Responses to Reviewers

We would like to thank the Reviewers for taking the time to read our manuscript and provide thorough feedback to help improve our study and the presentation of the results. Using your valuable feedback, we have made several substantial changes to our study to mitigate some sources of uncertainty, and we have updated the manuscript text to better convey the introduction and results of our study. In particular, we have standardized the duration of a given "scene" to be 10 minutes around a given dropsonde. Following this change, we removed the histogram of scene durations previously presented as Fig. 1 and have added a new Fig. 1 which shows relevant data from a single example scene. The Purpose and Background (section 1) has been revised to give a more-thorough overview of the prior literature and the motivation for our study, with information about specific environmental parameters moved to section 2. We now also utilize the 95th percentile, rather than the maximum, of each AMPR and APR-3 convective parameter in a given scene when comparing with the mean aerosol parameters to avoid potential influences from outlier values. We've also normalized all (modified) CAPE values by dropsonde altitude and presented this new variable as $NCAPE_{mod}$. This inclusion of the 95th percentile threshold, $NCAPE_{mod}$, and the standardized scene durations have significantly modified our results, which we present throughout sections 3–5.

Our responses to specific Reviewer comments can be found below. In the following discussion, Reviewer comments are presented in italicized font and our response immediately follows in standard font.

Reviewer #1:

Review of "Observed impacts of aerosol concentration on maritime tropical convection within constrained environments using airborne radiometer, radar, lidar, and dropsondes" by Amiot et al.

Recommendation: accept after revisions

Overall evaluation: This paper investigates links between various observed atmospheric parameters and the moist convection strength with the overall goal to understand aspects of the so-called convection invigoration. In other words, the motivation is to explore links between factors that in theory can affect moist convection and observed convection strength. The observations come from the CAMP2Ex experiment.

First, I have to say that I am not the right person to review this submission. Although I was involved over the last decade in the discussions of the invigoration conundrum (and for that reason I am signing my review), I feel someone with more expertise in atmospheric observations should also be involved. In particular, I feel the observations lack estimates of their uncertainty. Since this is not my area, I am not sure what to suggest. One possible suggestion is to use a spread of the observations near (in space and time) of the convective event. But this aspect can only address the sampling problem, that is, an uncertainty of a single observation. Another important aspect is an accuracy of the observation itself or uncertainty of a retrieval algorithm. I am not capable to assess if that aspect is appropriately addressed in the submission.

Below I discussed two major points concerning this submission, and follow with several specific comments that require authors' attention.

We sincerely thank you for taking the time to review our manuscript and provide helpful feedback. We have included more information about the uncertainties associated with each instrument utilized in our study on Lines 153–161.

Specific major comments.

I have two main issues concerning the motivation and interpretation of results. For the motivation, the convective invigoration discussed in the introduction is poorly explained. There are important published studies that discuss and criticize the original convective invigoration proposal of Andreae et al. (2004) and Resenfeld et al. (2008), including my own papers, that should be included in the introduction. The recent review paper by Varble et al. (ACP 2023, https://doi.org/10.5194/acp-23-13791-2023) should definitely be cited for that.

We have refined our literature review of convective invigoration from increased aerosol concentrations throughout section 1, including an additional paragraph starting on Line 82 to discuss studies that have presented counterarguments for this convective invigoration, especially above the environmental freezing level. This updated paragraph includes references to Grabowski and Morrison, (2016, 2017, 2020) and Varble et al., (2023).

For the interpretation, I strongly object the suggestion in the summary section 5 that the results support the notion of enhanced aerosol concentrations may invigorate convection (e.g., lines 501 and 510). The key problem is that the correlation does not imply causality. Is it possible that higher aerosol concentrations simply occur in conditions supporting stronger convection (e.g., higher CAPE)? The text in lines 512-515 seems to suggest that such a conclusion might be valid. One suggestion would be to explain the correlation versus causality conundrum in the introduction, and then try to use the results to shed some light on the problem.

