Author’s response to reviewer 2, revision 2

We thank reviewer #2 for their detailed feedback on the revised manuscript. We have responded to each comment in detail below. A simple “Ok” means we have changed it, for larger changes, the sentence or figure in question has been included. Figures shown in this document are labelled using Roman numerals and referrals to the original figures are included when appropriate.

Review 2

After its first round of revisions, this manuscript is improved relative to the first version. There are fewer speculative statements, and some details have been described better. When pushed by both reviewers to add depth via additional simulations and analyses, the authors have instead argued for why these are not necessary. Fair enough; the authors can ultimately determine the scope of their study. At the same time, its impact on the community will be similarly scaled. The topic of ice processes and SIP in Arctic clouds is important, and the community does need more studies of this nature to help make the needed advances in understanding and modeling. Thus, perhaps this study is a baseline for very warm Arctic mixed phase and multi-layered clouds that can motivate the modeling community to dive further into the relevant processes. In this regard, it is useful to eventually see this manuscript published. There are, however, still some important details for the authors to address before the paper is ready for publication. All of the comments below can be addressed via text modifications, thus I would characterize these as minor revisions.

→ We thank you for your thorough comments. We will answer your points as they come up in the next section.

The Main Challenge

The main persisting challenge I have with this paper is encapsulated in the first paragraph of the discussion/conclusion section. The text outlines the two foci of the paper: 1) “to show that multilayer clouds can be accurately modelled….” and 2) “to evaluate the microphysical sensitivities…..while constraining the parameterisations to the observed ground-based measurements to better represent cloud phase.” The first of these statements suggests that in this manuscript the multilayer clouds have been accurately modelled. This depends strongly on the definition of accurate. If it is defined as simulating clouds at approximately the right times and heights, then perhaps it can be argued that these simulations were indeed accurate. But I argue that this type of accuracy is simply due to the advective forcing of the model (i.e., where and when is moisture advected at the large-scale). In my mind the main point of the paper, as conveyed in the title, is rather about microphysics, and if we refer to Figure 10 we see that the clouds are not accurately simulated.

→We understand the ambiguity of the word “accurate”. Our background to using this word even though the clouds are not perfectly captured comes from the fact that multilayer clouds have not been thoroughly evaluated in the high-Arctic with a realistic setup. We wanted to first make a point that they can to a certain degree be accurately captured with ICON to support further studies on these clouds.

We will add some distinctions to make the point that we do not perfectly capture the clouds clearer. We will split the focus into a thermodynamic and microphysics part. Thermodynamically, the 1st of September is accurately modelled while the 3rd was not, due to its initialisation as you stated. Microphysically, we do not capture the clouds very well. This is hopefully still encapsulated in the second point.

We change the line;

“.. 1) to show that the thermodynamical structure of real cases of multilayer clouds can be (to a certain degree) accurately simulated…”

Arguably the most important aspect of these clouds is the super-cooled liquid water, which ultimately determines the impact of these clouds on the environment and defines the cloud life cycle. The broader Arctic mixed-phase modeling community also appears to agree that the liquid component is the most important, based on the existing modeling studies in the literature. These simulations do not get the liquid component of the clouds correct and do not show any simulations that significantly impact the liquid component, on average (e.g., Fig. 10). Thus, in my mind, these simulations cannot be considered “accurate.” The second focus is on sensitivities in the microphysics. The paper has indeed examined sensitivities for the ice component, but interestingly, the simulations show no corresponding sensitivities for the liquid component. Thus, while in the most extreme sensitivity tests the observed ice component is indeed approximately reproduced by the model, there is no change in the poor representation of the liquid. So technically the cloud phase is “better represented,” going from very bad to not quite as bad, but it is not well represented.

