Response to Reviewer #2

1. Microphysical retrievals of radar volumes containing a mixture of pristine ice particles and aggregates are challenging, since larger particles tend to dominate the signal. The manuscript presents a method for retrieving PSD parameters separately for crystals and aggregates from polarimetric radar observables based on an ensemble retrieval framework. The framework is constructed using a prior PSD parameter distribution and forward modeled radar observables from the assumed PSD based on scattering database results of a number of different kinds of pristine crystals and aggregates. The method is evaluated first with synthetic observations and then against in-situ aircraft measurements. The in-situ comparisons show an overall improvement over existing methods.

   The text is generally well written and structured. It involves adequate analysis and discussion of related uncertainties and the figures are clear and demonstrative. I expect the presented method to help advance the use of radar polarimetry in studying snow microphysics. I recommend the manuscript to be accepted for publication with minor revisions.

   We thank the reviewer for thoughtful comments.

General comments

2. My only general comment is that I would have liked to see some discussion related to the possibility of taking rimed particles in to account in similar retrievals. Riming may have great significance depending on climate and is expected to have a very similar polarimetric radar fingerprint as aggregation. Do you expect that riming could have affected your evaluation results?

   Some modifications need to be made for the retrieval method to work for rimed particles. We have provided our thoughts in the revised manuscript:

> Line 575–582:
>
> Currently, our method is designed to work for conditions with a mixture of pristine ice and aggregates. In the presence of rimed particles, the state vector should be expanded to include additional variables that can accommodate and inform the degree of rimming, e.g., the riming factor described in Masson et al. (2018), or to include appropriate rimed species explicitly. When triple-frequency measurements are available and can be used to distinguish particle types effectively (e.g., Kneifel et al., 2015; Barrett et al., 2019), such information on particle types can also be incorporated into our method to provide retrievals for off-zenith radar scans that are more challenging for triple-frequency techniques. It is also possible to expand the observation vector with other radar observables at multiple wavelengths, providing further constraints on retrieval if added information exists.

Specific comments

3. 77: Instead of "spheroidal morphology" did you mean to make a statement on the aspect ratios of aggregates?

We meant to talk about their shape and orientation. For clarity, we have replaced "spheroidal morphology" with the following:

> Lines 75–76:
> Snow aggregates yield low Z_{DR} (about 0–0.6 dB; see Hogan et al., 2012) as a result of their sparse and irregular morphology, with the component crystals oriented at a wide range of angles.
>
> Lines 739–740:
> Hogan, R. J., Tian, L., Brown, P. R. A., Westbrook, C., Heymsfield, A. J. and Eastment, J. D.: Radar scattering from ice aggregates using the horizontally aligned oblate spheroid approximation. J. Appl. Meteorol. Clim., 51, 655–671, doi: https://doi.org/10.1175/JAMC-D-11-074.1, 2012.

4. 86: Did you mean that these variables are simply less widely adopted or that there is more work to be done connecting characteristics in the retrievals of these variables to snow processes? Please rephrase

   We apologize for the confusing wording. We meant that ρ_{HV} and K_{DP} observations have not been used extensively in quantitative retrievals. Since the other reviewer also has the same question, and the information provided here was not critical, we have removed this sentence.

5. 451-452: It is not evident to the reader what kind of temperature dip we are talking about since it seems to be excluded from the figure and not described here.

   Apologies. We excluded the dip in the figure, so have reworded the text to explain the data gap:

> Lines 466–468:
> The flight height was maintained at ~2 km from 6:30 to 6:40 UTC, suggesting that the missing temperatures due to a data glitch at ~6:40 UTC are likely to be about –5°C.

6. 510: I'm not sure if I understood this sentence. Did you mean that these radar signatures might represent only a subset of the aircraft-collected sample? Or that there might be a spatial mismatch? Please rephrase and discuss the possible implications.

   Thanks for pointing out the possibility of a spatial mismatch, which has helped us to look more carefully about the collocation.

   To understand how well the collocation is, we have produced 2D histograms of occurrences of the vertical and horizontal distance in the collocated in-situ and radar dataset, as shown in Fig. R2. The distance was calculated with respect to radar gate, i.e., the positive vertical distance represents that the flight altitude is higher than the radar gate of interest.

   Checking the collocation in Fig. R2, we found that in-situ samples were taken largely at radar scan heights or below in Clusters 1, 2 and 6. It is likely that both in-situ and radar have sampled the same regime with notable aggregations, and that's why the observed D_{eff} is close to the retrieved D_{eff} of aggregates. In contrast, in-situ samples were taken at higher altitudes over the radar scans in

Clusters 3–5. We believe that in Clusters 3–5, aircraft may have sampled a pristine ice growth zone aloft, but the radar gates below sampled the subsequent aggregations, which explains why the observed D_{eff} is closer to the retrieved D_{eff} of pristine ice, rather than aggregates.

