How does riming influence the observed spatial variability of ice water in mixed-phase clouds?

N. Maherndl, M. Moser, I. Schirmacher, A. Bansemer, J. Lucke, C. Voigt, and M. Maahn

August 15, 2024

Original Referee comments are in italic

> manuscript text is indented, with added text underlined and ~~removed text crossed out.~~

We would like to thank the reviewers for their helpful comments. We revised the manuscript and responded to all of the reviewers' comments.

In addition, we updated the IMPACTS W-band reflectivity Z_e data and the normalized rime mass M results obtained from the combined method for W-band, because a new version of the W-band dataset was published. Z_e was adjusted downward by about 0.9 dB leading to slightly lower M results. The positive bias due to saturation effects for Z_e values associated with large particles at W-band remains, but is less pronounced (Fig. 3c&d, Fig. 5).

Reviewer I

The authors explore the spatial variability of riming and its contribution to the clustering of ice within clouds associated with wintertime precipitation in the midlatitudes (IMPACTS) and marine cold air outbreaks over the high latitudes (HALO-(AC)3). They use a synergistic radar and in situ product to produce estimates of ice water content (IWC) with and without the influence of riming. By applying pairwise correlation functions to bulk microphysical parameters for a number of long flight segments (26 segments),

the authors aim to capture length scales associated with IWC clustering and can further compare the functions separately between those applied to IWC including riming and those excluding riming. I particularly like the section where long swaths of in situ observations are broken up into smaller segments to essentially maximize the sample size of environments, thereby producing a more robust statistical analysis. This analysis reveals clear modes in the spatial clustering of IWC. The paper follows logically towards its conclusions, and the figures are easily discernable and readable, and for that I thank the authors. However, I do have concerns with the robustness of the analysis. Additionally, I had confusion understanding the derived rimed and unrimed IWC, which hopefully could be better articulated/reorganized to improve upon the paper (mentioned in major comments). I recommend this paper be reconsidered with major revisions.

We thank the reviewer for the positive review and the constructive comments, which helped to improve the manuscript.

Major comments

Concerning the robustness of the results, I have two major points. First, I worry "artificial" positive pairwise correlation values are being produced by applying the moving average. This moving average is on scales of $\tilde{2}$ km, which is on the order of the largest observed positive correlations values (less than this value). I would propose sensitivity tests whereby varying the window size of the moving average. While I understand more robust measurements are obtained by averaging the in situ observations, it is very common to examine ice microphysical properties at 1 Hz scales (100m). It would be especially prudent to use smaller windows for moving averages especially when looking at lags below a few km.

Thank you for the comment. This is a valid concern, which we investigated using a sensitivity study as suggested. Because we calculate IWC using normalized rime mass M dependent mass-size relations for each time step and we need to use running averages to get a reliable M product, we can't investigate smaller running average window sizes for IWC. However, for the total number of ice particles N_i, we can go down to 1 Hz scales. We find that increasing the window size for computing running averages smooths peaks in the original signal and therefore the pair correlation function η gets closer to 0 the larger the averaging window. Spatial scales where $\eta > 0$ do not change significantly as long as the window size is reasonably small. Figure R.1 and Fig. R.2 show the same analysis as Fig. 9 of the manuscript for different window sizes for IMPACTS and HALO-(AC)^3, respectively. The dashed lines at 2 km and 4 km highlight the similar spacial scales where $\eta > 0$ for all window sizes, except larger 20 s for IMPACTS. We assume that due to the slow flight speed and generally lower N_i during HALO-(AC)^3, the 1 Hz results are noisy.

