Referee3

We appreciate your review and critique of the manuscript. Thank you.

This manuscript advertises observational evidence from combined ground-based snowfall rate (S) and airborne W-band radar reflectivity (Z) measurements that rimed frozen hydrometeors are associated with somewhat unique Z-S relationships. These types of studies are desperately need to more accurately characterize the sensitivity of W-band reflectivity to different particle microphysical characteristics, so I laud the authors on their attempts to constrain Z-S relationships for rimed situations using observational assets. My main concern is the lack of data points presented in this analysis - are the results meaningful since the sample size is so small? I am not sure how to suggest solving this issue other than collecting and analysing more data. Conversely, I am very cognizant of how difficult it is to match spatiotemporally disparate datasets like airborne radar to point source measurements of precipitation rates at the ground, so I can appreciate how this study might still be valuable to the community by demonstrating the "atmospheric measurement technique" used so it can be replicated and improved in the future. The manuscript could probably be improved greatly if the narrative leaned more heavily into this aspect of the study. Addressing this issue might be as simple as more forcefully advertising how difficult it is to make such measurements combined with how important it is to collect observational Z-S evidence under rimed conditions in both the introduction and conclusions. I might be able to offer more impactful suggestions in the future when I digest the manuscript again, but I encourage the authors to think about how to creatively make the narrative more impactful.

The revision has improved explanations of the approach we took (Sect. 1); we also improved on descriptions of our method for acquiring S/Z measurements (Sect. 3.5). Discussion of how our S/Z pairs compare to computed S/Z relationships is also revised (Sect. 4 and Sect. 5). In the revised Sect. 5, we added discussion of possible paths for future studies of S/Z relationships. In sum, we think the revised manuscript is improved in terms of how we describe what we did, how we describe our findings, and in terms of our descriptions of future research needed to better refine S/Z relationships for rimed snow particles.

Specific comments:

Introduction: I think it's important to note sooner in the introduction that some of the initial S/Z studies performed for W-band radars were purely modeling (i.e., using backscatter calculations from idealised models of frozen ice habits combined with parametrised particle size distributions) studies. This is a very simple way to accentuate the methodological differences (and importance) of observationally-based studies to assess the veracity of idealised modeling studies.

We added a paragraph to the revised Sect. 1. This encapsulates the connections between our observational approach and the computational work of others.

“The goals of this paper are as follows: 1) to describe measurements of undercatch-corrected liquid-equivalent snowfall rate (S, mm h^{-1}) and how those were paired with W-band measurements of reflectivity (Z, mm^6 m^{-3}) ; 2) to contrast the measurement-based S/Z pairs against calculated S/Z relationships commonly applied in retrievals of S based on reflectivity; and 3) to investigate why the acquired data set deviates from predictions of some calculated S/Z relationships.”

Two further studies of interest (and there are likely more) are Hiley et al. (2011) and Kneifel et al. (2015). Both highlight W-band radar applications for snowfall estimation and also provide analyses that either hint at or explicitly demonstrate how the existence of supercooled water and associated riming complicate Z-S relationships.

When writing the original submission, we were not aware Hiley et al. (2011). The latter is now one of the computational studies we compare to in the revision. Kneifel et al. (2015) is also included in the revision.

Battaglia and Delanoe (2013) and Battaglia and Panegrossi (2020) also demonstrate the global occurrence of snowfall events with supercooled liquid water and Z-S implications. These studies might provide additional context to frame this study’s importance, including W-band attenuation.

The second of these is referenced (revised Sect. 5) because it synergizes lidar, radiometer, and active W-band remote sensing with a views toward retrieving the spatial distribution of supercooled liquid and diagnosing where riming is occurring. Also, the paper's discussion of attenuation helped us in formulating our assessment of attenuation.

I am not very familiar with the hotplate and its history of accurate snowfall rate measurements. While the authors provide some background on previous studies that have been published using hotplates, mostly related to various hotplate precipitation estimates due to various issues (e.g., catch efficiencies, wind speed measurement height, etc.), I still do not see any evidence that this instrument is effective at accurately measuring snowfall rates under various environmental conditions. I would greatly appreciate at least a few more sentences that describe hotplate performance based on previous studies, including uncertainty estimates. No snowfall rate measurement device is perfect, but it would nice to see more details regarding the hotplate since this instrument is such an important component of this study.

