Response to RC2

The authors would like to thank Anonymous Referee #2 for their insightful comments, particularly with regard to the flow of the text, and their attention to details. Below we list their comments, our responses and subsequent changes to the text and figures.

General comments:

Overall, the manuscript is a bit long-winded owed to the fact of an extensive analysis and many links to existing literature – what the reviewer highly appreciates. At some points the manuscript seemed repetitive (see specific comments below) and the readability/overview of the manuscript should be improved introducing more subsections (so far there are only two subsection titles for 17 pages of results) guiding the reader through the manuscript. Furthermore, (to the reviewer) the climax of the manuscript is reached with the conclusion of different parameterizations needed for fresh and aged dust, respectively (L620-623). Afterwards, the results continue, however, with the analysis of the sea water samples, which are also used to assess the partitioning of INPs arising from dessert and sea sources, a discussion that had previously been addressed using different proxies. The authors may want to cover the discussion of the sea samples earlier in the manuscript when discussion the source portioning and in order to work towards the conclusion L620-623. This could potentially increase the readability of the manuscript, but the reviewer leaves the decision of changing the manuscript structure up to the authors.

Thank you very much for the helpful suggestions on how to improve the manuscript flow and readability. We have updated the methods and results sections with additional subsections including:

- L267: Methods 2.4 INP Measurements in Aerosol
- L398: Methods 2.5 INP Measurements in SSW
- L429: Results 3.1 INP Concentrations in Aerosol
- L472: Results 3.2 Seawater Source Potential
- L515: Results 3.3 Ice-active Surface Site Densities in Aerosol
- L627: Results 3.4 Characterization of INPs in Aerosol
- L675: Results 3.5 Characterization of INPs in a Soil Dust Sample
- L701: 4 Discussion

We have also reordered the results section as suggested, placing the subsection on the SSW measurements earlier on to help build evidence toward the conclusion that marine INPs were not likely observed while avoiding repetition.

The reviewer highly appreciates the effort of the authors to apply their methodology also to a Saharan dust sample of the previous study by Niemand et al. 2012. What would be valuable is to let the reader know how the results in this study compare to the initial measurements of Niemand et al. 2012. Was the same ice nucleation (IN) activity

observed? Could storage effects (e.g., Beall et al. 2020, Stopelli et al. 2014) have altered the observed IN activity? From the reviewer's viewpoint this is essential information to assess how representative the reduction in IN activity due to the two treatments is, which is later used to support the conclusion for the need of different dust parameterizations. In addition, the sample preparation for analysis is not described in the manuscript and should be added in the methods section. Furthermore, the sample's association to the study by Niemand et al. 2012 should be indicated at all instances in the text and in the caption of Figure 5 for clarity. The authors may consider plotting the original INP concentration observed by Niemand et al. 2012 to Figure 5.

We agree that the impacts of storage on measurements are an essential consideration and have added several lines to the text about (see responses to specific comments below).

The N12 SD sample we used for this manuscript was generated during the DeMott et al. (2018) Fifth International Workshop on Ice Nucleation phase 2 (FIN-02) intercomparison experiment (see Table 3). While the dust in the archived aliquot we used is the same as featured in Niemand et al. (2012), the experimental design differed and no surface area based measures were available from DeMott et al. (2018), and thus we unfortunately cannot directly compare our results with N12 or consider the effects of frozen storage on the sample since the Niemand et al. (2012) experiments.

Another consideration is how the SD sample may have changed since it was originally collected for N12 (i.e., from 50 km N of Cairo), because if storage affects the IN-properties of soil dust, this would be the point at which it potentially was most comparable to the sources encountered during the campaign. However, I think this kind of discussion would be too speculative for the manuscript for multiple reasons and is likely beyond scope. Niemand et al. (2012) provides few details about how and when the sample was originally collected or how it was subsequently stored. Similarly, (DeMott et al., 2018) provides no details on N12-SD storage before use. Hence, this sample already had an unknown storage history, even before its use in the FIN-02 campaign.

Moreover, there are still too few papers on storage impacts on INP samples. The findings in e.g. Stopelli et al. (2014) or Beall et al. (2020) are useful for precipitation samples, but it is yet unclear how consistent these results are across different sample types. We think this is an important topic for future studies to investigate.

We have added the following clarification about the SD sample to the text. We also added the following text to more clearly link the SD sample to Niemand et al. (2012).

Sect. 2.4, L391: "INP concentrations were additionally measured in in untreated, heat-treated, and peroxide-treated subsamples from an archived suspension of the soil dust sample N12 SD for comparison with this study (DeMott et al., 2018; hereafter referred to as "N12-SD"). Briefly, the sample was generated during the recent laboratory intercomparison of INP measurements (DeMott et al., 2018) collected on a 0.2 µm Nuclepore polycarbonate membrane filter (Whatman ®,

Chicago, Illinois, USA) and stored frozen at -20 °C until processed, as described in DeMott et al (2018)."

Sect. 3.5, 680: "For comparison with this study, we measured INPs in untreated, heat-treated, and peroxide-treated subsamples of an archived suspension of the N12-SD sample (Methods Sect. 2.4; DeMott et al., 2018) ."

Sect. 3.5, L682: "Sample N12-SD exhibits sensitivity to both heat and peroxide at temperatures > -16 °C, indicating biological composition of INPs at high freezing temperatures. Multiple AQABA samples influenced by desert air mass sources show similar sensitivities at higher temperatures: f006, f007, f019, and f020."

Sect. 3.5, L686: "Overall, the heat and peroxide sensitivities in the N12-SD sample indicate that desert dusts may contribute…."

Sect. 3.5, Figure 6 caption: "Measured concentrations of INPs in an aerosolized soil dust sample **"N12-SD"…"**

Generally, storage effects on both the obtained filter and seawater samples seem to not be addressed in the manuscript. Given the storage of the filters of up to more than three years (L251), had a change in IN activity been observed, e.g., a decrease in average IN activity with storage time? Could a potential decrease due to storage also explain the reduced INP concentrations in the samples compare to previous work? This should be added to the discussion and be potentially mentioned in the conclusions, where the possible explanations for the reduced observed INP concentrations are listed. Furthermore, how long have the sea water samples been stored before analysis?

We thank the reviewer for pointing out these omissions. It is reasonable to discuss how the storage impacts described Beall et al. (2020) may have affected SSW samples. We think it is too speculative to consider how the Beall et al. (2020) findings relate to aerosol collected on filters in detail, though we could mention the range of changes demonstrated in that study and that this would be too small to explain the discrepancy in n_s shown here. Desert dusts in this region would have already experienced extreme heat and radiation prior to being lofted and sampled, and these conditions would likely have already denatured labile INPs. To speculate briefly here that the Beall et al. (2020) findings are consistent with storage impacts on aerosol collected on filters, then it would be important to interpret results from the treatment experiments as a whole, rather than focusing on treatment impacts to any one sample. The average changes to INP concentrations reported in Beall et al. (2020) were relatively small for untreated and heat-treated samples (within a factor of 2), but changes within the 95% CI derived from the lognormal distribution of changes exceed 10x for some freezing temperatures. In Sect. 3.4, we focus on what is consistent across the 12 samples tested, namely the sensitivity to peroxide at T > -15 °C and the frequent heat sensitivity

at T > -10 °C. We think this approach is appropriate given the uncertainty of storage effects the reviewer points out. As mentioned above, we think impacts of storage on INP collected in aerosol on filters is an important topic for future studies.

Previous tests on nascent sea spray generated with a Marine Aerosol Reference Tank (MART), and collected on the same filter type as in AQABA (McCluskey et al., 2017) showed no effect of freezing storage upon INP concentrations (Fig. 1).

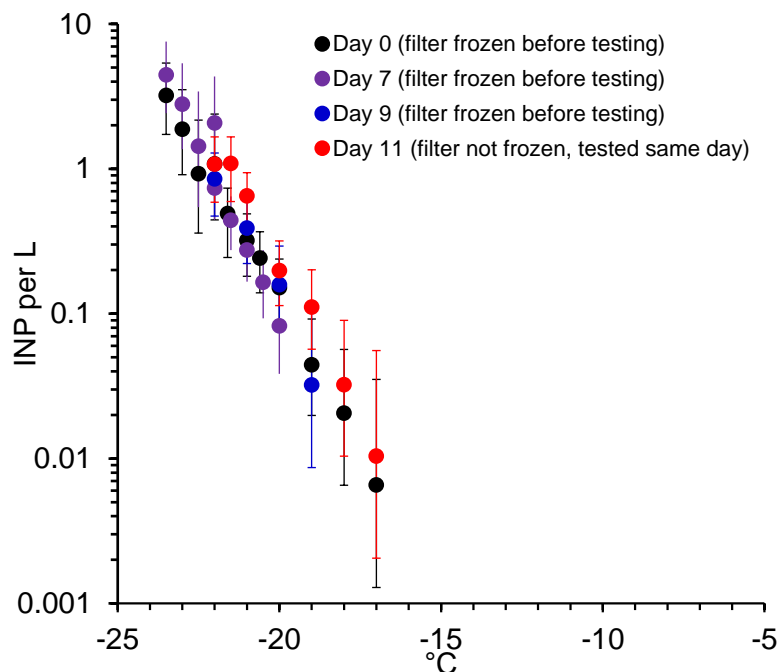



Figure 1. Comparability of INP spectra of aerosol collected from a MART experiment that were stored frozen before testing (black, blue, purple) and tested without storage freezing (red). This MART's INP emissions remained steady for an extended period before gradually declining.

