Review of EGUSPHERE-2025-5149: Aerosol-deep convection interaction based on joint cell-thermal tracking in Large Eddy Simulations during the TRACER campaign

General Comments:

In this study, the authors run two simulations at high-resolution (dx=dy=200m) of two cases during the TRACER field campaign. They initialize these simulations with clean and polluted aerosol conditions, based on observations from TRACER. The thermals within these simulations are tracked to assess the impacts of aerosol impacts on deep convective cloud processes on a thermal-centric level. Generally, their figures and discussion were well-presented and easy-to-follow. Their thermal-centric view is very compelling and allows for a more detailed assessment of aerosol impacts on finer scales than prior studies, which adds new insights to specific processes, particularly with respect to the locations of hydrometeors with respect to the thermal centers. This focus on hydrometeors is particularly helpful given that this study also focuses on lightning. The authors' conclusions generally support results shown in the literature with different modeling frameworks, which is encouraging. While I appreciate the authors' additional focus on environmental feedbacks, which are important, most of my more significant questions and concerns are related to their analyses and experimental set-up with regards to the environmental feedbacks. Additionally, some other clarifications, including details about their thermal sample composites and the aerosol evolution, would strengthen the manuscript.

We thank Reviewer #1 for this general positive overview of our manuscript. In the following we will provide replies to all specific comments one by one in italics and blue color. A track-changes file with the new version of the manuscript is also available, with all new additions in blue text and supplementary material at the end.

Specific Comments:

- L100: I appreciate that the authors constrain their analyses with realistic aerosol conditions from observations.

Thank you. Indeed, this makes our results more robust.

- L105-107: How well do the log-normal distributions with the parameters listed in Table 1 fit the TRACER observations? A quantification of this fit or an image that shows the fit would be helpful for the reader.

We now include Figure R1-0 below as supplementary material (Fig. S1). It shows how the lognormal fits capture the overall structure of the observed size distributions. The total number concentration derived from the modal fits is 1003 cm^{-3} for the clean regime and 15434 cm^{-3} for the polluted regime, with differences on the order of ~5-6%. (see ***lines 107-110)***

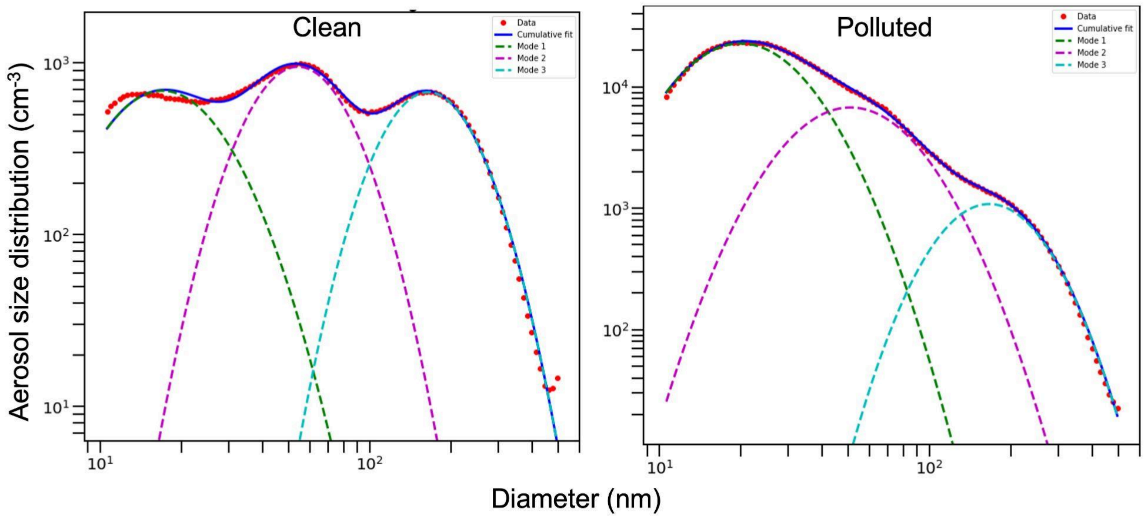

Figure R1-0 (same as Fig. S1 in supplementary material): Multimodal aerosol size distributions at the TRACER main site for clean (<10th percentile) and polluted (>90th percentile) conditions during 17 June to 07 August 2022. Red markers represent observations from scanning mobility particle sizer (SMPS) measurements, the blue line shows the cumulative lognormal fit, and dashed lines indicate the three fitted modes.

- L126: Can the authors provide an estimate of the vertical grid spacing, as some readers may not be able to easily estimate how 91 model levels translate to vertical distance?

Since the WRF configuration used in this study employed a typical sigma terrain-following vertical coordinate, the geopotential height of each model vertical level varies according to terrain height. The grid spacing (resolution) increases roughly with height. Tables R1-1 and R1-2 below show an example of the baseline geopotential height and layer thickness of the mass grid in the Arakawa-C staggered grid over terrain with a height of 7.5 meters near the TRACER AMF1 site.

Table R1-1. Geopotential heights of the model mass grid points over terrain height of 7.50 meters near the TRACER AMF1 site.

Layer 1: 36.18 m	*Layer 31: 5203.43 m*	*Layer 61: 11018.94 m*
Layer 2: 106.54 m	*Layer 32: 5396.62 m*	*Layer 62: 11212.40 m*
Layer 3: 202.33 m	*Layer 33: 5589.79 m*	*Layer 63: 11405.99 m*
Layer 4: 324.16 m	*Layer 34: 5782.98 m*	*Layer 64: 11599.72 m*
Layer 5: 477.04 m	*Layer 35: 5976.23 m*	*Layer 65: 11793.57 m*
Layer 6: 670.74 m	*Layer 36: 6169.57 m*	*Layer 66: 11987.47 m*
Layer 7: 907.05 m	*Layer 37: 6363.03 m*	*Layer 67: 12181.38 m*
Layer 8: 1083.19 m	*Layer 38: 6556.62 m*	*Layer 68: 12375.29 m*
Layer 9: 1179.89 m	*Layer 39: 6750.35 m*	*Layer 69: 12569.16 m*
Layer 10: 1277.42 m	*Layer 40: 6944.22 m*	*Layer 70: 12762.98 m*
Layer 11: 1375.80 m	*Layer 41: 7138.21 m*	*Layer 71: 12956.77 m*
Layer 12: 1522.61 m	*Layer 42: 7332.33 m*	*Layer 72: 13150.58 m*
Layer 13: 1717.34 m	*Layer 43: 7526.56 m*	*Layer 73: 13344.45 m*
Layer 14: 1911.89 m	*Layer 44: 7720.89 m*	*Layer 74: 13538.51 m*
Layer 15: 2106.22 m	*Layer 45: 7915.34 m*	*Layer 75: 13732.49 m*
Layer 16: 2300.35 m	*Layer 46: 8109.88 m*	*Layer 76: 13925.76 m*
Layer 17: 2494.31 m	*Layer 47: 8304.50 m*	*Layer 77: 14118.05 m*
Layer 18: 2688.11 m	*Layer 48: 8499.16 m*	*Layer 78: 14309.54 m*
Layer 19: 2881.74 m	*Layer 49: 8693.81 m*	*Layer 79: 14500.45 m*
Layer 20: 3075.23 m	*Layer 50: 8888.35 m*	*Layer 80: 14690.92 m*
Layer 21: 3268.63 m	*Layer 51: 9082.75 m*	*Layer 81: 14881.05 m*
Layer 22: 3462.00 m	*Layer 52: 9276.97 m*	*Layer 82: 15070.89 m*
Layer 23: 3655.38 m	*Layer 53: 9471.01 m*	*Layer 83: 15260.53 m*
Layer 24: 3848.81 m	*Layer 54: 9664.85 m*	*Layer 84: 15449.94 m*
Layer 25: 4042.32 m	*Layer 55: 9858.52 m*	*Layer 85: 15639.02 m*
Layer 26: 4235.92 m	*Layer 56: 10052.06 m*	*Layer 86: 15827.73 m*
Layer 27: 4429.58 m	*Layer 57: 10245.51 m*	*Layer 87: 16016.13 m*
Layer 28: 4623.22 m	*Layer 58: 10438.90 m*	*Layer 88: 16204.24 m*
Layer 29: 4816.76 m	*Layer 59: 10632.24 m*	*Layer 89: 16392.12 m*
Layer 30: 5010.16 m	*Layer 60: 10825.58 m*	*Layer 90: 16579.79 m*

Table R1-2. Layer thicknesses for the model mass grid points over terrain height of 7.50 meters near the TRACER AMF1 site.