We have added a sentence to section 1 (Lines 134–136) to emphasize that correlation does not imply causality, and we now refer back to this point throughout the manuscript. We have also pulled back somewhat from asserting such conclusions in the summary section, and we now briefly mention these potential implications but focus more on the correlations as indicators of potentially interesting trends that are worth examining in future work, rather than providing definitive conclusions on their own. We have highlighted some of the most-interesting trends in the scatterplots more clearly in the updated summary as well. You are correct that, despite our environmental stratifications, there is still a chance that there is a convective dependence on the values within each group; e.g., “low high” CAPE versus “high high” CAPE for the high CAPE group. We acknowledge that correlation does not indicate causality as noted in the manuscript, and this is meant to inspire additional work rather than be a conclusive/final study.

However, I have to admit that the discussion of the results in sections 3 and 4 are difficult for me to follow. Specifically, I do not see any trends in figs. 3, 6 and 9, just scattered data points. Does that suggest that the overall outcome of the study is inconclusive? Is there a better way to present the results? See 7 below.

We have included additional emphasis in our updated manuscript that correlation does not indicate causation, and thus we aren’t able to make definitive conclusions about aerosol impacts on convection solely based on our results. However, the conclusions we present in the updated manuscript may inform future studies about potentially interesting relationships between a wide range of convective and aerosol parameters that may be worthy of further investigation. There were indeed several comparisons that yielded little to no correlation, but others provided statistically significant correlations (though, many were hindered by a relatively small sample size, and some trends did not stand out very clearly, as you note). Please see our response to your comment #7 below regarding the presentation of the results.

Specific comments.

1. The paragraph starting in l. 68 and the reference to Mulholland et al (2021) in particular. The argument here is wrong as discussed in Grabowski (QJRMS 2023), see the summary Fig. 17 there in particular. Please revise keeping in mind that the argument in my paper applies to convective boundary layer over land. Such an argument might not always apply to oceanic convection with weak surface buoyancy forcing. The above comment is also relevant to the discussion around l. 156. Specifically, it is unclear to me why higher LCL may favor stronger convection.

We have modified our discussion in this paragraph (now starting on Line 115) to note that the conclusions of Mulholland et al., (2021) have been debated in subsequent literature (specifically citing Grabowski, 2023). We still examine LCL altitude in this paper to see what correlations result between the convective and aerosol parameters when binning the environments according to their LCL altitude (now with additional acknowledgements of the shortcomings associated with environmental binning, as discussed in Varble et al., 2023), but we note on Lines 131–132. that the exact impact of LCL altitude on convective intensity is still up for debate.

2. L. 156: It is unclear what you mean by "frequency" here and in other places (e.g., l. 214). Do you mean higher cloud cover? Please explain or remove to focus on the convective intensity alone.

You are correct; "frequency" was meant to refer to the relative areal coverage of convective storms during a given flight scene. We have modified this wording to "prevalence" throughout the manuscript to avoid confusion, especially since we use "frequency" for other contexts (e.g., e.g., when discussing microwave frequencies).

3. L. 109 and later in the text. I was confused by units of CLW and I though it should be kg m^{-3} (i.e., content). That was until I realized that this is the CLW path. I suggest to use CLWP throughout the manuscript to avoid confusion.

We have retained the use of CLW to refer to integrated cloud liquid water path in the updated manuscript to maintain consistency with past AMPR-related literature, but we acknowledge the differences between CLW and CLWP and have now fully spelled out our definition of CLW on Line 120 as integrated cloud liquid water path.

4. The discussion in paragraph starting in l. 119. Please see me major comment above. Perhaps referring here to postulated "warm" and "cold" invigorations would be appropriate here. However, see the discussion in section 2 in Grabowski and Morrison (JAS 2020, p. 2567) and the review of Varble et al. (2023) already mentioned above.

We have now included a discussion of the differences between convective invigoration from enhanced aerosol concentrations in warm-phase and cold-phase regions in the new paragraph that begins on Line 82.

