

Response to the reviewer comments on the manuscript:
“Arctic low-level boundary layer clouds: in-situ measurements and simulations of mono- and bimodal supercooled droplet size distributions at the cloud top layer”
[acp-2014-304]

We thank the reviewers for diligently reading and carefully reviewing our manuscript and providing us with useful comments and suggestions to improve the quality of the manuscript. A list of all reviewer comments and questions (written in italics) as well as our response (written in regular and beginning with >) is given below. Whenever we provide information in which line changes were made we refer to the line numbering of the revised manuscript.

Answer to Reviewer #1

1. *There is an inadequate comparison against the results of previous studies. The authors have ignored the results of two of the most detailed studies of the properties of mixed-phase and supercooled boundary layer clouds collected in the Arctic, namely the Mixed-Phase Arctic Cloud Experiment (M-PACE, Verlinde et al. 2007, Bull. Amer. Meteor. Soc.), and the Indirect and Semi-Direct Aerosol Campaign (ISDAC, McFarquhar et al. 2011, Bull. Amer. Meteor. Soc.), both conducted over the DOE ARM North Slope of Alaska Site. The authors' descriptions of the changes in cloud properties in their vertical profiles mirror many of the features that have been uncovered in the plethora of studies published using data from these two field campaigns. Some of the observations during those campaigns can also assist in explaining some of the phenomena seen in the VERDI data. For example, Fig. 12 in McFarquhar et al. (2011) shows the small-scale structure and inhomogeneities in the cloud microphysical characteristics using an airborne Doppler X-band radar, showing the close proximity of updrafts and downdrafts in arctic boundary-layer clouds. They also used these observations and modeling studies (Korolev and Isaac 2003; Korolev and Field 2008) to show that a quasi-steady state driven by cloud top radiative cooling could be reached under appropriate conditions, consistent with the findings in this study. They also showed that the turbulence and mixing was responsible for the observed cloud distributions and mixing. They showed that vertical mixing had a major role in determining the particle size distributions, and is a good reason why their profiles were not adiabatic. Other analysis of the microphysical data using M-PACE and ISDAC data reached similar conclusions and should be referenced in this study. For example, McFarquhar et al. (2007, JGR) looked at 53 profiles within single-layer stratus clouds in the Arctic during M-PACE. Their analysis (e.g., Figure 11, 12 and 14 along with Table 2) showed many of the same features shown here, with increases in particle size and liquid water content with height, but with a relatively constant liquid droplet number. However, their presented size distributions do not seem to show evidence of a bimodal peak. Using ISDAC data, Jackson et al. (2012, JGR) compared cloud microphysical properties to aerosol concentrations measured above and below clouds. Correlation analysis showed that aerosols were being mixed into the cloud from above. And, Earle et al. (2011, JGR) examined correlations between aerosols measured below clouds and liquid cloud properties, with calculations from an adiabatic parcel model showing*

increases in aerosol concentration can limit activation to larger and more hygroscopic particles, explaining some differences. As all these studies can explain, to various extents, results obtained in the current study, they should be referenced at the appropriate place in the manuscript to show which of the findings in the current manuscript have already been seen by others. Even better, it would be nice if the current authors were able to explain reasons for similarities and differences between their study and the previous ones, so that our understanding of aerosol-cloud interactions in boundary layer clouds was further enhanced.

- It certainly was not our intent to “ignore” the important publications which recently appeared in the literature. (If it came across like this, we apologize.) We tried to provide a short, focused paper with as little as possible text and only few figures on the “simple” questions: “Are our size distributions REALLY bimodal or only a sampling artefact from two different droplet populations with monomodal peaks of different median size like the lower cloud part?” and “Where could such truly bimodal size-distributions come from?” (The underlying observations do not even necessarily have to occur only in Arctic low level clouds.) These are the reasons why we centered our manuscript mainly on the occurrence of bimodal size distributions in the cloud top region and accordingly researched the part of the literature where such SDs are treated. From the observational side there were not so many publications. However, we agree with the reviewer and we thank for bringing the previous studies of more broadened scope to our attention. In the revised manuscript we now describe the M-PACE and the ISDAC campaigns in the introduction and included some of the similarities in the appropriate sections of our manuscript.
2. *I am a bit puzzled as to why such a limited set of data was presented in this study. The authors state that there were 13 research flights conducted during VERDI, but they only seem to present flights from the 14 May and 15 May 2012. I think their results and conclusions would be more convincing if they were to incorporate a more complete analysis of their data, showing relationships between cloud and aerosol concentrations, doing HYSPLIT trajectory analysis to identify where the aerosols were coming from on different days, and showing a more complete data set of how often, when and under what conditions the bimodal distributions actually occurred. This is important because the April-May season is typically a transition season in the Arctic, where both pristine and polluted conditions can occur. The authors might want to review papers of Glenn Shaw and Patricia Quinn that have examined the seasonal distributions of aerosols in the Arctic to further explore this topic. Thus, if aerosols are having an influence on the measurements, it is possible that different results will be obtained on different days. I should also mention that I found the description of the flight profiles unsatisfying. Can you present a 3-d flight track or give some description of the different profiles that were executed on the different days? I do not quite understand how the vertical profiles were obtained. Were the legs shown in Figure 1 flown sequentially at different altitudes, or was the altitude varied throughout the flight leg indicated in red? If it is the later case, this could be a bit of a concern. See Figure 13 in McFarquhar et al. (2011) that gives an example of how much a flight leg can vary during an individual leg. For that matter, I am also curious how the location*

of cloud top was identified? I don't think the lidar would be giving good observations of cloud top when the aircraft was ascending through cloud top? And, with respect to analysis of observations from several different flight days, McFarquhar et al. (2007) combined data from several different flights by plotting cloud parameters as a function of a normalized altitude (zero at cloud base, one at cloud top) so that patterns from multiple flights could be seen on the same plot. Such type of analysis would be helpful for doing a more comprehensive analysis of the complete set of data collected. With regards to Figure 6, a classic example of a single-layer arctic boundary layer cloud is shown. Were such patterns observed on all 13 flights? This is certainly worthy of a paragraph discussion. In the ISDAC and M-PACE field campaigns, the single-layer nature of boundary clouds was a rarity, with clouds typically occurring in multiple layers. Was this ever seen in VERDI?