It is also worth noting that the dust samples used to derive the N12 and U17 parameterizations were stored for several years prior to analysis.

Sect. 2.5, L402: "SSW samples for INP analysis were collected in 15 mL sterile centrifuge tubes (Falcon™, ThermoFisher Scientific, Waltham, Massachusetts, USA) and stored frozen at -20 °C until they could be shipped in a dry shipper via Cryoport® (-180 °C) and ultimately stored at -80 °C as for aerosol samples **until processed as described above (Sect. 2.4),** within 18 to 38 months of collection. **Storage duration was not correlated with INP concentration changes in frozen marine and coastal precipitation samples (Beall et al., 2020)."**

Sect. 3.2, L503: "These results are indicative of the particulate organic carbon (POC) type of marine INP defined in McCluskey et al. (2018a) (Fig. S13), though this result should be interpreted with caution as storage could potentially have increased sensitivity to filtering treatments.

Understanding of storage impacts on INPs measured in SSW is lacking. However, Beall et al., (2020) showed that average INP concentration changes for untreated coastal precipitation samples due to frozen storage were within 2× of INP concentrations measured in fresh samples, with changes at the upper or lower end of the 95% CI exceeding 10× for some freezing temperatures. If SSW samples are similarly sensitive to storage, we would expect INP concentration changes to be within 2× on average, but up to > 10× for any particular untreated sample. Beall et al. (2020) also reported similar changes INPs < 0.45 µm with a greater tendency toward losses, which indicates that storage may have caused increased sensitivity to the filter treatments applied to stored samples."

Sect. 3.3, L569: "However, considering that SSA is associated with 1000 times fewer IN sites per unit surface area than dust (i.e. 1000× lower n_s) (McCluskey et al., 2018b), the characteristically low IN activity of untreated SSW (even in light of the modest changes expected from storage, Sect. 3.2), and the frequency of dust events during AQABA, our findings suggest it is unlikely that the observed INPs originated from SSA."

Sect. 4, L791: "Storage protocol represents another difference between Price et al. (2018) and the two surface-based studies. Gong et al. (2020) and the present study stored samples frozen prior to analysis, whereas Price et al. (2018) processed samples immediately after collection. The understanding of storage impacts on INPs collected on filters is lacking (Wex et al., 2019), but we note that the discrepancies in n_s between the two surface-based studies and Price et al. (2018) exceed the range of INP concentration changes reported in INP precipitation samples stored frozen (Beall et al., 2020)."

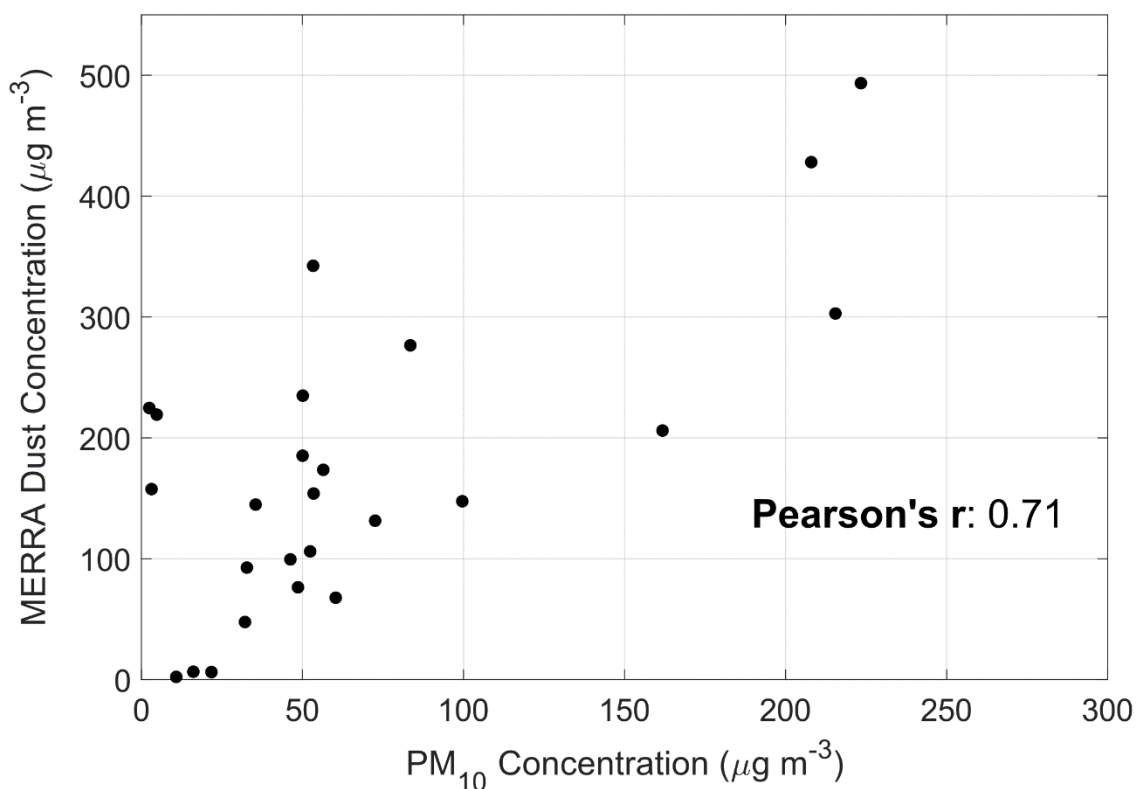
It is not entirely clear to the reviewer why the MERRA re-analysis data was chosen to assess the dust mass concentration albeit aerosol in situ observations were available on the ship. It is understood that using MERRA source regions could potentially be identified and the aerosol classes allow for the discrimination of different sources. However, these capabilities seem to not be used (for the source apportionment FLEXPART is used, for discrimination the sources it appears an AMS and a MARGA was used). Furthermore, from Table 1 it seems that the re-analysis average dust concentration did not necessarily relate to the in situ aerosol surface area. If the re-analysis is representative, a potential relation should exist to PM10 derived from the in situ measurements assuming a representative dust density. To support the representativity of the re-analysis data, this relation should be explored and supported with e.g., a correlation coefficient. Would the usage of an in situ derived PM10 compare better to the INP concentration (cf. Figure S7)? If so, the usage of an in situ derived PM10 might be more appropriate. The authors may have addressed

this approach in their analysis already. I would suggest elaborating briefly on their decision from where to take the dust concentration in the manuscript.

Thank you for these suggestions. We have added a figure showing the correlation coefficient with PM_{10} as suggested. Dust concentrations were not measured during the campaign and this was the motivation for retrieving this data from MERRA-2.

Sect. 2.3, L248: Buchard et al. (2017) showed a high degree of correlation between MERRA-2 and surface dust concentration observations ($r \geq 0.69$), particularly during dust storms ($r \geq 0.92$). MERRA-2 surface dust mass concentrations also correlated well with PM_{10} observed during AQABA ($r \geq 0.71$) (Fig. S6).

Sect. 2.3, L245: "Since dust concentrations were not measured during the campaign, hourly dust surface mass concentrations along the cruise track…"

Supplementary Figure 6. Comparison of MERRA-2 surface dust mass concentrations and observed PM_{10} during INP sampling periods.

We checked for correlation between INP concentrations and *in situ* PM_{10} and found that this too was not well correlated (Pearson's $r < 0.34$ across all freezing temperatures).

Throughout the manuscript, the authors need to perform a delicate balance between the analysis of individual samples and the implications to the overall big picture. At some locations in the manuscript very specific and detailed information of individual samples are given without putting it into context with other samples or the big picture (e.g., for L661-663, while the relation to the study of McCluskey et al. 2018 is very interesting to an INP researcher, indication in only one observed sample does not fully allow general conclusions). The authors may consider reading critically through their manuscript, removing side thoughts to streamline the manuscript. The reviewer acknowledges that the importance of such small connecting statements is subjective and leaves it up to the authors to assess their importance.

This is a good point. We removed the statement referenced here about s001. We also moved several details about MERRA-2 from the main text to the supplement so that Sect. 2.3 is brief and focused on the simulation of dust.

We have left the discussion of sample f010 because the anomalous results were consistent between the heat-treated and peroxide-treated subsamples and have been reported in multiple prior studies (McCluskey et al., 2017; Martin et al., 2019; Paramonov et al., 2018). Another reviewer suggested that we include this as one of the manuscript's outlook study topics.