Layer 1: 57.77 m	*Layer 31: 193.21 m*	*Layer 61: 193.39 m*
Layer 2: 82.94 m	*Layer 32: 193.17 m*	*Layer 62: 193.52 m*
Layer 3: 108.64 m	*Layer 33: 193.17 m*	*Layer 63: 193.66 m*
Layer 4: 135.00 m	*Layer 34: 193.21 m*	*Layer 64: 193.80 m*
Layer 5: 170.77 m	*Layer 35: 193.29 m*	*Layer 65: 193.89 m*
Layer 6: 216.63 m	*Layer 36: 193.40 m*	*Layer 66: 193.91 m*
Layer 7: 255.98 m	*Layer 37: 193.52 m*	*Layer 67: 193.92 m*
Layer 8: 96.30 m	*Layer 38: 193.66 m*	*Layer 68: 193.90 m*
Layer 9: 97.11 m	*Layer 39: 193.80 m*	*Layer 69: 193.84 m*
Layer 10: 97.95 m	*Layer 40: 193.93 m*	*Layer 70: 193.79 m*
Layer 11: 98.81 m	*Layer 41: 194.06 m*	*Layer 71: 193.79 m*
Layer 12: 194.81 m	*Layer 42: 194.17 m*	*Layer 72: 193.83 m*
Layer 13: 194.65 m	*Layer 43: 194.28 m*	*Layer 73: 193.93 m*
Layer 14: 194.44 m	*Layer 44: 194.39 m*	*Layer 74: 194.18 m*
Layer 15: 194.22 m	*Layer 45: 194.50 m*	*Layer 75: 193.79 m*
Layer 16: 194.04 m	*Layer 46: 194.58 m*	*Layer 76: 192.75 m*
Layer 17: 193.88 m	*Layer 47: 194.64 m*	*Layer 77: 191.84 m*
Layer 18: 193.71 m	*Layer 48: 194.68 m*	*Layer 78: 191.13 m*
Layer 19: 193.55 m	*Layer 49: 194.62 m*	*Layer 79: 190.67 m*
Layer 20: 193.43 m	*Layer 50: 194.47 m*	*Layer 80: 190.28 m*
Layer 21: 193.37 m	*Layer 51: 194.31 m*	*Layer 81: 189.97 m*
Layer 22: 193.37 m	*Layer 52: 194.14 m*	*Layer 82: 189.72 m*
Layer 23: 193.40 m	*Layer 53: 193.95 m*	*Layer 83: 189.55 m*
Layer 24: 193.46 m	*Layer 54: 193.74 m*	*Layer 84: 189.27 m*
Layer 25: 193.56 m	*Layer 55: 193.59 m*	*Layer 85: 188.89 m*
Layer 26: 193.63 m	*Layer 56: 193.49 m*	*Layer 86: 188.55 m*
Layer 27: 193.69 m	*Layer 57: 193.41 m*	*Layer 87: 188.25 m*
Layer 28: 193.60 m	*Layer 58: 193.36 m*	*Layer 88: 187.98 m*
Layer 29: 193.47 m	*Layer 59: 193.33 m*	*Layer 89: 187.76 m*
Layer 30: 193.33 m	*Layer 60: 193.34 m*	*Layer 90: 187.58 m*

- L149-150: How was this vertical distribution of aerosol decided? Was it also based on observations from TRACER?

When we started our work, we had no observation data for the vertical profile of the aerosol. Thus, we set the initial vertical profile with a constant value from the surface up to 1,500 meters (m) of geopotential height, decreasing exponentially with a scale height of 800 m above 1,500 m. This would represent a typical vertical profile at 7 p.m. local time, when the well-developed daytime planetary boundary layer (PBL) switches to the nighttime PBL, as described in Section 2.2. We are now aware that the Texas A&M group published Chen et al. (2025) regarding the aerosol vertical profiles during the TRACER campaign. The plot below compares the vertical profiles of normalized aerosol concentrations from the observations and our model simulations at the location of AMF1 *site. The normalized aerosol concentrations from the observations on August 7, 2022, are actually derived from those prepared for the TRACER Model Intercomparison Project (https://arm-synergy.github.io/tracer-mip/Roadmap.html; https://github.com/ARM-Synergy/tracer-mip/blob/59c0249519558d79ee4130eee52f2d5cf1148033/Pyplt.TRACER_Aerosol_Profiles_MIP.ipynb), based on Chen et al. (2025). Here, we compare normalized concentrations rather than raw values because our model simulations employed typical concentrations for polluted and clean conditions that differ from those on the given day. The base concentration for the observational data normalization is one at the surface. The base concentration for the simulation data normalization is one at the surface in the initial condition, i.e., the concentrations shown in Table 1. As the simulation forecasting time progresses, from the initial condition (blue line), the normalized vertical profiles (red lines) get closer to the observed profiles (black line). The vertical reduction rates appear similar except around the PBL top.*

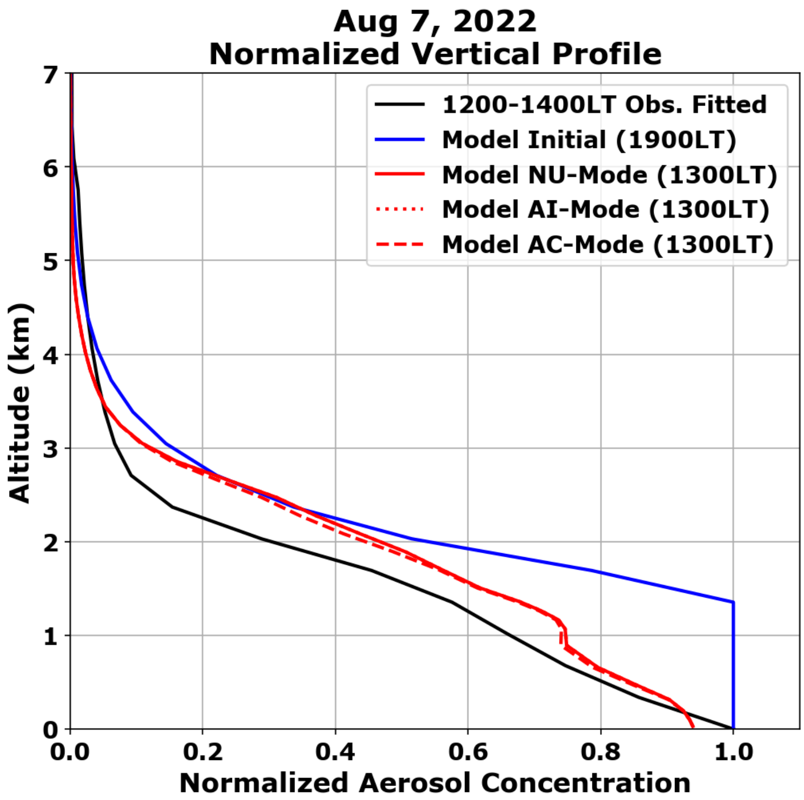

Figure R1-1: Vertical profiles of the normalized aerosol concentration. Observation values (black solid line) are derived from lidar measurements at the AMF1 site from 1200 to 1400 local time (LT) and are normalized by the concentration observed at the surface. The blue line shows the initial model simulation profile, which is normalized by the surface concentration (i.e., the concentration listed in Table 2). The red lines show the simulated profiles at 1300 LT.

