**General comments**

Ambient measurements from Spitzbergen during spring and summer 2018 are reported and analysed. The dataset is a welcome addition to the growing body of ice nucleation concentration measurements from the Arctic region. For the analysis the authors correlate the measured INP concentrations to bulk particle properties, season, meteorology and airmass trajectory. Unfortunately, the analysis is not well motivated by hypotheses, and the results not presented clearly. Additionally, I suspect that the set of measurements is too small to perform a robust analysis and the found correlations might be random. By ignoring this, the authors got mislead to overinterpretations and speculative conclusions. The interpretation of data (e.g. concerning INP size and land vs. marine contributions) agrees with previous studies and no new insights are obtained. The one interesting finding is that there was no seasonal variation observed in 2018. The authors need to be more specific in their descriptions, quantitative within reason for the interpretation and visualize their findings clearer to turn this manuscript into a valuable contribution to the field of ambient ice nucleation measurements.

**Specific comments**

Major changes are suggested for each section. Minor corrections or clarification requests at specific line numbers are listed below.

**Structure**

The manuscript could be structured better. Adding a Discussion section instead of including the discussion with the results would help the organization of the paper. Currently the Result section has mixed-in discussion, interpretation, literature comparison and some method description. Isolating the Results, visualizing and explaining them more specifically would be helpful to judge the interpretation.

**Citations**

A high number of citations are given, but it is often unspecific what information can be found in which citation or why several citations are listed. Better integration in the text would be helpful instead listing multiple citations at the end of a sentence.

**Title**

The title does not fit the manuscript. There was never doubt that multiple sources contribute INP at different temperatures. Something along the line of “Concentration of ice nucleating particles measured at Ny-Ålesund during 2018” would be more accurate.

**Abstract**

The abstract uses wordy, euphemistic language. In the interest of clarity, the tone should be revised to be strait forward. At the moment the abstract consists of too many unsupported statements and vague conclusions that are incomprehensive before reading the manuscript in detail.

Line 15 ff: INP sources, INP concentrations and ice nucleation properties should be distinguished clearly. They are not the same.
Line 15: INP sources are unknown not “inadequately understood”
Line 18: specify what properties were characterized
Line 22: why are the temperatures (-15°C, -18°C, -22°C) probed with the DFPC not evenly spaced?
Line 24-25: The dependence on ice nucleation mode is not investigated in the manuscript and a sampling issues is a more reasonable cause for the discrepancy.
Line 26: specify how the increase in coarse INP was observed.
Line 27: Explain why increase in coarse INP fraction suggests local sources.
Line 28: Speculative. INP active at -15°C are not exclusively biological particles. The source of the particles active at -15°C in this study are unknown.
Line 30: specify “distinct behaviours of particles”
Line 31: How was the inter-annual variability of local INP sources previously considered? Specify the evidence for an inter-annual variability based on the current dataset and all available data from Spitzbergen. Consider that the 2012 data set is only 12 measurements and not a strong dataset to compare to.
Line 33: It seems trivial that the INP population can be contributed by terrestrial and marine sources on an island.
Line 35: specify nucleation ability. Higher particle load can explain higher INP concentrations without increased ice nucleation activity.
Line 37-40: requires reading the paper to understand this description.

**Introduction**

The introduction is not tailored enough to the subject of this study. To make the introduction more effective at explaining and putting into perspective what follows, I suggest to focus on: ice nucleation mechanisms (condensation and immersion mode), INP-cloud interaction in the Arctic without going into detail on radiative effects, previous INP measurements in the Arctic and what has been learned about potential sources, tends, dependencies of nINP in the Arctic. A lot of literature is currently discussed in the Result section. Better to include a concise discussion of relevant literature to the introduction to develop the hypotheses which are then addressed in this study. Currently literature review is used inefficiently in the Result section to point to similar conclusions found elsewhere in literature.

