Referee1

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

This manuscript describes a study to relate snow fall rate and W-band reflectivity based on two observational events. Overall, the results of this study may add some new incremental knowledge of mm-wavelength radar-based snowfall remote sensing. However, some revisions are needed.

Main comments.

You should, probably add some information about radar calibration. How well is the radar calibrated?

This was added to the revised Sect. 2.3:

"Ground-based calibrations of the WCR's up-looking antenna and correlations between in-flight retrievals acquired using its up-looking and down-looking antennas were used to estimate the absolute accuracy of the WCR-derived values of dBZ. This is ±2.5 dBZ (PV11)."

Did you account for the two-way radar signal attenuation by gases and hydrometers between the aircraft and the radar resolution gate, which was used?

We did in the revised Sect. 3.2.

What are the uncertainties of the hot plate for measuring snowfall rate? Given that sometimes you are getting negative snowfall rates as much as -0.3 mm/h (Fig.8), these uncertainties can be substantial.

The revised Sect. 2.4 includes a description of the hotplate precision. This was based on a comparison between the hotplate and 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 sites in Figs. 1a-b of the revised manuscript. In the revised Sect. 4, we apply the S precision when considering the departure between our S measurements and computation-based values of S. Marlow et al. (2023) was reviewed at AMS/JAMC; we submitted revisions back to the journal two months ago.

As I understand your results are shown only by a couple of points representing mean Z and S values. Why do not you show more detailed information on the S-Z correspondence?

We do not completely understand your question.

Perhaps you are saying this: Why didn't you consider time intervals smaller than 60 s (one minute) for averaging of the hotplate data? If that is correct, then our rationale is in Sect. 3.5:

"We temporally and spatially averaged the values of Z we compared with time-averaged values of S. There are two reasons for this: 1) As discussed in Sect. 3.1, the WCR did not sample Z exactly over the hotplate, and furthermore, the width of radar beam at 1500 m range - roughly the distance between the aircraft and the ground at the overflight times - is 30 m and thus considerably smaller than the minimum horizontal distance between the aircraft and the HP. 2) Compared to the WCR, the hotplate is a relatively slow-response measurement system whose output is commonly averaged over one-minute intervals (Z18)."

Or, perhaps you are saying this: Why didn't you average further forward in time (hotplate) and further backwards in time (WCR)? We addressed this in the revision, Sect. 3.5:

"As discussed earlier in this section, the averaging scheme initializes with 60-second blocks of HP data between t_o and t_o +120 s. When we applied the scheme to data from 3 January 2017, but outside the specified time range, an inconsistency was documented. This is apparent in Fig. 8, where the t_o +120 s to t_o +180 s interval (i.e., the $i = 2$ interval) has negligible average S, while in Fig. 10, the $i = 2$ interval has a non-negligible average Z (~ 0.3 mm^6 m^{-3}). A firm explanation is not available for the inconsistency, but a factor may be the convective nature of the fields in Figs. 10a-b. Because of the inconsistency, only averages corresponding to the $i = 0$ and $i = 1$ intervals were analyzed further."

Note that the Matrosov (2007) relation was derived for Z > 0 dBZ. It needs to be stated in the paper and shown in Fig. 12 (like it is done in the PV11 paper).

Yes. In fact, some of Matrosov's points (his Fig. 5b) plot slightly smaller below 0 dBZ. Also, some of his low-Z points are for dendritic crystals while most points in his figure are for aggregates. In the revised Sect. 4, we discussed the relevance of Matrosov's calculations as a comparator for our measurements:

"Figure 12 shows our S/Z measurements after we corrected the reflectivities for attenuation. Below we compare those plotted S/Z pairs to calculations reported Hiley et al.