- The measurements during the VERDI campaign were mainly concerned with the ice and the mixed phase inside the clouds and the plane consequently spent most of the time in the main cloud parts. At the transition zone at the cloud tops during the campaign, the aircraft often crossed this altitude band very fast (also because we were unaware at this time), which gives us a limited data set to use for the presented analysis. Thus we studied the bimodal size distributions in the transition zone of arctic boundary layer clouds in more detail for the few available cases because we suspected the bimodality to be a more common phenomenon. For this paper it was not a goal to give an overview of the whole campaign. Also a special issue in ACP was opened dedicated to the VERDI campaign and an overview paper will be submitted probably soon, as well as different other manuscripts on radiative aspects and the observations of the ice phase.

The altitude profile for the Flight leg in Figure 1 is the same altitude profile which is shown in Figure 3. We added a sentence to make this more obvious. A remarkable change in the cloud height is not the case during this flight leg. The LiDAR backscatter signal in Figure 6a shows the area of the penetrated cloud at 20:41UTC. In this region the cloud top height was nearly constant.

This “classical” example of a boundary layer cloud as shown in Figure 6 was not typical for all 13 flights. Mostly we had more disjoint cloud layers, and in the beginning of the campaign (end of April), sometimes the layers consisted of mixed phase or glaciated clouds. We added a sentence to make this more clear (Page 6, Line 30). For these reasons the data set for the described entrainment scenarios was not very extensive. (In the revised manuscript the data of Figure 6 have been recalculated. Now an additional correction has been applied to remove the contribution of the molecular backscatter.

3. *There are a couple of issues with regards to the instrumentation that could be described better. How good of an overlap is obtained between the CDP and CIP? This could be shown in an Appendix to give more confidence in the quality of the data. There is a potentially big problem with the quality of the number distribution functions estimated from the CIP data in the size range 50-125 microns due to the small and poorly defined depth of field that has been noted by Baumgardner and Korolev (1985). What correction was used for the depth of field. The authors do comment that the modified Korolev-tips were used on the probes, which is important given their note that ice crystals could occur in the clouds. However, Korolev et al. (2013) has shown that such tips do not remove all shattered artifacts. Were the Field et al. (2003, 2006) artifact rejection algorithms, based on interarrival times, also applied? This should be discussed. Also, what algorithm was used to identify the phase of the clouds? The authors mention that needles were sometimes seen, but I do not know whether this was frequent enough that all clouds were mixed-phase, or there were time periods with only liquid. What fraction of the 1 s times were liquid as opposed to mixed-phase? Could there be small non-spherical ice particles that are affecting the forward scatter signal from the CDP probe? For example, McFarquhar et al. (2013, JAOT) found that even in mixed-phase clouds, some of the small quasi-spherical particles could actually be ice particles.*
- During the VERDI-Campaign four cloud particle spectrometers were simultaneously operated side by side. Instead of providing an additional ACP Supplement to our submission we currently draft a second manuscript with detailed instrument inter-comparisons of all instruments deployed during VERDI. Because we focus mainly on the CDP Instrument here, we stayed away from include an instrument intercomparison here. The figure below demonstrates how well the CDP instrument –for example- agrees with the Small Ice Detector (SID) instrument inside a liquid phase cloud. Of course the Korolev-tips do not remove all the shattered artifacts. For this reason we applied the interarrival time method very carefully (i.e. by operator inspection of individual data sets) especially for the bimodal SDs. We added a sentence concerning this point to the manuscript (Page 5, Line 3). We used the NIXE-CAS instrument which uses polarization to identify the phase. This instrument shows that we have only a liquid phase in the size range of the CDP instrument. This information is now included in the manuscript (Page 5, Line 25).

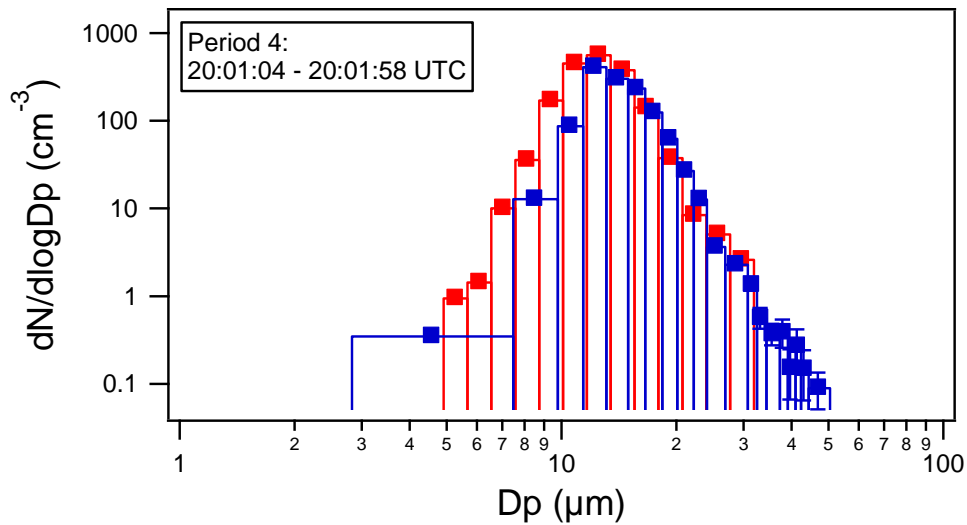


Figure 1: Size distributions from the CDP Instrument (blue) in comparison with the SID Instrument (red). Below 40 μm the error bars (counting statistics) are smaller than the solid symbols.