The presented supplementary information is appropriately chosen to support the understanding of the manuscript while not interfering with the main storyline of the manuscript. Despite being supplementary, the material should be presented in a more structured way, i.e., all figures should feature a corresponding caption beneath, and – in the reviewer's opinion – a brief explanatory description as text should be given. Additionally, the reviewer encourages the authors to revise some of the figures in terms of readability, consistency, and quality. Specific comments follow below.

The Supplement has been reorganized as suggested, and the figures have been checked and updated for readability, consistency and quality.

Specific Comments:

No finding related to the SSW samples is mentioned in the abstract. The authors may consider adding an obtained insight from these samples.

We have added the following to the abstract:

Abstract, L53: INP concentrations in seawater samples ranged between 3 and 46 mL^{-1} at -19 °C, demonstrating the relatively low INP source potential of seawater in the region as compared to seawater from multiple other regions.

L49: It should be annotated that the dust concentration was obtained from re-analysis data in contrast to the in situ measured INP samples.

Abstract, L48: Over half of the samples (at least 14 of 26) were collected during dust storms with average dust mass concentrations between 150 and 490 µg m^{-3} (PM$_{10}$), as simulated by the Modern-Era Retrospective analysis for Research and Application, version 2 (MERRA-2).

L120: The reviewer enjoyed the rather extensive but nicely written introduction to INP parameterizations.

Thank you!

L174: Different instruments were connected to different inlets. How do the different inlets compare, were the sampling lines similar per instrument? Furthermore, it would be very beneficial to have an overview picture/schematic of the different sampling sites to see their relative locations, e.g., are there any obstacles between the different inlets?

Thank you for this suggestion. We have added a figure to the supplement to illustrate the locations of the different sampling sites and inlets. We have also added the following clarification on the different sampling line and inlet set-ups and a comment on what the differences mean for our results.

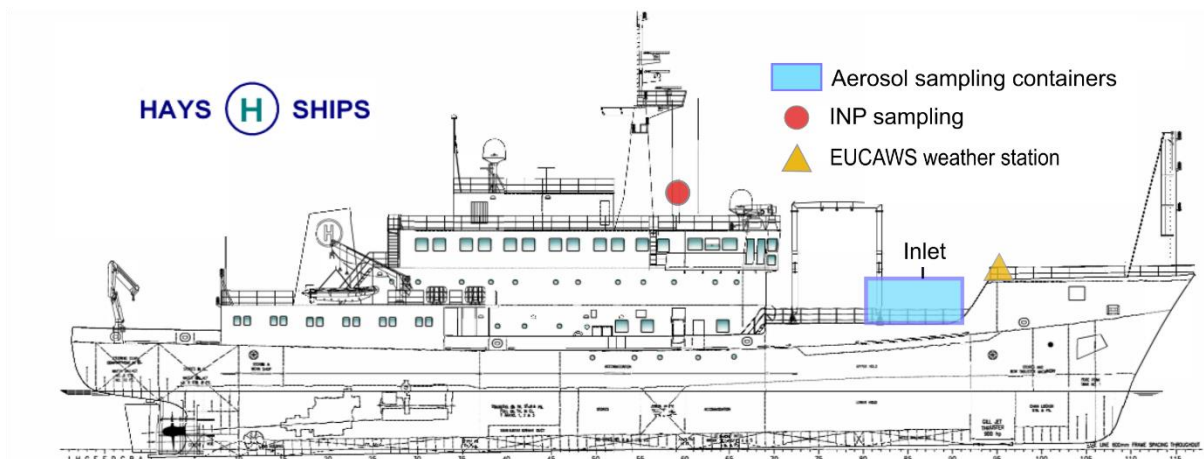
Sect. 2.4, L268: Ambient aerosol sampling for offline measurement of INPs was conducted from 5 Jul – 31 Aug 2017 on the *Kommandor Iona's* wheelhouse top (platform above the bridge), ~15m from the online aerosol measurements inlet and ~15 m from the ocean surface (Figs. S3-4).

Sect. 2.2, L187: "The inlet for the aerosol instrumentation was located at the top of a measurement container at a horizontal distance of about 15 m from the INP filter sampling unit **(Fig. S3)**."

(the details about the drying system in the sampling line for aerosol instrumentation remain in the following line)

Sect. 2.2, L235: "The MARGA sampling line was equipped with a PM$_{10}$ cyclone but the sample was not dried as the instrument is not prone to condensation."

Sect. 2.4, L285: "Lacking a size-selective inlet for INP sampling, it is possible that aerosols > 10 µm were present in INP samples during dust events. Surface area may be underestimated for these samples due to the PM$_{10}$ threshold applied to aerosol sizing (Sect. 2.2 and S3), but we do not expect this to affect our overall conclusions as increased aerosol surface area would further depress n_s (see Results Sect. 3.3 and Discussion Sec**t. 4**)."

Supplementary Figure 3. Aerosol sampling, INP sampling and weather station locations.

L176: The authors may want to consider adding the range of typical RHs during the campaign.

Sect. 2.2, L191: "**Ambient RH ranged between 67 and 81% during INP sampling periods.**"

L177-L182: It would be beneficial to briefly explain the used thresholds/logical arguments of the filter flag.

The following clarification was added.

Sect. 2.2, L196: "The flag was set when the apparent wind direction was from the direction of the stack (± 30°) as seen from the aerosol inlet position (Fig. S3) and strong fluctuations of black carbon and/or particle number concentrations were observed relative to background levels."

This flag was developed by manual inspection of the time series of apparent wind direction, particle number concentration and black carbon. This is not a very "strict" definition of self contamination in terms of exact numbers. However, when you look at the plotted time series it becomes clear that self contamination is very clearly detected with this combination of conditions. Even if only the edge of the exhaust plume was measured, particle number concentrations and BC concentrations immediately go through the roof. If this is associated with wind from approximately the direction of the stack, it is quite clear that this was from the stack.

L187-L188: Is there a reference for the mass ratio assumption? The authors may consider adding the values of used refractive indices and shapes, including a reference if possible.

We have added this information to the Supplement S1 and S2.

L191: Is there a reference for the AMS data obtained during AQABA?

So far there is no reference for the AQABA AMS data.

L198: How was the uncertainty of 30% derived? Please elaborate. We have added the following details to the main text and Supplement.

Sect. 2.2 L226: The overall uncertainty of derived particle surface area concentrations is estimated to be 30%, including the uncertainty due to particle losses (see Sect. S3 for details).

Supplement S3:

S3 Estimation of aerosol surface area concentration uncertainty

As described in Sect. 2.2, the particle surface area concentrations were calculated from FMPS (dp = 5.6 nm – 560 nm) and OPC (dp = 250 nm – 32 µm) measurements. The final particle surface area concentrations have an uncertainty that is the result of several contributions: inlet and transport losses, re-binning uncertainty, FMPS scaling uncertainty and uncertainty due to counting statistics. The details on uncertainty estimation for each component are provided below.

S3.1 Inlet and transport losses

Inlet and transport losses were calculated using the Particle Loss Calculator (von der Weiden et al., 2009). FMPS losses were below 10% for particles larger than 15 nm and below 2% for particles larger than 30 nm. For the OPC the losses were well below 1% up to particles of 3 µm diameter and reach 10% for particles of 6 µm diameter.

Since most (~75%) of the particle surface area concentration is found between 30 nm and 1 µm (where losses are between 2% and 0.2%) and most of the rest is found between 1 µm and 7 µm (where losses increase from 0.2% to 20%), overall losses of the particle surface concentration are less than 3%.

S3.2 Instrumental measurement uncertainty of FMPS and OPC

Measurement uncertainty of the FMPS and OPC, e.g. as a consequence of calibration uncertainties or uncertainties of the flow rates, have been determined by co-located measurements with other instruments over extended time intervals. The OPC and FMPS uncertainty per size bin are 10 and 15%, respectively.

S3.3 Re-binning uncertainty

For calculation of the surface area concentrations, geometric diameters have been calculated from optical particle diameters for the OPC measurements. This re-binning requires making assumptions about the composition of the coarse particles (fraction of dust and sea salt, respectively) and about the optical properties of the fine particles (chemical composition taken from AMS and BC measurements; see Sects. S1-2). The ensuing uncertainty in the calculated particle diameter of the re-binned size distribution results in uncertainties of the calculated surface

concentration. From sensitivity tests we estimate the resulting uncertainties from re-binning for the particle surface concentration to be around 5%.

S3.4 FMPS scaling uncertainty

A scaling factor of 1.3 (with an uncertainty of 10%) was applied to PS1 to correct for under-measurement of particles in the upper measurement range of the FMPS (above ca. 100 nm). As a consequence of low-biased particle concentration measurements in the FMPS, the calculated surface concentrations were too low. To correct for this effect, the FMPS and OPC data were merged again, not averaging the overlap size bins but interpolating from the upper end of the correct FMPS data up to the lower end of the OPC. From comparison of the result of this exercise with the previous calculations, a correction factor for particulate surface area $< 1 \,\mu m$ (PS$_1$) of 1.3 (with an uncertainty of 10%) was determined. The particulate surface area $< 10 \,\mu m$ (PS$_{10}$) concentrations were corrected accordingly by accounting for the corrected PS$_1$ contribution. From temporal variability of this scaling factor, we estimate the uncertainty of this effect to result in an uncertainty of the particle surface concentration of 10%.

