RC1 review of "The Fifth International Workshop on Ice Nucleation Phase 3 (FIN-03): Field Intercomparison of Ice Nucleation Measurements," by P. J. DeMott et al. 2024

RC1: This manuscript is well written. The authors conducted a fair intercomparison of online and offline INP-measuring instruments in the field. Despite the challenging environment at SPL, invaluable outcomes and lessons are reported in a neutral and unbiased manner. Furthermore, the authors include a list of limitations (e.g., deviation in sampling particle sizes etc.) and things to be further explored in this manuscript for more understanding of aerosol-cloud interactions (e.g., a need for online/direct deposition ice nucleation measurements), which are important messages to the INP research community. This reviewer agrees that more research is necessary to predict and explain the temporal variation of biological INPs (perhaps in a predominantly biogenic environment). While the authors found a predominant contribution of mineral and/or other inorganic particles to INP abundance in this study, they also note the need for in situ mixed INP detection and characterizations, especially for Soil & BBA INPs, which is important. The study topic is relevant to the journal scope of AMT. This reviewer supports the publication of this paper in AMT after the authors address several questions below.
We thank the reviewer for this overall positive assessment of the paper.

Questions

[1] Figure 7: This reviewer wonders if using 3-hr INP median or log-average changes any conclusions of this intercomparison study. The ratio in Fig. 7 is computed by using time averages, which is reasonable. But, since the reported N_{INP} spans a log range at a majority of freezing temperatures examined in this study, the average can be biased by high N_{INP} values at the given temperature, such as the ones from FRIDGE-CS and CSU-IS. Perhaps, using the median may overall result in better agreement for NC State(F), NC State(I), and CSU-CFDC? The same average vs. median argument applies Figs. 8 & 9.

We thank the reviewer for this comment and these suggestions. They brought to the fore that we have not properly described the somewhat varied nature of data in Figure 7, which we do now in the revised manuscript. Primarily, it was erroneous to say that all data were averages. In fact, most data are single temperature measurements during cooling rate scans conducted on suspensions from single few-hour filters in the case of standard immersion freezing devices, and only in the case of the CFDC data are multiple point measurements averaged. Hence, converting to median values is not possible. Former Figure 7 (now 8) has been revised and other figures have been added or revised based on comments to Dr. Gabor Vali's review.

Figures 8 and 9 represent instrument averages and so are amenable to a median analysis. However, the number of overlapping periods of more than 2-3 instruments is minor and in analysis of medians for these cases, they differed from means by less than 20%. Hence, we retain the analysis as shown in the present figures.

[2] Figure 5: This reviewer wonders why NC State CS(F) shows a lower detection of N_{INP} (~6 x 10^{-3} L^{-1}) than NC State CS(I) (~10^{-1} L^{-1}). The sampling air flow rate seems similar for these two methods as described in Sect. 2.2.2. The sampling interval was shorter for impinger sampling?

Or it may be due to the difference in collected particle sizes (L836-839; L846-848; L855-858)? This reviewer is aware of a general statement in L865-870.

We thank the reviewer for helping us to clarify what is already apparent in the noted figure, that is, the detection limit is different for the two NC State measurement methods, and this is a consequence of differences in the liquid and air volumes used. To make this clear before results are shown, we write at the end of the Section 2.2.2 subsection on the NC State methods: "Note that due to the greater V_{liquid} *used in the impinger for the stated air collection volumes means that the filter samples were more concentrated by a factor ≈11. Thus, the filter technique is more sensitive and has a lower limit of detection (LOD). The precise ratio for any given experiment depends on the exact sampling times of filter and impinger, and the exact number of droplets for the filter, impinger sampling, averaging across repeats, and binning into 1-degree intervals. For this reason, the ratio of LOD for the averaged samples may differ slightly from this estimate"*

[3] P31L649-655: Low AE (<1) seen in 9/14-16 in Fig. 3b may be due to the predominance of large dust seen in Fig. 4? The authors also report that the submicron particles dominated during the study period (L-637-638). The effective aerosol scattering efficiency from SPL during this intercomparison campaign can be similar to what is reported in Testa et al. (2021)?

Refs.

Russell. P. B. et al., ACP, 10, 1155-1169, https://doi.org/10.5194/acp-10-1155-2010, 2010.

Testa, B. et al. JGR-A, 126, e2021JD035186. https://doi.org/10.1029/2021JD035186, 2021.

We believe that the reviewer meant Fig. 2c instead of 3b, and Fig. 3 for Fig. 4. While a general preponderance of dust was possible during the period noted, dust seen in Fig. 4 on the 15th is from a very short period of PALMS operation and likely represents an anomaly due to generation of road dust near the site at that time, as discussed already in Section 3.2.2. There was scarce data from mass spectrometry during that period and so we choose not to emphasize it. While scattering efficiency might be like Testa et al. (2021), we do not wish to emphasize derivation of surface area from nephelometer data in this publication, as was done in Testa et al.

[4] P49L971-973: Can the authors clarify this part?

Yes, the intended meaning was that it would be unrealistic to believe that all particles in the size regime larger than 500 nm are soil-sourced only. The intention of the discussion is to say, what if they somehow were? We have revised this to read, "In this case, a somewhat unrealistic maximum assumption on soil dust numbers and surface area that considers all particles and compositions in this size range as emanating from dust, Niemand 2012 estimates a dust source for 50% and DeMott 2015 estimates 25% of observed INPs on average."

Comments

P37L749-750: This is good. Comparability of impinger and filter-based methods shown in this work implies that ambient particles collected on filters are well-scrubbed in liquid suspension for

freezing tests on NC State CS, resulting in comparable N_{INP} to that from directly suspended impinger samples, for this field study at least.

We add to reflect this point by expanding the sentence to say, "…, suggesting that particle removal from filters can be highly effective for immersion freezing measurements of ambient particles."

P44L885-887: This recaps that the link between aerosol chemical composition and INP is not straightforward and underscores the importance of ice residual composition data.

Yes, although ice residual composition data is difficult to obtain due to the low INP concentrations one attempts to assess via that method, and the low efficiency of doing this by mass spectrometry, as discussed in limited publications on this topic since 2004. There can also be pitfalls for identifying particle types via SEM and TEM. It is hard work, though we agree that it must continue. We feel that this paper is not the venue for emphasizing this point though.

P64L1249-1255: This reviewer agrees. The ultimate future INP instrument intercomparison may be performed on the aircraft platform in cirrus and/or pyrocumulonimbus cloud regimes with collocated aerosol instruments suggested by Burrows et al. onboard then.

We appreciate the point the reviewer is making also but will only comment in this response. Aircraft campaigns are notorious for not providing enough signal to noise in comparison to ground based efforts. This is true for both INPs (typically lower, except perhaps in a pyrocu, though few pilots will fly into them) and with compositional measurements. Hence, while we agree that such intercomparisons would be ideal, they may be a work in progress over many years.