

Responses to Reviewer #2
REVISIONS DUE: September 28, 2026

We sincerely thank the reviewer for their time, careful evaluation of our manuscript, and constructive feedback. We have addressed the suggestions intended to improve the impact of the manuscript. Please refer to our detailed responses to the major and minor comments in red below.

Major comments:

1. There are significant differences between GCCN concentrations from the two instruments used in the analysis: the Cloud Droplet Probe (CDP) and the Cloud-Aerosol Spectrometer (CAS). This difference is not explained, it is not clear which one could be more precise, and GCCN concentration is crucial for the study. This difference propagates into the entire analysis (e.g., discrepancies in Fig. 7). Including data from the third instrument available (Fast Cloud Droplet Probe, FCDP) might reveal which of the results is closer to reality. FCDP was not included because of a higher threshold of droplet detection than that of the other instruments. Still, its results could at least be used in Fig. 2 to give confidence in CAS or CDP. A comparison with data from the Giant Nucleus Impactor (GNI) used in other campaigns (e.g., Jensen and Nugent 2017) would also be welcome.

We thank the reviewer for this suggestion and have added the dataset mean FCDP distribution to Fig. 2. Over the overlapping size range, the CAS and CDP distributions generally agree more closely with each other than either does with the FCDP, particularly at larger diameters. There is an approximate factor of 2 discrepancy between the GCCN concentrations from the CDP and the CAS, which is significantly smaller than the spread across prior studies presented in Lewis and Schwartz (2004), their Figure 22b. Furthermore, a quantitative comparison with GNI measurements is not currently feasible because a comprehensive publication or dataset for the GNI collected under comparable conditions is not available. Currently, a synthesis of the open ocean GNI data, including relations with wind speed, is not yet available in the scientific literature for comparison. We have clarified these limitations and reiterate that comparison with suitable GNI measurements would be valuable when such data become available.

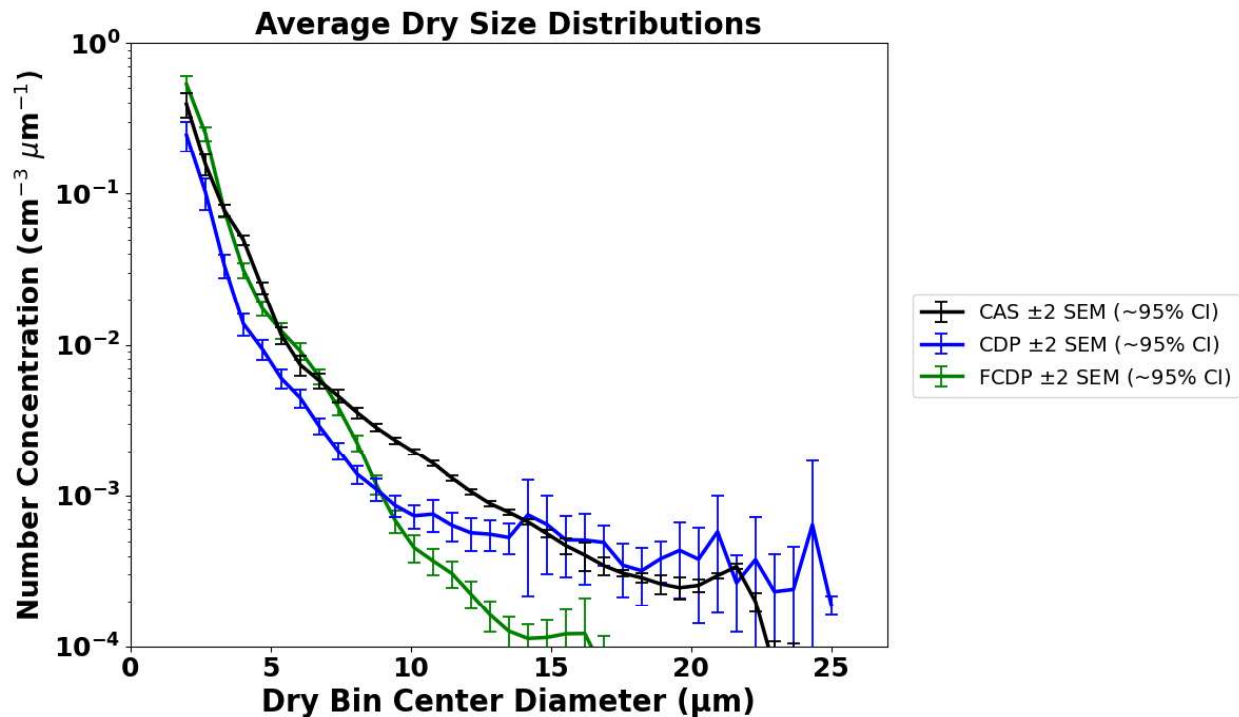


Figure R1 Average dry aerosol size distributions derived from CAS (black), CDP (blue), and FCDP (green) measurements. Error bars represent ± 2 standard errors of the mean (95% confidence interval) based on variability across contributing flight legs (see section 2.3 for details) for January-June 2022 below cloud base flight legs.

2. Authors conclude that GCCN concentration cannot be explained by wind speed only, but that other unknown factors are also important. This comes from looking at correlations between CCN concentration and wind speed, which are weaker than in existing aerosol flux parameterizations. However, they measure GCCN concentration and wind speed well above the sea surface (around 700 m on average). This can lead to errors in the analysis. First, GCCN at these altitudes can come from prior emissions at different wind speeds or from advection from other regions (given GCCN lifetime could be between hours and days). Second, there could be significant error in deriving near-surface wind speed from measurements at 700 m using a logarithmic correction. These caveats need to be addressed.

