 32% of the

995 data, which indicates the presence of moderately absorbing aerosol species such as aged smoke and dust. The observed low hygroscopicity in many of the retrievals also indicates organic aerosol species are present (Petters and Kreidenweis, 2007). Only 20% of the data had >0.2 , which is the upper limit of the range of for organic aerosol species (Massoli et al., 2010).

1000 – L432: misspelling with extra “i” in scattering
Thank you, we have fixed the spelling accordingly.

1005 – Figure 3: this does not seem like a correct representation of the calculation process. On L415-420, you mentioned that $C_{RH, meas}$ was actually calculated from $C_{dry, meas}$ and estimated gamma from Eq. 7. Really, the authors were using a single measured $C_{RH, meas(85\%)}$ and $C_{dry, meas}$ to estimate gamma. Then, they used gamma to derive $C_{RH, meas}$ at 80% - which now is an estimated value and not a measurement. Why didn’t the authors use $C_{RH, meas(85\%)}$ from the nephelometer instead?
Thank you for the comment. We do not use $C_{RH, meas(85\%)}$ directly as the RH associated with that measurement is not reported. To reduce the uncertainty with what exact RH to prescribe, we opted to used the parameterization to fix the RH to 80% for the retrievals. We feel no changes are needed.

1010 – L452: extra word “aerosol” or “particle”
Thank you for the comment. We removed “particle”.

1015 – L461-462: typo with “,” between number and μm
Thank you for catching this typo. It has been corrected.

1020 – L458-464: the use of CAS size distribution to derive coarse-mode particle properties contradict with L320-325 in section 2.3.2, where the authors said to use CDP mainly and CAS only if the other 2 datasets (CDP and FCDP) are not available. Please review and correct the appropriate section. Also, if the author chose CDP over CAS, what is the reasoning when CAS provides a wider range of available aerosol diameters?
Thank you for the comment. To clarify, we do favor the CAS for coarse-mode aerosol properties. For cloud filtering, we favor the CDP.

1025 – L506: the reported % does not need to be accurate to 2 decimal points (e.g., 26.49%), especially when the authors reported the biases as full % point (L587).
Thank you for the comment. While we do feel it is appropriate to truncate the bias estimate, we feel that significant figures are important when considering the length of the data set. In the updated version of this manuscript the number is now 9.26% this translates to the actual number of failures which is 919. If we truncate 9.26% to 9.3%, we get a 923 that is not the true number. We feel no changes are needed.

1030 – L515-517: why did the authors remove HSRL-2 AOD < 0.08 where HSRL-2 AOD uncertainty is 0.02 (stated in Table 1)?
Thank you for the comment. To clarify, The HSRL-2 lowest limit of detection is 0.08. The following text has been added: “, which is the detection limit for the HSRL-2”.

1035 – L521-524: I did not follow the logic. What is this 50% AOD difference criteria [between HSRL-2 and RSP derived AODs] based on? Similarly, how did the authors arrive at the filtering criteria of coarse-mode AOD being <0.1 ?
Thank you for the comment. This 50% AOD difference criteria criteria has been used in previous works (see Schlosser

et al., 2024), and limits the possibility that the HSRL-2 profile is not comparable to the closest RSP data point. In this case the data point is thrown out, which removes it from the data set. In response to the second part of your comment, in the most recent draft we removed the coarse-mode AOD threshold as it did not seem needed for this study which can benefit from having more coarse-mode aerosol comparisons. As such, the external consistency process has been updated.

- L555: extra “respectively”

Thank you for catching this error. We removed this extra instance of “respectively”.

- Figure 5: the authors could add lines that represent the range of measured data from Figure 1 on Figure 4 to show the range in synthetic data is smaller than the actual measurement (L578-579). Also, there are currently 2 panels (b) in this figure.

Thank you for the comment. We have added dashed magenta lines for the 10th and 90th percentiles from the ACTIVATE data set. We also corrected the panel labels.

- L587: did the authors mean “respectively” instead of “respectfully”?

Thank you for this comment. We replaced “respectfully” with “respectively”.

- Table 3: update the table title since it also includes information for Table 4. Also, the authors repeated the same information in the first paragraph immediately following Table 3.

Thank you for catching these errors. We updated the table’s title as follows:

Ancillary data for each of the 10 case studies. Ancillary information includes the case number, the profile start and stop times, the associated RSP sample time, the minimum and maximum altitudes sampled by the Falcon (i.e., in-situ) aircraft, the number of smoke counts above 2.5 km identified by the HSRL-2 typing product, the HSRL-2-derived LDR, the horizontal separation between the Falcon and King Air, and the latitude and longitude of the center point of the Falcon’s vertical profile. All dates and times are provided in coordinated universal time (UTC) and in the format “year-month-day” and “hour:minute:second”, respectively.

- L699: remote sensors or remote sensing retrievals?

Thank you for this comment. We changed “remote sensors” to “remote sensing retrievals” for clarity.

- Figure 11: why don’t we have $LR > 2$ km for HSRL-2 data? Both C_{ext} and C_{bsc} are available at these altitudes for HSRL-2 (panel b, e). Also, panel (i) seems to be missing from Figure 11.

Thank you for the comment. This is because LR is not reported at those low C_{ext} and C_{bsc} .

- L745: compared to case 7, case 9 with high smoke counts and high HSRL AOD would be interesting to investigate further compared to case 7.

Thank you for the comment. We have added a set of figures and tables for case 9 to contrast with case 7. We have also added corresponding results.

- Figure 12: the authors should add uncertainty to the calculated kappa, RRI, IRI that were derived with ISARA.

Thank you for the comment. We added error bars for the uncertainties determined in the synthetic consistency analysis.

The error bars associated with RRI are too small to be noticeable.

- Conclusion: I am not sure why the authors defined ISARA, CRI, etc. again when they had been mentioned all throughout the manuscript. Please remove this redundant information.

1085

Thank you for the comment. We removed the redundant information.

- L792: Similar to the work...

Thank you for catching this error. We corrected this.