



The spatio-temporal variability of ice-nucleating particle concentrations quantified with a new benchmark dataset

Kristoffer Kvist^{1,2}, Eirik Mikal Samuelson^{1,2,3}, and Rune Grand Graversen¹

¹Department of Physics and Technology, The Arctic University of Norway (UiT), Tromsø, 9019, Norway

²The Norwegian Air Ambulance Foundation, Oslo, 0184, Norway

³Norwegian Meteorological Institute, P.O. Box 6314 Langnes, Tromsø, 9037, Norway

Correspondence: Kristoffer Kvist (kristoffer.kvist@uit.no)

Abstract. Ice-nucleating particles (INPs) catalyze primary ice formation, thereby raising the glaciation temperature of mixed-phase clouds. Glaciation substantially alters cloud microphysical and radiative properties. Therefore, the development and validation of models that predict INP concentrations receive considerable attention. Current model validations rely on sparse observational datasets. These validations compare modeled and measured INP concentrations at individual measurement points, but do not consider their temporal variability. This variability reflects the physical processes controlling INP abundance, and it therefore offers a largely unexploited constraint on the model physics. This study presents a multi-campaign dataset consisting of 110 000 immersion-mode INP observations. Statistical adjustments for differing sampling intervals and campaign durations are derived that allow the heterogeneous dataset to be synthesized into a global characterization of INP concentrations and their variation. The campaign means of INP concentrations at -23°C span more than four orders of magnitude, from Antarctica to agricultural continental sites. The within-campaign spread is around half an order of magnitude and increases with temperature. A multi-campaign time-series analysis shows that almost half the variance of daily values resides within a week, and roughly 20 % beyond 3 months. Across the resolved temperatures (-28 to -15°C), organic material accounts for up to 90 % of the INP population in the sampled environments. Earlier modeling studies that attribute most INPs to mineral dust therefore likely overestimate the dust contribution. The extensive collection of INP observations and the derived statistics of INP variation provide a strong foundation for improved model evaluation.

1 Introduction

Ice-nucleating particles (INPs) enable primary ice formation through heterogeneous nucleation, allowing supercooled droplets to freeze at temperatures warmer than the homogeneous freezing limit of approximately -38°C (Murray et al., 2012). In supercooled liquid clouds and mixed-phase clouds, primary ice formation initiates an array of microphysical processes: the ice crystals grow at the expense of the surrounding liquid droplets through the Wegener–Bergeron–Findeisen process (Wegener, 1911; Bergeron, 1935; Findeisen, 1938), and by riming (Fan et al., 2011), as long as liquid water remains present. Together with ice-ice aggregation, these INP-dependent pathways allow mixed-phase clouds to form large hydrometeors more readily than purely liquid clouds. The increased size of hydrometeors affects cloud radiative properties, precipitation onset, and cloud



lifetime (Korolev et al., 2017; Hawker et al., 2021). Atmospheric models therefore require realistic INP concentrations to
25 correctly simulate cloud phase, precipitation, and energy budgets that depend on it.

Many climate models nevertheless diagnose primary ice from temperature alone or from dust-based schemes tuned to mid-latitude observations (Gjelsvik et al., 2025; Silva et al., 2025). The resulting errors in cloud phase are a leading source of uncertainty in projected warming (Armour et al., 2021; Murray et al., 2021). The consequences are greatest for the low-level mixed-phase stratocumulus and shallow cumulus of marine cold-air outbreaks, in which polar air streams over the open ocean
30 (Bodas-Salcedo et al., 2016; Murray et al., 2021). In a warming climate, part of the ice in these clouds is replaced by liquid water. Liquid clouds, composed of many small droplets, reflect more sunlight and precipitate less readily than glaciated clouds, so albedo over the dark ocean increases. This negative cloud-phase feedback scales with the ice available for replacement (Storelvmo et al., 2015). A model with too much ice in present-day cold-air-outbreak clouds thus overestimates the feedback and underestimates climate sensitivity (Tan et al., 2016; Frey and Kay, 2018). Southern Ocean low-cloud phase is consequently
35 a central target for model development (Vignon et al., 2021; Bodas-Salcedo et al., 2016; Gettelman et al., 2020). Over much of the Arctic, snow and sea ice make the surface itself reflective, and the sun is low or absent for a large part of the year. There, liquid water gained under warming mainly warms the surface by letting clouds trap more of the heat the surface emits. A model with too much ice in present-day Arctic clouds has more ice to convert to liquid under warming and tends to simulate stronger Arctic amplification (Tan and Storelvmo, 2019; Taylor et al., 2022; Gjelsvik et al., 2025).

40 On the shorter timescales of weather forecasting, INP representation influences the cloud albedo (Armour et al., 2021), precipitation (Shete et al., 2026; Hawker et al., 2021; Thompson et al., 2004, 2008), and the amount of supercooled liquid cloud water with a direct effect on aircraft icing forecasting (Dupont et al., 2024; Engdahl et al., 2020; Rosenfeld et al., 2013; Shete et al., 2026).

Despite their importance, INPs are exceedingly rare within the aerosol population. Only about one in 10^3 to 10^6 ambient
45 particles acts as an INP between -10 and -37°C (Zhao et al., 2019), a fraction which is even smaller in remote maritime air masses (McCluskey et al., 2018c) and in ice-covered polar regions (Wex et al., 2025; Sze et al., 2023). Whether a given particle belongs to this subset depends strongly on its composition. The aerosol types that contribute most to the INP population can be divided into dust, sea spray aerosols (SSAs), and primary biological particles.

Mineral dust, originating from arid and semi-arid regions, is generally regarded as the dominant INP source in supercooled
50 liquid clouds and mixed-phase clouds, owing to both its global abundance and its ability to nucleate ice through immersion freezing (Herbert et al., 2025; Hoose and Möhler, 2012; DeMott et al., 2010). The ice nucleation efficiency of mineral dust depends on its mineralogy (Atkinson et al., 2013; Chatziparaschos et al., 2025) and organic content. Soil, agricultural dusts, and dust from glacial outlet plains are frequently enriched in organic and biological material that raises their ice nucleation ability well above that of their mineral component alone (O'Sullivan et al., 2014; Hill et al., 2016; Zhou et al., 2026; Tobo
55 et al., 2019; Sanchez-Marroquin et al., 2020).

Over the oceans, SSAs produced by bubble bursting constitute the primary local INP source. Because sea salt is a poor ice nucleus, the ice-nucleating ability comes from the organic fraction of SSA produced by marine biological activity and concentrated in the sub-micron mode (Burrows et al., 2014, 2022a). In the pristine, dust-poor marine air of the Southern Ocean,



SSAs are often the dominant INP source, despite their limited ice nucleation activity (Herbert et al., 2025; Chatziparaschos et al., 2023; Zhao et al., 2021; McCluskey et al., 2018c; Sato and Inoue, 2021).

Primary biological particles, e.g., bacteria, fungal spores, and pollen, are the most efficient INPs known, with the ability to nucleate ice at temperatures warmer than mineral and marine particles (Hoose and Möhler, 2012). These aerosols, although scarce in number, contribute significantly to the INP population and dominate at high temperatures due to their high nucleation activity (Gratzl et al., 2025; Zhou et al., 2026; Schneider et al., 2021). Current reports on the influence of biological and other organic aerosols on atmospheric INP concentrations are largely confined to individual sites and campaigns. This study provides comprehensive, temperature-resolved statistics of the biological and non-biological organic fractions of INPs across locations spanning the Southern Ocean to the high Arctic.

The diversity of INP types, which differ in ice nucleation activity, source regions, and atmospheric lifetime, produces complex spatio-temporal variation in INP concentrations. Across a wide range of environments, INP concentrations at a fixed activation temperature vary by as much as an order of magnitude within days to weeks at a single site, and their frequency distribution is approximately lognormal (Welti et al., 2018, 2020; Schrod et al., 2020). The lognormal shape is attributed to successive random dilution during transport: each mixing step reduces the INP concentration in an air parcel by an independent random factor, and the product of many such factors is lognormally distributed (Ott, 1990). Whether the width of these distributions differs systematically between environments and INP populations has not been consistently assessed across campaigns. This study provides such an assessment by comparing within-campaign variability in INP concentration across contrasting regions at multiple activation temperatures.

The timescales at which variability in INP concentration occurs remain largely unconstrained. Several studies report on seasonal variation of INP concentrations, in environments ranging from boreal forest (Schneider et al., 2021) to continental mountains (Zhou et al., 2026) and remote Arctic locations (Tobo et al., 2024). In contrast, a study of 2-year measurement series of deposition and condensation INPs from four locations worldwide shows that short-term variation dominates long-term signals (Schrod et al., 2020). These diverging results highlight limited understanding of INP variation and the need for datasets that include long-term, continuous INP concentration measurements. The dataset in the current study includes 21 long-term campaigns of largely unbiased INP concentration measurements spanning at least 90 days and sampled continuously, from which the typical decorrelation time and timescales of variation in atmospheric INP concentration are extracted.

Campaign statistics depend on how the campaign sampled the atmosphere: the time over which each sample is collected and the campaign's total time span both affect the mean, variance, and autocorrelation of the logarithm of the INP concentration. To make these statistics comparable across campaigns sampled differently, this study derives adjustment formulas that convert the three statistics obtained with a campaign's native sampling to a common reference sampling. These adjustment formulas allow a global characterization of INP variability to be obtained from the large number of heterogeneously sampled INP campaigns from which data are already published. The resulting characterization rests on a far larger and more diverse set of observations than the single-site measurements and the sets of identically sampled campaigns on which assessments of INP variability have so far relied (Welti et al., 2018, 2020; Schrod et al., 2020).



Over the last 15 years, climate models have moved toward aerosol-aware, multi-species INP models attempting to represent the large differences in ice nucleation ability across the aerosol population, with the aim of reducing the uncertainty in cloud–aerosol interactions involving INPs (Hoose et al., 2010; Vergara-Temprado et al., 2017; Wang et al., 2014; Zhao et al., 2021; McCluskey et al., 2023; Herbert et al., 2025; Costa-Surós et al., 2026). By treating each aerosol type separately, these schemes generate spatial and temporal variability that the traditional temperature- and humidity-based schemes of Cooper (1986), Meyers et al. (1992), and Fletcher et al. (2011) cannot provide. Recent studies suggest that aerosol-aware approaches improve INP concentration predictions in the regions of lowest loading, such as the Southern Ocean (Zhao et al., 2021; McCluskey et al., 2023). However, the advantage of multi-species aerosol models for global INP prediction remains uncertain because these models are evaluated on sparse datasets (Herbert et al., 2025; Vergara-Temprado et al., 2017; Chatziparaschos et al., 2023, 2025). The global dataset of atmospheric INP concentrations presented in this study combines measurements from land stations, balloons, aircraft, and ships. The resulting dataset spans environments from heavily aerosol-loaded continental air to pristine high-latitude regions and appears considerably larger and more diverse than datasets previously used to evaluate multi-species INP schemes. The hope is that this dataset will significantly ease thorough validation of future INP models by providing an easily accessible common benchmark along with statistical characterizations of the magnitude and timescales of INP variability that can serve as additional model constraints.

2 INP dataset

The INP dataset compiled for this study uses observations from more than 50 global sites and covers a wide range of aerosol regimes, from pristine maritime air to continental conditions with high natural and anthropogenic loading. Heterogeneous ice nucleation proceeds through several modes, distinguished by how the INP initiates freezing: immersion freezing, in which a particle immersed in a supercooled droplet triggers freezing; contact freezing, in which freezing is triggered by a particle colliding with the droplet; and deposition and condensation freezing, in which ice forms on the particle directly from the ambient vapor or during droplet activation (Burrows et al., 2022b). Immersion freezing dominates in mixed-phase clouds (Hande and Hoose, 2017), whereas deposition, condensation, and homogeneous nucleation are the main contributors to primary ice formation in pure ice clouds (DeMott et al., 1997; Sassen and Benson, 2000). The dataset used in this paper consists of immersion-mode INP observations, each giving the number of particles per liter of air at standard temperature and pressure (denoted std L^{-1}) that are active as INPs at a specified activation temperature. Figure 1 shows a geographical overview of the campaigns included in the compiled dataset. The dataset includes campaigns conducted from July 2011 to August 2025, with the contributing campaigns and their references listed in Table A1. Details of each campaign, including its environmental setting and sampling methodology, are provided in Sect. S3 of the supplement. Based on the campaign details in Sect. S3, each campaign is classified as either representative or not representative of local INP conditions at the sampling site. A campaign is considered representative if it sampled ambient air for at least 4 weeks without targeting particular meteorological or aerosol conditions, so that the observations reflect the conditions that occurred rather than those selected for study. This distinction is



125 used in many of the analyses presented below so that the results represent the climatology of INPs to the best possible extent. Campaigns deemed representative are marked with boldface text in Table A1.

Some campaigns sampled from more than one platform, for example from a land station and a tethered balloon. Each unique combination of campaign and platform is referred to as a *deployment*, and the statistics of Sect. 4 are computed per deployment. Table 1 lists the distribution of observations across platforms, sample treatments, instrument types, and hemispheres. It also
130 gives two quantities that characterize the sampled air and the surface beneath it: the number of observations from samples collected above the boundary-layer top and the number of observations from land samples collected over snow-covered ground. The altitude of the boundary-layer top is estimated for each sample as the sum of the surface height and the boundary-layer height from the ERA5 reanalysis (fifth generation ECMWF reanalysis, Hersbach et al., 2020) at the sampling position and time. Snow cover is defined by a 10-day running mean of the ERA5 snow depth exceeding 0.005 m water equivalent. For
135 mountaintop stations, this is a conservative estimate of the snow cover because the coarse ERA5 topography does not resolve mountain peaks. In total, the dataset comprises 109 637 observations, of which 91 134 stem from untreated samples (Sect. 2.1). Most measurements come from land stations, with the remainder split among ship-, aircraft-, and balloon-based measurements. More than a quarter of the land-station observations stem from samples collected over snow-covered ground, and about 10 % from samples collected above the boundary-layer top. The latter come from mountain stations such as Puy de Dôme, Storm
140 Peak, Sonnblick, Mount Helmos, and Zeppelin, which are intermittently situated above the top of the regional boundary layer as represented in ERA5; in such cases, they are assumed to sample air representative of the free troposphere. Neither the snow-cover nor the above-boundary-layer flag is used in the analysis presented here; both are retained in the dataset because they illustrate the range of conditions sampled and are likely to be useful in other applications of the dataset.

2.1 Instruments

145 The measurements of the dataset are obtained using either filter samples analyzed with ice spectrometers (ISs) or the Portable Ice Nucleation Experiment (PINE). To obtain the INP concentration using ISs, filters are exposed to airflow at the site of interest and then transported to the laboratory, where they are washed to obtain a water solution with the collected INP candidates suspended. The solution is added to a temperature-controlled metal plate as an array of droplets. By decreasing the plate temperature stepwise and counting the number of droplets frozen, the number concentration of INPs is deduced as a
150 function of temperature (DeMott et al., 2018b; Hiranuma et al., 2015). Finally, the background spectrum of a blank filter that has not been exposed to airflow is subtracted. The uncertainty of filter measurements is multi-sourced, including unintentional filter contamination, temperature uncertainty during the cooling of the metal plate (roughly $\pm 0.5^\circ\text{C}$; Miller et al., 2021), and counting uncertainty (Vali, 2014), the latter of which is especially relevant at the warmest temperatures where frozen droplets are sparse. Among the campaigns using filter samples, 10 different ice spectrometers are applied: CSU IS (Hill et al., 2014),
155 NIPI (O’Sullivan et al., 2014), CRAFT (Tobo, 2016), DRINCZ/DRINCO (David et al., 2019), LINA/INDA (Chen et al., 2018), LINDA (Stopelli et al., 2014), SPICE (Bergner et al., 2026), and INSEKT (Schiebel, 2017). Table A1 denotes the ice spectrometer used in each campaign. Intercomparison studies of ice spectrometers measuring the ice nucleation active surface site (INAS) density of lignin (Miller et al., 2021) and illite NX (Hiranuma et al., 2015) in the temperature range -10°C to -25°C

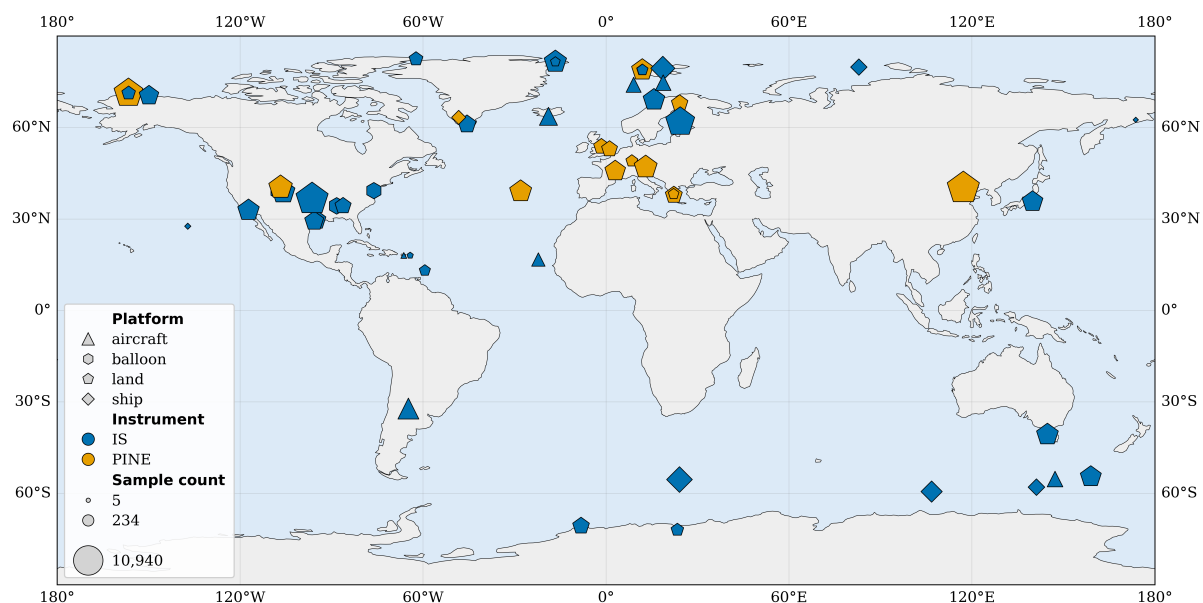


Figure 1. World map showing the average lon/lat location of each campaign contributing to the dataset. Marker shape illustrates the acquisition platform, while marker color indicates the instrument type, with IS denoting ice spectrometers (Hiranuma et al., 2015) and PINE denoting the Portable Ice Nucleation Experiment (Möhler et al., 2021). The marker area is proportional to the square root of the number of campaign samples. Table A1 summarizes the campaigns, with further details provided in Sect. S3.

and -10°C to -35°C , respectively, show that the ice spectrometers represented in the data rarely deviate by more than half an order of magnitude. Lacher et al. (2024) extend these laboratory results to ambient aerosol by deriving INP concentrations from filter samples collected at the mountain research station Puy de Dôme with seven different ice spectrometers, and find a comparable spread. Most INP spectra in the dataset have a resolution of either 0.5°C or 1°C . Depending on the purpose for which the dataset is applied, down-weighting or down-sampling the 0.5°C resolution spectra may be advantageous.

Exposing the solution obtained from filter washing to heat (95°C for 20 min) deteriorates the INP activity of biological particles, allowing their INP contribution to be quantified by comparing the results with those of an untreated sample. In addition to heat treatment, hydrogen peroxide can be added to the solution to obtain a 10 % H_2O_2 solution, which oxidizes and destroys organic compounds, leaving only mineral INPs in the solution (Huang et al., 2021). Some IS samples in the compiled dataset underwent such treatments, as shown in Table 1. These are used to investigate the biological and organic INP contributions at several sites and how these contributions differ across contrasting environments (Sect. 4.3).

The PINE instrument for immersion INP measurement, described in detail by Möhler et al. (2021), is an automated cloud expansion chamber that measures INP concentrations directly on site, in contrast to offline analysis of collected filters. Dried ambient air is drawn into a pre-cooled chamber and depressurized to cause adiabatic cooling, driving the air above water saturation and activating cloud droplets. An optical particle counter counts ice crystals formed from frozen droplets, and the immersion-mode INP concentration follows from the ice crystal count divided by the sampled air volume. Each expansion



Table 1. Observation counts by platform, treatment, and instrument over the entire dataset, together with the number of observations from samples collected above the boundary-layer top (Above BL) and over snow-covered ground. The presented counts reflect the dataset obtained after cleaning and clipping warm-tail plateaus (Sect. 3.2). The altitude of the regional boundary-layer top is estimated as the sum of ERA5 surface height and ERA5 boundary-layer height. A land sample counts as snow-covered when the 10-day running mean of the ERA5 snow depth at its position exceeds 0.005 m of water equivalent. NH (%) denotes the percentage of observations in the Northern Hemisphere.

Platform	Total	Treatment			Instrument		Above BL	Snow-covered ground	NH (%)
		Untreated	Heat	Peroxide	PINE	IS			
Land station	95 075	79 003	9120	6952	35 710	59 365	9772	26 442	91.3
Ship	8430	6946	1004	480	234	8196	–	–	35.2
Aircraft and balloon	6132	5185	577	370	0	6132	2231	–	64.7
Total	109 637	91 134	10 701	7802	35 944	73 693	12 003	26 442	85.5

175 takes 5–15 min, and the chamber temperature is ramped between expansions to obtain a temperature-resolved INP spectrum. The PINE data used in this study are from the large dataset GloPINE (Herbert et al., 2026), in which individual expansions are averaged into 6 h intervals and 2 °C temperature bins to lower the detection limit. The temperature uncertainty is ± 1 °C, dominated by the inhomogeneous temperature field developing during expansion, while the ice-crystal concentration uncertainty is 10–20 % (Möhler et al., 2021). As for filter methods, counting uncertainty dominates when INPs are sparse, i.e., in pristine
 180 regions and at high sampling temperatures. The previously mentioned study by Lacher et al. (2024) compares simultaneous in situ INP measurements from a series of ISs and PINE. They find that differences in INP concentrations obtained using different ISs are similar to the difference between the average INP concentration from ice spectrometry and the INP concentration measured by PINE. The comparison temperature range was -20 to -32 °C, covering about 94 % of the PINE data in the compiled dataset. The volume of sampled air per reported measurement is much lower for PINE than for typical filter approaches, forcing
 185 PINE to operate at low sampling temperatures (typically below -20 °C) to reach the minimum detection limit.

3 Methods

This section presents the methods applied to clean the data and to make the INP measurements comparable across the structured IS temperature spectra and the non-uniform spectra of the GloPINE dataset. It then describes how biological and non-biological organic contributions are obtained from treated spectra. Finally, the statistics reported for each campaign are adjusted to a
 190 common reference sampling, accounting for differing exposure times and campaign time spans of the included campaigns.

