We thank the reviewers for their thoughtful feedback, which has helped us to further improve and strengthen the manuscript. Our detailed responses to the final reviewer comments are provided below in blue. All revisions are visible in the track-changes version of the revised manuscript and shown below in green.

We note that, during the review process, we conducted an intercomparison between the original PUFIN system, a newly built duplicate unit, and our standard ARM ground-based INP filters as presented in Creamean et al. (2025). All filters were collected outside our building at CSU for 30 minutes. We added a figure showing the resulting spectra in the Supporting Information and included text describing this intercomparison in Section 4.3: Two identical PUFIN units are currently available, including the original system and a newly constructed duplicate. An intercomparison test was conducted using standard ARM ground-based INP filters (Creamean et al., 2025), in which ambient samples were collected outside CSU for 30 minutes, using the first sample valves on both PUFIN systems. The resulting INP spectra show good agreement between the two units, and the ground-based INP filters deployed routinely at ARM sites, demonstrating reproducibility and consistency in sampling and analysis (see Figure S5).

RC1

The manuscript presents a newly developed payload, "PUFIN," for vertically resolved sampling of ice-nucleating particles (INPs) using tethered balloon systems. The PUFIN system samples aerosol particles with a powerful scroll pump onto polycarbonate filters in three separate filter holders, each selected for sampling at a different altitude, using remote-controlled magnetic valves. INP abundance and temperature spectra are derived from offline filter analysis using an INP detection system (here, the Colorado State University INS). This is a very valuable approach to extending INP sampling into the vertical with tethered balloon systems, particularly because multiple filters allow contrasting altitudes, such as the free troposphere from the boundary layer. The paper is very well written and provides a sound description of the PUFIN system, its deployments across the US, and the INP analysis of the sampled filters. The INP community may also benefit from the author's open-source approach, which provides technical drawings and details about PUFIN in a public repository, as well as freely available INP data from their previous balloon deployments. I recommend accepting the manuscript for publication in AMT.

We thank the reviewer for their overall assessment and appreciate their perspective on how this work will benefit the INP community!

Some recommendations for improvements are listed below:

1. The manuscript could benefit from showing meteorological data from the iMet-XQ2, particularly given that the INP samples exhibit varying abundance at different altitudes, rather than simply assuming atmospheric layering as the sole explanation (e.g., lines 303-305, 320). This could further emphasize the need for vertically resolved INP sampling.

This is a good point. Unfortunately, for the February flights, the iMet-XQ2 data were not recorded in the INP flight data files, although they are available for the July flights. In addition, the accuracy of the iMet-XQ2 measurements is limited (see below), so we instead use the more reliable "TBSIMET" data to examine

equivalent potential temperature, calculated following Bolton (1980). This parameter is well suited for assessing atmospheric stratification and boundary layer mixing.

We added two new figures to the Supporting Information and included the following text at the beginning of Section 4.2: We also evaluated equivalent potential temperature (θ_e) during flight days using data collected from the TBS to assess whether the boundary layer in the PUFIN sampling region was well mixed, where INP concentrations would be expected to be similar across altitudes, or stratified, which can promote aerosol layering (Creamean et al., 2021; Griesche et al., 2021, 2025). Details of the θ_e calculations are provided in the Supporting Information.

Throughout Section 4.2, we added supporting evidence showing that, on days with two spectra that differed significantly, these differences were often associated with a stratified boundary layer based on θ_e profiles. Please refer to the tracked changes version of the manuscript below to see where these additions were made, as they are distributed throughout the section.

In the Supporting Information, we added the following methods:

Because the iMet-XQ2 accuracy is limited (see Summary section), we use the more reliable TBSIMET observations to examine equivalent potential temperature (θe), calculated following Bolton (1980), which is well suited for assessing atmospheric stratification and boundary-layer mixing. Additionally, for the CRG February flights the iMet-XQ2 data were not recorded in the INP flight data files, although iMet-XQ2 measurements were taken for the July flights. TBSIMET data are available from the ARM Data Center (https://adc.arm.gov/discovery/results/s::tbsimet).

To calculate θe, TBSIMET data (altitude, temperature, pressure, relative humidity) were first preprocessed by removing unrealistic, non-NaN spikes in altitude observations and converting altitude from kilometers to meters. The profiles were then averaged into 10-m vertical bins; within each bin the mean altitude, temperature (°C), pressure (hPa), and relative humidity (%) were computed. Dew point temperature (T_d) in each bin was estimated from the binned temperature and relative humidity using the Magnus–Tetens approximation. Saturation vapor pressure at T_d was computed and the mixing ratio r (kg kg^{-1}) obtained via $r = \varepsilon e_s/(p - e_s)$ with $\varepsilon = 0.622$. Potential temperature θ was calculated from the binned temperature and pressure, and equivalent potential temperature was computed using the Bolton (1980) formulation: $\theta e = \theta \exp((L_v r)/(c_p T_d(K)))$, where L_v is latent heat of vaporization ($\approx 2.5\times10^6$ J kg^{-1}), c_p is the dry-air specific heat at constant pressure (≈ 1005.7 J kg^{-1} K^{-1}), and $T_d(K)$ is T_d in kelvin. Uncertainty in θ_e was estimated as the one-standard-deviation of θ_e values computed for each original sample within a bin (i.e., compute θ_e per sample, then take the binwise standard deviation).

To determine whether the boundary layer at the PUFIN sampling altitudes was well mixed or stratified, we evaluated θ_e only within the altitude ranges where PUFIN was deployed. If the vertical variation (minimum to maximum) in θ_e was ≤ 1 K, the layer was considered well mixed; if > 1 K, it was classified as stratified. Variability in θ_e could manifest as discrete layers or as a continuous vertical gradient.

Bolton, D.: The Computation of Equivalent Potential Temperature, Mon. Wea. Rev., 108, 1046–1053, https://doi.org/10.1175/1520-0493(1980)108%3C1046:TCOEPT%3E2.0.CO;2, 1980.

Griesche, H. J., Engelmann, R., Radenz, M., Hofer, J., Althausen, D., Ansmann, A., Barry, K., Creamean, J., Jimenez, C., and Seifert, P.: Annual cycle of surface-coupling effects on Arctic mixed-phase clouds during MOSAiC, EGUsphere, 1–33, https://doi.org/10.5194/egusphere-2025-5708, 2025.

2. Some more background on the sampling strategy would be nice. For instance, why were the specific sampling altitudes chosen? Which atmospheric layers are of interest?