→ We thank you for bringing up the liquid water in this discussion. We did not focus on the liquid component of the clouds primarily due to the large discrepancy in cloud ice water content. The liquid water differs by a factor of 4 compared to more than two orders of magnitude difference in cloud ice. Furthermore, we find that reducing cloud liquid is difficult in a model based on saturation adjustment. Cloud liquid will always (as long as cloud droplets are present) form during supersaturated conditions. Thus, we do not focus on this but show the impact of reducing the CCN concentration, which does reduce the liquid water path by 20%. We also explain the reason this change is not more substantial (saturation adjustment and upstream changes to the clouds). On the sensitivities, yes cloud liquid does not drastically reduce but this cannot be expected as the cloud ice, even at the largest scaling of INP still only reaches 10 g/m^2. We do see some sensitivities to cloud liquid though, for example in Lines 371-372, we see a 77% reduction in cloud liquid due to an increase in, primarily, vapour deposition with the large INP scaling (1E6). The lack of glaciation may be surprising but the clouds are found at warm temperatures where glaciation is less likely (Line 318).

But you are right, it is in itself an interesting finding that the cloud liquid is so persistent and we add a further discussion to the discussion section regarding this (see further below).

Moreover, it is worth pondering if the sensitivity to ice processes would be different if there were less liquid water (like in the observations) or if there were more sensitivity of the liquid water to the ice (as is seen in many other model studies). So are these sensitivities to ice parameterizations representative?

→We do believe the sensitivities would change if cloud liquid was lower. However, we still believe these to be representative sensitivities as we do not have unreasonably high levels of cloud water. Future case studies of multilayer clouds with contrasting thermodynamic structures and within different large-scale flows would be beneficial to elucidate their sensitivities. We will, however, leave this for future work.

I still believe this paper has interesting results that contribute to the field, but it is important that the authors present the paper in the proper context. First, there needs to be some discussion of the liquid component of these clouds. Why does the liquid component show such little sensitivity to the perturbations in the ice? Why does the model produce too much liquid to start with? And why is the model seemingly unable to produce sustained liquid water clouds that have much lower (more realistic) liquid water paths? One main goal for why we care about the ice is to properly simulate the phase partitioning in these clouds, so this topic must be discussed more.

→It is difficult to trace the origin of too much liquid. We believe the saturation adjustment is likely a culprit, together with misrepresentations in the initial and boundary forcings (1-moment ICON) which initialised and updated the boundaries with poorly constrained data. Liquid introduced into the domain may continue the cycle through the saturation adjustment while limiting the cloud droplet activation hardly influences cloud liquid water content given a pre-existing liquid cloud and the use of the saturation adjustment. We find that the saturation adjustment is not recommended for extreme maritime cloud cases (Kogan & Martin 1994), where the low CCN (~25cm^-3) gave rise to up to 40% error in condensed water vapour. However, follow-up work, about to be submitted (published in the thesis available here; DOI: 10.5445/IR/1000179667), where we look at a long-term simulation over the same area, has a lower prescribed CCN concentration and systematically under-predicts LWP compared to retrievals. This case might thus be an outlier.

Plotted below are the domain (85°N-90°N) and time (08UTC-12UTC) average process rates contributing (nucleation, 1st and 2nd saturation adjustment) and reducing (rain formation (accretion+autoconversion), riming by cloud droplets, and vapour deposition (onto ice, snow, and graupel) which within mixed-phase clouds shows the impacts of the WBF process) cloud liquid water content. Condensation (1st (left) and 2nd (right) saturation adjustment) is strong in this case. The rate for the 2nd saturation adjustment (right) is two orders larger than both the other contributing and reducing rates. While it also shows large negative rates (evaporation) it serves as a large source of cloud liquid within the cloud layers. Rain formation as well as mixed-phase processes are negligible in comparison and thus cloud liquid persists.