(2011), but first, we consider the computational S/Z relationship reported by Matrosov (2007) and its relevance to our measurements. Since the particle images (Figs. 11a-b) reveal no compelling evidence for the aggregates modeled by Matrosov (2007), a model based on that particle type is not a useful comparator. Moreover, the overlap of PV11's S/Z measurements and Matrosov's S/Z calculations has already been discussed in the literature (PV11). However, before going forward, two clarifications will be made about PV11's data points in Fig. 12: 1) Presentation clarity was what guided our selection of the S and Z axis ranges in this figure but with the consequence that 32 of PV11's S/Z pairs are not shown at Z > 10 mm^6 m^{-3}. 2) The scatter of PV11 data at the largest values of Z in Fig. 12, combined with the fact that PV11 points at Z > 10 mm^6 m^{-3} are not shown, could lead to the interpretation that the slope describing the relationship at Z approximately > 2 mm^6 m^{-3} should be decreased relative to the slope of the PV11 best-fit line. Readers who view PV11's Fig. 11 will conclude that this interpretation is not warranted."

How the reflectivities were averaged? Did you average them in linear scale (mm^6/m^-3) or in the logarithmic scale (i.e., in dBZ units)?

In the original submission, and in the revision, we averaged the Z values (mm^6 m^{-3}). In the revision (Sect. 2.3), we explicitly state that.

How well the snowfall rate and reflectivity measurements were collocated? What was the vertical separation between radar Z and hotplate S measurements used in analysis of Z -S pairs?

Section 3.6 explains this:

"Figure 11a shows imagery from 12 s of measurements acquired near the end of the sequence in Fig. 9a (00:01:02 to 00:01:14). This time interval was selected by tracing forward from t_o, along the slope of the fall streaks, to the flight level."

From Fig. 9a you can see the vertical separation between flight level and the altitude of the hotplate. The hotplate is at the overflight time (~3010 m) and the flight level is at ~4550 m. The vertical separation is therefore 1540 m. That vertical separation is also equal to the pathlengths for vapor and snow particles in Table 3 (revised manuscript) where attenuation is estimated.

Section 2.3: How did you separate components of the Doppler velocities (i.e., the reflectivity-weighted fall speeds and vertical air motions)?

We did not do that. Rather, we averaged Doppler velocities in a WCR averaging interval/domain and used Eq. A8 to calculate v_p. The latter is our "maximum likely snow particle speed toward the ground. Details are in the revised Sect. 3.5 and in the revised Appendix.

Was your assertion that particles were rimed based for the most part only on the analysis of the 2DP particle images?

We used both optical array probes. This is stated in Sect. 3.6.

Did you utilize 2DS particle measurements?

Yes. This is stated in Sect. 3.6.

You suggest that the 2DP particle images are representative of those that fell from the flight level toward the hotplate. It might be not so since the height separation was very significant.

Yes, but we don't have ground measurements of particle shapes, so, Sect. 3.6 and Figs. 11a-b are the best we can do.

Minor comments

Line 91: what are rho_ 1 and rho_3 ?

We thought this was clear from Sect. 1. Since it wasn't, we added the following to the revised Sect. 3.7:

"…In the figure legend, results from PV11 are specified as $S(\rho_1)/Z$ because those authors applied the lower of two density-size functions (ρ_1) with airborne measurements of optical particle images to calculate the snowfall rates (Sect. 1). Our data pairs plot above the $S(\rho_1)/Z$ line but within the variability of PV11's measurements."

The manuscript could benefit from additional editing.

Yes. We worked on that.

I wonder if you need any permission to reproduce the figure from PV11 paper (their Fig. 11), which is copyrighted by the AMS.

We don't know. In the Acknowledgements, we do acknowledge Gabor Vali for providing data values published in Fig. 11 of PV11.

Referee2

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

The manuscript presents a field experiment in which airborne W-band reflectivity is matched with ground measurements of snowfall rate to investigate the Z-S relationship for rimed particles. The topic is very important for the precipitation community because the uncertainties in the microphysics still lead to very big uncertainties in the precipitation retrievals. The authors follow up from a series of previous papers, but in particular from the Pokharel and Vali 2011 (PV11) in which a full range of particle types is assumed and the precipitation rate is calculated from particle density assumptions. In this manuscript the authors focus on a specific particle type, rimed particles, for which precipitation rate is usually underestimated using "conventional" Z-S relationships.

Despite the great importance of the topic, the manuscript doesn't really provide a Z-S relationship for rimed particles as the title would suggest. Most of the manuscript is focused on the description of the methodology used to calculate the relationship, and very little space is dedicated to actual results. 4 points are really not enough to derive a Z-S relationship and the conclusions just state that the measurements of this field campaign fit within PV11 variability. The fact that rimed particles were not really well represented by published Z-S relationships was already known so the fact that this manuscript does not present a new Z-S relationship specific for rimed particles doesn't match with what the title suggests.

