Quantifying riming from airborne data during HALO-(AC)^3
Response to the reviewers

Nina Maherndl, Manuel Moser, Johannes Lucke, Mario Mech,
Nils Risse, Imke Schirmacher, and Maximilian Maahn

December 8, 2023

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 made slight changes to improve both methods, which resulted in a shift to slightly higher M values for both methods without changing the overall shape (location of maxima and minima) of M results. Here, we briefly describe the changes implemented. During a calibration of the MiRAC radar against ground based radars at the AWIPEV station in Ny-Alesund, it was found that MiRAC underestimated Z_e by 2 dB. Therefore the Optimal Estimation retrieval of the combined method was rerun with the corrected Z_e. We did not note this change in the manuscript, because the corrected Z_e will be published on PANGAEA instead of the original data. Further, we improved the in situ method by implementing thresholds for equations Eq. (6) to better account for unrealistically high or low $\log_{10}(M)$ results. Very low or very high $\log_{10}(M)$ can occur due to e.g. measured χ values outside of the range covered by the simulated rimed aggregates, we used to derive the relations. We added:

> χ is calculated from CIP and PIP measured P and A for each detected particle. M is then calculated from $D_{\max}$ and χ for each particle~~:~~. To avoid unrealistic values, we set all $\log_{10}(M) > 0$ to 0 and all $\log_{10}(M) < -3.5$ to -3.5. The latter threshold is chosen based on the minimum M of the combined method results.

2 Reviewer II

2.1 Summary

This manuscript presents two new methods to quantify riming on ice particles using the normal- ized rime mass M from airborne in situ and radar measurements. The first method combines in situ and remotely sensed radar observations in the configuration such that the radar carrying aircraft is overflying the in-situ probe carrying aircraft, while the second approach is based on in situ observations only (less demanding in terms of aircraft flight pattern). The two methods are shown to produce similar estimates of the normalized rime mass over the data collected during the HALO-(AC)3 campaign that took place in the Arctic near Svalbard in Spring 2022, in a statistical sense. Two case study are further investigated. The first one (based on combined radar and in situ data) suggests that the the regions characterized by a higher normalized rime mass are related to regions exhibiting higher reflectivity values. The second one (based on in situ data only) illustrates that riming may occur in regions with low liquid water path and hence suggests that riming may occur in layers above, containing more liquid water.

2.2 Recommendation

The manuscript is clear, the methods are properly described, as well as the associated assumptions and limitations. The topic is of interest to the community and readership of AMT. There are however a few questions/issues to be clarified (see list below), and I recommend to send the manuscript back to the authors for major revisions.

We thank the reviewer for the positive review and the constructive comments.

2.3 General comments

1. My main concern about the evaluation of the two methods is related to Fig.5 and also Fig.7 to some extent. My take from the scatter diagram in Fig.5 and the two curves corresponding to the two methods in Fig.7 is that the two methods agree on the overall shape of the distribution of the M values, in a statistical sense, but are not cofluctuating (also confirmed in case study 1, see Fig.10). In addition, most of the M values are low and the 2 methods do not seem to agree well on the very few high values, leading to a density of points that is not at all aligned with the 1-1 line in Fig.5. In Fig.7, the two methods apparently agree well for the low M value (in log scale...) but again not that much for the large M values. So I am wondering if the two methods are really in good agreement or if the data set is too unbalanced to provide a robust answer. I think *this issue should be clearly discussed in the paper, and the limitations of the comparison performed should be better emphasized.*

Thank you for the comment. We agree that we achieve agreement in a statistical sense, but are not in terms of temporal cofluctuations for all flight segments. We wanted to show under which condition the latter can be achieved using case study 1, but this was not stated clearly enough in the first manuscript version. We restructured Sect. 4, and added more discussion about the evaluation of the two methods throughout. In addition, we added uncertainty estimates in what was formerly Fig. 7 (now Fig. 9). The uncertainties stem from the Optimal Estimation output for the combined method, and 30 s running standard deviation calculation for the in situ method.

2. In section 3.1, it is mentioned (l.217-218) that the prefactor and the exponent of the mass-size relationship are taken for dendrites. What is the influence of this choice on the retrieval of M in clouds with other habits than dendrites? I did not find the discussion on this assumption.

We added the Appendix section A (”Assumption on particle shape”) to discuss the dendrite assumption for both methods. We decided to compare to assuming columns or plates, due to the temperature range of the majority of HALO-(AC)^3 measurements (as is further discussed in this appendix section). We found that for the combined method, results for dendrites and plates agree within the uncertainty estimates. However, in clouds with column habits, we typically overestimate M by assuming dendrites. But only a small percentage of measurements (about 10% for collocated segments) were taken in temperature regimes that favor column- over plate-like shapes. In future studies, we plan to evaluate incorporating particle habit classification results into the retrieval schemes.

3. In section 4.1, the vertical profiles of M are discussed and linked to environmental conditions (temperature, LWC, LWP...) but I am wondering what the uncertainties associated with the retrieved M values (and subsequently on the rimed fraction) are. And in particular if the shapes of the curves are statistically significant. As Sa is taken about 1, my gut feeling (and I may well be wrong) is that the uncertainty associated with small M values (the vast majority of the cases) is relatively large and may induce limited significance. Such uncertainties are displayed in Fig.10 for instance, why not in Fig.7? This would strengthen the analysis of the shapes of those curves (or suggest that those are not statistically significant).

As discussed above, we included uncertainty estimates in the revised figure. The shape of the curves are still significant, especially the minima, we found at temperatures of about -15°C.