5. L. 186 versus l. 63. Overall, CAPE refers to the buoyancy integral from the level of free convection to the equilibrium level as given by (1). Integrating buoyancy to the aircraft altitude as in (4) gives only a fraction of CAPE as pointed out in the submission. I feel the authors may want to refer to what I call "cumulative CAPE" (cCAPE), that is, how CAPE builds up in an adiabatic parcel as the parcel rises through the atmosphere. I feel this is a useful concept as shown, for instance, in Thomas et al. (ACP 2018, Fig. 4 in particular) and in some of my papers concerning convective dynamics (e.g., Grabowski and Morrison, ACP 2021, see Fig. 3 there). Clearly cCAPE depends on the aircraft altitude, so it is unclear how important that aspect is for the analysis. Can some large-scale analysis be used to extend calculations above the aircraft altitude to get the total CAPE?

We have reevaluated our computation of "CAPE" in the revised manuscript given the concerns raised by all Reviewers. However, rather than focusing on modified CAPE or cumulative CAPE, though the cumulative CAPE is certainly an interesting idea, we have normalized (modified) CAPE by the dropsonde launch altitude, similar to the methods of Blanchard, (1998). This parameter, now labeled as $NCAPE_{mod}$ in the manuscript, is computed as

$$\mathrm{NCAPE_{mod}}\ (\mathrm{m\ s^{-2}}) = \frac{\mathrm{CAPE_{mod}}}{\mathrm{z}},$$

where z is the dropsonde launch altitude and

$$\mathrm{CAPE_{mod}}\left(\mathrm{J\ kg^{-1}}\right) = g\int_{z_{lfc}}^{z_{P3}} \frac{\left(T_v - T_{v,0}\right)}{T_{v,0}}\ dz,$$

which was originally used as the variable "CAPE" in the previous version of the manuscript, with the "mod" subscript added to distinguish it from true CAPE. The use of $NCAPE_{mod}$ allows for the vertical acceleration of the parcel to be evaluated directly, as discussed in Blanchard, (1998), and helps mitigate some of the influence of the varying dropsonde launch (i.e., P-3) altitude. This information has been added to the manuscript in section 2, and the updated results are presented throughout sections 3 and 4

6. Table 1 should include units for all symbols.

Added as suggested.

7. Figs. 5. 7, and 8. These are not figures, these are tables. It is difficult for me to draw any conclusions looking at them. Can they be shown as bar diagrams? I think the authors need to think about a better way to show those key results.

We refer to these as figures since: 1) they utilize color shading from a colorbar throughout, and 2) they have multiple horizontal lines to create the grid structure, both of which are specifically mentioned in ACP's guidelines to be avoided for final published tables (https://www.atmospheric-chemistry-and-physics.net/submission.html#figurestables). Since the color shading and grid pattern are essential for Figs. 2, 4, 5, 7, and 8, we classify them as figures.

In regard to the key results, we've considered several possible methods for presenting the results as concisely as possible, especially since there is a lot of information to include. For example, in Fig. 2, we include every comparison of AMPR CLW with the aerosol conditions across all stratified environments in this study, along with Pearson correlation coefficients, p-values, and sample size. We feel each of these is important to present to the reader to convey the full story of the study. Expanding to other variables and sensitivity tests creates even more of the colored figures (e.g., as in the supplemental material). We have attempted to highlight the key results in these figures using bolded text for the highest Pearson correlation coefficient magnitudes and strongest color shading for the most-significant p-values with the hope that it these would draw the reader's attention to the key relationships between the variables. To highlight some of the key results further, we employ the scatterplots and draw connections between the color figures and the scatterplots as much as possible. Utilizing other plot types (e.g., bar diagrams) without Figs. 2, 4, 5, 7, and 8 would limit the presentation of the full results.

Reviewer #2:

Review of Observed impacts of aerosol concentration on maritime tropical convection within constrained environments using airborne radiometer, radar, lidar, and dropsondes by Amiot et al. (2024).

Amiot et al. investigate the correlations between a number of aerosol, convective and environmental parameters observed on research flights during the CAMP2Ex campaign. They hypothesize that an increase in values of aerosol parameters leads to an increase of values of convective parameters under similar environmental conditions. I believe that the foundation of this manuscript has merit and would be a good addition to the convective invigoration literature. Especially interesting is the analysis of a large array of environmental and aerosol parameters and how they are correlated with convective metrics. However, I have some major concerns about the methodological approach and the structure of the manuscript that must be addressed. Given these concerns I cannot recommend this manuscript for publication at this time. I believe that addressing my concerns is likely to require a substantial amount of time exceeding that typical of a major revision. However, if the authors do address these concerns, and decide to resubmit, I would be happy to review a resubmission.

