We thank the reviewers for their thoughtful comments and careful attention to details that have improved this version of the manuscript. Our point-by-point responses may be found below in red font.

We also used this time to make minor changes to the wording and organization of parts of the manuscript not explicitly commented on by reviewers.

Our responses are in red and the original reviewer comments are in a black font. Text reproduced from the manuscript is given in a red italic font.

Reviewer 1:

The authors present a well-written manuscript that analyzes one specific closed-to-open cell transition in a marine cold air outbreak. The analysis appears thorough and detailed and contains appropriate figures. My comments below are mostly of minor character.

Thank you

General comments

The abstract reads quite technical, mostly due to the frequent use of specific numbers/values. Please consider giving a broader picture first before going into the details. E.g., it could be more interesting to mention that the studied case is a relatively polluted and cold one, compared to other studies, than only giving the numbers in the abstract. Also, you could mention in the abstract already that you identified the decrease in aerosol number due to precipitation and entrainment as a key factor for the timing of the closed-to-open cell transition (if I understood you correctly from the conclusions).

We have revised the abstract so that it is easier to read and summarizes the main findings of our manuscript. The abstract now reads as:

> *"Closed-to-open cell mixed-phase cloud transitions within marine cold air outbreaks remain poorly understood despite their importance to high-latitude climate. The Cold-Air outbreak Experiment in the Sub-Arctic Region (CAESAR) research aircraft sampled a cold and polluted closed-to-open cell transition with cloud-top temperatures below -20 °C and upstream cloud condensation nuclei concentrations surpassing 600 cm^{-3}. The aerosol likely originates from Siberian industrial emissions. As the boundary layer deepens prior to and across the transition, driven by strong turbulent surface fluxes and aided by a weakening cloud-top inversion, aerosol number concentrations decrease and supercooled droplet sizes increase, causing ice particles to shift from primarily vapor depositional growth towards riming. Precipitation shafts within open-cells contain abundant large graupel (> 5 mm diameter) with liquid-equivalent precipitation intensities*

reaching 3 mm h^{-1}. Ice precipitation co-exists with lidar-observed surface cold pools, modulated by entrainment events, juxtaposed with surface-based plumes of warm moist air. Updrafts support liquid water production along convergence boundaries at the upwind cold pool edge. Nonetheless, strong buoyancy fluxes exceeding 200 W m^{-2} prevent sub-cloud decoupling. Ice particle number concentrations increase tenfold across the transition, surpassing 10 L^{-1} within open-cells, more than 1000× ice nucleating particle number concentrations. Given that cloud temperatures are too cold for Hallett-Mossop rime splintering, this indicates the importance of other secondary ice production mechanisms. This case expands the observations needed to better understand mixed-phase Arctic cloud processes."

I am not entirely sure yet, what role the output of AROME-Arctic plays in their study. While it is definitely valuable to compare the archived model output to the observations, the simulations are not described in great detail in the manuscript. Specifically, I am missing information about the aerosol (CCN and INP) concentrations applied in the model. As I see it, the statement in the beginning of the conclusions that the transition is incorrectly depicted by the model (line 499), is both unclear (convective bands have not been introduced earlier) and lacks discussion beforehand. How do the CCN and INP number concentrations in the model compare with observations and may differences there explain differences in the depiction of the transition?

This is not a modeling study, and the only use of AROME-Arctic is to provide a high resolution reanalysis to fill in gaps where we lack observations. I have clarified this in lines 171-173:

"Archived initializations of AROME-Arctic, a convection-allowing 2.5 km spatial resolution model developed by the Norwegian Meteorological Institute, serve as a high-resolution reanalysis (Müller et al., 2017)."

I have removed mention of the convective banding shown in the AROME-Arctic initialization. Line 462-465 now reads as:

"In situ thermodynamic measurements confirm the synoptic pattern depicted by AROME-Arctic, supporting the inference that a cooling FT weakens the cloud-capping inversion, aiding the transition, as documented for warm-cloud transitions (Sandu and Stevens, 2011)."

The CCN concentrations in Fig. 13 and 14 (h) show an irregular frequency of scattered points. Sometimes points with larger distance than 3-4 km are connected by lines, and for larger distances between points the line is interrupted. Why does the distance of the points vary (different aircraft speed, values not shown due to not overcome thresholds, instrument failure or other reasons?) and how are larger distances/gaps treated? Please also confirm whether the concentrations shown in (h) and (i) are all from in-situ observations.

