## **Response to the comments**

We would like to thank all the reviews for their contributions, as we feel the quality of the manuscript has substantially improved due to their comments. We have addressed all the comments, and we have added some extra changes, that include:

- Pressure altitude has been changed to GPS altitude, which is more precise.
- Colours in the SEM graphs changed to be more intuitive and distinguishable. Bins of SEM analysis simplified, numbers of surface area have been updated (the changes are very minor).
- The back trajectories have been run as ensembles and moved to the Supplementary Information.

### RC1

### General comments

1. A bit more discussion of aerosol-cloud interaction is needed. I.e. what do the results mean for mixed phased clouds and the overall INP "budget" in this region?

We held off extending this paper to a study of aerosol-cloud interactions and purposefully kept it to a study of the INP and aerosol populations in the springtime Arctic. In the case of this particular study, we piggy backed on a campaign and made opportunistic measurements of INP and aerosol, so there was minimal opportunity to link our results to cloud properties. In more recent campaigns conducted in 2022 we have won funding to define aerosol-cloud flights where the INP we measure are of direct relevance to cloud systems we probed in the same flights.

2. The authors suspect low-latitude sources for the mineral dust, is it possible to be more specific in that regard? It is known that mineral dust sources have compositional fingerprints (Scheuvens & Kandler 2014, and references therein, https://doi.org/10.1007/978-94-017-8978-3\_2). Could the authors use the SEM-EDS results to constrain possible source regions? Could the authors also use satellite products to at least qualitatively verify the transport of dust from low to high latitudes over the relevant time period?

SEM: Unfortunately, so few particles were observed due to the low concentrations that the statistics are poor. However, the analysis in Sanchez-Marroquin et al. (2020) shows that the Alaskan dust looks more like Barbados dust than Icelandic dust. This does not mean the dust is from Africa, but it probably isn't from a location with similar mineralogy to Iceland. A more complete study would be interesting, but beyond the scope of what is feasible.

Using satellite products: The dust concentrations we are talking about here are very low and do not show up in typical satellite products. We are not talking about a well-defined plume, but instead the pervasive accumulation mode dust that has a lifetime of many days to months in the atmosphere.

3. Several details on the flights and the sampling strategy are missing in my view, but are needed for a reader to assess the conditions in which the samples were collected. Hence I suggest the addition of the following information to the manuscript (at least in the appendix) • Individual Flight tracks and/or height profiles with the sections when a samples was collected highlighted.

• Indicating if a samples was collected within the PBL/MBL or in the free troposhere; Indicating if sample was collected below/above a cloud layer. This could be done as additional columns in Table A1

• If available, vertical profiles of aerosol number concentration (e.g. derived by integrating the OPC size distribution) and basic meteorological parameters (temperature, humidity)

The referee makes some good suggestions here. We have added the mean temperature, dew point and number concentration (for particles 0.1 to  $3\mu$ m) into Table 1, which is now in the main text. PBL/MBL has also been added to the table. Flight tracks have been added to Fig. 1.

4. For details on the inlet, the authors refer to one of their previous publications, which is fine, but nevertheless the main key should be repeated in this manuscript. That includes cut-off diameter, is the inlet heated or is the aerosol stream dried in some way to avoid the collection of cloud droplets, is the sampling isokinetic, and where on the aircraft is the inlet located?

We have added some pertinent details of the inlet system in the methodology.

5. Analogous to my previous comment, the main key points of the SEM-EDS should be repeated in this manuscript and not just referenced. This includes whether there was any form of sample preparation, such as gold sputter coating, what accelerating voltage was used, how long was the collection time for an EDS spectrum, whether the analysis was manual or automatic, (if the

latter, whether particle detection was based on contrast in the BSE image or something else, and whether possible artifacts were removed before further analysis).

We have added some pertinent details of the SEM technique in a concise manner to help the reader understand what we have done.