The title was revised, and the abstract was revised. Readers of the abstract will see that the number of S/Z pairs in our analysis is smaller than in PV11.

In the revision, we distinguish our work against the studies of PV11. We made direct measurements of S while PV11 derived S using particle imagery. We think this makes our contribution significant, despite the smaller number of points.

Probably the use of a ground based W-band pointing radar would have helped with the availability of Z-S points, aided by the aircraft overpass to confirm the presence of riming with the cloud probes.

We agree. At the end of the revised Sect. 5, we state the following:

"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."

Given the availability of data (I assume no more aircraft overpasses are available at the site, otherwise they would have been used),…

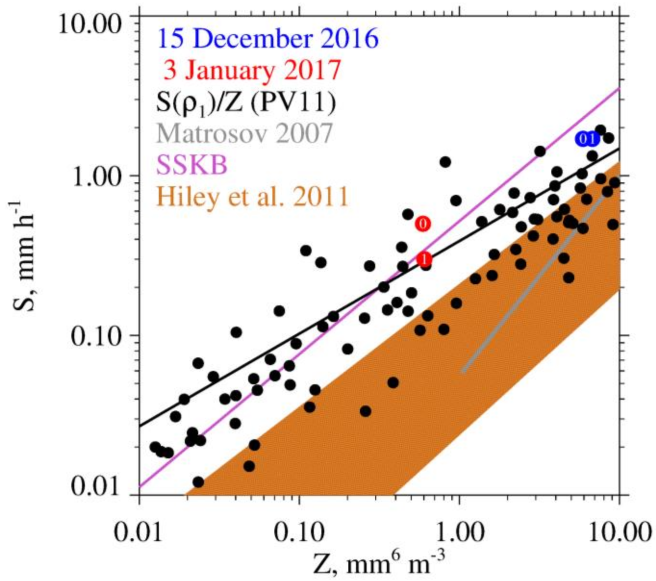
The two flights analyzed were two of three test flights flown from Laramie in preparation for the SNOWIE campaign (Tessendorf et al. 2019). The other test flight did not fly over the ground site.

I suggest to stress more the position of the Z-S points in fig. 12, trying to figure out what differentiates these 4 points from all the other points under the black best fit line or from the Matrosov 2011 range.

Following your critique, and that of Referee3, who brought Hiley et al. (2011) to our attention, we revised this section. In the revised text, we compare our measurements to Matrosov's (2007) calculation, as in the original submission, and we also compare our measurements to Hiley et al. (2011).

Attached here is revised text, from Sect. 3.7, relevant to your criticism:

"Our S/Z pairs are presented in Table 5 where the indexes ($i = 0$ and $i = 1$) are used to indicate results derived for the averaging intervals. Here, the reflectivities are not corrected for attenuation, however, in Fig. 12, the attenuation-corrected reflectivities are plotted. Uncorrected-reflectivities from Table 5, attenuations from Table 3, and Eq. 1 were used to calculate the corrected reflectivities…."



"Figure 12 – Snowfall rate versus radar reflectivity. Colored circles indicate attenuation-corrected reflectivities (Table 3, Table 5, and Eq. 1) for the $i = 0$ and $i = 1$ averaging intervals. The $S(\rho_1)/Z$ points are a subset from PV11's Fig. 11 ($0.01 < Z < 10$ mm^6 mm^{-3}). Also plotted is the PV11 best-fit line (black), the S/Z relationship from Matrosov (2007), the S/Z relationship abbreviated SSKB (Sect. 1), and the swath of S/Z relationships, for crystals, from Hiley et al. (2011)."

Here, from the revised Sect. 4, is discussion of Fig. 12. This is also relevant to your criticism.