PUFIN sampling is conducted concurrently with other airborne measurements that provide particle size distributions and multimodal micro-spectroscopy, yielding single-particle and molecular-level information on atmospheric aerosols at different altitudes. These complementary measurements predate PUFIN deployment on the ARM TBS and are typically performed within predefined altitude ranges to facilitate comparisons across flight days, seasons, and locations. Although altitude bins are adjusted based on flight conditions and available time, they are generally targeted at ~0–250 m, 250–500 m, and >500 m AGL. As shown in Figures 5 and 6, these ranges were not always strictly followed during the initial PUFIN flights due to operational constraints. Along these lines, we updated the figure legends to reflect the correct start and end altitudes based on TBSIMET data, which are more accurate and available through the ARM Data Center (Dexheimer et al., 2024).

We added new text on sampling strategy in Section 2.4.2: PUFIN sampling is often conducted in coordination with concurrent airborne measurements, including aerosol size distributions and multimodal micro-spectroscopy, to enable complementary characterization of aerosol properties across altitudes. Sampling is performed within predefined altitude ranges to facilitate comparison across flights, seasons, and locations, targeting roughly 0–250 m, 250–500 m, and >500 m AGL. These altitude bins are adjusted as needed based on flight duration and atmospheric conditions. They can be modified in real time to target layers of interest identified from on-site aerosol and meteorological measurements.

We also added the following to Section 2.2: However, these measurements are only accurate to approximately ±20 m; therefore, altitude data are taken from the TBS iMet system, available as the "TBSIMET" datastream from the ARM Data Center (Dexheimer et al., 2024).

3. Is it possible to derive a recommendation for the minimum sampling time in differing environments or seasons from the provided results in Section 4.2? A brief "lessons-learned" could also foster PUFIN rebuilds and deployments by other groups.

This is an excellent question. While we would like to provide recommendations for minimum sampling times, we are still building the necessary experience to reliably and robustly do so. To date, PUFIN measurements have been conducted at only two sites (CRG and BNF). Although additional sites have been sampled using IcePuck (with the combined deployments spanning urban, forest, agricultural, and mountain environments) these differed in terms of flow rates, sampling altitudes, and seasons, limiting our ability to draw generalized conclusions. We anticipate that, as we expand observations across a broader range of sites and conditions, we will be better positioned to offer such recommendations in the future.

That said, we agree that including a brief "lessons learned" discussion is a valuable suggestion, and we now added the following text to the Summary section: As PUFIN has only been deployed a limited number of

times to date, we are still actively identifying best practices. While we have not yet fully evaluated performance under extreme environmental conditions, future deployments will require careful consideration of such factors. Initial experience highlights the importance of maximizing sampling volume for robust INP detection and the value of sending test samples from new or unfamiliar environments for immediate analysis to inform subsequent sampling strategies. We also note that inlet components (e.g., hose barbs) may introduce minor collection efficiency losses, though further testing is needed to assess potential trade-offs with background contamination during ascent and descent. Maintaining consistent altitude ranges where possible improves comparability, and future dedicated flights are planned to include longer-duration sampling at fixed "loitering" altitudes to better resolve vertical structure and potentially achieve even lower detection limits.

4. Figure 5 is difficult to interpret due to overlapping dots. Is each data point shown, or could one point be entirely covered by another? A separation into seasons might be helpful, as the temperature spectra are not only sampling time-dependent.

Figure 5 generated confusion among both reviewers and did not substantially contribute to the overall narrative. It was only briefly described in the text and was not essential for conveying the main points. Therefore, we chose to remove the figure to improve clarity.

5. Some more interpretation and discussion of the derived temperature spectra in section 4.2 would be interesting, e.g., line 309: Which INP types are representative of the winter spectra?

We appreciate the reviewer's interest in further interpretation of the derived temperature spectra and agree that exploring which INP types are represented would be scientifically valuable. However, a detailed analysis and interpretation of these features falls beyond the primary scope of this manuscript and the focus of *Atmospheric Measurement Techniques*, which emphasizes instrumentation, methods, and data availability. Our intent in Section 4.2 was to provide an initial presentation of the temperature spectra to demonstrate the capabilities of the dataset and to highlight its potential for future scientific investigation. In this sense, the spectra are meant to serve as a preview that encourages the community to further explore these data in combination with complementary meteorological and aerosol measurements. A more in-depth discussion of the underlying processes and attribution to specific INP types would be better suited for a study centered on scientific interpretation, such as one aligned with the scope of a journal like *Atmospheric Chemistry and Physics*.

We chose not to expand substantially on the interpretation in order to maintain focus and coherence with the goals of the paper. We clarified this intent by adding the following at the beginning of Section 4.2: The objective of this section is to provide an initial presentation of the temperature spectra, demonstrating the capabilities of the dataset and highlighting its potential for future scientific investigation. A more in-depth scientific interpretation of these data falls beyond the primary scope of this manuscript; however, we encourage the community to further explore these observations in combination with complementary meteorological and aerosol measurements.

RC2

General:

Creamean et al., 2025, describes the development of an aerosols sampling system (PUFIN) with focus on deployment on tethered balloon systems. The authors successfully built a system capable of sampling ice nucleating particles (INPs) in vertical profiles. Ice nucleation measurements were conducted downstream after extracting INPs from the filters of PUFIN. The technology is of great value for the scientific community, and the paper is very well suited for publication in AMT. Especially the accessibility to the data and possibility to request PUFIN for field measurements is highly appreciated. Nevertheless, the paper could be significantly improved by comparing the results and the sampling technique with literature and explain the example dataset in more detail. A discussion of the advantages and disadvantages of the sampling technique in comparison with other UAS and balloon systems and sampling techniques (e.g. impinger or impactor) could further improve the study. Therefore, I suggest the publication of the study after minor revisions.

We thank the reviewer for recognizing the accessibility of the dataset and for highlighting these shortcomings in our manuscript.

Key Considerations:

Introduction

The introduction is well written and summarizes important knowledge and state of the art about aerosol and INP measurement at different heights in the troposphere. The authors also discuss the advantages of tethered balloon systems (TBS) over uncrewed aerial systems (UAS), such as longer sampling times. However, I was wondering how the flight path and position of sampling is different between TBS and UAS. For example, using fixed-wing drones[1] or rotary wing drones[2–4] could allow for a more precise flight path and location than a balloon system which also relies on the wind conditions and ground accessibility of the sampling locations. Furthermore, I was wondering what the key advantages of the PUFIN system are over the SHARK kit.[5] I wish the authors could address the key differences in the introduction or in the discussion of the paper. A table comparing different unoccupied sampling tools for INPs[3,5,6] could be helpful in understanding the need for PUFIN.