CCN were sampled by the W-CCN two times per minute. Only valid W-CCN measurements were plotted in Figures 13 and 14. The following tests were applied to establish if a W-CCN measurement was plotted.

1) All three requirements, explained in the W-CCN readme document (ReadMe), were satisfied:

> The following three conditions are requirements for a CCN concentration measurement to be written into an output file: 1) A CCN detection interval must contain 28 detector signal (VDET) samples with 26 of these from a scattering detection interval designated by the instrument status (ASTAT), and in addition, one sample of VDET from the chamber flush before the scattering detection interval, and one sample of VDET from the chamber flush after the scattering detection interval. 2) The absolute value of the rate of change of the 28 VDET samples must be less than 0.0095 Volt per second. This rate was evaluated from the beginning to the end of the 28 samples. 3) At least 10 of the 28 samples must lie above a straight line drawn from the beginning to the end of the 28 VDET samples.

2) A valid W-CCN measurement could not coincide with a LWC measurement > 0.01 g m^{-3} (lines 130-131)

3) A valid W-CCN measurement could not coincide with IWC > 0.1 g m^{-3} (lines 130-131).

A plotting error is responsible for why points were sometimes connected by lines and sometimes not. We have decided to remove all lines connecting observations to avoid a reader misinterpreting the lower temporal resolution W-CCN data in Figures 13h and 14h.

Confirmation is now added to the aerosol methods section that all concentrations shown in h) and i) are in situ in lines : "*All aerosol and CCN data were collected in situ, by instrumentation inside the aircraft cabin and the wing-mounted PCASP cloud probe.*" (lines 132-133)

The caption in Figure 13 has been updated to clarify what is plotted in (h) and (i):

"...(h) W-CCN N_{CCN} (black diamonds) and UHSAS/MPSS N_a (dark green line and orange filled circles, respectively); (i) In situ N_i (L^{-1}, blue line) and CFDC derived N_{INP20} (L^{-1}, magenta cross indicating sample duration and error bounds) … "

Specific/minor comments

Line 187: Typo: should be "Scandinavia", not Scandanavia. *corrected*

Fig. 4, 13, 14, 15 and others: The "rainbow"/jet colormap is not perceptually uniform. Please consider shifting to a colormap that is perceptually uniform (this refers to the gradients inside the colormap).

Figure 4: RH, temperature, colormaps have been modified to be perceptively uniform. Thank you for this suggestion as they also look nicer now!

Figures 13,14: Modified IWC, thetaV, thetaE, and WVMR colormaps to be perceptively uniform

Figures 15: Modified wind speed colormap to be perceptively uniform

Line 224: What do you mean by “not yet shown”? *Changed “not yet shown” to “discussed in Section 5”*

Line 249: Could you add information here, what causes the higher aerosol concentrations downwind? Which sources are there between Villum and the flight path?

Revised lines 262-266 to the following:

“BL CO concentrations of 140-150 ppbv measured during vertical profiles are consistent with values measured at Villum between 21:00 UTC 28 February and 01:00 UTC 29 February (Fig. 6b, Fig. 8a-c). This suggests new aerosol sources, such as from sea spray production during transit over open water can help explain the N_a increase. Sampling differences between the two locations (surface observations in a vertically stratified environment at Villum vs. airborne observations in a mixed BL during RF02) may also contribute.”

Fig. 7, 11 and 17: Do you need the color bar? If the curves are of only one color each, you could also consider to just give the distance in the legend or caption.

While the colorbar is not necessary, having the color of each of the lines determined by a colormap can help the reader interpret the results when comparing multiple plots. We chose to keep the colorbar in Figure 7, but remove the repetition in Figs 11, 17, and C1 in light of the reviewer’s comment.

Line 263: Consider adding that you talk about “vertical” profiles here.

Changed line 273 to: *“The four vertical profiles bounding SL1 and SL3 straddle 450 km (Fig. 1), …”*

Line 272/3 and 290: How certain are the altitude estimations for inversion and cloud top height? Is the overshoot robust?

The altitude estimations for cloud top height during vertical profiles are determined by identifying the first (for descents) or last (for ascents) altitude that CDP LWC > 0.01 g m^{-3}

The altitude for inversion height is determined by identifying the altitude of the minimum temperature.

Within Section 2.1 (Methods: In-situ Cloud) we have clarified how each of these metrics are determined:

Lines 112-114: "*During profiles, the cloud-top height is determined by identifying the highest altitude with LWC > 0.01 g m^{-3}. The inversion level is defined as the altitude with the minimum temperature.* "

Both of these metrics utilize 1 Hz data. The average ascent/descent rate during spiral profiles was ~5 m/s.