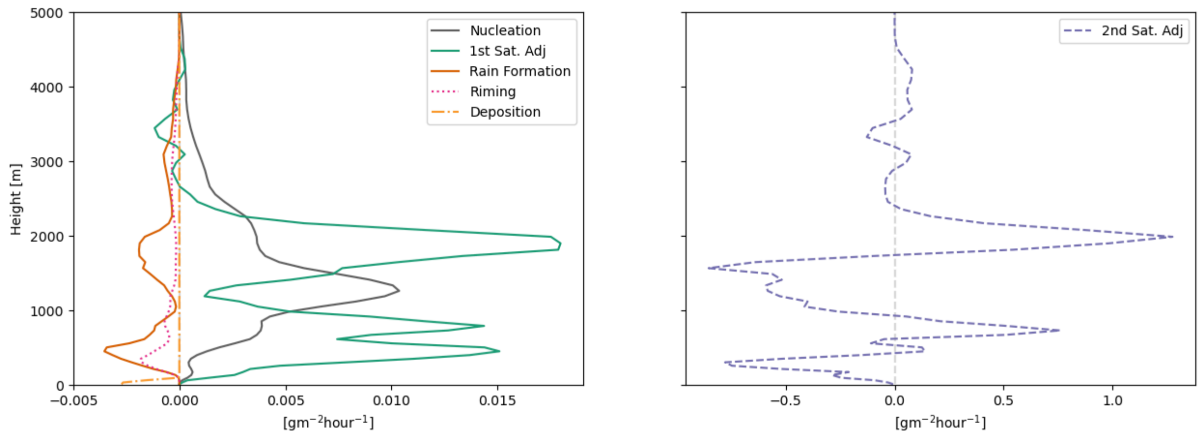

Fig I. Domain (85°N-90°N) and time (08UTC-12UTC) average process rates during the 3rd of September (1.6km simulations) contributing to (nucleation, 1st and 2nd saturation adjustment) and reducing (rain formation (accretion+autoconversion), riming by cloud droplets, and deposition (onto ice, snow, and graupel)) cloud liquid water content.

Future work should try to disentangle the complicated processes leading to this large impact from the saturation adjustment, including possible issues in CCN-limited regimes. Yet, we consider this out of the scope of this paper as the saturation adjustment in ICON has been thoroughly evaluated across the world including the Arctic (Schemann & Ebell 2020, Kretzschmar et al. 2020, Kiszler et al. 2023).

Considering the mixed-phase processes, an inefficient WBF process may also be blamed, this has been recently seen in ICON (Omanovic et al. 2024).

We add a discussion point on this subject in the discussion:

“The saturation adjustment (condensation) in the model ensures that during supersaturated conditions (with respect to water), condensation occurs, making a reduction in cloud liquid difficult to obtain. The initialisation of cloud liquid (using 1-moment ICON at a coarser resolution) enables the saturation adjustment to act throughout the simulation (as long as nucleation still supplies newly activated droplets and/or cloud liquid is not entirely removed). Overall, cloud liquid is persistent. While it responds to changes in CCN and large perturbations of INPs, we find it difficult to reduce the modelled liquid water content to observed levels. This can potentially indicate a lack of efficiency in the WBF process as seen by previous studies (Omanovic et al. 2024) or a possible issue using a saturation adjustment scheme in a low CCN environment (Kogan & Martin, 1993).”

Second, the authors need to provide the justification for why it is still relevant to examine the sensitivity to various ice processes in these cases where the liquid water is so far from accurate. Are these sensitivities generally representative and therefore relevant for thinner (more accurate) liquid water clouds? i.e., is it OK that the liquid component is not well simulated? I think the authors can make these points, and they need to make these points.

→As with all case studies, the applicability to other cases is limited. But we believe the sensitivities shown here bring up the important point that these clouds respond very similarly to single-layer clouds and this may be a transferable generalisation.
We add this to the discussion;

“The sensitivities shown here highlight the point that these clouds respond very similarly to single-layer clouds.”

The main purpose of the sensitivity study is to try to achieve the observed state. The sensitivities are introduced through the text as the next step to perturb the system into something observed rather than a study to test how the clouds react to sensitivities. Thus, we believe it’s highly relevant to do the sensitivity tests introduced in the paper. Your point makes us realise we may not have made this clear enough and add this to the discussion as well partially in the introduction:

“The model in its original form, severely underestimates the cloud ice mass concentration (by two orders of magnitude) while the cloud liquid water content is over-predicted by a factor of four. We perturb the aerosol parameterisations in an effort to 1) obtain the observed state, and 2) understand sensitivities in the cloud response to aerosols.”