4. I am mot sure I understand what it is added value of the case study 2: rimed particles are detected in regions with rather low LWC, therefore there must be layers with higher LWC above, or more generally there is not enough information about the context above the aircraft to draw any solid conclusion. So nothing original here, and I do not think it *is worth being mentioned in the conclusions (see l.493-495). If this is the case, I suggest to remove the 2nd case study.*

We agree, case study 2 was not necessary and we therefore removed it. We thank the reviewer for this suggestion as we feel it really helped us to restructure Sect. 4 and improve readability of the manuscript.

2.4 Specific comments

1. P.5, l.108: what is the influence of the choice of those parameters for the time and space consistency on the optima estimation parameters (e.g. covariances)?

We do not include the uncertainty stemming from non-exact collocation between the aircraft in the Optimal Estimation retrieval. We added:

> ~~In Appendix A~~Uncertainties due to non-exact collocation between Polar 5 and 6 are neglected here. The average standard deviation of Z_e is 0.7 dB over distances of 555 m, which corresponds to the mean horizontal distance between the aircraft, and therefore smaller than the assumed uncertainty of 1.5 dB. In Sect. 4 we discuss implications of the non-exact collocation on the presented results.

2. P.7, l.171-175: I did not understand how the LWC values were estimated along the radar beam (in order to quantify the attenuation), this should be better explained.

We apologize for keeping this analysis step confusingly brief. We extended:

> To estimate attenuation due to liquid water, we took LWC measurements from the Nevzorov probe operated onboard Polar 6 during the temporally closest vertical cloud profile. To obtain information on the vertical structure of clouds, Polar 6 flew vertical profiles in so-called ”saw-tooth patterns”. These patterns were flown in addition to straight legs at constant altitudes. Saw-tooth patterns are not well suited for good quality collocated measurements with Polar 5, where straight legs are preferred. Therefore a limited number of vertical profiles are available for each flight with collocation. During each flight analyzed in this study, at least three of such ”saw-tooth patterns” were collected. Whenever Nevzorov probe measurements were not available, LWC was calculated by integrating the particle size distribution (PSD) of liquid particles (¡ 50 µm) measured with the cloud probes on board Polar 6. In both cases, LWC measurements were averaged to be on a regular vertical grid with a resolution of 10 m. Here, we neglect the distance traveled by Polar 6 during the profile, assuming LWC to be constant at each height bin. This assumption likely does not hold in reality, however, no measurements

> with more precise information on horizontal and vertical LWC distributions are available.

3. P.7, l.186: I think it should be "for" instead of "to" before "our results".

Thank you.

4. P.9, l.235: is Z_e expressed in dBz or $mm^6 m^{-3}$?

In dBZ, we added:

> ...at Polar 6 flight altitude in dBZ and...

5. P.9, l.237: "to make S_a more Gaussian": maybe showing a distribution (in appendix?) would strengthen the claim?

We decided to remove this statement to avoid confusion. What we wanted to refer to was that the mass size parameter a_m and b_m from Maherndl et al. (2023) are a function of M and that choosing M in a logarithmic scale makes it possible to cover the whole range of mass size parameter (see Fig. 9 (a) from Maherndl et al. (2023)). However, since we only have one state space and one measurement parameter, S_a is reduced to a scalar and and the original sentence does not make much sense.

6. P.9, l.239: given that S_y and S_a are of the same magnitude, does the insensitivity to Sa imply that the 1st term in Eq(3) is dominant and hence that F() is strongly conditioning the retrieved values?

Not really, because S_y and S_a have different units and are therefore not comparable. Additionally, our problem is well constrained and the cost function is expected to have one only minimum in most cases.

7. P.11 l.286: it should be Fig4.b, no?

Yes, thank you.

8. P.13, l.307: it seems that "a" in between "point" and "perfect" should be removed.

Done.

9. P.13, l.312: the RMSE value seems much larger than the mean value, which suggests strong uncertainty no?

We agree that it shows strong uncertainty in terms of temporal confluctuations, however the low ME shows good agreement in a statistical sense. We added:

> Mean error (ME) and root mean square error (RMSE) are ~~-0.0077 and 0.032~~

> 0.0026 and 0.031 for the point by point comparison. While we do not achieve good point by point agreement (large RMSE), both methods agree in a statistical sense (small ME).

10. P.14, l.14: “simIlar”

Done.

11. P. 16, l.360: ‘ ‘Figure 8 analyses the dependence of temperature and LWC on M ”: should it be the other way around?

Yes, we changed the sentence structure.

12. P.18, l.411: I suggest to add “in terms of temporal cofluctuations” after “agreement” to clearly emphasize on what this agreement is.

Good point, done.

13. P.20, Fig.10: the dashed line in plots (b), (d)... is not explained in the caption.

Thanks, we added the explanation in the caption.

14. P.23, l.498: “depended”: should it be “depending on the”?

Yes, changed.

15. P.23, l.513: the units of N and N_0 should be mm^{-3}mm^{-1}, as $N(D)dD$ is the concentration of drops of size between D and $D + dD$. This is consistent with the definition of Λ 3 lines after (in its current version, Λ would be dimensionless).

Yes, thank you for catching the mistake. We changed:

> To approximate errors of the combined method M retrieval, we present results obtained for synthetic data. We use the simulated rimed dendrite aggregates from Maherndl et al. (2023) binned into 10 logarithmic M bins from 10^{-2} to 10^{0} (”true” M) and linear D_{max} bins from 0 to 10 mm with bin widths of 200 µm. We apply exponential PSDs $N(D) = N_0\ \exp(-\Lambda D)$ to each M bin, where N is the number concentration in ~~m^{-3}~~ m^{-4} of particles of size D in m, the intercept parameter N_0 (in ~~m^{-3}~~ m^{-4}) describes the overall scaling and the slope parameter Λ controls the shape.