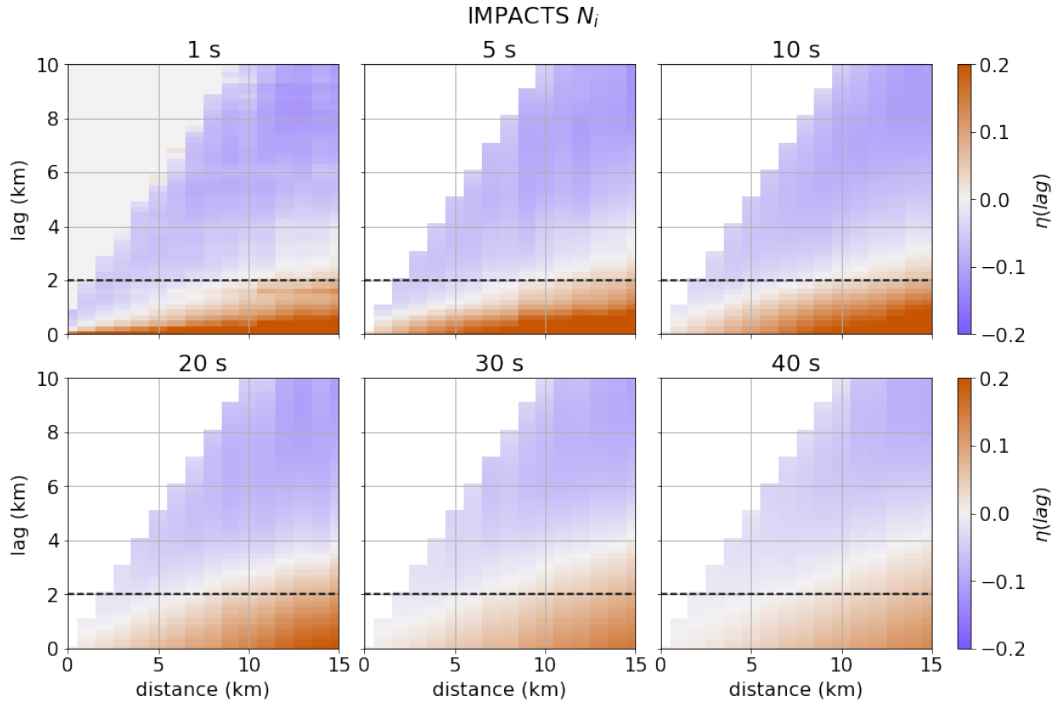

Figure R.1: Average pair correlation function (PCF) η as a function of distance and lag calculated using all IMPACTS flight segments for N_i for different running average window sizes. The dashed line shows the mentioned 2 km scale.

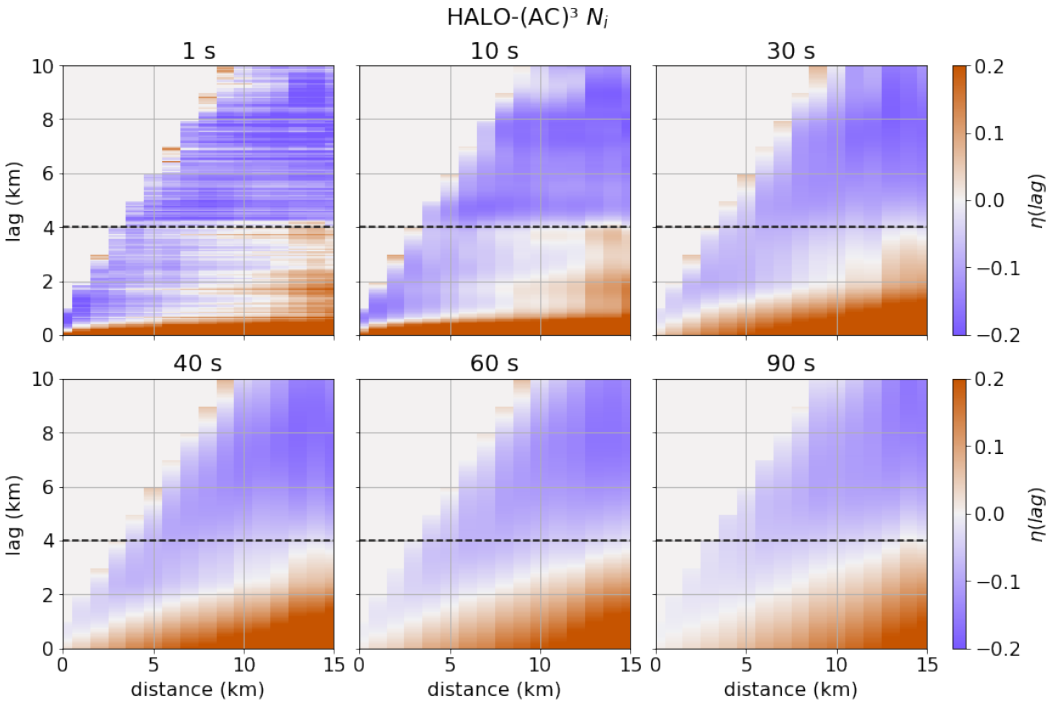

Figure R.2: Average pair correlation function (PCF) η as a function of distance and lag calculated using all HALO-(AC)^3 flight segments for N_i for different running average window sizes. Here, the dashed line is drawn at 4 km.

In the revised manuscript, we added:

> Because IWC is derived using running averages of 10 s and 30 s for IMPACTS and HALO-(AC)^3 data, respectively, we investigated the impact of the window size of the moving average on $\eta(r)$. We found that while increasing the window size from 1 s to 10 (30) s for IMPACTS (HALO-(AC)^3) decreases absolute values of $\eta(r)$, at which lags r $\eta(r)$ is positive does not change (not shown). This is because applying a moving average smooths peaks in the 1 Hz signal, but does not necessarily change their periodicity as long as the window size is reasonably small.