4. *The modeling study is one of the strengths of this paper, as it gives a better explanation of the observed bimodal size distributions than I have previously seen. However, in order to make this section more convincing I would recommend that the authors perform some sensitivity studies to determine how sensitive their results are to some of the assumed values. Are the similar qualitative conclusions reached as long as reasonable values are chosen for LWC, temperature, relative humidity, turbulence and mixing parameters? Also, some non-standard terminology is used when describing the simulation parameters. The authors state a LWC of 0.25 and TWC of 3.15 g/kg is used. Typically, I think of $TWC = LWC + IWC$, where IWC is ice water content of cloud. I think in this context the authors are thinking of TWC as liquid water plus vapor?? If so, I recommend rewording. I think it also might be appropriate to reference the studies of Korolev and Mazin and Korolev and Field here, as well as some other studies. For example, on p. 14614 the authors state that they neglect microphysics, surface fluxes, shear or rain. But, there have been many studies showing the nature of the circulations and boundary layer clouds in the Arctic, and whether or not they can be maintained, depend critically on some of the parameters chosen to represent these processes. This issue should be discussed in more detail.*
- A sensitivity study as the one proposed by the referee is indeed desirable but this is just too costly (each simulation takes 40000 core hours). However we can still answer, at least partially, many of the referee's concerns. In Mellado (2010) and in de Lozar and Mellado (2013b) they have already shown the mixing in our simulations does not change considerably by varying the resolution or the mixing parameters (since DNS do not use any subgrid mixing model, we understand the Reynolds number as the only mixing parameter). Furthermore, in order to check the robustness of our explanation we have at least performed one second simulation in which we used new data from the

Flight 13, in which also a bimodal SD was observed. Although the properties of both simulated clouds are very different (see captions of the figure below and the values in the manuscript), the simulation also shows the bimodal SD. The results of the simulations are shown in the following pictures. Rather than providing data and figures for the second study in the manuscript we decided to provide only a further sentence (Page 12, Line 17).

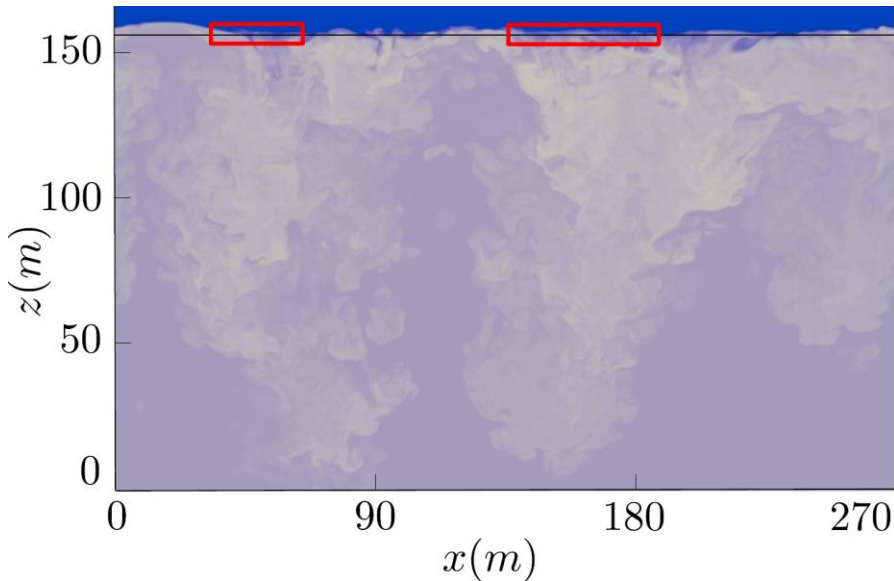


Figure 2: DNS Model for Flight13 with the following parameters. Cloud: $q_l=0.22\text{g/kg}$, $q_t=3.38\text{g/kg}$, $T=270.62\text{K}$, $\Theta_l=269.95\text{K}$. Dry air: $\text{RH}=63.7\%$, $q_t=3.2\text{g/kg}$, $T=277.01\text{K}$, $\Theta_l=276.9\text{K}$. The total width of the Figure is 270m.

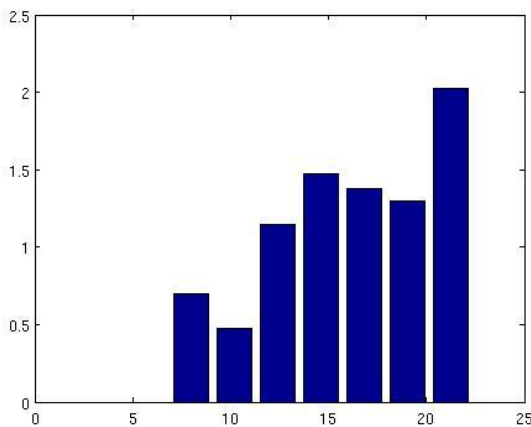


Figure 3: Bimodal SD as a result from the DNS Model. The simulated flight path represents a length of 50m, beginning from the middle point in Figure 2.

We are sorry for the non-standard terminology. In the manuscript we do not use TWC anymore. In the new version we use, q_t , which is the sum of the vapor and liquid water content (Page 11, Line 21).

Our hypothesis is that the fast turbulent mixing is responsible for the bimodal SD. In order to investigate this hypothesis we use a mixing-layer model, which is designed to properly solve the turbulent mixing dynamics at the cloud top. Direct numerical simulations of turbulence are generally expensive because they require to solve many length scales at the same time. In order to find a balance between saving computational time and keeping a realistic setup, we decided to keep only the two strongest driving mechanisms in stratocumuli: radiative and evaporative cooling. Besides, these two mechanisms act directly at the cloud top and have the largest impact on the DSD. This means that we neglect all other processes that might modulate the short-time mixing dynamics, but probably not alter the mixing behavior qualitatively because their contribution to the turbulence and droplet growth is usually less important. The only factor that could change the mixing behavior during a limited time is a strong shear at the cloud top, but unfortunately the shear was not measured in the campaign.

