Author’s response to two anonymous reviews for ACP-2020-1096

Review #2

We thank the two referees for taking their time again to provide us with helpful comments which improve the quality of the manuscript. We have thoroughly discussed the addressed issues. Please find below our responses to the raised points (Reviewer comments are given in blue-italic).

Report #1

1. I don't know how much time aerosols of different types lifted from the surface to the lower troposphere can remain airborne before settling but I doubt it is typically on the order of 10-days. The authors did not provide any reference or reasonable justification for this selection nor have they provided sensitivity test results.

Typical periods for back-trajectory analysis in the Arctic are on the order of 5-10 days (e.g. Freud (2017, 10 days), Schmeisser (2018, 7 days), Stock (2008, 5 days)). Also a recently published model analysis showed that the aerosol transport from midlatitude sources into the Arctic can be on the order of 8 days (Zheng, 2021).

2. Air parcels can generally speaking travel enormous distances over a period of 10-days, much longer than the great circle distance between Svalbard and Leipzig. Thus, I think that without any other contextual analysis (e.g., probability of a common geographic source), from a statistical perspective, it is not surprising that the results from Svalbard could be rather similar to those from Leipzig.

It is well accepted that the aerosol in the free troposphere over the Arctic is dominated by aged aerosol pollution mixed with dust and wildfire smoke from all the continents around the Arctic (see Law, 2014, and the recent reviews of Abbatt, 2019 and Willis, 2018). However, we acknowledge the hint that our analysis lacked context. Therefore we added a comparison of a similar aerosol source analysis published in Radenz (2021a) for Krauthausen, Germany, to the manuscript. The distribution of the possible source regions above an arrival height of 3 km show a comparable pattern to what we have found for the analyzed period in the Arctic. One exception is the importance of ‘barren’, which contributes in the Arctic only for trajectories arriving above 6 km altitude.

3. I understand that a reception height of 2 km is widely used in the literature. That does not suggest that it serves as a reasonable assumption in every case, and specifically in this case, where most clouds are much closer to the surface, and hence, an altitude of 2 km is not representative and lacks context.

The reception height is used along the trajectory to identify time periods when the air parcel was in the vicinity of the surface and hence the air mass was possibly influenced by aerosol sources on the ground. The height distribution of the clouds at the trajectory destination should play a minor role in this matter. As an alternative to a fixed reception height, the model-derived local mixing depth can be applied. Yet, models often do not consider the residual layer and hence they tend to underestimate the mixing layer depth (Vivone, 2021), even in high resolution models like WRF (Banks, 2015). This is why the approach of Radenz et al. (2021) focuses on fixed reception height. Nevertheless, we performed the same analysis as presented in the manuscript using the model-derived mixing layer depth as reception height. The respective results are shown in Fig. 1. The largest contribution of ‘snow/ice’ as a possible source region has shifted from trajectories reaching the position of Polarstern at heights around 3 km altitudes to below 2 km. Also, the contribution of 'savanna/shrubland' has become more frequent in all heights. Overall the attribution of possible source regions to the aerosol burden at the analyzed location qualitatively still compares well to what has been published Radenz (2021a) for a campaign performed in Krauthausen, Germany, confirming continental conditions in the free troposphere over RV Polarstern during the time period investigated in the frame of our study.

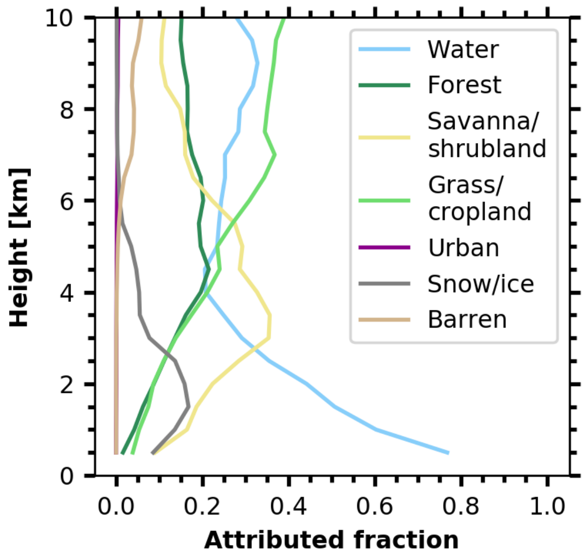

Figure 1. Fraction of residence time of air parcels arriving at heights between 0 and 10 km below the model-derived local mixed layer depth based on a FLEXPART 10 days back-trajectory analysis.