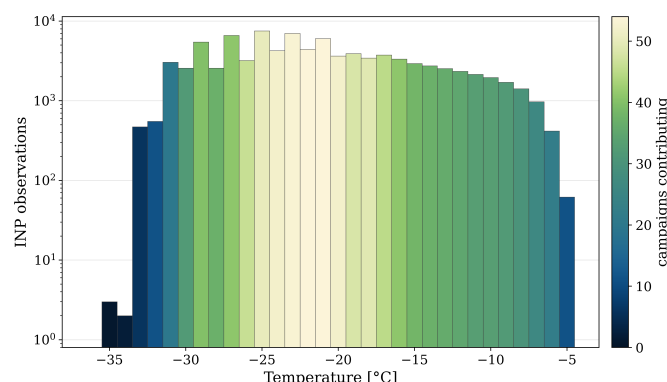


Figure 2. Temperature coverage of the INP dataset. The bars show the number of observations per 1 °C bin of activation temperature. The bar color indicates the number of distinct campaigns contributing to each bin.

3.1 Removal of negative and near-zero INP concentrations

Both PINE and IS measurements occasionally yield negative or near-zero INP concentrations. As described in Sect. 2.1, GloPINE records average INP concentrations over 6 h windows with activation temperature binned to a 2 °C resolution. An INP concentration of zero is reported when all measurements within a 6 h window contributing to a single temperature bin yield no ice formation. INP spectra from IS measurements are *blank-subtracted* and thus report near-zero values when the air-exposed filter yields an INP concentration similar to that of the blank filter. In the compiled dataset, near-zero values ($\leq 10^{-11}$ std L⁻¹) are removed before further processing. Two causes guide this choice: firstly, INP concentrations are analyzed in log₁₀-space in which near-zero values diverge, and secondly, the zero or near-zero values are misleading due to their high uncertainty. The maximum retrievable information from these readings is an upper bound on INP concentration set by the sampling volume (Vali, 1971; Herbert et al., 2026). Removing these uncertain outliers introduces a small positive bias near the instrument detection limits. For the filter-based methods, this is partly mitigated by the warm-tail clipping described next.

3.2 Warm-tail clipping of IS spectra

At the warm end of the IS spectra, where counting uncertainty is large, many spectra feature an initial plateau formed by a single *stray* droplet freezing at a high temperature, followed by no freezing activity for several degrees. The INP concentration of a filter-derived spectrum can even decrease despite a temperature decrement. This occurs when the INP concentration derived from the corresponding blank filter rises faster with decreasing temperature than that obtained from the air-exposed filter, which is not uncommon for heat-treated or peroxide-treated samples at sampling temperatures where the INP concentration is close to the detection limit. To remove such unphysical warm-tail plateaus, each IS spectrum is traversed from its warmest reading toward colder temperatures, and readings are discarded until a clear rise is evident, defined as an increase exceeding 0.02 decimal exponents (dex); this rather low threshold is set to avoid unnecessary loss of data, but with the trade-off that some



erroneous plateaus remain. This *clipping* is applied before any derived quantity is computed, so that the stored concentrations of the dataset are consistently based on the retained readings only. The crosses of Fig. 3b exemplify clipped warm-tail plateaus.

3.3 Projecting INP to a reference temperature

Temperature is the dominant variable controlling INP concentration, so comparisons of measured INP concentrations are sensitive to differences in sampling temperatures between measurements. Virtually all IS measurements in the presented dataset have spectra defined at integer temperatures in degrees Celsius. The GloPINE data (Herbert et al., 2026), by contrast, are binned in 6 h windows and 2 °C temperature intervals. Within each bin, both the activation temperature and the INP concentration are averaged over all expansions, so the reported temperature is a bin mean rather than a fixed value. The GloPINE spectra are therefore defined on an unstructured temperature grid.

To compare statistics of IS and PINE observations across campaigns, the INP concentrations are provided for a set of reference temperatures: $[-30^{\circ}\text{C}, -25^{\circ}\text{C}, -23^{\circ}\text{C}, -20^{\circ}\text{C}, -15^{\circ}\text{C}, -10^{\circ}\text{C}]$. These estimated INP concentrations are referred to as $\text{INP}^{\text{proxy}}(T_{\text{ref}})$. The 5 °C stride of reference temperatures is chosen because modeling studies often report the INP concentration at temperatures that are integer multiples of 5 °C (e.g., Vergara-Temprado et al., 2017; Chatziparaschos et al., 2023, 2025; Herbert et al., 2025); hence, the recorded reference concentrations are readily available for comparison with published model outputs. The reference temperature -23°C is included because it is the integer temperature that features the highest number of observations while still representing the maximum number of campaigns with overlapping temperature ranges, as shown in Fig. 2. For each spectrum, $\text{INP}^{\text{proxy}}(T_{\text{ref}})$ is obtained by the first applicable of the following methods, illustrated in Fig. 3a. No estimate is made if the reading nearest to T_{ref} lies more than 1.5 °C (IS) or 2.5 °C (PINE) from T_{ref} .

1. *Direct reading*. If the spectrum includes a reading within 0.1 °C of T_{ref} , that reading is used.
2. *Local linear fit*. Otherwise, a window centered on T_{ref} is widened in steps of 0.5 °C, starting from a width of 1 °C, until it includes at least four readings for IS or three readings for PINE, or reaches the maximum width of 6 °C. When enough readings lie within the window, a straight line is fitted to $\log_{10}(\text{INP})$ versus T within the window and evaluated at T_{ref} . If the window never holds enough readings, but at least three readings within $\pm 3^{\circ}\text{C}$ all lie on one side of T_{ref} , the line is fitted to these readings and extrapolated to T_{ref} . A fitted slope steeper than 1 dex per °C yields no estimate, since such slopes arise from nearly vertical stretches of a spectrum rather than from its local trend.
3. *Slope-shifted average*. Otherwise, the readings within $\pm 3^{\circ}\text{C}$ of T_{ref} are shifted to T_{ref} along a spectral slope and averaged. The slope is fitted separately for each campaign, platform, and treatment to the readings within $\pm 3^{\circ}\text{C}$ of T_{ref} , after the mean temperature and the mean $\log_{10}(\text{INP})$ of each spectrum have been subtracted. The slope thus describes how INP changes with temperature within the spectra, rather than how spectra measured at different temperatures differ. If fewer than 200 readings from spectra with at least two readings in this range are available, the slope of Cooper (1986), 0.132 dex per °C of supercooling, is used instead.

None of the presented results are sensitive to moderate adjustment of these settings.

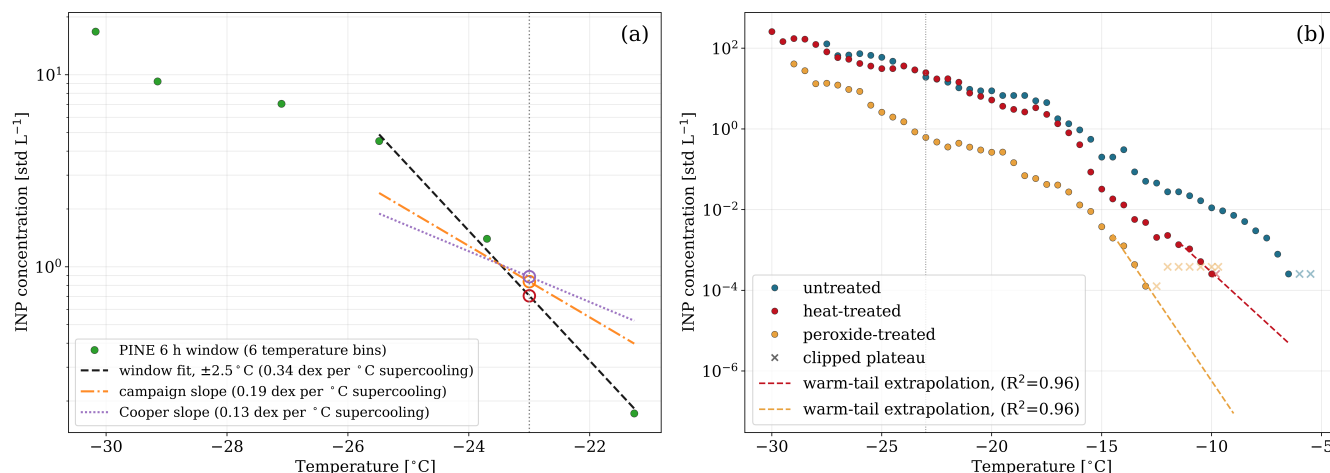


Figure 3. (a) Interpolation of the INP concentration to the reference temperature $T_{ref} = -23$ °C (Sect. 3.3) for one 6 h GloPINE spectrum of the ExINP-ENA campaign (28 December 2020). The dashed line is the local straight-line fit in $\log_{10}(\text{INP})$ within ± 2.5 °C of T_{ref} , which sets the stored value (red circle). The dash-dotted and dotted lines show the two fallback estimates, which shift all readings within ± 3 °C of T_{ref} to T_{ref} along the campaign slope and the slope of Cooper (1986), respectively, and average them. The legend gives each slope in dex per °C of supercooling. (b) Freezing spectra of one IS filter (GUC, 28 September 2021). Circles are the retained readings of the untreated, heat-treated, and peroxide-treated spectra, and crosses are the warm-side readings removed by the warm-tail clipping (Sect. 3.2), which reflect the detection limit rather than ambient INP. Dashed lines show the warm-tail extrapolations used for Fig. 7 (Sect. 3.4).

3.4 Temperature-resolved biological and organic INP contributions

To quantify how heat and peroxide treatments of filter-derived water solutions alter the INP concentration as a function of activation temperature, the quantities $\log_{10}(\text{INP}_{\text{heat}}/\text{INP}_{\text{untr}})$ and $\log_{10}(\text{INP}_{\text{perox}}/\text{INP}_{\text{heat}})$ are calculated. They express the fraction of the INP activity of the untreated (heat-treated) sample that remains after treatment with heat (peroxide). These quantities are calculated by forming treatment pairs from the subset of IS filters that include treated spectra. Each untreated (heat-treated) observation is matched with the heat-treated (peroxide-treated) observation closest in activation temperature, only accepting matches within 1 °C. To remove the temperature-induced offsets between matched pairs sampled at different temperatures, each heat-treated (peroxide-treated) value is projected to its untreated (heat-treated) counterpart's activation temperature using the $\log_{10}$ -linear projection strategy of Sect. 3.3.

Each freezing spectrum (untreated, heat-treated, peroxide-treated) terminates at the warmest temperature where freezing is detected, typically causing the heat-treated (peroxide-treated) spectra to terminate at colder temperatures than the untreated (heat-treated) spectra. If excluding pairs where a finite INP concentration is reported for the untreated (heat-treated) sample and an observation is lacking for the heat-treated (peroxide-treated) sample, the warmest temperature bins will have an overrepresentation of heat-treated (peroxide-treated) spectra with unusually warm freezing temperatures, since such abnormal measurements are the only ones that exceed the detection limit at high temperatures. This effect causes a systematic under-



estimation of the treatment effect at warm temperatures. To avoid this bias, while preserving the available information about the treatment effect at warm temperatures, an extrapolation of treated spectra beyond their warmest reported sampling temperature is performed. A straight line is fitted to the four warmest readings of the treated spectrum in $\log_{10} \text{INP}(T)$, and is accepted only if it describes the warm tail well ($R^2 > 0.9$) and its slope is steeper than half the Cooper (1986) slope (0.066 dex per $^{\circ}\text{C}$ of supercooling), to avoid extrapolating warm-tail plateaus that survived the clipping described in Sect. 3.2. The accepted line is evaluated at the temperature of each untreated (heat-treated) observation warmer than the warmest reading of the heat-treated (peroxide-treated) spectrum from the same filter. Extrapolation is capped at 4°C beyond the warmest reading of a spectrum, in an attempt to balance information extraction and uncertainty divergence. In the few cases where the heat-treated (peroxide-treated) spectrum of a filter terminates at a warmer temperature than the untreated (heat-treated) spectrum, an analogous extrapolation is applied in the opposite direction. Figure 3b shows an example filter with accepted warm-tail extrapolations of both the heat- and peroxide-treated spectra. Multiple values of the number of points applied in the extrapolation fit (here set to four points) and the maximum extrapolation distance along the temperature axis (here set to 4°C) have been tested. These tests reveal that the conclusions drawn are not sensitive to these parameters. The exact values were set to balance information loss and approximation uncertainty. Fitting the extrapolation line to a small number of readings and allowing long-range extrapolation conserves more information, reduces warm-tail bias, and increases uncertainty; conversely, fitting to a large number of readings while allowing only minimal extrapolation does the opposite.

3.5 Adjusting statistics to a reference sampling routine

At a fixed site and activation temperature, the ambient INP concentration varies continuously in time. This instantaneous concentration cannot be observed directly because every INP measurement collects air over a finite time interval and reports the mean concentration over that interval. Such a measurement is referred to as a *sample*, and its averaging interval as its *exposure time*. For IS data, the exposure time is the time during which ambient air is pumped through the respective filter; for GloPINE data, it is the 6 h averaging window. A campaign is a sequence of samples, and its *campaign time span* runs from the start of the first sample's exposure to the end of the last sample's exposure. The qualifier *measured* on a statistic (e.g., the measured variance) denotes that statistic computed from the samples of a campaign, as opposed to the corresponding statistic of the continuous atmospheric INP concentration.

Across the campaigns in the dataset, the exposure time ranges from less than an hour to 2 weeks, and the campaign time span ranges from days to several years. A campaign's measured statistics depend on its sampling routine (L_0, R_0), where L_0 denotes the exposure time and R_0 the campaign time span. For instance, a campaign with a long exposure time misses short-term variability, whereas a campaign with a short time span misses long-term variability. Consequently, for any campaign with $L_0 > 0$ and finite R_0 , the measured variance is expected to underestimate the variance of the continuous atmospheric INP concentration.

Because INP concentrations vary by several orders of magnitude in space and time, all analyses use $\log_{10}$ INP. Since each sample is a linear average and the logarithm is a concave function, Jensen's inequality implies that the logarithm of a long-



exposure sample is equal to or larger than the mean of the logarithms of shorter samples covering the same period. The measured $\log_{10}$ -space mean therefore increases with L_0 .

To remove these sampling effects and compare campaigns on an equal footing, all statistics in Sects. 4.1 and 4.2, Tables 2 and 3, and Figs. 5 and 6 are adjusted to a common reference sampling routine $(L_{\text{ref}}, R_{\text{ref}})$, consisting of consecutive samples with an exposure time of $L_{\text{ref}} = 24$ h over a campaign time span of $R_{\text{ref}} = 365$ d. The remainder of this section describes that adjustment.

3.5.1 Formulas for adjusting between different sampling routines

The adjustment maps the measured statistics of a campaign, sampled with (L_0, R_0) , to the statistics expected had the same site been sampled with $(L_{\text{ref}}, R_{\text{ref}})$. Four measured statistics are adjusted in this way, namely the mean, the variance, the normalized accumulated variance, and the autocorrelation. All four are computed from the logarithms of the sample concentrations, $y_i = \log_{10} \text{INP}_i$, where INP_i is the concentration reported by sample i . Let $\bar{y}$ and s^2 denote the measured mean and variance of y_i over a campaign. For a time window of width Δt , let $w(\Delta t)$ denote the variance of the samples y_i within the window, averaged over all such windows in the campaign. This quantity is referred to as the measured *accumulated variance*, and $f(\Delta t) = w(\Delta t)/s^2$ as the measured *normalized accumulated variance*. Let $c(\tau)$ denote the measured *autocovariance*, i.e., the covariance of samples y_i and y_{i+k} separated by a time lag $\tau = kL_0$, and $r(\tau) = c(\tau)/s^2$ the measured *autocorrelation*.

The measured statistics of a campaign depend on its start time, e.g., October 1, 2026, at 8 a.m. or February 22, 2024, at 2 p.m. This dependence is strongest when the INP concentration varies on timescales longer than the campaign time span, for example through the seasonal cycle or irregular dust events. The start time then determines which part of this variability the campaign captures. Two otherwise identical campaigns with different start times can therefore yield different measured statistics. A measured statistic is thus one realization among many possible ones. Its expected value is defined as the average over all possible start times at the site and thus depends only on the site and the sampling routine (L_0, R_0) . Because only the expected values are uniquely determined by the site and the sampling routine, the adjustment is derived for expected values rather than for measured statistics.

In analogy to the measured accumulated variance $w(\Delta t)$, the accumulated variance $V(\Delta t)$ is defined as the variance of the continuous log concentration $y(t) = \log_{10} \text{INP}(t)$ within a time window of width Δt , averaged over all window positions within a time interval Θ , that is much longer than any timescale on which $y(t)$ varies; i.e., unlike $w(\Delta t)$, it is not confined to a campaign time span. The adjustment of all four statistics can be formulated in terms of the single site-property $V(\Delta t)$. A central result to the derivations that follow is the relation that links $V(\Delta t)$ to the expected measured variance of a campaign with sampling routine (L_0, R_0) ; i.e.,

$$\langle s^2 \rangle = V(R_0) - V(L_0), \quad (1)$$

where $\langle \cdot \rangle$ denotes the expected value when averaging over all campaign start times. The expression is formally derived in Sect. B2. Equation (1) intuitively states that the expected measured variance $\langle s^2 \rangle$ equals the expected variance $V(R_0)$ within the campaign time span minus the expected variance $V(L_0)$ within the exposure time. Appendix B uses Eq. (1) and the



definitions of $\bar{y}$, s^2 , $f(\Delta t)$, and $r(\tau)$ to derive formulas for converting the expected values $\langle \bar{y} \rangle$ and $\langle s^2 \rangle$, the expected normalized accumulated variance $F(\Delta t) \equiv \langle w(\Delta t) \rangle / \langle s^2 \rangle$, and the expected autocorrelation $\rho(\tau) \equiv \langle c(\tau) \rangle / \langle s^2 \rangle$ from the native sampling (L_0, R_0) to the reference sampling $(L_{\text{ref}}, R_{\text{ref}})$. The derived formulas are Eqs. (B12), (B34), (B16), and (B26) of Appendix B. Applied to a single campaign, these formulas read

$$s_{\text{ref}}^2 = \frac{s_0^2}{A}, \quad \bar{y}_{\text{ref}} = \bar{y}_0 - \kappa B s_{\text{ref}}^2, \quad f_{\text{ref}}(\Delta t) = A f_0(\Delta t) + B, \quad r_{\text{ref}}(\tau) = A r_0(\tau) + S, \quad (2)$$

where

$$A = \frac{V(R_0) - V(L_0)}{V(R_{\text{ref}}) - V(L_{\text{ref}})}, \quad B = \frac{V(L_0) - V(L_{\text{ref}})}{V(R_{\text{ref}}) - V(L_{\text{ref}})}, \quad S = \frac{V(R_{\text{ref}}) - V(R_0)}{V(R_{\text{ref}}) - V(L_{\text{ref}})}, \quad \text{and} \quad \kappa = \ln(10)/2. \quad (3)$$

Here, the subscript 0 denotes a statistic measured with the campaign sampling routine (L_0, R_0) , and the subscript ref denotes the same statistic adjusted to the reference sampling routine $(L_{\text{ref}}, R_{\text{ref}})$. The exposure time L_0 is treated as a constant shared by all samples of a campaign, and Sect. 3.5.4 describes how L_0 is set for campaigns with varying exposure times.

The accumulated variance $V(\Delta t)$ is a property of the continuous INP concentration and thus differs from its measured analogue $w(\Delta t)$; i.e., $w(\Delta t)$ is a measured quantity while $V(\Delta t)$ is unknown because the continuous INP concentration is unavailable. However, $V(\Delta t)$ enters the three constants A , B , and S only as ratios of differences. These constants are thus unchanged when $V(\Delta t)$ is replaced by $\alpha V(\Delta t) + \beta$ for any constants α and β . The adjustment procedure therefore requires only the *shape* of $V(\Delta t)$, and the fundamental assumption is that this shape is the same at every site for a given activation temperature. Formally, the accumulated variances $V_p(\Delta t)$ and $V_q(\Delta t)$ of any two sites p and q are assumed to be related by

$$V_p(\Delta t) = \alpha_{pq} V_q(\Delta t) + \beta_{pq}, \quad (4)$$

where α_{pq} and β_{pq} are constants that depend on the two sites but not on Δt . Equation (4) is assumed separately for each activation temperature. Under Eq. (4), the constants A , B , and S of Eq. (3) are the same for every site at a given activation temperature.

3.5.2 Replacing $V(\Delta t)$ with a measurable normalized shape

Consider the exposure time L_0 and a normalization width $\Delta t_n > L_0$ and let

$$\mathcal{W}(\Delta t) \equiv \frac{V(\Delta t) - V(L_0)}{V(\Delta t_n) - V(L_0)}, \quad (5)$$

so that $\mathcal{W}(L_0) = 0$, $\mathcal{W}(\Delta t_n) = 1$, and $\mathcal{W}(\Delta t)$ is site-independent under the assumption of Eq. (4), but depends on L_0 and the choice of Δt_n . Equation (5) is of the form $\alpha V(\Delta t) + \beta$, so $\mathcal{W}(\Delta t)$ can replace $V(\Delta t)$ in Eq. (3) without altering the constants A , B , and S ; hence, the adjustments of Eq. (2) can be calculated from $\mathcal{W}(\Delta t)$ rather than $V(\Delta t)$, and $\mathcal{W}(\Delta t)$ can be measured, as shown next.

Let $w_p(\Delta t)$ denote the measured accumulated variance $w(\Delta t)$ of campaign p . Equation (1) holds for any set of consecutive samples, so for a window of width Δt filled with samples of exposure time L_0 , the expected value of $w_p(\Delta t)$ is

$$\langle w_p(\Delta t) \rangle = V_p(\Delta t) - V_p(L_0) = [V_p(\Delta t_n) - V_p(L_0)] \mathcal{W}(\Delta t) \Rightarrow \mathcal{W}(\Delta t) = \frac{\langle w_p(\Delta t) \rangle}{\langle w_p(\Delta t_n) \rangle}, \quad (6)$$



using Eqs. (5) and (1). By Eq. (6), $\mathcal{W}(\Delta t)$ is obtained from any campaign p as the expected value of the measured accumulated
355 variance $\langle w_p(\Delta t) \rangle$ divided by the constant $\langle w_p(\Delta t_n) \rangle$, where Δt_n is any chosen time interval $L_0 < \Delta t_n < R_0$. Typically,
only one campaign is available from each site, so the expected values $\langle w_p(\Delta t) \rangle$ and $\langle w_p(\Delta t_n) \rangle$ are unknown. To estimate
 $\mathcal{W}(\Delta t)$ it is assumed that the single-campaign quantity $w_p(\Delta t)/w_p(\Delta t_n)$ approximates $\mathcal{W}(\Delta t)$; i.e.,

$$\mathcal{W}(\Delta t) \approx \frac{w_p(\Delta t)}{w_p(\Delta t_n)}. \quad (7)$$

Because $\mathcal{W}(\Delta t)$ is site-independent under the assumption of Eq. (4), an improved estimate of $\mathcal{W}(\Delta t)$ is obtained by averaging
360 Eq. (7) over multiple campaigns at different sites. The curve $\mathcal{W}(\Delta t)$ as defined by Eq. (5) depends on L_0 and Δt_n , and thus
Eq. (7) must be averaged over campaigns with the same exposure time L_0 , and the same chosen value of Δt_n must be used
consistently.