The introduction and methods sections were honestly quite confusing and did not serve as good introduction to the science at hand or provide enough information to build trust in the data. In its current state this manuscript is of little value to anyone not directly involved with CAMP2Ex. While the results should be reproducible with the descriptions in the manuscript, I am not sure whether the methodological approach provides robust results. It seems that some of these issues might be due this manuscript being mostly the same as Chapter 4 in Amiot (2023), with little changes to make it suitable for journal publication. I.e., things that might have been described in preceding chapters of Amiot (2023) are missing from this manuscript.

Below I list comments with some suggestions that the authors might want to consider to improve the manuscript.

We sincerely thank you for taking the time to review our manuscript and provide helpful feedback.

Major Comments

1. The introduction needs to be reworked. For instance, the three paragraphs from ll. 51-110 do not really belong in the introduction and do not introduce the reader to the general scientific topic. I suggest that the authors move this description to the methods section.

We have moved these paragraphs to the Data and Methods section as suggested, and we have included additional information throughout the Purpose and Background section to bolster it.

Furthermore, the literature review is very brief. This study tries to 'contribute knowledge to long-standing questions of aerosol influences on convection' through concerns modeling and theoretical work. I suggest that the authors include more observational studies in their introduction. Some examples are: Lin et al. (2006), Fan et al. (2018), Veals et al. (2022), and Zang et al. (2023).

We have added these suggested references to the manuscript and referred to them in a new paragraph that we have added to the Purpose and Background section starting on Line 76.

2. Currently, the manuscript lacks any information about the CAMP2Ex campaign. It is not even mentioned when the campaign took place or what type of convection is investigated ('maritime tropical convection' is very broad)! The authors should not expect every reader to be familiar with this campaign. I suggest starting section 2 with a general description of the CAMP2Ex campaign. This description does not need to be overly long since the campaign has been described in detail in the referenced literature, but some basic description is needed. When and where did it take place? Some details about the P3, the instruments and the SFs. What were the meteorological conditions, are they comparable between the analyzed SFs? What types of clouds are we looking at? Why are SF 1-4 not included? These are all very basic questions that are needed to fully understand this manuscript but are not addressed at all.

These omissions were definitely in error and occurred during the transition of Chapter 4 in Amiot (2023) to this manuscript, as you noted in one of your previous paragraphs. We have included additional details about the CAMP^{2}Ex campaign on Lines 44–52, including descriptions of when the campaign took place, the types of clouds investigated, and the instruments that participated in the campaign. Details regarding the exclusion of SFs 1–4 have been added on Lines 256–258.

3. A possible major concern with the data that needs to be addressed is the different scene length. Figure 1 suggests a bimodal distribution of scene lengths with 'short' scenes under 5 min long and longer scenes around 10 min long. Assuming that the P3 flies completely straight in each scene and at similar speed (impossible to know see comment 2), the longer scenes inevitably capture more meteorological variability than the short scenes. Any potential issues with differing scene times is not acknowledged until l. 452. This could lead to major sampling issues. For instance, the convective parameters could be representative of a cloud that is at one end of the scene, but the aerosol parameters could be significantly influenced by data at the other end of the scene. In such a case the scene's aerosol values do not represent the actual aerosol conditions at the location of the convection. Such issues become increasingly likely with longer scenes. With the authors providing no example data from a scene it is impossible to determine if this actually was an issue. I suggest that the authors investigate the sensitivity of their results to limiting the scene lengths to different times. For instance, for all scenes longer than 5 minutes, only the 5 minutes around the dropsonde are used (5 min serves as an example, I *cannot reasonably say whether this is a good threshold). Or the authors could set a maximum horizontal length for scenes appropriate for the convection being analyzed. Alternatively, the authors should demonstrate that different scene lengths do not impact the results. See the minor comment below on including a figure of a 'scene' that might help with this issue as well.*

