2nd Referee report on "Ice-nucleating particle concentration measurements from Ny-Ålesund during the Arctic Spring-Summer in 2018" by Rinaldi et al.

The authors have addressed most of the previous comments and improved the manuscript. However, I would like to suggest some further revisions to improve the manuscript before it can continue the review process.

Main comment

The number of samples this study is based on is small and might not allow to characterize the population of INP. It needs to be highlighted in the abstract and conclusion that results are preliminary because of the small sample size. Findings in the abstract and conclusion should be limited to the strong signals that are expected to be reproduceable in future investigations. For less clear results it should be stressed that more observations are needed.
The limitations of the applied analysis due to the size of the available dataset should be stated clearly in each section. This includes stating the number of measurements used and the assumption on the structure of the data, e.g. normal distribution of nINP (not typical) to compare spring to summer concentrations or contribution of coarse and fine particle fraction to nINP.
Because correlation analysis is prominent throughout the paper, I suggest to add a subsection detailing the statistical analysis, including how significance in dependence of number of samples is determined (t-test), to the Method Section. This would help to understand, e.g. why correlation coefficients are high and why 0.5 is significantly non-zero for DFPC but 0.8 can be not significant for WT-CRAFT (Tabs.3 and 4).

Specific comments

Line 27f, 101f, 609ff.: The difference in nINP due to two different ice nucleation modes does not emerge from the data. To make this conclusion, filter samples from one sampling setup should be analysed with both DFPC and WT-CRAFT. More plausible are differences in the samples used by the two methods.

Line 29: name the "several important indications"

Line 29f: This is inconsistent with what is reported on line 419ff

Line 31: specify "subset of our data"

Line 33: Explain how higher AF can be interpreted as larger freezing efficiency of large particles.

Line 34: inconsistent to line 417 where it is stated that no clear relation emerges between nINP and meteorology.

Line 48: what other mechanisms beside Bergeron-Findeisen play an important part?

Line 56, 57: Clarify the difference between water vapour deposition and water vapour condensation.

Line 60f.: If condensation and immersion are the same process how can the difference in nINP between DFPC and WT-CRAFT be explained by it? To make this case, provide evidence that the two mechanisms can be different.

Line 96: Why are Santl-Temkiv et al. and Wex et al. cited for the interpretation of data in Tobo et al. and not the Tobo et al. paper itself?

line 104f: Explain why INP measurements are the key to verify that immersion freezing is the most relevant ice nucleation mechanism in Arctic mixed-phase clouds and how such measurements can be used to do that.

Line 117ff.: state the number of samples collected in spring, summer, PM1, PM10.

Line 128ff.: on line 131 it says the 5.4lpm are a subset of the airflow through the TSP inlet. Specify the total flow through the inlet and describe more clearly how the 5.4lpm are extracted from the higher total flow. Was there a pressure drop in the flow from which the subset flow was taken? Was this considered to determine the sample flow? Was there an online measurement of the flow through the membrane filter (how are flow rates in Tab.S1 measured) and why did the flow vary from 4.8-5.6lpm between samples?

Line 165: what is meant by "reasonably matches"? The 95% confidence interval is inherent in the cited formula.

Line 178: point to sec. 2.2.3. for details on how nINP are estimated

Line 179: mention already here that half of each filter was used and explain how the used water volume was calculated. Volumes given in Tab.S1, row 2-6 are off by -0.1 to 0.2ml from calculated values.

Line 202ff: I couldn't find the size-range in Kanji et al., 2017. Pruppacher & Klett, 2010 section 9.2.3.2 suggest 0.1um as a lower size limit, which I would recommend. The choice needs to be motivated better. How does the lower size limit effect the results in sec. 3.3?

Line 215ff: Clarify what size-range was used to calculate AF. From line 203 it would seem that only APS data (0.5-10um) was used.

Line 246: provide a reference for the marine boundary layer height in the Arctic.

Line 270: It could be already mentioned here that nINP at -15°C, PM1 are used for this exercise.

Line 280: to be consistent throughout the paper, consider using nINP instead of Ct for the INP concentration.