3.5.3 Obtaining an approximation of $\mathcal{W}$

The GloPINE campaigns report INP concentrations from consecutive samples with an identical exposure time of 6 h and
365 include several campaigns with time spans exceeding 1 year. The consecutive sampling with a constant L_0 across all campaigns
in the GloPINE dataset and the inclusion of campaigns with long time spans make the campaigns of GloPINE ideal for
estimating $\mathcal{W}(\Delta t)$ with $L_0 = 6$ h over a wide range of window widths Δt . To define a universal $\mathcal{W}$, a normalization width Δt_n
must be chosen. Here, the normalization width $\Delta t_n = 7$ d is chosen as a compromise between the measurement noise in the
measured accumulated variance, which grows toward short widths, and the uncertainty of the average over windows, which
370 grows toward long widths as fewer independent windows fit into the campaigns.

To obtain the estimate of $\mathcal{W}(\Delta t)$ used for adjustment in the rest of the paper, $w_p(\Delta t)/w_p(7 \text{ d})$ is averaged over the 10
GloPINE campaigns that are judged representative (Sect. S3), span at least 1 month, take at least 80 % of their samples con-
secutively, and sample their campaign time span continuously rather than in a few separated blocks. The gray curves in Fig. 4
are $w_p(\Delta t)/w_p(7 \text{ d})$ of these 10 GloPINE campaigns, calculated using the INP concentrations at the reference temperature of
375 -23°C (Sect. 3.3). Three sources cause $w_p(\Delta t)/w_p(7 \text{ d})$ to differ between campaigns: (i) the campaigns contributing to Fig. 4
have blocks of several days to weeks without samples due to technical problems which influence $w_p(\Delta t)$; (ii) Eq. (7) is only
approximate because each site is only sampled during one campaign with a limited time span; and (iii) the assumption of a
universal shape of $V(\Delta t)$ formalized by Eq. (4) is an a priori assumption not expected to be exact.

The mean across the gray curves in Fig. 4 is plotted in blue and estimates the $\mathcal{W}(\Delta t)$ used in place of $V(\Delta t)$ in Eq. (3)
380 to calculate the constants A , B , and S that enter the adjustment formulas of Eq. (2) to adjust statistics derived from INP
concentrations at the reference temperature -23°C . The mean of the gray curves is only used directly in Eq. (3) when at least
five different campaigns (three at -20°C) provide an estimate of $\mathcal{W}$ for a given window width Δt . From the largest width
 Δt_{max} at which at least five curves contribute to the approximation of $\mathcal{W}$, and up to the reference campaign time span R_{ref} of
1 year, $\mathcal{W}$ is assumed to follow a straight line in $\log \Delta t$. The straight line is obtained by fitting to the mean of $w_p(\Delta t)/w_p(7 \text{ d})$
385 from 7 d up to Δt_{max} . In Fig. 4 this straight line is plotted in black for the reference temperature -23°C . As for the reference
temperature -23°C , the means of $w_p(\Delta t)/w_p(7 \text{ d})$ evaluated at reference temperatures -30 , -25 , and -20°C , shown as thin



colored lines in Fig. 4, are used to adjust statistics derived from INP concentrations at the respective reference temperatures. According to Fig. 4, $\mathcal{W}$ is almost identical at the evaluated temperatures (-30 , -25 , -23 , and -20°C), and it seems fair to assume that $\mathcal{W}$ is independent of temperature. However, this assumption is avoided because the results of Sects. 4.1 and 4.2 indicate that $\mathcal{W}$ does depend on temperature.

The curve $\mathcal{W}(\Delta t)$ is estimated from GloPINE campaigns only, but it is applied to all campaigns, including IS campaigns at other sites and with exposure times of 8 h to 2 weeks. Figure S1 tests whether the IS campaigns share the shape of the accumulated variance assumed in Eq. (4). The figure compares the normalized accumulated variance $f_{\text{ref}}(\Delta t)$ at -23°C of the 10 GloPINE and 11 IS *deployments* (unique combinations of campaign and platform, Sect. 2) of Fig. 6. All curves in Fig. S1 are adjusted to the reference sampling using Eq. (2) with $\mathcal{W}(\Delta t)$ at -23°C . If the IS campaigns followed a different shape, their adjusted curves would deviate systematically from those of the GloPINE campaigns. Instead, the two means agree within their 90 % confidence bands at every width Δt from about 5 days, the shortest width for which at least three IS campaigns contribute.

3.5.4 Implementation and limitations

Two additional choices complete the adjustment procedure. First, the variance is assumed not to accumulate further beyond 1 year, i.e., $V(\Delta t) = V(R_{\text{ref}})$ for $\Delta t > R_{\text{ref}}$, so that campaigns longer than 1 year are not adjusted for their campaign time span. This assumption is supported by the multi-year campaigns in the dataset, since a 1-year window captures their full variance to within a few percent. Second, the L_0 and R_0 used in Eq. (3) are defined for each campaign from the subset of its samples that report a value at the temperature in question: L_0 is the median exposure time of this subset, and R_0 is the time from the start of the subset's first sample to the end of its last. Eckermann-PrEliS, which consists of two austral summers a year apart, uses the summed length of its two periods for R_0 , so R_0 is about 4 rather than 14 months.

The formulas in Eq. (2) are derived for expected values, i.e., averages over all possible campaign start times, but they are applied to statistics measured during a single campaign with a single start time. A measured statistic deviates from its expected value under the native sampling by an amount specific to the campaign, for instance because the campaign fell in a calm or variable period. Section S6 relates the deviation of each measured statistic from its expected value under the native sampling to the deviation of the adjusted statistic from its expected value under the reference sampling. Concretely, Sect. S6 shows that the adjustments in Eq. (2) conserve the relative deviation of s^2 , $f(\Delta t)$, and $r(\tau)$, and conserve the absolute deviation of the mean except for a term proportional to the deviation of the measured variance from the expected variance at the native sampling; i.e., the adjustments fix the systematic sampling error of each statistic while preserving the deviation of the campaign from the expected statistics of the native sampling.

One caveat concerns campaigns that sample regularly but sparsely, such as the IS campaigns at the ARM sites SGP, OLI, and EPC, which collect about one 24 h filter per week. With this sampling routine, a window of width Δt has relatively few samples, and its mean deviates from the window's true mean because the reported mean absorbs part of the window variance that denser sampling would resolve. Therefore, the normalized accumulated variance $f_0(\Delta t)$ of such a campaign is artificially low. A numerical test, in which densely sampled GloPINE campaigns are reduced to one 24 h sample per week, shows that

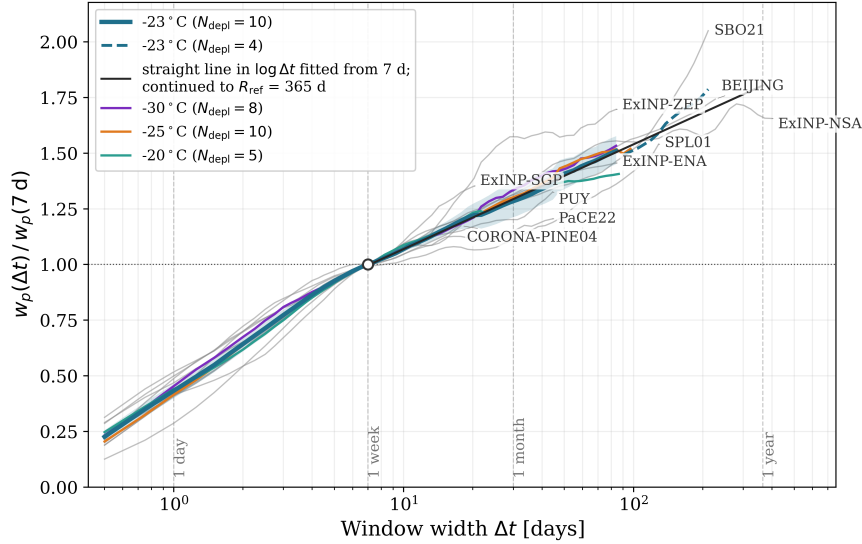


Figure 4. Estimates of $\mathcal{W}(\Delta t)$ (Eq. (5)) with $L_0 = 6$ h and $\Delta t_n = 7$ d; i.e., the site-independent shape of the accumulated variance $V(\Delta t)$ of the $\log_{10}$ INP concentration obtained from the 10 GloPINE campaigns that are judged representative (Sect. S3), span at least 1 month, take at least 80 % of their samples consecutively, and sample their campaign time span continuously rather than in a few separated blocks. Each gray line shows $w_p(\Delta t)/w_p(7 \text{ d})$ at the reference temperature -23°C (Sect. 3.3) for one campaign p (Eq. (7)), drawn up to half the campaign time span and labeled with the campaign name. The thick blue line is the mean of the gray lines, drawn solid where at least five campaigns contribute and dashed where at least three contribute; the shading is the 90 % confidence interval obtained from bootstrap resampling of the 10 campaigns (Wilks, 2006). The black line is a straight line in $\log \Delta t$ fitted to the mean of the gray curves between 7 d and $\Delta t_{\max}$, the largest width with at least five contributing campaigns, and extended to $R_{\text{ref}} = 365$ d. The thin colored lines show the corresponding means at -30 , -25 , and -20°C , drawn where at least five campaigns contribute (three at -20°C). In the legend, N_{depl} is the number of *deployments* (unique combinations of campaign and platform) contributing to the respective mean, which here equals the number of campaigns.

$f_0(\Delta t)$ is underestimated by $\approx 20\%$ at the shortest width evaluated, where a window holds three samples. As the window width grows, $f_0(\Delta t)$ converges to the value of the densely sampled campaign.

The curves of $\mathcal{W}(\Delta t)$ estimated from GloPINE campaigns with $L_0 = 6$ h (Fig. 4) are zero at $\Delta t = 6$ h by definition and has no information on variance accumulation for $\Delta t < 6$ h. Therefore, campaigns with $L_0 < 6$ h, which include all aircraft and
425 balloon deployments, cannot be adjusted with $\mathcal{W}$ estimated from GloPINE campaigns; hence, these are not included in the statistics of Tables 2 and 3, drawn unadjusted and semi-transparent in Fig. 5, and excluded from Fig. 6.

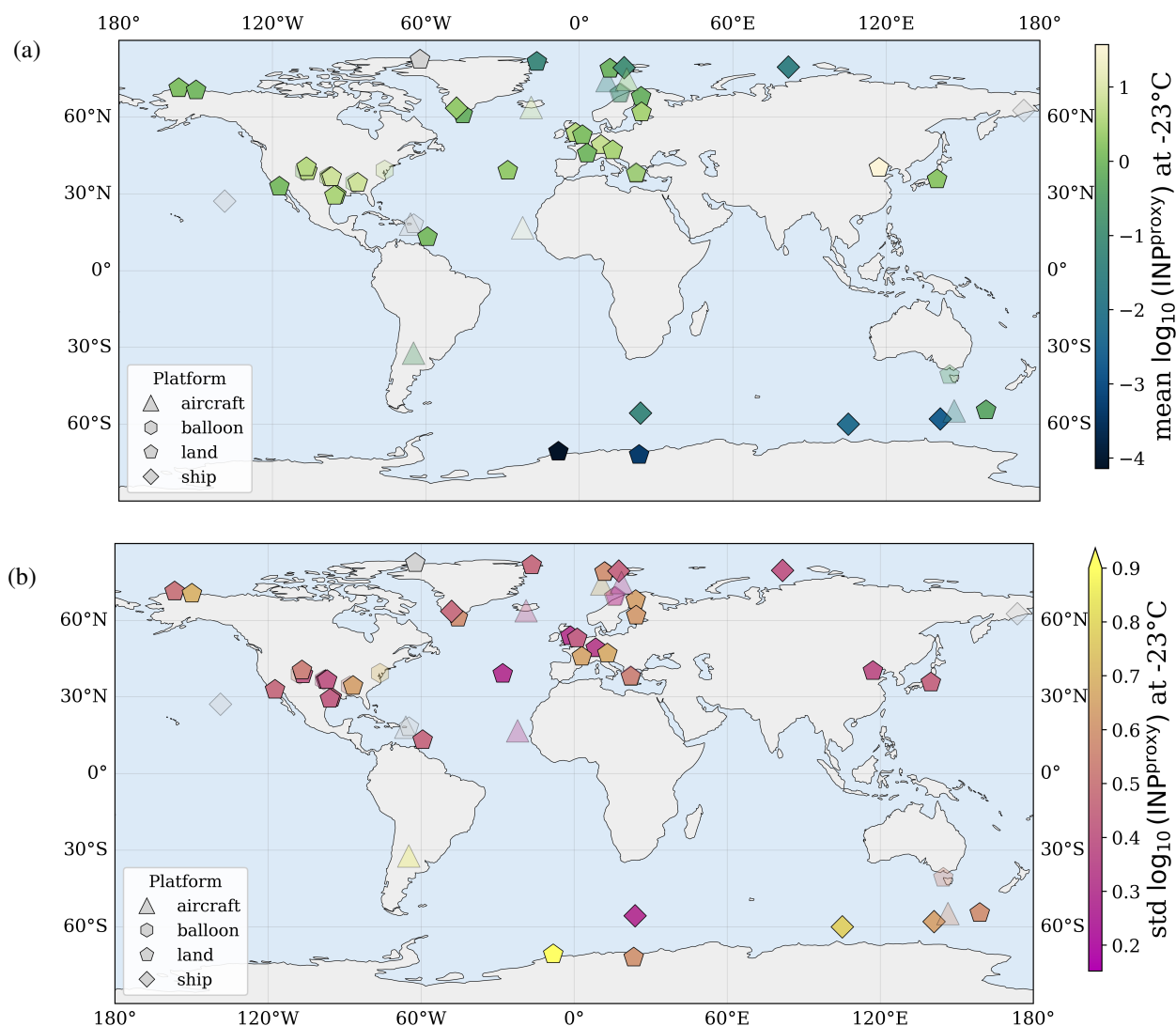


Figure 5. (a) The $\log_{10}$ -space mean of observed INP concentrations reported per *deployment*, i.e., grouped by each unique combination of campaign and platform. Each spectrum has been evaluated at a common reference temperature $T_{\text{ref}} = -23^\circ\text{C}$ using the estimation hierarchy described in Sect. 3.3, which ensures that each sample (filter or 6 h PINE window) only contributes one value to the mean. Deployments with fewer than 10 samples with a value at T_{ref} are shown in gray and are excluded from the statistics provided in Sect. 4.1. The same applies for deployments judged not to be representative of ambient conditions at their site (Sect. S3), which are drawn at half opacity. (b) Same as (a) but with the color mapped to the within-deployment standard deviation of $\log_{10}(\text{INP}_{\text{proxy}})$ at -23°C , thus reflecting the within-deployment variability in INP loading. The mean (a) and spread (b) of each deployment are adjusted to the reference sampling routine with an exposure time of 24 h and a campaign time span of 365 d, as described in Sect. 3.5. Balloon and aircraft measurements, whose exposure times are shorter than 6 h, cannot be adjusted (Sect. 3.5.4) and are shown unadjusted, also at half opacity. Statistics over the deployments represented by opaque markers are displayed in Table 2.



4 Results and Discussion

4.1 Observed INP concentrations and their variation

This section provides a global view of INP concentrations and their local variability. Based on INP measurements from more than 50 locations and a wide range of aerosol regimes, Fig. 5a provides a global overview of INP loading based on the dataset described in Sect. 2 and used throughout Sect. 4. The INP concentrations in Fig. 5a span more than four orders of magnitude, with the lowest concentrations on the Antarctic ice sheet, followed by pristine maritime environments in the Southern Ocean and the Arctic. Among the *deployments* (unique combinations of campaign and platform, Sect. 2) judged representative in Sect. S3, the one with highest observed average INP loading at an activation temperature of -23°C is the GloPINE campaign near Beijing, yielding a geometric mean of 1.57 dex, equivalent to 37.4 INP per liter at standard pressure (1 bar) and temperature (0°C), which is written 37.4stdL^{-1} from hereon. In comparison, the Antarctic campaigns report average values of less than 10^{-3}stdL^{-1} at an activation temperature of -23°C .

Figure 5b is analogous to Fig. 5a, but maps the within-deployment spread, i.e., the standard deviation of $\log_{10}(\text{INP}^{\text{proxy}}(-23^{\circ}\text{C}))$ of each deployment. Whereas Fig. 5a shows a clear geographical pattern with lower INP concentrations toward the poles, no such meridional dependence is evident from Fig. 5b. The observations of Fig. 5 are quantified in Table 2, providing the mean and spread of the INP concentration at -23°C , across five regional groups (*Antarctica*, *Southern Ocean*, *Pristine Arctic*, *Contiguous US*, and *Central Europe*) and the whole *Global* set. The table clearly outlines the well-established ordering of geographical zones by average INP concentrations, with Antarctica having the lowest INP concentrations, followed by the Southern Ocean and the high Arctic. The campaign MICRE is an outlier in the Southern Ocean, with INP concentrations significantly higher than the three ship campaigns in the group, thus inflating the group's spread. The high values likely reflect locally sourced aerosols from Macquarie Island, but investigating the exact cause of the discrepancy is beyond the scope of this work.

Except for the Antarctic campaigns, the regional groups of Table 2 all feature a within-deployment spread of around 0.5 dex. Despite the groups being characterized by wildly differing environmental characteristics with equally different INP populations – some locally sourced, some having drifted from afar – the scale of variation in INP concentration at -23°C appears very similar.

A possible explanation is provided by the theory of successive random dilution (Ott, 1990), which has been invoked to explain the lognormal frequency distributions of INP concentration observed across a wide range of environments (Welti et al., 2018, 2020; Schrod et al., 2020). In this framework, an air mass leaves an INP source with some initial INP concentration and is then diluted in a series of independent mixing events during transport. Ott (1990) shows that the log concentration at the measurement site, $y = \log_{10} \text{INP}$, is the sum of the log concentration at the source, y_{src} , and the logarithms of the dilution factors of the mixing events. Let the sum of the logarithms of the dilution factors be denoted y_{dil} , so that

$$y = y_{\text{src}} + y_{\text{dil}}, \quad (8)$$



where $y_{\text{dil}} \leq 0$ because the dilution factors are in the range $(0, 1]$. Given that y_{src} and y_{dil} are independent, the within-
 460 deployment variance of y splits into a source term and a dilution term, $s^2 = s_{\text{src}}^2 + s_{\text{dil}}^2$. The source term s_{src}^2 is the variance of
 y_{src} among the air masses reaching the site during the deployment; it reflects both changes in emissions from a given source
 over time and alternation between source regions. The dilution term s_{dil}^2 is the variance of the accumulated log dilution y_{dil}
 among the air masses reaching the site. Each mixing event lowers y_{dil} of an air mass by a random amount, and these random
 decrements add up along the way, so that air sampled far from its source is both more diluted and more variably diluted than air
 465 sampled close to it (Ott, 1990). By Eq. (8), stronger sources (a higher mean y_{src}) raise $\bar{y}$ without changing s^2 (Ott, 1990; Welti
 et al., 2018), more variable sources (a larger s_{src}^2) increase s^2 , and more dilution (more mixing events) lowers $\bar{y}$ and increases
 s^2 .

In this framework, the more than four orders of magnitude separating the deployment means of Table 2 are compatible with
 within-deployment spreads of about 0.5 dex in every region except Antarctica. Differences in source strength between regions
 470 change $\bar{y}$ but not s^2 . The similar spreads instead indicate that $s_{\text{src}}^2 + s_{\text{dil}}^2$ is similar across regions, i.e., that source variability
 and dilution variability are similar, or differ in compensating ways.

Antarctica deviates from the other groups in both mean and spread. The mean INP concentration of the Antarctic campaigns
 is two orders of magnitude below that of the Southern Ocean and four orders of magnitude below those of the continental
 groups. At the same time, the average within-deployment spread of the Antarctic campaigns is 0.82 dex; i.e., the largest of all
 475 groups. The Antarctic ice sheet is the most remote of the sampled environments from any INP source, and its extremely low
 mean reflects strong dilution, perhaps combined with weak sources, e.g., marine organic aerosol from the Southern Ocean.
 Since more dilution also raises s_{dil}^2 , the Antarctic spread is attributed to dilution rather than to more variable sources.

To further investigate the inter-regional similarity of within-deployment spread in INP concentration, Table 3 reports the
 average within-deployment spread across regional groups at three different temperatures, only including deployments that are
 480 representative according to the evaluation in Sect. S3 and have at least 10 samples with a value at *all* three temperatures, i.e.,
 -25°C , -20°C , and -15°C . As for Table 2, each spread is adjusted to the reference sampling (Sect. 3.5) before averaging
 across deployments. Across all groups with at least three contributing deployments (Southern Ocean, Pristine Arctic, Contigu-
 ous US, and the whole globe), the within-deployment spread of $\log_{10}$ INP becomes larger at higher temperatures, growing
 monotonically from about 0.4 dex at -25°C to values of 0.6 to 0.7 dex at -15°C , while remaining similar across the three
 485 regions at every fixed temperature.

Because the spread remains similar across regions at every temperature, the variance added toward warmer temperatures
 is also similar across regions, pointing to a change in the INP population common to the sampled environments. The most
 pronounced such change is the increasing fraction of biological INPs toward warmer temperatures, found in most campaigns
 of the dataset that include treated samples (Sect. 4.3) and in several earlier studies (e.g., Zhou et al., 2026; Lata et al., 2026;
 490 Testa et al., 2021); hence, the growth in spread is attributed to the increased biological contribution at higher activation tem-
 peratures. At the coldest temperatures, a relatively large and diverse fraction of the aerosol acts as INPs, causing many sources
 to contribute to y_{src} of an air mass, making the fluctuations of these sources partly average out. Toward warmer temperatures,
 the active population narrows to rare, predominantly biological INPs from episodic and spatially patchy sources (Sect. 4.3),



Table 2. Regional summary of Fig. 5. Each *deployment*, i.e., each unique combination of campaign and platform, contributes its mean (markers of Fig. 5a) and standard deviation (markers of Fig. 5b) of $\log_{10}(\text{INP}^{\text{PROXY}}(-23^\circ\text{C}))$, and every deployment in a region is weighted equally. N is the number of contributing deployments. Only deployments judged representative (Sect. S3) and having at least 10 samples with a value at -23°C are included, so the dimmed and grayed markers of Fig. 5 contribute nothing here. Every deployment mean and spread is adjusted to a common exposure time of 24 h and a common campaign time span of 365 d as described in Sect. 3.5, and deployments for which that adjustment cannot be applied are excluded. Uncertainties are one standard deviation over 1000 bootstrap resamplings of the contributing deployments. All quantities are in dex. The groups are *Antarctica*: Eckermann-NM, Eckermann-PrElis. *Southern Ocean*: ACE, CAPRICORN_II, MARCUS, MICRE. *Pristine Arctic*: ExINP-NSA, ExINP-ZEP, Li-2025, OLI, PS106, Sze-2023. *Contiguous US*: BNF, EPC, ExINP-SGP, GUC, HOU-M1, HOU-S3, SGP, SPL01. *Central Europe*: CALISHTO, CORONA-PINE04, CountIce-pt1, CountIce-pt2, PUY, SBO21. *Global* includes every contributing deployment; that is, the five groups above together with BEIJING, Bergner-2026, ExINP-ENA, Harrison-2022, M-Phase, PaCE22, Schneider-2021, Tobo-2020, WAO. Table A1 summarizes the campaign details.