While UAS platforms indeed can offer more precise control over flight paths and sampling locations, they are often constrained by shorter flight durations, payload limitations, and regulatory restrictions. In contrast, TBS deployments can operate for extended periods (up to ~6–8 hours), enabling longer sampling times and sufficient particle collection at multiple altitude levels (up to three) within a single flight. This extended sampling capability is particularly advantageous for INP measurements, where sufficient loading is critical.

We expanded Section 4.2 to include a comparison of INP concentrations from recent balloon-borne and UAS-based studies: At −20 °C, the INP concentrations observed at CRG (0.01–1 L^{-1} in winter; 0.01–12 L^{-1} in summer) and BNF (0.01–1 L^{-1} in spring; 0.02–11 L^{-1} in summer) fall within the broad range reported in recent studies. For example, Creamean et al. (2018) reported 1–11 L^{-1} in springtime agricultural regions in Colorado from vertically-resolved filters via a launched and retrieved balloon system, while similar concentrations (~0.02–12 L^{-1}) were observed by Bieber et al. (2020) and Seifried et al. (2021) near a lake in Austria in summer via UAS. Böhmländer et al. (2025) reported lower values of ~0.2–0.5 L^{-1} during

spring and autumn UAS test flights in a boreal environment in Finland. In April in Cyprus, Marinou et al. (2019) reported ~3 L^{-1} from UAS measurements. Values ranging from ~0.2–10 L^{-1} from size-resolved measurements have been reported by Porter et al. (2020) across multiple sites in Europe and the Arctic using balloon-based sampling. Overall, the concentrations reported here are consistent with the range of environments sampled in the literature, spanning relatively clean to more biologically or terrestrially influenced regions.

As noted in the introduction, previous work such as Porter et al. (2020) and Böhmländer et al. (2025) highlights the strengths of alternative approaches. We agree that instruments such as SHARK provide valuable measurements, but all techniques, including PUFIN, have inherent limitations. While SHARK offers important size-resolved information, it is currently limited to sampling at a single altitude. However, these methods and PUFIN are complementary rather than competing. One key aspect that distinguishes PUFIN is that, unlike existing systems typically developed within individual PI-led efforts, PUFIN is designed for deployment within the ARM user facility to provide accessible, community-driven measurements. Users can request the PUFIN systems for deployments and access data without the need to build or operate specialized instrumentation or process the samples via drop freezing assays themselves.

In response to the reviewer's suggestion, we also broadened the discussion to include additional relevant studies in the Introduction to better contextualize PUFIN within the existing literature: Bieber et al. (2020) and Seifried et al. (2021) describe a developed system, the DAPSI (Drone-based Aerosol Particle Sampling Impinger/Impactor), for collecting INP samples using UAS. To date, results have been limited to test flights and a short (3-day) campaign conducted near a lake in Austria during summer.

Bieber, P., Seifried, T. M., Burkart, J., Gratzl, J., Kasper-Giebl, A., Schmale, D. G., and Grothe, H.: A Drone-Based Bioaerosol Sampling System to Monitor Ice Nucleation Particles in the Lower Atmosphere, Remote Sensing, 12, 552, https://doi.org/10.3390/rs12030552, 2020.

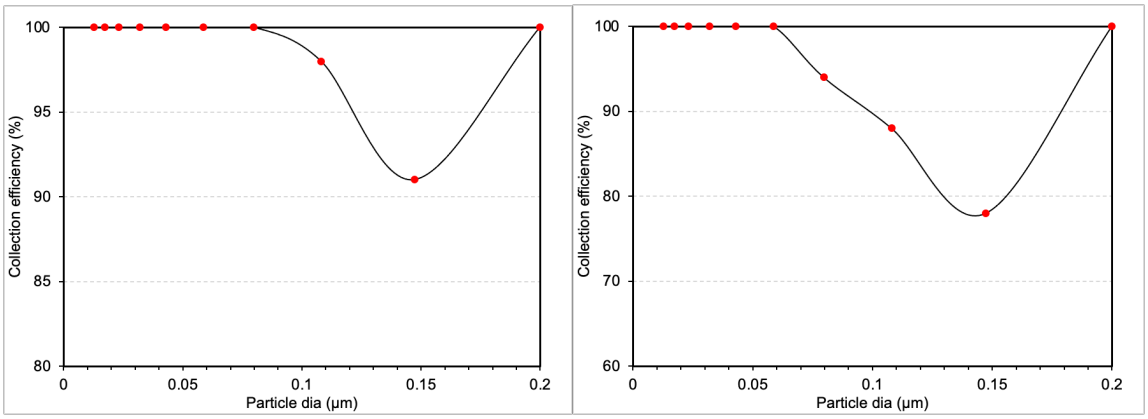
Seifried, T. M., Bieber, P., Kunert, A. T., Iii, D. G. S., Whitmore, K., Fröhlich-Nowoisky, J., and Grothe, H.: Ice Nucleation Activity of Alpine Bioaerosol Emitted in Vicinity of a Birch Forest, Atmosphere, 12, https://doi.org/10.3390/atmos12060779, 2021.

Methods

PUFIN samples INPs on filters which can be further extracted for ice nucleation experiments. Their workflow is clever and straightforward, the cleaning procedures and the immediate cooling of the samples to -80°C prior INP measurements with the CSU INS is thorough. However, there are some remaining questions which I think deserve attention prior publication: What is the size cutoff and sampling efficiency for aerosol sampling on the filter? How do the authors make sure that all INPs are washed of the filter? What is the advantage of sampling INPs on filters over impingers[7] or impactors[5]? Does the developed sampling procedure limit downstream analysis of INPs, for example cultivation of ice nucleating microbes[7]? What is the total weight of the setup? Could it in theory also be applied for other balloons or UAS?

Based on the calculations of Spurny and Lodge (1972), the collection efficiencies of the 0.2-µm filters across the flow rate range at which PUFIN has operated at CRG and BNF (4-11 L min^{-1}) are relatively high.

As shown in the plots below, the sampling efficiency decreases around 150 nm, reaching minimum values of approximately 78–91%.



Collection efficiency plots at (left) 4 L min^{-1} and (right) 11 L min^{-1}.