Line 46: provide reference and explain how aerosol affect cloud properties
Line49-57: specify how the numerous local processes and feedbacks interact to affect structure, phase and persistence of clouds in the Arctic. Beyond what is generally true for INP-cloud interactions, explain why Arctic clouds are sensitive to INP concentrations. Provide a reference for the uncertainty associated with nINP, eg. DeBoer et al., 2018.
Line 59: the references are not “recently”
Line 61-62: specify “most” ice nucleation processes. It would be good to introduce ice nucleation mechanism in more detail.
Line 63-64: It is not generally true that biogenic INP nucleate ice at temperatures above -15°C. It is a bit of a stretch from ice nucleation properties to rain. Provide a more coherent explanation and provide the link to INP in the Arctic region.
Line 67-86: Specify that the literature review is separated into condensation and immersion mode measurements as well as separated into short and year around observations.
Line 67: specify “short periods of time”
Line 72: add some information how Hartmann et al. confirmed this.
Line 74: How did Bigg et al. identify the Ocean was the main source? There was a third Arctic cruise in 2001 and a more recent expedition in 2017. They are reported in Welti et al., 2020.
Line 76-86: Mention at what temperatures Mason et al., 2016, Si et al., 2018, Creamean
et al., 2018, Irish et al., 2019 reported data. This section seems to contradict line 61-64 where it is argued that INP active at T>-15°C are biological and not mineral dust.

Line 88: Quantify the increase. Contrast to the fact that often ambient nINP measurements scatter within 1-2 order of magnitude within less than a day, highlighting that caution should be used when interpreting variations smaller than one order of magnitude in such dataset.

Line 90: A time-series showing measured nINP at -15°C as function of DOY from all the listed measurements would be a helpful addition to illustrate the discussion and to show if the Arctic region as a whole experiences seasonal variations or if these are local phenomena.

Line 94: As your literature review exemplifies, there is no general “gap” of INP measurements in the Arctic.

**Methods**

The method of how INP concentrations are determined with the WT-CRAFT starting from the air volume sampled through filter to counting the number of frozen aliquots, should be explained in more detail, focusing on how nINP can be derived step by step from sample volume, water volume, droplet volume. How the preparation is done practically is of secondary interest.

In section 2.3 it should be clarified (by a short explanation at the beginning of each subsection) for what purpose the measurement or analysis is performed or used in the context of this paper. Here it would be helpful to already know from the introduction what the aim of the analysis is or what hypotheses are going to be tested with these data. Give context in the introduction section: why are you investigating ground type, trajectories, chlorophyll,...

Line 99: Refer to the location as Gruebadet station throughout the paper and introduce the abbreviation (GVB) here.

Line 100: Point to fig. S1 showing the location on a map.

Line 102: Can Longyearbyen in the SE of the GVB station contribute to the aerosol population?

Line 110: Is filter overloading an issue in the clean Arctic air? The WT-CRAFT filters were sampled longer, with higher flow and on filter with smaller pore size. Where these filters potentially overloaded?

Line 115: Pumping 150lpm through filter with 0.2um pore size creates a huge pressure drop. Can you comment on how sampling was possible without fracturing the filter? How was the flow monitored? In this setup the filter probably acted as flow limiter rather than the critical orifice before the pump. An overestimation of the sample flow would explain the offset between WT-CRAFT and DFPC.

Line 116: specify pump model

Line 118: Is filter overloading (line 110) an issue for 4-day samples? The volume sampled is more than 100-times larger for WT-CRAFT than for the DFPC filters.

Line 127: Specify how uncertainties in T and Sw convert into uncertainties in nINP.

Line 128f: Has a systematic difference between condensation and immersion mode ice nucleation been observed in these inter-comparisons?

Line 132: The large sample volume of over 800m³ would allow to detect approximately 100-times lower INP concentrations than 1 m⁻³. Why was the analysis not performed in the full range?

Line 133: State how the two CRAFT systems are different. All I could find in Hiranuma et al., 2019 was that they used different sizes of droplets. This is not an instrumental difference.

Line 135: Explain how the uncertainty in ice nucleation efficiency is derived.