Region	N	Deployment mean, Fig. 5a		Within-deployment spread, Fig. 5b	
		Mean	Spread	Mean	Spread
Antarctica	2	-3.74 ± 0.27	0.56 ± 0.28	0.82 ± 0.15	0.32 ± 0.16
Southern Ocean	4	-1.67 ± 0.46	1.04 ± 0.29	0.57 ± 0.09	0.21 ± 0.07
Pristine Arctic	6	-0.65 ± 0.27	0.72 ± 0.13	0.51 ± 0.04	0.11 ± 0.03
Contiguous US	8	0.57 ± 0.09	0.29 ± 0.07	0.46 ± 0.03	0.09 ± 0.03
Central Europe	6	0.42 ± 0.11	0.30 ± 0.08	0.49 ± 0.07	0.18 ± 0.03
Global	35	-0.24 ± 0.20	1.25 ± 0.21	0.51 ± 0.03	0.16 ± 0.02

so few sources contribute to y_{src} , and s_{src}^2 increases. Transport-induced dilution cannot cause the increase. Primary biological particles are short-lived compared to other INP types, so the biological INPs sampled at a site originate from sources closer to the site than the INPs active at colder temperatures. Air from nearby sources is less variably diluted, so s_{dil}^2 decreases toward warmer temperatures and counteracts the increase in spread rather than contributing to it.

The source variability of biological aerosols driving increased within-deployment spreads can operate on both short and long timescales. On timescales of days, the heat-labile contribution to Southern Ocean INP populations varies widely (McCluskey et al., 2018a), and at terrestrial sites rainfall releases biological INPs in bursts that raise warm-temperature INP concentrations within hours (Huffman et al., 2013; Prenni et al., 2013) and affect week-to-week variation (Conen et al., 2017). On longer timescales, the seasonal cycles of INP concentration, commonly reported to be strongest for warm-temperature biological INPs (Schneider et al., 2021; Zhou et al., 2026; Tobo et al., 2020, 2024), slowly shift the mean of y_{src} over the year; hence, a deployment spanning several seasons samples a mixture of distributions with different means and accumulates additional spread.

Measurement uncertainty adds an instrumental contribution to s^2 on top of s_{src}^2 and s_{dil}^2 , so the spreads reported in this section are upper bounds on the atmospheric variability. For the GloPINE campaigns, Sect. S7 bounds this contribution with



Table 3. The within-deployment spread of $\log_{10}(\text{INP}^{\text{proxy}}(T_{\text{ref}}))$ compared at three reference temperatures: -25°C , -20°C , -15°C . Only *deployments* (unique combinations of campaign and platform) judged representative (Sect. S3) and having at least 10 samples with a value at *every* one of those temperatures are included, so the columns describe one identical set of deployments, with the number of deployments included in each group denoted N . Regions contributing fewer than two deployments are omitted. Every spread is adjusted to a common exposure time of 24 h and a common campaign time span of 365 d as described in Sect. 3.5. At -15°C $\mathcal{W}(\Delta t)$ at -20°C is used, because too few GloPINE campaigns report values at this temperature for $\mathcal{W}(\Delta t)$ at -15°C to be estimated within accepted uncertainty. Uncertainties are one standard deviation over 1000 bootstrap resamplings (Wilks, 2006) of the contributing deployments: the same resampling is used in every column. All quantities are in dex. The groups are *Southern Ocean*: ACE, MARCUS, MICRE. *Pristine Arctic*: ExINP-NSA, Li-2025, OLI, PS106, Sze-2023. *Contiguous US*: BNF, EPC, ExINP-SGP, GUC, HOU-M1, HOU-S3, SGP. *Global* includes every contributing deployment; that is, the groups above together with Bergner-2026, Eckermann-NM, M-Phase, SBO21, Schneider-2021, Tobo-2020, WAO. Table A1 summarizes the campaign details.

Region	N	Mean within-deployment spread, Fig. 5b		
		-25	-20	-15
Southern Ocean	3	0.41 ± 0.09	0.55 ± 0.13	0.62 ± 0.04
Pristine Arctic	5	0.45 ± 0.05	0.56 ± 0.04	0.65 ± 0.06
Contiguous US	7	0.42 ± 0.02	0.54 ± 0.05	0.69 ± 0.07
Global	22	0.46 ± 0.02	0.56 ± 0.03	0.64 ± 0.04

the accumulated variance of Fig. 4 and shows that measurement noise accounts for at most 4 % of the spread at -23°C under the reference sampling. An analogous bound for the IS campaigns would require consecutive short-exposure samples, which these campaigns rarely provide. The noise of an IS sample is nevertheless expected to be smaller than that of a 6 h PINE sample: the temperature uncertainty is $\approx 0.5^{\circ}\text{C}$ for IS vs. $\approx 1.0^{\circ}\text{C}$ for PINE, and the much larger air volume sampled per IS filter reduces the counting uncertainty (Sect. 2.1). The spreads of Table 2 are therefore dominated by real atmospheric variation for both instruments. Toward warmer temperatures, however, the counting uncertainty grows as INPs become sparse (Sect. 2.1), so a measurement contribution to the increase in spread with temperature in Table 3 cannot be excluded. However, the clipping of Sect. 3.2 limits this contribution, and discarding INP concentrations at or below 10^{-11} std L^{-1} (Sect. 3.1) reduces the observed spread near the detection limit, i.e., at warmer temperatures.

4.2 Temporal structure of variations in INP concentration

As shown in the previous section, INP concentration varies widely, with a spread of roughly half an order of magnitude at -23°C (Table 2). The fluctuating nature of INP concentrations is well established in the INP literature, but the timescales on which variations occur are still not well quantified. Multiple studies report seasonal variation of the INP concentration (Schneider et al., 2021; Zhou et al., 2026; Tobo et al., 2020; Wex et al., 2019; Creamean et al., 2025; Tobo et al., 2024), while Schrod et al. (2020) use nearly 2 years of condensation- and deposition-mode INP measurements at four globally distributed



stations to argue that the short-term variation of INP concentrations dominates the total variance. Furthermore, Canzi et al. (2025) couple Saharan dust events at the alpine French research station Puy de Dôme directly to increased INP concentration and show that these dust events usually last only a few days or less. These studies show that processes with characteristic timescales ranging from a single day to several months shape INP concentrations, e.g., local variations in wind or precipitation, large-scale atmospheric circulation, and seasonal variations in snow cover and biological activity. However, the literature has not quantified the distribution of observed variance in INP concentrations across these timescales. This study quantifies this using a time-series analysis of the campaigns with a long time span in the presented dataset. Figures 6a and b show the results of this analysis, presenting two complementary views of the characteristic timescales of variation.

All deployments in Fig. 6 are adjusted to the reference sampling (Sect. 3.5), which removes the systematic effects of their differing exposure times and campaign time spans. Figure 6a shows the autocorrelation $r_{\text{ref}}(\tau)$ of $\log_{10}(\text{INP}^{\text{proxy}})$ against the time lag τ for the individual deployments at -23°C , their mean, and the deployment means at -30 , -25 , and -20°C . To compare the autocorrelation between deployments with different sampling routines, while resolving the autocorrelation of each deployment at the best obtainable accuracy, the lags between the samples of each deployment are sorted into a tailored set of bins. The bins are centered on the lags at which the sampling schedule concentrates the sample pairs (Sect. S4).

The autocorrelation of Fig. 6a illustrates the rate at which the time series loses memory of its current state, with full decorrelation occurring when the autocorrelation reaches zero. The mean autocorrelation at -23°C drops from 0.62 at a lag of 1 day to 0.31 at 1 week, so decorrelation is rapid in the first few days. It then decays slowly, remains significantly positive for lags up to about 2 months apart from a brief dip near 35 days, and reaches zero at about 3 months. Thus, while most of the memory is lost within a few days, some is retained for up to about 3 months. The autocorrelation is significantly negative at lags around half a year and rises again to about 0.08 near 1 year. This minimum and the subsequent rise reflect seasonal dependence in INP concentration, i.e., an annual cycle, although only five deployments are long enough to contribute at lags near 1 year.

A complementary view is provided by Fig. 6b, which shows the normalized accumulated variance $f_{\text{ref}}(\Delta t)$ (Sect. 3.5) for the same deployments and temperatures, i.e., the fraction of the variance under the reference sampling that is found within windows of width Δt . The windows slide through each deployment, starting at each sample in turn, and a window contributes only if it lies within the campaign time span of the deployment and holds at least three samples. For the mean at -23°C , a window of 1 week holds about 45 % of the variance, a window of 1 month about 65 %, and a window of 3 months about 80 %. Thus, almost half of the variance of daily samples resides on timescales of a week or shorter, and the rest accumulates steadily across weeks and months without a plateau, indicating that no single dominant timescale exists. About a fifth of the variance resides at timescales longer than 3 months, in line with the annual cycle indicated by Fig. 6a.

The means at the other temperatures in Fig. 6 are computed over different sets of deployments, since not every deployment carries values at every temperature. At -30°C , 12 deployments contribute, most of them GloPINE campaigns, and at -20°C , 15 deployments contribute, most of them IS campaigns, whereas the set at -25°C is nearly identical to that at -23°C . At all four temperatures, the autocorrelation is significantly negative at lags around half a year. The means at -30 , -25 , and -23°C largely agree within their confidence bands, which are not shown for -30 and -25°C . At widths of 1 to 3 months, the normalized accumulated variance at -30°C lies below that at -23°C , for example 0.70 against 0.80 at 3 months, but the same

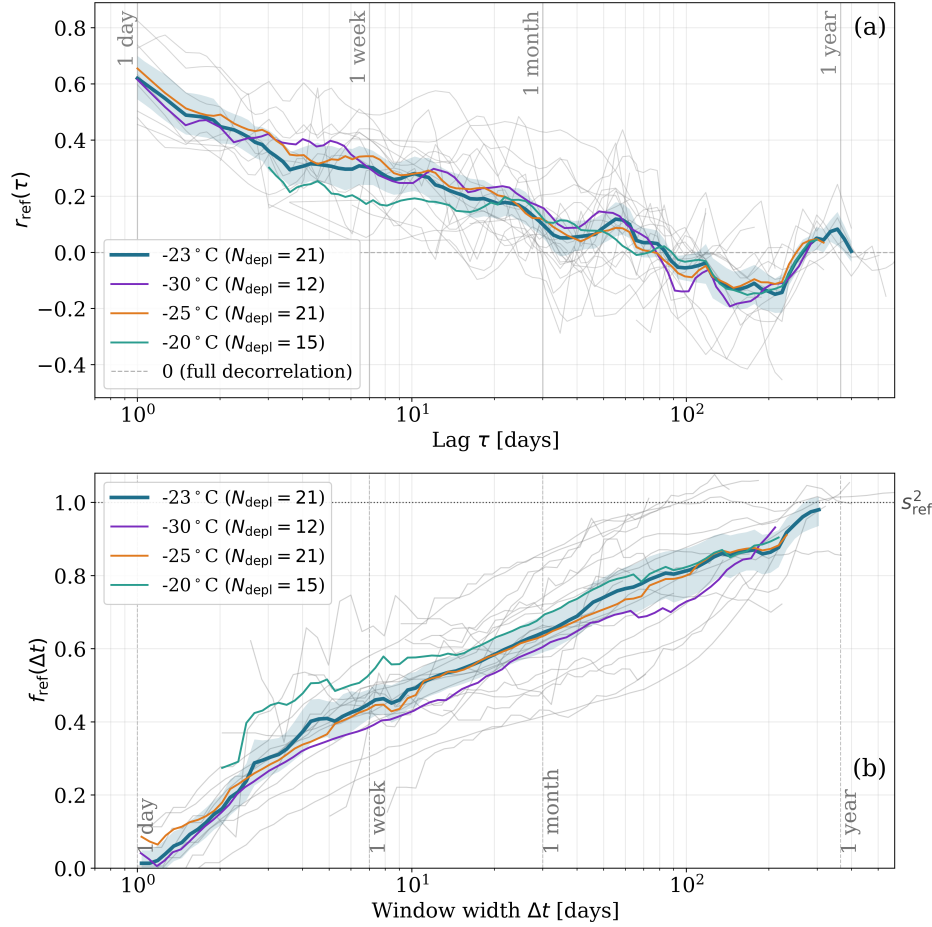


Figure 6. Temporal structure of the INP anomaly of the 21 *deployments* (unique combinations of campaign and platform) that span at least 90 days, sample their campaign time span continuously rather than in a few separated blocks, and are considered representative of the local conditions (Sect. S3). Ordered from north to south, these are Sze-2023, ExINP-ZEP, ExINP-NSA, OLI, PaCE22, Schneider-2021, CountIce-pt1, CountIce-pt2, SBO21, PUY, SPL01, BEIJING, ExINP-ENA, GUC, SGP, Tobo-2020, EPC, HOU-M1, MICRE, MARCUS, and Eckermann-NM. All curves are adjusted to the reference sampling routine with an exposure time of 24 h and a campaign time span of 365 d (Sect. 3.5), so lags and window widths below 1 d are not shown. (a) Autocorrelation $r_{\text{ref}}(\tau)$ of $\log_{10}(\text{INP}^{\text{PROXY}}(-23^\circ\text{C}))$ against the time lag τ . (b) Normalized accumulated variance $f_{\text{ref}}(\Delta t)$, i.e., the mean variance of the samples inside a window of width Δt relative to the variance under the reference sampling. In both panels, gray lines show the individual deployments, and the thick blue line is their mean, which gives every deployment equal weight and is drawn where at least five deployments contribute. Bands are 90 % confidence intervals from bootstrap resampling of the deployments (Wilks, 2006). Thin colored lines are the means at -30 , -25 , and -20°C , each adjusted with $\mathcal{W}(\Delta t)$ at its own temperature and computed over the deployments that carry values there. In the legend, N_{depl} is the number of deployments contributing to the respective mean.



deployments give 0.72 at -23°C , so this difference reflects the set of deployments rather than the temperature. At -20°C , by contrast, a larger share of the variance resides at short timescales than at -23°C for the same deployments, for example 0.45 against 0.33 within windows of 3 days, and the autocorrelation at a lag of 1 week is lower at -20° than -23° (0.17 against 0.25). This shift toward short-term variability at the warmer temperature is consistent with the growth of the within-deployment spread toward warmer temperatures (Table 3) and its attribution to episodic biological sources in Sect. 4.1. The difference lies, however, within the wide confidence band of the -20°C mean (not shown), which rests on fewer deployments.

4.3 Biogenic and organic INPs

As explained in Sect. 2.1, the solutions that result from washing filter samples can be subjected to two treatments. Heating the solution (95°C for 20 min) deactivates biogenic INPs, e.g., proteinaceous matter from bacteria, fungal spores, and vegetation (Hill et al., 2014, 2016). Adding peroxide (10 % H_2O_2 at 95°C for 20 min, followed by exposure to UV light) additionally removes non-biological organic INPs, primarily associated with soil organic matter (O’Sullivan et al., 2014). Comparison of treated and untreated spectra from the same filter thus reveals the contributions of these two aerosol classes to the INP population. Numerous studies use these techniques to investigate the nature of the INP population at a single site or region of interest. The current dataset allows for a broader investigation, as 15 of its campaigns include treated samples. This treated subset spans the high Arctic, the Southern Ocean, boreal forest in Finland, southern Greenland, and continental North and South America, and therefore covers most of the environmental range of the full dataset. However, the dataset does not feature treated samples from the major dust source and outflow regions of northern Africa, the Middle East, and central and eastern Asia. The composition statistics presented in this section therefore characterize the INP population of the sampled mid- and high-latitude environments, and the mineral fraction of the INP population may be larger in dust-laden regions than reported here.

Treatment results are usually displayed in one of two ways: either the spectra from a single filter are plotted at a time (e.g., McCluskey et al., 2018a; Lata et al., 2026; Gratzl et al., 2025; Schneider et al., 2021; McCluskey et al., 2018b), or the treatment-colored spectra from all sampled filters are shown together (e.g., Zhou et al., 2026; DeMott et al., 2025; Testa et al., 2021). Neither convention allows the treatment results from 15 campaigns to be summarized and compared compactly.

Therefore, an alternative to these established methods is applied, which pairs equal-temperature INP concentrations from spectra of different treatments derived from the same filter to track the ratios $\log_{10}(\text{INP}_{\text{heat}}/\text{INP}_{\text{untr}})$ and $\log_{10}(\text{INP}_{\text{perox}}/\text{INP}_{\text{heat}})$ as a function of temperature, as described in detail in Sect. 3.4. The ratio $\log_{10}(\text{INP}_{\text{heat}}/\text{INP}_{\text{untr}})$ gives the fraction of INP activity left after the biogenic INPs have deteriorated, while $\log_{10}(\text{INP}_{\text{perox}}/\text{INP}_{\text{heat}})$ provides the fraction that non-organic INPs contribute to the population of heat-stable INPs. Figure 7 shows box plots of the two ratios binned by activation temperature. The temperature at which the INP concentration exceeds the detection limit usually differs between treatments; for example, freezing occurs at warmer temperatures in an untreated INP water solution than when the solution receives heat treatment before measuring $\text{INP}(T)$, since heat treatment deteriorates warm biogenic INPs. Discarding samples where the untreated solution has detectable INP concentrations while the treated solution does not, systematically overestimates the INP concentration of treated samples, as explained in Sect. 3.4. The same section explains the spectrum extrapolation procedure



introduced to limit this bias. In Fig. 7, boxes are colored gray in proportion to the fraction of contributing data points that rely on extrapolation. Temperature bins with fewer than 20 ratios show the raw data points, with those relying on extrapolation colored gray. The extrapolation procedure allows $\log_{10}(\text{INP}_{\text{heat}}/\text{INP}_{\text{untr}})$ and $\log_{10}(\text{INP}_{\text{perox}}/\text{INP}_{\text{heat}})$ to be resolved at higher temperatures, but also adds uncertainty. The most conservative reading would therefore be focused on boxes where no, or only a few, measurements depend on extrapolation.

At cold temperatures, most $\log_{10}(\text{INP}_{\text{heat}}/\text{INP}_{\text{untr}})$ ratios are close to zero, i.e., heating has little effect on the INP concentration. Above a campaign-dependent transition temperature, typically between -20°C and -15°C , the heat-labile ratio declines increasingly steeply with temperature, marking the transition to a regime in which heat-labile INPs dominate. Such heat-labile dominance at high temperatures has been shown earlier across diverse environments (Zhou et al., 2026; Lata et al., 2026; Testa et al., 2021). A notable exception is Schneider-2021 (Fig. 7b) showing ratios $\log_{10}(\text{INP}_{\text{heat}}/\text{INP}_{\text{untr}})$ that remain below -1 dex down to -25°C rather than converging toward zero. Schneider et al. (2021) attribute this to a remote boreal forest INP population dominated by heat-labile biogenic material with a negligible mineral dust contribution; hence, the heat-stable floor that sets the cold end of the spectrum elsewhere is largely absent here. Also, MICRE and MARCUS (Figs. 7y, 7z, 7aa, and 7ab), both from the Southern Ocean, show temperature dependencies diverging from the remaining campaigns by lacking the characteristic increase in the heat-labile INP fraction with increasing temperature. Instead, $\log_{10}(\text{INP}_{\text{heat}}/\text{INP}_{\text{untr}})$ has a large spread at high temperatures with many cases of $\log_{10}(\text{INP}_{\text{heat}}/\text{INP}_{\text{untr}}) > 0$. As shown in Fig. S2, the clean-sector part of KCG, aimed at sampling air masses from the Southern Ocean (Sect. S3), also shows small heat-labile contributions within the resolved temperature range (up to -12°C). This behavior is consistent with McCluskey et al. (2018a), who found heat-stable organic material to be a common contributor to Southern Ocean INPs during CAPRICORN, with heat-labile INPs present in only some samples. Based on mesocosm experiments, McCluskey et al. (2018b) proposed that marine organic INPs comprise two populations: (i) dissolved organic carbon INPs, which are heat-stable and consist of molecules capable of nucleating ice that are coating or embedded in other particles, and (ii) particulate organic carbon INPs, which are heat-labile and consist of intact cells or cell fragments. The weak heat lability of MICRE and MARCUS thus points to a dominance of the former type of marine organic INPs. McCluskey et al. (2018b) also observed frequent cases of $\log_{10}(\text{INP}_{\text{heat}}/\text{INP}_{\text{untr}}) > 0$, i.e., an increase in INP concentration upon heating. They hypothesize that heating redistributes IN-active organic coatings or releases IN-active dissolved organic carbon from microbes. Interestingly, the results from the high Arctic ship campaign Li-2025 (Fig. 7a) show a clear difference from those obtained in the Southern Ocean, indicating a significant and temperature-dependent contribution to the INP population by heat-labile aerosols. In addition, Lata et al. (2026) recently documented a significant heat-labile component of the INP population in mid-winter samples from Ny-Ålesund in Svalbard, even in air masses originating from the polar icecap, illustrating the large difference in INP characteristics between the two polar regions.

The $\log_{10}(\text{INP}_{\text{perox}}/\text{INP}_{\text{heat}})$ ratios have a mean of -1.00 dex, a spread of 0.30 dex, and a weaker temperature dependence than $\log_{10}(\text{INP}_{\text{heat}}/\text{INP}_{\text{untr}})$. Across the temperature range that most campaigns cover densely, -28°C to -15°C , roughly 90 % of the IN-active particles that survive heat treatment are removed by oxidation in most larger campaigns (Fig. 7). Hence, heat-stable organic material seems to dominate the heat-stable INP population at every temperature resolved here, leaving only a small residual for the inorganic, presumably mineral, fraction.



Assuming that the heat and peroxide treatments do not reduce the ice nucleation potential of mineral INPs, the results of Fig. 7 are incompatible with several global modeling studies. The INP predictions of Vergara-Temprado et al. (2017) based on aerosols from GLOMAP (Mann et al., 2010) reproduce or exceed observed INP concentrations over the full temperature range using only the desert dust INP parameterization of Niemand et al. (2012), while overestimation is limited to temperatures below -25°C when the K-feldspar parameterization of Atkinson et al. (2013) is applied as the sole INP source. Similarly, the INP model of Chatziparaschos et al. (2023), based on aerosols from TM4-ECPL (Myriokefalitakis et al., 2015), modestly overestimates INP concentration below -25°C when using the K-feldspar and quartz parameterizations of Harrison et al. (2019) to predict INP concentration. These evaluations are based primarily on surface INP observations with a geographical footprint similar to the current dataset, making the comparison fair but not direct. If mineral dust accounts for only a tenth of the heat-stable INPs, as Fig. 7 suggests, then dust-only parameterizations like those of Vergara-Temprado et al. (2017) and Chatziparaschos et al. (2023) would be expected to underestimate INP concentration significantly. This discrepancy is confined to the environments represented in the treated subset, and dust-based descriptions may remain adequate in the dust-laden regions that the subset does not cover.

It is important to note that the ice nucleating ability of mineral INPs is not completely unaffected by heat and peroxide treatments. Daily et al. (2022) report that wet heating at 95°C leaves K-feldspar unaffected but deactivates quartz, plagioclase feldspar, and Arizona Test Dust by 2°C to 7°C in median freezing temperature. Furthermore, H_2O_2 has been found to deactivate microcline that heat alone did not affect (Peckhaus et al., 2016) and to reduce the activity of Arizona Test Dust (Perkins et al., 2020). Both effects lower the INP potential of mineral aerosols and bias the inferred mineral INP fraction downward. Therefore, both $\log_{10}(\text{INP}_{\text{heat}}/\text{INP}_{\text{untr}})$ (blue in Fig. 7) and $\log_{10}(\text{INP}_{\text{perox}}/\text{INP}_{\text{heat}})$ (orange in Fig. 7) should be considered lower bounds.