In principle, these filters sample total suspended particulates (TSP); however, the use of in-line filter units introduces potential inlet-related biases, particularly for the largest particles. The transmission of larger particles into the inlet is likely influenced by wind speed and associated inertial and gravitational effects, which can reduce sampling efficiency for coarse-mode aerosols. From a particle transport perspective, such behavior would be governed by size-dependent losses characterized by parameters such as the Stokes number and settling velocity, suggesting that very large particles may be underrepresented. That said, these effects have not yet been empirically constrained for this system and remain an area for future characterization.

Regarding the removal method, in the past we tested rotation versus vortexing for the same extraction duration. Although vortexing is more aggressive in removing particles, it did not produce different results compared to rotation. Additionally, this technique has also been intercompared with other methods employing different particle removal approaches and shows good agreement (DeMott et al., 2017, 2025; Lasher et al., 2024). However, future work should explore a range of rotation times to confirm that 20 minutes is sufficient across filters with varying particle loadings. We added the following to the Summary section: For particle removal from the filters, this approach has been intercompared with other techniques employing different extraction methods and showed good agreement (DeMott et al., 2017, 2025; Lasher et al., 2024). Future work will evaluate a range of rotation times to ensure that the current 20-minute extraction is sufficient across filters with varying particle loadings.

Filters offer several advantages over impingers, particularly in terms of simplicity and robustness. They do not require liquid handling during sampling, are less sensitive to environmental conditions (e.g., temperature, humidity), and are easier to deploy in remote or harsh environments. In addition, impinger samples generally need to be analyzed more quickly, as INPs can evolve during storage in aqueous media, even when frozen, whereas INPs collected on filters tend to be more stable when stored frozen (e.g., Beall et al., 2020). Compared to impactors, filters provide a bulk-integrated sample rather than size-segregated fractions, which is advantageous when the goal is to quantify total INP concentrations across all particle sizes. Although this comes at the expense of size resolution, it simplifies analysis and reduces uncertainties

associated with particle bounce, re-entrainment, and stage-specific collection efficiencies. That said, each method has its strengths: impingers can better preserve particle hydration state and collect semi-volatile material, while impactors provide size-resolved information. Ultimately, the choice depends on the scientific objective. In general, filters offer a versatile, robust, and widely adopted approach for INP sampling, facilitating greater intercomparability across studies.

Downstream microbial analyses may be possible, but are unlikely with the sampling volumes and flow rates currently used, which limit the amount of biomass collected. Achieving sufficient loading for such analyses would require higher-flow systems like impingers or SASS samplers (https://www.resrchintl.com/products/sass-3100.php) operating at several hundred liters per minute, particularly over the relatively short sampling durations employed here. The primary objective of PUFIN is to obtain INP-resolved measurements at up to three altitudes within a single flight, which constrains available sampling time and precludes the longer durations that would be needed to accumulate adequate biomass. Extending flight times to meet these requirements would be operationally impractical. That said, this is a compelling direction. We are exploring opportunities with ARM users to incorporate miniaturized bioaerosol sampling alongside PUFIN in upcoming campaigns, including the DUSTIEAIM (https://armgov.svcs.arm.gov/research/campaigns/amf2026dustieaim) study in the coming year.

The total weight of PUFIN is 6.1 kg as stated on line 134.

To address your final question, the PUFIN systems are currently configured for deployment on the ARM TBS. However, in principle, the setups are flexible and could be adapted for use on other balloon platforms or UASs. If there is interest from the user community, we would welcome the opportunity to explore expanding PUFIN beyond its current platform. We added the following text to Section 4.3, which we also changed the title to "Requesting PUFIN for future ARM TBS campaigns and beyond": PUFIN is currently configured for deployment on the ARM user facility TBS system, but the design is flexible and could be adapted for use on other balloon platforms or UASs, contingent on community interest.

To address these, we now changed the title of section 2.4.1 to "Filter material and holder preparation" and added the following paragraph: The 0.2-µm filters used in PUFIN provide a simple, robust, and widely adopted approach for INP sampling, enabling bulk-integrated collection across particle sizes and facilitating intercomparability across studies. Compared to impingers, filters do not require liquid handling, are less sensitive to environmental conditions, and allow for more stable storage of INPs when frozen (e.g., Beall et al., 2020). Compared to impactors, they provide a bulk-integrated sample rather than size-segregated fractions, simplifying analysis and reducing uncertainties associated with particle bounce and stage-specific collection efficiencies. Based on calculations from Spurny and Lodge (1972), these filters exhibit relatively high collection efficiencies across PUFIN's operating flow rate range (4-11 L min^{-1}), with a decrease near 150 nm but still maintaining efficiencies on the order of 78–91%. In principle, the filters sample total suspended particulates. However, inlet-related biases, particularly for coarse particles, may arise due to wind speed-dependent inertial and gravitational losses (e.g., governed by Stokes number and settling velocity), such that the largest particles may be underrepresented, although this has not yet been empirically quantified for this system.

Beall, C. M., Lucero, D., Hill, T. C., DeMott, P. J., Stokes, M. D., and Prather, K. 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, P. J., Hill, T. C. J., Petters, M. D., Bertram, A. K., Tobo, Y., Mason, R. H., Suski, K. J., McCluskey, C. S., Levin, E. J. T., Schill, G. P., Boose, Y., Rauker, A. M., Miller, A. J., Zaragoza, J., Rocci, K., Rothfuss, N. E., Taylor, H. P., Hader, J. D., Chou, C., Huffman, J. A., Pöschl, U., Prenni, A. J., and Kreidenweis, S. M.: Comparative measurements of ambient atmospheric concentrations of ice nucleating particles using multiple immersion freezing methods and a continuous flow diffusion chamber, Atmospheric Chem. Phys., 17, 11227–11245, https://doi.org/10.5194/acp-17-11227-2017, 2017.

DeMott, P. J., Mirrielees, J. A., Petters, S. S., Cziczo, D. J., Petters, M. D., Bingemer, H. G., Hill, T. C. J., Froyd, K., Garimella, S., Hallar, A. G., Levin, E. J. T., McCubbin, I. B., Perring, A. E., Rapp, C. N., Schiebel, T., Schrod, J., Suski, K. J., Weber, D., Wolf, M. J., Zawadowicz, M., Zenker, J., Möhler, O., and Brooks, S. D.: Field intercomparison of ice nucleation measurements: the Fifth International Workshop on Ice Nucleation Phase 3 (FIN-03), Atmospheric Measurement Techniques, 18, 639–672, https://doi.org/10.5194/amt-18-639-2025, 2025.

