Review of ACP-2021-888: "**Aerosol-cloud-turbulence interactions in well-coupled Arctic boundary layers over open water**" by Jan Chylik, Dmitry Chechin, Regis Dupuy, Birte S. Kulla, Christof Lüpkes, Stephan Mertes, Mario Mech, and Roel A. J. Neggers

Ian Brooks

Overview

This paper examines a large eddy simulation of a convective cloud system in the Fram Strait, based on and validated against airborne measurements from the Polar 5 and Polar 6 aircraft. It is a not untypical case for the marine environment on the margins of the Arctic sea ice, but perhaps less studied than the near-neutral boundary-layer capped by stratus/stratocumulus that dominates over the sea ice. Given the rapid warming of the Arctic and increasing loss of sea ice, the conditions represented here are arguably likely to be more prevalent over a wider area within the Arctic in the future, motivating a more detailed study.

The nominal aim of the paper, encapsulated in the title, is a study of 'aerosol-cloud-turbulence interactions'; the analysis of these interactions is, however, rather cursory and for the most part does little to push forward the scientific understanding of these convective cloud conditions. The analysis of aerosol interactions really does little more than demonstrate that the model's CCN sinks (snow and graupel formation) do indeed remove CCN.

I get the strong impression from mentions of planned or potential future studies, that this paper is primarily intended as a description of a baseline case (the only model run discussed is referred to a a 'control simulation'), demonstrating that the model successfully represents the observed conditions, with the real science – process sensitivity studies, etc – expected to be covered by future papers. In that case the earlier parts of the paper (notably section 4: 'Evaluation') covering the basic validation against the aircraft observations, might be better presented as supplementary material to a paper more strongly focused on a scientific question.

In many places the details required to properly assess the data visualised is not given, e.g.: times of model outputs, periods over which averages/PDFs are generated. Specific cases are noted below. A good paper should provide enough information for a reader to repeat the analysis – we might be able to download the data, but are not always given sufficient information to know which bit of it is needed to reproduce a plot.

The analysis focuses on a single convective cluster, located within a region of about ¼ of the larger domain, and over a short period of time (~1.5 hours). While it is not unreasonable to assume that the behaviour of this convective cluster is representative of such features in general, we are not given any evidence to support that assumption. Is this the only region of convection over this time period? If so, then the choice of what portion of the domain to assess variables over may bias the results since the convection takes up a larger fraction of the sub-domain than of the whole domain. Similar issues pertain to the time averaging…how often do such convective events initiate? Are the results presented truly representative of the wider environment, or biased by the selection of one event? There is much to be learned from case studies of single events, but it they give no idea of the natural variability of such systems, and it may be misleading to try and generalise from them.

The analysis in figures 14, 15, and 16 is informative, but far more could be revealed with conditional sampling. For example in 14a, three points representing layer means are shown. Those for the near-surface and 'entrainment source' are well defined (although we can't tell where the entrainment source actually is in the vertical), but the mean for the layer from which all the individual points are taken has a value that is likely to depend on the ratio of cloud to clear-air. It would be instructive to also see the means for cloudy and clear air, with the latter being more representative of the background environment in which the convection is embedded.

Similarly, a lot could be learned from conditional sampling of points in updrafts, downdrafts, by liquid or ice content (beyond the simple extremes already indicated). And by looking at the time evolution of the system.

I am left rather frustrated by what is clearly a very extensive LES simulation by a very capable model, but for which the analysis presented is very limited. I'm not sure what, of the results shown, would actually be worth citing other than by future studies using this model and base case, as a demonstration that the model accurately represents the observations.

Detailed comments

Line 30-31: The statement "A large variety of cloud forms occurs at high latitudes; one possible way of classifying these can be based on the properties of the underlying surface" mixes up different things. The surface below the clouds is more relevant to the boundary layer environment within which the clouds exist than as a means of classifying the clouds.

The authors then define 3 regimes (pack ice, marginal ice zone, and open water) and then state that only the open water case will be studied. The whole bit about cloud forms and classification regimes seems rather unnecessary – as if they feel it necessary to justify looking at clouds over open water rather than ice.