"Departures between our S measurements (Fig. 12) and S/Z calculations from Hiley et al. (2011) were evaluated as the vertical distance between the top of the orange region and our S/Z data points. Reflectivities at the top of the orange region were calculated using attenuation-corrected reflectivities (Eq. 1 and Table 5) and the upper-limit S/Z equation from Hiley et al. (2011) ($S = 0.21 \cdot (Z')^{0.77}$; Sect. 1). The departures were evaluated as a relative difference expressed as $(S_{HP}-S)/S$ with S_{HP} one of four snowfall rates from Table 5. The relative difference is no smaller than 0.9 and 1.1 on 15 December and 3 January, respectively. These minimum relative differences exceed the hotplate precision (Sect. 2.4) by approximately a factor of three. We therefore conclude that our paired values of surface-measured precipitation rate and aircraft-measured radar reflectivity, after correcting for attenuation, provide evidence that a calculation of S based on the Hiley et al. (2011) upper-limit, when applied to rimed snow particles, is associated with a low-biased estimate of S."

On the other hand, I understand that this journal is about atmospheric measurement techniques, so if the goal is to describe the methodology to match aircraft with ground based observations, that is not really clear from the title and the abstract. As I said earlier, my expectation here is to find a new Z-S relationship for rimed particles. Based on what you decide the goal of the manuscript is, please revise accordingly.

In addition to modifying the title and abstract, we addressed this by adding goals to the revised Sect. 1.

"The goals of this paper are as follows: 1) to describe measurements of undercatch-corrected liquid-equivalent snowfall rate (S, mm h^{-1}) that 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."

Also as a general comment, there are too many not needed figures in this manuscript, I provided some suggestions to consolidate them.

Figures 7a and 8a are removed from the revised manuscript.

Specific comments:

Section 2.1 and in general when you mention AF environmental data. It is not clear to me when you actually use this dataset in your analysis since HP already has the data needed to calculate precipitation rate. Probably I missed it, but I would suggest to be more clear so it could be more obvious.

This is clarified in the revision. The AF data was used to derive the following: Absolute humidity (Sect. 3.2), cloud base altitude (Sect. 3.2), horizontal wind advection speed (Sect. 3.5), and adiabatic cloud liquid water path (Sect. 3.7). We used AF measurements for these properties because the hotplate T measurement is known to be high biased during daytime (Marlow et al. 2023). Marlow et al. (2023) was reviewed at AMS/JAMC; we submitted revisions back to the journal two months ago.

But on the other side, how far are the two sites? we know environmental conditions change a lot, especially in mountain environment, could the conditions be very different in this case?

AF and HP were separated horizontally by 2000 m and vertically by 190 m. SN and HP were separated horizontally by 1200 m and vertically by 110 m. Site altitudes are in Fig. 1a.

Is it actually reliable to use that data as it was at HP? And the same is for the SNOTEL site, would it actually reflect the HP situation?

The AF thermodynamic measurements (T/RH/P) were acquired on a tower at a long-term climate monitoring site (AmeriFlux). The exact altitude of that measurement is in the footnotes of Table 2. Relevant to your question, here is what we know about the ground sites: 1) The vertical separation of AF and HP, and 2) that the winter-season wind flow is nearly always directed approximately from AF to HP. From those characteristics, and the dry adiabatic temperature lapse rate, we expect the temperature difference AF - HP to be no smaller than -2 K. If you look at the sequences from HP and AF (Data Availability Statement; https://doi.org/10.15786/20247870), you will see that the AF - HP temperature difference, at night (see above discussion of the HP's daytime temperature measurement bias), conforms to our expectation. Hence, we think it is reasonable to assume the AF thermodynamic measurements are representative of the region surrounding the three ground sites (AF/SN/HP). This region is shown in Figs. 3a-b.

The consistency of the SN and HP snowfall measurements is discussed in Sect. 2.4 (revised manuscript) and in Marlow et al. (2023).

Regarding the AF-derived horizontal wind velocity, we do not have a check on how representative that is for the AF/SN/HP region. We do know that the measurement was made above the tree tops (the anemometer was/is deployed at the top of a tower) and that the measurement system (propeller anemometer) is reliable.

Section 2.4, you describe the hotplate and all the bias corrections needed, included a comparison with a fenced precipitation gauge. Why isn't the HP inside a fence?

We apply an algorithm which assumes the hotplate is _not_ within a fence. This is discussed in Sect. 2.4 of the revised manuscript.

