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.

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.

   Furthermore, the literature review is very brief. This study tries to 'contribute knowledge to long-standing questions of aerosol influences on convection' through the analysis of observations. However, the referenced literature almost exclusively 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).

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.

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.

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.

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.

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.

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!)

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'.

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).

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.

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.

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.

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.

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’.

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

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.

17. Table 1: Please add units.

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

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., 99^{th} and/or 95^{th}) of these observations or demonstrating that maximum values do represent the actual environment well.

20. From my understanding the parameter $Pixels_{Ku}$ 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 $Pixels_{Ku}$, the analysis of $Pixels_{Ku}$ 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.

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.

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.

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.

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'?

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.

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.

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).

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

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.

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.

References

- Amiot, C. G.: Airborne passive microwave geophysical retrievals and applications in assessing environmental and aerosol impacts on maritime convection. Ph.D. dissertation, Dept. of Atmospheric and Earth Science, The University of Alabama in Huntsville, Huntsville, AL, 176 pp, https://louis.uah.edu/uah-dissertations/278/, 2023
- Fan J, Rosenfeld D, Zhang Y, Giangrande SE, Li Z, Machado LAT, Martin ST, Yang Y, Wang J, Artaxo P, Barbosa HMJ, Braga RC, Comstock JM, Feng Z, Gao W, Gomes HB,

Mei F, Pöhlker C, Pöhlker ML, Pöschl U, de Souza RAF. Substantial convection and precipitation enhancements by ultrafine aerosol particles. Science. 2018 Jan 26;359(6374):411-418. doi: 10.1126/science.aan8461

- Lin, J. C., T. Matsui, R. A. Pielke Sr., and C. Kummerow (2006), Effects of biomass-burning-derived aerosols on precipitation and clouds in the Amazon Basin: a satellite-based empirical study, *J. Geophys. Res.*, 111, D19204, doi:10.1029/2005JD006884
- Reid, J. S., and Coauthors: The coupling between tropical meteorology, aerosol lifecycle, convection and the energy budget: The Cloud, Aerosol and Monsoon Processes Philippines Experiment (CAMP2Ex). B. Am. Meteorol. Soc., 106, E1179–E1205, https://doi.org/10.1175/BAMS-D-21-0285.1, 2023
- Varble, A. C., Igel, A. L., Morrison, H., Grabowski, W. W., and Lebo, Z. J.: Opinion: A critical evaluation of the evidence for aerosol invigoration of deep convection, Atmos. Chem. Phys., 23, 13791–13808, https://doi.org/10.5194/acp-23-13791-2023, 2023
- Veals, P. G., A. C. Varble, J. O. H. Russell, J. C. Hardin, and E. J. Zipser, 2022: Indications of a Decrease in the Depth of Deep Convective Cores with Increasing Aerosol Concentration during the CACTI Campaign. J. Atmos. Sci., 79, 705–722, https://doi.org/10.1175/JAS-D-21-0119.1
- Zang, L., Rosenfeld, D., Pan, Z., Mao, F., Zhu, Y., Lu, X., & Gong, W. (2023). Observing aerosol primary convective invigoration and its meteorological feedback. *Geophysical Research Letters*, 50, e2023GL104151. https://doi.org/10.1029/2023GL104151