Reviewer #1:

In this study the authors use detailed multi-frequency radar observations in order to constrain key parameters in a 2-moment bulk microphysics scheme that are important for the parameterization of snow aggregation. The authors examine the problem in detail in a simple 1D context and then expand their tests using a 3D LES model. Overall, I think this is a great paper and a solid piece of scientific work. I liked the authors' initial premise that tuning of physics parameterizations based on "large-scale" results can be deceiving due to the potential for compensating model errors and their strategy of attacking the problem on an observation-based process level study. This paper is an excellent illustration of how to tune (constrain) a microphysics scheme on a process level – which is a difficult task – using observations.

I really do not have any constructive comments to add regarding things that could be improved in the paper, which I think is essentially publishable in its current form. The comments I made below are simply offered as food for thought for the authors, which they may wish to comment in the paper (as they see fit). Overall, great paper.

General comment by the authors to Reviewer #1:

We thank the reviewer for the time and effort in reading the manuscript and providing constructive comments and suggestions. In the following, we address these comments and suggestions point by point. Line numbers refer to the non-revised manuscript.

Point-by-point reply of the authors to the questions and comments raised by reviewer#1

Specific Comments

This study makes comparisons between direct radar measurements (i.e. variables in radar space) to comparable model variables that are computed using instrument simulators. If I understand correctly, the alternative

approach would be to apply retrievals to the radar observations and convert those to fields that are more directly comparable to model fields. Could you please comment on the relative strengths and weaknesses of the two approaches? For modelers, the second approach seems more intuitive. (I may be getting confused with dual-polarization retrievals.)

A: We believe that both retrieval (observation-to-model approach) and forward operation (model-to-observation approach) have their advantages. Retrievals make the comparison in model space possible, which might be more intuitive, as the reviewer suggests.
The main advantages of forward operation might be that uncertainties can be evaluated in the observational space and it might be easier to trace uncertainties arising from the assumptions which have to be made in the forward operator.
An example, that illustrates the advantage of considering uncertainties in the observational space is the interpretation of the biases between model and observation in Figure 11. Although at lower temperatures only small crystal are present, DWRs have a spread of about 1-2dB, which can be due to natural variability and measurement uncertainty. This spread of DWR can be used to asses which deviations are significant and which should not be overinterpreted.

To remind the reader of the motivation why we use the forward operation approach given in Ori et al. 2020, we included some additional sentences at line 81:

"The comparison of model and observations in the radar space using a radar forward operator simplifies the assessment of uncertainties because the deviations between model and observation can be directly compared to the variability of the observation.
The alternative approach of applying a retrieval to the observations might seem more intuitive because microphysical variables, such as number density, can be compared directly. However, assuring consistency between model and retrieval as well as tracing the propagation of uncertainties, for example in the observables or the forward model, is often more complicated to obtain (e.g., Reitter, 2011)."

Could (and should) the approach used in this study be applied to parameterize tendency rates for the spectral width (i.e. for triple-moment treatment of snow) for aggregation (and other processes, such as break-up)? Could you please comment on whether triple-moment snow would improve the representation of the effects of these processes?

A three-moment representation of the snow could be advantageous since the width of the size distribution (spectral width) is a critical parameter for linking the modeled mean mass to the observed dual-wavelength ratios (DWRs). However, the observables considered (reflectivity, DWRs, mean Doppler velocity) mainly constrain the mean values of the particle population, such as integrated mass, mean size, and mean velocity. In contrast, these observables contain only limited constraints on the variance of the particle population, such as size. Therefore, we argue that additional observables from radar (e.g., Doppler spectrum width) or in situ observations should be considered when applying the approach with a three-moment microphysics scheme.

We added the following sentences to the discussion in line 668:

"Therefore, using a microphysical scheme that explicitly simulates the width of the size distribution (e.g., a three-moment scheme) would provide a more consistent link between model and observation. However, additional observational constraints from radar (Doppler spectral width) and in situ observations should be considered in this case."

> In the collection kernels, is there not a slight "break-down" for the situation where the collector and collectee particles have similar sizes, and hence fall speeds, for the analytic solutions (that is, is the kernel values are underestimated)? It was my understanding that that was one of the reasons some schemes use numerical calculations and lookup tables to compute collection rates. On the other hand, I guess the good results summarized in Fig. A1 speak for themselves.

A: In contrast to the Wisner approximation, collision rates are non-zero in the variance approximation of Seifert and Beheng (2006) also if the bulk velocities of the collector and the collectee are the same (this also includes selfcollection). Fig. A1, indeed, demonstrates that the variance approximation deviates only little from the numerical solution, even in this difficult case of similar bulk velocities. However, one can easily imagine that assumptions about the size distribution widths are especially critical in this case. We added the following sentence to motivate better the approximation of Seifert and Beheng (2006) given in Appendix A already in the main text at line 180.

"The latter is made possible by considering the square root of the second moment of the velocity differences, which also has the advantage over the approximation by Wisner (1972) that the collision rates between different particles are non-zero even if their bulk velocities are equal."

Minor Points

1. Line 139 and beyond: Probably at this point you could stop using quotation marks when writing "snowshaft".

A: Done. Thanks for the comment.

2. Line 160: "horizontal resolution" should be "horizontal grid spacing"; "vertical resolution" should be "vertical grid spacing".

A: Corrected.

3. Line 293: On the other hand, the use of look-up tables allows for accurate and numerically efficient run-time integration, both of which are non-trivial advantages.

A: We have added the following sentence to the discussion to illuminate better the advantages and disadvantages of look-up tables at line 178.

"Lookup tables can accurately store precomputed process rates and might be numerically more efficient than analytical solutions, depending on the computer architecture, size of the lookup table, and complexity of the analytical solution."