We acknowledge the important points regarding the altitude of the GCCN and wind speed measurements and agree that extrapolation of aircraft wind measurements from several hundred meters altitude to 10m introduces uncertainty. We examine both the dependence of GCCN number concentration on sampling altitude and the sensitivity of the estimated 10m wind speed to the assumed surface roughness length. We find a weak decrease in GCCN number concentration with altitude, with substantial variability in GCCN number concentration occurring at similar altitudes. Sampling altitude within the boundary layer does not appear to be a primary control on the observed GCCN variability. This behavior is consistent with previous airborne observations (Gonzalez et al., 2022; Ackerman et al., 2023). We also tested the sensitivity of the logarithmic wind correction to the assumed surface roughness length. Using our original $Z_0=0.02$ m (a relatively high value for open ocean water) gives a mean estimated U_{10} of 5.31 m s^{-1} . Decreasing Z_0 to 0.001 and 0.0001m (increasingly calm water) increases the mean U_{10} to 6.10 and 6.50 m s^{-1} .

s^{-1} , respectively, corresponding to mean differences of 0.79 and 1.2 m s^{-1} . Thus, the typical uncertainty associated with the choice of roughness length is approximately 1 m s^{-1} . This is small relative to the 0-16 m s^{-1} range of wind speeds represented in our analysis, and more importantly, is smaller than the standard deviation of the wind speed estimates (2.65 m s^{-1}), as well as narrower than the bin widths we use when deriving the concentration (or mass) wind speed relationship (Figure 4 in the revised manuscript). Furthermore, the corrected wind speeds remain highly correlated with those obtained using the original roughness length ($r=0.999$ and 0.998 , respectively). We have provided a few statements regarding this in the updated manuscript.

Linear regression confirms a statistically significant decrease in GCCN number concentration with altitude for both instruments. The fitted slopes are $-0.67 \pm 0.09 \text{ cm}^{-3} \text{ km}^{-1}$ for the CAS and $-0.19 \pm 0.05 \text{ cm}^{-3} \text{ km}^{-1}$ for CDP. This corresponds to an estimated decrease between the surface and 700m of 0.47 ± 0.06 and $0.14 \pm 0.03 \text{ cm}^{-3}$, respectively. However, altitude accounts for only 13.4% and 4.4% of the variability in number concentration for CAS and CDP, respectively. While there is a small altitude dependence, sampling altitude alone cannot account for the large observed variability in GCCN concentration. We therefore conclude that, although there is uncertainty in our estimates of 10m wind speed, this uncertainty is unlikely to be the primary source of the weak observed dependence of GCCN number concentration on wind speed.

3. On multiple occasions, authors conclude that rain enhancement is better correlated with GCCN mass than GCCN concentration. That is true for the modeling data only. In observations (Fig.7), CAS data show the opposite (that concentration has a stronger influence), and CDP data show similar effects of GCCN mass and concentration. This could be related to higher uncertainty of GCCN mass measurements due to small sampled volume (rare, large GCCN are rarely sampled, and they affect mass more strongly than concentration).

We agree that the stronger relationship between rainfall and GCCN mass is primarily a modeling result and is not reproduced in the observational analysis. To test the robustness of this result, we revised the observational analysis in several ways. First, rather than estimating GCCN mass concentration and number concentration independently from either instrument as was originally done in the manuscript, we averaged the CAS and CDP GCCN size distributions to create single flight-mean estimates of $M_{d>2}$ and $N_{d>2}$. We used our previously calculated leg-means (section 2.3) for each instrument and then combined them across the CAS and CDP for each flight. This approach incorporates information from both probes while allowing the analysis to center on the distinction between the effects of $M_{d>2}$ and $N_{d>2}$, rather than the differences between the instruments. We then repeated the RWC/LWC analysis using cloud CWC and droplet N_d derived from (a) CAS, (b) the CDP, and (c) the mean of CAS and CDP.

These sensitivity analyses are shown below in Figures R2- R4. These analyses show that higher GCCN conditions are generally associated with enhanced RWC/LWC for both GCCN mass concentration and GCCN number concentration. The enhancement is strongest and most statistically robust in the central, more densely sampled LWC- N_d bins, where the percentage of bootstrap realizations greater than unity is near 100%, and the 90% confidence intervals lie entirely above unity. However, the magnitude of the High/Low RWC/LWC ratio varies among individual bins, and neither GCCN variable produces a consistently larger response. It is important to note that the 90% confidence intervals for the $M_{d>2}$ and $N_{d>2}$ enhancement responses overlap for most corresponding bins. We cannot statistically distinguish whether GCCN mass concentration or GCCN number concentration is more strongly associated with RWC/LWC

in the observational dataset. The individual CAS and CDP analyses, as well as the combined CAS and CDP sensitivity analyses, all support this same conclusion. These sensitivity analyses are shown below in Figures R2-R4.

We have revised Section 3.3 of the manuscript to state this result explicitly. The revised text now reflects that the overall RWC/LWC enhancement is statistically similar for $M_{d>2}$ and $N_{d>2}$, that neither variable produces a consistently larger response, and that the 90% confidence intervals for most corresponding bins overlap. We have removed the previous interpretation that $M_{d>2}$ produces a stronger observational precipitation response. The new manuscript distinguishes between the modeling and observational results more clearly. We explicitly state that the model shows a stronger and more consistent rainfall response to $M_{d>2}$, whereas the observations show broadly similar RWC/LWC enhancement for $M_{d>2}$ and higher $N_{d>2}$. We retain the CAS-based cloud CWC and N_d analysis in the main manuscript because the CDP-only and combined CAS-CDP sensitivity analyses all lead to the same overall conclusion.