Please notice that the model focus in the short-time dynamics (we simulate 12 minutes of the cloud evolution) and therefore we do not aim to investigate the mechanisms or circulations that are important for maintaining the clouds. We agree with the referee that a more complex model would be necessary for such a study. We want to stress that it would be computationally very expensive to solve at the same time the short-time turbulent mixing dynamics that lead to the bimodal SD, and the boundary layer circulations which are important for the cloud lifetime.

In order to clarify the simplifications introduced in our model, the introductory paragraph (old version Page 14614, Line 11) has been rewritten (Page 10, Line 40).

Specific Comments:

5. *p. 14602, line 29: dominates rather than dominated*
6. *p. 14603: Sentence beginning on line 6 is awkward. Reword*

➤ Thank you, both was changed.

7. *p. 14606, line 9: what probes are used to characterize the size distributions in which size ranges?*

➤ Thank you for this comment. We had forgotten this and added it (Page 4, Line 18).

8. *p. 14606, line 18: Are humidification effects of aerosols dealt with in deriving the size distributions?*

➤ No, the humidification effects were not considered in the derivation of the size distributions. Nevertheless, because an inlet and a long tube system are used inside the aircraft, the detected aerosol is much drier than the aerosol outside the aircraft. Since only particles smaller than 1 μm were considered and the temperature before the OPC was at least 20°C above outside ambient temperatures. Inside the OPC further heating

occurs and we assume the small particles to be dry. Therefore, calculations like the k-Köhler Theory from Petters and Kreidenweis (2007), which was also used in Earle et al (2011) cannot be applied easily.

Moreover, the Sky-OPC integrates a scattering angle from 60° up to 120° . This geometry is less sensitive to the refractive index changes than a narrow-cone forward scattering instrument. We added this information to the manuscript (Page 5, Line 15).

9. *p. 14607, line 11: See McFarquhar et al. (2011), Figure 13, which shows there can be substantial changes of the cloud top height during the course of a flight. Was any change of cloud top height occurring during this flight?*

➤ Looking at the LiDAR backscatter signal in Figure 6a shows that at 20:41 UTC we flew over this cloud region where we took the measurements of the cloud top. In this region, the cloud top height is nearly constant. But further in the south, the cloud top high is 250 m lower. Regarding the reviewer's second comment, this aspect of changing cloud top heights and/or disintegration into several broken cloud layers while flying along the cloud top entrainment zone also limited the available data from sufficiently homogeneous stretches. We included this information to the manuscript (Page 7, Line 41).

10. *p. 14608, line 13: How did the sum of the cloud and aerosol concentration change with height? Might be interesting to compare with Fig. 13 in McFarquhar et al. (2011) who showed this sum was constant both within and below cloud?*

➤ Thank you very much for this comment. We indeed looked at Figure 13 in McFarquhar et al. (2011) and find it very interesting that the sum of the cloud and aerosol concentration was constant. Unfortunately, we had to use a different size range ($0.25\mu\text{m} - 1\mu\text{m}$) for the aerosol optical particle counter because of the inlet design and the fact that the aerosol was measured in the sampling line inside the aircraft. A wing pod instrument like the UHSAS would have been much more appropriate. Therefore, our aerosol concentrations are very low. In Figure 3a you can see a part of the aerosol concentration, which is lower than 10cm^{-3} . The cloud droplet concentrations are one magnitude higher so that we cannot see the effect of the aerosol number concentrations if we sum up both concentrations. Moreover, we very rarely flew vertical profiles below the clouds, often because of ground proximity limitations. However, we added a sentence in the manuscript referencing this result of the McFarquhar et al. (2011) paper and giving a reason why we could not calculate such a "closure" (Page 10, Line 22).

11. p. 14608, line 28: *Do you have any information about the strengths of updrafts in these clouds? Although your stated growth times of 7-37 minutes are well within the lifetimes of such clouds, in 10 minutes with a 1 m/s updraft, the particles would ascend 600 m, which is greater than the depth of your cloud. Thus, there must be some sort of circulation like that in the Korolev and Field model to support this growth as well.*

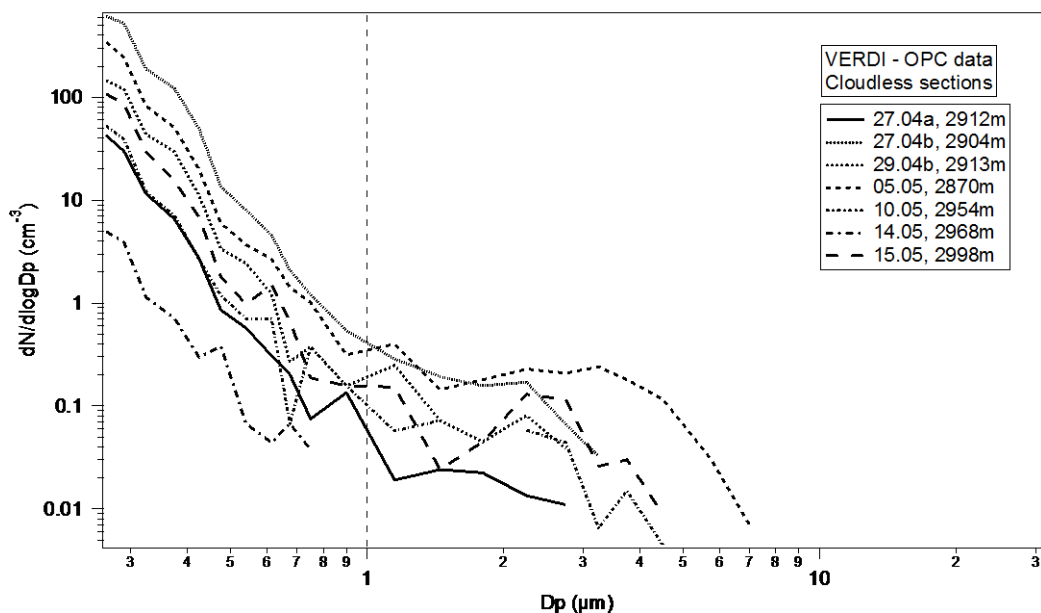
- No, unfortunately we do not have such data and yes, circulation like that in the Korolev and Field model is possible and we added this to the paragraph. Another option is that the vertical updrafts are smaller than 1m/s as next to no turbulence/bumpiness could be felt in the aircraft (Page 7, Line 1).