5 Conclusions

This study presents a dataset of immersion-mode INP concentration measurements. It comprises roughly 110 000 observations from 59 campaigns, obtained by filter-based ice spectrometry (Hiranuma et al., 2015) and using the PINE expansion chamber (Möhler et al., 2021). The observations are standardized to common reference activation temperatures, allowing consistent statistics to be derived for the concentrations of INPs active at these temperatures.

Campaigns collect each sample over periods ranging from hours to weeks, and they last from weeks to years. Both the collection time of a sample and the duration of a campaign affect the mean, the variance, and the autocorrelation of $\log_{10}$ INP, so these statistics are not directly comparable between campaigns that differ in sample collection time or campaign duration. This study derives formulas to convert each campaign's statistics to those expected for a common reference sampling, namely 24 h samples collected over 1 year (Sect. 3.5 and Appendix B). With this conversion, campaigns that sampled the atmosphere differently can be analyzed together. The statistics of INP variability presented here therefore draw on a far larger and more diverse set of observations than previous assessments, which were confined to single sites or to sets of campaigns sampled in the same way (Welti et al., 2018, 2020; Schrod et al., 2020).



660 At -23°C , mean INP concentrations vary by more than four orders of magnitude between *deployments* (unique combinations of campaign and platform, Sect. 2). By contrast, the within-deployment spread is about half an order of magnitude at -23°C . Because deployments differ widely in environment and INP sources, the uniform spread suggests that a large part of the local INP variation is set by processes common to all regions, such as the random dilution of the sampled air masses during transport (Ott, 1990; Welti et al., 2018). Toward warmer activation temperatures, the within-deployment spread grows in every
665 region with sufficient data, from about 0.4 dex at -25°C to 0.6–0.7 dex at -15°C , while remaining similar across regions. This growth reflects an increasing contribution of sources to the variance, driven by the increasing fraction of biological INPs toward warmer temperatures, which are short-lived and supplied by fewer and more episodic sources.

The timescales variation of INP concentrations are quantified at -23°C from the 21 deployments that are judged representative (Sect. S3), span at least 90 days, and sample continuously. The autocorrelation of daily mean values of $\log_{10}$ -space INP
670 concentration decays rapidly within the first few days, remains significantly positive for lags up to about 2 months, and turns negative at lags around half a year, which indicates an annual cycle. A window of 1 week holds about 45 % of the variance and a window of 1 month about 65 %. The remainder accumulates gradually across months, and about a fifth resides beyond 3 months. No dominant timescale emerges. This reconciles reports of pronounced seasonal cycles (Schneider et al., 2021; Tobo et al., 2024) with the dominance of short-term variation found by Schrod et al. (2020). The absence of a dominant timescale
675 also means that neither short campaigns nor seasonal averages characterize a site on their own; the local INP climatology, including its variability, requires long-duration, continuous measurements with daily or sub-daily sampling to resolve both the seasonal and the daily components of the variance.

Heat and peroxide treatments of filter samples from 15 campaigns separate the biological, organic, and inorganic contributions to the INP population. In most environments, the heat-labile and presumably proteinaceous fraction is small at the coldest
680 temperatures and grows rapidly above a campaign-dependent transition temperature, typically between -20°C and -15°C , confirming the dominance of biological INPs at warm temperatures. Across the densely covered range of -28°C to -15°C , the peroxide treatment removes roughly 90 % of the INPs that survive heating. Heat-stable organic material thus appears to dominate the heat-stable INP population at every resolved temperature. The mineral INP residual that survives both treatments amounts to roughly 10 % of the heat-stable INPs. This fraction is a lower bound, since heat and peroxide have been shown
685 to lower the ice nucleation potential of some mineral species (Daily et al., 2022; Peckhaus et al., 2016; Perkins et al., 2020). Yet, these results appear in opposition to global modeling studies that reproduce observed INP concentrations from mineral dust alone at moderate to cold activation temperatures (e.g., Vergara-Temprado et al., 2017; Chatziparaschos et al., 2023); hence, these studies appear to overestimate the mineral dust contribution by roughly an order of magnitude in the environments represented here. This strengthens the case for including organic INP sources in global INP models.

690 Several limitations should be kept in mind. Inter-instrument differences of up to half an order of magnitude remain after standardization (Sect. 2.1). Since the instrument is fixed within each deployment, these offsets do not affect the within-deployment spread. They are, however, comparable to the spread of deployment means within regions (Table 2), and thus add uncertainty to comparisons of mean INP concentrations between regions. Measurement noise inflates the within-deployment spread, but an upper bound derived for the GloPINE campaigns (Sect. S7) shows that it accounts for at most 4 % of their spread under



695 the reference sampling, and the IS samples are expected to be less noisy; hence, the reported spreads are dominated by real atmospheric variation. The geographic coverage is weighted toward the Northern Hemisphere mid- and high-latitudes, and only around 11 % of the observations represent the free troposphere. The reported statistics thus primarily constrain the INP field near the surface, whose relation to INP concentrations at the level of supercooled liquid and mixed-phase clouds remains poorly quantified. Finally, the treated subset includes no campaigns near strong dust sources. The dominance of organic INPs
700 thus describes the sampled mid- and high-latitude environments and may weaken in dust-laden regions.

This dataset provides a benchmark for INP models that is considerably larger and more environmentally diverse than the datasets used in previous evaluations (e.g., Vergara-Temprado et al., 2017; Chatziparaschos et al., 2023; Herbert et al., 2025). Each observation carries position, altitude, platform, and sampling interval as metadata, so model output can be sampled at the exact times and locations of the observations. The statistical characterization presented here constitutes a model constraint
705 in its own right. Beyond reproducing observed concentrations, a credible INP model should reproduce the magnitude of local variability and the distribution of variance across timescales. The hope is that this combination of data and diagnostics will ease thorough, statistically grounded validation of the next generation of INP prediction systems.

Code and data availability. The sources from which the INP dataset has been compiled are linked in the references of Table A1. Python scripts for data processing and visualization are available at <https://doi.org/10.5281/zenodo.23054649> (Kvist, 2026)

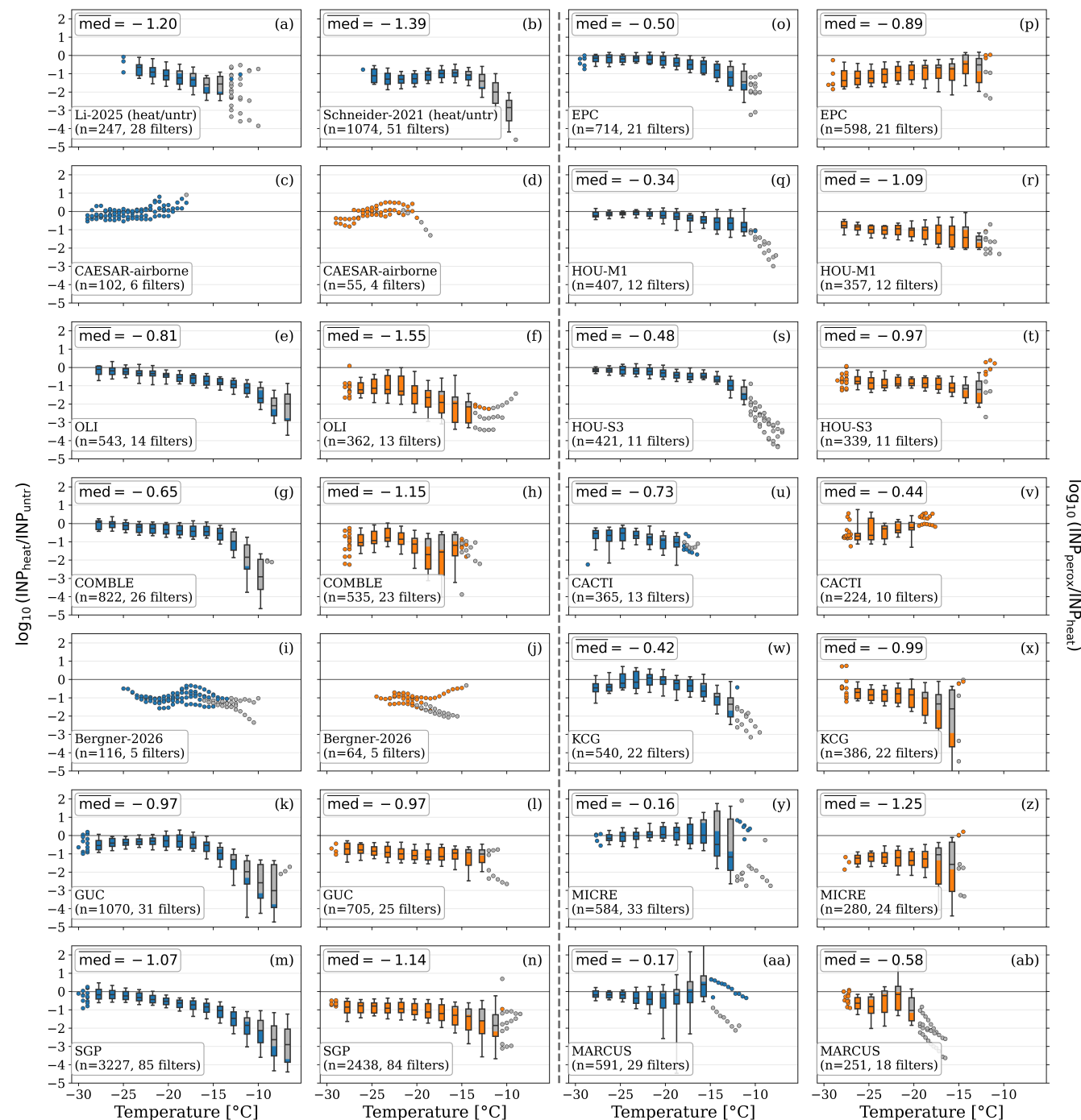


Figure 7. Temperature-resolved ratios $\log_{10}(\text{INP}_{\text{heat}}/\text{INP}_{\text{untr}})$ (blue) and $\log_{10}(\text{INP}_{\text{perox}}/\text{INP}_{\text{heat}})$ (orange), each computed from two treatments of the same filter (Sect. 3.4). Boxes show the 25th–75th percentile (center line: median) and whiskers the 10th–90th percentile of ratios in each temperature bin. The boxes are shaded gray from the top in proportion to the fraction of contributing samples using warm-tail extrapolation; n is the number of contributing ratios per panel. The legend in the top left corner reports the average median across the temperature bins. In temperature bins with fewer than 20 ratios, a point scatter of individual ratios is plotted instead of a box plot.



710 Appendix A: Summary of the campaigns contributing to the compiled dataset

Table A1: Overview of the INP measurement campaigns applied in this study, listed alphabetically. For each campaign, the table gives the geographic location, the measurement technique, the temporal coverage, the spatial extent of the observations, and the number of untreated observations N , i.e., INP concentrations of untreated samples (Sect. 2.1) at individual activation temperatures, split by platform: L land station, S ship, A aircraft, B tethered balloon. Campaign names set in bold are those judged representative of ambient conditions over the period covered. The judgment is based on the environmental conditions and sampling practices of each campaign as described in Sect. S3. The measurement technique (Instr.) is either IS, offline immersion-freezing assays of collected filter samples, or PINE, the Portable Ice Nucleation Experiment expansion-chamber instrument. For stationary sites, Lat/Lon reports the fixed coordinates, whereas for ship- and aircraft-based campaigns, the latitude and longitude ranges spanned by the platform along its track during the campaign time span are provided. For Bergner-2026 and GUC, which sampled at two nearby sites, the ranges span both sites. Reference provides the dataset or publication from which the measurements were obtained.

Campaign	Geographic location	Instr.	Time span	Lat/Lon	N	Reference
ACE	Antarctic Circumnavigation	IS (LINDA)	2016-12-21	73.77° S – 34.23° S	2579 (S)	Welti et al. (2020)
Andøya-2021	Andøya, Northern Norway	IS (DRINCO)	2017-03-18	179.98° W – 178.67° E	662 (L)	Gjelsvik et al. (2025)
			2021-03-15	69.30° N		
			2021-03-30	16.12° E		
BELJING	Pinggu, near Beijing, China	PINE	2022-03-28	40.17° N	10 940 (L)	Herbert et al. (2026)
Bergner-2026	Narsarsuaq / Narsaq, southern Greenland	IS (SPICE)	2023-07-16	117.12° E	1039 (L)	Bergner et al. (2026)
			2023-06-18	60.92° N – 61.21° N		
BNF	Bankhead National Forest, Alabama, USA	IS (CSU IS)	2023-08-06	46.05° W – 45.31° W	765 (B), 697 (L)	Creamean et al. (2020, 2025)
			2024-10-08	34.34° N		
CACTI	Sierras de Córdoba, Argentina	IS (CSU IS)	2025-06-15	87.34° W	1131 (A)	DeMott and Hill (2020)
			2018-11-06	32.30° S – 32.05° S		
CAESAR-airborne	Svalbard / Greenland-Sea aircraft transects	IS (CSU IS)	2018-12-08	65.11° W – 64.67° W	304 (A)	Perkins and Barry (2025)
			2024-02-28	71.39° N – 77.13° N		
CALISHTO	Mount Helmos, Peloponnese, Greece	PINE	2024-04-03	12.29° W – 23.04° E	842 (L)	Herbert et al. (2026)
			2021-10-12	37.90° N		
CAPRICORN_II	Southern Ocean	IS (CSU IS)	2022-05-07	22.20° E	452 (S)	Moore et al. (2024)
			2018-01-13	66.04° S – 44.77° S		
COMBLE	Andenes, Andøya, Norway	IS (CSU IS)	2018-02-20	132.05° E – 149.99° E	2246 (L)	DeMott et al. (2025)
			2019-12-01	69.14° N		
CORONA-PINE01	KIT-Karlsruhe, Germany	PINE	2020-05-26	15.68° E	120 (L)	Herbert et al. (2026)
			2021-04-09	49.09° N		
CORONA-PINE04	KIT-Karlsruhe, Germany	PINE	2021-04-28	8.43° E	205 (L)	Herbert et al. (2026)
			2021-03-25	49.09° N		
CountIce-pt1	University of Leeds, UK	PINE	2021-05-01	8.43° E	461 (L)	Herbert et al. (2026)
			2021-10-07	53.81° N		
CountIce-pt2	University of Leeds, UK	PINE	2022-02-15	1.55° W	588 (L)	Herbert et al. (2026)
			2023-03-09	53.81° N		
CRG	Baltimore region, Maryland, USA	IS (CSU IS)	2023-08-09	1.55° W	709 (B)	Creamean et al. (2020, 2025)
			2025-02-19	39.30° N		
Eckermann-NM	Neumayer Station III, Antarctica	IS (LINA + INDIA)	2025-07-28	76.17° W	783 (L)	Wex et al. (2025)
			2019-12-13	70.65° S		
			2021-12-03	8.25° W		



Campaign	Geographic location	Instr.	Time span	Lat/Lon	<i>N</i>	Reference
Eckermann-PrElis	Princess Elisabeth Station, Antarctica	IS (LINA + INDA)	2020-12-14 2022-01-28	71.95° S 23.35° E	257 (L)	Wex et al. (2025)
EPC	Scripps Pier, La Jolla, California, USA	IS (CSU IS)	2023-02-17 2024-02-14	32.87° N 117.26° W	2183 (L)	Creamean et al. (2020)
ExINP-ENA	Eastern North Atlantic, Graciosa, Azores	PINE	2020-10-02 2021-03-26	39.09° N 28.03° W	2365 (L)	Herbert et al. (2026)
ExINP-GVB	Gruvebadet, Ny-Ålesund, Svalbard	PINE	2023-03-15 2023-04-09	78.92° N 11.93° E	516 (L)	Herbert et al. (2026)
ExINP-NSA	North Slope of Alaska, Utqiagvik, USA	PINE	2021-10-21 2023-12-31	71.32° N 156.62° W	6988 (L)	Herbert et al. (2026)
ExINP-SGP	Southern Great Plains, Oklahoma, USA	PINE	2019-10-01 2019-11-13	36.61° N 97.49° W	1056 (L)	Herbert et al. (2026)
ExINP-ZEP	Zeppelin Mountain, Ny-Ålesund, Svalbard	PINE	2023-04-26 2023-10-22	78.90° N 11.88° E	2146 (L)	Herbert et al. (2026)
GUC	Storm Peak / Crested Butte, Colorado, USA	IS (CSU IS)	2021-09-10 2023-06-13	38.90° N – 38.92° N 106.95° W – 106.94° W	3086 (L), 339 (B)	Creamean et al. (2020, 2025)
Harrison-2022	Ragged Point, Barbados	IS (NIPI)	2017-07-24 2017-08-24	13.17° N 59.43° W	158 (L)	Harrison et al. (2022)
HOU-M1	Houston, Texas, USA	IS (CSU IS)	2022-06-05 2022-09-29	29.67° N 95.06° W	1373 (L)	Creamean et al. (2020)
HOU-S3	Houston, Texas, USA	IS (CSU IS)	2022-06-08 2022-09-30	29.33° N 95.74° W	1324 (L)	Creamean et al. (2020)
HyICE2018	Hyttälä SMEAR II, Finland	PINE	2018-03-22 2018-04-28	61.85° N 24.29° E	9 (L)	Herbert et al. (2026)
ICE-T	St. Croix, US Virgin Islands	IS (CSU IS)	2011-07-15 2011-07-23	17.50° N – 18.38° N 65.75° W – 64.50° W	16 (L), 7 (A)	DeMott et al. (2016)
KCG	Cape Grim, Tasmania, Australia	IS (CSU IS)	2024-04-18 2025-08-25	40.68° S 144.69° E	2275 (L)	Creamean et al. (2020)
Li-2025	Arctic Ocean (Russian high Arctic transect)	IS (DRINCZ)	2021-08-06 2021-09-03	71.37° N – 82.69° N 38.18° E – 106.38° E	431 (S)	Li et al. (2025)
LIFE- FROSTDEFEND	Aigio, Greece	PINE	2023-03-27 2023-04-21	38.23° N 22.13° E	136 (L)	Herbert et al. (2026)
M-Phase	Labrador Sea	PINE	2022-05-20 2022-06-23	52.42° N – 68.84° N 57.48° W – 18.19° W	234 (S)	Herbert et al. (2026)
MAGIC	NE Pacific (LA ↔ Honolulu cargo ship)	IS (CSU IS)	2013-07-21 2013-07-23	23.10° N – 31.30° N 151.70° W – 125.60° W	9 (S)	DeMott et al. (2016)
MARCUS	Southern Ocean transect (Hobart ↔ Antarctica)	IS (CSU IS)	2017-11-02 2018-03-23	68.57° S – 45.16° S 62.76° E – 158.94° E	1206 (S)	McCluskey et al. (2018a)
MICRE	Macquarie Island, Southern Ocean	IS (CSU IS)	2017-04-01 2018-03-03	54.50° S 158.93° E	1969 (L)	DeMott et al. (2018a)
OLI	Oliktok Point, North Slope of Alaska, USA	IS (CSU IS)	2020-08-21 2021-06-02	70.50° N 149.89° W	1496 (L)	Creamean et al. (2020)
PaCE22	Pallas, Finnish Lapland	PINE	2022-09-23 2022-12-22	67.97° N 24.11° E	753 (L)	Herbert et al. (2026)
PICNIC	Puy de Dôme, France	PINE	2018-10-09 2018-10-24	45.77° N 2.96° E	131 (L)	Herbert et al. (2026)
Price-2018	Cape Verde region	IS (NIPI)	2015-08-07 2015-08-20	13.60° N – 24.37° N 25.11° W – 17.84° W	190 (A)	Price et al. (2018)
PS106	Fram Strait / Arctic Ocean	IS (LINA)	2017-05-24 2017-07-17	55.24° N – 83.68° N 2.57° E – 34.11° E	2030 (S)	Hartmann et al. (2021)
PUY	Puy de Dôme, France	PINE	2022-12-15 2023-03-17	45.77° N 2.96° E	1861 (L)	Herbert et al. (2026)
Raif-2024	Svalbard / Norwegian-Sea aircraft transects	IS (NIPI)	2022-03-11 2022-03-30	70.99° N – 78.49° N 3.51° E – 28.00° E	348 (A)	Raif et al. (2024)



Campaign	Geographic location	Instr.	Time span	Lat/Lon	<i>N</i>	Reference
Sanchez-Marroquin-2017	Iceland aircraft transects	IS (NIPI)	2017-10-02	63.23°N – 64.83°N	625 (A)	Sanchez-Marroquin et al. (2020)
SBO21	Sonnblick Observatory, Austrian Alps	PINE	2017-10-05 2021-07-28	21.22°W – 16.46°W 47.05°N	2816 (L)	Herbert et al. (2026)
Schneider-2021	Hyytiälä SMEAR II, Finland	IS (INSEKT)	2022-10-13 2018-03-11	12.96°E 61.85°N	7148 (L)	Schneider et al. (2021)
SGP	Southern Great Plains, Oklahoma, USA	IS (CSU IS)	2019-05-30 2020-10-20	24.30°E 36.60°N	10 902 (L), 416 (B)	Creamean et al. (2020, 2025)
SHIPPO	Western Bering Sea	IS (CSU IS)	2025-02-04 2012-02-27	97.49°W 62.57°N	5 (S)	DeMott et al. (2016)
SOCRATES	Southern Ocean	IS (CSU IS)	173.70°E 2018-01-16	61.11°S – 45.94°S	351 (A)	Moore et al. (2024)
SPL01	Storm Peak Laboratory, Colorado, USA	PINE	2018-02-24 2021-10-12	135.93°E – 158.82°E 40.45°N	3127 (L)	Herbert et al. (2026)
Sze-2023	Villum / Station Nord, NE Greenland	IS (LINA + INDA)	2022-07-11 2018-07-12	106.74°W 81.60°N	2538 (L)	Sze et al. (2023)
Tobo-2016	Zeppelin Observatory, Ny-Ålesund, Svalbard	IS (CRAFT)	2020-09-12 2016-07-14	16.67°W 78.91°N	165 (L)	Tobo et al. (2019)
Tobo-2017	Zeppelin Observatory, Ny-Ålesund, Svalbard	IS (CRAFT)	2016-07-19 2017-03-13	11.88°E 78.91°N	121 (L)	Tobo et al. (2019)
Tobo-2020	Tokyo, Japan	IS (CRAFT)	2017-03-19 2016-08-02	11.88°E 35.71°N	1886 (L)	Tobo et al. (2020)
WAO	Norfolk, UK	PINE	2017-07-31 2023-10-27	139.81°E 52.95°N	650 (L)	Herbert et al. (2026)
Wex-2019-Alert	Alert, Ellesmere Island, Canada	IS (INDA)	2023-12-30 2015-04-26	1.12°E 82.50°N	376 (L)	Wex et al. (2019)
Wex-2019-Ny-Alesund	Ny-Ålesund, Svalbard	IS (INDA)	2016-04-29 2012-03-24	62.33°W 78.92°N	154 (L)	Wex et al. (2019)
Wex-2019-Utqiagvik	Utqiagvik, North Slope of Alaska, USA	IS (INDA)	2012-09-05 2012-07-02	11.93°E 71.30°N	348 (L)	Wex et al. (2019)
Wex-2019-Villum	Villum / Station Nord, NE Greenland	IS (INDA)	2013-06-02 2013-12-12	156.60°W 81.60°N	91 (L)	Wex et al. (2019)
			2015-12-10	16.67°W		



Appendix B: Deriving formulas for sampling adjustment

Table B1. Symbols used in this appendix, with the equations or sections in which they are defined.

