The paper presents the development of emperical IWC-Ze and SR-Ze relationships for obtaining snowfall rate and ice water content from W-band radar observations and accounting for the effects of riming. The work documents the performance of the relationships both when the normalized rime mass is known and when liquid water path is used as a proxy for rime mass. The normalized root-mean-square errors are somewhat large for small IWCs and SRs, but decrease as IWCs and SRs increase, giving performance generally better than a number of previously published studies.

The presentation of the work is clear and well organized. The conclusions are supported by the results. I have some requests for clarification or additional explanation that I consider mostly minor; see the comments below. I'm recommending acceptance with minor revisions.

My most significant concern is that it took a fair amount of examination of prior work (Seifert et al., 2019 and the earlier works by Maherndl et al.) to resolve in my mind the implications of changes in normalized rime mass. For example, does an increase in normalized rime mass mean that mass and D_max are unchanged,but that some of the mass is converted to lower-density rime mass? It would be helpful for the authors to provide a brief explanation to provide background for the readers, then allow them to consult the references if they desire a deeper understanding.

Science and technical comments
================================

L 21-22: Most gauges do not measure snowfall rate directly. Normally, they measure accumulation over a time interval, which is then used to compute mean liquid-equivalent snowfall rate over the time interval.

L 28-29: I don't think that W-band radars are commonly used for snowfall. There are complications introduced by attenuation and non-Rayleigh scattering by snow particles. For many operational applications, C- or S-band weather radars are the most commonly used for estimating snowfall rates, usually using empirical Ze-S relationships. Space-based operational missions like Global Precipitation Measurement and Tropical Rainfall Measurement Mission use Ku- and Ka-band radars, although they do suffer from a lack of the sensitivity needed for lighter precipitation. Research programs like CloudSat and US Department of Energy's Atmospheric Radiation Measurement do employ W-band and Ka-band radars which are used for snowfall retrieval, but W-band radars are probably the least common types used for snowfall.

L 34-35: Do you have a reference for the information content of satellite observaions for snowfall?

L 56-58: After reviewing Scarsi et al. (2024), I don't see any claims or documentation that the 800 km swath of WIVERN produces reduced uncertainty in polar snowfall estimates versus CloudSat. Please be more clear about what your are claiming here.

L 59-60: I'm concerned about this assertion since it seems at odds with what has been found for the scanning Ku- and Ka-band radars used by GPM (scattering from the beam side lobes causing a deeper blind zone at larger angles of incidence). See Kubota et al. (2016, JTECH, doi:10.1175/JTECH-D-15-0202.1). The Coppola et al. reference is not available for examination.

L 97-98: Do you have a reference for the statistical retrieval itself? Is this the LWP from the two-channel or from the three-channel radiometer? How was the Pluvio gauge fenced?

L 160: It seems rather arbitrary to remove light snowfall cases. What is the justification for doing this?

L 173-174: When you say "horizontally aligned", do you mean perfectly horizontally aligned, or are the particle orientations allowed to vary randomly by some maximum angle relative to horizontal (i.e., 'flutter')? My experience with discrete dipole calculations indicates the first approach can overestimate backscatter cross-sections if the radar is vertically pointing. This may not be a significant concern with a 40-degree viewing angle, though.

L 180: Can you provide a reference for this assertion about the dominance of riming for particle mass?

L 219-224: So, to be sure that I understand, you compared the relationship between Ze and IWC for 90-degree observations from SAIL against the corresponding relationship derived for 40-degree observations from SAIL? And the m(D) parameters by which IWCs were determined were obtained by fitting modeled reflectivities (obtained from the observed PSDs) against the observed 40-degree slant path observed Ze?

L 232-249: There are a few things to clarify here. First, although the relationships given in (6), (7), (8), and (9) are general, is it your intention to fit these relationships using the 40-degree Ze observations? And I'm fairly certain the 'LWP' used here is the LWP measured in the vertical direction, but it would be helpful to explicitly say so. The reason to clarify this is that at least one of the DOE microwave radiometers (the two-channel) produces both a liquid water path in the zenith direction and a liquid water path that is along the line-of-sight path of the radiometer.

L 273-275, Figure 4: Distinct point shapes are not apparent, particularly in the locations where the scatter plot point density is high. Would 2D histograms be more useful for panels (a) and (c)?

Also, in Figure 4, there are some interesting behaviors. In panel (a), if one follows a vertical, downward trajectory of constant Ze, normalized rime mass M increases while IWC decreases. It would be interesting to see what this means in terms of size distribution changes, but that is beyond the scope of this work.

L 297-299: This explanation seems a bit simplistic. Could there not be differences versus particle habit assumptions or versus the population of PSDs used by Hogan et al. to produce their relationship? It's not apparent to me that this difference is attributable solely to the use of unrimed particles by Hogan et al.

L 314-317: At the SAIL site, it appears the LIMRAD94 was limited to a range of 2000m, or a height above ground of about 1300 m at the 40-degree observation angle. Similarly the maximum range was 2000m at Eriswil. The maximum heights for the RPG-FMCW-94-DP and JOYRAD-94 are not specified.

For this WIVERN-like comparisons, did you simulate an atmospheric column? What were the depths of the columns and how did you treat 94-GHz attenuation? Also, how did you approximate the 1 km horizontal resolution using the ground-based observations? Please explain more about how the WIVERN observations are simulated from the ground observations.

Finally, given the empirical nature of the retrieval, it is not clear how these uncertainties are introduced into the retrieval process and how their effects are evaluated. Please explain.

L 323-326: See my previous comment regarding how the effects of uncertainties are evaluated.

L 356-360: Wind effects ('pumping') are probably the most significant source of error for heated Pluvio guages. Did you consider filtering out cases with high wind conditions to see if the validation results improved?

L 380-383: See my earlier comment. At face value, increased riming of particles should increase IWC. But in this case, you are describing the changes given a constant Ze. So, what physical processes are happening if Ze is constant but riming increases and IWC decreases? It would be helpful to most readers to provide some insight here.

L 391-395: Se my earlier comment. More details are needed in the discussion section regarding how these WIVERN-like measurements were produced and how observational errors were propagated into the empirical retrieval.

Minor language corrections
==========================

L 40: Should be 'particle size distributions'.

L 58: Should be 'signficantly reducing the uncertainty'.

L 295: Should be 'site'.

L 298: Should be 'their'.

L 357: Should be 'HYY is operated as'.