Line 136: repetition, delete
Line 139: provide camera model specifics
Line 140: define how INP concentrations are derived.
Line 141f: specify water volumes, for 1 INP per m$^3$ that would be 90mL for 4 day samples and 180mL for 8 day samples.
Line 143f: specify water volume used for soaking
Line 144: specify how mechanical vibration was applied. By sonication?
Line 145: How were droplets prepared?
Line 146: The method by how much the sample was diluted is not explained clearly. Specify the dilution water volume and how dilution was considered for the derivation of INP concentrations.
Line 149: How was the stitching performed? At what temperature were the spectra stitched? In Fig.1 a a jump in max concentration appears at -23°C. Is this the range of the diluted measurements? To show the “absence of failure” of this technique, it would be helpful to show the individual measurements in Fig.1 and not just the range.
Line 153f: Derivation of nINP must be defined clearer. Dividing the number of INP by the total sample volume is incorrect. The concentration is calculated from the filtered air volume, dilution water volume, droplet volume and number.
Line 154: As you state on line 379, only a small, T-dependent fraction of ambient aerosol are INP. Therefore, dividing the number of INP by the bulk particle surface has no physical meaning for a heterogeneous aerosol population. I recommend changing this approach to deriving only the fraction of particles that are ice active, by dividing the INP concentration by the particle concentration.
Line 161: specify APS measurement range
Line 163: what are the references pointing at? How were the size distributions averaged for the sampling interval of the filters?
Line 164: What substance is assumed for a density of 1.95 gcm$^{-3}$? Mineral dust and sea salt have higher densities.
Line 173: Where was this analysis performed? While handling during the analysis is relevant, handling before and after sampling, storage and transport are equally important and could be described.
Line 181-196 (Section 2.3.4.): Very similar to the text in Wex et al., 2019. Sentences in line 191-195 are copied from Sec. 2.7 in Wex et al., 2019. It is difficult to understand without consulting the original description. Section 2.3.4. should be rewritten entirely, explaining more clearly how ground types were categorized and how trajectories were merged to the filter sampling intervals. I suggest (instead of the analysis in 2.3.5.) to include high chlorophyll concentration as a fifth ground type in this analysis. Additionally, precipitation along the trajectories should be considered.
Line 201: state the temporal resolution of the dataset.
Line 202: Shift the description of how INP concentrations and chlorophyll maps were merged from the Result section to here. The DFPC summer data consists of only 17 measurement days and 3 are excluded because land influence, leaving 14 data points. Demonstrate that correlations are robust by showing some scatterplots of grid cells with a strong correlation as a supplement.
Line 203: Explain why a relationship between INP and chlorophyll concentration is expected.
Line 203: Excluding the samples with land input is mentioned several times. Elaborate why this is important.
Line 212, 213: specify, concentration of INP
Line 217: Specify how many trajectories were used and demonstrate that this is a large enough sample to draw conclusions. Looking at the figures it seems that higher CWT is found where more trajectories passed.
Line 227: Justify that longer residence time in a grid box is related to higher INP concentration. I would expect high windspeed to generate more particles, but also less
endpoints at the location because the trajectory moves faster. Discuss assumptions made for this analysis.

Line 227: What uncertainties are avoided by weighting? Motivate the application of a weighting factor. This methodology makes no sense to me and needs a clearer explanation.

Results
Line 237-238, 244ff: Discuss different ice nucleation modes in the introduction section.
Remove here.

Line 240: specify “sharper”

Line 241: Provide a more detailed explanation how “time resolution” and “sampling activities” can explain these differences. Calculate how much of the difference can be explained by the uncertainty introduced by the ice nucleation analysis and how much from uncertainties in sample volume.

Line 242: unclear what the references point at

Line 247: Vali 1975 is a better reference for ice nucleation modes

Line 247-255: This section is speculative. Provide an explanation how the different mechanisms can exert an influence on nINP and why in particular on mixed particles. Much more probable would be an uncertainty in the sample volume.

Line 257ff (Sec. 3.2): For a field study as this, aiming to learn something about the abundance and nature or source of INP, I would consider the differences in concentration of minor importance. Focus should be on the big picture, on trends while being cautious not to overinterpret the data.