Flight-mean $M_{d>2}$ and $N_{d>2}$ are moderately positively correlated ($r = 0.47$), which shows that flights with higher $M_{d>2}$ also tend to have higher $N_{d>2}$. This covariance reduces the extent to which the observational analysis can independently distinguish the associations of $M_{d>2}$ and $N_{d>2}$ with RWC/LWC.

We also assessed whether the high or low GCCN flight categories for $M_{d>2}$ and $N_{d>2}$ are distinct enough compared to the in-flight variability; we computed the difference between the mean GCCN mass concentration and number concentration values of the high and low flight groups with the median in-flight standard deviation calculated from the combined instrument leg-mean measurements. For $M_{d>2}$, the high and low groups differed by $11.25 \mu\text{g m}^{-3}$, compared with the median in-flight standard deviation of $5.57 \mu\text{g m}^{-3}$. The split separation between the High and Low groups was approximately 2 times the typical in-flight variability. For $N_{d>2}$, the high and low groups differed by 0.40 cm^{-3} , compared with a median in-flight standard deviation of 0.17 cm^{-3} . The separation between the High and Low groups was approximately 2.4 times the typical in-flight variability. These comparisons show that the separation between High and Low categories is distinct enough between higher and lower GCCN environments despite variability among individual legs within flights.

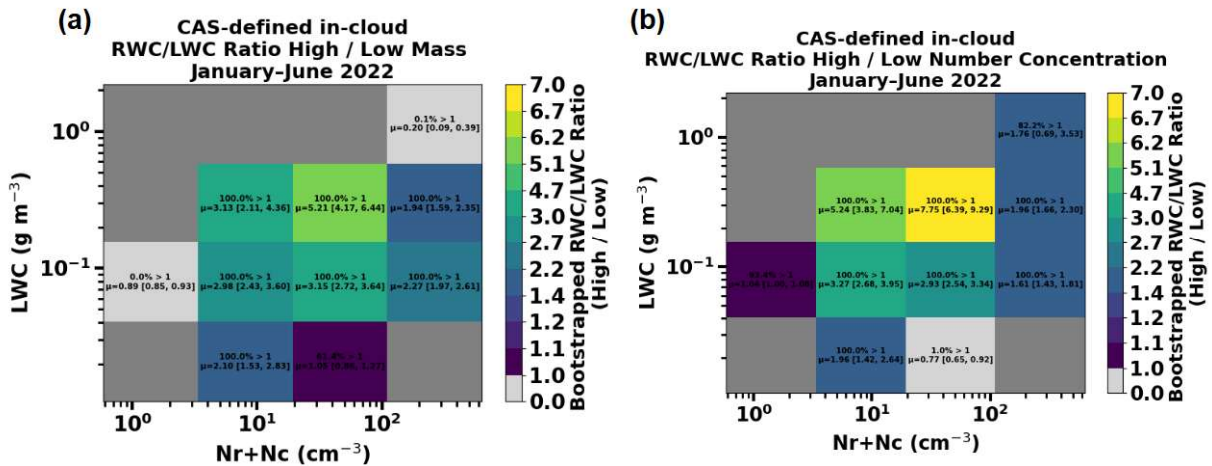


Figure R2 Observationally derived precipitation efficiency responses, defined as RWC/LWC , across bins (boxes) of LWC and $N_r + N_c$ for cases with high versus low $M_{d>2}$ (a) and high versus low $N_{d>2}$ (b). GCCN mass and number concentration are derived from the combined CAS and CDP measurements, while the CAS is used to define in-cloud conditions. Colors indicate the bootstrapped ratio of mean RWC/LWC in high-GCCN cases relative to low-GCCN cases within each bin (High/Low), while light gray bins indicate High/Low ratios below unity. Text within each box shows the percentage of bootstrap realizations producing ratios greater than 1, the mean ratio, and the 90% confidence interval of the bootstrap distribution. Light gray bins indicate bins where fewer than 100 valid data points are available across the compared subsets and are excluded from analysis. Dark gray bins met the minimum sampling threshold but yielded bootstrap distributions whose mean ratio did not exceed 1, indicating a statistically robust *reduction* of RWC/LWC under higher GCCN conditions in those bins.

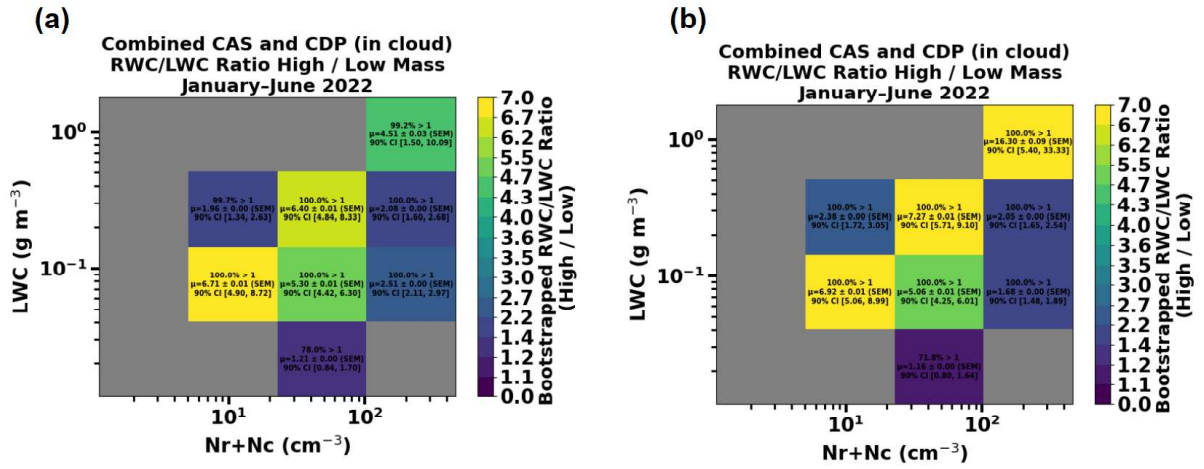


Figure R3 Observationally derived. Same as Fig. R2, both CAS and CDP are used to define in-cloud conditions by averaging both instruments at each 1 Hz.