Since an overshoot was identified in both ascending and descending profiles, differing response times between the CDP and temperature could not explain the observed overshoot. This gives us confidence that the identification of overshooting tops is robust, based on the metrics we have applied. If the reviewer believes we should use different metrics we could try those as well.

Fig. 9: For the sake of colorblind-friendliness, I recommend changing one of the colors red or green.

Figure 9: Changed green to dark green which makes each line distinguishable when input into the colorblind simulator while maintaining the intuitive color choice of green representing moisture and red representing temperature

Figure 10: Changed IWC to a darker blue to make the contrast with the magenta LWC more colorblind-friendly

Figures 7, 11, 17, C1: Modified colormap to be more colorblind friendly

Fig. 13/14(f): Could you use the same color scale for both figures anyway? SL1 goes hardly ever below 1 g kg^{-1} anyway, as far as I can see.

Modified Figure 13f to use the same colorscale as 14f.

Fig. 14: Typo in caption: Locations where cold pools (entrainment) are identified "are" underlined in blue (gold) in (d)-(f).

Corrected to: *"Locations where cold pools (entrainment) are identified are underlined in blue (gold) in (d)-(f)."*

Line 411: Typo "Aitken" mode, not Aiken. Corrected

Section 7: The answer to the question proposed in the section title seems hidden in line 435-437. You could consider repeating that at the end of the section.

The answer is now reiterated in lines 594-596:

"While the closed-cell convection is likely primarily maintained by the high concentrations of BL accumulation mode aerosol, further modeling work is needed to assess the contribution of FT Aitken mode aerosol for the transition taking place later (e.g., Wyant et al., 2022)."

Note that the discussion of Aitken mode aerosols was moved to the Appendix D as the quantitative analysis it contained did not fit well in the body of the manuscript. A brief summary of the findings of Appendix D are mentioned in lines 407-411.

Line 485-489: Do you mean that collisional breakup is followed by increased transfer from the liquid to the ice phase through the Wegener-Bergeron-Findeisen process? The ice-ice collisions can't be causing the depletion of LWC themselves.

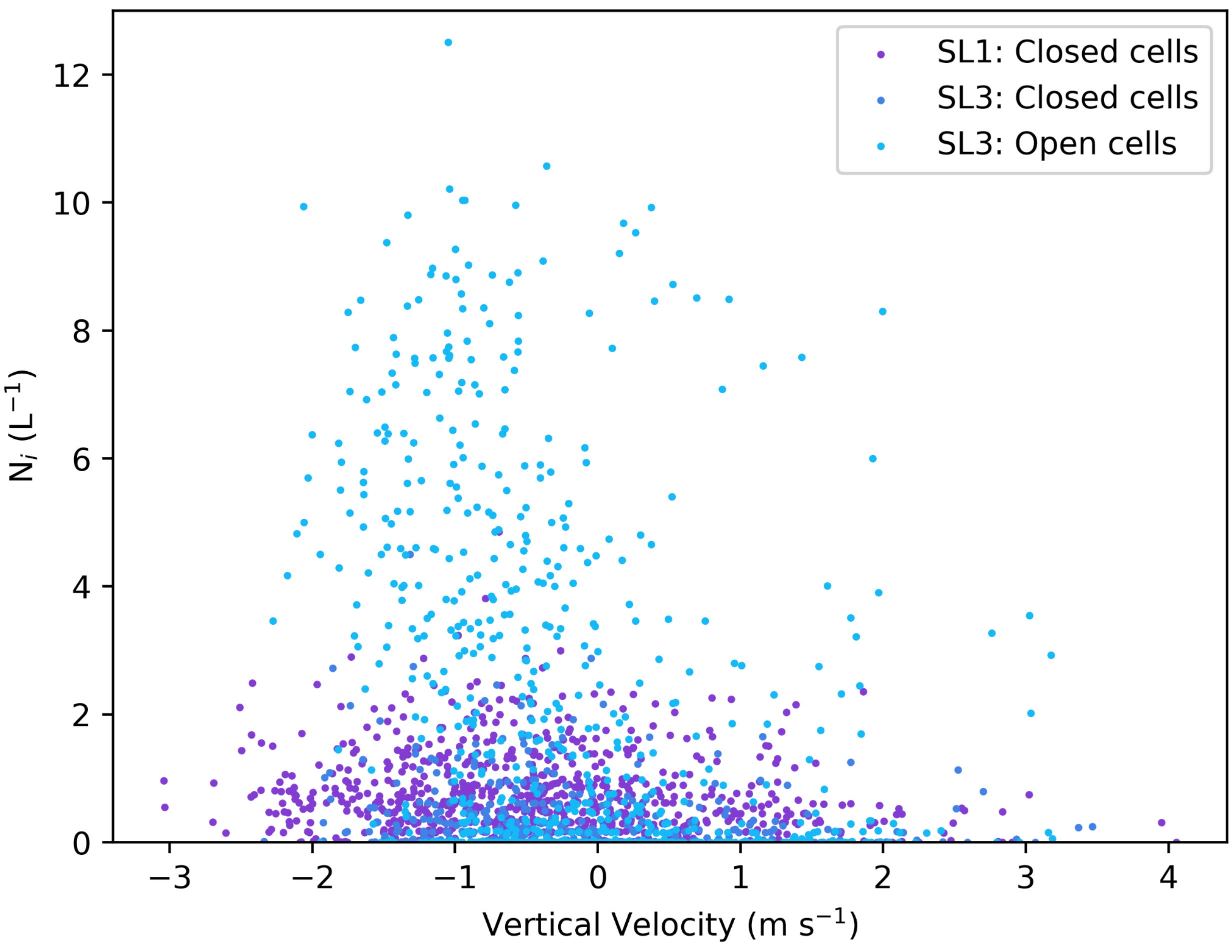
We clarify that while the minimum of LWC is not because of ice-ice collisions, the sharp local minima of LWC is consistent with liquid depletion through WBF processes and/or riming in the presence of enhanced ice number concentrations.

