Review of “Dual-frequency spectral radar retrieval of snowfall microphysics: a physically constrained deep learning approach” by Billault-Roux et al., AMT-2022-199

The manuscript presents a unique approach to utilize dual-frequency radar Doppler spectra and a two-step deep learning framework to retrieve a number of snow microphysical parameters. The method is applied to spectra collected during a recent ground-based campaign where the retrieval results could be compared to in-situ data obtained by aircraft overpasses.

Overall I congratulate the authors to this comprehensive approach using machine learning to extract information from multi-frequency Doppler spectra. In my opinion, this will help a lot to better exploit the information about snow microphysics in Doppler spectra. Given the complexity of the method itself and the challenge of selecting appropriate input parameters for the training dataset I find the results very impressive and encouraging. As the authors discuss properly, the method has still some shortcomings as any retrieval probably has. I found the paper very well structured and clearly written which is important in order to help non-expert readers to understand the different steps of the approach. I also appreciate that the code for training the NN is openly accessible. I think this is very important given the fact that the NN has to be trained separately for any other radar combination.

My comments are mostly minor and many of them meant as suggestions to make the description even more clear and complete. I completely understand that certain compromises in the assumptions had to be made in order to keep the computational costs feasible. The references to previous literature could be a bit more extended; I provided some references in my specific comments.

I clearly suggest the manuscript to being published in AMT after my comments and questions have been addressed.

General comments:

Forward model assumptions, Sect. 3.1.1: I am a bit surprised to see that you used only inverse exponential PSDs for generating the synthetic Doppler spectra dataset. Several in-situ and remote sensing studies showed that a modified Gamm distribution is a much better function to fit snow PSDs. As shown by Mason et al., the PSD shape (for example mu) impacts also DWRs. My concern is that with your inverse exponential PSD, you are unable to represent for example cases of intense aggregation, where one expects the small particle number to decrease (high mu). Or imagine secondary ice processes, which can lead to super-exponential PSD by enhancing the number of small particles. If it is too much effort to implement a more general PSD in your framework I would strongly suggest to add this aspect to the discussion.

What is general a bit unclear to me is how the retrieval is able to handle attenuation effects (especially at W-band) which can be expected especially when more super-cooled liquid water is in the ice part of the cloud or if even melting layer or rain is present. From the description I find it hard to understand how attenuation was handled. PAMTRA is able to also simulate attenuated spectra, so maybe this information could even be implemented as a variable to be retrieved in the future? (I see a small comment in L. 553 on this topic)

Another aspect is still unclear to me but maybe I misunderstood it: You mention in the beginning that it is extremely challenging to match spectra of two radars due to various effects, for example slight antenna mis-pointing. It is unclear to me how your retrieval deals with those effects. For example, a mis-alignment of the two radars would cause a different velocity shift due to different vertical and horizontal wind components in the radial velocity. This will cause “fake” sDWR simply due to the different shifting of the spectra. As such effects are probably not included in the training, I wonder what effect those artefacts might have on your retrieval?

General question: Maybe it could be worth to comment in the outlook a bit on how easily your method could be also applied to multi-frequency spectra obtained in rain. There has be quite some work done on that using mostly OE but I wonder how much of the “pre-selection” work that needs to be done for those OE retrievals can maybe be avoided using your approach.

Minor comments:

L. 2: I think radar polarimetry should be mentioned here as well. Admittedly, the signals are weak for snowflakes or rimed particles but overall it provides a lot of information about ice processes which are related to the generation of snow. In the future, I guess your method could also be extended to utilize also spectral polarimetry, right?

L. 56-57: I think you should somehow indicate that there are many more studies than Tetoni et al. which explored radar polarimetry for studying ice and snow processes. There is also lots of literature on improving snowfall retrievals by using polarimetric information such as for example Bukovcic et al., 2018 (https://journals.ametsoc.org/view/journals/apme/57/1/jamc-d-17-0090.1.xml).

L. 63-64: For the non-expert reader I suggest to explain more what is the underlying idea of using Doppler spectra: If the particle’s terminal fall velocity differs, we can assign the backscattered signal to different populations and hence refine our retrievals. If ice, snow, rimed particles would all fall with the same velocity, the spectra would be actually quite useless.

L. 73-76: Also differential attenuation at the two frequencies is an issue. It shifts the two spectra up and down. This effect can also be used if one knows that certain parts of the spectra must match. Li and Moisseev, JGR, 2019 (https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2019JD030316) used this effect to estimate melting layer attenuation. I think it would make sense to mention this study here.

L. 114-116: As the spectra are quite essential for this work I suggest to expand a bit on it in order to enable to reader to better understand the background. I would split the sentence and explain first that the underlying microphysics defines a “perfect” microphysical spectrum which only depends on v-D relation, backscattering properties and PSD. In order to estimate the impact of the dynamics on the spectrum, I already need to know the radar parameters (e.g. beam width). Finally, the radar itself adds noise to the spectrum. In order to reduce the noise effect, one needs to average over several spectra but this leads usually to a smear-out effect of the microphysical information. A study which explores and discusses those different effects w.r.t drizzle signatures is Acquistapace et al., AMT, 2017. Maybe also a reference to a textbook or review article explaining the various influences on Doppler spectra might be good to add.

L. 193: It’s not clear to me how you assign aspect ratio if you only have mass, size and cross sectional area. I assume you apply some ellipsoidal fitting? Please specify or provide reference.