Lacher, L., Adams, M. P., Barry, K., Bertozzi, B., Bingemer, H., Boffo, C., Bras, Y., Büttner, N., Castarede, D., Cziczo, D. J., DeMott, P. J., Fösig, R., Goodell, M., Höhler, K., Hill, T. C. J., Jentzsch, C., Ladino, L. A., Levin, E. J. T., Mertes, S., Möhler, O., Moore, K. A., Murray, B. J., Nadolny, J., Pfeuffer, T., Picard, D., Ramírez-Romero, C., Ribeiro, M., Richter, S., Schrod, J., Sellegri, K., Stratmann, F., Swanson, B. E., Thomson, E. S., Wex, H., Wolf, M. J., and Freney, E.: The Puy de Dôme ICe Nucleation Intercomparison Campaign (PICNIC): comparison between online and offline methods in ambient air, Atmospheric Chem. Phys., 24, 2651–2678, https://doi.org/10.5194/acp-24-2651-2024, 2024.

Spurny, K. R. and Lodge, J. P.: Collection Efficiency Tables for Membrane Filters Used in the Sampling and Analysis of Aerosols and Hydrosols, Laboratory of Atmospheric Science, National Center for Atmospheric Research, 56 pp., 1972.

Results

Figure 5 could be discussed with more detail. What is the main message of this figure? That PUFIN has the advantage over IcePuck with the sampling time? This is not obvious to me as a reader.

Figure 5 generated confusion among both reviewers and did not substantially contribute to the overall narrative. It was only briefly described in the text and was not essential for conveying the main points. Therefore, we chose to remove the figure to improve clarity.

The resulting INP data (Figure 6 and 7) show a wide range of onset temperatures and INP concentrations. How does this range compare to other studies that measured vertical profiles of INPs?

While direct comparison of onset temperatures is not feasible given their dependence on blank corrections and variability across different techniques, we can compare our results with INP concentrations reported for similar environments. We added the following text to Section 4.2: At −20 °C, the INP concentrations

observed at CRG (0.01–1 L^{-1} in winter; 0.01–12 L^{-1} in summer) and BNF (0.01–1 L^{-1} in spring; 0.02–11 L^{-1} in summer) fall within the broad range reported in recent studies. For example, Creamean et al. (2018) reported 1–11 L^{-1} in springtime agricultural regions in Colorado from vertically-resolved filters via a launched and retrieved balloon system, while similar concentrations (~0.02–12 L^{-1}) were observed by Bieber et al. (2020) and Seifried et al. (2021) near a lake in Austria in summer via UAS. Böhmländer et al. (2025) reported lower values of ~0.2–0.5 L^{-1} during spring and autumn UAS test flights in a boreal environment in Finland. In April in Cyprus, Marinou et al. (2019) reported ~3 L^{-1} from UAS measurements. Values ranging from ~0.2–10 L^{-1} from size-resolved measurements have been reported by Porter et al. (2020) across multiple sites in Europe and the Arctic using balloon-based sampling. Overall, the concentrations reported here are consistent with the range of environments sampled in the literature, spanning relatively clean to more biologically or terrestrially influenced regions.

Bieber, P., Seifried, T. M., Burkart, J., Gratzl, J., Kasper-Giebl, A., Schmale, D. G., and Grothe, H.: A Drone-Based Bioaerosol Sampling System to Monitor Ice Nucleation Particles in the Lower Atmosphere, Remote Sensing, 12, 552, https://doi.org/10.3390/rs12030552, 2020.

Böhmländer, A., Lacher, L., Brus, D., Doulgeris, K.-M., Brasseur, Z., Boyer, M., Kuula, J., Leisner, T., and Möhler, O.: A novel aerosol filter sampler for measuring the vertical distribution of ice-nucleating particles via fixed-wing uncrewed aerial vehicles, Atmospheric Measurement Techniques, 18, 3959–3971, https://doi.org/10.5194/amt-18-3959-2025, 2025.

Creamean, J. M., Primm, K. M., Tolbert, M. A., Hall, E. G., Wendell, J., Jordan, A., Sheridan, P. J., Smith, J., and Schnell, R. C.: HOVERCAT: a novel aerial system for evaluation of aerosol–cloud interactions, Atmospheric Measurement Techniques, 11, 3969–3985, https://doi.org/10.5194/amt-11-3969-2018, 2018.

Marinou, E., Tesche, M., Nenes, A., Ansmann, A., Schrod, J., Mamali, D., Tsekeri, A., Pikridas, M., Baars, H., Engelmann, R., Voudouri, K.-A., Solomos, S., Sciare, J., Groß, S., Ewald, F., and Amiridis, V.: Retrieval of ice-nucleating particle concentrations from lidar observations and comparison with UAV in situ measurements, Atmos. Chem. Phys., 19, 11315–11342, https://doi.org/10.5194/acp-19-11315-2019, 2019.

Porter, G. C. E., Sikora, S. N. F., Adams, M. P., Proske, U., Harrison, A. D., Tarn, M. D., Brooks, I. M., and Murray, B. J.: Resolving the size of ice-nucleating particles with a balloon deployable aerosol sampler: the SHARK, Atmos. Meas. Tech., 13, 2905–2921, https://doi.org/10.5194/amt-13-2905-2020, 2020.

Seifried, T. M., Bieber, P., Kunert, A. T., Iii, D. G. S., Whitmore, K., Fröhlich-Nowoisky, J., and Grothe, H.: Ice Nucleation Activity of Alpine Bioaerosol Emitted in Vicinity of a Birch Forest, Atmosphere, 12, https://doi.org/10.3390/atmos12060779, 2021.

The February vs July comparison in Figure 6 is particularly interesting. Could that correlate with the agricultural activity the authors were discussing earlier? Where any additional test performed on the sample (e.g. heat test) to test for the composition and origin of the INPs?

The characteristic “bio-hump” observed at temperatures warmer than –20 °C in the July spectra, relative to February, is likely associated with agricultural sources. The TBS was deployed in a region bordered by farmland to the east and the Chesapeake Bay to the west, and according to the 2022 USDA Agricultural Census, the dominant crops in Kent County are grains, oilseeds, dry beans, and dry peas (USDA, 2022). The resulting INP spectra are consistent with those observed at SGP, which is similarly influenced by surrounding agricultural activity. Although heat treatments were not performed on these samples, such analyses could help further constrain INP types and may be requested by users through ARM (https://www.arm.gov/guidance/campaign-guidelines/small-campaigns).