<i>Sampling</i>	
exposure time	time interval of duration L over which one sample integrates $\text{INP}(t)$
L, R, K	exposure time, time span, and number of samples of a record, $R = KL$
$\Delta t, \tau$	width of a window $[t, t + \Delta t]$ and time lag
subscript 0	the sampling of the record as measured, with exposure time L_0 and time span R_0
subscript ref	the reference sampling, with exposure time L_{ref} and time span R_{ref}
<i>Ideal series, single realization</i>	
$\text{INP}(t)$	the concentration of INPs at time t at some fixed activation temperature
$y(t)$	$\log_{10} \text{INP}(t)$, Sect. B1
$a(t) = y(t) - \langle y \rangle$	anomaly of $y(t)$, where $\langle y \rangle$ is the long-term mean of $y(t)$, Sect. B1
$\bar{a}_{\Delta t}(t), v_{\Delta t}(t)$	mean and internal variance of a over a window of width Δt , Eq. (B3)
$\nu(t')$	deviation of $a(t')$ from its mean over an exposure, Eq. (B29)
<i>Ideal series, average over all placements</i>	
Θ	duration of the long time interval over which $\langle \cdot \rangle$ averages, Eq. (B1)
$\sigma^2, C(\tau)$	variance of a and autocovariance of a at lag τ , Eq. (B2)
$C_L(\tau)$	autocovariance at lag τ of the means of a over windows of width L , Eq. (B21)
$V(\Delta t)$	accumulated variance, Eq. (B4)
<i>Samples of a record, single realization</i>	
b_i, t_i	sample anomaly of the i -th sample of a record, and the start of its exposure, Eq. (B6)
m, s^2	mean and variance of the sample anomalies of a record, Eq. (B7)
$s_{\Delta t}^2$	variance of the sample anomalies inside a window of width Δt , Sect. B4
$c(\tau)$	autocovariance of the sample anomalies of a record, Eq. (B17)
$w(\Delta t)$	measured accumulated variance, the mean of $s_{\Delta t}^2$ over all windows of width Δt in a record, Sect. B4
$f(\Delta t), r(\tau)$	normalized accumulated variance $w(\Delta t)/s^2$ and autocorrelation $c(\tau)/s^2$ of a single record, Sects. B4 and B5
$y_L(t_i), \bar{y}$	logarithm of the concentration reported by a sample with exposure time L starting at t_i , and its mean over the samples of a record, Sect. B6
<i>Samples of a record, average over all placements</i>	
$\langle s^2 \rangle, \langle c(\tau) \rangle, \langle y_L(t_i) \rangle$	expected values of s^2 , $c(\tau)$, and $y_L(t_i)$, Eq. (B9)
$F(\Delta t), \rho(\tau)$	expected normalized accumulated variance and autocorrelation of a record, Eqs. (B14) and (B23)
<i>Constants</i>	
A, B, S	ratios of differences of V , Eqs. (B12), (B16), and (B26)
κ	$\ln(10)/2$, Eq. (B31)



A campaign yields a record of K consecutive samples. Each sample integrates over a time interval of length L , referred to as its *exposure time* below, and the total time span of the record is denoted R . In the derivations below, it is assumed that samples are taken consecutively so that $R = KL$. The statistics of a record depend on its sampling routine (L, R) ; i.e., the same atmospheric process sampled with different (L, R) yields different statistics. This appendix derives the relations between the statistics of a record sampled with (L_0, R_0) and those of a record of the same site sampled with a reference sampling $(L_{\text{ref}}, R_{\text{ref}})$. Table B1 lists the symbols used in this appendix.

B1 Definitions

Let $g(t)$ be any quantity that depends on time t , and let $\langle \cdot \rangle$ denote its temporal average over a time interval of duration Θ , where Θ is much longer than any timescale on which $g(t)$ varies. Then,

$$\langle g \rangle \equiv \frac{1}{\Theta} \int_0^{\Theta} g(t) dt. \quad (\text{B1})$$

Let the logarithm of the INP concentration at a site be denoted $y(t) = \log_{10} \text{INP}(t)$. Its anomaly, $a(t) = y(t) - \langle y \rangle$, is assumed to be stationary; i.e., its statistical properties taken over the long time interval Θ do not depend on the start time of that interval. The variance of $a(t)$, σ^2 , and the autocovariance of $a(t)$ at lag τ , $C(\tau)$, are defined by

$$\sigma^2 \equiv \langle a(t)^2 \rangle, \quad C(\tau) \equiv \langle a(t) a(t + \tau) \rangle \Rightarrow C(0) = \sigma^2. \quad (\text{B2})$$

Define a window as an interval $[t, t + \Delta t]$ of width Δt that starts at position t . The mean of $a(t)$ over the window, $\bar{a}_{\Delta t}(t)$, and the internal variance of $a(t)$ over the window, $v_{\Delta t}(t)$, are then given by

$$\bar{a}_{\Delta t}(t) \equiv \frac{1}{\Delta t} \int_t^{t+\Delta t} a(t') dt', \quad v_{\Delta t}(t) \equiv \frac{1}{\Delta t} \int_t^{t+\Delta t} [a(t') - \bar{a}_{\Delta t}(t)]^2 dt' = \frac{1}{\Delta t} \int_t^{t+\Delta t} a(t')^2 dt' - \bar{a}_{\Delta t}(t)^2. \quad (\text{B3})$$

Let the accumulated variance, $V(\Delta t)$, be defined as the average of the internal variance over all window positions t within the long time interval Θ :

$$V(\Delta t) \equiv \langle v_{\Delta t}(t) \rangle. \quad (\text{B4})$$

Along with Eqs. (B1) and (B3), the definition in Eq. (B4) implies

$$\begin{aligned} V(\Delta t) &= \frac{1}{\Theta} \int_0^{\Theta} \frac{1}{\Delta t} \int_t^{t+\Delta t} a(t')^2 dt' dt - \langle \bar{a}_{\Delta t}(t)^2 \rangle = \frac{1}{\Theta} \int_0^{\Theta} \frac{1}{\Delta t} \int_0^{\Delta t} a(t+s)^2 ds dt - \langle \bar{a}_{\Delta t}(t)^2 \rangle \\ &= \frac{1}{\Delta t} \int_0^{\Delta t} \frac{1}{\Theta} \int_0^{\Theta} a(t+s)^2 dt ds - \langle \bar{a}_{\Delta t}(t)^2 \rangle = \frac{1}{\Delta t} \int_0^{\Delta t} \sigma^2 ds - \langle \bar{a}_{\Delta t}(t)^2 \rangle = \sigma^2 - \langle \bar{a}_{\Delta t}(t)^2 \rangle \\ &\Rightarrow \sigma^2 = V(\Delta t) + \langle \bar{a}_{\Delta t}(t)^2 \rangle, \quad \text{with } 0 \leq V(\Delta t) \leq \sigma^2, \end{aligned} \quad (\text{B5})$$



where the substitution $t' = t + s$ was applied, the order of integration was swapped, and it was used that the inner integral equals σ^2 by stationarity, since $\Theta \gg \Delta t$ by assumption. Equation (B5) shows that the variance of $a(t)$ can be decomposed into the expected internal variance within a window of width Δt , i.e., $V(\Delta t)$, and the variance of the mean of $a(t)$ within such windows, $\langle \bar{a}_{\Delta t}(t)^2 \rangle$. This term is a variance because $\langle \bar{a}_{\Delta t}(t) \rangle = 0$, which follows from $a(t)$ being an anomaly.

740 Each sample collects air during its exposure time L and reports the average INP concentration over that time. For the sample whose exposure starts at t_i , the *sample anomaly* b_i is obtained by taking $\log_{10}$ of this average concentration and subtracting $\langle y \rangle$, i.e., the long-term mean of $\log_{10} \text{INP}(t)$. A *record* consists of K consecutive samples whose exposures start at the times

$$t_i = t_1 + (i - 1)L, \quad i = 1, \dots, K, \quad (\text{B6})$$

and together span the window $[t_1, t_1 + R]$ with $R = KL$. The mean, m , and the variance, s^2 , of the sample anomalies of a
745 record are

$$m \equiv \frac{1}{K} \sum_{i=1}^K b_i, \quad s^2 \equiv \frac{1}{K} \sum_{i=1}^K (b_i - m)^2. \quad (\text{B7})$$

By Jensen's inequality, $b_i \geq \bar{a}_L(t_i)$ with $\bar{a}_L(t_i)$ as defined by Eq. (B3), i.e., the logarithm of the exposure-mean concentration is larger than or equal to the exposure mean of the logarithm of the concentration, $y(t)$. The difference between b_i and $\bar{a}_L(t_i)$ grows with the variability of $\text{INP}(t)$ during the exposure. With K samples, the mean of $a(t)$ over the record can be written
750 $\bar{a}_R(t_1) = \frac{1}{K} \sum_{i=1}^K \bar{a}_L(t_i)$, which together with $b_i \geq \bar{a}_L(t_i)$ causes the mean of the sample anomalies, m , to overestimate the mean of $a(t)$ over the record, $\bar{a}_R(t_1)$. Section B6 quantifies this shift. The variance and the autocorrelation of the sample anomalies are affected only by the differences between the shifts of the individual samples, which are small when all samples share one exposure time L , as is the case here. Therefore

$$b_i = \bar{a}_L(t_i) \quad \text{and hence} \quad m = \bar{a}_R(t_1) \quad (\text{B8})$$

755 are assumed in Sects. B2 to B5, which derive the adjustment formulas for the expected variance, $\langle s^2 \rangle$, the expected normalized accumulated variance, $F(\Delta t)$, and the expected autocorrelation, $\rho(\tau)$. The assumption of Eq. (B8) is relaxed when deriving the adjustment formula for the expected sample mean, $\langle y_L(t_i) \rangle$, in Sect. B6.

By Eq. (B6), all samples of a record are fixed by its start time t_1 , so m , s^2 , and every other quantity computed from the sample anomalies are functions of t_1 . The expected value of such a quantity is defined as its average taken over all start times
760 t_1 within the long time interval Θ , i.e., the expected value is obtained by applying Eq. (B1) to a statistical quantity, e.g., the measured variance s^2 , with the start time of the record t_1 used as the integration variable:

$$\langle s^2 \rangle = \frac{1}{\Theta} \int_0^\Theta s^2(t_1) dt_1. \quad (\text{B9})$$

A record corresponds to the measurement period of a campaign at a fixed site, with t_1 being its start date. The campaign yields one variance, $s^2(t_1)$, which depends on whether it occurred in a calm or variable period. Averaging over t_1 removes
765 dependence on campaign timing and leaves the variance that a record with exposure time L and time span R measures on



average at the site in question, which depends only on the site and the sampling routine (L, R) . The relations derived below hold between expected values of the adjusted statistics; e.g., $\langle s^2 \rangle$. Applied to a single record, the derived adjustments remove the systematic effect of the sampling (L, R) while preserving the campaign-specific deviation between the measured and expected values, as shown in Sect. S6 of the supplement.

770 B2 Variance of a record

Expanding the square in Eq. (B7) and using the definition of m ,

$$s^2 = \frac{1}{K} \sum_{i=1}^K b_i^2 - m^2 \Rightarrow \langle s^2 \rangle = \frac{1}{K} \sum_{i=1}^K \langle b_i^2 \rangle - \langle m^2 \rangle. \quad (\text{B10})$$

By Eq. (B8), $b_i = \bar{a}_L(t_i)$, which together with Eq. (B5) using $\Delta t = L$ yields $\langle b_i^2 \rangle = \langle \bar{a}_L(t)^2 \rangle = \sigma^2 - V(L)$. By Eq. (B8), $m = \bar{a}_R(t_1)$ is the mean of $a(t)$ over a window of width R , and Eq. (B5) with $\Delta t = R$ gives $\langle m^2 \rangle = \sigma^2 - V(R)$. Hence

$$775 \quad \langle s^2 \rangle = [\sigma^2 - V(L)] - [\sigma^2 - V(R)] = V(R) - V(L). \quad (\text{B11})$$

The expected variance of a record is the variance accumulated between its exposure time and its time span.

B3 Adjustment of the measured variance to reference sampling conditions

Let the subscript 0 denote the native sampling of a campaign, so that s_0^2 is the measured variance of a record sampling the INP concentration with the sampling routine (L_0, R_0) . Based on s_0^2 , R_0 , and L_0 , the goal is to estimate the variance s_{ref}^2 that would have been measured if the INP concentration had been sampled through a record of time span R_{ref} with exposure time L_{ref} . By Eq. (B11),

$$\begin{aligned} \langle s_0^2 \rangle &= V(R_0) - V(L_0) \quad \text{and} \quad \langle s_{\text{ref}}^2 \rangle = V(R_{\text{ref}}) - V(L_{\text{ref}}), \\ \Rightarrow \quad \boxed{\langle s_{\text{ref}}^2 \rangle} &= \frac{\langle s_0^2 \rangle}{A}, \quad \text{where} \quad A = \frac{V(R_0) - V(L_0)}{V(R_{\text{ref}}) - V(L_{\text{ref}})}. \end{aligned} \quad (\text{B12})$$

For a single record, s_{ref}^2 is estimated by s_0^2/A .

785 B4 Adjustment of the normalized accumulated variance

Consider a window of width $\Delta t = nL_0$ inside a record sampled with (L_0, R_0) , so that the window holds n consecutive samples. Let $s_{\Delta t}^2$ denote the measured variance of their sample anomalies, computed as in Eq. (B7). The n samples form a record sampled with $(L_0, \Delta t)$, so by Eq. (B11)

$$\langle s_{\Delta t}^2 \rangle = V(\Delta t) - V(L_0). \quad (\text{B13})$$

790 Let $w_0(\Delta t)$ denote the measured accumulated variance of the record, i.e., the mean of $s_{\Delta t}^2$ over all windows of width Δt inside the record (Sect. 3.5), so that $\langle w_0(\Delta t) \rangle = \langle s_{\Delta t}^2 \rangle$. The expected normalized accumulated variance of the record, $F_0(\Delta t)$, is



defined as the expected measured accumulated variance relative to the expected measured variance of the whole record $\langle s_0^2 \rangle$:

$$F_0(\Delta t) \equiv \frac{\langle w_0(\Delta t) \rangle}{\langle s_0^2 \rangle} = \frac{V(\Delta t) - V(L_0)}{V(R_0) - V(L_0)}, \quad L_0 \leq \Delta t \leq R_0. \quad (\text{B14})$$

Thus by definition $F_0(L_0) = 0$ and $F_0(R_0) = 1$. Had the same site instead been sampled with the reference sampling $(L_{\text{ref}}, R_{\text{ref}})$,
795 the expected normalized accumulated variance would be

$$F_{\text{ref}}(\Delta t) = \frac{V(\Delta t) - V(L_{\text{ref}})}{V(R_{\text{ref}}) - V(L_{\text{ref}})}, \quad L_{\text{ref}} \leq \Delta t \leq R_{\text{ref}}. \quad (\text{B15})$$

Inserting $V(\Delta t) - V(L_{\text{ref}}) = [V(\Delta t) - V(L_0)] + [V(L_0) - V(L_{\text{ref}})]$ and using Eqs. (B12) and (B14) gives the relation between $F_0(\Delta t)$ and $F_{\text{ref}}(\Delta t)$,

$$\boxed{F_{\text{ref}}(\Delta t) = A F_0(\Delta t) + B}, \quad \text{where} \quad B = \frac{V(L_0) - V(L_{\text{ref}})}{V(R_{\text{ref}}) - V(L_{\text{ref}})}, \quad \text{and} \quad \max(L_0, L_{\text{ref}}) \leq \Delta t \leq \min(R_0, R_{\text{ref}}). \quad (\text{B16})$$

800 For a single record, the normalized accumulated variance adjusted to the reference sampling, $f_{\text{ref}}(\Delta t)$, is estimated by $f_{\text{ref}}(\Delta t) = A f_0(\Delta t) + B$, where $f_0(\Delta t) = w_0(\Delta t)/s_0^2$ is the measured normalized accumulated variance (Sect. 3.5).

B5 Adjustment of the autocorrelation

Consider a record, sampled with (L_0, R_0) and holding $K = R_0/L_0$ samples. Let τ denote the lag between two samples, and consider a lag of k samples, i.e., $\tau = kL_0$. The record then holds $K - k$ pairs of sample anomalies (b_i, b_{i+k}) separated by the
805 lag τ . The autocovariance of the record's sample anomalies at that lag is

$$c_0(\tau) \equiv \frac{1}{K-k} \sum_{i=1}^{K-k} (b_i - m)(b_{i+k} - m) = \frac{1}{K-k} \sum_{i=1}^{K-k} b_i b_{i+k} - m^2 - m[(m^- - m) + (m^+ - m)], \quad (\text{B17})$$

where m^- is the mean of the first $K - k$ sample anomalies and m^+ the mean of the last $K - k$ sample anomalies. Averaging over all start times t_1 of the record gives

$$\langle c_0(\tau) \rangle = \frac{1}{K-k} \sum_{i=1}^{K-k} \langle b_i b_{i+k} \rangle - \langle m^2 \rangle - \langle m[(m^- - m) + (m^+ - m)] \rangle. \quad (\text{B18})$$

810 Let $\hat{m}^-$ denote the mean of the last k sample anomalies of the record, which are the sample anomalies that m^- omits, and let $\hat{m}^+$ denote the mean of the first k sample anomalies, which are the sample anomalies that m^+ omits. Since m is the mean of all K sample anomalies,

$$Km = (K-k)m^- + k\hat{m}^- \Rightarrow m^- - m = \frac{k}{K-k}(m - \hat{m}^-), \quad (\text{B19})$$

and likewise $m^+ - m = \frac{k}{K-k}(m - \hat{m}^+)$. Adding the two yields

$$815 (m^- - m) + (m^+ - m) = \frac{2k}{K-k}(m - \hat{m}), \quad \text{where} \quad \hat{m} = \frac{\hat{m}^- + \hat{m}^+}{2} \quad (\text{B20})$$

is the mean of the $2k$ sample anomalies at the ends of the record. The last term of Eq. (B18) is thus $-2k/(K-k)$ times $\langle m(m - \hat{m}) \rangle$, and it is neglected for the following reasons:



- The factor $2k/(K - k)$ is small for lags short compared with the time span, $\tau \ll R_0$.
- $\hat{m}$ covers $2k$ of the K sample anomalies, so $m - \hat{m}$ tends to zero as k tends to $K/2$ where $\hat{m} = m$.
- Since $\langle m \rangle = 0$ by Eq. (B8), $\langle m(m - \hat{m}) \rangle$ is the covariance between the record mean and the deviation of the record mean from the mean of the ends of the record. This covariance is small when only a small share of the variance of a record stems from variability on timescales comparable to or longer than the record time span, which is assumed for the campaigns of Fig. 6.

By Eq. (B8), $b_i = \bar{a}_{L_0}(t_i)$ and $b_{i+k} = \bar{a}_{L_0}(t_i + \tau)$. Let $C_L(\tau)$ denote the autocovariance of the means of $a(t)$ over windows of width L ,

$$C_L(\tau) \equiv \langle \bar{a}_L(t) \bar{a}_L(t + \tau) \rangle. \quad (\text{B21})$$

By this definition, $\langle b_i b_{i+k} \rangle = C_{L_0}(\tau)$ for every i . Furthermore, Eqs. (B8) and (B5) give $\langle m^2 \rangle = \sigma^2 - V(R_0)$. Neglecting the last term, Eq. (B18) thus becomes

$$\langle c_0(\tau) \rangle = C_{L_0}(\tau) - \sigma^2 + V(R_0). \quad (\text{B22})$$

The expected autocorrelation of the record, $\rho_0(\tau)$, is defined as the ratio of the expected autocovariance to the expected variance, Eq. (B11),

$$\rho_0(\tau) \equiv \frac{\langle c_0(\tau) \rangle}{\langle s_0^2 \rangle} = \frac{C_{L_0}(\tau) - \sigma^2 + V(R_0)}{V(R_0) - V(L_0)}. \quad (\text{B23})$$

Under the reference sampling the autocorrelation, $\rho_{\text{ref}}(\tau)$, is

$$\rho_{\text{ref}}(\tau) = \frac{C_{L_{\text{ref}}}(\tau) - \sigma^2 + V(R_{\text{ref}})}{V(R_{\text{ref}}) - V(L_{\text{ref}})}. \quad (\text{B24})$$

To relate $\rho_0(\tau)$ to $\rho_{\text{ref}}(\tau)$ it is assumed that

$$C_{L_{\text{ref}}}(\tau) = C_{L_0}(\tau), \quad (\text{B25})$$

i.e., the covariance between two sample anomalies a lag τ apart is assumed not to depend on the exposure time. Section S5 shows that Eq. (B25) holds for $\tau > \max(2L_0, 2L_{\text{ref}})$, provided that the variance of periodic variability with periods up to $\max(L_0, L_{\text{ref}})$, in practice the diurnal cycle, is small compared with the variance of the reference record $\langle s_{\text{ref}}^2 \rangle$. With Eq. (B25),

writing $C_{L_0}(\tau) - \sigma^2 + V(R_{\text{ref}}) = [C_{L_0}(\tau) - \sigma^2 + V(R_0)] + [V(R_{\text{ref}}) - V(R_0)]$ and using Eqs. (B12) and (B23) gives the relation between $\rho_0(\tau)$ and $\rho_{\text{ref}}(\tau)$:

$$\boxed{\rho_{\text{ref}}(\tau) = A \rho_0(\tau) + S}, \quad \text{where} \quad S = \frac{V(R_{\text{ref}}) - V(R_0)}{V(R_{\text{ref}}) - V(L_{\text{ref}})}. \quad (\text{B26})$$

Equation (B26) is a good approximation for $\tau > \max(2L_0, 2L_{\text{ref}})$, provided that only a small share of the variance of a record stems from variability on timescales comparable to or longer than its time span, which allowed the last term of Eq. (B18) to be



845 neglected, and that the variance of periodic variability with periods up to $\max(L_0, L_{\text{ref}})$, in practice the diurnal cycle, is small compared with the variance of the reference record, which Eq. (B25) requires. Both are assumed for the campaigns in Fig. 6.

For a single record, let $r_0(\tau) \equiv c_0(\tau)/s_0^2$ denote the autocorrelation of its sample anomalies at lag τ , which estimates $\rho_0(\tau)$, and let $r_{\text{ref}}(\tau) = A r_0(\tau) + S$ denote the corresponding estimate of $\rho_{\text{ref}}(\tau)$.