4. The authors did not elaborate on this issue, but HYSPLIT (assuming that this is the model used by the authors) often tends to continue the back-trajectory calculations even after the parcel reaches the surface. Did the authors remove such cases in which the parcel apparently reached its source?

Our analysis was actually based on FLEXPART trajectories. FLEXPART uses a more sophisticated treatment of turbulence (Stohl, 2005) and hence the particles do not 'stick' to the surface for longer periods, as they can do in HYSPLIT. Terminating the particles at surface contact would then imply the disappearance of air parcels and hence a violation of continuity. Additionally, contacts with the surface between the output time steps cannot be diagnosed afterwards.

However, a sensitivity analysis with the HYSPLIT variant of the air mass source estimate revealed no significant difference in the campaign averaged residence times.

5. In general, a single paragraph is definitely insufficient to introduce, describe, and conclude from a new analysis of the data, but on the other hand, I doubt how much further the authors would like to elaborate on this analysis because it may cause the manuscript to lose some of its focus.

We thank the reviewer for pointing out the weaknesses of our analysis and the way it was presented. The respective paragraph was reworded and complemented by missing information addressed in this review (see page 15, line 24 and the following in the diff-version).

Minor comments:

1. Recommenced swapping sec. 4.1 with 4.2. It seems to me like sec. 4.2 better fits the beginning of the Discussion right after the results are presented, while sec. 4.1 could provide a smoother transition to the Summary & Conclusions.

We swapped the respective sections.

2. P. 15 l. 2 - Suggest changing "lowest detection limit" to "lowest radar range gate" because detection limit often insinuates intensity limitation.

Done

3. Caption of figure 8: perception --> reception.

Done

Report #2

After a detailed reading of the authors' comments and of the updated version of the manuscript, I am still concerned about the approach adopted by the authors. I must acknowledge that the authors extended the presented data analysis to support their conclusions. Nevertheless, I regret to note there are still not fully transparent aspects and not obvious choices in the data analysis which does not allow the presented analysis to be complete and to improve the accuracy of results.

My concerns can be summarized in the two following major points.

The authors state that they initially "did not provide quantitative thresholds about how we separated ice and liquid clouds because of the challenges in their estimation. As a consequence their preferred to use manpower to manually analyze the data set and decide where ice is or not." Nevertheless, In their response, the authors also provides an approach fundamentally based on the definition of a threshold on lidar volume depolarization ratio for the detection of ice-containing clouds.The comparison between the "manual" data selection and those based on the automatic data classification, independently on the difficulties intrinsic to the definition of credible thresholds for the automatic algorithm, opens the way to the following thought: considering that with the automatic selection the difference in the percentage of surface-coupled and decoupled clouds is lower compared to the manual analysis (i.e a factor of about 1.5-4 vs a factor of 2-6 on average in between 0°C and -10°C) can we consider this difference as the results of the level of subjectivity of the analysis? Or is this an indication that the analysis is largely affected by the irreducible uncertainties due to the assumptions done in the retrieval of lidar products? Assuming the error bars in Fig.5 are a good estimation of statistical uncertainty according to Seifert et al. (2010), the variability between the "manual" and "automatic" data processing may be representative of a bias which may affect your manual approach. Although the results demonstrates that majority of clouds in the height corresponding to the interval 0°C and -10°C are surface-coupled, the quantification of their fraction must be as accurate as possible and potential systematic effects, such as those due to a manual data analysis, should be discussed in the manuscript.