- L168: How are the cell merger and splits accounted for. Are they removed? The authors state that the tracks are “refined” based on the evaluation of mergers and splits, but this is vague. Can the authors be more explicit here?

*We have revised this fragment of the text to (see **lines 172-173**),*

“… subsequently linked across time steps into cells. A final processing step allowed for cell splitting and merging (Sokolowski et al. 2024, Sec. 3.2) by aggregating the paths of cells that came within 25 km and five min into storm tracks.”

- L170: What is a neighboring feature? This was not defined, so it is hard to understand this limitation on their tracks.

*First, we have corrected a minor error in the original version of the manuscript: the reflectivity threshold we use is 35dBZ, and not 25dBZ. Second, regarding the definition of a “neighboring feature”, the maximum number of features within a 5-km radius is set to one in order to identify isolated cells. Under this criterion, if multiple cells exceeding the 35-dBZ threshold are located within 5 km of one another, they are treated as a single isolated cell. This clarification has been added to the revised manuscript (**lines 176-178**).*

- L206: It is great that the authors consider more than one case to create more robust results.

Thank you for acknowledging this.

- L225: The authors use comparative language (e.g., relatively), but it is unclear what they are comparing against.

Before addressing this, we should point out that by looking in detail into this comment, we realized that the correct mass flux units should be kg m^{-2} s^{-1}, and not kg m s^{-1}. We have now corrected this in Fig. 4b (previously Fig. 3b), and in Fig. 6g (previously Fig. 5g) by dividing by each thermal’s volume.

Regarding the use of comparative language, indeed, this is unclear, in particular regarding mass flux, since there is no unique reference that tells us what is high or low mass flux, and this strongly depends on the area through which one computes it, which is arbitrary. However, we see that the mass flux of individual thermals is typically ~2-3 orders of magnitude smaller than the total mass flux (based on the sum of all thermals) of each case, as shown in Figure R1-2.

*This reflects the fact that deep convective clouds are made up of many thermals, with these thermals contributing similar amounts of mass flux, as opposed to the classical plume view, in which just a few strong updrafts add up to the total mass flux. We have now adjusted this part of the text, making this point clear regarding mass flux, and including comparative measures for lifetime and travel distances. That is, thermals live short times (5 min on average, with 90% living less than 9 min) compared to cell lifetime (24.4 min on average), and travel short distances (<2km) compared to the total cloud depth (>~10km). (See **lines 240-247**.)*

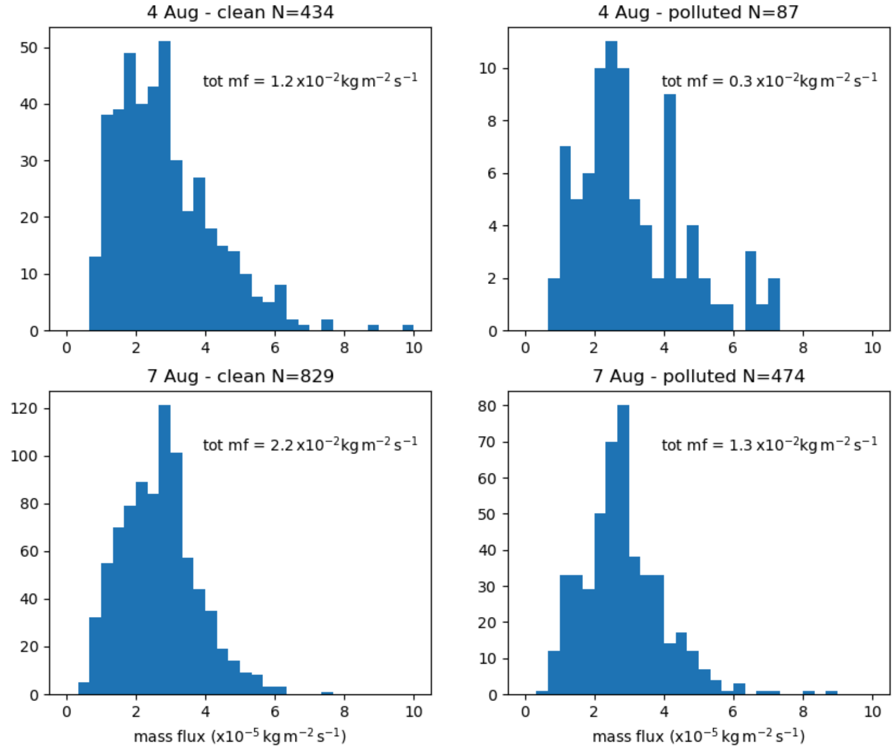

Fig. R1-2: Histograms of mass flux of thermals from the simulations of 4 August (top panels) and 7 August (bottom panels), for the clean cases (left panels) and the polluted cases (right panels). The total mass flux, computed as the sum of all contributions from thermals, is annotated in each panel.

- L238-240: The authors track thermals within the tracked cells for each of the two cases. For the clean simulations, ~33% of thermals come from 08/04 case and 67% of thermals come from the 08/07 case. However, for the polluted simulations, only 15% of the thermals come from the 08/04 case, while 85% are from the 08/07 case. As such, differences in polluted and clean simulations may be impacted by the different data weights from these cases. Have the authors checked this sampling impact on their results?

*Thank you for pointing this out. Indeed, the number of thermals from each case may impact the results, given that there is a larger number of thermals from the 08/07 case. To verify this, we have performed a bootstrap technique in which we use the same number of thermals from each case to redo the analysis. To do this, we reduce the number of thermals used from the 08/07 case to match the same number of thermals from the 08/04 case. We choose the 08/07 subset of thermals randomly. We repeated this procedure 6 times, and compared these 6 new versions of the different figures (Figs. 5, 6, 7, 8 and 9) with the original ones where we use all thermals from the 08/07 case. Overall, we find that our results are robust against this sampling. However, we do see that the response regarding ice, snow, graupel and hail is weaker when using this more equilibrated sampling. This is most evident from Fig. 5, from which we show here the 6 different versions of it using random sampling of thermals in the 08/07 case (Fig. R1-3 and Fig. S2 in the supplementary material). In conclusion, our results are robust against this sampling of thermals, but we must state that the response of ice, snow, graupel and hail may be slightly overestimated by this "unbalanced" sampling. (see **lines 261-264 and 375-379**)*

