Dear Authors,
the manuscript “Observations of Coherent L-Band Emission from Snow-Covered Arctic Sea Ice” investigates whether a coherent or incoherent L-Band emission model is more consistent with in-situ observations from a 2024 measurement campaign in Cambridge Bay. The manuscript is clearly structured and well written. It finds that only the coherent model can reproduce the small polarization ratios (PR) found in some measurements.

However, I am concerned about the relevance of this study. It is not new that coherent emission models are required for good predictions of TB emissions in idealized settings. This has also been shown for the effect of snow on sea ice (e.g. Huntemann 2015, Fig. 4.20). The fact that this is confirmed by field measurements is good to see. This might well be the first time such low PD is observed in L-Band radiometry in the field in conjunction with consistent snow thickness measurements (but I am not an expert on this). These low PD measurements seems to be completely driving the improved performance of the coherent model in the manuscript and all further interpretation (which is recognized in the manuscript).

I am skeptical regarding the motivation, that this is relevant for satellite algorithms. For this to be the case the snow thickness variability would, for my understanding, have to be below 5cm within satellite L-Band footprints. I am not convinced, that those conditions exist in a relevant number of cases (considering the large variability of snow thickness and very large satellite L-band footprints). The argument, that these conditions might be found for new ice or land-fast ice, would have to be backed up by literature or data (though I agree, that this would be the most likely places to find them, if they exist).

Of course, the results do not have to be relevant for satellite retrievals in order to be a valuable contribution. Though, in that case the study requires in my view some refocusing. Two ideas come to mind, but of course others are welcome.

(1): Focus on local L-Band measurements entirely, illustrating the potential of ARIEL (-type) sensors mounted on ships, sledges, maybe helicopters. To show its (ARIELs) potential to derive sea-ice and snow parameters it would require to adjust the study setup. Currently you show, how well (or not) the models can be optimized to fit a set of sea-ice and TB measurements. Here you would want to show that one/some of the parameters can be inferred from the others, e.g. by applying your optimization on all but one of the parameters (such as snow thickness) and validate whether the predicted (=optimized) left-out parameter matches the in situ observations. How sensitive would that retrieval be to the other input parameters?

(2): Focus on validation and development of L-Band emission models. You show that one coherent model is more consistent with observations than one incoherent. This seems to be a good dataset to compare a number of available L-Band models to each other and see which ones are most consistent with real field measurements. Is the focus on thin snow layers in MEMLS a real limitation or is coherence lost in the data above 5 cm snow thickness anyways (as one could argue the equal performance of coherent and incoherent model for snow thicknesses near 15cm in your figure 5b suggests). If other coherent L-Band emission models are available, one would want to include them, too. The methodology presented here seems adequate for this investigation as it is. The description of the emission models and their differences would need to be expanded. It would be important to make sure that the comparison is not simply selecting the most flexible model.

As this will require a major revision, I will not go into detail regarding specific comments, but want to give some areas which will need attention:

-Description of the field sites: Ideally with a map of the measurements or at least a more general description. It is not clear what it means if measurements are described as ‘close’, it is not clear how far TB measurements are from each other, what it means if properties are described as ‘homogeneous’, etc.

-The naming of TB and sea-ice conditions (‘measurements’ and ‘observations’ respectively) is very confusing. The optimized x vector is repeatedly called observation even though \alpha is not measured at all and the other variables are optimized to match a self-consistend model state including the TB. TB and x are also called prior which is typically what we call expectations before they are updated by observations. I understand it refers here to the values before optimization, but since it is also the target (i.e. a small difference between optimized and its prior is considered a measure of model quality here) the naming is confusing.

-You never show simple forward simulations. What TBs do the models predict for x_a? This should be an intuitive first measure of how well the emission models perform to accompany by your optimization results.

-Figure 3: For TBV the incoherent value is clearly above the average of the coherent values. This seems to directly contradicting your statement in line 148. Can this be explained by the differences in models (besides coherence?)

-The used uncertainties are:

-Not always provided exactly, but sometimes as range. This is not reproducible and a reference to the code is not particular reader-friendly.

-base on not exact criteria. What is considered ‘homogeneous’, what ‘close’, what ‘few measurements’?

-surprisingly small for snow thickness and ice thickness. At least for thicker snow it is common to have more than 2cm difference within measurements of the same snow-pit. Please double-check that uncertainties are consistent across parameters. For example, the difference snow-thickness uncertainties of 1 cm and 0.5 cm might seem small, but it is effectively doubling the weight of snow thickness in your optimization.

I believe that you present a unique combination of measurements and a genarally robust methodology. But for the before-mentioned reasons I have to recommend major revisions.

All the best