Line 295: It could be explained here how the CHL correlation analysis and CWT were overlayed (Fig.8b).

Line 303: Can you provide an interpretation of the difference in slope? It could suggest that there is a specific type of high temperature INP, that was only detected in the DFPC measurement. Could storing the filters at different conditions cause such an effect?

Line 304ff: Consider referring to Tab.2 instead of listing the concentrations.

Line 307: give number of samples instead of "<50%".

Line 311: give T-range for the nINP, see comment on Tab.1

Line 312ff: Can the higher nINP from condensation mode measurements compared to immersion mode be confirmed from the studies listed in Tab.1?

Line 320: give number of samples that span the 3-200m-3 nINP range.

Line 323ff: The difference in nINP depending on the sensitivity of methods indicates that the INP population is highly variable and the variation is underestimated at the detection limits of methods.

Line 325: Clarify for what the nINP range is substantial. Below what concentration would it be negligible?

Line 328: give nINP at -22°C to compare to the range at -22°C on line 325.

Line 333ff: add nINP at -22°C measured at the cited locations

Line 344: As I understand a normal distribution is assumed to obtain the difference at $p<0.005$. I doubt that the dataset (<20 measurements per season) is large enough to obtain a valid distribution to perform statistics.

Line 346: From Tab.2 it can be seen that he median nINP,PM10 is the same in summer and spring. The nINP,PM1 is lower in summer thus contradicting the conclusion that coarse INP from exposed surface cause the difference. The higher nINP,PM1 in spring could indicate INP from the arctic haze.

Line 347: on line 344 the contribution of coarse INP is described as significant to a high level of certainty. How were the "substantial uncertainties" considered to estimate the significance of an increase in coarse INP contribution?

Line 349: On the PM1 filters not only particles larger than 0.5um are sampled, as was assumed for the fractions shown in Fig. S2. I suggest using the full size-range below 1um for the fine particle concentration to calculate the change in the particle population.

Line 350: The coarse INP fraction doesn't dominate. Looking at Tab.2, the fine fraction makes up 80% of nINP in spring and 50% in summer.

Line 352: As mentioned on line 350, Mason et al. also reported a spring to summer dataset, making the results not "unique".

Line 360: Repeat here how coarse and fine AF are calculated, i.e. particle size-range used and nINP coarse as difference of PM10-PM1 nINP. Also mention it in the caption of Fig.3. What particle size-range was used for the WT-CRAFT (sampling through a TSP inlet) AF?

Line 360, 414, 544, 573, 579, 585, 587: Ice nucleation efficiency/ability/activity/capability was not measured in this study, but concentrations of INP. Ice nucleation efficiency requires knowing the concentration of the ice active species. To be more consistent replace them throughout the paper with nINP or AF where appropriate.

Line 362: the difference would be much larger if fine particles would not be limited to particles larger than 0.5um.

Line 363ff: instead of listing AF, point to Fig.3. Disentangle if the AF is governed by changes in particle concentration or nINP.

Line 373: mention what particle size-ranges Si et al. used.

Sec.3.4: This section needs to be structured better. It is currently unclear what the important results are.

Line 387, 391ff: factors of 1.5 or 1.6 are very small differences to draw conclusions. Additionally, the number of samples could be too small to determine the underlying nINP distributions on which the comparisons in this section are based on.

Line 395: State the number of June samples and explain how it was determined that they represent a significant peak.

Line 409ff: The effect of particle concentration and nINP need to be disentangled. The absence of a change in nINP with season points to the particle concentration causing the change in the AF.

Line 430: the 50-120nm size range mentioned here provides additional evidence against the chosen lower size limit of 500nm for this study.

Line 462, Fig.6: The trajectories with land contact are not visible in Fig.6. Removing the not used grey trajectories could help.

Line 468f: specify that the -15°C, PM1 dataset was used. It is unclear why the coarse and fine data is an advantage for DFPC. There is no comparison between coarse and fine nINP shown here, except Figs. S7 and S8. Do these maps confirm Mansour et al. and McCluskey et al.?

