We thank the reviewer for their comments, and our responses are given below.

Concerning the presented radiative transfer simulations with different particle types, I do not fully agree with the authors that the measured and simulated brightness temperatures are in the same range and represent the same variability. Strictly I would say, this is only the case for a few frequencies or parts of the flight legs. Although, they mention it is not within the scope of the study, one could consider varying the particle type along the legs or when flying in different altitudes as indicated by the particle images. The extensive description of the particle habits in the database indicates the possibility of doing so. Here the shape information of the insitu probes could have been taken more into account.

In summary, the closure study did not succeed to find a match between observations and simulations over the whole measured spectrum and IWP range presented here.

We agree that no single particle habit is capable of reproducing all the observations at different frequencies and locations, and this may partly be caused by variations in particle habit spatially and through the depth of the cloud, combined with the different sensitivities of the various channels considered. However, as discussed in our response to reviewer 1 we do not think that the in-situ observations are representative enough to generate sufficiently accurate layered cloud models. Rather our approach is to use the in-situ observations to eliminate implausible particle habits (through the mass-dimension relationship), and consider whether the remaining shapes are capable of simulating the observed brightness temperatures. However, even with the wide range of particles available in the ARTS database this is not always possible and our attempt to generate a rosette-type particle consistent with the in-situ observations for B895 was not able to reproduce the observations at 664GHz. We will expand the discussion around this in the revised manuscript.

I would recommend using the scan information provided by ISMAR if there are measurements under different angles during these flights. By this the study can be brought closer to ICI and could give information about orientation. It is too bad that interesting receiver channels did not work properly to perform a more in depth investigation of particle orientation.

Please see our response to reviewer 1 regarding using off-nadir scan angles closer to the viewing geometry of ICI. We have focused on the nadir views to be consistent with the lidar viewing direction and to reduce the impact of horizontally aligned particles which are not currently present in the ARTS scattering database. Since the main difference between nadir and off-nadir views is expected to be polarisation effects from oriented particles, detailed consideration of off-nadir views is beyond the capabilities of the current simulations.

448 +/- 1.4 Ghz is left out because the weighting function peaks very high in the atmosphere. Since the flights are very close to the clouds and high in the atmosphere, it might be worth taking them into account, eventhough the signal due to ice particles scattering might be even smaller than in the other channels.

We do not think that including the 448 ± 1.4 GHz results would add anything to the paper as the cloud signals are very small - the simulated brightness temperature depressions for this channel are generally below 2K except for the SectorSnowflake during the initial parts of B939. This channel is also very sensitive to the upper troposhperic water vapour profile which is not well constrained by the available 183GHz observations, leading to larger uncertainties in the clear-sky simulations that are used as a baseline to determine the cloud-induced signals. The plots for the two flights for this channel are shown in figs 1 and 2 for reference; the large scatter in the observations is probably due to uncertainties in the water vapour profile.

The influence of the surface in 243 GHz could be reduced by slanted simulations and observations. Over ocean it should be anyway possible to estimate the influence of the surface to a good degree. Especially in comparison between clear and cloudy sky, surface signal might not play a big role.

At 243GHz, Prigent et al. [2016] showed that the mean error in surface upwelling brighntess temperature using ERA-Interim atmospheric and surface fields and the FASTEM ocean emissivity model is $\sim$4K. Since the clear-sky transmission at 243GHz is $\sim$0.5 at nadir (and $\sim$0.36 at 50 degrees) this could introduce a significant bias in the clear-sky simulations that is of similar magnitude to the expected cloud signals. Although the surface signal may not be significant when calculating simulated cloud-induced brightness temperature depressions (i.e. cloudy simulation - clear sky simulation), the calculation of the observed cloud-induced brightness temperature depression (i.e. cloudy observation - clear sky simulation) requires unbiased clear-sky simulations for the comparison to be representative. For cases over land the surface emissivity may also be rather variable and is not well constrained. We therefore do not feel that the 243GHz results are worth including in the paper. Nevertheless, we include the relevant plots in fig 1 for reference, and it can be seen that the scatter in the observations hides any underlying variations in the cloud signal in response to changing IWP.

The derivation of the profiels of ice water content is not fully clear to me. I would appreciate of (average) profiles or time series of the IWC or IWP as utilized in the radiative transfer could be shown.

The profiles of ice-water content (and hence the integrated ice water path) are estimated from the profiles of lidar extinction using flight-dependent empirical relationships derived from the in-situ observations as described briefly on p15 l16. Note that this ice mass profile is not directly used in the radiative transfer simulations which instead use lidar-derived profiles of the second moment of the PSD to fix the size distribution. However, they are used to calculate the IWP used to plot Fig 9 and Fig 10. We will expand the description of the method and include plots of the lidar-derived IWC (shown below in fig 2) in the revised manuscript.

