

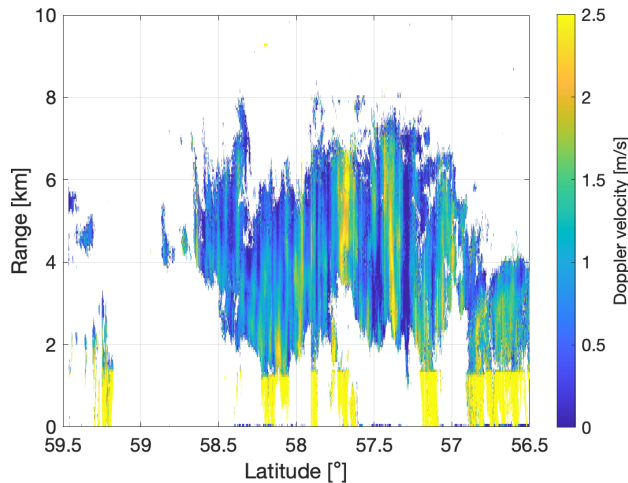
Reviewer 2:

This manuscript provides a closure study on cross-instrument consistency of the observed ice microphysics from in-situ and inferred ice microphysics from remote sensing measurements using suborbital campaign collected data. The main science goal is to understand to what degree the V-H polarization difference (PD) signal from sub-mm channels, in particular, 243 and 664 GHz from the ISMAR instrument, is induced by oriented ice particles that are realistically observed by in-situ cloud probes and Ka band radars. With the fully polarized simulation realized by ARTS and its scattering database using the observed particle shapes, the authors found the 243 GHz PD and TB can be largely reproduced in different cloud regimes, but the largest PD signals have to involve a few percent (10-50% of the bottom layer) of dendrite monomers. Given the identical set-up, the simulated 664 GHz PDs however are too large and TBv is too warm. The authors speculated several possible reasons to explain such a discrepancy (too noisy; ice might not be dominantly horizontally oriented; particle habit incorrect, etc.) Overall this is a very nice and informative paper that I strongly encourage publication on AMT. The experiment design was thoughtfully crafted to make sure to use as much as collocated data as possible, and the level of details paid toward the execution and documentation are highly appreciated. The major conclusions are solidly supported by evidence. Before publication, I think some minor issues can be fixed or improved. I'll explicitly say "optional" if the additional work is not necessary to complete this paper, but would be otherwise "nice to have" to help enhance the science impact of this paper.

Major suggestions:

1. The falling speed (i.e., terminal velocity) should be different for dendrite monomer vs. dendrite aggregates. You have W-band Doppler radar but the vertical velocity data were never used. Can you check to validate your finding using the Doppler velocity?

Thank you for this suggestion. We have looked at the velocity data (see below) but there is a lot of spatial variability, which is a result of fluctuations in vertical air motion, and this makes it difficult to identify the microphysically-driven differences from the observed data. Since the best results are found by replacing only about 10% of the layer with single crystals (e.g. Fig. 9), their impact on the net Doppler velocity is very marginal, and any such signal is swamped by other heterogeneities in the field (as is evident in the figure).



2. I scrutinized your Fig. 7d and found for TBv in the range of 245 to 235 K, there are two groups of PDs. The larger PD group seems to correspond to latitude = 58 to 58.5 deg and latitude = 57.5 – 57.2 deg. You have ground precipitation radar (Fig. 1) that you can check against, e.g., Kdp, Zdr, for orientation signal, as well as the connection to surface precipitation type and intensity. It's very interesting that we can see the cloud touches ground in these two latitude bands (i.e., precipitation) that I don't know if can help you infer more connections between oriented dendrite (as opposed to dendrite aggregates) and surface precipitation properties.

Unfortunately, the dual-polarisation upgrade to the Stornoway radar wasn't completed until 2017 so these data aren't available for this flight.

3. For interpretation of the 664 GHz PD discrepancy, I think other than instrument noise issue (that I had an impression was fixed for 664 GHz for some other flights?), the observed TB-PD relationship is much more scattered compared to the tight 243 GHz relationship. For smaller particles up in Layer 1-3, they tend to be less impacted by the aerodynamics but more by temperature and humidity. This is supported by findings using CALIPSO lidar, which sees much fewer oriented ice than microwave sensors (e.g., Noel and Chepfer, 2010; Zhou et al., 2012). References: Zhou et al. (2012): <https://doi.org/10.1175/JAMC-D-11-0265.1> Noel and Chepfer (2010): <https://doi.org/10.1029/2009JD012365>

The 664GHz instrument noise issue hasn't gone away sadly (some of the bias issues have improved, but we applied a correction to this flight for that anyway).