Lines 447-451: "*One supporting observation is that, at an altitude of 750 m during the descent to SL1, N_i reaches a concentration of 6.5 L^{-1} (500× higher than N_{INP20}), collocated with a spike in IWC to 0.25 g m^{-3} (and a depletion in LWC consistent with the enhanced removal of supercooled liquid through WBF processes and/or riming in the presence of higher N_i) at a temperature of −16.3 °C (Fig. 10a).*"

Line 493/4: This is an interesting suggestion. Do you have observational evidence for how frequent the recirculation of falling into updrafts occurs or on which scales this process is relevant in Arctic marine cold-air outbreaks? Does this e.g. depend on the width or intensity of the cold pool?

While we are unsure of the frequency/scale of recirculation within cold air outbreaks, this comment encouraged us to look more closely at whether we have evidence of recirculation during RF02. We have examined the in-situ ice particle number concentration as a function of in-situ vertical velocity, to see if N_i are indeed contained within updrafts. This is now the new Fig. S2, included here as well. While N_i is higher within downdrafts, as might be expected, N_i > 1 L^{-1} are also present in updrafts, particularly in the open-cell region. The updraft vertical velocities can exceed the fall speed of the ice particles, estimated crudely to be 1 m s^{-1} for this exercise. These N_i can potentially be brought up further to the cloud base, where the N_i can easily interact with any developing liquid cloud. We have added the following:

Lines 454-459: "*Ice is also present in updrafts with in situ vertical velocities > 1 m s^{-1}, crudely on par or exceeding the ice particle fall velocity (Fig. S2). N_i is greater than 0.1 L^{-1} (1 L^{-1}) during 43% (21%) of the flight-level samples below open-cells with in situ vertical velocities > 1 m s^{-1}. N_i values exceeding 100× ambient N_{INP20} within updrafts indicates that their recirculation into open-cell convection could increase further N_i production (Korolev and Leisner, 2020; Lasher-Trapp et al., 2021).*"



Line 537: Typo "THREDDS" server, not TREDDS. Corrected

Line 570/1: There seems to be a slight underestimation of droplet number concentration in the model compared to observations, even though the ranges overlap (as you correctly state). Could you comment on that possible underestimation?

We are not sure on why there is a slight underestimation of modeled N_d with respect to observed N_d during SL1. We already account for the uncertainty in total aerosol concentrations (20%) used as an input into the activation model along with the uncertainty of measured droplet number concentrations (20%). Remaining sources of uncertainty are our determination of w*, our choice of kappa, any model error, and the presence of a spatial offset between surface leg aerosol measurements and observation of N_d during spiral profiles.

w*: Since modeled N_d approaches its asymptote with increasing w*. Even a large (1 m s^{-1}) increase in w* only results in a 5-10% increase in N_d

Kappa: Even if kappa is increased from 0.8 to 1.3, the hygroscopicity of sea salt, modeled N_d only increases by 5%.