S3.5 Uncertainty due to counting statistics

Relative uncertainties due to counting statistics were calculated from the average number of particles per size bin as measured during a sampling interval of 60 s (probed volume = 1200 cm^3) for different size bins. Smaller sizes have smaller relative uncertainties due to larger numbers of counted particles (see examples in Table below).

Particle diameter (µm)	Average number concentration (cm^{-3})	Average sampling rate (min^{-1})	Relative uncertainty (%)
1	1	1200	3
5	0.04	50	14
10	0.004	5	45

From the average surface distribution, we find that particles smaller than 1 µm in diameter with a counting statistics uncertainty below 3% contribute ~75% of the surface area; 25% of the surface area is found for particles between about 1 µm and 8 µm with an estimated average counting statistics uncertainty of about 15%. Therefore, we estimate that the overall counting statistics-related uncertainty of the particle surface concentration is in the order of 5-7% for 1-minute averages.

Considering all the components to the overall particle surface concentration uncertainty described above, we conservatively estimate the uncertainty of the particle surface concentration to be 30%.

The only systematic bias is due to the inlet losses, which were estimated to have an effect of less than 3% on the total surface concentration.

L208: Here again, an overview graphic (in the main text, supplementary, or referenced from a different publication) would strongly help to understand the relative locations.

Thank you for this suggestion. As mentioned above, the suggested figure has been added to the supplement (Fig. S3).

L211: What is meant by "area-averaged"? Please elaborate.

The following details have been added to the text:

Sect. 2.3, L245: "Since dust concentrations were not measured during the campaign, hourly dust surface mass concentrations along the cruise track were obtained from the (0.5×0.625 °) Modern-Era Retrospective analysis for Research and Application, version 2 (MERRA-2; Gelaro et al., 2017) and were averaged over the region covered during each sampling period."

L230: Are the ~25 m displacement vertically or horizontally? Horizontally.

Sect. 2.4, L268: "Ambient aerosol sampling for offline measurement of INPs was conducted from 5 Jul – 31 Aug 2017 on the *Kommandor Iona's* wheelhouse top (platform above the bridge), ~15 m **horizontally** from the online aerosol measurements inlet and ~15 m from the ocean surface (Figs. S3-4)."

L233: What determined the sampling length? What determined the sampling frequency? Please elaborate.

Sampling intervals and frequency were determined by multiple factors including OPC data, frequent stack sampling during Leg 1, and the evolving security situation due to pirate activity, particularly near Bab-el-Mandeb. Access to the sampling location was restricted from time to time during periods of big swell or elevated security risk. For the sake of brevity, we can just say we aimed to collect > 5000 L during dust events and >10,000 L when particle counts were relatively low (e.g. during sampling periods f044) as conditions allowed.

Sect. 2.4, L275: "**Sampling intervals and frequency were chosen with the aim of collecting > 5000 L during dust events and > 10,000 L when OPC particle counts were relatively low (e.g., during sampling periods f040-44), as conditions allowed.**"

L238-L242: Just to ensure: The relative wind direction towards the sampling unit has not been used to switch off the sampling pump?

Correct. Relative wind direction would also have been equally useful for this, though NO is a more direct measurement of stack emissions.

L253: How was the shaking realized?
The following clarification was added.

Sect. 2.4 L298: To release collected particles, filters were immersed in 5-8 mL ultrapure water (Cat. Number W4502, Sigma-Aldrich®, St. Louis, MO, USA) and shaken by hand for 20 minutes just prior to measurement.

Eq. (1) + E1. (2): Consider adding a reference for the equations.

A reference was added.

L315: "…**where V_d is the volume of the sample in each well (Vali, 1971).**"

L287-L289: Given the spread of two orders of magnitude of the background INP concentration, the reviewer does not fully agree with the used background correction methodology. The degree of contamination of the field blanks when inserting them in the sampling unit and removing them could likely be a function of the ambient aerosol concentration while taking the field blank. Thus, subtracting the timely closest background from a sample might be more meaningful. However, the reviewer acknowledges that this change would not substantially affect the background corrected INP concentrations and does not need to be implemented.

We thank the reviewer for bringing up this point on field blank corrections. There were multiple factors to consider in the approach to the correction. An additional consideration was that the vessel surfaces and hardware got salty and dusty at times. The surfaces around the sampling unit were frequently wiped down, but build-up in between these periods could potentially have affected the field blanks and samples. Thus, we chose to consider the average of all field blank measurements rather than correct for the closest one temporally. For example, the closest blank temporally for a given sample may have been taken after a heavy dust event, conditions which may not have been as representative as another field blank. Moreover, the amount of contamination from ambient aerosol is insignificant compared to those attached to and released from surfaces during unit handling, and the removal and storage of the filter. We also agree that changing the correction method would have little impact on the results based on our experience trying multiple approaches before arriving at this one.

L293: When overlapping, were the INP concentrations from diluted samples combined with the INP concentrations from the undiluted samples? If so, please specify.

We consider both values from the diluted sample and the undiluted sample, and these are shown in Figs. 2-3. For the freezing temperature ranges where there is overlap between the diluted and

undiluted spectra, most of the 95% CI overlap. There are one or two data points in f035, f042, f044 where the confidence intervals do not overlap. In the case of Fig. 1, the value for INPs at -15 °C as measured in the undiluted sample is shown, and thus could be considered conservative for f035 and f042. Since -15 °C was beyond the limit of detection in the undiluted f044 sample, the diluted sample value for INP concentration was used.

We have made the following updates to the text.

Figure 2 caption: "For the 8 samples on which a dilution was performed (Fig. S8), n_s for both the raw undiluted and diluted sample are shown."

Figure 3 caption: "For the 8 samples on which a dilution was performed (Fig. S8), n_s for both the raw undiluted and diluted sample are shown."

Eq. (3) + Eq (4): While Eq. 3 and Eq. 4 are presented in Kanji et al. 2017, a reference like e.g., Hiranuma et al. 2015 explaining the approximation to Atot might be more helpful to the reader.

Thank you for this suggestion.

Sect. 2.4, L353: The difference between the n_s approximation (Eq. 4) and n_s (Eq. 3) is that many particle types are typically included in the n_s approximation, and in an ambient aerosol measurement most of these are not IN-active (see also Hiranuma et al., (2015) Sect. 2.4).

L315: What was the motivation/criteria for the selection of the specific 12 samples? Please elaborate.

The following has been added:

Sect. 2.4, L366: "The 12 samples were selected based on sampling location with the aim of getting a representative measurement from each region."

L315-L338: Has this procedure/protocol of heating and hydrogen peroxide treatment been previously described? If yes, the authors may want to shorten the description and cite the existing literature.

Yes. The following clarification has been added. We have shortened the section on the procedure from the methods.

Sect. 2.4, L364: "Heat and peroxide treatments were applied to a subset of samples (12 of 26) to test for heat-labile biological (e.g., proteinaceous) and organic INP composition, respectively, following the procedure described in McCluskey et al. (2018b) and Suski et al. (2018)."

Sect. 2.4, L368: "For each heat-treated sample, a 2 mL aliquot of the original ultrapure water suspension was heated to 95 °C for 20 min in a water bath and re-tested to assess the reduction in INP concentrations. For peroxide treatments, 1.6 mL of the original suspension was combined with 0.8 mL of 30% H_2O_2 (Sigma Aldrich®, St. Louis, Missouri, USA) to achieve a final concentration of 10%, then the mixture was heated to 95 °C for 20 min while being illuminated with two 26-W UVB fluorescent bulbs. To remove residual H_2O_2, to prevent otherwise significant freezing point depression, the solution was cooled and catalase (Cat. number IC10042910, MP Biomedicals, Santa Ana, California, USA) was added."

L345-L346: Also here, how were the 5 samples selected? Please elaborate.

The following clarification has been added. Sect. 2.5, L407: "Heat and peroxide treatments as described above were applied to 5 SSW samples from the Arabian Sea and the Gulf of Aden. The focus on these regions was motivated by the detection of marine aerosol originating from the upwelling region in Somalia reported in Edtbauer et al., (2020; see Sect. 3.3)."

Line 349ff: How frequent were back trajectories released, how many were investigated per filter sample, and which time (e.g., start or end time of sampling) was considered for choosing the displayed trajectories? Is there a reason for choosing 72 hours? Furthermore,

the authors may consider indicating the trajectory locations in plots S8 and S9 where the trajectories were below a certain threshold height indicating potential air mass uptake.

The figures were edited as suggested, and the details requested were added to the text. After making the suggested changes with regard to the threshold height, the choice of 72 hours is no longer relevant to the description of figures since the time at which each back-trajectory crosses the height threshold was less than 72-hours. 2000 particles were released for these simulations continuously over the duration of the sampling period (meaning 2000 back trajectories were simulated per sampling period).