Specific Comments:

Line 0: The title is much improved.

→ Thank you, we agree

Line 94: “ice floe” not “ice shelf”
→ Ok, changed.

Line 96: "….taken from level 3 files produced from Vaisala…."
→ Ok, changed.

Line 102-103: Thanks for mentioning the cloud classification is challenging. However, the newly included statement is misleading. Yes, having radar and lidar helps with cloud classification. Yet, in multilayered clouds (the focus of this study), the lidar signal is often attenuated by the lower cloud, rendering the phase classification in the upper clouds very unreliable. This is a distinct weakness of the CloudNet approach and adds significant uncertainty.
→Yes, the uncertainty is significant. We change the sentence to the one below to make this more clear.

"...During the days investigated here, both radar and lidar products were available for Cloudnet retrievals. In general, the highest classification confidence is given for clouds that are detected by both lidar and radar. Confidence decreases for the upper clouds in the MLC systems considered here as the lidar signal is attenuated by the lowermost cloud. For identifying MLCs, however, we utilise soundings and confirm the presence of clouds in saturated layers through lidar observations."

Line 103-104, 162: CCN and INP acronyms have already been introduced in the first section.
→ Thanks for noticing. Removed.

Line 106: "…..and from the US Atmospheric Radiation Measurement (ARM) facility…."
→ Ok, changed.

Line 108: Introduce the CSU acronym.

→ Thank you for noticing, changed.

Line 153-155: I understand the computational cost of high resolution. The manuscript mentions that a simulation was done at higher resolution, but with the baseline (default) approach to ice processes. According to the statement given here, this higher resolution did not have as big of an effect on the ice as the liquid. However, many of the other sensitivity simulations employ different details about the ice that might be sensitive to resolution. If the authors simply do not want to do further high-resolution simulations, that is acceptable, but a general statement like provided here that cloud ice is less effected by resolution seems to be inappropriate. It is possibly true for the one tested model set up, which itself already has many other problems, but unless the simulations are run, we do not know if it is true for some of the other ice processes that are tested and might produce more ice. Moreover, the fact that the high-resolution simulation gets some change in the liquid, when the other simulations conducted in this paper apparently do not, is very useful for addressing a major deficiency of the current paper – the fact that there is little impact on the poorly represented liquid phase by changing the ice processes.

→ Yes, we only tried one setup as it is very expensive to run and we found a lack of immediate improvement in the cloud phase partitioning. Of course, this does not show any sensitivities at finer resolutions, but simply that the baseline simulation at a fine grid spacing is similar to simulations at a coarser one.
We add for clarity:

“The cloud ice is less affected than the cloud liquid in this specific case.”

We believe the induced differences in cloud liquid are quite well understood with a higher updraft velocity as more cloud droplet activation occurs. The reduction of cloud liquid is really where the challenge lies. We refer you to our answer above for this discussion.

Line 352-353: I believe this is an incorrect assumption. In the observations there is at least a 1-2 km gap between the upper and lower clouds. It takes time for ice to fall across this depth. The lower cloud layer is less than 1 km thick. Thus, the timing for seeding vs fallout from the lower cloud layer does not work out for this hypothesized mechanism for forming ice in the observed lower cloud. Moreover, why would there have been seeding ONLY upstream for this whole time period? There is no clear reason why that would happen upstream consistently over time but not at the observation site. Typically, if there were seeding upstream the whole system would advect in time and eventually be observed at the Eulerian observation point. It is much more likely that either: 1) The “observations,” which are really retrievals, are wrong about the presence of ice in the lower cloud, or 2) there is a mechanism for forming ice in the lower cloud that does not involve seeding.

→The IWC retrieval (Hogan et al. 2006) in Cloudnet comes with high uncertainty. Furthermore, in this specific case, with temperatures close to zero we are close to the boundary of the valid retrieval range which adds additional uncertainty.