L64 Since Tobo et al. 2019 has already been cited above, and because it is consistent with one of your conclusions, it should be explicitly mentioned that dust can also be a carrier for biological INPs We have added an explicit statement in the introduction: '*Biogenic material attached to dust particles could be an important part of these terrestrial INPs (O'Sullivan et al., 2014;O'Sullivan et al., 2015;Tobo et al., 2019).*'

L72-73 This statement about Rinaldi et al. in relation to Wex et al. is correct, but also a bit misleading. Wex et al. attribute the seasonal increase to biological INP and see their effect primarily at high temperatures, with a less pronounced seasonality at lower temperatures. Because of the setup used by Rinaldi et al. they can measure at lower temperatures/higher concentrations. So even if there were a seasonality, Rinaldi et al. would not see it as pronounced in their data, as Wex et al. do. The wording here should take this into account.

The measurement range of Rinaldi and Wex do overlap, with Rinaldi's between 10<sup>-3</sup> and 1 L<sup>-1</sup> and Wex's between 10<sup>-4</sup> and 3x10<sup>-2</sup> L<sup>-1.</sup> However, the range of dates over which measurements were made by Rinaldi is smaller than Wex and we should mention this. The pertinent sentence has been adjusted to: 'However, a recent study did not find strong seasonality of Arctic INPs at Ny-Ålesund, although their measurements were limited to being between April and August 2018'

# L116-117 What is meant by "subset"? Were not all PC filters used for SEM-EDS? If so, what happened to the rest and how was this "subset" selected?

Not all the PC filters were analysed using SEM-EDS for two main reasons. One reason is that the aerosol concentrations were very low, hence samples with low sampling volumes did not have enough particles for the analysis. The second reason is that the technique is very time-consuming and the main author lost access to the SEM-EDS facilities when the COVID-19 pandemic started. Therefore, only a selection of samples could be analysed before the project finished. The criteria to choose samples was getting a sample with a relatively large sampling volume from each flight.

L153-155 I would suggest also including a more recent publication on aircraft INP measurements in the Arctic such as Hartmann et al. (2020), which you cited earlier, in the comparison. They differ in location but are very similar in time (March/April 2018). Borys (1989) may be more comparable in location, but his measurements took place more than 30 years ago, before the extraordinary warming of the Arctic was observed (Arctic Amplification), and some might argue that the differences are related to this.

Our objective here was to first specifically to compare our results to measurements in close locations and timings, and then comparing our measurements to the range of all INP measurements carried out in the Arctic (which includes the Hartmann dataset). We have split figure 3 so now it includes a detailed comparison to the measurements in a close by location and a full comparison to all the existing Arctic measurements.

L155 When making the comparison here, it should also be mentioned which method was used in the respective studies to determine [INP] (freezing assay, CFDC, expansion chamber). *Added to Sec. 3.* 

L172 There are many ways to obtain particle diameters in SEM images, it should be mentioned what diameter was used here

Added to Sect. 4: "The equivalent circular diameter size distributions obtained with the SEM-EDS technique were compared with..."

L173 More details about the optical probes are needed, such as the name of the instrument, the manufacturer, and where it is mounted. Before reading in the caption of Fig. 4 that it was a PCASP-CDP, I wondered if the OPC might have been fed through the same inlet as the filter sampler. PCASP-CDP, on the other hand, indicates that it is a wing-mounted device. In addition, it should be mentioned whether the OPC size distributions shown in Fig. 4 represent the average over the entire period in which the SEM sample was taken or whether it is one point measurement during that sampling period. *These OPCs were all under-wing probes, so not behind an inlet. We have added some details at the beginning of section 4.* 

L175 (+ Fig 4) Is there a reason why only four samples were analyzed with SEM-EDX? An explanation should be given in the text. We addressed this above.

L190-192 PCASP-CDP suggests the combination of aerosol and cloud probe. Doesn't this make it possible to confirm rather than speculate whether the artifacts are actually cloud droplets?

We could not confirm that fully for two reasons. The first is that the shape of the artefacts is rather aphysical, as the concentration decreases more than four orders of magnitude in about 5  $\mu$ m. This does not look like a log normally distributed size distribution of cloud droplets. The second reason is that our sampling strategy deliberately avoided sampling in cloud and rain conditions, so we do not expect a very large amount of droplets detected across the run.

L201 Because of the scale of the y-axis in Fig. 1b, it is difficult to tell when the air masses are below 500m as described in the text. A horizontal line indicating the 500m mark might help, and perhaps a logarithmic axis.