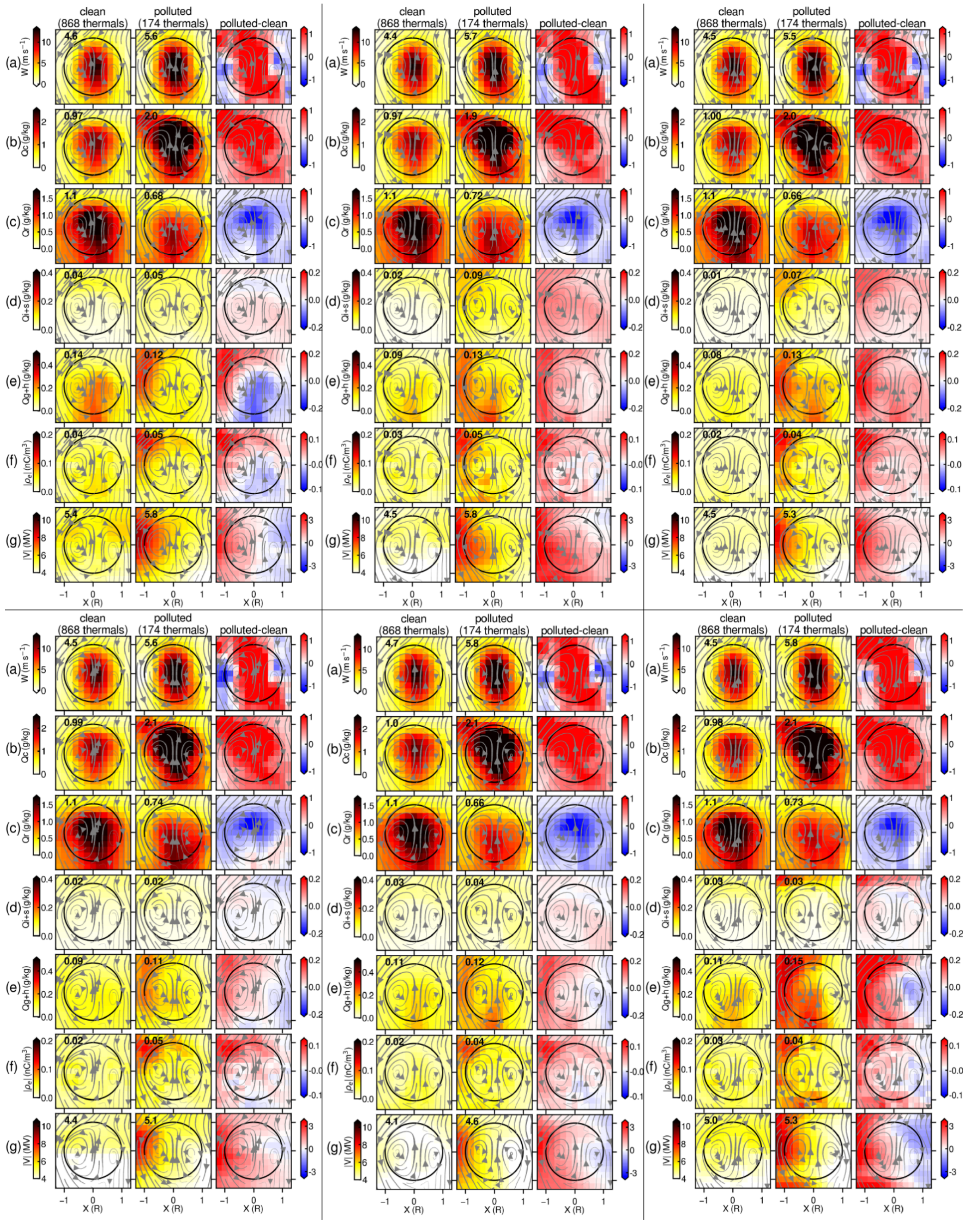

Figure R1-3 (Fig. S2 in supplementary material): Composites of different variables for thermals in the clean and polluted cases (as in Fig. 5 in the revised manuscript), in which each panel uses a different random sample of thermals from the 08/07 case, in order to have an equal size sample from both cases (08/04 and 08/07).

- Figure 4: Before diving into the results of clean and polluted thermals, it would be helpful to understand how the aerosol concentration profiles evolved during the simulation. It seems like much of the analysis is focused on the last 6-12 hours of their simulation, after 12-18 hours of model integration. As such, during these last 6-12 hours, do the differences in clean versus polluted match those shown in Table 2 from the initialization, and if not, what are the aerosol differences at these later times? This would be helpful in terms of understanding the actual aerosol perturbation being analyzed in this study and help with the sensitivity experiments described later.

The plots below (Fig. R1-4), which are the new Fig. 3 in the manuscript, show the time series of normalized aerosol concentrations averaged horizontally at a height of 1,500 meters above mean sea level for each simulation case. The base concentrations used for normalization are those listed in Table 1. As described in Section 2.2, the concentration at 1,500 meters is equal to the surface concentration in the initial condition. Concentrations at 1,500 meters are used as the reference because this height is consistent with the cloud base height shown in Fig. 6. The plots demonstrate that the normalized concentrations remain nearly consistent regardless of clean or polluted conditions, as well as the aerosol mode. This indicates that the normalized concentration represents the scale of differences in aerosol concentrations between clean and polluted conditions at 1,500 meters compared to the differences in concentrations in Table 1, which are derived from observations at the surface. Approximately, the normalized concentrations range from 0.75 to 0.95 over the last six hours on 4 August and from 0.75 to 0.85 over the last six hours on 7 August. The thermals analyzed in the subsequent sections are expected to be affected by the aerosol perturbation scaled by these range numbers rather than by the raw values shown in Table 2 of the surface observations. The normalized concentration at 1800 UTC (1300 CDT) on 7 August differs from the vertical profile plot because the former is a horizontal average, whereas the latter is sampled only at the AMF1 location. The vertical profile derived from observations at the AMF1 site indicates that the simulated normalized concentration at a height of 1,500 m could be an overprediction. However, we note that there is large uncertainty in how representative the observed vertical profile is of vertical profiles across the simulation domain, which includes various types of land use.

*We added these plots (Fig. 3 in the new manuscript) and new text to the end of Section 2.4 of the manuscript (see **lines 226-236**).*