We fully agree about the benefits of using standardized scene durations and acknowledge the shortcomings of using varying scene durations. As a result, we have now standardized all scene durations to be 10 minutes long, ±5 minutes from the associated dropsonde launch time. This allows time for ample observations around the dropsonde location and gives time for the dropsonde to descend (most CAMP^{2}Ex dropsonde descent durations were 5–10 minutes; Vömel et al., 2020). Since these new standardized scene durations involve the combination of multiple APR-3 scan files in many instances, there were some cases where data were unavailable at the start and/or end of a given scene (e.g., P-3 was in a turn). To accommodate some of these cases where reliable data became available shortly before and/or after the ±5-minute window, we have allowed up to 1-minute (i.e., 10%) grace in the scene duration. That is, any scene durations < 9 minutes or > 11 minutes were discarded from the analysis, which amounted to 47 of the 144 dropsondes previously considered. We have included a description of this uncertainty on Lines 259–263. When combined with five additional dropsondes excluded due to potential cloud contamination (please see our response on page 12), our new dropsonde sample size is 92.

Your comment about the possibility for convective metrics to be influenced by one end of the scene while aerosol metrics are influenced by the other end of the scene are possible, and we have not specifically examined spatial offsets between the convective and aerosol metrics in our scenes. However, we attempt to mitigate these effects through the use of a scene-averaged mean value for each aerosol parameter, rather than relying on maximum values, and we now use the 95th percentile of the convective metrics, following your recommendation in a different comment. As you've correctly stated, the potential impact of these spatial offsets would increase with increasing scene duration, and our new standardization of the scene durations helps remove the longest scenes, where were previously ~18 minutes long (albeit, while also increasing the duration of shorter scenes that were previously ~2 minutes long, for which the effects of spatial offsets were relatively limited).

4. During CAMP2Ex a wide range of shallow to deep convection was observed. Aerosol impacts in these regimes might be vastly different. There is no description of a separation between these regimes. Are the stratified environmental parameters a good predictor of the type of convection? I understand that sample sizes are already small, however, I believe the authors should at least look at their results separated by different types of convection. Whether this should be included in the manuscript depends on the results.

We agree that such an analysis would be beneficial, but we have kept our focus on the campaign-wide statistics for the time being, largely due to the limited sample size as you note (which was further reduced from 144 scenes to 92 scenes in our updated manuscript). However, we have noted the different types of clouds observed during CAMP^{2}Ex on Lines 47–48 and added this separation as future work on Lines 595–596.

5. ll. 152-153: This first hypothesis does not seem to be directly addressed anywhere. There is only some description based on the stratification in the scatter plots but at best this is a weak indicator of correlations. I think it would be beneficial if the authors test how the chosen environmental parameters correlate with the chosen convective (aerosol) parameters for their data set as was promised by this hypothesis.

We agree that this type of analysis would be interesting and beneficial, but our focus in this study was on the convective-aerosol correlations within the stratified environmental groups, along with some (secondary) discussion of the observed increase or decrease in convective intensity/prevalence associated with the high, medium, and low groups for each environmental parameter in the scatterplots. We have updated the wording around Lines 127–128 and 133–134 in the updated manuscript to convey this more clearly.

6. Based on some of the comments that I have classified as minor below and my concerns about the methodological approach, I believe some of the conclusions are only weakly supported by the results. Some conclusions are very speculative based on just a few data points. This issue could be overcome by focusing on specific types of convection and using a more standardized definition of a scene. Currently, it seems that it is quite possible that outliers on which the authors base some conclusions are at least partially due to different scene length. Furthermore, the manuscript could benefit from being more focused on which specific environmental and aerosol parameters are actually good predictors of the convective intensity/frequency. The most we learn is that some convective parameters correlate with some aerosol parameters, for lapse rate and K-index stratifications. I suggest expanding these types of analysis.

In our updated manuscript, we have placed less emphasis on the outlier values, instead focusing on more-general trends and discussing the influences outliers may have on some of the results, especially those associated with a relatively small sample size. While we have standardized the scene durations in the updated manuscript, which helps to mitigate some of the uncertainty, we have still shifted the focus away from the outlier values compared to the previous version of the manuscript. Our main goal was to begin with this investigation and all results of our analyses are included for the interested reader. However, we agree that expanded analyses of key environmental parameters would be beneficial, and we have listed this as an avenue for future work.