Sect. S5: "Particle releases (n = 2000) from 35 m above sea level (ASL) followed the vessel track using vessel position information from the European Common Automatic Weather Station (EUCAWS; http://eumetnet.eu/; last access Sept. 2021) for the duration of each sampling period."

Sect. 3.1, L446: Figs. S9-S10 show the extent of k-means clustered FLEXPART back-trajectories below the altitude of 1500 m. This threshold was applied to eliminate most of the free tropospheric parts of the back-trajectories and was selected based on the MERRA-2 monthly average planetary boundary layer (PBL) heights during the campaign period, which were 200-700 m over the ocean and up to 1700 m over land."

L415: It could be insightful to add a flag to Table 1 to indicate which samples are thought to be sampled during dust events. Otherwise consider changing "(Table 1)" to "(dust concentration > 50 µg m-3, see Table 1)

Thank you for this suggestion.

Sect. 3.3, L546: "Considering the frequency of dust events encountered (dust concentration > 150 µg m^{-3}, see Table 1), …"

L415-L416: How was a high probability of dominant dust contributions determined? To the reviewer's understanding this assessment is based on increased PM10 loadings of the reanalysis data. However, it seems that the reanalysis data does not necessarily correlate with the in situ observed particle surface area concentration (Figure S7). Has the relation between re-analysis data and in situ observations been investigated to support the assessment of dust contributions solely on re-analysis data? If no relation prevails the reviewer would suggest tempering the statement from high probability to likely given a potential absence of in situ increased concentrations.

Thank you for this suggestion. As mentioned above in the general comments section, we have added a figure to the Supplement showing the correlation coefficient with PM$_{10}$ as suggested.

L428+L429: In Table 1 the seasalt concentration is given. If this is used as proxy for SSA, the authors should specify this and explain their reasoning for choosing this proxy.

This information has been added.

Sect. 2.2, L230: "Sea salt concentrations were estimated by scaling measured soluble Na$^+$ concentrations by 3.27 following Manders et al. (2009) and were used as a proxy for SSA number concentrations. Size-resolved single particle chemical composition measurements have shown that sea salt represents 50-70% of SSA particles by number ($d_p > 0.5\mu$m) (Collins et al., 2014)**."**

L437-L439: The higher relative abundance of dust over sea salt concentration seems to be the essential part of the last two paragraphs. If the authors see a possibility, they could abridge their argumentation in L415-L443 to reach this statement.

This is a good point. To shorten this section and focus on this point, we have removed the 4-5 lines discussing agricultural soil dusts, as this is mentioned in what is now Sect. 4 Discussion.

L450:   What   is   meant   by   efficiencies?   Please   elaborate.

Updated to "IN activity" for consistency with the rest of the text:

Sect. 3.3, L585: "Figures 4(a) and (b) show overlap in n_{INP} and n_s observed in samples collected in low dust and high dust conditions, indicating that the INP populations observed during AQABA exhibited similar **IN activity** despite variation in total aerosol composition and dust loading."

L455: The exclusion of the sample by Price et al. 2018 seems a little bit arbitrary. Could the authors add the observed surface area of the excluded samples and the maximal observed surface area during the present study to relate the two observational ranges?

We have added a comparison of the maximum aerosol surface area as suggested.

Sect. 4, L590: **"Price et al. (2018) reported higher maximum aerosol surface area concentrations of ~1500 μm cm^{-3} from three samples collected in an exceptionally optically thick layer, compared to 965 μm cm^{-3} in the present study (Table 1).** Yet **overall**, the aerosol surface area concentrations compare very well, indicative of comparable dustiness in the two studies. Excluding **the case mentioned above,** the average aerosol surface area was 227 ± 68 μm^2 cm^{-3} vs. 226 ± 26 μm^2 cm^{-3} for the present study."

L459: Is there also a possibility for regional difference affecting the aerosol/INPs differently (Price et al. 2018 sampled at the western end of the Sahara)?

Yes, we agree. This thought was the motivation for the statement a few lines below this one, that region-specific parameterizations or an alternative to the n_s-based parameterization may be needed for T > -20 °C.

L492+: In the following, is sensitivity referred to as degradation? The authors may specify that.

This clarification was added.

Sect. 3.5, L640: "Sensitivity to peroxide in most samples (**i.e., INP degradation**) demonstrate the consistent presence of stable organic INPs at temperatures ≥ -15 °C."

L543-L546: Is there a plot supporting this argument?

Thank you for pointing out this omission. The reference to the figure has been added.

Sect. 4, L702: "Considering the high freezing temperatures observed, evidence of organic composition, and FLEXPART back trajectories showing that aerosol sources included populous regions and at least one agriculturally active region (the Nile River Delta**; Figs. S9-S10**)…"

L550-L560: This passage has more introductory character. The authors may consider moving it to the introduction.

Thank you for this suggestion. Since the reviewer commented on the length of the introduction, we have shortened this paragraph as follows:

Sect. 4, L702: "Considering the high freezing temperatures observed, evidence of organic composition, and FLEXPART back trajectories showing that aerosol sources included populous regions and at least one agriculturally active region (the Nile River Delta**; Figs. S9-S10**), it is possible that agricultural soil dusts contributed to some of the relatively higher n_s, n_{INP}, and heat and peroxide sensitivity observed during AQABA. A range of n_s has been reported in studies of agricultural soil dusts, the lower end of which agrees with the n_s observed in the present study between -8 and -25 °C (**Fig. 3;** Steinke et al., 2016; Tobo et al., 2014; O'Sullivan et al., 2014). Samples from air masses influenced by the Nile River Delta or Southern Europe (f007-8, f010, f038, f042, f044) show a higher fraction of heat-sensitive INPs (Fig. 4). Heat-sensitivity is indicative of biological INPs, which have been associated with agricultural soil dusts in prior studies (Hill et al., 2016; O'Sullivan et al., 2014). Hill et al. (2016) and O'Sullivan et al. (2014) showed peroxide sensitivity in agricultural soil dusts at temperatures > -18 to -15 °C, respectively, a range which aligns with the peroxide sensitivity exhibited in the present study."

L585: Cite the referred studies.

Sect. 4, 747: "In summary, it has proven difficult to determine any consistent impact of atmospheric processing on the IN activity of dust in model systems such as ATD (**Perkins et al., 2020**), and few studies have investigated impacts of aging on ambient desert dust, especially at modest supercooling (**Boose et al., 2016**)."

L626: Does Figure 5 refer to the soil sample used previously in the Niemand et al. 2012 study? What about the ambient aerosol observations during the research cruise? The authors may consider including the results presented in Figure 4 to the discussion here.

Updated, thank you.

Sect. 4, L826: The apparent decreased n_s observed in this study between -18 and -12 °C could potentially be related to a plateau in n_s through the transition between the mineral and organic "modes" (see **untreated samples in Figs. 5-**6)."

L644: Add a reference for the SSW samples collected at the Scripps Memorial Pier. If the data hast not been published yet, the authors should add a description of the sampling and analysis in the methods section

This information has been added to Methods Sect. 2.5

Sect. 2.5, L414: "INP concentrations in SSW collected at the Ellen Browning Scripps Memorial Pier at Scripps Institution of Oceanography (SIO; 32.8662 N, 117.2544 W) were assessed in 17 samples for comparison with SSW collected during AQABA. Samples were collected between 31 Jan and 7 May 2016 in 15-30 mL sterile centrifuge tubes (Falcon™, ThermoFisher Scientific, Waltham, Massachusetts, USA) at depths of 1-3 m and processed immediately using the SIO-AIS as described above."

L657: Why were not all samples processed? Please elaborate here or in the methods section.

For the seawater samples, there were access and researcher availability constraints due to pandemic-related disruptions. Fortunately, most aerosol samples had been processed prior to

March 2020, but some of the seawater samples remained. These details are probably not necessary to include in the text.

L661-L663: The reviewer is not fully sure what this result adds to the discussion. The authors should consider elaborating on this and highlight how the observation in an individual sample relates to a bigger implication.

This line has been removed from the text.

L670-671: The authors may consider adding the total number of observed samples.

This detail has been added.

Sect. 5, L875: INP concentrations measured in 26 aerosol samples spanned two or more orders of magnitude (0.002 to 0.5 L^{-1} at -15 °C).

L678: The authors should add a quantitative statement on what is meant by "showed agreement" for the different data they refer to.

This line has been updated as suggested.

Sect. 5, L884: "**Observed n_s for some samples was equivalent to that of** the A13 parameterization for K-feldspar (Atkinson et al., 2013), an ice-active component of desert dust, observations within the marine boundary layer (DeMott et al., 2016; Yang et al., 2020), and the Price et al.'s (2018) measurements of INP concentrations in dust-laden air masses over the Tropical Atlantic."

L685: The authors may consider stating explicitly the temperature ranges they refer to in the conclusions.

This line has been updated as suggested.

Sect. 5, L890: Heat-sensitivity further suggested the presence of biological (e.g., proteinaceous) INPs in a subset of samples, particularly at high freezing temperatures **> -10 °C**.