In lidar research, a manual analysis is by far the most accurate approach as this allows the possibility to check the reliability of the basic, fundamental signal profiles as well as the retrieval products by your experienced eye (one can do that back and forth several times in the data analysis). This is well known and accepted in the lidar community. There are numerous examples available, in which manual analyses reveal deficiencies in automatic retrievals. Just to name some studies of our group, related to evaluation of measurements of the spaceborne lidar CALIOP: Wandinger et al. (2012) and Kanitz et al. (2014). All the attempts to introduce automated data analysis schemes (as for example in the case of EARLINET, D'Amico et al., 2015) were motivated by the fact that more and more lidars deliver continuous observations. But all these products obtained with automated analysis schemes have to be convincingly compared with manually analyzed products, before the (lidar) community trusts them. So, there is no doubt: If the chance is given to apply best knowledge in a manual analysis, there is no better alternative. In our case, the chance was given, as the temporal extent of the dataset allowed us to do this with justifiable efforts, and given the unprecedented measurement capabilities (not available so far for observations in the marine environment of the summer Arctic). In our analysis we thus followed the well-established methodology derived and refined by Ansmann et al. (2009), Seifert et al. (2010, 2011, 2015) and Kanitz et al. (2011).

We once more take the opportunity of this reply letter to highlight to Reviewer #2 and the Editor that our studies are the very first approach of a detailed analysis of the structure of extremely low-level Arctic mixed-phase clouds and their response to surface coupling. This was - to our knowledge - not done before, also because of technical caveats of measurement systems deployed so far in the Arctic (and in the marginal sea ice zone over the open ocean). To our knowledge, there was no polarization-sensitive lidar deployed so far in the marine Arctic, which can provide a liquid/ice separation at heights starting as close as 50 m above ground. Please consider, meanwhile new studies are underway which go into a similar direction with respect to surface coupling effects on ice formation. E.g., Radenz et al. (2021b) found surface-coupling effects on heterogeneous freezing in clouds observed in Southern Chile. They, however, did not have to deal with the very low cloud layers, as they are subject to our Arctic study discussed here. These low clouds provide a great challenge to lidar observations, as they frequently occur at heights within the incomplete laser-beam receiver-field-of-view overlap, even in the case of a lidar with near field capabilities like the PollyXT.

The reliability of an automated data analysis is strongly constrained by an accurate base height of the liquid dominated layer. If this base is set too high, multiple scattering present in the liquid dominated layer might be classified as ice occurrence. A base height located too low might omit depolarization signals from ice crystals below the cloud and hence cause a misclassification of ice containing clouds as liquid clouds. It requires further investigations to carefully consider such effects in automatic retrievals. Especially in the absence of appropriate radar measurements.

We recommend to not elaborate further on the automatic retrieval of our statistics within this manuscript. The approach shown in the reply letter #1 was supposed to briefly evaluate any systematic effects on the approach we present in the manuscript. Even though the automatic retrieval appears of certain value and triggers certain interests, issues resulting from multiple-scattering effects cannot be neglected and need further investigation, which is outside the scope of this study. The application of thresholds in the automated approach combined with the frequently occurring low-level clouds is likely contributing to the difference between the statistics derived using the automated and the manual analysis (see: Figure 4 in reply letter #1). In addition, it is not possible to separate ice-containing from liquid-only clouds below a height of 80m, as the minimum height of the lidar signal is 50m and the minimum ice layer thickness used to identify ice in the depolarization profile was set to 30m. The difference in frequency of occurrence of ice-containing clouds between the manual and the automated approach for clouds with a cloud minimum temperature above -5 °C is likely a consequence. Moreover, this method underestimates pure ice clouds which are lacking a liquid layer and clouds where the ice is located above the liquid layer. This is likely a cause of the differences in ice containing clouds below -10°C.

Thus, evaluating the differences between the automatic and the manual approaches is to our opinion no suitable approach for estimating the level of subjectivity of the analysis. The presented automatic approach comes with its own deficiencies which cause further ambiguities. We hope and wish that our study encourages future studies which are based on enhanced instrumental and thus methodological capabilities.

