Riming-dependent Snowfall Rate and Ice Water Content Retrievals for W-band cloud radar

N. Maherndl, A. Battaglia, A. Kötsche, and M. Maahn

March 24, 2025

Original Referee comments are in italic

> manuscript text is indented, with added text underlined and ~~removed text crossed out.~~

We would like to thank the reviewers for their helpful comments. We revised the manuscript and responded to all of the reviewers' comments.

Reviewer I

Short summary

The authors derive relations between radar reflectivity (Ze) and ice water content (IWC), as well as Ze and snowfall rate (SR). Compared to existing relations, they take into account the normalized rime mass (M) into their relations. The relations are trained on data from one field site and evaluated on data from three other field campaigns. Additionally, since M is often unknown, they also provide relations based on liquid water path (LWP). Those relations could be applied to space-borne measurements.

General Comment

I think the manuscript is generally well written and addresses a relevant scientific question within the scope of the journal. Therefore, it is likely worth of publication after all comments are addressed.

I like that the authors train their relations based on data from one site and then test the performance at three other locations on very different latitudes. This greatly increases the credibility into the universality and robustness of the method. The only issue I see is the following: The relations are established as a fit of a polynomial function between Ze and IWC. However, IWC can not be observed directly. Since the in situ imager can only measure the particle size distribution N(D), we need an estimate of the snowflake mass. This mass information depends on estimates of the normalized rime mass M, which in turn is derived from N(D) and Ze. Therefore, in the training, the "labels" (IWC) are not independent from the "features" (Ze)!

This is the case for the train and test sites. The only truly independent evaluation, if I am correct, is coming from the SR gauge.

I think one has to discuss carefully what this double use of Ze implies. Would a bias in Ze go unnoticed, since it affects both sides of the fit equation equally? Can it explain the weaker correlation with gauge measurements? Generally, for me, it seems like the retrievals of IWC and M face a similar problem: In both cases, Ze, which depends on number and mass, has not enough information to fully constrain the values of interest. For IWC, we need additional information about the mass or density (like M).

For the retrieval of M, we need information about the particle number, which is coming from the particle imager N(D). Then, a forward operator in combination with an optimal estimation approach is used to find the M, which matches the observed Ze best.

So why not using the same approach directly for the retrieval of IWC? I.e., given N(D), *using a forward operator and vary IWC until it matches the observed Ze? Otherwise, an interesting study on a unique dataset!*

We thank the reviewer for the positive review and the constructive comments, which helped to improve the manuscript.

Regarding the comment about IWC not being independent from Z_e, we have added the following in the revised manuscript:

> It must be noted that our reference IWC and SR data are not fully independent of Z_e because we derive the particle mass from the retrieved normalized rime mass M. This is a necessary limitation because IWC and SR cannot be inferred from the available in situ measurements alone. For SR, we evaluate our approach with completely independent SR gauge measurements.

To give a more detailed answer here, yes, our reference data of IWC and SR are not independent from Z_e. We are aware that this is a limitation of the approach, however this way, we get the best estimate of IWC given the available data. An independent IWC estimate is not possible because we currently don't have size resolved measurements of particle mass. The radar data we use is quality controlled and we assume Z_e is not biased. If Z_e would have a positive bias, then M would have a positive bias as well, resulting in a positive bias of IWC.

The suggestion to directly retrieve IWC would result in the excact same outcome as the M retrieval. This is because in the forward operator, particle mass is a function of M (because we use the mass-size relation form Maherndl et al. (2023)).

Other comments

line 36-42: I believe error cancellation happens if the relations are trained on enough data to capture the full snowfall climatology with all its diversity in shape, habit, etc. Then, a relation can be wrong in a single case, but the overall average might be correct. However, if the relations themselves differ by about one order of magnitude, why would you assume that the error cancels out on seasonal time scales? If one relation is systematically higher than another (compare e.g. in Fig. 5 i)), it will also lead to systematically higher results for any reflectivity time series!

We agree that error cancellation results when relations are trained on a large amount of data capturing the snowfall climatology. We added:

> ~~Although these relations~~ Relations can have significant uncertainties for individual cases, ~~they~~ but are successfully applied to space-borne radar data sets because the random errors cancel ~~partly out~~ out partly in seasonal time

> scales assuming they are trained on a large enough data set to capture the full snowfall climatology (Kulie, Bennartz, 2009).

Line 205: Why not discarding NaN cases? (Would it reduce the number of data points too much?)

Yes, the number of data points would be heaviliy reduced. There are often size bins with a number concentration but no fall velocity in the VISSS products due to the different methods in deriving number and velocity distributions. We added:

> ... To avoid unrealistic behavior at the edges of the size spectrum, NaN values of v are filled with v from the closest available size bin. Removing cases with NaN values would greatly reduce the number of data points.

Sec. 3.4, Equations 6,7,8,9: Can you motivate, why you chose exactly this functional form as a basis for the regression? For example, when I look at the measurement data in Fig. 4a+c, it seems like the cone of data points is getting more narrow towards the right (the spread of IWC and SR is becoming less with higher Ze). This is not visible in your fit functions in Fig. 4b+d. Why not including a coefficient in the fit function which would allow for this flexibility?