In the revised manuscript, we now have included Fig. R2 as Fig. 11 and added the following text:

Lines 528–536:
The third scenario is that the discrepancy in D_{eff} is due to a sampling issue. Figure 11 shows two-dimensional histograms of occurrences of the vertical and horizontal distance in the collocated in-situ and radar dataset. The distance was calculated with respect to radar gate, i.e., the positive vertical distance represents that the flight altitude is higher than the radar gate of interest. Interestingly, for Clusters 1, 2 and 6, in-situ samples were taken largely at radar scan heights or below. It is likely that both in-situ and radar have sampled the same regime with notable aggregations, which explains why the observed D_{eff} is close to the retrieved D_{eff} of aggregates. In contrast, in-situ samples were taken at higher altitudes over the radar scans for Clusters 3–5. In these cases, aircraft may have sampled a pristine ice growth zone aloft, but the radar gates below sampled the subsequent aggregations, which explains why the observed D_{eff} is closer to the retrieved D_{eff} of pristine ice, rather than aggregates. Further studies using more datasets and retrievals would be needed to assess the third scenario.

Lines 569–574 in the Summary Section:
... In other clusters, the observed effective mean diameters agree better with the retrieved size of pristine ice, likely because the aircraft sampled pristine ice growth zones aloft instead of aggregation zones that radar sampled. Since planar crystal growth and subsequent aggregation can lead to zones with distinct ice bulk properties, taking frequent aircraft measurements at multiple vertical layers around the radar location would be particularly helpful to improve collocations and allow us to analyse individual rays in more detail.

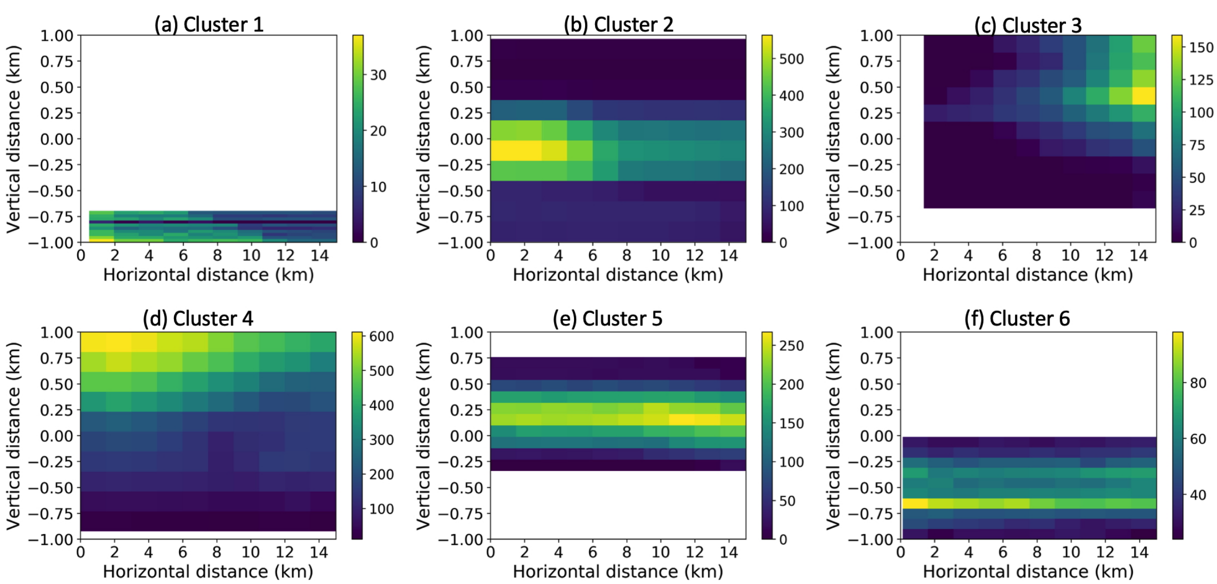

Figure R2: *2D histograms of occurrences of distances in the vertical and horizontal between in-situ measurements and radar gates for Cluster 1–6 in (a)–(f), respectively. Note that occurrences are counted for all pairs of in-situ data point and radar gate. In calculations of retrieval errors, selected in-situ data points and radar gates are only used once with equal weights.*

Technical comments

7. 424-426: This could be rephrased to avoid repetition.

   Thank you. We have rephrased it to the following:

   > Lines 437–438:
   > (k) and (l) represent the individual and combined effective mean diameters using the maximum particle dimension and the equivalent melted particle as the size descriptor, respectively.

8. 483: (f) should be (d).

   Done – corrected (f) to (d) on Line 501.

9. 512: particles

   The sentence (on Line 536) has been rewritten, and thus this error is removed.