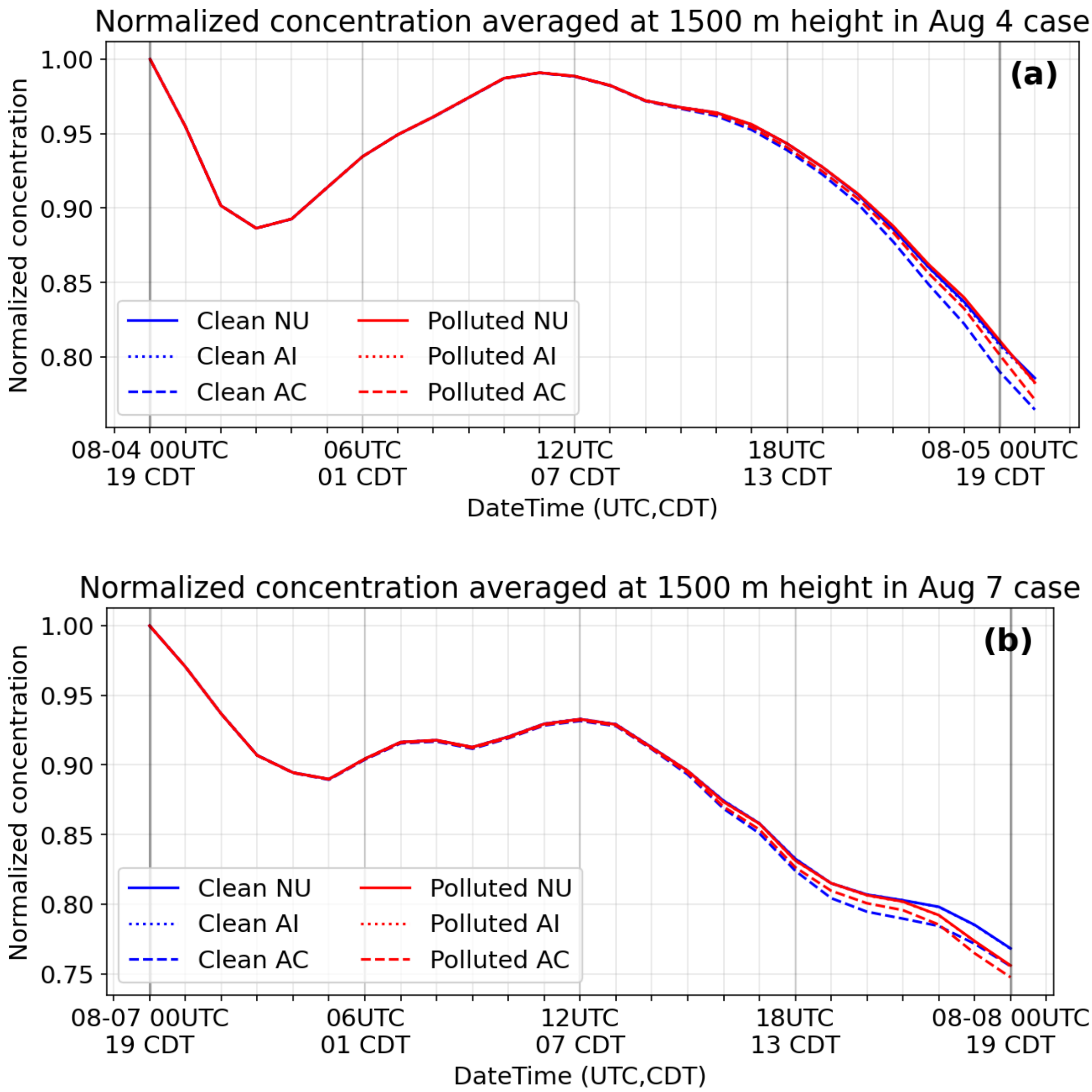

Fig. R1-4 (same as new Fig. 3 in manuscript): Temporal evolution of the horizontally averaged normalized aerosol concentration at a height of 1,500 meters above sea level from the start to the end of the simulation in each case. In the legend, NU, AI, and AC denote the aerosol nucleation, Aitken, and accumulation modes, respectively (Table 1). UTC and CDT represent coordinated universal time and central daylight time, respectively

- Figure 5: There seems to be very few thermals that initiate below 2 km (panel d). I understand their reflectivity threshold on cells may miss the early stages of the convective lifecycle; however, for more mature cells, are thermals not originating in the boundary layer? Can the authors explain the lack of thermals below 2 km AGL?

One of the criteria in order to identify thermals is a minimum liquid or ice water content of at least 1×10^{-5} g/kg. This helps avoid capturing updrafts that may be more related to *turbulence (both in-cloud and below-cloud), which would introduce other types of features different from cumulus thermals. So even if a cumulus thermal may have originated below the boundary layer, our algorithm will not identify it before it actually reaches the cloud (and is thus made of "cloudy-air"). Being able to trace back the origin of these cumulus thermals within the boundary layer would require significant additions to our algorithm, which is out of the scope of this study. However, it is a good idea for future studies.*

- L252 / L312: The thermal analyses weight their composite by mass flux, which emphasizes the strongest thermals. I am curious how sensitive their findings are to this weighting and whether an equal weighting (e.g., regular mean and median) produces similar results or provides different insights on aerosol impacts on thermals.

In order to investigate the sensitivity of our results to this mass flux weighting, we have re-processed the thermal composites using non-weighted averages, and generated the same figures (Figs. 5 & 9 in new manuscript). We find slightly different numbers, but overall the same features when comparing clean and polluted conditions. We show both versions of these figures below (Figs. R1-5 and R1-6).

Notice that all signs of the differences between clean and polluted are the same whether using mass flux weights or not, with some of them slightly smaller but others slightly larger. Thus, mass flux weights do not affect our results. This is consistent with the fact that most thermals contribute similar amounts of mass flux (see Fig. R1-2 above).

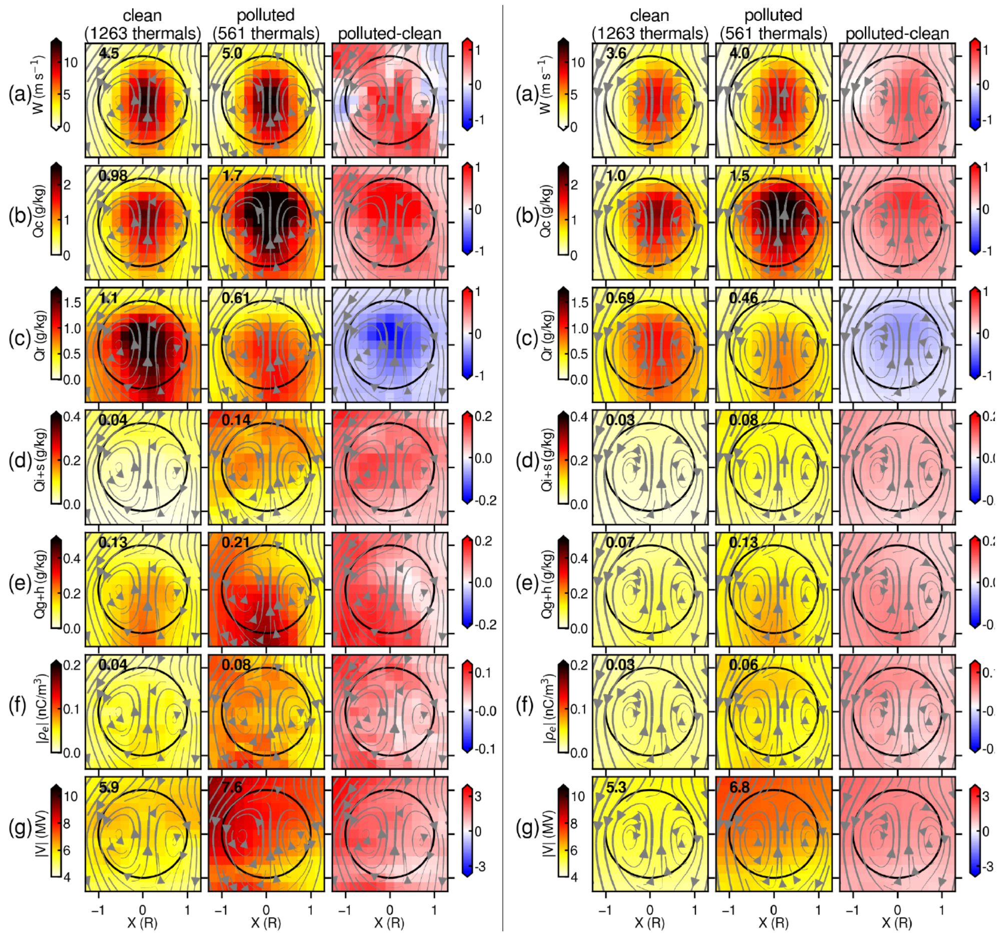

Fig. R1-5: Thermal composites of different variables as in Fig. 5 of the new manuscript. Left panel uses mass flux weights, whereas the right panel weighs all thermals equally.