We will add a point on the uncertainty of the retrieval together with this statement:

“Uncertainties in the cloud ice retrieval may also be causing this discrepancy.”

Your second point is difficult to believe as the temperatures are so high during the radiosonde launches. The only other possibility here than seeding would be that the cloud has previously been colder and thus formed ice that is then advected to the site. We can add this point:

“...or colder temperatures that allowed for primary ice nucleation upstream.”

Line 354: “redundant” seems like the wrong word here.
→ The word redundant is simply to state that it does not make any sense to do further simulations, in regard to cloud ice sensitivity, when the temperature is too high for any of the parameterisations to be active. We will replace this with the synonym “unnecessary”.

Line 425-437: This paragraph talks about impacts on frozen hydrometeors in a couple of ways. While there are some sensitivities, I would argue that these are rather inconsequential. First, in Fig. 10 the logarithmic Y axis means that much of the range and apparent changes are so small as to not even be detectable from the observational perspective. For the most extreme case of INPx1E6 the FWP only gets up to 1 g/m2! And the differences among the other cases are all very small and well below 1 g/m2. Thus, while indeed the model suggests that there is SOME ice present, it is inconsequential in most cases. Similarly, towards the end of the paragraph there is discussion of ice particle collisions. However, at the highest concentration there is 1 particle per 50 L, and for the 1E4 simulation it is 1 particle per 500L! Again, the number of collisions resulting from these low concentrations is vanishingly small and inconsequential for the cloud processes. It seems to me that this paragraph could be summarized by the statement that while some of these model changes have slight impacts, ice processes are not significant in any of them.
→The changes between the 1.6km simulation and 1E6 are not inconsequential as the

increase in FWP exceeds 3 orders of magnitude. Thus, we are confused as to which changes you deem inconsequential here.

We may agree with you that some values are very small in the discussion in the section surrounding Fig. 10. To this effect, we exclude Line 360-365 to remove some discussion surrounding very low values of snow/graupel production from changes in CCN concentration. However, we want to keep the discussion about the high INP scalings to explain the production of quite appreciable snow (1gm^-2) from the lower layer caused by the seeding from above (the collision discussion). We restructure the sentences to make it clear that the collision discussion is in regard to the seeding from above and not within the lower layer.

Line 457-458: As noted in the text, the key transition in the lower cloud occurs just after 12, but the process rates are given for 6-12 UTC. Thus, do the processes described in the figure actually represent the apparently abrupt shift that occurs just after 12?

→ No they do not, that is true. After 12, very rapid glaciation occurs which of course would increase the ice mass concentration greatly, however, this period is not the focus to us as, in effect, the mixed-phase cloud is destroyed. Furthermore, after 12UTC, in all simulations, the upper cloud dissipates before 13UTC which is why we limit the analysis to 12UTC.

Line 459-461: Perhaps I am missing a key detail here, but how do these SIP mechanisms impact the ice mass? These processes are ones in which individual ice crystals turn into multiple ice crystals, such that there would be an impact on ice number but not on mass (other than the conversion of liquid droplets to ice in some of the processes). Ice-ice collisions, for example, should have no mass change, just a number change. Thus, I’m confused by how these are put in terms of mass rates. I believe they should instead be number rates. If not, please provide a clear explanation for the reader.
→ It is correct that SIP does not directly increase the mass of the system as only higher numbers are generated through SIP. However, these fragments have an associated mass which is what is shown in Figure 11.

These fragments are then free to interact with their environment. We find that they impact the ice mass concentrations through higher vapour deposition rates, and subsequently, snow through higher aggregational rates. For clarity in terms of the SIP impacts, we will change the figure into rates of number concentrations (see below).

