The authors are performing an analysis of Multi-Layer Clouds (MLCs) in the Arctic using ICON model simulations over August 22 to September 23, 2020, and compare the model results to observations from the MOSAiC campaign. The authors use a newly developed *Arctic fit* immersion freezing parameterization, which includes an exponential for INP for colder temperatures and a second-degree polynomial fit for warmer temperatures. The model results show high sensitivity to mass thresholds in occurrences of MLCs and cloud seeding. The cloud occurrences and macrophysical quantities compare well with observations, even though the microphysical quantities are off often by orders of magnitude.

The manuscript is well written and presents novel work that contributes significantly to scientific progress. The manuscript is suitable for publication with minor revisions. The manuscript needs to address three aspects in detail.

1. The significant difference in Liquid Water Path and Frozen Water Path between the model and observations. Due to the high sensitivity on cloud mass thresholds, are the authors just coincidentally matching the observations on cloud occurrences, since the simulations have severe differences on quantities as fundamental as FWP and LWP?
2. The cloud seeding mechanism discussion could be expanded. The authors compare 1st layer MLCs and SLCs. A similar comparison could be made between seeded and non-seeded MLCs.
3. It would be helpful to include mean vertical profiles of clear-sky vs. SLCs vs. MLCs for more intercomparison. Also, differences in radiation between the SLC layers and MLCs could be reported, as that could be one of the main causes for the differences between 1st layer MLCs and SLCs.

Specific comments:

In figs. 3 & B1, why are the mean values outside the range? The values outside 1.5 times the IQR are excluded to simplify the interpretation, hence the mean should be within that range

The manuscript switches back and forth between using Celsius and Kelvin scale for temperatures. The readability could be improved by using a consistent unit (C?) for temperature throughout, and in cases where the other unit (K) needs to be used, provide the corresponding (C) values in parentheses.

Since the novelty of the work is focused on the newly developed *Arctic fit* immersion freezing parameterization, the authors should directly compare results from the ICON model using the Arctic fit parametrization and ICON model using the H15 (Hande et al. 2015) parametrization.

In Fig. 6, the temperature and qv are flipped in the figure compared to the caption and discussion.

Also, QV is referred to as specific humidity. QV is the water vapor mixing ratio, the ratio of water vapor mass to dry air mass, while specific humidity would be the ratio of water vapor to moist air mass.

In figs. 7 and 9, the authors should have 2-layer MLCs next to SLC, as SLCs are similar to 2-layer MLCs than >4 layer MLCs. I understand the authors want to focus on the robustness of the 2-layer MLCs across observations and models with different thresholds, but it's more sensible for the order to be CS, SLC, MLC (2), MLC (3), MLC (4), MLC (>4). The colors of CS and SLC could be switched for readability to have clear sky be represented by blue.

If the authors want to stress the robustness of 2-layer MLCs as more than just a coincidence, they need to do further analysis on what happens to the 3+ layer MLC regions when increasing the cloud mass threshold. What percentage of the 3+ layer MLC regions become CS / SLC / 2-layer MLCs upon increasing the threshold? What percentage of 2-layer MLC regions become CS / SLC upon increasing the threshold?

Line 291 "With a larger number of MLCs, either the SLCs or the clear-sky fraction has to decrease" - the lines before and after this mention the opposite, a decrease in MLCs and increase in SLC and clear-sky

One might argue that the "RS+Radar" observations are the best available data for comparison, as they are further validated by radar data. To that end, the authors should consider a cloud mass threshold of 10^-5 kg/kg, which following the trends in threshold would be closer to the RS+Radar data

Lines 327-329 - "This may be explained by a 12% (25%) occurrence of modelled thin MLC (SLC) layers that are less than 100m thick. These cloud layers would not be included in the observational algorithm." - The authors should consider using the same cut-off thickness for the clouds as the observation algorithm.

Line 332 – "(39m and 22m difference in model and observations, respectively)" – It would be better to report these differences in median thicknesses as percentages.

Figs. 3, 5, 8 and B1 – Violin plots would give the readers a better understanding of the distribution of the microphysical and macrophysical quantities.