Line 48: "LES stands for the numerical simulation of the discretized set of primitive equations for geophysical flow…" – I know what the authors mean here, but the words 'stands for' give the wrong impression. 'LES' stands for (ie is an acronym for) Large Eddy Simulation…which is all ready established. "LES represents…" or "LES models perform a numerical…" might better convey the intended meaning.

Lines 80-86: there's quite a bit of repetition here from the previous section. Recommend providing the details in one or the other, but not both.

Line 97-98: "The air mass in the Fram Strait was slow-moving on this day, being situated over relatively warm open water." – the phrasing here, 'being situated…' implies a causal link: that the air mass was slow moving *because* it was overlying warm water. That is not the case, change 'being' to 'and'.

Figure 7: the shading in the plots, representing the spread in values, are very narrow. It really isn't worth including the contoured 5,25,50,75,95% intervals – the really can't be seen, and provide no useful information here.

Figure 9, and line 328-331: what time periods are LES PDFs of cloud top height generated over? Is this a single time or aggregated over many output times?

Figure 11, and line 360: "LES shown as a distribution of (resolved) values occurring at each height. Note that these distributions reflect single-gridbox values" – it isn't clear exactly what is meant here. Do these profiles represent the mean (and shading the percentile ranges) of all the grid boxes at a specific altitude at one time, of a single column through the model over a period of time, or of all the grid points on a level over multiple output times? If there time averaging, over what period?

Figure 11, lines 366-367. The authors state “Both the simulation and the noseboom observations suggest… featuring a concave w’^{2} structure, a convex T’^{2} structure…”, this doesn’t seem to match the figure. The w’^{2} results below ~250m are rather different for LES and obs, with the LES decreasing towards the surface while the observations increase to a maximum at their lowest altitude (~50m). The curvature of the observations below 500m appears convex for both w’^{2} and T’^{2} contradicting the statement, and for T’^{2} the LES and observations have opposite curvatures, not the same, as implied by the text.

Then at line 372 the text claims the T’^{2} profile has a concave structure, contradicting the previous statement.

Figure 12. At what time are the vertical cross sections obtained?

Figure 13: The colorbar text is far too small – I can’t read either the variable units or scale values without zooming in a long way. (same applies to a less extent on figure 12).

Line 391, Figure 13: at what times are the model outputs obtained? We’re told they are at 15 minute intervals, but no more. How do these times relate to the time of the vertical cross sections in figure 12? I assume (hope) on of them is at the same time.

It would be useful here if the vertical level of the figure 13 horizontal cross sections were indicated on figure 12 to help put them into context. And similarly if the location of the vertical cross sections in fig 12 was indicated on figure 13 (maybe just on the panels from the same time…assuming they are at the same time).

Line 394: the text referring to figure 13 states “…showing up as narrow maxima in w and q_t and…” but ‘w’ is not plotted in figure 13.

Line 397: “This ring forms some times after cluster onset, and over time radiates out…” how long after cluster onset? How long does it take to move outwards? The lack of any information on temporal evolution of this system is a weakness of the analysis.
Also ‘times’ should be ‘time’

Line 404 and elsewhere: The plots in figures 14, 15, and 16 are referred to throughout the text as representing ‘joint-pdf’ analyses. These are not joint-PDFs, there are no probability distributions presented here. These are simple scatter plots of individual data points. Figure 14a is formally a mixing diagram (as per Paluch, 1979) since the variables plotted are conserved under adiabatic processes and phase changes of water. The remaining panels in all three figures are simply scatter plots rather than mixing diagrams, since at least one variable in each case is not a conserved quantity. This distinction is not made clear, and confusion might ensue for a reader unfamiliar with this, since significant discussion is made of 14a as a mixing diagram, and then the caption for figure 15 states ‘Same as Fig. 14’, which is true in that they are both scatter plots, but misleading if one thinks of fig 14 as a mixing diagram.

Line 406: “the 3d time point” – should this be “the 3rd time point”?

Line 407: “The width of these probability density functions (pdfs) represents the variance at the…” – no, the width of the scattered points represents the full range of data values. We cannot judge how many points might cluster on top of each other near the centre of the distribution – the more of them there are, the smaller the variance. We need an actual PDF to be able to judge the variance.