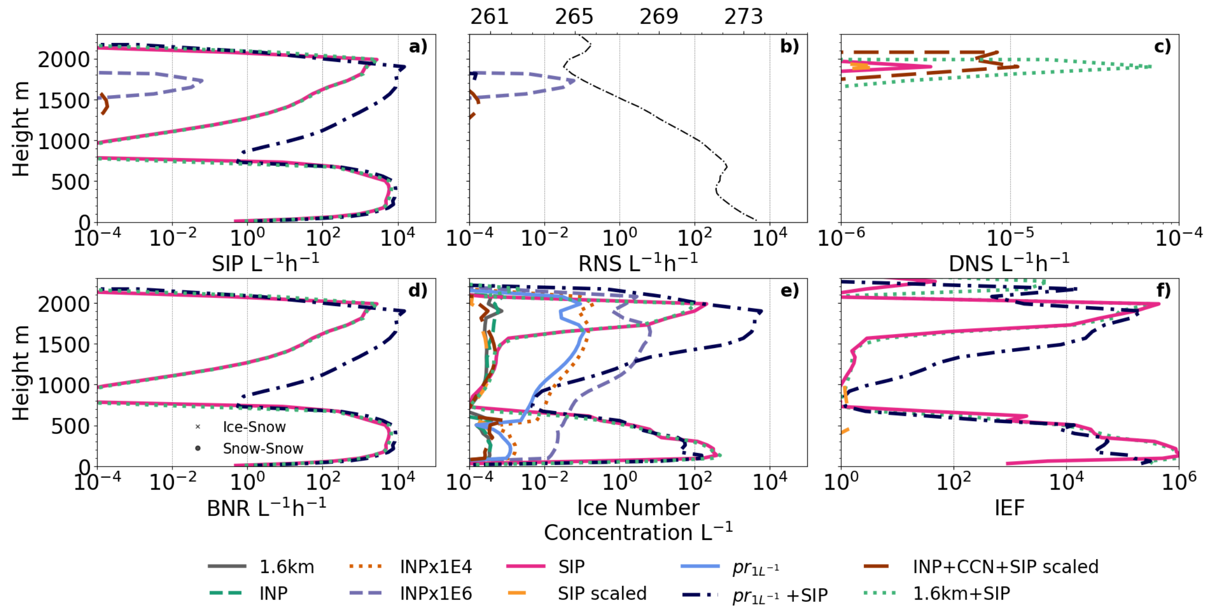

Fig. II: New Figure 11 with number concentrations of SIP rates. Please note the different ranges of the x-axis in panels c and f.

Line 468: Again, ice-ice breakup does not directly impact mass. The subsequent vapor growth of the new, higher number of ice crystals could impact mass growth, but that is no longer breakup but rather depositional growth.

→Please refer to the answer above.

Line 471-473: How much of a drop in temperature? This would be useful to know. Presumably this drop in temperature is due to advective tendencies because radiative cooling in the cloud layer should be minimal.

→We find an approximate 0.5°C reduction in temperature between 11-12UTC at cloud top. As the upper cloud is glaciating and dissipating at the same time, an increase in radiative cloud-top cooling could be expected in the lower layer.

Line 490: This is a nearly identical sentence to one that was included a couple of paragraphs prior.

→Yes thank you for noticing. However, this sentence is used to transition into the scalings of the breakup mechanism and thus we'd like to keep it for the justification of the scaling.

Line 491: Why is this not already "realistic"? What is the basis for thinking that breakup needs to be reduced?

→ Previous studies, as mentioned in Lines 186-191, believe this implementation of the breakup is not atmospherically relevant. For clarity, we may re-iterate this in the result section:

We add to Line 428:

"Previous modelling studies have deemed the contribution from the Takahashi et al. 1995 scheme too substantial and therefore introduce the concept of scaling this mechanism (Sotiropoulou et al. 2021, Georgakaki et al. 2022, Han et al. 2024). The scaled simulation (SIP scaled), whereby the breakup is scaled by the colliding particle diameter divided by the original particle diameter used in the experiments by Takahashi et al. 1995, nullifies the breakup contribution (Fig. 10 and contours in Fig.A1b)."