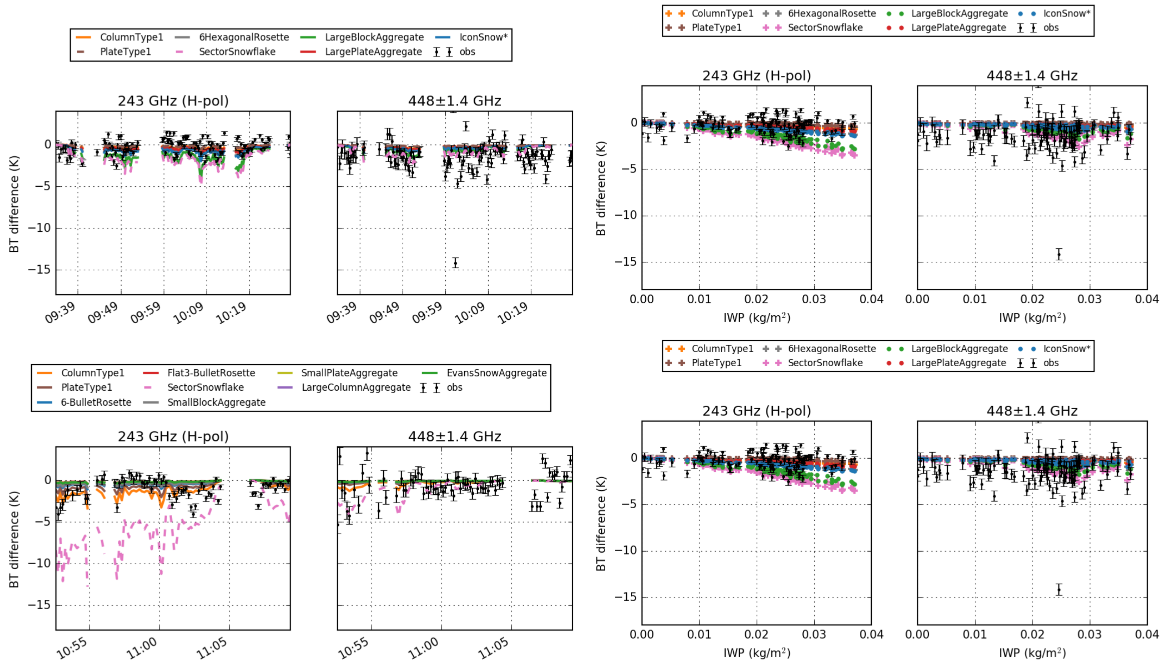

Figure 1: Cloud-induced brightness temperature differences as a function of time and IWP at 243 and 448±1.4 GHz for B895 (top) and B939 (bottom)

To my knowledge, there are coordinated flights of the BAe-146 with ISMAR on board with other aircraft like the HALO carrying water vapor lidar, radar and additional passive microwave instruments. Could these measurements help to constrain further the atmosphere and therefore the vertical distribution of ice water?

A co-ordinated flight between the FAAM BAe-146, HALO and the Safire Falcon was performed in October 2016 bringing together ISMAR, multi-frequency cloud radar, lidar, additional passive microwave instruments and in-situ observations of a deeper cloud system than the cases used in this study. This flight will be the subject of future studies.

References

C. Prigent, F. Aires, D. Wang, S. Fox, and C. Harlow. Sea-surface emissivity parametrization from microwaves to millimetre waves. *Quarterly Journal of the Royal Meteorological Society*, pages n/a–n/a, 2016. ISSN 1477-870X. doi: 10.1002/qj.2953. URL `http://dx.doi.org/10.1002/qj.2953`.

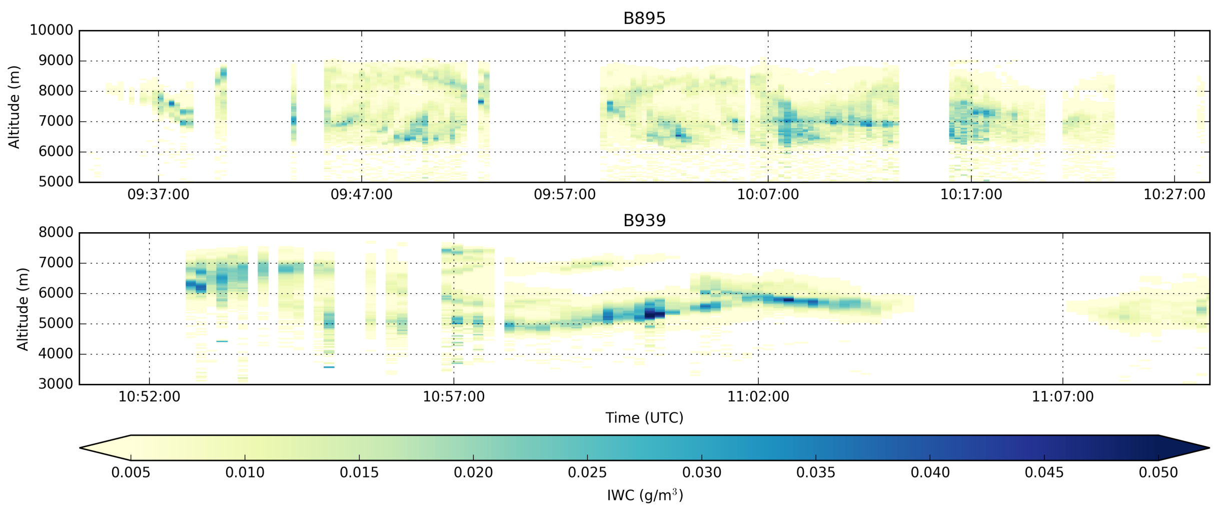

Figure 2: Lidar-derived ice water content profiles for B895 (top) and B939 (bottom)