Second, there is no testing for the statistical significance of the pairwise correlation functions. This is especially a concern as standard deviations of the functions mostly overlap values equal to 0 (values expected of a homogeneously distributed system; Figure 7a,b). Further, some of the results of the rimed and "assumed-unrimed" IWC spatial inhomogeneity are nearly identical. If it's possible, applying some sort of bounds for rejection testing using white noise at some XX percentile could be helpful.

Thank you for raising this point. Not including significance testing was clearly an oversight from us. We are now using a Student's t test with a 95 % significance threshold

and reworked Fig.9. In panels (d) and (h), we only plot differences, where η_{IWC} is positive and highlight significant positive differences with hatching. We also only show the respective $\eta = 0$ line in (a)-(c) and (e)-(g) to make the plot easier to read.

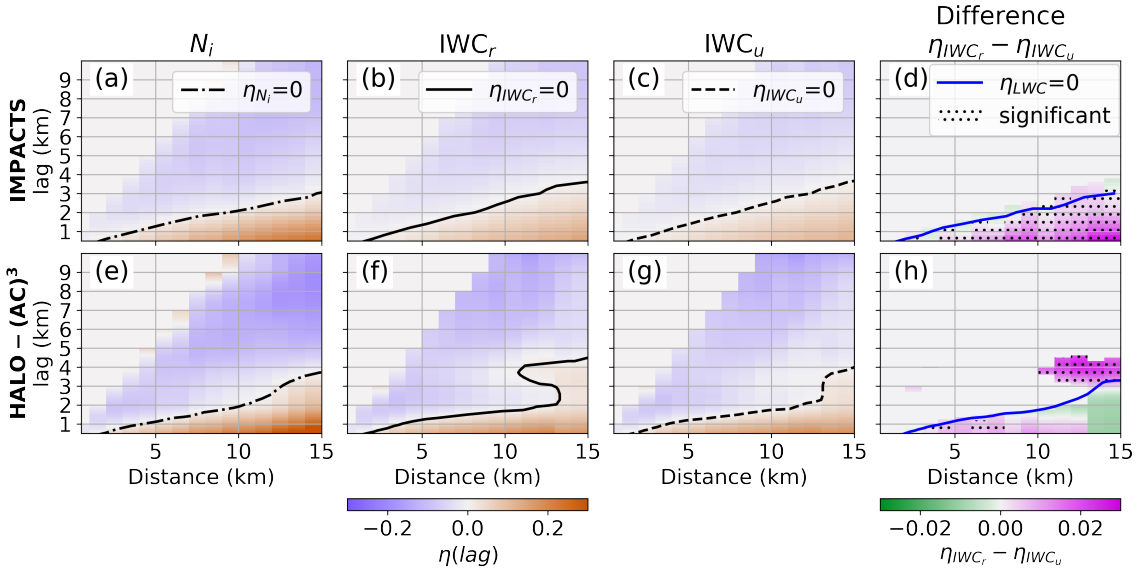

Figure R.3: Average pair correlation function (PCF) η as a function of distance and lag calculated using all (a-c) IMPACTS and (e-g) HALO-(AC)³ flight segments for (a)&(e) N_i, (b)&(f) ice water content (IWC) accounting for riming IWC$_r$, and (c)&(g) IWC assuming no riming IWC$_u$. The Difference between (b) and (c) are shown in (d); difference between (f) and (g) in (h). Differences in (d) and (h) are only shown, where $\eta_{IWC_r} > 0$. Areas, where differences are significant according to a Student's t-test (95 % significance threshold) are hatched. $\eta = 0$ is drawn as shaded lines for the ice number concentration N_i (~~dotted~~ dash-dotted black), IWC$_r$ (solid black), IWC$_u$ (dashed ~~grey~~black), and liquid water content (LWC, ~~dash-dotted~~ solid blue), where LWC measurements from King probe (Nevzorov ~~)~~ probe) measurements obtained during IMPACTS (HALO-(AC)³) are used.

In the text, we added:

> Differences between positive values of IWC$_r$ ~~,~~ and IWC$_u$ (Fig. R.3d) reveal that riming enhances the probability of ice clusters for distances larger 6 km for lags from about 1 km to 10 km (at distances of 12 km). To show the statistical significance of this enhancement, a one-sided Student's t-test with a significance threshold of 95 % is used. Areas where differences are significant are hatched (Fig. 9d).

I also experienced confusion in the methodology in deriving the rimed and unrimed IWC. Concerning the organizational comment, for example, separating section 3.3 and 4.2 confused me. The derivation of IWC influenced by riming and IWC not influenced by riming seems to be separated into multiple sections (3.3 & 4.3), when section 3.1 is titled "Retrieving ice particle riming". I'm also still not sure how IWC can be separately obtained assuming riming and no riming. Are you simply using different coefficients in the mass-diameter relationship for the two variables (which I assume would be an issue since riming would in theory impact the diameter of "unrimed ice")? I'm sure it's explained in the text, however, it's difficult to determine.