The revised Sect. 2.4 includes a description of the measurement precision. This was based on a comparison between the hotplate and the SNOTEL pillow systems (Marlow et al. 2023). The gauge comparison has 57 paired measurements from the HP (hotplate) and SN (SNOTEL pillow) gauges operated at the HP and SN locations in Figs. 1a-b. In the revised Sect. 4, we apply the S precision in a discussion of the departure between our measurements and the computational S/Z relationships. Marlow et al. (2023) was reviewed at AMS/JAMC; we submitted revisions back to the journal two months ago.

Somewhat related to the last point, can the authors further quantify (or at least qualitatively describe) the uncertainties related to their spatiotemporal averaging methodology for both airborne radar and ground-based snowfall rate measurements? What is the sensitivity of the results for slight changes in averaging methodology?

There is discussion of this in the revised manuscript. The following is from Sect. 3.5.

"The HP measurements were averaged over two adjacent 60 s intervals. The first extends from t_O to t_O + 60 s (Fig. 6a) and the second from t_O + 60 s to t_O + 120 s (Fig. 6c). In Fig. 6a and in Fig. 6c, $t_{HP,B}$ symbolizes an interval's beginning time and $t_{HP,E}$ symbolizes an interval's ending time. Formulas describing how these times were related to the beginning and ending time of a corresponding WCR averaging interval are in the Appendix. Fig. 6b is a schematic of the first WCR averaging interval and Fig. 6d is a schematic of the second. Again, the subscripts "B" and "E" are used to indicate averaging beginning and ending times. Figures 6b and 6d both have lines at the top of an averaging interval/domain. The slopes of these lines are proportional to the ratio of two speeds. These speeds are a maximum likely snow particle speed toward the ground (v_p) and a horizontal wind advection speed (v_w). The v_p was calculated using averaged vertical-component Doppler velocities and v_w was calculated using a vertical profile of horizontal winds, based on WKA horizontal wind measurements and AF horizontal wind measurements (Figs. A1a-b), and using the WKA track vector (Table 2). An altitude ($z' = 3400$ m) was assumed in the calculation of v_w . This is the altitude of the ridges west and northwest of the HP site (Figs. 3a-b). Picking the altitude to be either $z' = 3200$ m or $z' = 3600$ m does not alter our findings."

The radar blind zone, and what happens within that layer, is incredibly important. The 200 m WCR blind zone is mentioned in this study in a few locations, but I think the authors need to mention more prominently that a tacit assumption used in this study (similar to a host of other airborne or spaceborne radar studies) is that microphysical evolution within the blind zone could be a major source of uncertainty. I do not recall any studies that conclusively document how rimed particle density evolves in the lowest few hundred meters of the atmosphere – presumably not much – but this is an important to note within this manuscript. It at least warrants a topic that should be studied in the future in the conclusion or discussion sections. It would have been nice to have additional microphysical measurements at the surface to assess the microphysical evolution, but I completely understand how difficult it is to procure instrument suites for fieldwork.

We agree. There is the 200 m deep radar blind zone that encompasses the flight track and the blind zone immediately above the terrain. The latter is a consequence of ground clutter, and in our opinion, is more important for our analysis. Given this, we wrote this in the revised Sect. 5:

"New research can also refine the S/Z relationship for rimed snow particles. This could be computational – exploring the utility of parameterizing S in terms of both Z and density – or could be observational. Unlike the investigation of PV11, where only an airborne platform was employed, we have demonstrated how useful information can be obtained with ground-based and airborne systems. Another approach would be with collocated ground-based instrumentation, for density and particle imaging, and for measuring wind, snowfall rate, and radar reflectivity. This would avoid some of the complications encountered in this study, including W-band attenuation and a reliance on particle imagery acquired aloft. A close-range measuring radar might also allow retrievals closer to the surface than in this work. Improvement of methods that remotely sense supercooled cloud water are also needed."

I will likely add further comments later in the review cycle. But I would like to see the above comments addressed by the authors before I devote more time to more specific comments.

I think this manuscript has potential and could be publishable. But I encourage the authors to fine tune it further to make it more impactful.