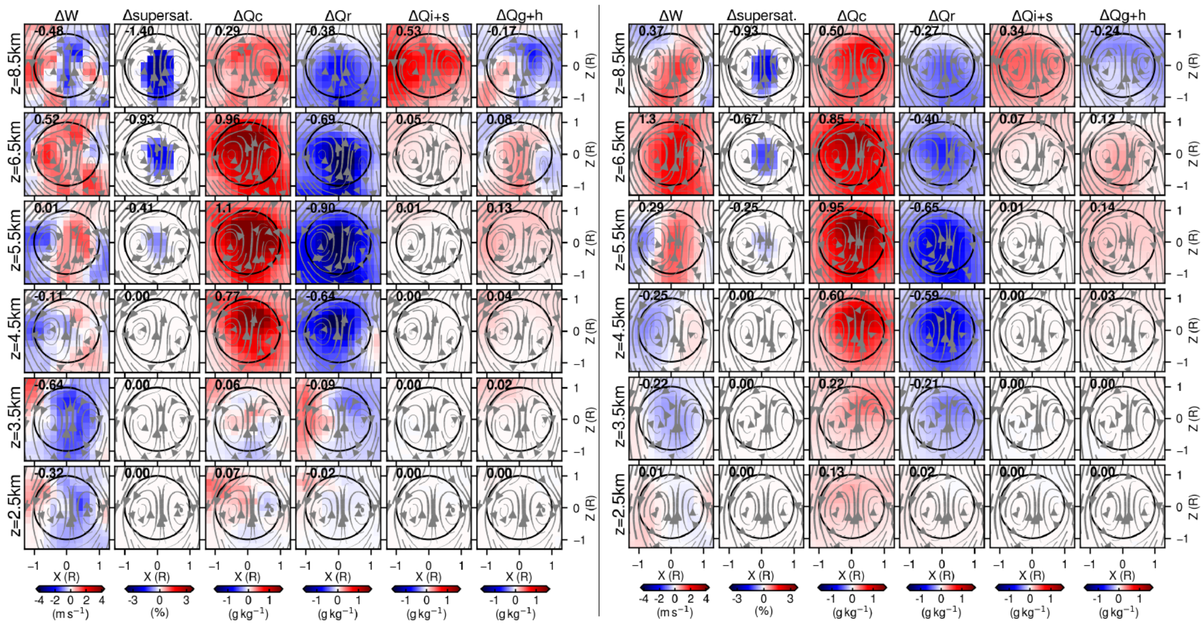

Fig. R1-6: Difference of thermal composites by altitude as in Fig. 9 of the new manuscript. Left panel uses mass flux weights, whereas the right panel weighs all thermals equally.

- Figure 8: I really like the analyses in and evolution of Figures 4, 7 and 8.

Thank you.

- Figure 11: It seems like having a time series of a reflectivity or rain rate or updraft contour in panels a, b, e, f, such that the time evolution of updraft or precipitation would be more comparable to the time evolution of the Meso-LLC in panels c, d, g, h.

We have revised Figure 11, which now also depicts a time series of rainfall contours (>1mm/hr). We have divided it into two figures, one for the 4 August case (now ***Fig. 12****) and one for the 7 August case (now* ***Fig. 13****).*

*We have added the following sentence (****lines 437-439****):*
"The timing of rainfall nearly coincides with the meso-LLC. More specifically, positive meso-LLC (deep convergence) develops first in association with strong convection (ensemble of thermals), while surface rainfall is slightly delayed."

- L373-381: The Aerosol-Cloud-Precipitation-Climate Model Intercomparison Project, which focused on the Houston region with similar sea breeze convection, also found a very similar trend in the environmental response of temperature and moisture at the same vertical levels described in this section. As such, I think it would be helpful to cite that work (Marinescu et al., 2021) here, which will help demonstrate the robustness of this result.

Thank you for pointing this out. We have added more references to this study (see ***lines 411-413****)*

- L391: 0 and 5 km AGL seems like a very deep layer in terms of understanding low-level wind convergence. Why do the authors use 5km AGL as an upper limit? Could other thermal circulations above the bottom (inflow) of the thermal be included in these results due to using such a high limit?

We first conducted both visual and statistical comparisons of grid-level low-level convergence (LLC; 250 m resolution) integrated over depths of 2.5 km and 5 km. The results show that the 5-km integration depth is much more strongly correlated with the column hydrometeor path (i.e., the sum of all hydrometeor species), which is largely governed by the evolution of vigorous cumulus thermals in deep convection. As shown in Fig. 3b in the manuscript, LLC of the 2.5-km depth is primarily related to the convergence induced by shallow, weak thermals, whereas the 5-km depth better represents *convergence due to deeper, longer-lived, and stronger thermals that are typically evolved from the middle of the clouds. Strong middle-level thermals tend to produce substantial hydrometeor mass. This is clear from the scatter-density plot below (Fig. R1-7), which compares LLC integrated over 2.5 km and 5 km with the total hydrometeor path.*

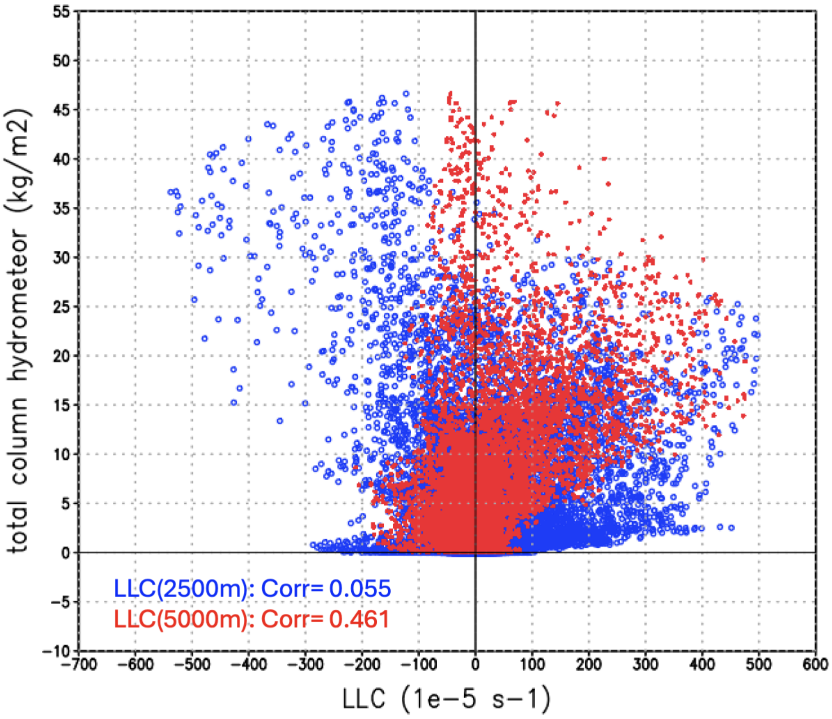

Fig. R1-7: Scatter plot of total column hydrometeor per unit area versus low level wind convergence (LLC) integrated over 2.5km (blue) and over 5km (red).

- L394: Are the mesoscale low-level convergence features discussed here referring to cold pools, a sea breeze front or some other low-level convergence boundary? Some clarification here would be helpful.

LLC integrated to 5 km (LLC_5000) primarily represents deep convergence associated with strong convective thermals. As shown in Fig. R1-7 above, a negatively correlated regime between hydrometeor amount and LLC integrated to 2.5 km (LLC_2500) appears on the left side of the distribution; this feature is not evident in LL_C5000. This negative correlation is most likely associated with downburst or microburst processes driven by intense raindrop evaporation and precipitation drag, which generate strong low-level divergence near the surface. When LLC is integrated over a deeper layer (5 km), these shallow, near-surface effects are smoothed out, yielding a clearer signal of deep convective convergence. Particularly the 10km-averaged field can clearly indicate the location and timing of deep convection (Fig. 12).