Line 264: give concentration ranges at -15°C, -18°C, -22°C to compare to DFPC instead. Specify what can be learned from these concentration ranges.

Line 265: repetition from introduction line 67-68.

Line 272: It is implied that Borys, 1983, Bigg 1996, Bigg 2001 did not measure in the immersion mode. This should be clarified. I recommend merging the literature review here into the introduction.

Line 265-290: Consider presenting the comparison to literature in form of a table and to shift it into a Discussion section. Point out and discuss any systematic differences between marine and land influenced data from the Arctic region at specific temperatures.

Line 279-280: Explain how parameters intervene with INP concentrations. Specify what is meant by “particle activation modality”. Quantify the conclusion that the data are generally consistent to literature.

Line 282: Quantify "reasonable agreement”

Line 285: Quantify “overlaps well”

Line 286: Quantify “wider range”

Line 288: What other factors can explain the differences? It would be helpful to specify the upper and lower detection limits of the methods used here for a comparison to Wex et al., 2019.

Line 291-299: Wide reached and speculative. Sec. 3.7.2. does not provide quantitative evidence on the contribution of continental particles.

Line 299: Name the locations of the high-altitude and coastal measurements in Rinaldi et al., 2017, 2019.

Line 300: Ice formation is usually observed at -15°C in filter based INP measurements and not unique. It is also present in dust rich environments. Provide references for examples showing otherwise.

Line 303: Specify the “special feature”
Line 304: This is the only reference to Fig.2. The figure is not relevant and can be removed.
Line 309: Two size ranges do not qualify as "size distribution".
Line 309: Instead of Table 1, provide a figure showing a scatterplot of INP concentrations measured on PM1 versus PM10 filter in the same time interval (day). All 3 temperatures can be included. Use different colours for spring and summer data.
Line 311: Specify why long distance to source is suggested. Quantify "long distance".
Line 314: To substantiate this interpretation, compare to by much the concentration of particles in the coarse fraction change from spring to summer, based on the measured size distribution.
Line 316: Speculative, the coarse particles could be dust particles. It is not clear to what "above considerations" this is liked to.
Line 308: PM1 data is not depicted in any figure but used in the analysis. Include DFPC PM1 data to Fig.1 and Fig.3.
Line 308-318: the difference between PM1 and PM10 samples is not obvious from this section. Please provide a figure showing both time series together at -15°C, -18°C, -22°C as well as a scatterplot comparing PM1 to PM10 INP concentrations.
Line 319-325: INP concentrations from PM1 and PM10 should be compared to particle concentrations <1um and <10um including all smaller sizes instead of only supermicrometre particles. Otherwise the comparison is not objective and only implies that INP concentrations were similar for both cut-offs and the difference is introduced by the choice through what size range was divided.
Line 327-332: shift to Discussion or introduction. Show how the current data compares to trends found in other studies, e.g., Wex et al., 2019 by plotting the data (at -15°C, -18°C, -22°C) as a function of DOY into the same time series.
Line 333-343: The main findings need to be worked out clearer in this section. Listing a lot of factors at random temperatures in the text is not helpful to understand the situation. Show that the small dataset can be used to determine robust trends. Factors on the order of 2 are small and should not be overinterpreted. Scattering within a season is much higher.
Line 337: What is the reason for limiting the WT-CRAFT dataset to the same period?
Line 340: Why only at -17.5°C and -21.5°C? A plot showing the individual measured T-spectra would be helpful to show how relevant this increase is.
Line 340: quantify "clear nINP peak"
Line 344: Why was the last sample excluded?
Line 345: It is an often-misinterpretation of DeMott et al., 2010. The concentration of particles >0.5um are simply used to parameterize INPs of all size, not an actual size fraction of them.
Line 346: This is incorrect. Aerosol were not more ice active, there were only more INP.
Line 347-350: Speculative. Maybe the nINP is higher because more activity at the station towards the end of sampling.
Line 352: How was the statistical significance of a seasonal trend determined?
Line 353: Quantify "peaked mainly"
Line 355: Speculative. One order of magnitude scattering occurs also on short timescales.
Line 360-371: Explain the relevance of scavenging values for the interpretation of measurements here or delete these lines.
Line 372: Covariance with particle concentration was not shown. This could be an interesting addition to discuss the ice active particle fraction.
Line 374: Quantify "even more accentuated".
Line 375-376: Sentence fragment.
Line 376-383: Unclear what this discussion is aiming for. Clarify main point.
Line 385ff (Sec. 3.6) Converting nINP to ns doesn’t yield new insights. As stated in line
379, INP are only a small fraction of total particles and the total surface area from all different particle types is not related to the number of INP. I recommend deleting the section and Fig. 4 and Fig.5. Instead include the ice active particle fraction at different temperatures and the spring-summer contrast.