We opted for a log-log IWC-Z and SR-Z functional form due to its common practice. Temperature in degree Celsius must be included linearly due to negative values. M is included logarithmically, because the distribution of M often follows a Gaussian shape in logarithmic space. The smaller spread of IWC and SR at higher Z_e could be due to the rarer occurrence of such high Z_e in our data. It could be the case that if we were to have a larger data set, the spread at high Z_e were also larger. In the revised text, we added:

> We chose the functional forms of Eq. 6-9, because power law relations are commonly used for IWC-Z_e and SR-Z_e (Fuller et al., 2023). T in degree Celsius must be input linearly due to negative values. The logarithm of M and LWP are used, because the logarithm of both variables often follows a Gaussian shape.

Line 224: Why not using the SAIL-derived 2.25 dB as offset for the viewing angle correction, instead of the 2.29 dB derived from inter-site comparison (line 218)? That way, the data evaluation from the validation sites would be completely independent from the training site.

Because of the small data set when only using SAIL data and the limitation of which M threshold we can use for unrimed particles (see lines 229-231 of the revised manuscript). Also, the standard deviation for SAIL only is much larger due to the small data amount.

Fig. 3 caption: I thought the offset between SAIL and HYY+NYA is 2.29 dB +-0.39?

Yes, this was a mistake. We fixed it.

Line 276: Does this mean the relations become uncertain/invalid beyond 20 dBZ? I would recommend to plot only the range where data points are available, or at least indicate the range where the relations are based on extrapolation.

We have added shading for Z_e larger 15 dBZ indicating the range of available data in Fig. 4.

Line 289: "based on in situ data" − > not only (see main comment)

We added:

> Reference IWC and SR based on in situ data and retrieved normalized rime mass M (Eq. 4 and Eq. 5) are denoted $\text{IWC}_{\text{reference}}$ and $\text{SR}_{\text{reference}}$, respectively.

Line 307: Site specific effects − > what could those be?

Orographically induced turbulence at the SAIL site, for example. We added:

> While the IWC relations developed for SAIL perform similarly well for the other sites, the SR relations performs noticeable worse indicating site-specific effects (e.g., orographically induced turbulence might affect snowfall at the SAIL site).

Sec. 4.2.2.: You trained the retrieval based on air temperature measurements and, if I understand correctly, also use air temperature in your "simulated" satellite retrieval (+-2K uncertainty). For the case of a real satellite, which temperature would you use? Would it be valid to use brightness temperature? Would this limit the retrieval to cloud top?

For WIVERN, we suggest using auxiliary data from ECMWF as input for air temperature. The retreival can therefore be applied to the whole column, not only cloud top. This is common praxis also for other satellite missions such as CloudSat.

Line 325: Interesting. I would have assumed, since the inclusion of LWP (as proxy for M) is the main improvement over existing relations, the influence of LWP would be bigger. What are the other error sources you mention?

Including LWP leads to an improvement as opposed to not including LWP even when assuming a large uncertainty. However, it is no surprise that the literature relations perform worse than our relations, because they were derived from different data sets. Errors result from the natural variability in IWC-Z_e space, the random errors we included to approximate WIVERN observations, and the regridding to the coarser WIVERN resolution. We added:

Doubling the LWP error from 30 $\mathrm{g\,m^{-2}}$ to 60 $\mathrm{g\,m^{-2}}$ has barely any impact, because other error sources (e.g., from averaging to the WIVERN resolution) dominate the resulting variability.

Fig. 7: Instead of SAIL vs All sites, I would probably split SAIL vs Other sites (clear Train vs Test distinction). Also in other figures.

We changed the figure row title from *All Sites* to *Other sites* and changed *all sites* to *all other sites* in the text.

Line 352: A positive bias increases the NRMSE, a negative has no effect. Why is this asymmetry?

The asymmetry is due to the logarithmic scale.

Fig. 11, 12: It is interesting to see that the gauge error (R2, RMSE, ME) seems to be generally less for HYY than for SAIL, even though the relations are developed on SAIL data.

Yes, this is interesting and could be due to lower variability of snowfall at HYY than SAIL, but this is only speculation.

Conclusion 1 (line 376 following): Since the viewing angle offset was derived for unrimed particles, high and low IWC situations are mainly different in number concentration and particle size, I assume. Since riming is a main point of this study, it is interesting to ask what happens for riming. Intuitively, at least for strong riming, I would expect the particles to become rounder and therefore, the offset between vertical and slanted observations to become smaller?

Yes, that is expected for very heavily rimed graupel particle, but not necessarily for light riming. We don't have enough data with very strong riming to prove this hypothesis.

Technical Comments

Line 42: "cancel partly out" − > "cancel out partly"

Line 166: "Windows corresponds"− > Windows correspond

Line 218: "between the vertically pointing" − > remove?

Line 221: "when they when" − > when they where

Line 298: there − > their

Thank you, we fixed the mistakes.

Figure 4, panel b): What is the black dashed line? Add a legend.

Thanks for noticing, the line should not be part of the plot and was removed. (It was a leftover from testing.)

Figure 6: Even though histogram units are often w.r.t. density or similar, a colorbar would be nice.

We have added colorbars to Fig. 6-9 and Fig. 11-12.

Fig. 11 & 12 could optionally be combined into one figure, like Fig. 9.

We decided to show Fig. 12 separately, because it shows HYY data only. The combined figures all depict SAIL vs. all other sites.

Line 414, last sentence: incomplete?

Yes, we fixed it to:

> Then, the IWC and SR relations including M can be used, which have lower uncertainties than the ones based on LWP.