Figure 14 caption: “Scatterplot of various state variables at a horizontal cross-section at a selected level and time point…” – what time? Which point(s)?

Line 422: “and a point on the mean profile (black plus) that is situated just above the local mean state at this height (black circle)” this is…unhelpful. What is meant by ‘just above’…how far above. what constitutes 'just above'...how far is that?

The figure legend indicates that the black + is the "entrainment source" and the caption that it is the "mixing source". 'Mixing source' doesn't help identify *where* it is, and while 'entrainment source' implies a point within or just above the inversion, exactly where, and how is this level selected? Give a specific, quantitative definition.

Line 425: "A second cluster of datapoints can be distinguished, situated around the local mean state in a fairly horizontal direction." – 'in a fairly horizontal direction' is rather vague, what is 'horizontal' on a plot so q_t vs θ_{li}? You mean that it has a wide range of θ_{li} while clustering around a narrow range of q_t.

Line 426: "The widths of this pdf in both directions reflects the turbulence in the stratiform cloud layer" – the range of values of both variables represents something of the mixing history, but not necessarily of active turbulence. If turbulence stopped, this variability would remain.

Line 439: "This height is about 100 m above the level of diagnosis" – what is the 'level of diagnosis'? Do you mean the level at which the horizontal cross section is taken?

Line 447: "Apparently the highest cloud ice values are not found in the rising cores" – we already know this, it was observed in figure 13.

Line 449: "Accordingly, the chevron shape seems caused by the onset of glaciation in updrafts," – isn't this inconsistent with the updrafts being 'almost ice free'?

Line 454: "This confirms the idea that the process of glaciation mainly takes place in air that was until recently part of an updraft" – we might infer this, but without information on the time evolution of the system it is perhaps rather overstating the case to say it 'confirms' this.

Line 459: "the snow always sits close to the cloud ice points, being their source" – the grammar inverts the causal relationship, as stated this implies the snow is the source of ice, rather than the other way around.

Line 466: "which is almost non-moving vertically" – yes the centre of the distribution is at w ~=0, but it spans the range $\pm$0.5 m/s with outlies to ~$\pm$1 m/s – so most individual grid points do have an appreciable vertical motion. Typical for a turbulent layer.

Line 477: "Snow and graupel formation removes glaciated water mass, a process that also increases gridbox buoyancy" – having estimated the buoyancy contribution from latent heating, could you not also estimate the contribution from snow and graupel formation? The model presumable outputs all the terms in the heat budget, so you could find all the fractional contributions to the buoyancy perturbation.

Line 483: "the highly correlated cluster of points in the low CCN-high buoyancy quadrant. Interestingly, all three subsets of glaciated hydrometeors sit in this tail, " – this is exactly what you expect given that the model is formulated so that formation of snow and graupel is a sink of CCN. I'm not sure that this counts as a result.

Line 485-487: "We speculate that the strong correlation with buoyancy is probably not causal, but reflects that CCN depletion and buoyancy boosting are both driven by the same process (glaciation and frozen precipitation formation), and happen at the same location (the ice ring)." – this could be examined by conditionally/spatially subsampling the domain.

Line 516: "The analysis confirms that the mixing-line paradigm as well-known from warm convection also holds for mixed-phase clouds" – doesn't it have to, it is a fundamental property of conserved quantities representing heat and moisture.

Line 517: "suggesting that the mixing source of rising convective updrafts is about 100 m above them" – this is a bit vague. What is meant by 'mixing source', and 'about 100m above them' (ie above convective updrafts) is very vague…what is your definition of the top of the updraft from which that '100m above' starts?

Line 525: “It seems particularly suited to investigate interactions between aerosol, hydrometeors and turbulent dynamics” – this is a bit misleading. ‘It’ refers back to the subject of the previous sentence, which was conserved variable mixing diagrams, but none of the variable here are conserved. Yes scatterplots can be informative here, but they are not the same as mixing diagrams.

Line 531: “One wonders how often the stagnant wind conditions as encountered during ACLOUD RF20 actually occur in the region,…” – I appreciate this is outside the scope of this study, but there are archives of reanalysis data which would answer this question.

Figure B1. The overlapping titles on the panels make them more or less illegible. Move this information into the figure caption.