L. 197: If I am not wrong, you use Heymsfield+Westbrook to calculate the terminal velocity based on your assumed m-D and A-D relation, right? In the way you write it it sounds like another independent v-D relation is used.

L. 195-197: Sorry if I missed it but I would like to see a table (maybe in the Appendix) which all relevant radar parameters assumed in PAMTRA such as noise level, number of spectral averages, velocity resolution, beam width, etc. I could imagine different choice of those parameters might affect your results quite a bit. I assume you have adjusted those parameters to the radars you use later for applying the retrieval? I think making the assumptions in Pamtra consistent with the "real" radar settings is quite key because you will get for example quite different spectral broadening if beam widths or number of spectral averages don't match. I would like to see these aspects clarified/discussed more deeply.

L. 198: How do you deal with attenuation? Do you use PAMTRA spectra which include attenuation?

L. 216: Did you use SSRGA also for graupel in Pamtra?

L. 243-244: Why did you not assume a range of Eddy dissipation rates? I think this would represent real conditions much better than a single EDR.

L. 260: Similar to my previous question: Wouldn't you need to do the Pamtra simulations separately for X- and W-Band simply because of the much larger beam width of the X-band? I mean, the broadening due to turbulence must be quite different, or? I think on L. 301 I found the answer. Maybe mention this aspect somewhat earlier.

Table 2: Information about the radar sensitivity/noise level would be good to add.

L. 281: So that means you don't have information about particles larger than 6.4mm? Isn't that a problem for studying snow which can reach sizes of a few cm? I guess missing this large-particle-tail will impact the derived PSD and moments. I think this aspect must be discussed.

L. 340: Could you expand a bit more on this aspect of spatial consistency? First I would like to understand better what the term means. I guess it means that certain spectral features (e.g. a second spectral mode) appear to be connected throughout certain range gates? 3 range gates appear to me quite a small range if I just look at your measured spectrum in Fig. 4a.

L. 388-389, Fig. 3: Couldn't it maybe also be related to insufficient range of assumptions in the training? For example, the range of densities assumed, the simple inverse exponential PSD? Do the two panels in Fig. 3 actually have the same velocity axis? Then the X-band spectrum is from rimed particles, while the W-band is for slight updraft conditions, right? Wouldn't it make more sense to show the spectra for the same time-height of a specific event? Both spectra look quite smooth. I would be interested in seeing the original spectra.

Figure 4 looks indeed very impressive! But for me it is hard to understand how you are able to reconstruct also areas of various level of turbulence so well given that you only use one value of turbulence (EDR) in the Pamtra simulations for training.

Sect. 5.2.2: Correct me if I am wrong, but I guess you could also retrieve eddy dissipation rate, right? I think it could be nice to compare your EDR estimate with simpler methods such as presented by Borque et al., JGR, 2016 (https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2015JD024543)

L. 495: Looks like the in-situ confirm that the inverse exponential is not an ideal PSD shape. I would be interested to know if the higher spectral moments (for example skewness) are changing if you use the in-situ PSD instead of your fitted exponential. While certainly the larger particles dominate Ze and DWR, I am not sure that they don't affect the spectral shape.

Sect. 5.3.4: I guess some of the Ar discrepancy could also result from the averaging over particle 2D projections done in common in-situ probes. I guess mentioning the study by Jiang et al., JAMC, 2017 (https://journals.ametsoc.org/view/journals/apme/56/3/jamc-d-16-0248.1.xml) could make sense here.

Fig. 12: Would be great to also have a similar plot showing the impact of absolute and relative velocity offset (I guess it will impact m-D parameters quite a bit).

Sect. 6.3: I like the discussion of the impact of the SSRGA parameters on the D0 estimate. However, I think the fact that you are sticking to inverse exponential PSD might also explain some of the D0 bias. Mason et al., AMT, 2019 (https://doi.org/10.5194/amt-12-4993-2019) demonstrate this effect very nicely and I think this aspect should be included in the discussion.

Figure 9 and D0 discussion: What I find quite puzzling is that your retrieval is never producing D0 values lower than 1mm despite very low DFRs and Ze at some time periods where you have overpasses. Couldn't it maybe be a relative bias in Ze? Can you show reproduced and measured X and W-band spectra plotted over each other once for a low DFR and once for a high DFR region? Ideally there should be no big offset in the y-direction except the different noise levels.

L. 635: Just a comment: Airborne triple-frequency data as used in Mroz et al., AMT, 2021 might not be very useful for your approach as the spectra are probably affected by the aircraft motion. There are ground-based triple-frequency datasets available which provide quite acceptable 3f-spectra (see for example as shown in another article by Mroz et al., AMT, 2021 (https://amt.copernicus.org/articles/14/511/2021/amt-14-511-2021.html) or von Terzi et al., 2022 (https://doi.org/10.5194/acp-2022-263).

Typos:

L. 58: "studies, however, rely"

L. 206: "We, however, believe"

Table 1: I think there is a mistake in the description of am/bm and alpha_a/beta_a. It says both "… of the mass-size power law"

L. 348: in this way

L. 360: I guess the "1" with the double line in the equation denotes a vector of ones if the condition in parenthesis is fulfilled?

Caption Fig. 5: Change DFR into DFR_X,W as in the Figure

Figure 5: If possible enlarge font size of the labels on the axis and color bars. Also information on Time should be added on the x-axis (I guess it's time in UTC); also in Fig. 7

L. 486: with, however, the

L. 513: This is not