L685-L689: To the reviewer this statement has rather introductory than concluding character. The authors may consider either moving the sentence to the introduction or shortening the statement and directly combining it with observation of the study.

This line has been removed, as these points were made in the discussion.

L689-L691: The authors may add that the analyzed sample originates from the analyzed samples in Niemand et al. 2012.

This detail has been added.

Sect. 5, L896: "A soil dust sample from North Africa **(originally from N12)** exhibited heat and peroxide sensitivity between -5 and -16 °C, further demonstrating that the IN activity of mineral dust could be associated with organic and/or biological material."

L698: Also here, what is considered a well agreement? Agreement within an order of magnitude? Please support with a quantitative statement.

The line has been updated as suggested.

Sect. 5, L906: "The observed n_{INP} **for SSW samples were equivalent to those of** Gong et al. (2020) at Cabo Verde within the 95% binomial sampling confidence intervals (Agresti and Coull, 1998)."

L700-L707: In the reviewer's opinion, the possibility of sample degradation and the potential impact of sampling close to the ground (as discussed in the results) should be added.

These points have been added.

Sect. 5, L915: "Vertical profiles of n_s in dust-laden air masses are also needed to determine whether n_s is consistently lower at the surface and the variability of n_s with altitude. Potential storage impacts on INPs collected on filters are an additional factor worthy of future investigation, though storage alone does not likely explain the relatively decreased n_s compared to parameterizations observed in this study, as U17 and N12 were both derived from stored dust samples."

L709: The authors may want to add a brief explanation of the working principle/data basis of the methods by Gong et al. 2020 to the sentence.

We have added a few words to describe the principle behind this method:

Sect. 5, L922: "In addition to providing observations at high to moderate freezing temperatures, future studies could apply the methods developed in Gong et al. (2020) to estimate the contribution of marine INPs to the aerosol sampled by assuming equivalent distributions of sea salt and INPs between seawater and air."

Editorial notes:

L37: The authors may consider changing "reflectivity, and precipitation efficiency" to "radiative properties, and precipitation initiation efficiency", to avoid using reflectivity, which is often connotated to radar observations, and be more specific about the role of INPs towards precipitation. This comment also applies to other occurrences in the manuscript.

This line has been rephrased as suggested. Other instances of "reflectivity" have also been updated to "radiative properties."

Abstract, L36: "…lifetime, radiative properties and precipitation initiation efficiency…"

One instance of "reflectivity" remains in the manuscript because we want to be consistent with the reference.

Sect. 1, L99: "Results from a follow-on cloud-resolving model study showed that Southern Ocean cloud reflectivity is strongly modulated by INP concentrations…"

L47: Increase consistency: sometimes e.g., "1-3 orders of magnitude" is used, sometimes "one to three orders of magnitude".

The phrasing is now consistent.

L50: Despite standard in our community, the authors may want to specify the usage of "hydrogen peroxide". In the manuscript, different versions (peroxide, hydrogen peroxide, H2O2) – the authors may use one version for consistency.

Thank you for bringing this to our attention. This line and all other instances have been updated to peroxide, which is defined as H2O2 on L276:

Sect. 2.4, L276: Prior to sampling, filters were cleaned by soaking in 10 % **peroxide** (H_2O_2) for 10 min.

L51: Consistency: here ≥ is used, whereas in L56 and L58 >. Abstract, L60 has been updated for consistency. The other line was removed.

Abstract, L60: "Future efforts to develop or improve representations of dust INPs at modest supercooling (≥-15 °C) would benefit from a characterization of the specific organic species associated with dust INPs."

L57: Consistency: modest (three instances) or moderate?

We use modest with respect to supercooling for consistency with prior publications (e.g., (DeMott et al., 2018; Kanji et al., 2017). There are no instances "moderate supercooling" that I can find, at least after revisions.

L73: The long-term INP observations of Brunner et al. 2021 may be an additional suitable reference for this statement

Thank you for this reference! Added.

Sect. 1, L70: "As initiators of ice formation and related phase-partitioning processes, INPs affect multiple cloud properties and exert a strong influence on cloud lifetime, radiative properties and precipitation initiation efficiency (e.g. Lohmann and Feichter, 2005; Vergara-Temprado et al., 2018; Brunner et al., 2021)."

L76: Consider adding "average" to "global dust loading".

Added.

Sect. 1, L80: "Of the average global dust loading in the atmosphere (20-29 Tg), …"

L79: emissions → emission.
Corrected.

Sect. 1, L83: "Analysis of satellite products indicates that dust emission rates…"

L89: Consider adding "by number" or "by mass" after "dominant".
Added.

Sect. 1, L92: "Though mineral dust is considered to be the dominant INP source in many regions, multiple modeling and observational studies suggest that marine INPs are frequently dominant by number in remote ocean regions in air masses with low concentrations of terrestrial aerosol (McCluskey et al., 2018b, 2018c; Vergara-Temprado et al., 2017; Wilson et al., 2015; DeMott et al., 2016)."

L111: Consider introducing A13 here as well.
We have added an introduction.

Sect. 1, L113: "There are, additionally, multiple mineral-specific INP parameterizations including illite (Broadley et al., 2012), kaolinite (Welti et al., 2012), quartz (Harrison et al., 2019) and K-feldspar (Atkinson et al., 2013, hereafter, "A13")."

L112+L114: Add 'hereafter,' before introducing the abbreviations of N12 and D15.
Added.

Sect. 1, L113: "There are, additionally, multiple mineral-specific INP parameterizations including illite (Broadley et al., 2012), kaolinite (Welti et al., 2012), quartz (Harrison et al., 2019) and K-feldspar (Atkinson et al., 2013; hereafter, "A13"). The parameterizations by Ullrich et al. (2017; hereafter, "U17") and Niemand et al. (2012; hereafter, "N12") were developed using dust samples from multiple deserts, and both found little variability in the IN activity between dusts from locations as disparate as the Sahara and Asia. DeMott et al. (2015; hereafter, "D15") found agreement between their observations-based parameterization and N12, supporting the validity of laboratory-based parameterizations."

L121: Consider adding "in situ" before "INP measurements".
Added.

Sect. 1, L126: "By contrast, few studies report *in situ* INP measurements near…"

L126: Consider adding "west" before "of the Sahara".
Added.

Sect. 1, L130: "Boose et al. (2016) found that D15 overpredicted INPs observed during Saharan dust events at a location within 100s of km west of the Sahara (Izaña, Tenerife, Spain)…"

L144: "SSA" not defined yet.
Definition added, thank you.

Sect. 1, L146: "There are now two parameterizations available for the estimation of atmospheric concentrations of marine INPs emitted from the ocean surface: Wilson et al. (2015), which estimates cumulative INPs from total organic carbon (TOC) concentrations in simulated sea spray aerosol (SSA),…"

L149-L159: If further subsections are added to the manuscript, the different sections could be added here after the individual analysis steps to allow easy finding of specific parts of the manuscript.

Added, thank you for this suggestion.

Sect. 1, L160: "The rest of this study will be structured as follows. We present an overview of measurements and data sources in Sect. 2 Methods. In Results Sect. 3.1, an overview of INP concentrations observed is presented, followed by an assessment of subsurface seawater (SSW) source potential (Sect. 3.2). Observed n_s are compared to dust and marine INP parameterizations in Sect. 3.3, followed by an analysis of thecontributions of heat-labile (e.g., proteinaceous) and heat-stable organic compounds to observed INP populations in aerosol (Sect. 3.4). The same analysis is applied to a soil dust sample from a likely source region in Sect. 3.5. We discuss the findings, potential INP sources and compare with prior studies in the Discussion Sect. 5. Finally, in Sect. 5 we offer strategies to address the challenges of evaluating dust-specific INP parameterizations and recommend measurements needed to develop predictive understanding of dust INPs at modest supercooling (T ≥ 15°C)."

L152: Consider adding "ambient" before "aerosol sample".

Added.

Sect. 1, L156: "INP concentrations were measured in 26 ambient aerosol samples collected during…"

L156: Leave out " ' "

Removed.

L190: dopt, dgeo, and PM1 are not defined yet.

These definitions have been added.

Sect. 2.2, L217: "Fine particle (d_p < 700 nm) size **was** converted from **optical diameter (d_{opt})** into **geometric diameter (d_{geo})** using the optical properties calculated from the chemical composition **of particles < 1 µm** (PM$_1$) as measured by an Aerosol Mass Spectrometer (Aerodyne HR-ToF-AMS),…."

L198: Consider adding "geometric diameter" after "µm". Added.

Sect. 2.2, L224: "Total particle surface concentrations were determined by integrating the surface area distribution for particles up to 10 µm (d_{geo})."

L254: At different instances "-fold", "x", or "-times" are used. Consider using one version throughout the manuscript.

The text has been checked for consistency, and all instances of these terms have been replaced with "×" if necessary.

L256: SIO not defined.

The definition has been added.