12. p. 14609, line 19: *Can this argument be made more quantitative? Given the number of particles in each bin needed to have about 10% uncertainty (something Hallett (2003) lists as statistical significance), I would think that not a very large averaging time would be required to get adequate statistics at cloud top. How thick is the relevant cloud level? If you only need 1 s or so to get adequate statistics, I'm having trouble visualizing that the airplane would be ascending or descending so fast that you would not get adequate statistics.*

- The flight strategy during the VERDI Campaign was mainly focused to the inside of the clouds and the cloud top layer was mostly neglected. Also, it seems not sufficient to just “statistically resolve” one or a few bimodal size distributions right after each other. Especially when considering the comparison with the model calculations longer homogeneous cloud top layer “stretches” may be needed in order to have entrainment effects of a variety of eddy sizes included. A second argument is that for the statistical analyses of droplet PAIR occurrences (i.e., one from the smaller mode, one from the larger mode) longer flight stretches may be needed than for just constraining the counting statistics error for one particular size bin. For this reason in Figure 8 sample periods of 4 seconds are used to get adequate statistics. During those 4 seconds periods one would have to fly on constant level. This did not occur during many flights which lead to a further reduction of the periods available for analysis. The bimodal size distribution in Figure 5 occurs in an altitude range between 1033 and 1045m. In this 12m thick layer we saw the bimodal SDs clearly. The reason is, that in this cloud section the pilots decreased/increased very slowly. For the other cases, when we entered a cloud from the top, the occurrence of bimodal SDs happened randomly. Here we do not clearly know whether the layer in which the bimodal SDs occurs was thinner or the entrainment processes were not so much advanced. For these reasons we limited the analyses on carefully selected cases where we think the conditions at the cloud top entrainment zone were homogeneous enough. Since the DNS is so time consuming we rather preferred to identify two cases where the data situation was most clear, homogeneous and unambiguous. A follow up campaign (RACEPAC) was performed in 2014, where we encountered similar bimodal SDs again.

13. p. 14610, line 12: *The spring is the transition season in the Arctic and there can be a lot of day-to-day variation in the aerosol amount. Further, a lot of inhomogeneity in aerosol amount in the vertical and horizontal was noted during ISDAC aircraft missions. Were similar uniform distributions and assumption of arctic background aerosols also valid on other days?*

- The following Figure shows OPC data from different days during the campaign. Since the inlet system was designed mostly for gas phase instrumentation inside the aircraft, there are limitations inherent for (small) particle sampling and we decided to use only the size range between 0.25 μm and 1 μm . At larger sizes the sampling losses become large and size distribution measurements are error prone. However, the figure below shows the range in the aerosol concentration that we measured outside of clouds and aerosol layers. The concentration level was mostly in the range of the arctic background aerosol as measured on 15.05.2012.



14. p. 14610, line 18. *Was the HYSPLIT analysis done on the other days as well to show where the aerosols were coming from on those days?*

- Yes, HYSPLIT analyses were performed as well. Especially between the 05.05.2012 and the 10.05.2012 we saw aerosol layers comparable to the one shown in Figure 6b. During this time the air masses were mostly advected from the Canadian mainland. We added this to the manuscript (Page 7, Line 49).

p. 14610, line 22: See comments above. Can more observations to be shown to confirm that clear arctic atmosphere with low number densities were consistently seen?

- Please see the answer –especially the additional figure- to question 13

15. p. 14611, lines 12-13: Recommend removing this. Given the large variations in the transition season noted in the Arctic in other regions (e.g., Alaska), I think comparisons of measurements 4 years apart are too uncertain to be included in the paper. The paper is sufficiently strong if that data analysis section is expanded without doing such a comparison.

- We see your point and we discussed it a lot before writing the manuscript. For this reason we included that the measurements are “four years apart” which leads to “uncertainties”. Nevertheless, we would like to show by means of the old data that the measured concentrations and size distributions in Figure 7 are well within the range of what can be expected, and not –for example- significantly higher than the adapted measurements, which were categorized as arctic background aerosol. (Also from a measurement perspective it seems well worthwhile to us demonstrating that our data are not “way off”.) The implication basically is that the presented case (Flight 11) is not influenced by higher particle concentrations than the arctic background aerosol and therefore the occurrence of bimodal SD is not connected with very high or unusual aerosol particle concentrations. We added a few sentences to the text cautioning the reader to not mistake the juxtaposition as comparison, which can be interpreted (Page 8, Line 10).

The previous, older measurements also are a composite of data for a regional average. This should include local variabilities with higher and lower concentrations and single (station) measurements thus can appreciably differ from the regional average. However the regional average provides a “variability context” against which single measurements –or measurements later in time- can be imaged in order to see by how much a single measurement can be off the regional(or previous) averages. Had we compared our single measurement against “any” single measurement from the Latham et al (2012) publication, such a comparison indeed would be meaningless.

16. p. 14613, lines 11-15: Might also be worth comparing with the findings of Jackson et al. (2012) who looked at role of entrainment of aerosols from the top of the cloud on ice nucleation, and also comparing with the study of Earle et al. (2011).

- Thank you, we added these references to the paragraph (Page 10, Line 10).

17. p. 14615, line 11: Add “the” between “as” and “next”

- This is changed in the revised manuscript.