7. Are the authors employing the sensitivity tests of their environmental binning in response to the concerns raised about environmental binning techniques in Varble et al. (2023)? The authors might want to make this connection since at least some of the results appear to be robust across the sensitivity tests. (And those that are not highlight the importance of such sensitivity tests!)

The initial idea for performing the sensitivity tests was developed in 2022 [i.e., prior to the publication of Varble et al., (2023)], but we agree that the ideas raised in Varble et al., (2023) should be referenced in our manuscript and the results herein should be connected to their study. We have added references to Varble et al., (2023) throughout the updated manuscript.

Minor comments

8. In the short summary, the terms 'microwave-frequency' and 'K-index' are too technical for the more general audience. I suggest removing 'microwave-frequency' altogether and replacing 'K-index' with something like 'convective potential'.

Modified as suggested.

9. l. 82: The authors acknowledge that convective intensity refers to peak updraft velocity. They should also acknowledge that the convective parameters they analyze (in particular reflectivity) can have changes unrelated to convective intensity as described in Varble et al. (2023).

This is an important point to consider. We have added a note on Lines 136–138 about how the radar- and radiometer-based convective metrics may vary due to factors not specifically tied to peak updraft intensity.

10. Does Figure 1 show the length of all scenes during all flights? The number of occurrences adds up to significantly more than 144. Even counting the bars conservatively, there seem to be more than 300 scene lengths in this figure. Why do the authors include the length of scenes that are not analyzed? I suggest reducing the data in this figure to the 144 analyzed scenes. If the authors standardize their scenes this figure could be removed altogether.

This was the result of a coding typo in our previous draft, but we have removed this figure from the updated manuscript due to the standardization of scene times in the new results, following your recommendation.

11. I suggest a figure displaying a 'scene'. This could help addressing some of the major comments. There are a lot of measurements that could be shown in such a figure, and they cannot all reasonably be included. However, I suggest time-height series (curtains) of reflectivity of at least one APR-3 frequency (at nadir, see Fig. 7 in Reid et al., 2023), time-height series of HSRL2 backscatter for one wavelength, and some timeseries from Table 1. Maybe a dropsonde profile could be included as well (see Fig. 7 in Reid et al., 2023), with corresponding values for the environmental parameters.

We have developed a new figure illustrating a single 10-minute "scene," which is presented as the new Fig. 1 in the manuscript. Since our study focuses on APR-3 composite Z_H throughout a given scene, we have presented Ku-band composite Z_H in Fig. 1 rather than a time-height series of reflectivity. Likewise, we have included AMPR CLW values from each AMPR pixel throughout the scene. Following your suggestion, we included the corresponding dropsonde profile and time-height series of HSRL2 532-nm backscatter for this scene.

12. Was any testing done whether dropsondes passed through clouds? Such dropsondes are unlikely to be representative of the environmental background and should be used with caution when determining convective parameters.

We had not screened dropsondes for the presence of clouds in our prior submission, but we have done so for the updated manuscript. Since the uncertainty in a given AVAPS relative humidity (RH) measurement is approximately 3% (Freeman et al., 2020), we examined the presence of RH values > 97% in each profile. While some dropsondes did indeed pass through clouds, most of them were relatively brief (e.g., < 5% of the total sounding). However, we identified five dropsondes where more than 20% of the dropsonde profile occurred in-cloud, and we have removed these from our updated analyses to avoid their potential contamination of the results, as you note. We have added a sentence on Lines 151–154 explaining this. When combined with 47 additional dropsondes excluded due to limitations with defining a 10-minute scene around them (please see our response on page 9), our new dropsonde sample size is 92.

13. The authors acknowledge the inconsistency of 'modified CAPE' due to dropsondes launching at different altitudes, but do not seem to address this issue afterwards. Currently, it seems that CAPE would most strongly correlate with launch altitude and not convective parameters. The authors could consider using (modified) normalized CAPE instead, i.e. normalize CAPE by the altitude of the P3/dropsonde launch.