We added the following to section 4.2: However, the Baltimore region is bordered by farmland to the east where the dominant crops are grains, oilseeds, dry beans, and dry peas (USDA, 2022), which may have influenced the sampled INP population. Heat treatment

U.S. Department of Agriculture (USDA), National Agricultural Statistics Service (NASS): 2022 Census of Agriculture: County Profile – Kent County, Maryland, available at: https://www.nass.usda.gov/Publications/AgCensus/2022/Online_Resources/County_Profiles/Maryland/cp24029.pdf, last access: 3 April 2026.

We also added the following to section 4.3: Users may request sample collection for total INP concentrations, as well as the application of thermal and peroxide treatments to infer INP types, including heat-labile (likely biological), heat-stable (likely organic), and inorganic (likely mineral) INPs (Barry et al., 2023, 2025; DeMott et al., 2025; Hill et al., 2016; McCluskey et al., 2018; Schiebel et al., 2016; Suski et al., 2018; Testa et al., 2021; Tobo et al., 2019).

Barry, K. R., Hill, T. C. J., Kreidenweis, S. M., DeMott, P. J., Tobo, Y., and Creamean, J. M.: Bioaerosols as indicators of central Arctic ice nucleating particle sources, Atmospheric Chemistry and Physics, 25, 11919–11933, https://doi.org/10.5194/acp-25-11919-2025, 2025.

DeMott, P. J., Mirrielees, J. A., Petters, S. S., Cziczo, D. J., Petters, M. D., Bingemer, H. G., Hill, T. C. J., Froyd, K., Garimella, S., Hallar, A. G., Levin, E. J. T., McCubbin, I. B., Perring, A. E., Rapp, C. N., Schiebel, T., Schrod, J., Suski, K. J., Weber, D., Wolf, M. J., Zawadowicz, M., Zenker, J., Möhler, O., and Brooks, S. D.: Field intercomparison of ice nucleation measurements: the Fifth International Workshop on Ice Nucleation Phase 3 (FIN-03), Atmospheric Measurement Techniques, 18, 639–672, https://doi.org/10.5194/amt-18-639-2025, 2025.

McCluskey, C. S., Hill, T. C. J., Humphries, R. S., Rauker, A. M., Moreau, S., Strutton, P. G., Chambers, S. D., Williams, A. G., McRobert, I., Ward, J., Keywood, M. D., Harnwell, J., Ponsonby, W., Loh, Z. M., Krummel, P. B., Protat, A., Kreidenweis, S. M., and DeMott, P. J.: Observations of ice nucleating particles over Southern Ocean waters, Geophysical Research Letters, 45, 11,989-11,997, https://doi.org/10.1029/2018GL079981, 2018.

Schiebel, T., Höhler, K., Funk, R., Hill, T. C. J., Levin, E. J. T., Nadolny, J., Steinke, I., Suski, K. J., Ullrich, R., Wagner, R., Weber, I., DeMott, P. J., and Möhler, O.: Ice nucleation activity of various agricultural soil

dust aerosol particles, European Geosciences Union General Assembly, Wien, A, April 17-22, 2016. Geophysical Research Abstracts, 18(2016) EGU2016-13422, 2016.

Testa, B., Hill, T. C. J., Marsden, N. A., Barry, K. R., Hume, C. C., Bian, Q., Uetake, J., Hare, H., Perkins, R. J., Möhler, O., Kreidenweis, S. M., and DeMott, P. J.: Ice Nucleating Particle Connections to Regional Argentinian Land Surface Emissions and Weather During the Cloud, Aerosol, and Complex Terrain Interactions Experiment, JGR Atmospheres, 126, https://doi.org/10.1029/2021JD035186, 2021.

Tobo, Y., Adachi, K., DeMott, P. J., Hill, T. C. J., Hamilton, D. S., Mahowald, N. M., Nagatsuka, N., Ohata, S., Uetake, J., Kondo, Y., and Koike, M.: Glacially sourced dust as a potentially significant source of ice nucleating particles, Nat. Geosci., 12, 253–258, https://doi.org/10.1038/s41561-019-0314-x, 2019.

I wish the authors could spend more time discussing the data evaluation of the results from their freezing experiments: It might be worth (i) showing frozen fractions for some samples, at least in the SI, (ii) discussing how the blank background nucleated ice and why background subtraction was applied, and (iii) discussing the choice of calculating confidence intervals using Agresti & Coull[8] over using the approach of Fahy et al.[9] or others.

We agree that additional clarification is valuable.

(i) We do not present frozen fraction curves, as our analysis is based on serial dilutions, with INP concentrations calculated independently for each dilution. As such, frozen fractions are specific to each dilution step rather than representing a single continuous curve, and we instead report INP concentrations derived from these dilution-resolved measurements.

(ii) Blank subtraction is applied to account for residual and procedural sources of INPs. Some INPs may remain on filters even after cleaning and can also be introduced during handling and connection of the sampling units. Additionally, while DI water is filtered (0.1-µm), trace INPs may still remain or be introduced during experimental setup. We also note that defects in PCR trays used for droplet freezing assays can occasionally initiate freezing. Together, these factors contribute to background freezing and necessitate subtraction to isolate the atmospheric INP signal.

(iii) We acknowledge the alternative bootstrap approach proposed by Fahy et al.; however, this method is best suited for datasets with larger droplet counts per distribution (~150–200 droplets). In our case, although ~160 droplets are analyzed per sample, they are partitioned into serial dilutions of 32 wells each, and INP concentrations are derived independently for each dilution. Applying the Fahy et al. approach under these conditions would likely yield overly large and unrepresentative uncertainty bounds. Instead, we use Agresti and Coull confidence intervals, which are appropriate for binomial proportions with sample sizes on the order of ~30, where the central limit theorem supports an approximate normal approximation.

It seems that the authors use the two-sample t-test for determining statistical differences between that samples at different altitudes. Is that test suitable for testing cumulative nucleation spectra? If yes, what values were used for the test? Alternatively, the authors could just discuss the overlapping confidence

intervals or also use a bootstrapping algorithm[9] to back calculate the variations of the ice nucleation measurements.

Because the method of Fahy et al. is best suited for datasets with larger droplet counts per distribution (~150–200 droplets), and the reviewer raised a valid question that a two-sample t-test may not well suited for intercomparison of INP spectra, we now instead use overlapping confidence intervals to assess differences between samples at different altitudes or sites as suggested. The text has been revised accordingly throughout Section 4.2, which can be viewed in the track changes version below.

Summary

What are future improvements that could be made for PUFIN?