18. p. 14614, line 15: *Other studies have found processes like microphysics and surface fluxes are very important in determining the mixing dynamics. See aforementioned papers by Korolev and collaborators.*

- We agree with the referee that these mechanisms will have some effect on the mixing dynamics but these are probably of second order when comparing with the evaporative and radiative cooling (see Reply 4 to Reviewer 1). For example, the sensible heat flux at the surface in the flight RF01 of the DYCOMS-II campaign is around 15 W/m^2 , which is considerably smaller than the 70 W/m^2 estimated for the radiative cooling (Stevens et al. 2005). Unfortunately we do not have measurements for the VERDI campaign, but we would expect a lower sensible heat flux than in the tropics. Related to the microphysics: the most important microphysical impact on mixing is settling (see e.g. Wood 2012). Bretherton et al. (2007) estimated that settling reduces mixing by 3% to 7%, which is probably not strong enough to alter the mixing dynamics considerably.

19. p. 14615: *what does TWC mean? In cloud microphysics this frequently refers to sum of ice water content and liquid water content. Clarify and possibly use alternate terminology for consistency with previous papers.*

- Sorry for the misunderstanding. We have change the terminology it as explained in comment 4 from Reviewer 1.

Answer to Reviewer #2

1. *Page 14601: I'm not sure what is meant by this statement: "Satellite measurements (CALIPSO and ICESat) between 2003 and 2008 showed that cloud cover and optical depth reach a maximum over ice-free waters in the Arctic (Palm et al., 2010)" I figured out later that this statement was meant to be understood as "In the Arctic, there are more clouds and thicker clouds over water than over ice", but at first I wondered if it meant that there were more clouds and thicker clouds in the Arctic than in other locations. I think "in the Arctic" should be moved from the end of the sentence to clarify the important parameter that is being evaluated ("ice-free waters" versus "iced-covered waters")*
 - Thank you very much for this comment. We agree and changed the sentence to "In the Arctic, satellite measurements (CALIPSO and ICESat) between 2003 and 2008 showed that cloud cover and optical depth reach a maximum over ice-free waters (Palm et al., 2010)" (Page 2, Line 12).
2. *Page 14601: "The microphysical characteristics of clouds, e.g., particle phase, size, number concentration, and shape determine the radiative properties which influence the atmospheric radiation budget ... On the other hand climate simulations indicate that the sea salt emissions may increase with receding ice coverage leading to..." It would help to have a transition statement between these two sentences, to clarify what is being compared. Something like "Therefore, reductions in sea-ice leading to greater cloud coverage and cloud optical thickness may have a substantial radiative impact."*
 - In this case we meant to point out that the aerosols also have an effect, but due to uncertainties, it is still unclear what the relative importance of cloud- and aerosol-effects is on the radiation budget. In an effort to make this clearer we modified this sentence (Page 2, Line 19).
3. *Page 14602: "However, in the Arctic, boundary layer clouds mostly warm the below cloud atmosphere." This depends upon the season (See Curry et al, 1993).*
 - That's true, but in this case we considered the net effect and explained it in the following sentence. Nevertheless, we changed this section to: "However, in the Arctic, boundary layer clouds lead to a net warming effect of the below-cloud atmosphere depending on the season. This is due to the generally low sun elevation, the long-lasting polar night, and the high solar surface albedo of land/sea ice and snow (Wendisch et al., 2013; Shupe and Intieri, 2004)." (Page 2, Line 41)
4. *Page 14603: "... showed that a small fraction of ice crystals in the cloud top layer may change the cloud top reflectivity significantly what may also has consequences for the accuracy of cloud remote sensing..." Poor grammar.*
 - Thank you, we corrected the sentence.

5. *Page 14604: At the same time as the NASA ARCTAS project, there was the DOE ISDAC project and the NOAA ARCPAC project. Lance et al (2011) showed bimodal drop distributions at cloud top in mixed-phase stratus during ARCPAC (only observed in less polluted clouds, i.e. clouds with lower CO concentrations). The DOE MPACE project took place in 2004 <http://www.arm.gov/campaigns/nsa2004arcticld>.*
- In Lance et al (2011) there are no bimodal droplet SDs shown that are similar to what we found in our study. For example, the SD in the CDP range from Figure 4 in Lance et al (2011) shows a high number concentration between 10 and 20 μm (white) at the cloud top in the clean case. At the edges of the CDP range, the number concentration is much lower and the size distribution does not exhibit a second mode.
6. *Page 14605: "...and an optical particle counter..." do you mean an aerosol optical particle counter? CDP is an optical particle counter too.*
- Yes, we mean an aerosol optical particle counter. We changed it here and in the following paragraph (Page 4, Line 24).
7. *Page 14607: "While the measured and calculated LWCs are fairly close within the lowest 90m of the cloud the influence of the entrainment processes becomes increasingly evident by the deviations between the two in the upper regions of the cloud"*
It has not been established yet in the paper that this cloud system does not contain appreciable ice, which could of course also deplete water vapor and LWC due to the Wegener Bergeron Findeison process. Also, is there a bulk LWC measurement available? LWC from size distributions is highly dependent on accurate sizing, which was not mentioned in the methods section (only sample area calibrations were reported).
- To identify the phase of the cloud particles we used the polarization method from the NIXE-CAPS instrument. It shows that we have only water droplets in the CDP size range.
 Of course, the LWC is highly dependent on the accurate sizing. We have already performed some instrument intercomparisons and are currently drafting a technical/instrument related publication on this subject. Please see also our reply to Comment 3 from Reviewer 1 for more information.

8. Page 14609: *“As shown in Fig. 2d, deviations between the measurements and the expected adiabatic LWC due to an adiabatic ascent are obvious and indicate that the incloud air does not experience fully adiabatic conditions in our case.” I don’t believe this until you do an uncertainty analysis of LWC derived from the droplet size distribution measurements.*

➤ The error bars provided in Figure 2d are calculated as error propagation and are not statistical variabilities within the data bins. The LWC is calculated from the droplet volume size distribution. Especially in the upper cloud part, where the droplets have a size around 20 μm , the sizing uncertainties are very low. The reason for this is that this size is not close to the upper or lower detection limit of the instrument and the sizes can be well estimated because the measured particles are spherical.