Line 472f: Has the analysis been tried for -18°C and -22°C? It is mentioned on line 494ff that experiments showed INP active at -22°C are generated.

Line 483: Provide an interpretation of significant, negative correlated areas. Are they caused by elevated CHL without the expected nINP response?

Line 487: see comment on Fig.7.

Line 534ff: There could be an increase in biological, high T INP from growing biota on CFPC filter during storage at room T.

Line 538: It could be added that similar air volumes were collected on the filter for CFPC and WT-CRAFT.

Line 549: It could be expected that soluble compounds suppress ice nucleation. Why would the opposite be observed for these samples?

Line 552: A detailed comparison is not needed. To claim sensitivity of arctic nINP on the ice nucleation mode, exemplary DFPC filter need to be analysed with the WT-CRAFT method and vice versa. Otherwise it is speculative.

Line 566: It seems plausible that due to small number of samples, the INP population was not well characterized in aforementioned studies, leading to misinterpretation.

Line 578: state which results

Line 591: where is shown that land sources have this potential? PM10 nINP do not change much even when more land and open sea are along trajectories and PM1 nINP decrease.

Line 591, 592: at what T have mineral dusts a higher activity and at what T are marine INP less active?

Line 598: specify the results, e.g. “location of hot spots for marine INPs”

Line 610f: The presented analysis does not support this conclusion. A more plausible reason could be differences in the DFPC and WT-CRAFT samples. As mentioned above, to arrive at this conclusion the same samples must be analysed with both methods. Replace “undeniable” by “potential”.

Line 616: The dataset is not unique, other studies have collected Arctic INP covering different seasons e.g. Manson et al., Schrod et al.

Line 624: how did the back-trajectories show this separation? They were done for summer only.

Line 629: repetition of line 625

Tab.1: Consider making 3 columns showing the nINP at -15°C, -18°C, -22°C, relevant for this study, instead of the current last two columns. A minus sign is missing in Bigg, 1996 T range.

Tab.2: Subtracting 1-3 standard deviations from the average values gives negative concentrations. nINP can not be negative, the real variation is clearly asymmetric. Consider reporting the range and average instead and refer to the table instead of listing these values in the text.

Tab.3, 4: Scatterplots for all 39 significant correlations reported in these tables could be shown in the supplement and investigated for outliers to exclude false positives.

Fig.3: mention the particle size range used to calculate AF in a), b) c)

Fig.4: indicate what is shown in a), b), c) in the caption. Specify that uncertainties can be found in sec.2.2.1 (CFPC) and 2.2.2 (WT-CRAFT)

Fig.6: consider removing grey trajectories.

Fig.7: specify which region belongs to which time lag. Consider overlaying the trajectories as thin lines to indicate upwind locations and clarify the choice of regions.

Fig. 8 b): consider showing the trajectories corresponding to the 14 INP samples as thin lines

Tab. S1: Column headers "Total Flow (optimized for 50% of filter)" and "Suspension water volume (First frozen drop=0.001 INP L-1)" are unclear. Give a description in the legend and rename the columns, e.g. "air volume" and "suspension water volume". Check the calculation of the suspension volumes.

Fig.S3: consider using the same y-axis scale for subplots showing the same temperature. Even if there are only 3 spring datapoints in Wex et al. at this T, add -22°C for DFPC and WT-CRAFT as the seasonal change is discussed in the main text. Give a description of the subplots a)-f) in the figure legend.

Fig. S4: Describe y-axis in the figure legend. Is PM1 or PM10 shown for DFPC? Mark the increase in AF due to the decrease in particle concentration from spring to summer (= particle conc. summer/ particle conc. spring) as horizontal line. Spring to summer AF increase close to this line indicate no change in nINP.

Technical corrections

Provide more links/references between sections and to tables and figures to navigate the paper.

Line 27: remove "trustful"

Line 79: "suggested *them* to be…"

Line 200: remove "now"

Line 210: replace "side-by-side" with "in parallel".

Line 413: replace "point sin" with "points in"

Tab. 4b: remove "I" before nINP in the caption.

Fig. S1: change (a) -18°C to (b) -18°C in the second figure.