A longer-lasting dataset would have been of advantage for our study. However, on the one hand this dataset is what we had at hand for our study. It is the first one ever with the PollyXT near-field capabilities plus collocated cloud radar observations. On the other hand, the rather short duration also opened up the possibility of the manual best-knowledge analysis. In our opinion, we provide a thorough discussion of the possibilities for future refinements which can be applied to future studies, hopefully covering longer time periods.

2. The authors removed from the manuscript the analysis on the estimation of INP concentration and added a new analysis to demonstrate the hypothesis that the aerosol source acting as INPs in Arctic are the similar to those of a continental site, i.e. Leipzig, which is used as a comparison term in the data analysis.

First of all it is not clear to me why, using a multi-wavelength Raman lidar, aerosol extinction profiles in clear sky, presented in the updated manuscript version, and consequently, lidar ratio profiles are not retrieved from Raman channels. The capability of a multi-wavelength Raman lidar are fully neglected with consequent increase of the uncertainties in the retrieved products. Assuming a constant value of the lidar ratio may be considered acceptable only for ice cloud, although also I that case the variability of lidar ratio could affect at smaller extent the selected thresholds. The retrieval of the aerosol lidar ratio from Raman lidar measurements from could definitely, coupled with the particle depolarization ratio and air mass back-trajectories, clarify the role of different aerosol types involved the ice clouds formation (several example from literature can be provided, but I am sure the authors knows all of them very well). This could also avoid to include the comparison with Leipzig in support of the authors interpretation of a major role for the continental aerosol in the ice cloud formation in the Arctic and increasing the credibility of the presented analysis, which appears still too speculative.

An in-depth lidar analysis using the full range of the capabilities a Raman lidar offers could have been helpful for a detailed characterization of the lofted aerosol layers. Yet, for our contrasting analysis we need robust (i.e. not influenced by overlap problems) quantities, which are the backscatter coefficient and the depolarization ratio. Both quantities are determined from signal ratios so that incomplete overlap effects cancel out (in case of a well adjusted lidar). We continuously ensured the good performance of the lidar during our manual data analysis, as well as already on-board of RV Polarstern (in presence of the first author of this study, Hannes Griesche) during acquisition of the measurements.

Backscatter coefficient and depolarization ratio alone already allow the identification of dust. Extinction coefficients and lidar ratios are useful to have but it is usually not sufficient to unambiguously determine the aerosol type. In our study, the focus is on the INP efficiency for aerosol in the lowest heights, where only dust and biological particles are of importance (e.g. Abbatt., 2019; Willis, 2018). In these low altitudes, extinction and lidar ratio information are of no advantage. Additionally, since the analyzed cruise was conducted in the Arctic summer in the Arctic ocean we were measuring under continuous daylight conditions. The high background signal prevented us from performing a multi-wavelength Raman analysis of aerosol extinction and lidar ratios.

I want also to add that aerosol backscatter profiles in Fig. 7 must be shown over a longer vertical range to ensure the reader can have a clear idea of the calibration accuracy of lidar profiles.

We have developed and applied sophisticated methods to accurately calibrate lidar signals and signal ratios to obtain quality assured backscatter coefficients. The applied methods, such as Rayleigh fitting, determining the lidar constant to get even get backscatter profiles below cloud decks, when Rayleigh calibration in clear skies in the upper troposphere is not possible, are described in Baars (2016), Hofer (2017), Haarig (2017) and Jimenez (2020).

There is thus no need to show higher-reaching profiles, as it would distract the reader from the conditions in the vicinity of the coupling height.

About the presented investigation of the aerosol sources using FLEXPART model, if it is true that marine fraction decrease with height, above 2 km, it remains the major source at all levels, while the other aerosol types slightly increases with the height, except for "grass cropland" aerosol which has a bit larger increase. I do not see any reference to support the authors' statement of similarity with the aerosol composition typical for the Leipzig site, which is not a marine site. For this part, the analysis looks carried in out in hurry and must be deeper.