Done

L204-207 (+ Fig 1a) Looking at sea ice concentration (e.g. https://seaice.unibremen.de/databrowser/#day=11&month=2&year=2018&img=%7B"image"%3A"image-

1"%2C"product"%3A"AMSR"%2C"type"%3A"nic"%2C"region"%3A"Arctic3125"%7D ) instead of sea ice extent or satellite imagery in general (e.g., https://go.nasa.gov/3TevAJn), it is clear that the ocean around the north coast of Alaska is not completely covered by ice, but is highly fragmented with numerous open leads. This makes the marine particle source mentioned in L207 more relevant in my eyes and should be discussed accordingly in the text.

We did not intend to rule sea spray INP production out and the referee is correct in stating that there were open leads. Other work from our group indicates open leads at the North Pole were a weak source of INP, but not necessarily negligible, and also this was the North Pole not the Alaskan Arctic

which might be different. We have emphasised this possibility in Sect. 4: "Hence, it is possible that the sea spray particles had been emitted from open leads in the sea ice, as this is thought to be a common source of sea spray aerosol in the region {May, 2016 #873;Kirpes, 2019 #758;Chen, 2022 #869}."

L210-211 Defining mineral as described here is reasonable in many cases. However, considering that there are algae with Si or Ca-based skeletons and the question of the importance of marine or terrestrial INP sources is still debated, I wonder if algal skeletal fragments were misclassified as mineral dust. Can the authors comment on this? Was visual screening done to determine the amount of non-dusty Si and Ca containing particles?

Although the majority of the dust is in the Al-Si rich category (contains Al) or Si-rich categories (contains other elements apart from Si other than Ca), there is a chance that algae skeleton fragments could end up in the Si rich or Ca rich categories, as in any other dust measurement done with this technique (or mass spectroscopy). In this study, a qualitative visual inspection of some of these particles did not reveal any diatom structures, but we cannot categorically rule them out.

L219-222 Indeed, studies such as Huang et al. (2015) show that dust from lower latitudes can reach the Arctic within 5 days. The back trajectory analysis in this study also states that the air masses containing the dust particles circulated for at least 5 days. Meaning that the found mineral dust particles with a size of 1-5  $\mu$ m had to remain suspended in the air for a minimum of 10 days. Can the authors support the claim this typically happens in the atmosphere?

The residence time of dust aerosols in the atmosphere can range from a few days to weeks (Di Biagio., 2021 10.1016/B978-0-12-818234- 5.00033-X, Uno et al., 2009; Huneeus et al., 2011; Ménégoz et al., 2012) and in the Arctic this can be months (Carslaw 2022). Additionally, not all the air masses had been circulating around the Arctic; a couple of them came from the south, and some mixing can occur. We have added a short statement to this effect in the SEM discussion section 4:

"Almost all the mineral dust particles found in this study had sizes below 5  $\mu$ m and it is known that the lifetime of accumulation mode aerosol typically has a liftime of many days or weeks so can be transported to Alaska from distant sources {Huneeus, 2011 #637; Ménégoz, 2012 #835}. Once in the Arctic, accumulation mode aerosol has a lifetime extending to months during winter and spring, when removal processes are weak (Carslaw 2022)"

#### L234 Fig. 5 (instead of Fig. 3)?

Thanks for spotting the typo.

L235-236 Is it not possible to use the SEM-EDS analysis to derive a K-feldspar amount that is more representative of the samples of this study? You already have the "AI-Si rich" category, which should contain mainly the feldspars and their weathering products (illite, kaolinite, etc.). If this category is filtered for appropriate amounts of K, an estimate for the K-feldspar content could be derived. I think studies by K. Kandler and co-authors have covered the categorization of mineral dust with SEM-EDS in detail.

Yes, we tried the approach by Kandler et al, however, the number of dust particles analysed in this study is relatively low due to the low aerosol concentrations of the Arctic (always under two thousand per filter), so the statistics of K-feldspar particles were even lower, hence we do not report them.

