Review of *"An intercomparison of EarthCARE cloud, aerosol and precipitation retrieval products",* by Mason and coauthors, egusphere-2023-1682.

The approach used in this study, to evaluate the forthcoming EarthCARE cloud products by testing them against GEMS model output, is interesting and novel. Three test scenes from the model output are used, not unlike the Qu et al. (2022, amt-2022-300) article. As the authors note, the use of the model output provides the opportunity to perform a detailed evaluation of EarthCARE's instruments and retrieval algorithms when the model "truth" is known.

Microphysical properties-size distributions, ice and liquid water contents, etc, are generated in the model using Milbrandt and Yau's (2005) double-moment bulk cloud microphysics Scheme, which predicts mass and number mixing ratio for each of six hydrometeors classes: non-precipitating liquid droplets; ice crystals; rain; snow; graupel; and hail. These are the fundamental microphysical properties used in the evaluation of the EarthCARE instrument algorithms.

1. How good are the model variables and how well do they represent ice clouds? The retrievals strongly depend on the model data, and the fundamental question to ask is how good is the model data? There are many good field program data sets available, and have the model data become compared to recently collected microphysical observations from field programs? This obviously could be readily done. I looked back at the Milbrandt and Yau article and find the ice and snow categories and assumptions are based on the Ferrier and Ivanova et al. articles. The latter article uses FSSP particle probe data for the (bimodal) size distribution relationship(s). The FSSP data has been shown to significantly overestimate the concentrations of ice crystals because of instrument shattering. This could obviously affect the ATLID calculations, etc.

You note in Section 2 that apparent errors or biases in the retrievals presented in this paper may therefore be due to differences in assumptions underlying the model truth. In the concluding remarks you note that your evaluation has been carried out with three simulated EarthCARE granules. I suggest that you discuss my point about model uncertainties in detail below that sentence. You could potentially do sensitivity studies with the model output, increasing the concentration of small ice crystals by an order of magnitude, etc, to see how the retrieval products would be affected.

2. Some of the acronyms used to represent the different instrument combinations and retrievals are not intuitive (ACM-CAP, etc) and are difficult to follow. I made my own table representing the acronyms. I suggest making a table containing the acronyms
3. Lines 159-163. Non-Rayleigh effects at W-band are extremely significant at reflectivities of 12 dBZ or so and above. W-band radars do not measure reflectivities above about 18 dBZ-that is, increasing "real" reflectivity results in decreasing W-band reflectivity. This should be mentioned. I looked at the Mroz (2023) article and it did seem like non-Rayleigh effects were accounted for.
4. General comment. Figure 1 and subsequent figures and in the text. Use g/m3, not kg/m3. The former is what is used in the literature and the units are such that values are easier to "digest". Lines 211, 233. Do you mean 1 g/m^3? 1 kg/m^3 is physically implausible.
5. Figure 2 is very informative and useful.
6. Line 274 and elsewhere. When referring to ice, use ice water content, when liquid, use liquid water content.
7. Lines 291-292. This statement is not quite correct. For W band, complete attenuation of the radar beam can occur in regions of very high radar reflectivity.
8. Section 3.2 What is not mentioned in your article is that a lidar beam is fully attenuated at an optical depth of about 3. Thus, in liquid cloud, penetration into the cloud layer would be a very short vertical distance. The relationship between optical depth and liquid water content (path) can be found in https://atmos.uw.edu/~robwood/papers/chilean_plume/optical_depth_relations.pdf.

This optical depth limitation applies to ice cloud as well. I suggest you mention this point in the text.

9. Looking at the Qu et al. (2022) article, what would the results have been if you used the Hawaii rather than the Halifax test scene? The precipitation rates and cloud ice water contents and optical depths would be considerably higher.
10. Figure 15 is extremely informative. Maybe just below the top row describing the retrieval, put in which instruments are being used.