We thank the reviewer for this question and the opportunity to outline potential future improvements to PUFIN. At present, altitude measurements rely on the iMet-XQ2, which provides accuracy on the order of ±20 m. Thus, we often instead employ the TBSIMET data, which are more accurate, but it may not be clear to users that this is the best dataset to use. In future iterations, we plan to incorporate a pressure-based altitude reference, which would improve consistency with TBS system measurements and better align with flight logs. We are also considering integration of higher-accuracy GPS sensors (e.g., improving from ~±12 m to sub-meter or centimeter-scale precision available with more advanced systems), as well as implementing a standardized approach to report altitude relative to surface elevation (AGL) directly from flight logs to facilitate inter-site comparisons.

We aim to improve power management in the next-generation design. During some deployments, low battery levels in the iMet-XQ2 may have led to increased power draw as the system attempted to recharge it, potentially impacting overall instrument performance. Future designs will address this by enabling more reliable charging and power distribution to ensure stable operation throughout flights.

Additional, future work should explore a range of rotation times to confirm that 20 minutes is sufficient across filters with varying particle loadings. Our technique has been intercompared with other methods using different particle removal approaches and shows good agreement (DeMott et al., 2017, 2025; Lasher et al., 2024). Nonetheless, future testing to prove the methods is useful.

We have added the following to the Summary section: Future improvements to PUFIN will focus on enhancing altitude accuracy, overall system robustness, and evaluating particle removal methodologies to ensure the most effective approach is used. Planned upgrades include the incorporation of pressure-based altitude measurements and higher-accuracy GPS sensors, as well as standardized reporting of altitude relative to surface elevation (AGL) to facilitate inter-site comparisons. In addition, improvements to power management, particularly more reliable charging and distribution for onboard sensors, are anticipated to ensure stable operation during extended deployments. For particle removal from the filters, this approach has been intercompared with other techniques employing different extraction methods and shows good agreement (DeMott et al., 2017, 2025; Lasher et al., 2024). Future work will evaluate a range of rotation times to ensure that the current 20-minute extraction is sufficient across filters with varying particle loadings.

Minor suggestions:

Line 20: It might be worth dropping two or three concentration levels in the abstract to support the boundary layer stratification and the local vs. transported aerosol argument.

We are not entirely clear on the reviewer's request and have therefore not made changes at this time. We would be happy to address this point if the reviewer can provide further clarification.

Line 34: Change "freezing supercooled liquid …" to "freezing supercooled liquid water …"

Done.

Line 46: Could you perhaps give an example of INP measurements at "surface" and how high these measurements normally extend? It is worth noting that INP measurements from towers can reach sampling heights of 60 m above ground.[10]

We revised the text to read: "However, INP measurements are often made at ground level, which may not accurately represent the concentrations or composition of INP populations at cloud level where ice nucleation processes occur (Burrows et al., 2022)...". Burrows et al. directly addresses this issue, particularly in the context of numerical models, highlighting why this discrepancy is important and why improved representation is needed.

Line 112: It might be worth adding the list and schematic design to the SI of the paper.

While we appreciate the suggestion, we have chosen not to include the current version of the parts list and design in the Supporting Information, as these are subject to change over time. Instead, we prefer to provide the community with access to the most up-to-date parts list and PUFIN design via a GitHub repository, which we will update regularly.

Line 117: It might be worth defining standard liter in "sL min^{-1}".

Done.

Figure 1: Very minor edit, but it seems the mass flow meter symbol does not connect to the airflow arrows correctly.

Fixed.

Table 1: The authors give the altitude of sample in meters above mean sea level (AMSL). Although sampling from 0 to 1000 m AMSL is theoretically possible, it would mean flying from sea level (0 m) to 1000 m. Their sampling locations in Colorado, Alabama and Maryland seem to be already a couple of meters above sea level. The authors should consider adding more accurate sampling heights into Table 1 and also consider adding the GPS coordinates of the sampling locations.

This was a typo in the original submission. The values are AGL, not AMSL, and we have corrected this in Table 1.

Line 217: Consider changing “proportion” to “fraction”

Done.

Line 220: “INP spectra are corrected using DI negative controls and subsequently blank-subtracted.” Please specify what was subtracted, the DI blank or the sample blank (filter from the PUFIN setup).

Both DI and sample blanks are subtracted from the INP spectra. Each sample run includes a 32-well DI negative control. For each 0.5 °C temperature bin, the number of frozen wells in the DI control is subtracted from the number of frozen wells in each dilution. The corrected freezing counts are then converted to frozen fraction for each dilution and subsequently to INP concentration in a combined spectrum.

Sample blanks are processed in a similar manner as filter samples, with INP concentrations calculated per blank filter for each 0.5 °C temperature bin. All blanks from a given campaign are averaged to generate a representative blank spectrum (expressed as INP per blank filter). For each sample, concentrations are first converted to INP per filter, the average blank spectrum is subtracted, and the corrected values are then converted back to INP L^{-1}.

We added the following to Section 3.2: Specifically, each run includes a 32-well DI negative control. For each 0.5 °C temperature bin, the number of frozen wells in the DI control is subtracted from the number of frozen wells in each dilution. The corrected freezing counts are then converted to frozen fraction for each dilution and subsequently to INP concentration as a combined spectrum. Sample blanks are processed in a similar manner as filter samples, with INP concentrations calculated per blank filter for each 0.5 °C temperature bin. All blanks from a given campaign are averaged to generate a representative blank spectrum (expressed as INP per blank filter). For each sample, concentrations are first converted to INP per filter, the average blank spectrum is subtracted, and the corrected values are then converted back to INP L^{-1}.

Line 269: Worth citing the flow rate values of IcePuck for comparison with PUFIN.

We now include the value or range of flow rates for each instrument and deployment in Table 1.

Figure 5: Check the left-hand side of the label, it seems to be cut off. I also got confused with the caption saying “Temperatures at which INPs were detected”. Does that mean onset temperatures, mean freezing temperatures or all temperatures at which droplets froze in the INS experiment?

See comment above. Figure 5 has been removed.

Figure 6: A statement to why and when sampling over an altitude range versus sampling at a set altitude was applied might be useful.