B6 Adjustment of the mean

850 This section derives the upward shift of the measured mean of $\log_{10} \text{INP}(t)$ relative to $\langle y \rangle$, which results from averaging $\text{INP}(t)$ over the exposure time L of each sample before applying the logarithm. The assumption of Eq. (B8) is therefore not made in this section. By the definition of $a(t)$ in Sect. B1, $y(t) = \langle y \rangle + a(t)$, so that

$$\text{INP}(t) = 10^{y(t)} = 10^{\langle y \rangle + a(t)}. \quad (\text{B27})$$

A sample whose exposure starts at t_i and lasts L reports the average concentration over the exposure. Let $y_L(t_i)$ denote the
855 logarithm of this average so that

$$y_L(t_i) = \log_{10} \left(\frac{1}{L} \int_{t_i}^{t_i+L} 10^{\langle y \rangle + a(t')} dt' \right). \quad (\text{B28})$$

Write $a(t') = \bar{a}_L(t_i) + \nu(t')$ for $t_i \leq t' \leq t_i + L$, where $\bar{a}_L(t_i)$ is the mean of $a(t)$ over the exposure and $\nu(t')$ the deviation from it. Since $\langle y \rangle$ and $\bar{a}_L(t_i)$ do not depend on t' ,

$$y_L(t_i) = \log_{10} \left(10^{\langle y \rangle} 10^{\bar{a}_L(t_i)} \frac{1}{L} \int_{t_i}^{t_i+L} 10^{\nu(t')} dt' \right) = \langle y \rangle + \bar{a}_L(t_i) + \log_{10} \left(\frac{1}{L} \int_{t_i}^{t_i+L} 10^{\nu(t')} dt' \right). \quad (\text{B29})$$

860 Assume that $\nu(t')$ is normally distributed within the exposure, which amounts to assuming that $\text{INP}(t)$ is lognormally distributed within the exposure. By the definition of $\nu(t')$ it has a mean of zero, and by Eq. (B3) its variance is $v_L(t_i)$. The time average in Eq. (B29) is equal to the average of 10^ϵ with ϵ distributed as $\nu(t')$:

$$\frac{1}{L} \int_{t_i}^{t_i+L} 10^{\nu(t')} dt' = \frac{1}{\sqrt{2\pi v_L(t_i)}} \int_{-\infty}^{\infty} 10^\epsilon \exp \left(-\frac{\epsilon^2}{2v_L(t_i)} \right) d\epsilon = \exp \left[\frac{1}{2} v_L(t_i) \ln^2(10) \right]. \quad (\text{B30})$$

Equation (B29) then becomes

$$865 \quad y_L(t_i) = \langle y \rangle + \bar{a}_L(t_i) + \log_{10} \left(\exp \left[\frac{1}{2} v_L(t_i) \ln^2(10) \right] \right) = \langle y \rangle + \bar{a}_L(t_i) + \kappa v_L(t_i), \quad \kappa = \frac{\ln(10)}{2} \approx 1.15. \quad (\text{B31})$$

Averaging Eq. (B31) over all positions t_i of the exposure, with $\langle \bar{a}_L \rangle = 0$ because $a(t)$ is an anomaly and $\langle v_L \rangle = V(L)$ by Eq. (B4), gives the expected logarithm of a sample,

$$\langle y_L(t_i) \rangle = \langle y \rangle + \kappa V(L), \quad (\text{B32})$$



which is the same for every i . It exceeds $\langle y \rangle$ by $\kappa V(L)$, an amount that depends on the exposure time and not on the time span. Let the logarithms reported by samples with the native and the reference exposure time be denoted by the shorthand $y_0 = y_{L_0}(t_i)$ and $y_{\text{ref}} = y_{L_{\text{ref}}}(t_i)$, so that $\langle y_0 \rangle$ and $\langle y_{\text{ref}} \rangle$ stand for $\langle y_{L_0}(t_i) \rangle$ and $\langle y_{L_{\text{ref}}}(t_i) \rangle$. Equation (B32) then gives

$$\langle y_{\text{ref}} \rangle = \langle y_0 \rangle - \kappa [V(L_0) - V(L_{\text{ref}})]. \quad (\text{B33})$$

The constants A , B , and S of Eqs. (B12), (B16), and (B26) contain $V(\Delta t)$ only through ratios of differences, so these constants are unchanged for linear transformations of $V(\Delta t)$, meaning that any function $\alpha V(\Delta t) + \beta$ with constants α and β can be used to calculate them. To allow the adjustment of the mean of Eq. (B33) to be obtained from the same linear transformation of $V(\Delta t)$ as A , B , and S , it is recast in terms of B . By the definition of B in Eq. (B16) and $\langle s_{\text{ref}}^2 \rangle = V(R_{\text{ref}}) - V(L_{\text{ref}})$ from Eq. (B12), $V(L_0) - V(L_{\text{ref}}) = B \langle s_{\text{ref}}^2 \rangle$, and Eq. (B33) becomes

$$\langle y_{\text{ref}} \rangle = \langle y_0 \rangle - \kappa B \langle s_{\text{ref}}^2 \rangle. \quad (\text{B34})$$

This step uses Eq. (B11) and thus the assumption of Eq. (B8), that is, the shifts $\kappa v_L(t_i)$ of the individual samples are assumed not to contribute to the variance of a record, although they do contribute to its mean. For a single record, let $\bar{y}_0$ denote the mean of $y_{L_0}(t_i)$ over its K samples, which is the measured mean of $\log_{10} \text{INP}(t)$ and estimates $\langle y_0 \rangle$. Then $\langle y_{\text{ref}} \rangle$ is estimated by $\bar{y}_{\text{ref}} = \bar{y}_0 - \kappa B s_0^2 / A$, with s_0^2 / A estimating $\langle s_{\text{ref}}^2 \rangle$ as in Sect. B3.

Author contributions. **Kristoffer Kvist:** Conceptualization, Data curation, Methodology, Formal analysis, Software, Visualization, and Writing (original draft preparation). **Eirik Mikal Samuelsen:** Project administration, Supervision, and Writing (review and editing). **Rune Grand Graversen:** Project administration, Supervision, and Writing (review and editing).

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work benefited from data infrastructure maintained by several institutions. INP observations were obtained from the NSF National Center for Atmospheric Research (NCAR) Earth Observing Laboratory, the U.S. Department of Energy Atmospheric Radiation Measurement (ARM) User Facility, the ETH Library research collections, the Research Data Leeds Repository at the University of Leeds, and the PANGAEA data publisher for Earth and environmental science. The authors thank the measurement teams and data curators who collected, quality-controlled, and archived the observations compiled in this study, without which this work would not have been possible. AI tools were used for supervised code generation and testing and for proofreading the manuscript.



References

- Armour, K., Forster, P., Storelvmo, T., Collins, W., Dufresne, J.-L., Frame, D., Lunt, D., Mauritsen, T., Palmer, M., Watanabe, M., et al.: The Earths energy budget, climate feedbacks, and climate sensitivity, in: AGU Fall Meeting Abstracts, vol. 2021, pp. U13B–07, 2021.
- Atkinson, J. D., Murray, B. J., Woodhouse, M. T., Whale, T. F., Baustian, K. J., Carslaw, K. S., Dobbie, S., O’Sullivan, D., and Malkin, T. L.: The importance of feldspar for ice nucleation by mineral dust in mixed-phase clouds, *Nature*, 498, 355–358, <https://doi.org/10.1038/nature12278>, 2013.
- Bergeron, T.: On the physics of clouds and precipitation, *Proc. 5th Assembly UGGI, Lisbon, Portugal*, 1935, pp. 156–180, 1935.
- Bergner, N., Marsh, G., Barry, K., Lacher, L., Böhmländer, A., Alden, J., Ahlqvist, C., Althuler, I., Bröder, L., Farinotti, D., et al.: Ice-nucleating particles in Greenlandic glacial outwash plains, *Atmospheric Chemistry and Physics*, 26, 10 557–10 585, <https://doi.org/10.5194/acp-26-10557-2026>, 2026.
- Bodas-Salcedo, A., Hill, P., Furtado, K., Williams, K., Field, P., Manners, J., Hyder, P., and Kato, S.: Large contribution of supercooled liquid clouds to the solar radiation budget of the Southern Ocean, *Journal of Climate*, 29, 4213–4228, <https://doi.org/10.1175/jcli-d-15-0564.1>, 2016.
- Burrows, S. M., Ogunro, O., Frossard, A., Russell, L. M., Rasch, P. J., and Elliott, S.: A physically based framework for modeling the organic fractionation of sea spray aerosol from bubble film Langmuir equilibria, *Atmospheric Chemistry and Physics*, 14, 13 601–13 629, <https://doi.org/10.5194/acp-14-13601-2014>, 2014.
- Burrows, S. M., Easter, R. C., Liu, X., Ma, P.-L., Wang, H., Elliott, S. M., Singh, B., Zhang, K., and Rasch, P. J.: OCEANFILMS (Organic Compounds from Ecosystems to Aerosols: Natural Films and Interfaces via Langmuir Molecular Surfactants) sea spray organic aerosol emissions–implementation in a global climate model and impacts on clouds, *Atmospheric Chemistry and Physics*, 22, 5223–5251, <https://doi.org/10.5194/acp-22-5223-2022>, 2022a.
- Burrows, S. M., McCluskey, C. S., Cornwell, G., Steinke, I., Zhang, K., Zhao, B., Zawadowicz, M., Raman, A., Kulkarni, G., China, S., Zelenyuk, A., and DeMott, P. J.: Ice-Nucleating Particles That Impact Clouds and Climate: Observational and Modeling Research Needs, *Reviews of Geophysics*, 60, e2021RG000 745, <https://doi.org/10.1029/2021RG000745>, 2022b.
- Canzi, A., Freney, E., Grzegorzczak, P., Baray, J., Lacher, L., and Planche, C.: Unraveling ice nucleating particle concentration variability: Insights into source emissions origin and parameterizations, *Journal of Geophysical Research: Atmospheres*, 130, e2024JD041 258, <https://doi.org/10.1029/2024jd041258>, 2025.
- Chatziparaschos, M., Daskalakis, N., Myriokefalitakis, S., Kalivitis, N., Nenes, A., Gonçalves Ageitos, M., Costa-Surós, M., Pérez García-Pando, C., Zanolli, M., Vrekoussis, M., et al.: Role of K-feldspar and quartz in global ice nucleation by mineral dust in mixed-phase clouds, *Atmospheric Chemistry and Physics*, 23, 1785–1801, <https://doi.org/10.5194/acp-23-1785-2023>, 2023.
- Chatziparaschos, M., Myriokefalitakis, S., Kalivitis, N., Daskalakis, N., Nenes, A., Gonçalves Ageitos, M., Costa-Surós, M., Pérez García-Pando, C., Vrekoussis, M., and Kanakidou, M.: Assessing the global contribution of marine aerosols, terrestrial bioaerosols, and desert dust to ice-nucleating particle concentrations, *Atmospheric Chemistry and Physics*, 25, 9085–9111, <https://doi.org/10.5194/acp-25-9085-2025>, 2025.
- Chen, J., Wu, Z., Augustin-Bauditz, S., Grawe, S., Hartmann, M., Pei, X., Liu, Z., Ji, D., and Wex, H.: Ice-nucleating particle concentrations unaffected by urban air pollution in Beijing, China, *Atmospheric Chemistry and Physics*, 18, 3523–3539, <https://doi.org/10.5194/acp-18-3523-2018>, 2018.



- Conen, F., Eckhardt, S., Gundersen, H., Stohl, A., and Yttri, K. E.: Rainfall drives atmospheric ice-nucleating particles in the coastal climate
930 of southern Norway, *Atmospheric Chemistry and Physics*, 17, 11 065–11 073, <https://doi.org/10.5194/acp-17-11065-2017>, 2017.
- Cooper, W. A.: Ice initiation in natural clouds, in: *Precipitation enhancement—A scientific challenge*, pp. 29–32, American Meteorological
Society, <https://doi.org/10.1175/0065-9401-21.43.29>, 1986.
- Costa-Surós, M., Gonçalves Ageitos, M., Chatziparaschos, M., Georgakaki, P., Thomas, M. A., Montané Pinto, G., Myriokefalitakis, S., van
Noije, T., Le Sager, P., Kanakidou, M., et al.: Implementation of primary and secondary ice production in EC-Earth3-AerChem: global
935 impacts and insights, *Atmospheric Chemistry and Physics*, 26, 2667–2690, <https://doi.org/10.5194/acp-26-2667-2026>, 2026.
- Creamean, J., Hill, T., Hume, C., Vazquez, M., and Shi, Y.: Ice Nucleating Particles (INP) Raw Data, Tech. rep., Atmospheric Radiation
Measurement User Facility, <https://doi.org/10.5439/1889624>, 2020.
- Creamean, J. M., Hume, C. C., Vazquez, M., and Theisen, A.: Long-term measurements of ice nucleating particles at Atmospheric Radiation
Measurement (ARM) sites worldwide, *Earth System Science Data*, 17, 6943–6963, <https://doi.org/10.5194/essd-17-6943-2025>, 2025.
- 940 Daily, M. I., Tarn, M. D., Whale, T. F., and Murray, B. J.: An evaluation of the heat test for the ice-nucleating ability of minerals and biological
material, *Atmospheric Measurement Techniques*, 15, 2635–2665, <https://doi.org/10.5194/amt-15-2635-2022>, 2022.
- David, R. O., Cascajo-Castresana, M., Brennan, K. P., Rösch, M., Els, N., Werz, J., Weichlinger, V., Boynton, L. S., Bogler, S., Borduas-
Dedekind, N., et al.: Development of the DRoplet Ice Nuclei Counter Zurich (DRINCZ): validation and application to field-collected snow
samples, *Atmospheric Measurement Techniques*, 12, 6865–6888, <https://doi.org/10.5194/amt-12-6865-2019>, 2019.
- 945 DeMott, P. and Hill, T.: Cloud, Aerosol, and Complex Terrain Interactions (CACTI) ARM Mobile Facility (AMF) Mea-
surements of Ice Nucleating Particles Field Campaign Report. Tech. Doc, ARM user facility. DOE/SC-ARM-20-008.
<https://www.arm.gov/publications/programdocs/doe-sc-arm-20-008.pdf>, 2020.
- DeMott, P. J., Rogers, D. C., and Kreidenweis, S. M.: The susceptibility of ice formation in upper tropospheric clouds to insoluble aerosol
components, *Journal of Geophysical Research: Atmospheres*, 102, 19 575–19 584, <https://doi.org/10.1029/97jd01138>, 1997.
- 950 DeMott, P. J., Prenni, A. J., Liu, X., Kreidenweis, S. M., Petters, M. D., Twohy, C. H., Richardson, M. S., Eidhammer, T., and Rogers, D. C.:
Predicting global atmospheric ice nuclei distributions and their impacts on climate, *Proceedings of the National Academy of Sciences*,
107, 11 217–11 222, <https://doi.org/10.1073/pnas.0910818107>, 2010.
- DeMott, P. J., Hill, T. C., McCluskey, C. S., Prather, K. A., Collins, D. B., Sullivan, R. C., Ruppel, M. J., Mason, R. H., Irish, V. E., Lee, T.,
et al.: Sea spray aerosol as a unique source of ice nucleating particles, *Proceedings of the National Academy of Sciences*, 113, 5797–5803,
955 <https://doi.org/10.1073/pnas.1514034112>, 2016.
- DeMott, P. J., Hill, T. C., Marchand, R., and Alexander, S.: Macquarie Island Cloud and Radiation Experiment (MICRE) Ice Nucleating Par-
ticle Measurements Field Campaign Report, Tech. rep., DOE Office of Science Atmospheric Radiation Measurement (ARM) Program . . . ,
2018a.
- DeMott, P. J., Möhler, O., Cziczo, D. J., Hiranuma, N., Petters, M. D., Petters, S. S., Belosi, F., Bingemer, H. G., Brooks, S. D., Budke, C.,
960 et al.: The Fifth International Workshop on Ice Nucleation phase 2 (FIN-02): laboratory intercomparison of ice nucleation measurements,
Atmospheric Measurement Techniques, 11, 6231–6257, <https://doi.org/10.5194/amt-11-6231-2018>, 2018b.
- DeMott, P. J., Swanson, B. E., Creamean, J. M., Tobo, Y., Hill, T. C., Barry, K. R., Beck, I. F., Frietas, G. P., Heslin-Rees, D., Lackner, C. P.,
et al.: Ice nucleating particle sources and transports between the Central and Southern Arctic regions during winter cold air outbreaks,
Elem Sci Anth, 13, 00 063, <https://doi.org/10.1525/elementa.2024.00063>, 2025.
- 965 Dupont, R., Taymans, C., and Vié, B.: Improving supercooled liquid water representation in the microphysical scheme ICE3, *Quarterly*
Journal of the Royal Meteorological Society, 150, 4086–4112, <https://doi.org/10.1002/qj.4806>, 2024.



- Engdahl, B. J. K., Thompson, G., and Bengtsson, L.: Improving the representation of supercooled liquid water in the HARMONIE-AROME weather forecast model, *Tellus A: Dynamic Meteorology and Oceanography*, 72, 1–18, <https://doi.org/10.1080/16000870.2019.1697603>, 2020.
- 970 Fan, J., Ghan, S., Ovchinnikov, M., Liu, X., Rasch, P. J., and Korolev, A.: Representation of Arctic mixed-phase clouds and the Wegener-Bergeron-Findeisen process in climate models: Perspectives from a cloud-resolving study, *Journal of Geophysical Research: Atmospheres*, 116, <https://doi.org/10.1029/2010jd015375>, 2011.
- Findeisen, W.: *Kolloid-meteorologische vorgänge bei niederschlagsbildung*, (No Title), 55, 121, 1938.
- Fletcher, N. H. et al.: *The physics of rainclouds*, Cambridge University Press, 2011.
- 975 Frey, W. R. and Kay, J. E.: The influence of extratropical cloud phase and amount feedbacks on climate sensitivity, *Climate dynamics*, 50, 3097–3116, <https://doi.org/10.1007/s00382-017-3796-5>, 2018.
- Gettelman, A., Bardeen, C., McCluskey, C. S., Järvinen, E., Stith, J., Bretherton, C., McFarquhar, G., Twohy, C., D’Alessandro, J., and Wu, W.: Simulating observations of Southern Ocean clouds and implications for climate, *Journal of Geophysical Research: Atmospheres*, 125, e2020JD032 619, <https://doi.org/10.1029/2020jd032619>, 2020.
- 980 Gjelsvik, A. B., David, R. O., Carlsen, T., Hellmuth, F., Hofer, S., McGraw, Z., Sodemann, H., and Storelvmo, T.: Using a region-specific ice-nucleating particle parameterization improves the representation of Arctic clouds in a global climate model, *Atmospheric Chemistry and Physics*, 25, 1617–1637, <https://doi.org/10.5194/acp-25-1617-2025>, 2025.
- Gratzl, J., Böhmmländer, A., Pätz, S., Pogner, C.-E., Gorfer, M., Brus, D., Doulgeris, K. M., Wieland, F., Asmi, E., Saarto, A., Möhler, O., Stolzenburg, D., and Grothe, H.: Locally emitted fungal spores serve as high-temperature ice nucleating particles in the European sub-Arctic, *Atmospheric Chemistry and Physics*, 25, 12 007–12 035, <https://doi.org/10.5194/acp-25-12007-2025>, 2025.
- 985 Hande, L. B. and Hoose, C.: Partitioning the primary ice formation modes in large eddy simulations of mixed-phase clouds, *Atmospheric Chemistry and Physics*, 17, 14 105–14 118, <https://doi.org/10.5194/acp-17-14105-2017>, 2017.
- Harrison, A. D., Lever, K., Sanchez-Marroquin, A., Holden, M. A., Whale, T. F., Tarn, M. D., McQuaid, J. B., and Murray, B. J.: The ice-nucleating ability of quartz immersed in water and its atmospheric importance compared to K-feldspar, *Atmospheric Chemistry and Physics*, 19, 11 343–11 361, <https://doi.org/10.5194/acp-19-11343-2019>, 2019.
- 990 Harrison, A. D., O’Sullivan, D., Adams, M. P., Porter, G. C., Blades, E., Brathwaite, C., Chewitt-Lucas, R., Gaston, C., Hawker, R., Krüger, O. O., et al.: The ice-nucleating activity of African mineral dust in the Caribbean boundary layer, *Atmospheric Chemistry and Physics*, 22, 9663–9680, <https://doi.org/10.5194/acp-22-9663-2022>, 2022.
- Hartmann, M., Gong, X., Kecorius, S., van Pinxteren, M., Vogl, T., Welti, A., Wex, H., Zeppenfeld, S., Herrmann, H., Wiedensohler, A., et al.: Terrestrial or marine—indications towards the origin of ice-nucleating particles during melt season in the European Arctic up to 83.7° N, *Atmospheric Chemistry and Physics*, 21, 11 613–11 636, <https://doi.org/10.5194/acp-21-11613-2021>, 2021.
- 995 Hawker, R. E., Miltenberger, A. K., Wilkinson, J. M., Hill, A. A., Shipway, B. J., Cui, Z., Cotton, R. J., Carslaw, K. S., Field, P. R., and Murray, B. J.: The temperature dependence of ice-nucleating particle concentrations affects the radiative properties of tropical convective cloud systems, *Atmospheric Chemistry and Physics*, 21, 5439–5461, <https://doi.org/10.5194/acp-21-5439-2021>, 2021.
- 1000 Herbert, R. J., Sanchez-Marroquin, A., Grosvenor, D. P., Pringle, K. J., Arnold, S. R., Murray, B. J., and Carslaw, K. S.: Gaps in our understanding of ice-nucleating particle sources exposed by global simulation of the UK Earth System Model, *Atmospheric Chemistry and Physics*, 25, 291–325, <https://doi.org/10.5194/acp-25-291-2025>, 2025.