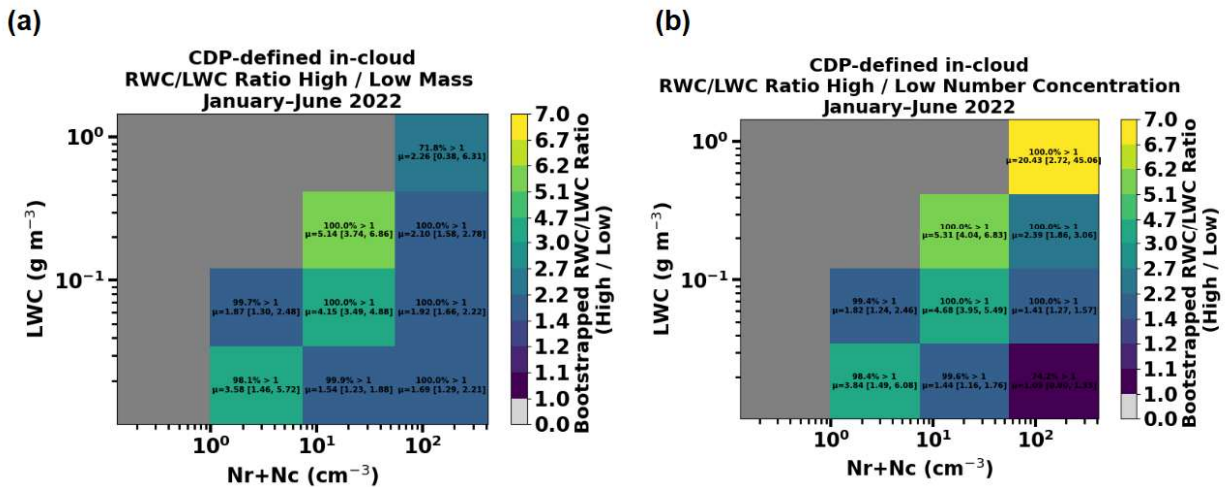
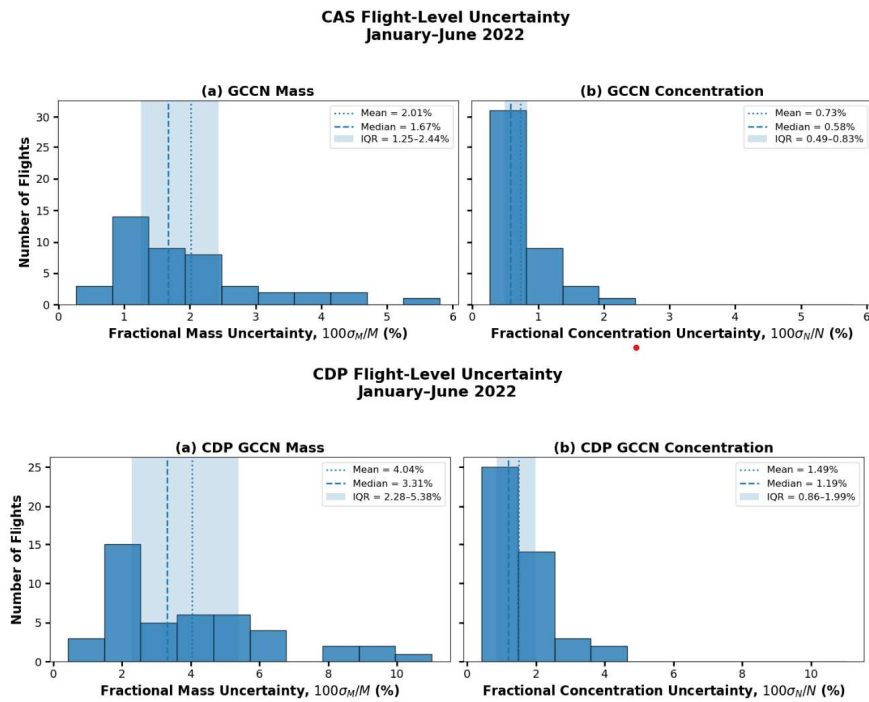


Figure R4 Same as Fig. R2, except CDP is used to define in-cloud conditions.

To evaluate whether sampling uncertainty may contribute to differences in results between mass and number, we estimated the Poisson counting uncertainties for $M_{d>2}$ and $N_{d>2}$ (Figure R2). The median fractional mass uncertainty is approximately three times larger than the number concentration uncertainty for both instruments: 1.67% versus 0.58% for CAS and 3.31% versus 1.19% for CDP. This is consistent with the weighting of mass, which makes it more sensitive to large particles that are less frequently sampled and therefore induce larger counting error. We have revised the manuscript to distinguish the modeling result from the observational relationships and to acknowledge the greater counting uncertainty associated with GCCN mass.