We apologize for the confusing structure. We have restructured the methodology section, which is now split into 3.1 "Retrieving ice particle riming", 3.2 "Deriving ice water content (IWC)", and 3.3 "Characterizing scales of ice water variability in clouds". The sensitivity study is now contained in section 4.2. Sect. 3.2 explains how we obtain IWC with and without accounting for riming in more detail, which was previously missing. We are indeed using different mass-size coefficients. When accounting for riming, we vary the mass-size coefficients depending on M for each time step. When neglecting riming, we keep the coefficients fixed at values for unrimed particles. This assumes, that the particles would have the same size, if they were unrimed. As you note, this assumption is likely not realistic, because riming typically increases particle sizes, as you mention. However, with this assumption, we underestimate the increase of IWC due to riming and therefore our findings. Sect. 3.2 reads

~~Characterizing scales of cloud variability~~Deriving ice water content (IWC)

IWC is calculated by summing the product of ice particle mass $m(D_{\max})$ and $N(D_{\max})$ for the probes' lower to upper size ranges D_{lower} to D_{upper}

$$IWC = \sum_{D_{lower}}^{D_{upper}} m(D_{\max}) N(D_{\max}) \Delta D_{\max}, \tag{1}$$

where $\Delta D_{\max}$ is the size bin width. $m(D_{\max})$ is approximated by a power law relation with prefactor a_m and exponent b_m

$$m(D_{\max}) = a_m D_{\max}^{b_m}. \tag{2}$$

a_m scales the density of ice particles (independent of particle size) and b_m modulates the size dependency of particle mass, which is related to particle shape and growth processes. a_m and b_m depend strongly on riming (e.g., Mitchell, 1996) and reported literature values range from 0.0058 to 466

for a_m and 1.8 to 3.0 for b_m in SI units (e.g., discussed by Mason et al., 2018). As shown by Maherndl et al. (2023b), a_m and b_m strongly depend on the amount of riming, which increases particle densities. Maherndl et al. (2023b) provide a_m and b_m values for discrete M, which are interpolate to obtain parameters for a continuous M in this study. We derive a_m and b_m for each time step as a function of the retrieved M. IWC is then calculated with Eq. 1 for each time step based on the measured PSD and the derived a_m and b_m parameters. We refer to this quantity as IWC_r (IWC accounting for riming).

To estimate the contribution of the riming process to IWC, we also calculate IWC using fixed mass-size parameters a_m and b_m for unrimed particles (also taken from Maherndl et al., 2023b), thereby neglecting density changes (e.g., due to riming). We refer to this quantity as IWC_u. IWC_u can be seen as the "theoretical" IWC, if the ice particles were unrimed so that the riming contribution can be estimated from the difference between IWC and IWC_u. However, this implies that riming does not impact the size of the unrimed ice particle, which is not necessarily the case in nature. Riming typically not only leads to an increase in ice particle density, but also ice particle size (Seifert et al., 2019). Therefore, we likely underestimate the contribution of riming to particle mass when comparing IWC_u with IWC. Since we are interested in the contribution of riming to IWC variability, this approach likely results in a conservative estimate of the contribution of riming to IWC variability.

Please see the revised manuscript for additional changes (e.g., the fusion of Sect. 3.3 and Sect. 4.2 into Sect. 4.2).

Additional comments

Line 6: delete comma after "understood"

Done.

Line 10: delete "closely" or rephrase

Done.

Line 39-40: Citation for this statement?

Line 43-44: What are the actual length scales of these smaller bands (also a citation speculating these processes would be nice).

We rewrote the Introduction due to the comments from Reviewer 2. Therefore this paragraph is no longer included.

Line 53-54: Why the long dashes?

Removed.

Line 55-56: Citation showing the P3 scheme still struggles with ice processes (I get there are still broad concerns but a citation would be good when specifying a specific microphysics scheme)?

We included one example study:

> Their representations are therefore likely incomplete, even in sophisticated cloud microphysics schemes (e.g., Cao et al., 2023), such as the predicted particle properties (P3) scheme proposed by Morrison, Milbrandt (2015)

Line 63: "space-borne radar" is more commonly accepted nomenclature.

The sentence was removed in the revised introduction.

Line 65-66: should specify why measurements of IWC remain challenging (since the ensuing text implies a synergistic remote sensing/in situ method reduces uncertainty in IWC, which is misleading).