Section 3.3, lines 287-291: why mentioning this previous attempt to compare wind speeds if data sets are difficult to interpret and they do not provide useful results for this work?

Because we reported, in a conference presentation, comparisons of hotplate-derived and Vaisala-derived wind speeds. We later found the problem with the Vaisala-derived speeds.

What is the point to show up- and down-looking reflectivities? Up-ward ones are not needed for this work…

There are three reasons for this. 1) In Sect. 3.6 we discuss the fall streaks at ~ z = 5500 m in Fig. 5a (i.e., above the flight level in the up-looking height-time crossection). 2) People would ask for what's above the flight level if we did not show that information. 3) To compare, on one page, the two weather systems (i.e., one has relatively large reflectivities, is deeper and stratiform, the other has smaller reflectivities, and is shallow and convective).

…actually these plots are a repetition of figures 9 and 10 (except for the up-ward reflectivities). Vertical winds can be consolidated into figs 9 and 10 too, focusing on the portion of the overpass that is actually of interest for the analysis.

We think we have crafted things effectively and logically. Please consider the revised manuscript. Here is how the presentation evolves from Figs. 5a-d, to particle imagery (Sect. 3.6), to Sect. 3.7 (S/Z Relationships), and to Fig. 12:

What is shown in Figures 5a-d (Sect. 3.5) ends at the overflight time. Figures 6a-d explain the averaging. Figures 7 and 8 show the ground measurements and ground-measurement averaging intervals. Nearly at the end of Sect. 3.5, we introduce Figures 9a-b and 10a-b. These show the WCR measurements prior to and after aircraft's overflight. We also state why the time axes are different in Figures 9a-b and 10a-b (compared to Figs. 5a-d), and that the WCR "structures" in Figs. 9a-b and 10a-b will be discussed in the following section (i.e., Sect. 3.6, Snow Particle Imagery). Section 3.5 ends with Table 5. The Table 5 has the averages. The averages are the basis for Fig. 12, Sect. 3.7 (S/Z Relationships), and Sect. 4 (Results).

Line 433-434, the meaning of the slopes is not really clear if the reader hasn't read the appendix yet. I would suggest to add a sentence explaining why the HP line is flat while the WCR one has a slope (and then refer to appendix for details).

We revised this portion of the manuscript and revised Fig. 6. Here is the revised text:

"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 times of the corresponding WCR averaging intervals 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 tops 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."

Figure 6: I am not sure this figure is needed or can probably be moved to the appendix. I find it a bit confusing.

We revised Fig. 6.

Figure 7b is the same as fig. 2, just extended to reflect the situation around the observation time. I would try to consolidate the figures.

Figures 7a and 8a (both had wind speed at the hotplate) were eliminated from the revision.

As I mentioned before, despite the presence of fig. 6, the averaging intervals are not clear and confusing. The appendix should be for details, not for the general understanding of what we are looking at. For example the difference between i=0 point being after t0 for HP and before for WCR should be stated somewhere in the text (not only in the appendix). Or the meaning of the WCR slope.

Figure 6 was revised.

Minor comments:

In the abstract you refer to ‘published Z-S relationship’ which sound like a very specific one (I assume you are referring to PV11). It is probably good to mention it.

Yes, in the revised abstract we did that.

line 309: add ‘forced through the origin, RED LINE’.

Yes, in the revised manuscript we did that.

Line 366: provide a time reference for the ridgeline as you did for the last 3 seconds.

Yes, in the revised manuscript we did that.

Figure 5, the plot at the end goes outside the axes (red line).

Yes, in the revised manuscript we fixed that.

Figures 7a and 8a are never mentioned in the text, either mention them or remove.

Yes, in the revised manuscript those two panels are removed.

Figures 9b and 10b, usually doppler velocity has a blue/red colormap, you might consider it for consistency with other publications or just for differentiating it from the reflectivity plot on figs 9a and 10a.

Yes. This was done in the Doppler velocity panels of Figs. 9 and 10.

Line 629: ‘within the variability’ – maybe in fig. 12 you can plot the PV11 variability to make it more clear.

We did not do that, but Fig. 12 was substantially modified in the revision.

Line 693: in Kulie et al the threshold is 0 dBZ.

That sentence was removed from the revision.

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.