Thank you for the references, it is a very interesting and complicated issue!

664 GHz is more biased towards the cold upper atmosphere, and we agree that the small particles in colder clouds have a weaker polarisation signal. Your suggestion is in agreement with our suggestion in the paper that the upper-layers have more quasi-random orientation. We have added the references you provided to the discussion.

4. The scan pattern of the in-situ cloud probes can be better elaborated. For example, you can overlay the flight level of the aircraft carrying the cloud probes.

If the particles are collected at different altitudes and cloud regimes throughout this leg, what lead you to think the columnar and dendritic aggregates are representative?

OK, we have added Fig. 1b, showing the altitude and latitude of the FAAM aircraft during the time of interest.

We used particle imagery from times of maximum IWC in each layer to choose the particle habits. There is a time difference between when the remote sensing and in-situ measurements were made, so there is definitely some uncertainty about whether the imaged particle habits were also present when the remote sensing measurements were made during the straight and level run. A more optimal set-up would be to sample both at the same time, but that data was not available for this case study.

5. For Fig. 10 and related text on 664 GHz discrepancy, I'd suggest you carry out a 100% random orientation simulation for columnar aggregates, but 100% horizontal orientation for dendrite aggregates below, as a reference to your possible TBv and PD range. That would be helpful to support some of your arguments.

We agree that this would be interesting, but since the computations are not currently set up in this way, it is beyond the scope of our present work. We have added to the discussion that it would be interesting to explore the sensitivity of the simulations to the degree of orientation of the particles.

6. I have some doubts regarding using 100 um as the cut-off. Although I agree with you that they shouldn't contribute much to the PD signal, 243 and especially 664 GHz are still sensitive to particles smaller than 100 um. See one of my simulations (not exactly the same frequency but close) below also using TC4 and 100% oriented columnar aggregates. The possibility on 664 GHz discrepancy induced by this cut-off should be discussed in the context.

OK, we have added this to the discussion.

Minor suggestions: Section 2.2, paragraph 1: some summary on in-situ instrument limitations and retrieval uncertainties needed.

As well as pointing out the major limitation that the in-situ and remote sensing measurements were not obtained at the same time, we have expanded on sizing uncertainties from the OAPs, and included the information below on the Nevzorov probe.

Line 208-209: are the Nevzorov probe sensitive to the same size range of ice particles compared to your Ka and/or W band radars?

The Nevzorov is OK for particles up to 4mm in size ([Improved Airborne Hot-Wire Measurements of Ice Water Content in Clouds in: Journal of Atmospheric and Oceanic Technology Volume 30 Issue 9 \(2013\) \(ametsoc.org\)](#)). It will still respond to larger particles, but may under or over-read depending on how they shatter and interact with the airflow. We have now mentioned this in the text.

[optional] IMA as an approximation to DDA: Since you have to use DDA to simulate dendrite monomer scattering, there's a discomfort feeling of inconsistency here. Although you cited your previous paper on demonstrating that IMA is a good approximation, is it for the same particle shape and size range and frequency? It's better to add a baseline comparison for DDA result compared to IMA for a simplified setting here.

We did tests comparing IMA and DDA for a simplified setting at 243GHz, and with smaller particles at 664 GHz, described in the second and third paragraphs of section 3.3. Although DDA for monomers seems inconsistent, it is actually not because IMA uses DDA for the monomers.

Line 252: explain the meaning of “k”.

OK, we have done this.

Line 372-375: This paragraph needs clarification. Why don't you use direct trajectory as the timeseries instead of using the latitude bin? I understand the original timestamp will lead to too many noise and non-robust signal, but your design is only valid when assuming the clouds don't change within a given latitude bin. Does this flight scans back and forth on the same trajectory which is perpendicular to latitudes? This was never explained very well and maybe it should be explained in more detail when you introduce Fig. 1.

The main issue here that the different instruments are not on the same platform. The HALO aircraft travels faster than the FAAM aircraft (I believe the speeds to be approximately 166 m/s for FAAM and 226 m/s for HALO). At the start (59.5-58.4 degrees latitude) the ISMAR sees a given latitude before the radar, then around 58.4 degrees they match and see the same latitude at the same time, then after that the radar sees a particular latitude before the ISMAR. Thus, we cannot compare time series so we choose to use latitude bins. We have added this detail to the text.

Line 512: this is contradictory to your statement in Line 425.

Yes, we have rephrased this.