*We added **lines 426-431** in the new manuscript to clarify this.*

- L394: Additionally, depending on what the feature is causing the mesoscale low-level conference in this case (e.g., sea-breezes, cold pools), tying in some of the seminal work in terms of aerosol impacts on those specific mesoscale features may be helpful in terms of understanding the physical mechanism at play.

*We have added following sentence **(see lines 432-435)**:*

"Meso-LLC in the Houston–Gulf region is primarily driven by the inland penetration and stalling of the sea breeze under weak synoptic flow, augmented by land–sea thermal contrasts and urban heat-island effects, which enhance convergence and convective initiation (e.g., Wang et al., 2022 & 2024, Choi & Lee, 2021, Mages et al., 2024). Cold-pool outflow boundaries from prior convection can merge with the sea breeze and further strengthen convergence (Thompson et al., 2026)."

- L394-400: What is the physical mechanism by which aerosol particles impact the clustering differences of convective features within the different simulations? Can the authors speculate about this and/or relate this finding to prior research?

*Initial aerosol-cloud interactions tend to create a warmer and drier boundary layer in the polluted case (Fig. 10). This environment and mesoscale dynamical process can generate a stronger cold-pool environment due to the raindrop evaporation, which can trigger convection clustering. We added the following sentences (**lines 445-448**) as well as an additional plot related to cold pools (Fig. 19):*

"Aerosols can impact the clustering or aggregation of convection. For example, Increasing aerosol loading can modify raindrop evaporation and downdraft latent cooling, thereby changing cold-pool strength and gust-front lifting so that secondary convection becomes either more or less organized depending on the altitude of the environmental dry layer (Tao et al. 2007; Grant and van den Heever 2015)."

- L397-398: The authors state that under polluted conditions their cells / forcing are more tightly clustered and more consolidated based on a *qualitative* assessment of Figure 11, which is a snapshot at one time. There are many quantitative assessments of organization in the literature (e.g., I_{org}, L_{org}), and perhaps these could be used to provide a more quantitative assessment.

In order to quantify more rigorously this clustering/aggregation, we have conducted an additional analysis using band-path spectral analysis of the low level convergence (LLC) field, and included an additional figure (Fig. 14 in the new manuscript). See ***lines 449-466*** *and Figure 14 in the new manuscript.*

- L398: The authors state that the low level convergence is "markedly stronger". Can the authors quantify this, as it is hard to determine this from Figure 11? Additionally, have the authors considered including convergence in their thermal-centric view plots (e.g., Figure 8), which I thought were very compelling analyses and would show the mesoscale dynamical feedbacks that more directly impact the thermals / updrafts?

Please see above plot and discussion related to LLC quantification.

- L402: In terms of understanding the mesoscale feedbacks and interactions, the sample size of what has been presented is effectively 2 simulation snapshots. While I recognize the computational limitations of running an ensemble at high resolution, I do think that it may be hard to draw conclusions on the impacts of mesoscale clustering based on this small sample size. This limitation should be clearly stated in this section and the conclusion.

*Yes, we totally agree. Our intent is NOT to generalize these findings beyond the present case. The results discussed here are specific to this particular LES configuration and event, and their interpretation should be viewed in that context. We now explicitly clarify this limitation here (**lines 466-468**) and in the conclusion section (**lines 638-641**).*

- L407: The authors state that appreciable strengthening only occurs in the late stages, but what do they mean by appreciable? And what is being strengthened? Can they quantify and be more explicit?

New LLC bandpath shows quantitative values of "appreciable strengthening" in mesoscale energetic counts. We have added the new Figure 14 in the revised text (see also comment above).

- L409: Can they authors be a little clear on what hypothesis they are testing here. They mention a time-sequential impact on the mesoscale environment on the prior line, but what exactly does that mean?

*We have added the following sentence (see **lines 476-477**):*

"These results suggest that later-stage deep convective invigoration is led by mesoscale dynamical feedback, rather than by local aerosol–cloud interactions. To verify this hypothesis..."

- L409-418: The authors conduct sensitivity simulations where they re-initialize their aerosol concentrations at the beginning of a 2-hour period at the end of their simulations. How does their model deal with such a strong perturbation from one time step to the next? For example, a grid point with clean aerosol conditions is replaced by one with polluted aerosol conditions instantaneously, which could lead to unphysical responses, particularly right after the change. Did the authors consider allowing for some spin-up time to allow for unphysical conditions to be alleviated by the model?

To test the above hypothesis, aerosol concentrations were instantaneously re-initialized while all other dynamical, thermodynamical, and cloud microphysical fields were held fixed. In other words, only the aerosol concentration was altered, with no accompanying changes to the background state. Although this approach is not physically realizable, it serves as a controlled sensitivity experiment designed to isolate the aerosol effect. No additional spin-up period was applied following the perturbation.

- L409-418: How sensitive are the authors results to the specific timing of these sensitivity experiments? For example, if the aerosol reinitialization occurred at 18 UTC or 20 UTC, as opposed to 22 UTC, how would that impact the results? Similarly, how sensitive are the results to the length of time of analysis after the reinitialization (2 hours)? I assume these results would be sensitive to the lifecycle stage of the convection. I think it would be helpful if the authors can spend more time motivating these experiments since they seem somewhat ill-constrained.

*At this point, we just want to simply test a focused hypothesis: that later-stage deep convective invigoration is driven primarily by mesoscale dynamical feedback rather than by local-scale aerosol–cloud interactions. Sensitivity experiments involving earlier reinitialization times and longer time lags are certainly of interest, but they are beyond the scope of the present manuscript and the amount of work, figures, and discussion won't fit in this article. This is left for future work or coordinated MIP-style experiments. We added the following sentence (see **lines 484-486**):*

"Additional sensitivity experiments examining time-lagged reinitialization with more stratified aerosol concentrations are beyond the scope of this study, but may be conducted in the future across a broader set of cases to further strengthen the discussion and scientific insights of this manuscript."

- L428-432. The authors show that the PollRun experiment that was reinitialized with "Clean" aerosol conditions have similar ice and charge conditions as the other PollRun experiments in their domain and time averaged profiles. The authors conclude that this suggests that the mesoscale environment development in the PollRun experiment plays a dominant role in impacting these hydrometeor and electrification profiles since the simulation with reinitialized clean aerosol shows a similar response. However, since these are based on domain mean profiles and the lifetimes of cloud ice can extend for many hours, could this finding be due to the fact that these simulations were run with polluted conditions for ~10 hours before this with polluted conditions. I am having a hard time understanding how the authors can isolate that this is due to the environmental response as opposed to the integrated microphysical response.

Full-domain statistics characterize how convection is distributed across the system as a whole; for example, aggregated convection can produce large total precipitation simply through increased spatial coverage of convective systems. In contrast, convective-core sampling isolates regions of active updrafts and therefore directly reflects local aerosol–cloud interactions.

These comparative analyses show that abruptly changing aerosol concentrations does not substantially alter domain-aggregated convection in the polluted–dynamic runs, but it does enhance convective intensity at the local (core) scale. In other words, while aerosol–cloud interactions can influence the characteristics of deep convection to some extent, gradually evolving mesoscale dynamical feedbacks exert a much stronger control on domain-wide convective activity in this experiment.