We have reevaluated our computation of "CAPE" in the revised manuscript given the concerns raised by all Reviewers. Following your suggestion, we have normalized (modified) CAPE by the dropsonde launch altitude, similar to the methods of Blanchard, (1998). This parameter, now labeled as $NCAPE_{mod}$ in the manuscript, is computed as

$$\mathrm{NCAPE_{mod}}\ (\mathrm{m\ s^{-2}}) = \frac{\mathrm{CAPE_{mod}}}{\mathrm{z}},$$

where z is the dropsonde launch altitude and

$$\mathrm{CAPE_{mod}}\ \left(\mathrm{J\ kg^{-1}}\right) = g\int_{z_{lfc}}^{z_{P3}} \frac{\left(T_v - T_{v,0}\right)}{T_{v,0}}\ dz,$$

which was originally used as the variable "CAPE" in the previous version of the manuscript, with the "mod" subscript added to distinguish it from true CAPE. The use of $NCAPE_{mod}$ allows for the vertical acceleration of the parcel to be evaluated directly, as discussed in Blanchard, (1998), and helps mitigate some of the influence of the varying dropsonde launch (i.e., P-3) altitude. This information has been added to the manuscript in section 2, and the updated results are presented throughout sections 3 and 4.

14. ll. 187-188: 'would likely be less than true CAPE', remove 'likely', since it would always be less than true CAPE , if 'the P3 did not fly above the EL during any science flight'.

Modified as suggested.

15. ll. 188-190: Without any description of the meteorology and cloud environment, it is impossible to understand the importance of this statement.

We have added additional descriptions of the cloud environments on Lines 47–48.

16. ll. 193,194,250 etc.: I personally would remove any mention of the Python packages from the main body of the manuscript. This might confuse some readers not familiar with Python (e.g., what is 'np.percentile'?). Instead, I suggest mentioning these packages and what they calculate in the data availability statement. The authors should then also mention what they use to calculate Pearson correlation coefficients and p-values.

The references to specific Python packages were moved to the data availability statement as suggested, and details were added therein regarding the calculations of Pearson correlation coefficients and the associated p-values.

17. Table 1: Please add units.

Added as suggested

18. ll. 217-220: '[…] due to their direct association with peak convective intensity' please add a reference for this statement.

We have added a reference to Kollias et al., (2001) on Line 273.

19. ll. 217: In my experience, maximum values in such observations can potentially represent significant outliers. I suggest testing the sensitivity of the results to high percentile values (e.g., 99th and/or 95th) of these observations or demonstrating that maximum values do represent the actual environment well.

We have updated our analysis to use the 95th percentile of the AMPR and APR-3 convective metrics from a given scene, rather than using the maximum value.

20. From my understanding the parameter PixelsKu should strongly correlate with scene length because longer scenes would inevitably contain more such pixels. This is later acknowledged in l. 452. Based on this understanding of PixelsKu, the analysis of PixelsKu appears to me to be largely meaningless. I suggest trying to normalize this parameter by scene length or standardizing scene length altogether as described above.

You are correct about the direct impact scene duration would have on $Pixels_{Ku}$. We have mitigated this effect with the new standardized scene durations of 10 minutes each (please see our response on page 9).

21. Figures 3, 6, and 9 subplots should be labeled with (a), (b) etc. and referred to as such in the text. Furthermore, I suggest increasing the label sizes in these figures and adding gridlines. The gridlines should make it easier to see which data points are actually within the ranges described in the text.

All modified as suggested.

22. Did the authors consider the measurement uncertainties when determining the correlation coefficients? Or are measurement uncertainties so small that they are of no concern? If not, the authors could use Monte Carlo type simulations to estimate the uncertainty of their correlation coefficients. I.e., simulate many random instances of the data set based on the uncertainties and determine a mean and variance of the correlation coefficient for all parameters.

The measurement and/or retrieval uncertainties in the AMPR, APR-3, and HSRL2 (i.e., the instruments considered in the correlation coefficient calculations) data were deemed negligible for the purposes of this study. We have listed the uncertainty values for each of these instruments, along with AVAPS, on Line 155–161 in the updated manuscript and noted how they are relatively small.