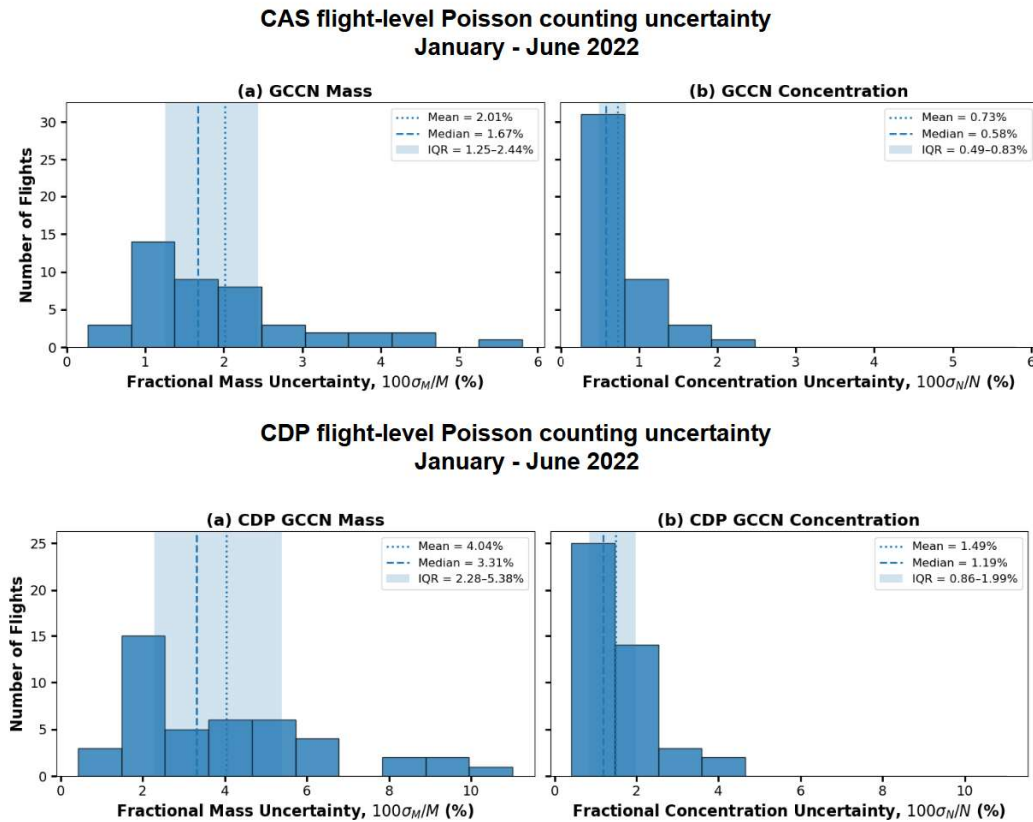


Figure R5 Flight-level Poisson counting uncertainties for GCCN mass and concentration derived from CAS and CDP measurements during January-June 2022. Shading denotes the interquartile range, and vertical lines indicate the mean and median. These estimates represent Poisson counting uncertainty only and do not include uncertainty associated with particle sizing.

Although the fractional Poisson counting uncertainty for GCCN mass is approximately three times larger than that for number concentration, the absolute magnitude of the counting uncertainty remains small, typically on the order of 1-4% for CAS and 2-8% for CDP. These uncertainties are small relative to the nearly two orders of magnitude range in mass concentration observed across the dataset. Therefore, Poisson counting uncertainty alone is unlikely to greatly affect the observed mass relationships. However, the Poisson estimates do not account for uncertainty associated with particle sizing. Because the derived GCCN mass depends strongly on particle diameter, sizing uncertainty can contribute much more to mass uncertainty than counting statistics. This would mean that the total uncertainty in GCCN mass measurements is likely much larger than the Poisson counting uncertainty. It is also likely larger than the uncertainty in number concentration, which we know to be dominated by counting statistics. The larger mass uncertainty would reduce the split between “high” and “low” mass groups used in the RWC analysis, and potentially weaken the split RWC/LWC results for mass. This would especially be true for flights with mass values right near or at the threshold that divides the two “high” and “low” populations. However, we should also keep in mind that uncertainty on the order of tens of percent remains much smaller than the overall range of GCCN mass across the dataset.

We have revised the manuscript to include a brief discussion of the greater uncertainty associated with $M_{d>2}$. Specifically, we now note that $M_{d>2}$ is more sensitive than $N_{d>2}$ to the sampling and sizing of relatively rare large particles. Although the Poisson counting uncertainties estimated here are small relative to the overall range of observed $M_{d>2}$, additional uncertainty associated with particle sizing is expected to affect mass more strongly than number concentration. We now acknowledge in the manuscript that greater uncertainty in $M_{d>2}$, could reduce the separation between the High and Low mass flights and potentially weaken the observed RWC/LWC response to $M_{d>2}$, especially for flights near the High/Low threshold. The detailed Poisson uncertainty analysis shown in Fig. R5 is provided here to address the reviewer's concern but is not included as a new figure in the manuscript.

4. In section 3.3 (observations of GCCN influence on precipitation), GCCN measurements cannot be co-located with RWC and LWC measurements. Therefore, flight-averaged GCCN concentration/mass is used. How big is the variability of GCCN concentration/mass during such a flight? If large, this adds a lot of uncertainty to the analysis.