9. Page 14610: *“Considering the clear Arctic atmosphere with generally low number densities in the measured size range this represents a significant difference.” Can you provide a reference for this? This statement seems to conflict with the statement on the following page: “Therefore, the detected aerosol above the cloud on 25 May 2012, might well be considered as dry Arctic background aerosol which seems to be continuously present in this region.”*

➤ We removed the first sentence to avoid any confusion. Nevertheless, we wanted to highlight the discrepancies in the aerosol concentrations between the in-cloud and the out-cloud measurements. We agree, if you compare both sentences then it looks like a conflict. Figure 7, which is discussed in the second sentence, shows aerosol size distributions from the same time section, which is presented in Figure 3a. That means that the high aerosol concentrations in Figure 3a agree with the background aerosol size distributions in Figure 7. (The first reviewer also raised concern about the juxtaposition of the older and our measurements. In response we changed the text as detailed in our reply to Comment 15 above.)

10. Page 14613: *“Further bimodal SDs from Flight 9 and Flight 16 were examined and also integrated in Fig. 9” Please label Figure 9 with the Flight # in addition to the date, to make it easier to compare the manuscript text and the information in the figure.*

➤ Thank you. We agree and added the flight number in addition to the date.

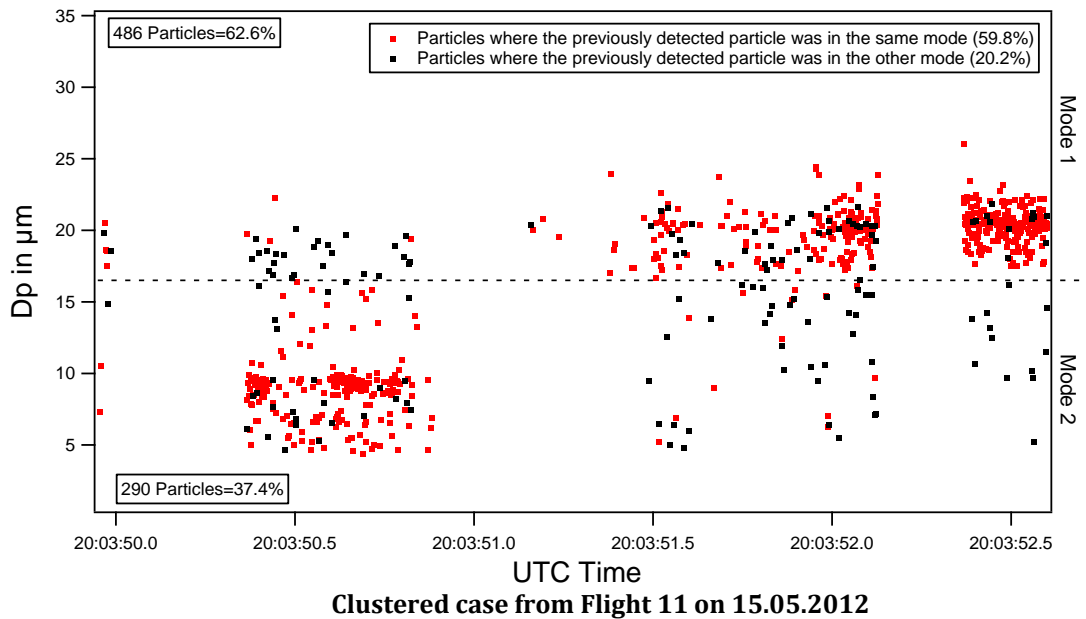
11. Page 14613: *“In summary, four bimodal SDs (marked by circles) seem to result from well mixed air parcels” Actually, all of the data falls below the “Uniform distribution” curve, which is interesting. Does that suggest there is always some degree of clustering (more so for the 3 pluses than for the circled data points)?*

➤ Yes, there is always some degree of clustering. The line represents an ideal mixing behavior which seems unlikely to be encountered in such in-situ measurements. This subtlety is inherent in the abscissa. If two successive particles are both from Mode 1 then the fraction of particles in Mode 2 is 0.0. If the two successive particles are from

Mode 2 then the corresponding fraction is 1.0. With this and equation 2 it follows that a maximum possible value occurs at $p_1=0.5$ being $P_{12+21}=0.5$.

12. Page 14613: “In these clustered cases, the cloud droplets essentially were sampled from two disjoint particle ensembles.” Wouldn’t that be true only if the joint probability in (y-axis in Figure 9) fell at zero? So the clustered particles are not entirely separated, and they aren’t perfectly mixed either. All points seem to fall between those limits.

➤ On the one hand you are right. In theory, the y-axis would fall at zero for fully clustered cases. On the other hand, we did not observe fully clustered cases in this study. When we referred to clustered cases we were describing “mostly” clustered cases. We produced the same kind of plot as the one given in Figure 8 but for the data point (labelled as cross) at [0.37, 0.20] in Figure 9 from Flight 11 on 15.05.2012. This hopefully highlights what we mean with clustered cases: In the beginning, around 20:03:50 UTC, there are a lot more particles in Mode 2. Later, between 20:03:51 and 20:03:52 UTC, the particles are concentrated in Mode 1 and they produce a cluster (red dots) here. To make this more clear in the manuscript we now write “In these mostly clustered cases,...”.



We added this Figure to Figure 8 in the manuscript.

13. Figure 10 Caption: “Isobaric condensation occurs when in mixing of two parcels results in a composition that lays below the condensation line, due to the nonlinearity

(or curvature) in the saturation-vapor-pressure function. The dashed line in the figure illustrates such a mixing process.” Instead of “below the condensation line”, this should read “above the red curve” or “above the saturation vapor pressure curve”. Isobaric condensation occurs when mixing of two parcels results in a composition that lies ON the “condensation line”, only because the line is above the saturation vapor pressure curve. The “condensation line” drawn on Figure 10 is only an example, and it is not mentioned why it is drawn here (it does not represent the actual situation for these observations).

