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),…

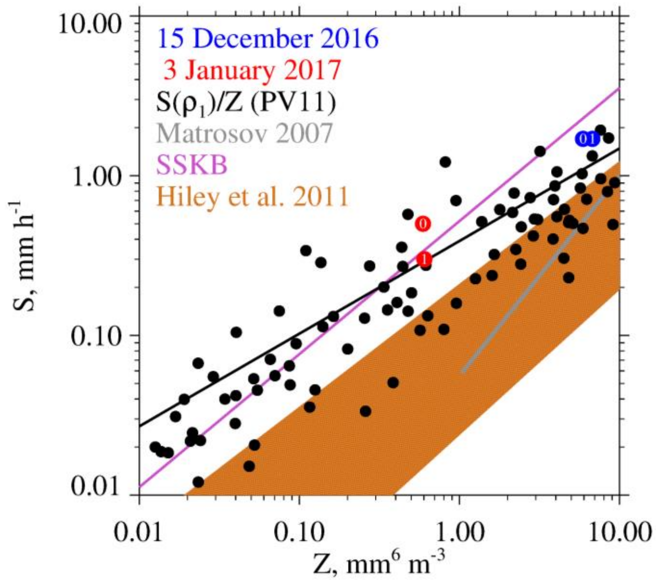
The two flights analyzed were two of three test flights flown from Laramie in preparation for the SNOWIE campaign (Tessendorf et al. 2019). The other test flight did not fly over the ground site.

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(ρ_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.

"We now evaluate departures between our S measurements and S/Z calculations from Hiley et al. (2011). Each of the departures will be evaluated as the vertical distance between the top of the orange region in Fig. 12 and our S/Z data points. Reflectivities at the top of the orange region were calculated using attenuation-corrected reflectivities (Eq. 1) and the upper-limit S/Z equation from Hiley et al. (2011) ($S = 0.21 \cdot (Z')^{0.77}$; Sect. 1 and Eq. 1). In terms of a relative difference, expressed as (S_{HP}-S)/S and with S_{HP} an attenuation-corrected snowfall rate, the departures are 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 and aircraft-measured radar reflectivity, after correcting for attenuation, provide evidence that a calculation of S based 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.