- Herbert, R. J., Lacher, L., Böhmmländer, A., Tarn, M. D., Canzi, A., Pantoya, A., Freney, E., Höhler, K., Bogert, P., Planche, C., et al.: GloPINE dataset: model-ready measurements of INP concentrations using PINE instruments, *Earth System Science Data*, 18, 6465–6483, <https://doi.org/10.5194/essd-18-6465-2026>, 2026.
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al.: The ERA5 global reanalysis, *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049, <https://doi.org/10.1002/qj.3803>, 2020.
- Hill, T. C., Moffett, B. F., DeMott, P. J., Georgakopoulos, D. G., Stump, W. L., and Franc, G. D.: Measurement of ice nucleation-active bacteria on plants and in precipitation by quantitative PCR, *Applied and environmental microbiology*, 80, 1256–1267, <https://doi.org/10.1128/aem.02967-13>, 2014.
- Hill, T. C. J., DeMott, P. J., Tobo, Y., Fröhlich-Nowoisky, J., Moffett, B. F., Franc, G. D., and Kreidenweis, S. M.: Sources of organic ice nucleating particles in soils, *Atmospheric Chemistry and Physics*, 16, 7195–7211, <https://doi.org/10.5194/acp-16-7195-2016>, 2016.
- Hiranuma, N., Augustin-Bauditz, S., Bingemer, H., Budke, C., Curtius, J., Danielczok, A., Diehl, K., Dreischmeier, K., Ebert, M., Frank, F., et al.: A comprehensive laboratory study on the immersion freezing behavior of illite NX particles: a comparison of 17 ice nucleation measurement techniques, *Atmospheric Chemistry and Physics*, 15, 2489–2518, <https://doi.org/10.5194/acp-15-2489-2015>, 2015.
- Hoose, C. and Möhler, O.: Heterogeneous ice nucleation on atmospheric aerosols: a review of results from laboratory experiments, *Atmospheric Chemistry and Physics*, 12, 9817–9854, <https://doi.org/10.5194/acp-12-9817-2012>, 2012.
- Hoose, C., Kristjánsson, J. E., Chen, J.-P., and Hazra, A.: A classical-theory-based parameterization of heterogeneous ice nucleation by mineral dust, soot, and biological particles in a global climate model, *Journal of the Atmospheric Sciences*, 67, 2483–2503, <https://doi.org/10.1175/2010jas3425.1>, 2010.
- Huang, S., Hu, W., Chen, J., Wu, Z., Zhang, D., and Fu, P.: Overview of biological ice nucleating particles in the atmosphere, *Environment International*, 146, 106 197, <https://doi.org/10.1016/j.envint.2020.106197>, 2021.
- Huffman, J. A., Prenni, A., DeMott, P., Pöhlker, C., Mason, R., Robinson, N., Fröhlich-Nowoisky, J., Tobo, Y., Després, V., Garcia, E., et al.: High concentrations of biological aerosol particles and ice nuclei during and after rain, *Atmospheric Chemistry and Physics*, 13, 6151–6164, <https://doi.org/10.5194/acp-13-6151-2013>, 2013.
- Korolev, A., McFarquhar, G., Field, P. R., Franklin, C., Lawson, P., Wang, Z., Williams, E., Abel, S. J., Axisa, D., Borrmann, S., et al.: Mixed-phase clouds: Progress and challenges, *Meteorological Monographs*, 58, 5.1–5.50, <https://doi.org/10.1175/amsmonographs-d-17-0001.1>, 2017.
- Kvist, K.: Ice-nucleating particle dataset of “The spatio-temporal variability of ice-nucleating particle concentrations quantified with a new benchmark dataset”, <https://doi.org/10.5281/zenodo.23054649>, 2026.
- Lacher, L., Adams, M. P., Barry, K., Bertozzi, B., Bingemer, H., Boffo, C., Bras, Y., Büttner, N., Castarede, D., Cziczo, D. J., et al.: The Puy de Dôme ICe Nucleation Intercomparison Campaign (PICNIC): comparison between online and offline methods in ambient air, *Atmospheric Chemistry and Physics*, 24, 2651–2678, <https://doi.org/10.5194/acp-24-2651-2024>, 2024.
- Lata, N. N., Diep, T., Gilardon, S., Mazzola, M., Cheng, Z., Rahman, A., Rogers, M. M., Fraund, M., Marcus, M. A., Hiranuma, N., et al.: Measurement report: Role of organic coating and chemical composition on ice nucleation potential of atmospheric particles in European Arctic, *Atmospheric Chemistry and Physics*, 26, 6611–6628, <https://doi.org/10.5194/acp-26-6611-2026>, 2026.
- Li, G., Welti, A., Rocchi, A., Fogwill, G. P., Dall’Osto, M., and Kanji, Z. A.: Terrestrial and marine sources of ice nucleating particles in the Eurasian Arctic, *Faraday Discussions*, 258, 94–119, <https://doi.org/10.1039/d4fd00160e>, 2025.



- Mann, G. W., Carslaw, K., Spracklen, D., Ridley, D., Manktelow, P., Chipperfield, M., Pickering, S., and Johnson, C.: Description and
1040 evaluation of GLOMAP-mode: A modal global aerosol microphysics model for the UKCA composition-climate model, *Geoscientific
Model Development*, 3, 519–551, <https://doi.org/10.5194/gmd-3-519-2010>, 2010.
- McCluskey, C. S., Hill, T. C., Humphries, R. S., Rauker, A. M., Moreau, S., Strutton, P., Chambers, S. D., Williams, A. G., McRobert,
I., Ward, J., et al.: Observations of ice nucleating particles over Southern Ocean waters, *Geophysical Research Letters*, 45, 11–989,
<https://doi.org/10.1029/2018gl079981>, 2018a.
- 1045 McCluskey, C. S., Hill, T. C., Sultana, C. M., Laskina, O., Trueblood, J., Santander, M. V., Beall, C. M., Michaud, J. M., Kreidenweis, S. M.,
Prather, K. A., et al.: A mesocosm double feature: Insights into the chemical makeup of marine ice nucleating particles, *Journal of the
Atmospheric Sciences*, 75, 2405–2423, <https://doi.org/10.1175/jas-d-17-0155.1>, 2018b.
- McCluskey, C. S., Ovadnevaite, J., Rinaldi, M., Atkinson, J., Belosi, F., Ceburnis, D., Marullo, S., Hill, T. C. J., Lohmann, U., Kanji,
Z. A., O’Dowd, C., Kreidenweis, S. M., and DeMott, P. J.: Marine and terrestrial organic ice-nucleating particles in pristine ma-
1050 rine to continentally influenced Northeast Atlantic air masses, *Journal of Geophysical Research: Atmospheres*, 123, 6196–6212,
<https://doi.org/10.1029/2017JD028033>, 2018c.
- McCluskey, C. S., Gettelman, A., Bardeen, C. G., DeMott, P. J., Moore, K. A., Kreidenweis, S. M., Hill, T. C., Barry, K. R., Twohy, C. H.,
Toohey, D. W., et al.: Simulating Southern Ocean aerosol and ice nucleating particles in the Community Earth System Model version 2,
Journal of Geophysical Research: Atmospheres, 128, e2022JD036955, <https://doi.org/10.1029/2022jd036955>, 2023.
- 1055 Meyers, M. P., DeMott, P. J., and Cotton, W. R.: New primary ice-nucleation parameterizations in an explicit cloud model, *Journal of Applied
Meteorology and Climatology*, 31, 708–721, [https://doi.org/10.1175/1520-0450\(1992\)031<0708:npinpi>2.0.co;2](https://doi.org/10.1175/1520-0450(1992)031<0708:npinpi>2.0.co;2), 1992.
- Miller, A. J., Brennan, K. P., Mignani, C., Wieder, J., David, R. O., and Borduas-Dedekind, N.: Development of the drop Freezing Ice Nuclei
Counter (FINC), intercomparison of droplet freezing techniques, and use of soluble lignin as an atmospheric ice nucleation standard,
Atmospheric Measurement Techniques, 14, 3131–3151, <https://doi.org/10.5194/amt-14-3131-2021>, 2021.
- 1060 Möhler, O., Adams, M., Lacher, L., Vogel, F., Nadolny, J., Ullrich, R., Boffo, C., Pfeuffer, T., Hobl, A., Weiß, M., et al.: The Portable
Ice Nucleation Experiment (PINE): a new online instrument for laboratory studies and automated long-term field observations of ice-
nucleating particles, *Atmospheric Measurement Techniques*, 14, 1143–1166, <https://doi.org/10.5194/amt-14-1143-2021>, 2021.
- Moore, K. A., Hill, T. C., McCluskey, C. S., Twohy, C. H., Rainwater, B., Toohey, D. W., Sanchez, K. J., Kreidenweis, S. M., and DeMott, P. J.:
Characterizing ice nucleating particles over the Southern Ocean using simultaneous aircraft and ship observations, *Journal of Geophysical
1065 Research: Atmospheres*, 129, e2023JD039543, <https://doi.org/10.1029/2023jd039543>, 2024.
- Murray, B., O’sullivan, D., Atkinson, J., and Webb, M.: Ice nucleation by particles immersed in supercooled cloud droplets, *Chemical Society
Reviews*, 41, 6519–6554, <https://doi.org/10.1039/c2cs35200a>, 2012.
- Murray, B. J., Carslaw, K. S., and Field, P. R.: Opinion: Cloud-phase climate feedback and the importance of ice-nucleating particles,
Atmospheric Chemistry and Physics, 21, 665–679, <https://doi.org/10.5194/acp-21-665-2021>, 2021.
- 1070 Myriokefalitakis, S., Daskalakis, N., Mihalopoulos, N., Baker, A., Nenes, A., and Kanakidou, M.: Changes in dissolved iron deposition to the
oceans driven by human activity: a 3-D global modelling study, *Biogeosciences*, 12, 3973–3992, <https://doi.org/10.5194/bg-12-3973-2015>,
2015.
- Niemand, M., Möhler, O., Vogel, B., Vogel, H., Hoose, C., Connolly, P., Klein, H., Bingemer, H., DeMott, P., Skrotzki, J., et al.: A particle-
surface-area-based parameterization of immersion freezing on desert dust particles, *Journal of the Atmospheric Sciences*, 69, 3077–3092,
1075 <https://doi.org/10.1175/jas-d-11-0249.1>, 2012.



- O'Sullivan, D., Murray, B., Malkin, T., Whale, T. F., Umo, N. S., Atkinson, J., Price, H., Baustian, K. J., Browse, J., and Webb, M.: Ice nucleation by fertile soil dusts: relative importance of mineral and biogenic components, *Atmospheric Chemistry and Physics*, 14, 1853–1867, <https://doi.org/10.5194/acp-14-1853-2014>, 2014.
- Ott, W. R.: A physical explanation of the lognormality of pollutant concentrations, *Journal of the Air & Waste Management Association*, 40, 1378–1383, <https://doi.org/10.1080/10473289.1990.10466789>, 1990.
- Peckhaus, A., Kiselev, A., Hiron, T., Ebert, M., and Leisner, T.: A comparative study of K-rich and Na/Ca-rich feldspar ice-nucleating particles in a nanoliter droplet freezing assay, *Atmospheric Chemistry and Physics*, 16, 11 477–11 496, <https://doi.org/10.5194/acp-16-11477-2016>, 2016.
- Perkins, R. and Barry, K.: CAESAR: C-130 Ice Spectrometer Measurements, Version 1.0, <https://doi.org/10.26023/P9VX-CD1Y-4Z03>, accessed: 2 June 2026, 2025.
- Perkins, R. J., Gillette, S. M., Hill, T. C., and DeMott, P. J.: The labile nature of ice nucleation by Arizona Test Dust, *ACS Earth and Space Chemistry*, 4, 133–141, <https://doi.org/10.1021/acsearthspacechem.9b00304>, 2020.
- Prenni, A., Tobo, Y., Garcia, E., DeMott, P., Huffman, J., McCluskey, C., Kreidenweis, S., Prenni, J., Pöhlker, C., and Pöschl, U.: The impact of rain on ice nuclei populations at a forested site in Colorado, *Geophysical Research Letters*, 40, 227–231, <https://doi.org/10.1029/2012GL053953>, 2013.
- Price, H., Baustian, K., McQuaid, J., Blyth, A., Bower, K., Choularton, T., Cotton, R., Cui, Z., Field, P., Gallagher, M., et al.: Atmospheric ice-nucleating particles in the dusty tropical Atlantic, *Journal of Geophysical Research: Atmospheres*, 123, 2175–2193, <https://doi.org/10.1002/2017jd027560>, 2018.
- Raif, E. N., Barr, S. L., Tarn, M. D., McQuaid, J. B., Daily, M. I., Abel, S. J., Barrett, P. A., Bower, K. N., Field, P. R., Carslaw, K. S., et al.: High ice-nucleating particle concentrations associated with Arctic haze in springtime cold-air outbreaks, *Atmospheric Chemistry and Physics*, 24, 14 045–14 072, <https://doi.org/10.5194/acp-24-14045-2024>, 2024.
- Rosenfeld, D., Chemke, R., DeMott, P., Sullivan, R. C., Rasmussen, R., McDonough, F., Comstock, J., Schmid, B., Tomlinson, J., Jonsson, H., et al.: The common occurrence of highly supercooled drizzle and rain near the coastal regions of the western United States, *Journal of Geophysical Research: Atmospheres*, 118, 9819–9833, <https://doi.org/10.1002/jgrd.50529>, 2013.
- Sanchez-Marroquin, A., Arnalds, O., Baustian-Dorsi, K., Browse, J., Dagsson-Waldhauserova, P., Harrison, A., Maters, E., Pringle, K., Vergara-Temprado, J., Burke, I., et al.: Iceland is an episodic source of atmospheric ice-nucleating particles relevant for mixed-phase clouds, *Science advances*, 6, eaba8137, <https://doi.org/10.1126/sciadv.aba8137>, 2020.
- Sassen, K. and Benson, S.: Ice nucleation in cirrus clouds: A model study of the homogeneous and heterogeneous modes, *Geophysical research letters*, 27, 521–524, <https://doi.org/10.1029/1999gl010883>, 2000.
- Sato, K. and Inoue, J.: Seasonal change in satellite-retrieved lower-tropospheric ice-cloud fraction over the Southern Ocean, *Geophysical Research Letters*, 48, e2021GL095 295, <https://doi.org/10.1029/2021GL095295>, 2021.
- Schiebel, T.: Ice nucleation activity of soil dust aerosols, Ph.D. thesis, Dissertation, Karlsruhe, Karlsruher Institut für Technologie (KIT), 2017, 2017.
- Schneider, J., Höhler, K., Heikkilä, P., Keskinen, J., Bertozzi, B., Bogert, P., Schorr, T., Umo, N. S., Vogel, F., Brasseur, Z., Wu, Y., Hakala, S., Duplissy, J., Moiseev, D., Kulmala, M., Adams, M. P., Murray, B. J., Korhonen, K., Hao, L., Thomson, E. S., Castarède, D., Leisner, T., Petäjä, T., and Möhler, O.: The seasonal cycle of ice-nucleating particles linked to the abundance of biogenic aerosol in boreal forests, *Atmospheric Chemistry and Physics*, 21, 3899–3918, <https://doi.org/10.5194/acp-21-3899-2021>, 2021.



- Schrod, J., Thomson, E. S., Weber, D., Kossmann, J., Pöhlker, C., Saturno, J., Ditas, F., Artaxo, P., Clouard, V., Saurel, J.-M., et al.: Long-term deposition and condensation ice-nucleating particle measurements from four stations across the globe, *Atmospheric Chemistry and Physics*, 20, 15 983–16 006, <https://doi.org/10.5194/acp-20-15983-2020>, 2020.
- 1115 Shete, S., Patade, S., Malap, N., Bankar, S., Jaykumar, S., Baumgardner, D., Axisa, D., Murugavel, P., and Prabhakaran, T.: Observations and simulations of immersion freezing INPs and their impact on cloud properties over the rain shadow region of India, *Atmospheric Research*, p. 109056, <https://doi.org/10.1016/j.atmosres.2026.109056>, 2026.
- Silva, K. M. R. d., Herdies, D. L., Kubota, P. Y., Bresciani, C., and Figueroa, S. N.: Impact of Aerosols on Cloud Microphysical Processes: A Theoretical Review., *Geosciences* (2076-3263), 15, <https://doi.org/10.3390/geosciences15080312>, 2025.
- 1120 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, *Atmospheric measurement techniques*, 7, 129–134, <https://doi.org/10.5194/amt-7-129-2014>, 2014.
- Storelvmo, T., Tan, I., and Korolev, A. V.: Cloud phase changes induced by CO₂ warming—A powerful yet poorly constrained cloud-climate feedback, *Current climate change reports*, 1, 288–296, <https://doi.org/10.1007/s40641-015-0026-2>, 2015.
- 1125 Sze, K. C., Wex, H., Hartmann, M., Skov, H., Massling, A., Villanueva, D., and Stratmann, F.: Ice-nucleating particles in northern Greenland: annual cycles, biological contribution and parameterizations, *Atmospheric Chemistry and Physics*, 23, 4741–4761, <https://doi.org/10.5194/acp-23-4741-2023>, 2023.
- Tan, I. and Storelvmo, T.: Evidence of strong contributions from mixed-phase clouds to Arctic climate change, *Geophysical Research Letters*, 46, 2894–2902, <https://doi.org/10.1029/2018gl081871>, 2019.
- 1130 Tan, I., Storelvmo, T., and Zelinka, M. D.: Observational constraints on mixed-phase clouds imply higher climate sensitivity, *Science*, 352, 224–227, <https://doi.org/10.1126/science.aad5300>, 2016.
- Taylor, P. C., Boeke, R. C., Boisvert, L. N., Feldl, N., Henry, M., Huang, Y., Langen, P. L., Liu, W., Pithan, F., Sejas, S. A., et al.: Process drivers, inter-model spread, and the path forward: A review of amplified Arctic warming, *Frontiers in Earth Science*, 9, 758 361, <https://doi.org/10.3389/feart.2021.758361>, 2022.
- 1135 Testa, B., Hill, T. C., Marsden, N. A., Barry, K. R., Hume, C. C., Bian, Q., Uetake, J., Hare, H., Perkins, R. J., Möhler, O., et al.: Ice nucleating particle connections to regional Argentinian land surface emissions and weather during the Cloud, Aerosol, and Complex Terrain Interactions experiment, *Journal of Geophysical Research: Atmospheres*, 126, e2021JD035 186, <https://doi.org/10.1029/2021jd035186>, 2021.
- Thompson, G., Rasmussen, R. M., and Manning, K.: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis, *Monthly Weather Review*, 132, 519–542, [https://doi.org/10.1175/1520-0493\(2004\)132<0519:efowpu>2.0.co;2](https://doi.org/10.1175/1520-0493(2004)132<0519:efowpu>2.0.co;2), 2004.
- 1140 Thompson, G., Field, P. R., Rasmussen, R. M., and Hall, W. D.: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization, *Monthly weather review*, 136, 5095–5115, <https://doi.org/10.1175/2008mwr2387.1>, 2008.
- 1145 Tobo, Y.: An improved approach for measuring immersion freezing in large droplets over a wide temperature range, *Scientific reports*, 6, 32 930, <https://doi.org/10.1038/srep32930>, 2016.
- 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, *Nature Geoscience*, 12, 253–258, <https://doi.org/10.1038/s41561-019-0314-x>, 2019.



- 1150 Tobo, Y., Uetake, J., Matsui, H., Moteki, N., Uji, Y., Iwamoto, Y., Miura, K., and Misumi, R.: Seasonal trends of atmospheric ice nucleating particles over Tokyo, *Journal of Geophysical Research: Atmospheres*, 125, e2020JD033 658, <https://doi.org/10.1029/2020jd033658>, 2020.
- Tobo, Y., Adachi, K., Kawai, K., Matsui, H., Ohata, S., Oshima, N., Kondo, Y., Hermansen, O., Uchida, M., Inoue, J., and Koike, M.: Surface warming in Svalbard may have led to increases in highly active ice-nucleating particles, *Communications Earth & Environment*, 5, 516, <https://doi.org/10.1038/s43247-024-01677-0>, 2024.
- 1155 Vali, G.: Quantitative evaluation of experimental results an the heterogeneous freezing nucleation of supercooled liquids, *Journal of Atmospheric Sciences*, 28, 402–409, [https://doi.org/10.1175/1520-0469\(1971\)028<0402:qeoera>2.0.co;2](https://doi.org/10.1175/1520-0469(1971)028<0402:qeoera>2.0.co;2), 1971.
- Vali, G.: Interpretation of freezing nucleation experiments: singular and stochastic; sites and surfaces, *Atmospheric Chemistry and Physics*, 14, 5271–5294, <https://doi.org/10.5194/acp-14-5271-2014>, 2014.
- Vergara-Temprado, J., Murray, B. J., Wilson, T. W., O’Sullivan, D., Browse, J., Pringle, K. J., Ardon-Dryer, K., Bertram, A. K., Burrows, S. M., Ceburnis, D., et al.: Contribution of feldspar and marine organic aerosols to global ice nucleating particle concentrations, *Atmospheric Chemistry and Physics*, 17, 3637–3658, <https://doi.org/10.5194/acp-17-3637-2017>, 2017.
- 1160 Vignon, É., Alexander, S., DeMott, P., Sotiropoulou, G., Gerber, F., Hill, T., Marchand, R., Nenes, A., and Berne, A.: Challenging and improving the simulation of mid-level mixed-phase clouds over the high-latitude Southern Ocean, *Journal of Geophysical Research: Atmospheres*, 126, e2020JD033 490, <https://doi.org/10.1029/2020jd033490>, 2021.
- 1165 Wang, Y., Liu, X., Hoose, C., and Wang, B.: Different contact angle distributions for heterogeneous ice nucleation in the Community Atmospheric Model version 5, *Atmospheric Chemistry and Physics*, 14, 10 411–10 430, <https://doi.org/10.5194/acp-14-10411-2014>, 2014.
- Wegener, A.: *Thermodynamik der atmosphäre*, JA Barth, 1911.
- Welti, A., Müller, K., Fleming, Z. L., and Stratmann, F.: Concentration and variability of ice nuclei in the subtropical maritime boundary layer, *Atmospheric Chemistry and Physics*, 18, 5307–5320, <https://doi.org/10.5194/acp-18-5307-2018>, 2018.
- 1170 Welti, A., Bigg, E. K., DeMott, P. J., Gong, X., Hartmann, M., Harvey, M., Henning, S., Herenz, P., Hill, T. C., Hornblow, B., et al.: Ship-based measurements of ice nuclei concentrations over the Arctic, Atlantic, Pacific and Southern oceans, *Atmospheric Chemistry and Physics*, 20, 15 191–15 206, <https://doi.org/10.5194/acp-20-15191-2020>, 2020.
- Wex, H., Huang, L., Zhang, W., Hung, H., Traversi, R., Becagli, S., Sheesley, R. J., Moffett, C. E., Barrett, T. E., Bossi, R., et al.: Annual variability of ice-nucleating particle concentrations at different Arctic locations, *Atmospheric Chemistry and Physics*, 19, 5293–5311, <https://doi.org/10.5194/acp-19-5293-2019>, 2019.
- 1175 Wex, H., Eckermann, O., Jurányi, Z., Weller, R., Mangold, A., Van Overmeiren, P., Zeppenfeld, S., van Pinxteren, M., Dall’Osto, M., and Henning, S.: Antarctica’s unique atmosphere: Really low INP concentrations, *Geophysical Research Letters*, 52, e2024GL112 583, <https://doi.org/10.1029/2024gl112583>, 2025.
- Wilks, D. S.: *Statistical methods in the atmospheric sciences*, vol. 627, Academic press London, 2006.
- 1180 Zhao, B., Wang, Y., Gu, Y., Liou, K.-N., Jiang, J. H., Fan, J., Liu, X., Huang, L., and Yung, Y. L.: Ice nucleation by aerosols from anthropogenic pollution, *Nature Geoscience*, 12, 602–607, <https://doi.org/10.1038/s41561-019-0389-4>, 2019.
- Zhao, X., Liu, X., Burrows, S. M., and Shi, Y.: Effects of marine organic aerosols as sources of immersion-mode ice-nucleating particles on high-latitude mixed-phase clouds, *Atmospheric Chemistry and Physics*, 21, 2305–2327, <https://doi.org/10.5194/acp-21-2305-2021>, 2021.
- Zhou, R., Perkins, R., Juergensen, D., Barry, K., Ayars, K., Dutton, O., DeMott, P., and Kreidenweis, S.: Seasonal variability, sources, and parameterization of ice-nucleating particles in the Rocky Mountain region: importance of soil dust and biological contributions, *Atmospheric Chemistry and Physics*, 26, 1515–1535, <https://doi.org/10.5194/acp-26-1515-2026>, 2026.