*We added the following sentences (**lines 504-508**):*

"Full-domain statistics capture the spatial extent and aggregation of convection across the system, while convective-core sampling isolates local updrafts and directly reflects aerosol–cloud interactions. Abrupt changes in aerosol concentration do not significantly affect domain-wide convection in polluted–dynamic runs, but they do intensify convection at the local core scale. Overall, mesoscale dynamical feedbacks dominate

domain-wide convective activity, with aerosol–cloud interactions playing a secondary, local role. "

- Figure 12: Given the focus on updraft thermals, it would be helpful to include vertical velocity as a panel in Figure 12.

*Vertical velocity and related discussion are added (see **Figures 15 and 16** and next comment).*

- L433: The authors conduct the same reinitialization experiments for their second case but do not show the results. What is the reasoning for this? The addition of these results would be valuable in terms of improving the robustness of their sensitivity experiment analyses.

*We now show the results for both cases in Figures 15 & 16, and have adjusted the text accordingly. Originally, 8/7 cases were focused to explain the mesoscale feedback in this case, but 8/4 are now added in Figure and discussion. Related discussions are updated within the revised manuscript. (see **Figures 15 and 16**, and **lines 490-523**)*

- L442: Are the differences/similarities in Figure 12 that are a result of the mesoscale environment due to changes in the temperature/moisture/stability or due to low-level convergence or both? The authors generally use the term "mesoscale environment," and I think it would be helpful to make it clearer what they are specifically referring to.

*We find that the nature of the feedback differs fundamentally between the August 4 and August 7 cases. In the August 4 simulation, the polluted run primarily enhances CAPE in the urban-downwind region, indicating a thermodynamic pathway to convective intensification. In contrast, the August 7 polluted run is characterized by strengthened low-level convergence and lifting in the same region, pointing to a predominantly dynamical pathway. While the precise mechanisms driving these changes remain uncertain, the results consistently suggest that aerosol–cumulus thermal interactions progressively modify the mesoscale environment toward the end of the simulations. We have added a thorough discussion regarding this (**lines 524-564**) and 3 new plots (**Figures 17, 18 and 19**).*

- L450: The authors refer to these cases as "golden." What do they mean by this?

Golden cases are selected based on three criteria: (1) sufficient observational data coverage, (2) frequent isolated deep convective events, and (3) robust model performance in reproducing precipitation.

- L483-491: As mentioned above, the ACPC MIP results from Marinescu et al., 2021 also found similar changes to the water vapor and temperature profiles in the environment (their Figure 5, panels i, j and resulting discussion) from a range of models for a similar case study (i.e., same region, type of convection). As such, it should be cited here to strengthen this finding.

Thank you once again for pointing this out. We have included new text referencing this study, as well as Saleeby et al. (2025) in the conclusion section (see ***lines 614-619****)*

- L492-494: Saleeby et al., 2025 focused on tracked updrafts and the microphysical process rates within aerosol perturbation experiments for similar convection in the same region with many different models in the ACPC MIP and could be used here to strengthen this finding about process rates.

*See previous comment (****lines 614-619****)*

- L509-510: Again, this finding about changes in the mesoscale environment using a regional LES domain within a 24-hour integration was also presented in the ACPC MIP studies (Marinescu et al., 2021).

See previous two comments and ***lines 614-619****.*

- L511-517. The authors state that "at least, quantifications of changes in the mesoscale environment will be requirement for future MIP activities." I apologize for bringing this up again, but it seems like the authors may not have been aware that this was done in the MIP studies that they reference here (Marinescu et al., 2021, Saleeby et al., 2025; and van den Heever et al., 2025). As noted in several comments in this review, Marinescu et al., 2021 does quantify changes to the environment within the ACPC MIP model simulations (their Figure 5) and discusses similar mesoscale, environmental feedbacks as discussed in this study to explain the aerosol-impacts on updrafts. Specifically, that study showed both warming and drying in the mesoscale environment boundary layer under more polluted conditions and stated explicitly in the conclusions section that "the consistent CCN-induced response [in updrafts] is likely related to an environmental feedback process where the High-CCN simulations have increased environmental instability as a result of warmer boundary layer temperatures and cooler cloud level temperatures." That study also shows and has discussion about the drier boundary layer environments and moister cloud level environments in polluted conditions. Given that the authors research focuses on the same region and with similar convective features (e.g., scattered, sea breeze convection), the consistent findings between the authors' work and the ACPC MIP strengthens the conclusion about aerosol impacts on the environment. Can the authors reword this section to accurately reflect this?

*Thank you for pointing this out. We have removed the sentence "At least, quantifications of changes in mesoscale environment will be required for future MIP activities" from the new manuscript, and reworded these paragraphs of the conclusions (see **lines 614-619**).*

Technical Corrections:

- Figure 7 caption – supersaturation with respect to liquid or ice?

Thanks for spotting this. It is with respect to liquid water. We have updated the figure caption accordingly (now Fig. 8).

- Some of the figures' text are small (e.g., Figure 9, 12), and I was wondering if the authors could make them larger.

Thank you, we have improved both Figures (now Figs. 10, 15 and 16) to make sure the text is larger.

References

Chen, B., Thompson, S. A., Matthews, B. H., Sharma, M., Li, R., Nowotarski, C. J., Rapp, A. D., and Brooks, S. D.: A new technique to retrieve aerosol vertical profiles using micropulse lidar and ground-based aerosol measurements, Atmos. Meas. Tech., 18, 5841–5859, https://doi.org/10.5194/amt-18-5841-2025, 2025.

Choi, Y., & Lee, Y.-H. (2021). Urban effect on sea-breeze-initiated rainfall: A case study for Seoul Metropolitan area. Atmosphere, 12(11), 1483. https://doi.org/10.3390/atmos12111483

*Grant, L. D., and S. C. van den Heever (2015), Cold pool and precipitation responses to aerosol loading: Modulation by dry layers, Journal of the Atmospheric Sciences, **72**(4), 1398–1408, https://doi.org/10.1175/JAS-D-14-0260.1.*

Mages, Z., Kollias, P., Treserras, B. P., Borque, P., & Oue, M. (2024). Shallow cloud variability in Houston, Texas during the ESCAPE and TRACER field experiments. Atmospheric Chemistry and Physics. https://doi.org/10.5194/egusphere-2024-2984

Tao, W.-K., X. Li, A. Khain, T. Matsui, S. Lang, and J. Simpson (2007), Role of atmospheric aerosol concentration on deep convective precipitation: Cloud-resolving model simulations, Journal of Geophysical Research, 112, D24S18, doi:10,1029/2007JD008728.

Thompson, et al. (2025). Understanding aerosol properties in convective outflows during TRACER. Aerosol Science and Technology. https://doi.org/10.1080/02786826.2025.2608352

Wang, D., Jensen, M. P., Taylor, D., Kowalski, G., Hogan, M., Wittemann, B. M., Rakotoarivony, A., Giangrande, S. E., & Park, J. M. (2022). Linking synoptic patterns to cloud properties and local circulations over Southeastern Texas. Journal of Geophysical Research: Atmospheres, 127. https://doi.org/10.1029/2021jd035920

Wang, D., Melvin, E. C., Smith, N., Jensen, M. P., Gupta, S., Abdullah-Smoot, A., Pszeniczny, N., & Hahn, T. (2024). TRACER perspectives on Gulf-Breeze and Bay-Breeze circulations and coastal convection. Monthly Weather Review, 152, 2207–2228. https://doi.org/10.1175/mwr-d-23-0292.1