We thank the reviewer for raising this point. We calculated variability using the standard deviation (s) of the leg-mean values for each variable within each flight and expressed it relative to the corresponding flight mean as $s/\bar{x}$. This is shown below in Figure R6 with the relevant median values displayed in the figure caption. There is substantial leg-to-leg variability in mass concentration and number concentration within individual flights, particularly for $M_{d>2}$. However, to place this variability in the context of the full dataset, we also compared the median in-flight standard deviation with the spread of the flight-mean values across all flights. For CAS $M_{d>2}$, the median in-flight s is $4.46 \mu\text{g m}^{-3}$, compared with a standard deviation of $5.44 \mu\text{g m}^{-3}$ among flight means; the median in-flight variability was therefore 82.1% across all flights. For CDP $M_{d>2}$, the corresponding values were 6.36 and $11.3 \mu\text{g m}^{-3}$, respectively, or 55.4%. For $N_{d>2}$, the median in-flight s was 0.21 cm^{-3} for CAS and 0.08 cm^{-3} for CDP, compared with across-flight standard deviations of 0.42 and 0.19 cm^{-3} , respectively. Thus, the median in-flight variability represented 49.5% and 40.6% of the across-flight spread for CAS and CDP $N_{d>2}$, respectively.

The leg-to-leg variability is substantial relative to the spread of flight-mean GCCN values across the campaign and therefore introduces uncertainty into the use of these mean values in the RWC/LWC analysis. This uncertainty is greater for $M_{d>2}$ than for $N_{d>2}$, consistent with the larger in-flight variability relative to the across-flight spread for mass concentration. To further evaluate how this variability affects the High/Low GCCN categories, we examined the distributions of individual leg-mean values from flights assigned to the High and Low groups (Fig. R7). The two populations overlap substantially for both GCCN variables, confirming that individual legs from flights classified as High can have values comparable to legs from Low flights. However, the modes remain separated by a factor of ~ 3 -5, the separation of the distributions somewhat greater for $N_{d>2}$ than for $M_{d>2}$, adding uncertainty but not invalidating the statistical approach used.

For the combined CAS+CDP measurements used to define the flight-mean GCCN groups, the difference between the mean High and Low populations is $11.3 \mu\text{g m}^{-3}$ for $M_{d>2}$, compared with a median within-flight standard deviation of $5.57 \mu\text{g m}^{-3}$, giving a separation of approximately 2.0 times the typical

in-flight spread. For $N_{d>2}$, the corresponding difference is 0.40 cm^{-3} , compared with a median in-flight standard deviation of 0.17 cm^{-3} , or approximately 2.4 times the typical in-flight spread. Thus, the flight-mean classification does distinguish relatively higher and lower GCCN environments on average, but the overlap among individual legs introduces larger uncertainty, particularly for the mass separation.

We have revised Section 3.3 of the manuscript, immediately following the description of the observational High/Low GCCN comparison, to directly acknowledge this limitation. We state that substantial leg-to-leg variability causes overlap between the High and Low flight populations and therefore adds uncertainty to the use of flight-mean GCCN for the cloud environment sampled during the RWC/LWC measurements. We further note that this uncertainty is greater for $M_{d>2}$ than for $N_{d>2}$, because the High/Low mass populations are less separated relative to the typical in-flight variability.

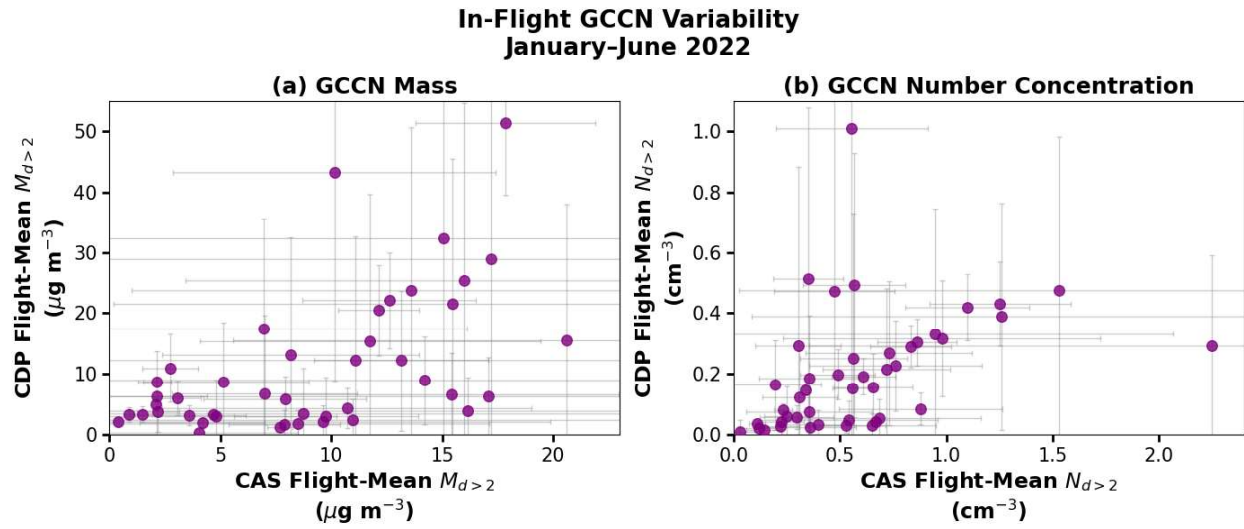


Figure R6 In-flight variability in GCCN mass concentration, $M_{d>2}$ (a), and number concentration, $N_{d>2}$ (b), measured by CAS and CDP during January-June 2022. Each purple scatter point represents the mean value calculated from the available legs within an individual flight. Horizontal bars represent ± 1 standard deviation (s) among the CAS leg means within each flight, while vertical bars represent $\pm 1s$ among the corresponding CDP leg means. The bars therefore indicate leg-to-leg variability within each flight rather than uncertainty in the flight mean. The median in-flight $s/\bar{x}$ was 70.7% and 105.6% for CAS and CDP $M_{d>2}$, respectively. The $s/\bar{x}$ was 47.9% and 89.2% for CAS and CDP $N_{d>2}$, respectively. The median in-flight s represented 82.1% and 55.4% of the across-flight standard deviation for CAS and CDP $M_{d>2}$, and 49.5% and 40.6% for CAS and CDP $N_{d>2}$, respectively.