We rewrote the Introduction due to the comments from Reviewer 2. This should be clearer now with the following paragraphs:

> Accurate in situ measurements of IWC remain challenging (Heymsfield et al., 2010; Baumgardner et al., 2017; Tridon et al., 2019), even though in situ cloud probes can provide reliable particle size distribution (PSD) data (Korolev et al., 2013; Moser et al., 2023). Lacking IWC measurements, Deng et al. (2024) calculated IWC from PSD observations assuming that ice particle mass as a function of ice particle size follows a power law relation . Because deriving size-resolved ice particle densities from in situ PSD alone is not possible yet (to our knowledge), Deng et al. (2024) used constant mass-size parameter from Heymsfield et al. (2010). Therefore, their analyses captures IWC variability due to ice number concentration and size, but not ice particle density, which is commonly linked to riming (Erfani, Mitchell, 2017; Seifert et al., 2019) .
>
> Combining collocated cloud radar and in situ PSD data allows to estimate IWC by not only showing great potential to gain better insight on microphysical processes (Nguyen et al., 2022; Mróz et al., 2021), but also to infer ice particle density changes due to riming (Maherndl et al., 2024). This way, IWC variability driven by riming-induced changes in ice particle density can

be studied. In recent years, the synergistic employment of both remote sensing and in situ instrumentation during airborne campaigns has become more common (Houze et al., 2017; McMurdie et al., 2022; Nguyen et al., 2022; Kirschler et al., 2023; Sorooshian et al., 2023; Wendisch et al., 2024; Maherndl et al., 2024).

Line 94: You never define normalized rime mass. Please do.

We now introduce and define the normalized rime mass M in Sect. 3.1.:

We ~~quantify riming using the two methods introduced in Maherndl et al. (2024). First, the combined method derives~~ use the normalized rime mass M ~~from a closure of in situ PSDs and collocated radar reflectivity Z_e. Second, the in situ method uses in situ measurements of ice particle area A, perimeter P, and D_{max} to derive~~ (Seifert et al., 2019) to describe riming. M ~~for individual ice particles from which an average~~ is defined as the particle's rime mass m_{rime} divided by the mass of a size-equivalent spherical graupel particle m_g, where we assume a rime density of $\rho_{rime} = 700$ kg m^{-3}:

$$M = \frac{m_{\text{rime}}}{m_g} \tag{3}$$

where

$$m_g = \frac{\pi}{6} \rho_{\text{rime}} D_{\text{max}}^3. \tag{4}$$

The maximum dimension D_{max} is defined as the diameter of the smallest circle encompassing the cloud particle in m and is used to parameterize particle sizes.

Line 95: I'm not sure what you mean "by closure" (this was also said in the abstract). Please specify.

We removed this sentence from the introduction. Instead, we describe the method we use to retrieve riming in more detail in Sect. 3.1:

We retrieve M ~~for the particle population is derived~~using the two methods introduced in Maherndl et al. (2024), which are termed the *combined method* and the *in situ method*. The methods in Maherndl et al. (2024) were developed for HALO-(AC)^3, but we apply ~~the same methods~~ them to IMPACTS data with slight adjustments due to different instrumentation. ~~For~~ In the following, we give a brief explanation of both methods and describe the adjustments for IMPACTS data. For more detail, we refer the reader to Maherndl et al. (2024).

> The combined method derives M along the flight track of the in situ airplane from collocated PSD and radar reflectivity Z_e measurements. It therefore relies on collocated in situ and remote sensing flights. An Optimal Estimation (Rodgers, 2000) algorithm is used to retrieve M by matching simulated radar reflectivities Z_e obtained from observed in situ PSD with the spatially and temporally closest measured Z_e. As forward operator we use the Passive and Active Microwave radiative TRAnsfer tool (PAMTRA, Mech et al., 2020) which includes empirical relationships Maherndl et al. (2023b) for estimating particle scattering properties as a function of M. For IMPACTS, the combined method is applied (separately) to X-, Ku-, Ka- and W-band Z_e (see Sect. 4.1.3). As in Maherndl et al. (2024), we use the riming dependent mass-size parameter relation for dendrites from Maherndl et al. (2023b) that were estimated for different degrees of riming, i.e., M values. Dendrites were chosen, because 86.2 % of data during the analyzed IMPACTS segments are within temperature ranges of -20 °C to -10 °C and -5 °C to 0 °C, where plate-like growth of ice crystals is preferred (only 13.8 % of the data lie between -10 °C and -5 °C, where column-like growth dominates). We assume dendrite shapes for the whole dataset, because of two reasons. First, Maherndl et al. (2024) found assuming plates or dendrites gives the same results within uncertainty estimates, and second, we want to keep the analysis of IMPACTS and HALO-(AC)³ data as consistent as possible.

In the abstract we include:

> We derive riming and IWC by combining cloud radar and in situ measurements.

Line 117: rephrase "... and sampled at different frequency rates producing different spatial resolutions" or something similar.

Done.

Line 118: change "fly" to "flew"

Done.

Line 121: What does "good collocation" mean?

We included:

> We selected these days because of the good collocation (which we define as maximum spatial offsets of 5 km and temporal offsets of 5 min; see Sect. 2.4) between the respective remote sensing and in situ aircraft as well as the data availability.