Line 389: Repetition of line 72
Line 397: Explain why significance is not found for -16°C. Contradicting results in a narrow T-range could indicate that this analysis is not robust.

Line 408: Quantify “substantial good agreement”
Line 409: The aerosol population at GVB is a mixture of many particle types and only a tiny fraction acts as INP. Interpreting ns compared to ns from well constrained particle types is speculative.

Line 417: Show scatterplots of significant correlations in the supplement.

Line 421: Explain why and how the results are in line with what considerations.

Line 422: If these are general tendencies they should agree with the PM10 data as well. Explain why the analysis is limited to PM1.

Line 423: Quantify “Less clear”
Line 427: Explain why these elements are not good tracers for the soil type. What would be good tracers for the local mineralogy of the soil?

Line 415-429 (Sec. 3.7.1): Add a conclusion, lesson learned from this exercise.

Line 430-441 (Sec. 3.7.2) Suggest some arguments why a larger land fraction (residence time) of a trajectory should linearly correlate to the INP concentration at -15°C. Why not at lower temperatures as well? The distance of land contact to the receptor, time past, precipitation formation along the trajectory and source strength in different land locations should make a large difference.

Line 433: Figure S1 is more informative than Fig.6 to show the overpassed ground types. I suggest changing Fig.6 for Fig.S1. In addition, Fig. S1 is referred to more often than Fig. 6 later in the manuscript.

Line 436: A scatterplot showing fLand versus nINP instead of timeseries would be more helpful than Fig.S3 and Table 3, to show the influence of land sources.

Line 435-438: If a fLand effect is found at -15°C it could be evidence against biological INPs dominating nINP at this temperature.

Line 440: specify what the “outcome” is and provide an overall conclusion from Sec. 3.7.2.

Line 443-482 (Sec. 3.7.3) Specify that this analysis was performed using 14 datapoints from PM1 DFPC. It should be demonstrated that the limited dataset yields robust correlations with CHL. Show some scatterplots. I suggest to include high CHL regions as a ground condition, subdividing the sea category, and include it in the analysis of sec. 3.7.2.

Line 443-459: The hypothesis and description of how INP and CHL maps are correlated fits better to the Method section.

Line 451: Why are trajectories with land contact excluded and why only some? A short land contact can have a large impact on nINP.

Line 460: The time lag doesn’t make sense to me. Why would the aerosol generating biochemical process not change location in 6 days or 16 days? The movement of the surface water should be considered.

Line 464: justify why 6- and 16-day time-lag was selected

Line 477-478: Explain how this can be seen in Fig. 7c?

Line 479-482: Consistency is not obvious. There seems to be even more negative correlations. The pattern looks random. I would expect some high productive areas based on ocean currents and biological factors that do not change rapidly.
Conclusion
Avoid euphemistic language.
Line 484-489: It seems that the paper gains little by including the WT-CRAFT dataset. It is only marginally relevant to discuss seasonality in sec. 3.4. In all other sections it is only mentioned that the data agrees with what was seen from analysing DFPC data. It there was a dependency on ice nucleation mechanism (condensation, immersion) it seems not to make a difference on size and source of INP.
Line 490-493: More, larger INP in summer seems contradictory to the absence of a seasonal trend.
Line 497: inter-annual variability is a trivial statement. The question is how large the variation is and why it happens.
Line 499: Explain the importance of this study in detail.
Line 502-505: This is a weak conclusion. There is no reason mentioned to assume that only one source contributes INP at all temperatures in the spectra.
Line 506-508: The relation has not been proven without doubt. It is a speculative interpretation.