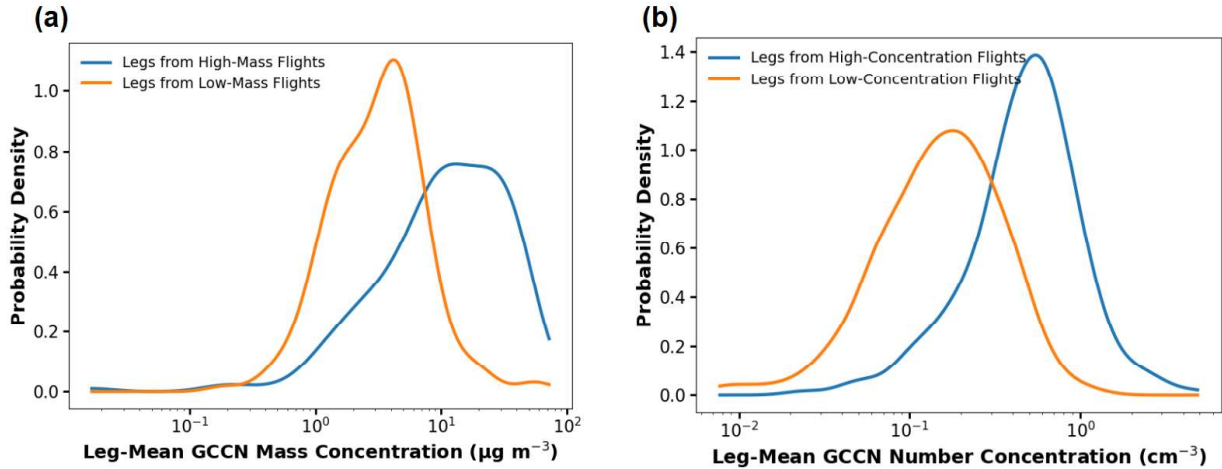


Figure R7 Probability density functions of $M_{d>2}$ (a), and number concentration, $N_{d>2}$ (b), for flight legs belonging to flights classified as High (blue) and Low (orange) based on the corresponding combined CAS+CDP flight-mean value. The distributions show substantial overlap between the High and Low flight categories, showing that leg-to-leg variability introduces uncertainty into the flight-mean group classification used in the RWC/LWC analysis.

5. There is significant variability in observations (e.g., Figs. 4, 8, 9). The analysis is based on 6 months, but the ACTIVATE campaign lasted 3 years. Including the entire period could smooth the variability and more reliably reveal trends. Why were only 6 months used?

We focused on data strictly from 2022 because it provides the most comprehensive ACTIVATE dataset, with more than 450 usable below cloud base flight legs, spanning multiple seasons and locations, whereas flight operations in 2020 and 2021 were substantially reduced by the COVID-19 pandemic. Each previous year contains fewer than one-third as many usable legs. Preliminary sensitivity analyses incorporating the available 2020 and 2021 observations do not greatly impact the results reported here; however, a full multi-year analysis is ongoing and is beyond the scope of this manuscript. We have clarified the rationale for selecting 2022 in the text.

6. Results of modeling should be quantitatively compared with the other cited modeling studies on the subject.

We thank the reviewer for this suggestion. A direct quantitative comparison is not entirely feasible because no cited study uses the same model framework and configuration (e.g., a fixed LWP of 385 g m^{-2}). Differences in model physics and environmental conditions would therefore prevent a like-for-like comparison. However, we attempt to compare quantitatively with Jensen and Lee (2008) because their modeled LWP is approximately 100 g m^{-2} . This is comparable to our thinner cloud case (yellow curve in Fig. R4). Using the approximate cloud droplet concentration of $N_c \approx 53 \text{ cm}^{-3}$ in their Figure 6, which is close to the 50 cm^{-3} case in our simulations, we compare their rain rates with our modeled rainfall as a function of GCCN mass. For GCCN masses of 0.9, 3, 10, and $40 \mu\text{g m}^{-3}$, Jensen and Lee (2008) give approximate rain rates of 0.0007, 0.0012, 0.0061, and 0.046 mm h^{-1} , respectively. Our corresponding 100 g m^{-2} LWP simulations yield approximately 0.0007-0.0009, 0.0026-0.0032, 0.0090-0.0107, and $0.0289\text{-}0.0324 \text{ mm}$ over the 1 hr simulation, depending on the presence of turbulence. Jensen and Lee (2008) modeling approach produces rainfall magnitudes that are generally comparable and agree within

approximately a factor of 2-3 (from a range that spans two orders of magnitude) over the overlapping GCCN mass range. We have added the Jensen and Lee (2008) values to Figure 5 to illustrate this quantitative comparison, shown below in Figure R4 and updated Figure 5 in the manuscript to reflect the Jensen and Lee (2008) comparison.