Similar to a comparable comment done by Reviewer 1 we like to point out that it is well accepted that the aerosol in the free troposphere over the Arctic is dominated by aged aerosol pollution mixed with dust and wildfire smoke from all the continents around the Arctic (see Law et al., 2014, and the recent reviews of Abbatt, 2019, and Willis, 2018).

Nevertheless we wanted to assess if our analysis is consistent with the literature. The results of the presented aerosol source analysis above 2 km height is comparable to a analysis of a multi-week campaign conducted in Krauthausen, Germany, presented in Radenz (2021a).

At this point, although the statistics presented in the manuscript on the surface-coupled and decoupled ice-containing clouds are interesting, I am not sure if the content of the manuscript is sufficient for publication on ACP. Major revisions are still required and if points above cannot be fulfilled in a short time by the authors, I'd suggest to re-submit the manuscript once data are more consolidated.

We see strong reasoning for considering this manuscript for publication in ACP:

- For the first time, we have analyzed Arctic lidar observations with focus on the relationship between the phase partitioning and surface coupling of the typically low-level Arctic cloud systems over the open ocean. This was technically just not possible before.
- We have used well established, well proven, extensively documented, reproducible methods accompanied by a careful data analysis.
- We show for the first time that heterogeneous ice formation in Arctic mixed-phase clouds depends, besides the minimum cloud temperature, on the liquid layer base height and the surface-coupling state.
- For the first time we showed that in the Arctic summer the thermodynamic linkage between the cloud and the surface increases the frequency of occurrence of surface-coupled ice-containing clouds by a factor of up to 3 compared to surface-decoupled clouds above a cloud minimum temperature of -10°C.
- For the first time, we found that the likelihood of occurrence of an ice-containing cloud is up to 6 times higher at a cloud minimum temperature above -10°C if the cloud is coupled to the surface.

- We acknowledge by means of an extensive discussion the potential and needs for future, extended studies. By doing so, we actively support the advancement of science.

We hope all these arguments are convincing enough to accept the paper.

Abbatt et al., ACP, 2019, https://doi.org/10.5194/acp-19-2527-2019

Ansmann et al., JGR, 2009, https://doi.org/10.1029/2008JD011659

Baars et al.,ACP, 2016, https://doi.org/10.5194/acp-16-5111-2016

Banks, Boundary-Layer Meteorol, 2015, https://doi.org/10.1007/s10546-015-0056-2

D'Amico et al., AMT, 2015, https://doi.org/10.5194/amt-8-4891-2015

Freud et al., ACP, 2017, https://doi.org/10.5194/acp-17-8101-2017

Haarig et al., ACP, 2017, https://doi.org/10.5194/acp-17-10767-2017

Hofer et al., ACP, 2017, https://doi.org/10.5194/acp-17-14559-2017

Jimenez et al., ACP, 2020, https://doi.org/10.5194/acp-20-15247-2020

Kanitz et al., GRL, 2011, https://doi.org/10.1029/2011GL048532

Kanitz et al., AMT, 2014, https://doi.org/10.5194/amt-7-2061-2014

Law et al., BAMS, 2014, https://doi.org/10.1175/BAMS-D-13-00017.1

Radenz et al., ACP, 2021a, https://doi.org/10.5194/acp-21-3015-2021

Radenz et al., ACPD, 2021b, https://doi.org/10.5194/acp-2021-360

Schmeisser et al., ACP, 2018, https://doi.org/10.5194/acp-18-11599-2018

Seifert et al., JGR, 2010, https://doi.org/10.1029/2009JD013222

Seifert et al., JGR, 2011, https://doi.org/10.1029/2011JD015702

Seifert et al., GRL, 2015, https://doi.org/10.1002/2015GL064068

Stock et al., Atmos Environ, 2008, https://doi.org/10.1016/j.atmosenv.2011.06.051

Vivone, ACP, 2021, https://doi.org/10.5194/acp-21-4249-2021

Willis et al., Rev. Geophys, 2018, https://doi.org/10.1029/2018RG000602

Wandinger et al., GRL, 2010, https://doi.org/10.1029/2010GL042815

Zheng et al., 2021, JGR, https://doi.org/10.1029/2020JD033811