The caption also states: “However, the temperature range covered by our measurements (shown by the continuous line) is so small that the saturation-vapor-pressure is almost linear in this range. As a consequence, condensational mixing is very unlikely.” I don’t understand what “continuous line” means, but I think the authors are referring to the line drawn between points (T_2, e_2) and (T_1, e_1) . All values on this line fall below the saturation vapor pressure curve, since the point (T_1, e_1) falls so far below the curve, which means that mixing between those two air masses will never produce condensation no matter what ratio of the two air masses is mixed. Basically the above cloud air is too dry to ever lead to condensation by isobaric mixing. This leads to my next question: why was an RH of 59% used, when the above cloud RH was reported earlier to be 80%? This likely won’t cause the “mixing line” to be supersaturated at any point, but it would increase the mixed relative humidity compared to using 59% for e_1 , so I wonder why this choice was made.

- Yes, the dashed “condensation line” is only an example and is not connected with the measurements. We started the line now at the point of the parcel (T_2, e_2) to highlight where a mixed parcel later would have to end up on this line to facilitate new nucleation of background aerosol. Also the caption was modified accordingly. Also, you are right, with “continuous line” we meant the drawn line (in black) between the points. We changed it in the caption to make this clearer. We also wrote “above the saturation vapor pressure curve” instead of “below the condensation line”. In regards to your last comment, we could have used a RH of 80%. In the transition zone of the cloud we observed this condition (RH=80%, $T=2^\circ\text{C}$) at places as well which resulted in $e=4.2\text{hPa}$. Even then the mixing line is still under the saturation vapor pressure curve. However, for the modeling analysis we needed measured values for the cloud layer and for the dry layer above. To be sure that we are above the cloud and not in the transition zone we used a RH of 59%. We also used this value in Figure 10 to maintain uniformity.

14. Page 14614: “...depending on the thermodynamically properties of the cloud and adjacent free atmosphere.” Typo.

- Thank you, we corrected it.

15. Page 14615: “We neglect the short-wave radiation due to its small warming contribution in Arctic clouds”. This should be substantiated. It is a mistake to always assume that shortwave radiation has negligible impact on Arctic clouds, because

whether it does or not depends significantly on the season/time of day (i.e. solar zenith angle) and the cloud optical depth. If you want to model the nighttime situation, then it's fine to simply assume that shortwave radiative effects are negligible.

- The referee is completely right. It was an error to state that SW radiation is negligible for Arctic clouds, and this has been deleted from the current version of the manuscript. However, under the assumption that SW heating will not change qualitatively the mixing behaviour at the cloud top we still neglect SW radiation in our simulations. This assumption is valid because the mixing behaviour is mainly dominated by radiative (in the IR) and evaporative cooling. This is justified by the lower warming rates generated by SW with respect to LW cooling. Typical values of SW warming in the tropics are around 2 K/h versus 7 K/h LW cooling (Wood (2012)). Probably this ratio is even larger in the Arctic. Quantitative results will be affected by SW warming, and this is one reason why we do not present any quantitative comparison in the paper, but only a qualitative comparison the SDs and mixing pattern.
A more extended justification of our simplified setup is included in Reply 4 for Referee 1. Besides, the simplifications in our model are now better explained when we introduce the model for the first time (old version page 14614, line 11).
Indeed further and more detailed simulations are necessary to clarify the importance of SW cooling in this context.

16. Page 14615: Again RH1 is set to 59%, though that does not seem to match the observations as reported in Figure 2b, or the text earlier in the manuscript.

- In Figure 2b only the in-cloud and the transition zone of the cloud is shown. The RH of 59% matches the measurements higher above the cloud in the free atmosphere. Please see also the answer to Question 13.

17. Page 14615: *“In comparison to the previously observed holes, the holes in our simulations contain more liquid water and stay closer to the cloud top.”* By *“previously observed holes”* are the authors referring to the work by Gerber et al (2005) in tropical clouds referenced in the preceding sentence? How do the LWCs from the eddy simulations compare with the observed and the adiabatic LWCs?

- Yes, we refer to Gerber et al. (2005). We have added the reference again to the manuscript to avoid misunderstanding.
Please be aware that we use Direct Numerical Simulations and not LES. In the simulations the initial liquid water content was chosen to match the measurements at the cloud top, and therefore it differs by the same amount from the adiabatic value than in the measurements. In the short time covered by the simulations the liquid water does not have enough time to depart appreciably from this value.

18. Page 14616: *You can also discuss the conclusions in terms of total droplet concentrations. If new activation of droplets was occurring at cloud top due to entrainment of fresh aerosols and supersaturation from isobaric mixing, then you'd expect the drop concentration to increase at cloud top. If the 2nd droplet size mode is instead due to drop evaporation, you'd expect the drop concentration to decrease at cloud top. Right? I suppose the two mechanisms could be happening at the same time, too (activating new droplets in some entrained filaments and evaporating droplets in other entrained filaments)... The argument was always posed as either/or.*

- This is true and you can see that by comparing the size distributions in Figure 5b and Figure 5c. In the transition zone the droplet concentration decreases significantly at Mode 1 (around 20 μm). This decrease in Mode 1 is caused by entrainment. In the manuscript, the second Mode (around 10 μm) is described as a result of eddy driven in-mixing of drier air with subsequent evaporation.
Also, we calculated the total droplet number concentration for Mode 2 from the SD in Figure 5c (in a range between 2.8 μm and 13 μm), which is 15.28 cm^{-3} . Then, we compared this value with the total aerosol number concentrations in Figure 3a, which are only around 7.5 cm^{-3} . That means that more droplets develop in the second Mode than there are available aerosols, which is a further hint of evaporation processes. Nevertheless, we cannot exclude the possibility that a few single aerosols were activated, but that seems very unlikely. We added this to the discussion (Page 10, Line 21).