Line 495-500: I do not understand the SIP scaling. Is this simply an implementation of a different parameterization for collision breakup? If so, just be clear that it is a different parameterization that is not modified here but simply tested (and found to render collision breakup ineffective). My problem is that the word “scaling” makes it sound like there is some parameter that is being adjusted (“scaled”) to modify the effectiveness of the collision process. If you believe the default collisional breakup is too aggressive (first that should be clarified and justified), it should certainly be possible to “scale” the collision process by adjusting one of multiple components of the collisional breakup parameterization. But I get the sense that the authors do not want to explore such details. In any case, the language used here is unclear. If I understand what is being said, perhaps it is most clear to just state that the so-called scaling adjustment to the collision parameterization effectively nullifies the process altogether.

→ The scaling is introduced in Line 191. The breakup parameterisation is scaled by the diameter of the colliding particles divided by the diameter used in the original experiment. Thus, we explore physical constraints to this parameterisation instead of tuning pre-factors that we cannot physically constrain. We further studied this scaling with a CCN constraint following comments from previous reviewers. The comment added above hopefully clarifies this point.

Line 534-536: Yes, I agree that the assimilation and/or advective tendencies for key parameters are likely problematic here. These may be why the liquid water is not well represented and not very responsive to the ice microphysics. It is worth noting this point explicitly here, unless there are other explanations for why the liquid is represented so poorly.

→ Yes, together with the saturation adjustment discussed above, this is likely a source of some errors. We will avoid further speculative sentences about the upstream processes as these were removed in a previous version of the paper in response to comments raised by reviewer #2 in the previous round of reviews.

We add: ‘Improvements to the 1-moment scheme may further improve the cloud mass partitioning.’

Line 559-563: These statements are speculative and not supported by any evidence presented in this paper. i.e., I do not see any ICON-ART results that are shown to be more realistic for this type of case. Moreover, it has not been shown that better representation of updrafts, which might impact droplet activation, enhances the predictability of cloud properties. These statements should be removed. All that can really be said is that in future studies it might make sense to explore if ICON-ART and/or better resolved vertical motions can help to improve the representation of this type of cloud.

→These sentences were added as part of a previous reviewer's comments. We may, however, restructure them, highlighting their speculative nature.

We change the lines:

‘Prognostic aerosols using a dynamic aerosol model such as ICON-ART (Aerosol and Reactive Trace gases module) *could potentially* improve the representation of the local CCN concentrations and provide a more realistic cloud droplet activation. Furthermore, accurate representation of updrafts *may* ensure a more realistic simulation of cloud droplet activation, which *could* lead to an enhanced predictability of cloud liquid and layering.

Line 585-586: This statement is not true. According to Fig. 10 the 1E6 simulation produces a reasonable amount of ice in both the upper and lower cloud layers, better than any of the SIP simulations. Thus, according to these simulations, SIP (at least the added SIP parameterizations) are not required to reach observed levels of ice.

→ Yes, it does but that high level of INPs in this region is also quite unrealistic. We make this point a bit more clear;

“We find that a large scaling of this immersion freezing (by 1E6) captures observed levels of ice, however, the required abundance of INPs in the high-Arctic is unrealistic. We conclude that secondary ice production combined with increased primary ice production is required to reach observed levels of ice.”

Line 606: The number of grid levels is increased, while the grid spacing is thus decreased.
→ Yes, thank you for noticing, we’ve changed it.

Appendix B: This appendix has no text, only figures. Does this conform to ACP style?

→We believe it does. Appendix A goes into detail about the high-resolution simulations, where some text is provided to give the full story while figures B1 and B2 serve as additional information to the main text. The moisture profiles in Fig. B1 are shown to remove doubts about vapour being the culprit of the high clouds during the 1st of September while B2 shows the process rates during the 3rd, these are sporadically mentioned to support the statements on the microphysical processes in the text.

Additional Note:
A typo was found where we state the IEF impact from the SIP. The correct values are now included in the revised MS. We have edited the IEF values from 10^5 to 10^6 and 10^4 to 10^5 which correspond to the values in Figure 11f.