Sect. 2.4, L303: "INP concentrations were measured using the Scripps Institution of Oceanography Automated Ice Spectrometer (SIO-AIS), …"

L265: In the reviewer's perception, the 'INP' subscript should not be in italics, i.e. nINP. This applies throughout the manuscript, including the figures. Further, the authors should consistently use nINP or INP concentration in the text.

The italics around INP have been removed for all instances in the text and figures. The text has been checked and updated for consistency on this term using "n_{INP}".

L267: Consider adding "drop" before "volume"

This clarification has been added, but in this case we are referring to the volume of ultrapure water used to resuspend the aerosol from the filter rather than the droplet volume.

Sect. 2.4, L315: "For aerosol filter samples, cumulative INP number concentrations are calculated using the ratio of the ultrapure water volume used for resuspension of the particles…."

L273-L290: Is nINP in bold font intended?

This was unintentional. The bold font formatting has been removed.

L282: When comparing to Figure S6 the concentrations seem different at -20 °C.

Thank you for catching this! This line has been corrected.

Sect. 2.4, L330: "Figure S7 shows the estimated n_{INP} across the 7 field blanks, which ranged between 1.0×10^{-4} and 7.0×10^{-3} L^{-1} at -20 °C."

L285: The linear regression seems to have been done on the log-transformed data. If so, please specify.

The regression was based on the geometric mean n_{INP} at each freezing temperature. The following clarification has been added.

Sect. 2.4, L332: "To correct n_{INP} measured in aerosol samples for background INPs from sample handling, a linear regression **based on the geometric mean** n_{INP} measured in field blank suspensions (mL^{-1} water) was used to estimate background concentrations of INPs in samples at all temperatures between -14.5 °C and -27 °C."

L314: °K → K

The degree symbol was removed here.

L328-L331: Maybe add that the results of each test are indicated in the corresponding figures.

Some thought was given to mentioning this, but it could also be distracting to the reader to get a reference to the figures at this point of a Methods section.

L362: The referred abbreviation PM10 does not appear in Table 1.

The PM10 clarification has been added to relevant variables in Table 1.

Table 1: Remove "unt" from sample ID to stay consistent with text. Consider adding (°N) and (°E) to latitude and longitude. At what sampling time are the presented latitudes and longitudes (start or stop datetime?). To enrich the information of the table, the authors may consider adding the INP concentration at an exemplary temperature.

The Table has been edited as suggested. The column headers are now looking a bit "squished" to fit the extra n_{INP} column in, but I think this will be adjusted during type-setting.

Table 1 caption: "Table 1. Summary of aerosol samples collected during AQABA. "—" indicates where data are missing; "NaN" indicates values below detection limit. Locations are given at the transect midpoint during each sampling period."

L392: Has M18 already been introduced?
Good point. The definition of M18 was added to the first reference of the parameterization in the introduction, for consistency with the other parameterizations.

Sect. 3.3, L519: "Overall, observations nearly bridge the full regime between the **M18** parameterization for marine INPs (McCluskey et al., 2018c), and multiple dust INP parameterizations..."

L396: Did the authors mean to refer to Yang et al. 2020? Corrected.

Sect. 3.3, L521: "At higher temperatures, between -5 and -12 °C, most observations show agreement with the composite spectrum of n_s observed in a range of marine and coastal environments from DeMott et al. (2016) and **Yang et al. (2020),** and/or the Atkinson et al. (2013) K-feldspar parameterization."

L397+L398+L401: The authors may consider adding the sample IDs of the mentioned samples/spectra

This level of detail could be too much for the main text because this would require listing each of the 26 samples, but of course the reader could refer to the data package if they wished to know how each sample compared to the parameterizations individually.

L403: It may not be entirely clear to which study "same region" refers to

Clarification was added.

Sect. 3.3, L528: "Multiple samples (~8) additionally agreed with Price et al.'s (2018) observations of INPs between 30-3500 m above the dusty Tropical Atlantic, and most agree with the Gong et al. (2020) surface-level observations, **measured at Cabo Verde** in the same region **as Price et al. (2018)**. "

L433: DMSO2 and VOCs remain undefined

Definitions added.

Sect. 3.3, L562:" Edtbauer et al. (2020) reported the detection of high levels of dimethyl sulfide (DMS, up to 800 ppt) in the Gulf of Aden associated with a local phytoplankton bloom during AQABA (as evidenced by visible bioluminescence around the ship at night) as well as high levels of **dimethyl sulfone** (DMSO$_2$) and other marine biogenic **volatile organic carbons** (VOCs) from the Somalian upwelling region."

L452: Figure S7 should be number S9.

The figures have now been renumbered throughout most of the text.

L467: DeMott et al. 2015a → DeMott et al. 2015

Corrected.

Sect. 4, L801: "Whereas DeMott et al. (2015) found that…"

L533+L595: Consistency: change "dusty" to "dust-laden".

Updated.

Sect. 3.5, L689: "Gong et al.'s (2020) results showing heat-sensitivity in INPs at temperatures > -10 °C further demonstrate the contribution of biological INPs at high temperatures in dust-laden air masses near N. Africa."

Sect. 4, L775**: "**A prior study that compared n_{INP} in dust-**laden** air masses at the surface with n_{INP} collected between 0.5 and 3 km above sea level found that…"

L534: "N." → "North"

Corrected.
Sect 3.5, L690: "…further demonstrate the contribution of biological INPs at high temperatures in dust-laden air masses near North Africa."

L565+L566: Consider rephrasing the citation text ("e.g.") as it is not directly evident what statement the citation supports.

This reference is probably unnecessary in this line and has been removed. It was meant to be an example of the measurements that would have been helpful for understanding the impacts of aging on observed INPs but is used for this purpose elsewhere in the text.

L569: The authors consider adding exemplary references for the mixed and contradictory results already here.

Since we don't go into detail on the "mixed and contradictory results" in this manuscript, we can simply refer to Perkins et al. (2020) and the references therein, which contains a nice summary of results on aging.

Sect. 4, L730: "In addition to field observations of $\boldsymbol{n_{INP}}$ demonstrating that aging increased the IN efficiency of desert dust INPs (see Introduction; Boose et al., 2016; Conen et al. 2015), prior studies of the effects of aging on mineral dust INPs have yielded mixed and sometimes contradictory results, indicating that the impact of aging on IN properties likely depends on multiple factors including the ice nucleation pathway, the type of aging process, surface morphology, and mineralogy (Perkins et al., 2020 **and references therein**)."

L581: Both ">" and "below" is used. For consistency the authors may want to use </> or below/above consistently throughout the manuscript.

The text has been checked and updated for consistency.

Sect. 4, L739: "Perkins et al. (2020) additionally reported that the lability of IN activity in ATD is temperature dependent, with large reductions evident at freezing temperature > 10 °C, yet little to no change at temperatures < -15 °C."

L611: Remove second "x".

Removed.

L612:   Remove   whitespace.

Removed.

L616:   Missing   unit   after   "(±65,000)".

The unit was added.

Sect. 4, L815: "Furthermore, a recent study that measured airborne concentrations of prokaryotic cells over the Red Sea characterized the region as a "global hot spot" with average concentrations of 155,000 (± 65,000) cells m^{-3}..."

L620: For Kanji et al. 2017 being a summary, consider putting it last in line and add "and references therein".

Edited as suggested.

Sect. 4, L820: "Organic and biological species have been shown to dominate IN activity at temperatures > ~-15 °C in many studies (e.g., Ladino et al., 2019; O'Sullivan et al., 2018, Kanji et al., 2017 **and references therein**)."

L629: What are the temperature ranges considered "high" or "moderate"? The authors should consider adding the temperature ranges here, and/or explicitly introduce the referred temperature ranges earlier in the manuscript.

The temperature range has been added. "High" has been defined as > -10 °C elsewhere in the text.

Sect. 4, L829: "This study underscores the need to characterize the IN-active organic species associated with dust from major source regions and to investigate the extent to which biological and/or organic particles contribute to INP populations in dust-laden air masses at high to moderate freezing temperatures ≥ **-15 °C**."

L660:   "POV"   was   already   introduced.
The definition was removed.

Sect. 3.2, L503: "These results are indicative of the **POC** type of marine INP defined in McCluskey et al. (2018a)**…**"

L670: Consider adding "Observed" before "INP concentration".
Added.

Sect. 5, L875: "Observed n_{INP} measured in 26 aerosol samples spanned 2 orders of magnitude (5×10^{-3} to 5×10^{-1} L^{-1} at -15 °C)."

L675: The authors may consider adding that the PM10 observations were obtained from reanalysis

Added.

Sect. 5, L880: "Despite proximity to major dust sources and a high frequency of dust events with **MERRA-2 simulated** mass concentrations up to 490 µg m^{-3} (PM$_{10}$)…"

Editorial notes on the figures

Generally, many of the figures (e.g., Figure 3, 4, and 6) feature strange gray lines not belonging to the plot itself. Further the indication of units is inconsistent, e.g., ""°C" after "Temperature" is never set in brackets. Legends that apply to different subplots could be put aside of the subplots and not within one subplot. The reviewer kindly asks the authors to check all their figures (including supplement) to enhance their quality.