Figures
The provided figures do not support the content of the manuscript. Fig.2 for example can be deleted is only referred to in a side note and supplementary figures are referred to more often than the actual figures included to the manuscript.
Helpful figures could include:
1. Temperature spectra with all nINP measured with WT-CRAFT and all datapoints (PM1 and PM10) measured with DFPC, with colour code for DOY of measurement and 3 symbols to differentiate the techniques.
2. Timeseries of nINP data at -15C, -18C, -22C as function of day of year including the 12 data points from Ny-Alesund in Wex et al, 2019,
3. Scatterplot of PM1 vs. PM10 nINP measured with DFPC
4. Timeseries showing activated fraction of particles (nINP divided by number of particles in PM1 for DFPC, in PM10 for DFPC and CRAFT) at -15C, -18C, -22C.
5. Fig. S1

Fig.1: Showing the individual measurements would be more informative than only median, min, max. Please change the figure accordingly and indicate summer, spring, PM1, PM10 samples in different colours and symbols. Indicate the detection limits of the DFPC and WT-CRAFT.
Fig.2: Comparing it to Fig. 4a in Irish et al., 2017 did not make it obvious how it was adapted. What assumptions are made to overlap the two y-axes (INP in water and INP in air)? This figure can be deleted.
Fig.3: It needs to be specified how the measurement uncertainty is determined from the sample volume and the analysis. It appears several times throughout the paper and is important.
Fig. 4: It would be more informative to scale the nINP with the total aerosol number. This would show that INP are not from the bulk aerosol population but rare exceptions. Summer, spring and PM1, PM10 can be contrasted.
Fig.5: redundant to Fig.1 no new information in this figure. Remove.
Fig.6: The main message from this figure seems to be that sea ice is melting in summer. This is trivial.
Fig. 7: c) It seems the colour bar shows nINP because it is written on top of it. Replace and label colourbar with units. Add a minimal explanation what can be interpreted from the patterns.
Fig. S1: Define ground types in the figure caption.
Fig. S2: why are some points connected by lines and others not? Homogenize all precipitation scales and nINP at the same temperatures. -18°C and -22°C plot in first column are switched. Second column last plot DFPC instead of FPC.
Fig. S3: use same fLand scale for all DFPC and WT-CRAFT subfigures and same nINP for same temperatures.
Fig S4, S5: Include chlorophyll as fifth land type in S1. Remove figures.
Fig. S6: use same colour-scale range for all subfigures. Use Fig. S1 map design to facilitate comparison. It seems regions where more trajectory points (Fig.S1) pass, also show higher CWT. This points to a problem with the small dataset size for this analysis.

Tables
The robustness of correlations in Tables 2 a, b and 3 would be clearer when shown as scatterplots. Due to the small size and structure of the data used, the derived linear correlation coefficients might be strongly biased by few outlier data points and be therefore misleading. Scatterplots help to visually judge correlations. Person’s R is sensitive to the data distribution and the R value can be generally misleading. Scatterplots of fLand and nINP would be helpful to investigate these issues.

Technical corrections
Delete “apparently”, “likely”, “noteworthy”, “worth highlighting” throughout the manuscript.
Line 73: icebreaker
Line 100, 168: km instead of Km
Line 104: Section instead of Par
Line 115: define TSP, define OD
Line 132, 142: per m³ not per m⁻³
Line 140: replace super-microliter with 3uL
Line 230: define $\bar{D}$
Line 239: nINP instead nIPN
Line 339: p<0.05 instead p<0.5

References


Murray, B. J., and Bertram, A. K.: Ice-nucleating ability of aerosol particles and possible sources at three coastal marine sites, Atmospheric Chemistry and Physics, 18, 15669-15685, 10.5194/acp-18-15669-2018, 2018.