Line 142-143: Why is a CDP and a Fast-CDP used? Was one in error for different flights? Also, what are these probes used for exactly? Is it for PSD measurements the radar uses for calibration? Results from these probes aren't shown anywhere in the paper.

For retrieving M, we use combined particle size distribution (PSD) data from the respective campaign (Bansemer et al., 2022; Moser et al., 2023), which are derived from the listed instruments including a CDP and a Fast-CDP for the size range below 50 µm for HALO-(AC)³ and IMPACTS, respectively. We assume the cloud particles in this size range to be liquid for the scattering simulations in the M retrieval. We only use the Fast-CDP for IMPACTS, we removed the erroneous double-mention:

> For IMPACTS, we use data from a ~~Cloud Droplet Probe (CDP, Lance et al., 2010) and a Fast-CDP (2-50)~~Fast-Cloud Droplet Probe (Fast-CDP, 2-50 µm, Lawson et al., 2017), a Two-Dimensional Stereo (2D-S, Lawson et al., 2006) probe (10-2000 µm, pixel resolution of 10 µm), one horizontally, and one vertically oriented High Volume Precipitation Spectrometer, version 3, (HVPS-3, Lawson et al., 1998) ~~probes~~ probe (0.3-19.2 mm, pixel resolution of 150 µm).

Line 149: Although understood to be somewhat common to assume 50um is all ice, it is possible droplets can get much larger than this. While the potential of icing is often the rationale for this assumption, kinetic heating of the aircraft can avoid icing at temperatures a few degrees less than 0C. In fact, I wonder if the large ice particle concentrations in IMPACTS might actually be large drops in the -5 to 0C range. To test this by doing a temperature dependent sensitivity test, I'm curious whether results overall might be sensitive to temperature ranges (possibly not, since you do the height analysis in Appendix B, but might be worth checking).

We agree that analyzing the temperature or height dependence of our results would be very interesting. However, we are limited to few flight segments and performing the pair correlation analysis for given temperature bins, reduces the data amount such that results are no longer trustworthy (too little data for statistical significance). To best remove periods with large droplets from our analysis, we did the following: we only include temperatures lower -1°C. In addition, we manually looked through cloud probe images for each segment and removed two IMPACTS segments with collocated data resulting in the 13 presented segments. We previously did not state this in the manuscript and therefore included:

> As in Maherndl et al. (2024), we only include data up to -1 °C to avoid melting effects. In addition, we manually looked through in situ images of all analyzed flight segments and removed two IMPACTS segments, where we could identify supercooled droplets larger 50 µm.

Line 179: the collocation of radar and in situ measurements can be as far as 5 km off? That seems pretty significant based on the spatial scales you're using the pcf analysis.

Yes, 5 km is the maximum spatial offset. On average the collocation is much closer with mean offsets below 2 km for both IMPACTS and HALO-(AC)³ segments. Fig. R.4 shows histograms of the horizontal distance between in situ and radar aircraft for IMPACTS and HALO-(AC)³. We use the same collocation criteria as in Maherndl et al. (2024), where we found that the standard deviation of Z_e over the average offset distances is smaller than the Z_e uncertainty of 1.5 dB assumed in the M retrieval (combined method). Ideally, in future studies when more collocated airborne radar and in situ data is available, the collocation criteria should be made stricter.

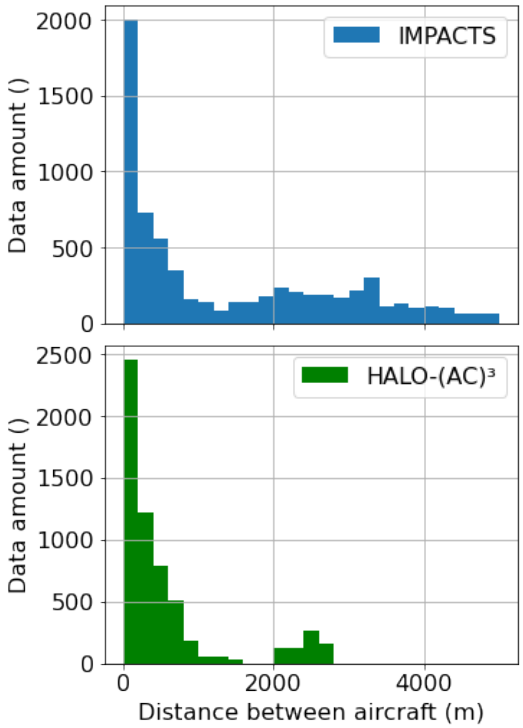

Figure R.4: Histograms of the horizontal distance in m between in situ and radar aircraft for IMPACTS (top) and HALO-(AC)³ (bottom).

Line 191-192: Perhaps I'm confused of what M really is, but isn't it possible to obtain a sum of total M over the particle population? Unsure why an average M is being obtained.