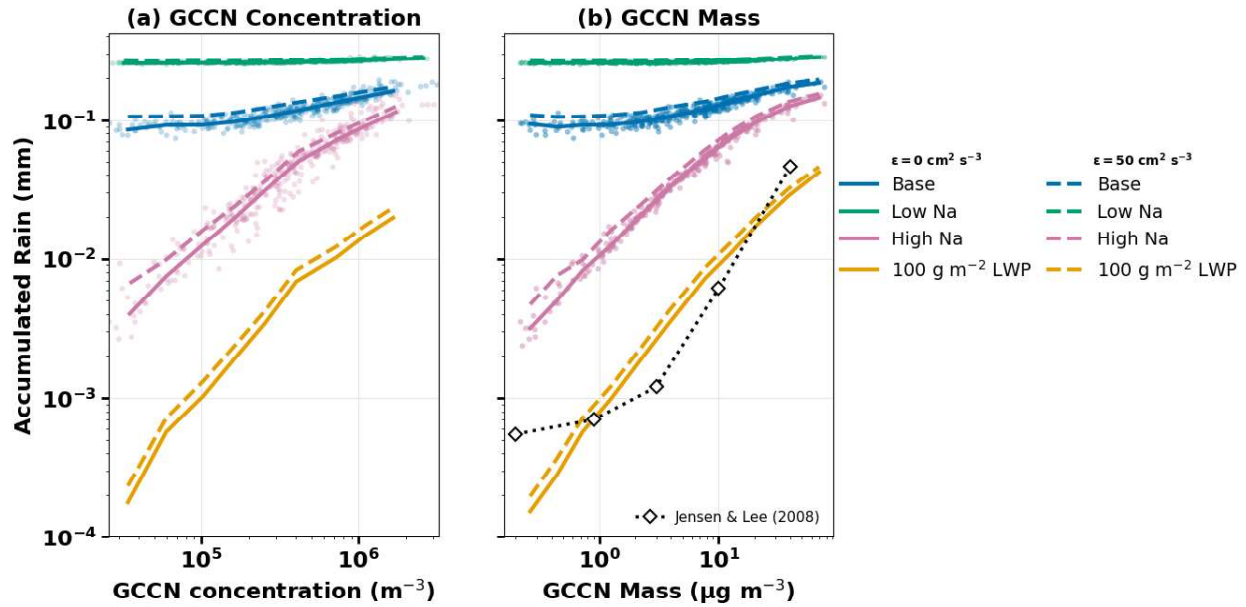


Figure R8 1-D kinematic super droplet microphysics model simulations showing accumulated rainfall over the 1 hr simulation, as a function of GCCN number concentration $N_{d>2}$ (a) and GCCN mass concentration $M_{d>2}$ (b). Model simulations were performed for all flight legs. For the “base” simulation ensemble (blue), the liquid water path is fixed at 385 g m^{-2} and the submicron aerosol concentration N_a is 50 cm^{-3} . Also shown are low LWP (100 g m^{-2}) simulations with the same N_a (yellow). Two levels of turbulent dissipation rate ($\epsilon = 0$ and $50 \text{ cm}^2 \text{ s}^{-3}$) are shown in each ensemble to represent increasing levels of turbulence. High N_a ensembles use a number concentration of 100 cm^{-3} and low N_a ensembles use a number concentration of cm^{-3} . Colored curves show the binned median responses for each ensemble, while the scatter points represent model estimates using individual flight legs. The median response provides a representative description of the ensemble behavior. The Jensen and Lee (2008) cases have an LWP of approximately 100 g m^{-2} , making them most directly comparable to our low-LWP simulation ensemble. The Jensen and Lee (2008) cases correspond to surface wind speeds of 1, 2, 4, 8, and 16 m s^{-1} , converted to mass concentrations for particles with $M_{d>2}$ of 0.2, 0.9, 3, 10, and $40 \mu\text{g m}^{-3}$, respectively. Jensen and Lee (2008) report rain rate in mm h^{-1} , while our simulations show accumulated rainfall (mm) over a 1 hr simulation.

Minor comments:

1. Why are error bars shown only for selected points, with more error bars for CAS than for CDP? Error bars represent ± 2 SEM and are displayed at every second interpolated radius for visual clarity. They have been updated to match the same interval for both instruments.
2. In Fig. 2: Plot aerosol distributions from other studies (e.g., Colón-Robles et al. 2006, Jensen and Nugent 2017) if possible.

We thank the reviewer for this suggestion. We elected not to add distributions from other studies to this figure due to differences in sampling duration, instrument characteristics, and data-processing methods that might limit the validity of a direct comparison with the approximately 3-min ACTIVATE flight legs.

Instead, we retain Fig. 2 as a comparison of the two selected ACTIVATE instruments and discuss the relevant literature, sampling limitations, and effects of counting statistics in detail in the accompanying text.

3. Lines 488–489: “the relative impact of microscale turbulence is greatest for the base N_a ensemble, and is also most important at the lower precipitation rates” —this is not clear to me looking at Fig. 5.

We thank the reviewer for pointing out that this interpretation was not evident from the figure. We have removed the statement that the relative turbulence effect is greatest for the base N_a ensemble and revised the text to emphasize that turbulence produces a comparatively modest increase in rainfall, with no consistent ranking among the aerosol ensembles.

4. Code used for simulations is not made available.

Simulation code has been made publicly available.

5. Author contributions need to be more specific—currently only 3 of 13 authors are mentioned.

Author contributions have been updated.

Editorial comments have been added.

References

Jensen, J. B. and Lee, S.: Giant sea-salt aerosols and warm rain formation in marine stratocumulus, *J. Atmos. Sci.*, 65, 3678–3694, doi:10.1175/2008JAS2617.1, 2008.