Thank you for pointing out these issues. We have checked and corrected all figures in the text and supplement and have also made the updates suggested.

Figure 1: Datapoints at 25°N/35°E are overlaying and barely visible. Change order or add some alpha to scatters. "µm" in colorbar.

We acknowledge that the points at this location in Fig. 1 are overlaying. Unfortunately, we cannot change the alpha without losing the coloring for dust concentration or marker size without losing the INP concentration information. This information is important for the overview figure. However, we do have multiple figures in the supplement in which all the data points are visible.

We have added a reference to some of these figures in the Fig. 1 Caption.

Figure 1 caption: "Map of the sample locations for 26 aerosol samples collected on the RV *Kommandor Iona* during Air Quality and climate change in the Arabian BAsin (AQABA; see also Figs. S1-S2)."

Figure 2: The readability of the plot may be increased when using generally open symbols in combination with some alpha.

As the reviewer may be able to see in the updated figure, the markers for the AQABA measurements have a low alpha of ~0.2. Another reviewer has asked the error bars be added to the AQABA markers, in which case we must keep the marker outlines, or at least plotting open

symbols with error bars is not a functionality that is available in the MATLAB version I am using to generate the figures.

Figure 3: Mention the errorbars in caption. Maybe mark sample f020 in plot (a) as reference. Consider stretching the y-axis limits in (b)

This detail has been added.

Figure 4 caption: "Error bars represent 95% binomial sampling confidence intervals (Agresti and Coull, 1998)."

Figure 4: The labels of the subfigures need to be increased. "INPs" to "INP concentrations" in caption.

The label sizes have been increased.

Figure 6: Shouldn't the y-label read nINP,L? "INP spectrum" to "INP spectra". Also name the number of samples from the other studies in the caption.

This is INP per mL seawater. INP measurements in seawater are often reported in terms of n_{INP} mL^{-1} (e.g., McCluskey et al. (2018a).

This information has been added to the Figure 2 caption.

Figure 2 caption: "Also shown are the composite INP spectrum of **14** coastal SSW samples collected on São Vincente Island, Cabo Verde (Gong et al., 2020), **17** coastal SSW samples collected at the Ellen Browning Scripps Pier (green shading), and **12** SSW samples collected in the Southern Ocean (McCluskey et al., 2018a)."

Figure S1: It seems that f006-f009 and f018+f019 overlap. Were several sampling units running at the same time? If so, this should be mentioned in the methods section.

No. During the first leg, the vessel would turn around frequently and spend the day heading upwind to increase sampling without stack contamination. Sometimes, this resulted in sampling roughly the same transect twice.

Figures S1 and S2 have been updated for better visibility of the overlapping transects.

Figure S3: The authors may consider reducing the dot size for readability

The Figure S4 dot size has been reduced.

Figure S5: In consistency to Figure 5, the dilutions could be indicated with a different marker than undiluted samples.

The markers have been updated for consistency with Fig. 6

Figure S7: The labels are rather small and could be increased in size. What correlation coefficient was used? To the reviewer's perception Spearman's correlation coefficient should be used. The reviewer is not fully sure about the expressiveness of the figure. Maybe a table with Spearman correlation coefficients between the INP concentration and both surface area and dust concentration at all temperatures could help for a more complete picture.

The labels have been increased. The correlation coefficients have been updated to Spearman's as suggested. A table is probably unnecessary. It is clear from the figure that there is no correlation between aerosol surface area and n_{INP}.

Figures S8+S9: The authors may consider indicating only locations of the trajectories below a certain height threshold.

Figures S9 and S10 have been updated as suggested. Some details have been added to the text about the height thresholds used.

Sect. 3.1, L446: "Figs. S9-S10 show the extent of k-means clustered FLEXPART back-trajectories below the altitude of 1500 m **(see Sect. S5 for details)**. This threshold was applied to eliminate most of the free tropospheric parts of the back-trajectories and was selected based on the MERRA-2 monthly average planetary boundary layer (PBL) heights during the campaign period, which were 200-700 m over the ocean and up to 1700 m over land."

Figure S9: To add information, the cross of the missing sample could be colored according to the observed dust concentration.

The X was updated as suggested.

Figure S11: There seems to be data gaps, which are indicated in black as is the land contour. Color either the land or data gaps in a different color. Remove gray lines around colorbar. Add shown quantity to the colorbar.

The updates were made to the color bar as suggested. The figure of Chl *a* was produced with the Giovanni online data system, and it would take some extra time to reproduce the figure myself to change the color of the land. If the reviewer prefers, this figure could be taken out of the supplement as it is not part of a central argument. It is a nice overview of a seawater property that has some relevance to INP measurements and seawater source potential.

Figure   S12:   Shouldn't   the   y-label   read   "$n_{INP,L}$ (mL-1)"?
Thank you for catching this. The y-label has been corrected.

Other changes:

A line was added to better connect the point about elevation to the discussion.

Sect. 4, L779: "The differences between Price et al. (2018) and the two surface-based studies draws attention to the need for vertical profiles of n_s > -25 °C in dust-laden air masses."

A member of the INP community who saw the pre-print kindly sent us a reference reporting INP measurements from the Red Sea and Indian Ocean made in 1979. We included the reference in Sect. 3.1

Sect. 3.1, L435: "This range agrees within an order of magnitude with that of Prodi et al. (1983) who measured n_{INP} in the Mediterranean, Red Sea, Gulf of Aden and Indian Ocean nearly 4 decades prior to the present study (4×10^{-2} to 2 **L^{-1} at -16 °C**)."

References

Beall, Charlotte M, Lucero, Dolan, Hill, Thomas C, DeMott, Paul J, Stokes, M Dale, and Prather, Kimberly A: Best practices for precipitation sample storage for offline studies of ice nucleation in marine and coastal environments, Atmos. Meas. Tech., 13, 6473–6486, https://doi.org/10.5194/amt-13-6473-2020, 2020.

DeMott, Paul J., Möhler, Ottmar, Cziczo, Daniel J., Hiranuma, Naruki, Petters, Markus D., Petters, Sarah S, Belosi, Franco, Bingemer, Heinz G., Brooks, Sarah D., Budke, Carsten, Burkert-Kohn, Monika, Collier, Kristen N, Danielczok, Anja, Eppers, Oliver, Felgitsch, Laura, Garimella, Sarvesh, Grothe, Hinrich, Herenz, Paul, Hill, Thomas C J, Höhler, Kristina, Kanji, Zamin A, Kiselev, Alexei, Koop, Thomas, Kristensen, Thomas B, Krüger, Konstantin, Kulkarni, Gourihar, Levin, Ezra J T, Murray, Benjamin J., Nicosia, Alessia, O'Sullivan, Daniel, Peckhaus, Andreas, Polen, Michael J, Price, Hannah C, Reicher, Naama, Rothenberg, Daniel A, Rudich, Yinon, Santachiara, Gianni, Schiebel, Thea, Schrod, Jann, Seifried, Teresa M, Stratmann, Frank, Sullivan, Ryan C., Suski, Kaitlyn J, Szakáll, Miklós, Taylor, Hans P, Ullrich, Romy, Vergara-Temprado, Jesus, Wagner, Robert, Whale, Thomas F., Weber, Daniel, Welti, André, Wilson, Theodore W, Wolf, Martin J, and Zenker, Jake: The Fifth International Workshop on Ice Nucleation phase 2 (FIN-02): laboratory intercomparison of ice nucleation measurements, Atmos. Meas. Tech., 11, 6231–6257, https://doi.org/10.5194/amt-11-6231-2018, 2018.

McCluskey, Christina S, Hill, Thomas C J, Malfatti, Francesca, Sultana, Camille M, Lee, Christopher, Santander, Mitchell V, Beall, Charlotte M, Moore, Kathryn A, Cornwell, Gavin C, Collins, Douglas B, Prather, Kimberly A, Jayarathne, Thilina, Stone, Elizabeth A, Azam, Farooq, Kreidenweis, Sonia M, and DeMott, Paul J: A Dynamic Link between Ice Nucleating Particles Released in Nascent Sea Spray Aerosol and Oceanic Biological Activity during Two Mesocosm Experiments, J. Atmos. Sci., 74, 151–166, https://doi.org/10.1175/JAS-D-16-0087.1, 2017.

Niemand, Monika, Möhler, Ottmar, Vogel, Bernhard, Vogel, Heike, Hoose, Corinna, Connolly, Paul, Klein, Holger, Bingemer, Heinz, Demott, Paul, Skrotzki, Julian, and Leisner, Thomas: A particle-surface-area-based parameterization of immersion freezing on desert dust particles, J. Atmos. Sci., 69, 3077–3092, https://doi.org/10.1175/JAS-D-11-0249.1, 2012.

Stopelli, E., Conen, F., Zimmermann, L., Alewell, C., and Morris, C. E.: Freezing nucleation

apparatus puts new slant on study of biological ice nucleators in precipitation, Atmos. Meas. Tech., 7, 129–134, https://doi.org/10.5194/amt-7-129-2014, 2014.