Using the combined method, only an reflectivity-weighted average M can be obtained. We included a more detailed explanation in Sect. 3.1 (see answer to comment about line 95) and kindly refer to Maherndl et al. (2024) for more details. We hope that the inclusion of the definition of M also helps.

Line 202: What are synthetic rimed aggregates?

Here, we meant the data set of simulated rimed aggregates from Maherndl et al. (2023a). To avoid confusion, we removed the term "synthetic" and only use "simulated" in the

revised version.

> ~~Synthetic rimed aggregates~~ Simulated rimed aggregates from Maherndl et al. (2023a) are used to derive empirical functions relating χ and ~~D_{max}~~ $D_{\max}$ to M, where χ and ~~D_{max}~~ $D_{\max}$ are derived using the same processing steps as for the respective cloud probes.

Equation 2: Are results being binned at some specified length(s) (i.e., should r be r+dr)?

Yes, but only when averaging η. There, we use the average flight speed of 200 m/s and 60 m/s for IMPACTS and HALO-(AC)³ respectively to bin into 200 m and 60 m bins. We kept the equation as is, but added:

> To perform the averaging, we bin η into 200 m and 60 m bins for IMPACTS and HALO-(AC)³, respectively, which corresponds to the respective distances covered in 1 s for the respective typical flight speeds.

Line 225: Define gaps. I'm actually unsure of what your in-cloud threshold is.

We use the radar sensitivity limits to define "in-cloud", meaning that when the radar sees a signal, we assume there is a cloud. By gaps, we mean measurement gaps, i.e., NaN values for radar reflectivity or in situ PSD. We added in Sect. 4.3.2:

> In this study, only straight flight segments with a minimum of 200 s of continuous in-cloudmeasurements are used to calculate $\eta(r)$. ~~We allow~~ The respective radar sensitivity limits are used to define "in-cloud". We allow measurement gaps with a maximum length of 5 s, which are linearly interpolated.

Line 257: "...and second, on spatial..."

Done.

Line 277: "...particles larger than 50 um..."

Done.

Line 280: Isn't this the definition of effective diameter (area weighted mean diameter)? Are these properties equivalent?

To our knowledge, there are different definitions of effective diameter as discussed in McFarquhar & Heymsfield, (1998). Here, we use the definition of "mass-weighted" mean diameter from Maahn et al. (2015).

Line 314: "...values associated with large particles..."

Done.

Line 321: "parameters"

Done.

Line 326: Did you mean by three orders of magnitude?

Yes, apologies. We changed to:

> By changing D_m from 1 to 8 mm, IWC changes by ~~four~~ three orders of magnitude.

Line 336-337: What are the observed ranges of M? Reflectivity seems to vary by up to 30 dBZ.

Observed ranges of M are shown in Fig. 5 and shaded in Fig.6. We added:

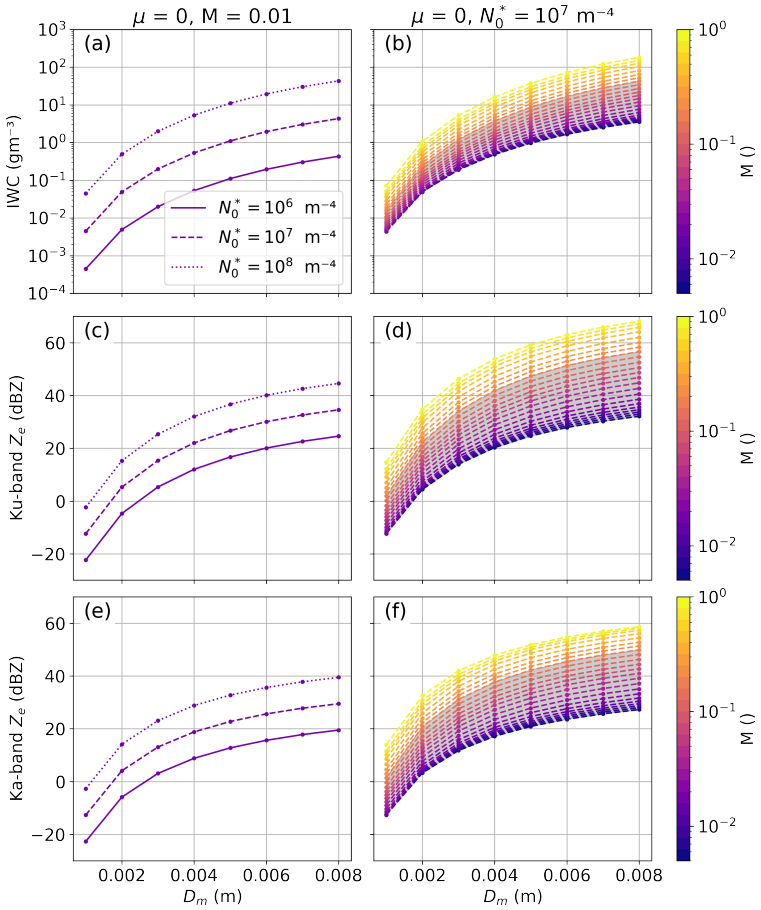

Figure R.5: Ice water content (IWC) (top), Ku-band Z_e (middle), and Ka-band Z_e (bottom) calculated from gamma particle size distributions as functions of D_m parameter. Results for varying N_0^* parameter are shown as solid and dashed lines in (a), (c), (e); for varying normalized rime mass M are color-coded in (b), (d), (f). Shaded areas in (b), (d), (f), (h) indicate M ranges observed during IMPACTS (90 % range: $0.005 < M < 0.15$).

Line 337: "dBZ"

Because we refer to differences, we mean dB not dBZ.

Line 343-347: What is meant by riming being "minimal" while also increasing IWC by about 2/3s?

This was misleading. We rephrased to:

We therefore conclude that ~~even at low amounts of riming, as were~~ for the

range of M observed during HALO-(AC)³ and IMPACTS, the effect of riming on IWC should not be neglected ~~and can cause~~ to avoid biases up to one order of magnitude for IWC.

Line 350-353: "Isn't this only true where positive values of pcf(IWCr) overlap this pcf difference?

Yes, we therefore now only show differences where $\eta_{IWC_r} > 0$ as discussed above.

Line 354: "...larger than zero."

Done.

Line 364: "lags"

Done.

Line 403-404: Can portions of sub-segments be resampled?

In principle, yes. However, the sampling is random.

Line 411: King probe LWC results aren't shown correct? If so state it.

Yes, sorry for the confusing presentation. We included:

> King probe-measured LWC ~~clusters behave similarly , increasing~~ cluster scales behave similarly to N_i (not shown) and maximum cluster scales increase from 0.6 km to 3.0 km.

Line 411: "suggests" not "indicates"

Done.

Line 412: "...supersaturation with respect to ice..." although this is somewhat dependent on whether you are within that -5 to 0C range or at colder temperatures.

Here, we refer to the HALO-(AC)³ results, where temperatures are lower -5 °C.

Line 414-416: It's difficult to tell how significant the differences are between IWCr and IWCu. I get doing the difference to highlight this but the subpanels for the respective quantities' pcfs are nearly identical (also seen in Figure 7). I think doing some statistical robustness testing would sell this point.

See answer to the second major comment.

Line 419-420: Again, LWC is not shown correct? I get the pcf(LWC)=0 contour is shown but the modes aren't. Worth at least specifying the modes are similar.

See answer above.

Line 427-428: Citation?

We added:

> The lidar detects small liquid droplets at cloud top, which follow vertical motions, therefore leading to higher CTH in updraft regions (Abel et al., 2017) .

Line 435-436: Please refer to the M panels in Figure 10.

We included:

> Given that the least (most) amount of riming (Fig. 10c,f,i) occurred on 4 (1) April, we conclude that in the studied MCAO clouds mesoscale updraft features likely enhance riming at spatial scales of 3-5 km.

Line 447: "deposition" not condensation.

Thanks, we changed to:

> Ice particles grow through ~~condensational~~ depositional growth and riming, which leads to enhanced probabilities of ice clusters at these scales.

Line 444-446 & Line 449-450: where did you show these IWC clustering results?

These results are shown in Fig. 9h and discussed in Sect. 4.3.2. Fig. 9h shows that riming influences IWC clustering at two spatial scales: 1. riming increases the probability of clustering at scales below 1-2 km, in the same range as the roll cloud circulation and updraft features found by Schirmacher, et al. (2024), and 2. riming leads to additional clustering at 3-5 km. In the revised manuscript, we added a reference to the figure:

> In the presence of additional mesoscale updraft features, IWC clusters also occur at spatial scales of 3-5 km (Fig. 9h).

Line 465: "...from cloud top."

Done.

Line 480: "...larger than 6 km..."

Done.

Appendices: would be nice to keep the figures within the respective appendices. Figure B1&C1: would be nice to show altitude or some sort of normalized cloud height rather

than distance below the higher aircraft (unless this higher aircraft is essentially flying at constant altitudes).

The figures are now within their respective appendix sections. Regarding Fig. B1 and C1: Here, we do show the distance to cloud top, not to the higher aircraft. We include the "radar" in brackets, because here we use cloud top height as derived by radar.

Additional References

McFarquhar, G. M., and A. J. Heymsfield, 1998: The Definition and Significance of an Effective Radius for Ice Clouds. J. Atmos. Sci., 55, 2039–2052, https://doi.org/10.1175/1520-0469(1998)055<2039:TDASOA>2.0.CO;2.