23. Maybe this is a misunderstanding on my part: In Figure 3, I counted the number of data points. They do not match what is shown in the corresponding figures (Figure 2 for the top row). For instance, in Figure 2 it says there should be 16 data points for medium K-index when comparing maximum AMPR CLW vs. 532 nm AOT. However, in Figure 3 top right I only count 13 data points for medium K-index (green). What is the reason for this discrepancy? This seems to be happening in almost all scatter plots. Are some data points hidden below others (although that seems unlikely)? It does not appear that ll. 240-244 explain this.

This may have been caused by an additional level of NaN-data masking that was present in the scatterplots but not the correlation figures. Based on a visual check, this discrepancy is not present in the updated manuscript draft (i.e., the parenthesized values in the correlation figures match the number of data points in the corresponding scatterplots).

24. l. 339: Figure 4 shows a low correlation coefficient and high p-value for medium LR850-500 and PCT19 vs. Bsc532. This statement seems to be incorrect. Figure 3 bottom right confirms that there is no correlation. Do the authors mean high instead of 'medium'?

Yes, thank you for catching this. The previous statement was mistakenly based on Bsc_{355}, rather than Bsc_{532}, versus PCT_{19} when binned by $LR_{850\text{-}500}$. However, this statement is no longer present in the updated discussion.

25. l. 340: There are only six data points in the described area and one of those is associated with low values. I personally would not draw the subsequent conclusion from such few data points. If one randomly draws 6 data points from 37 low, 44 medium, and 40 high values (as shown in Figure 4, although there are fewer points in figure 3 bottom left, see comment 23), the chance of there being exactly one low data point in those 6 is about 30%! Thus, it is not too unlikely that the shown behavior happened by pure chance.

The phrasing "$Bsc_{532} > 2$ Mm^{-1} sr^{-1} or $PCT_{19} > 240$ K" was meant to refer to all data points that met either condition, not necessarily both. That is, 39 data points were within the described region of the plot, which did indeed differ from Fig. 4 as you noted and as discussed in our response to comment 23. Of those 39 data points, 32 of them were associated with medium or high $LR_{850\text{-}700}$, which was the reason behind our original statement that a "vast majority" of the data points were associated with medium or high $LR_{850\text{-}700}$. However, as discussed in our response to comment 23, we have clarified the discrepancy between the values in Figs. 3 and 4 (along with all other similar figure pairs within the manuscript) and all results have been updated.

26. ll. 353-354: This sentence is incorrect because there is no correlation for the medium and high bins, i.e. there is no 'increase in PCT19 with increasing Bsc532'. The sentence afterwards describes better what is actually going on.

This statement is no longer present in the updated discussion.

27. Figure 6 bottom row: the correlations for medium and high environmental parameters appear to be high because of single outlier values. The authors might want to consider adding the actual regression lines as shown in Chapter 4 of Amiot (2023).

We have added these regression lines to Figs. 3, 6, and 9 and have updated their associated discussions.

28. ll. 439: Again, could this not just be because it is a longer scene, saying nothing about the actual convective frequency?

You are correct. The frequency of convection could be directly related to scene duration, which is a main reason why we now standardize scene duration (please see our response on page 9).

29. ll. 448-451: but there is also two data points with low-to-medium K-index and low backscatter but very high convective frequency. These data points do not support the conclusion in this sentence.

This statement is no longer present in the updated discussion.

Typographical:

30. l. 111: 'have been'
31. l. 137: 'that was deployed on'
32. l. 157: I suggest removing 'radiometer-retrieved'
33. l. 171: QC (quality control?) has not been defined.
34. l. 214: 'Nine remote-sensing parameters related to convective intensity', I believe that this should be 'eight' instead of 'nine' as shown in table 1 and mentioned elsewhere in the text.
35. l. 292: 'increase'
36. l. 293: 'is associated with', 'in association with' or 'associated with'?
37. ll. 292-293: something else seems wrong about this sentence, please correct.

All modified as suggested. Thank you for catching these!