Model error: For "moderate" updraft speeds (0.5-2 m s^{-1}) in "light" aerosol conditions (250-1000 cm^{-3}) model errors are expected to be less than 10% (Rothenburg and Wang, 2016).

Collocation: Since during SL1 we are comparing the activation model applied to aerosol observations along the surface leg to observations of drop number concentrations on profiles bounding both ends of SL1, this is not a major concern.

We add to lines 555-556: "*The modeled N_d skews towards the lower end of observed N_d (Table C1), for reasons that remain unclear to us.*"

We too wish we had more insight into what explains the slight bias.

Reviewer 2:

The manuscript presents observations from a single research flight comprising three low-level flight segments. The authors combine SL1 and SL3 to provide an approximate quasi-Lagrangian view of a closed-to-open cell transition in mixed-phase marine clouds. The dataset is rich, combining aerosol, cloud microphysical, thermodynamic, and remote-sensing measurements. The combination of these complementary measurements is a major strength of the study and it provides a useful basis for future modeling work. The figures are informative, the manuscript is well written, and the study is suitable for publication in ACP. I am not an expert in all the measurement techniques used and therefore cannot fully assess instrument-specific uncertainties. Nevertheless, the observations appear physically consistent, and the proposed interpretation is plausible.

The manuscript is in parts very technical and descriptive and contains perhaps too many numerical details in the main text, which sometimes obscure the key points. I suggest moving some to tables or to the Supplement and retaining only the most important numbers in the main text. The abstract could also be revised to improve readability and emphasize the main findings more clearly. In this respect, the conclusion section provides an effective and concise synthesis of the results.

We have made the following revisions to make the manuscript easier to read:

- We have revised the abstract so that it is easier to read and summarizes the main findings of our manuscript
- Removed mentions of times corresponding to different fetch distances/altitudes, and instead relocated that information into a supplement (Tables S1, S2)

- Restructured discussion of the vertical profiles bounding SL1 in lines 280-292 by removing unnecessary quantitative information.
- Relocated section discussing the Aitken aerosol budget to Appendix D, while summarizing our findings in lines 407-411.
- Removed other unnecessary quantitative information throughout the manuscript

The manuscript appropriately acknowledges that aerosol, microphysical, and boundary-layer processes are strongly coupled and that modeling will be required to quantify their relative roles. However, some statements appear to present the proposed sequence as causal. I suggest clarifying which links in the sequence from aerosol changes to enhanced precipitation, cold-pool formation, and open-cell organization are directly supported by the observations, and which remain interpretations or hypotheses.

Added the word "hypothesized" to the section title near line 316: "*Cold pools and entrainment are hypothesized to organize the closed-to-open cell transition*"

We added the words "likely" and "appears" in lines 433-435 to clarify that the link between riming and the transition is a hypothesis: "*The reduction in N_{CCN} accompanying the rapid BL growth prior to km 282 - likely brought about by entrainment facilitated by a weakening temperature inversion, subsequently accelerated by aerosol deposition - appears critical for setting up the internal conditions facilitating riming, more precipitation, and thereby the transition in cloud morphology.*"

Added the phrase "is associated with" in lines 335-336 to avoid speculating that the drying BL causes a decrease in LWP and IWP: "*The drying of the BL is associated with decreasing LWP and IWP (Fig. 14g), with the first clearing observed at km 350 (Fig. 14a).*"

Removed "ultimately impacting cold pool formation" from lines 412-413 to avoid speculation: "*Here we discuss how reductions in N_{CCN} and N_d, and the resulting increasing droplet size, increase the riming efficiency with fetch,~~ ultimately impacting cold pool formation~~.*"

Added the word "appear" to lines 478-479 to avoid speculating on whether cold pools are the cause of the transition: "*The cold pools appear to dynamically organize the convection through surface wind convergence, facilitating the transition to open-cell convection.*"

Figure 1 indicates the satellite acquisition time and the aircraft position at that time. However, because the flight segments were sampled over a substantially longer period, the authors should discuss to what extent the cloud morphology observed in the instantaneous satellite image is representative of the clouds sampled during the different parts of the flight. If suitable satellite observations are available, an image sequence spanning the flight period could help demonstrate the temporal evolution of the clouds.

Figure 1 was modified to add satellite images from before and after sampling the closed-to-open cell transition.

Line 245: the units appear to be incorrect.

Corrected line 251-252 to: *"A small enhancement in aerosol concentrations from 10–30 nm (Fig. 7) is potentially indicative of new particle formation."*