L238-242 It should be noted that the parameterization presented by McClusky et al. (2018) is meant to represent pristine SSA and they intentionally excluded events with elevated INP concentrations. They also state t during a period of elevated marine organic aerosol from offshore biological activity, [INP] at -15°C was 7.7x10-3 L-1. A value that is above the mineral dust and below the fertile soil in Fig. 5. This, of course, does not refute the author's conclusion that ice activity is related to mineral dust, but the possibility of biological INP of marine origin should be discussed. McClusky et al. (2018) also present a different parameterization based on TOC. Perhaps the authors can derive an estimate for TOC based on the "carbonaceous" category of the SEM-EDS data and add these predicted INP concentrations to Fig. 5 as well.

This is a good point regarding the McCluskey data. We have adjusted the discussion to take this into account and now state: 'It is possible that the sea spray in this location was more active than defined by McCluskey et al. (2018), however, the INP concentrations calculated based on the presence of dusts better explains the observed INP concentrations' and further down we now state 'hence this suggests that the samples from Alaska contained some biological ice nucleating material (either from marine or terrestrial sources).

Unfortunately, the carbonaceous category in the SEM analysis only indicates that those particles did not contain any other major element apart from the elements in the filter (carbon and oxygen), but we cannot quantify anything else with them. Additionally, some of the TOC would come in the sea spray particles, which are classified as Na rich. The fact that the filters are made of carbon and oxygen means we detect those elements everywhere in the analysis and we cannot quantify much of these elements.

L252-254 Following the author's train of thought, the sampled aerosol was transported for about 10 days, and as noted here, a decrease in ice activity would have been expected, yet Fig. 5 shows a very good match with fertile soils collected directly from the ground and brought to the laboratory quickly. This would suggest that the aerosol was not transported far, but may have more local sources. Can the authors comment on that? In this context, could the use of monthly average snowpack (Fig. 1a) or the sensitivity of the satellite product used for the ERA5 reanalysis be obscuring exposed soils?

In our opinion it is not obvious that there should be a decrease in activity on transport. While it has been shown in the lab that mixing with acids, for example, can deactivate some INP types it is not clear that this process occurs in the atmosphere to a significant extent.

L255-256 In the INP community, heat and  $H_2O_2$  treatments are commonly used to gain additional insight into the presence of proteinaceous and organic INPs. Could the authors perform such treatments as the results could further support the claim presented here?

Unfortunately, the droplet-on-filter INP measurement technique only allows us to perform the analysis once, hence, heat tests or  $H_2O_2$  treatments cannot be done in our samples. Also, it is not clear how the filter would respond to these treatments (although we have considered it and recognise the value). This is a disadvantage of the droplet-on-filter technique, but the greater sensitivity over the wash-off methods means we can make INP measurements in the 20 minutes or so that we have per filter leg. Hence, this compromise is well worth it.

Table A1, second column 2018 instead of 2017 *Corrected*.

Table A1 Are the presented altitude values averages over the sampling time? If yes, that should be mentioned in the text and caption.

Added to the caption of the table : "The given altitude values correspond to the average of each run."

Fig. 1 a) A reader is not able to recognize where a trajectory starts/ends. If the daily marker would shrink with time, the reader might be able to identify the direction of the trajectory and hence start and end. The authors also chose to show only one return trajectory per sample. For short sample times, a single back trajectory could be representative, but for longer sample times, the aircraft could travel long distances and enter air masses with very different histories. Can the authors comment on the representativeness of back trajectories they show. The authors might also consider adding a figure to the appendix showing high frequency back trajectories for each sample.

The fight tracks have been added to figure 1, while the back trajectory analysis has been sent to the supplementary material. The analysis has been separated per days to ensure it is possible to see each trajectory better and the single trajectories have been substituted by ensembles.

We have not performed a more detailed back trajectory analysis. We considered investing more time in a Flexpart study, but since the indication is that the sources are not local, we suspect there is little to be gained from a more detailed back trajectory approach. Also, our sampling periods tended to be focused on specific air masses, rather than sampling across air masses or fronts, hence we consider them to be qualitatively representative. In addition, detailed back trajectories would be required if we saw some filters with much greater INP concentrations than others, but this was not the case. Hence, we just didn't observe sufficient natural variability for source apportionment modelling.