PUFIN sampling is conducted concurrently with other airborne measurements that provide particle size distributions and multimodal micro-spectroscopy, yielding single-particle and molecular-level information on atmospheric aerosols at different altitudes. These complementary measurements predate PUFIN deployment on the ARM TBS and are typically performed within predefined altitude ranges to facilitate comparisons across flight days, seasons, and locations. Although altitude bins are adjusted based on flight conditions and available time, they are generally targeted at ~0–250 m, 250–500 m, and >500 m AGL. As shown in Figures 5 and 6, these ranges were not always strictly followed during the initial PUFIN flights due to operational constraints. We updated the figure legends to reflect the correct start and end altitudes based on TBSIMET data, which are more accurate and available through the ARM Data Center (Dexheimer et al., 2024).

We added new text on sampling strategy in Section 2.4.2: PUFIN sampling is often conducted in coordination with concurrent airborne measurements, including aerosol size distributions and multimodal micro-spectroscopy, to enable complementary characterization of aerosol properties across altitudes. Sampling is performed within predefined altitude ranges to facilitate comparison across flights, seasons, and locations, targeting roughly 0–250 m, 250–500 m, and >500 m AGL. These altitude bins are adjusted as needed based on flight duration and atmospheric conditions. They can be modified in real time to target layers of interest identified from on-site aerosol and meteorological measurements.

We also added the following to Section 2.2: However, these measurements are only accurate to approximately ±20 m; therefore, altitude data are taken from the TBS iMet system, available as the "TBSIMET" datastream from the ARM Data Center (Dexheimer et al., 2024).

Line 341: Is the "concentrations as low as 10^{-2} INP L-^{1} in as little as 28 minutes" also tied to a minimum onset temperature? I guess with the current workflow, 10^{-2} INP L^{-1} active below -30°C would not be detectable. Consider attaching a minimal onset temperature to this statement.

The minimum onset temperatures are constrained by INP contributions from blank filters and DI water backgrounds, in addition to the collection times. To accurately report and intercompare onset temperatures, all of these factors would need to be consistent. As this is not always the case, particularly with variability in the blanks, we are unable to robustly report this value.

This is very picky, but the authors use "warmer temperature" quite frequently. Technically, an object like a droplet can be warm, however, the temperature is a property of the object and therefore either high or low, not warm or cold.

We appreciate the reviewer's point and revised the six instances where we used "warmer freezing temperatures." We retained one instance of "warm-temperature INPs," as this is a commonly used term within the ice nucleation community.

It might be worth considering including PUFIN in the title for further papers referencing the technique. Suggestion: "Reaching new heights: Profiling Upper altitudes For Ice Nucleation (PUFIN) on the Atmospheric Radiation Measurement (ARM) tethered balloon systems"

Thank you for the suggestion – a valid point. We adopted the reviewer's proposed title.

Line 350: Again, great job and big step to help the ice nucleation community with access to INP measurements.

Thank you!

References:

1. Schmale Iii, D. G.; Dingus, B. R.; Reinholtz, C. Development and Application of an Autonomous Unmanned Aerial Vehicle for Precise Aerobiological Sampling above Agricultural Fields. J. Field Robot. 2008, 25 (3), 133–147. https://doi.org/10.1002/rob.20232.
2. Borchers, C.; Moormann, L.; Geil, B.; Karbach, N.; Wasserzier, D.; Hoffmann, T. Development and Use of a Lightweight Sampling System for Height-Selective UAV-Based Measurements of Organic Aerosol Particles. Atmospheric Meas. Tech. 2025, 18 (23), 7231–7242. https://doi.org/10.5194/amt-18-7231-2025.
3. Bieber, P.; Seifried, T. M.; Burkart, J.; Gratzl, J.; Kasper-Giebl, A.; Schmale, D. G.; Grothe, H. A Drone-Based Bioaerosol Sampling System to Monitor Ice Nucleation Particles in the Lower Atmosphere. Remote Sens. 2020, 12 (3), 552. https://doi.org/10.3390/rs12030552.
4. Seifried, T. M.; Bieber, P.; Kunert, A. T.; Schmale, D. G.; Whitmore, K.; Fröhlich-Nowoisky, J.; Grothe, H. Ice Nucleation Activity of Alpine Bioaerosol Emitted in Vicinity of a Birch Forest. Atmosphere 2021, 12 (6), 779. https://doi.org/10.3390/atmos12060779.
5. Porter, G. C. E.; Sikora, S. N. F.; Adams, M. P.; Proske, U.; Harrison, A. D.; Tarn, M. D.; Brooks, I. M.; Murray, B. J. Resolving the Size of Ice-Nucleating Particles with a Balloon Deployable Aerosol Sampler: The SHARK. Atmospheric Meas. Tech. 2020, 13 (6), 2905–2921. https://doi.org/10.5194/amt-13-2905-2020.
6. Böhmländer, A. J.; Lacher, L.; Brus, D.; Doulgeris, K.-M.; Brasseur, Z.; Boyer, M.; Kuula, J.; Leisner, T.; Möhler, O. A Novel Aerosol Filter Sampler for Measuring the Vertical Distribution of Ice-Nucleating Particles via Fixed-Wing Uncrewed Aerial Vehicles. September 4, 2024. https://doi.org/10.5194/amt-2024-120.
7. Šantl-Temkiv, T.; Amato, P.; Gosewinkel, U.; Thyrhaug, R.; Charton, A.; Chicot, B.; Finster, K.; Bratbak, G.; Löndahl, J. High-Flow-Rate Impinger for the Study of Concentration, Viability, Metabolic Activity, and Ice-Nucleation Activity of Airborne Bacteria. Environ. Sci. Technol. 2017, 51 (19), 11224–11234. https://doi.org/10.1021/acs.est.7b01480.
8. Agresti, A.; Coull, B. A. Approximate Is Better than "Exact" for Interval Estimation of Binomial Proportions. Am. Stat. 1998, 52 (2), 119–126. https://doi.org/10.1080/00031305.1998.10480550.
9. Fahy, W. D.; Shalizi, C. R.; Sullivan, R. C. A Universally Applicable Method of Calculating Confidence Bands for Ice Nucleation Spectra Derived from Droplet Freezing Experiments. Atmospheric Meas. Tech. 2022, 15 (22), 6819–6836. https://doi.org/10.5194/amt-15-6819-2022.
10. Schrod, J.; Thomson, E. S.; Weber, D.; Kossmann, J.; Pöhlker, C.; Saturno, J.; Ditas, F.; Artaxo, P.; Clouard, V.; Saurel, J.-M.; Ebert, M.; Curtius, J.; Bingemer, H. G. Long-Term Deposition and Condensation Ice-Nucleating Particle Measurements from Four Stations across the Globe. Atmospheric Chem. Phys. 2020, 20 (24), 15983–16006. https://doi.org/10.5194/acp-20-15983-2020.