https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

A Data-Driven Approach to Modelling a Global Distribution of Ice Nucleating Particles

Gabriella Wallentin[1], Alexander Böhmländer[2], Ross Herbert[3], Heike Wex[4], Charlotte M. Beall[5], Pia Bogert[2], Zoé Brasseur[6,7], Jie Chen[8], Jessie M. Creamean[9,10], Paul J. DeMott[9], Oliver Eckermann[4], Kunfeng Gao[8,11], Xianda Gong[12], Silvia Henning[4], Naruki Hiranuma[13], Christina S. McCluskey[14], Christos Mitsios[11,15,16], Ottmar Möhler[2], Athanasios Nenes[11,15,16,17], Mark D. Tarn[18], Christian Tatzelt[4], Ping Tian[19], Yutaka Tobo[20,21], Franziska Vogel[22], André Welti[23], Elise Wilbourn[24], Jennifer Winstone[1], and Corinna Hoose[1]

[1]Institute of Meteorology and Climate Research Troposphere Research (IMKTRO), Karlsruhe Institute of Technology (KIT), Germany
[2]Institute of Meteorology and Climate Research Atmospheric Aerosol Research (IMKAAF), Karlsruhe Institute of Technology (KIT), Germany
[3]Institute for Climate and Atmospheric Science, University of Leeds, UK
[4]Leibniz Institute for Tropospheric Research, Leibniz Institute, Germany
[5]Renaissance Philanthropy Fund, USA
[6]Institute for Atmospheric and Earth System Research/Physics, University of Helsinki, Finland
[7]SIOS Knowledge Centre, Norway
[8]Institute for Atmospheric and Climate Science, ETH Zürich, Switzerland
[9]Department of Atmospheric Science, Colorado State University, USA
[10]Department of Environmental Science, Stockholm University, Sweden
[11]Laboratory of Atmospheric Processes and Their Impacts, School of Architecture, Civil and Environmental Engineering, École Polytechnique Fédérale de Lausanne, Switzerland
[12]Research Center for Industries of the Future, School of Engineering, Westlake University, China
[13]Physics Department, University of Texas at El Paso, USA
[14]NSF National Center for Atmospheric Research, USA
[15]Institute for Chemical Engineering Sciences, Foundation for Research and Technology, Greece
[16]Department of Chemical Engineering, University of Patras, Greece
[17]Department of Chemistry, Aarhus University, Denmark
[18]School of Earth and Environment, University of Leeds, UK
[19]Beijing Weather Modification Center, China
[20]National Institute of Polar Research, Japan
[21]Graduate Institute for Advanced Studies, SOKENDAI, Japan
[22]Institute of Atmospheric Sciences and Climate (ISAC), National Research Council (CNR), Italy
[23]Finnish Meteorological Institute, Finland
[24]West Texas A&M University, USA

Correspondence: Gabriella Wallentin (gabriella.wallentin@kit.edu) and Corinna Hoose (corinna.hoose@kit.edu)

Abstract. Ice nucleating particles (INPs) are rare aerosols essential for cloud ice formation in the mixed-phase temperature range between -38 °C and 0 °C. Due to measurement challenges and limitations in instrument capabilities, the availability of atmospheric observations of INPs remains scarce in time and space. Consequently, no observation-based global distribution of INPs exists so far. This study applies a machine learning (gradient boosting) algorithm to predict the INP concentration over

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

the mixed-phase temperature range across the globe using aerosol mass concentration reanalysis and observed immersion-mode INPs. This proof-of-concept exercise demonstrates that even with limited measurements, the occurrence of INPs can be estimated, following the spatial pattern of key aerosol species. Point-based evaluation metrics, R^2 and log-based RMSE, reach 0.83 and 0.71 for the gradient boosting model, compared to 0.75 and 0.86 for the linear regression benchmark. Regional INP spectra with temperature extracted from the machine learning model agree within one order of magnitude with observations, except over Antarctica (mean bias factor of 150). INPs in continental (oceanic) regions are well-predicted within a mean bias of 2.4 (overestimated within a mean bias of 4) by the machine learning model. The prediction is driven by the strong temperature dependence followed by mid-sized dust particles. This approach can resolve a long-standing source of INP prediction uncertainty in regional weather and climate models.

1 Introduction

Clouds remain one of the most uncertain aspects of climate change (IPCC, 2021; Murray et al., 2021). Their composition alters the interaction with radiation, modulating the energy budget on Earth. In the mixed-phase regime, between -38 °C and 0 °C (Murray et al., 2012), clouds may contain both supercooled liquid droplets as well as cloud ice due to the presence of ice nucleating particles (INPs). These rare aerosols have the capacity to lower the energy barrier for phase change from liquid droplets to ice crystals and are necessary for cloud ice formation through immersion freezing in the mixed-phase regime. Immersion freezing describes the process of a cloud droplet freezing due to an immersed INP within, and is thought to be the main ice nucleation pathway in mixed-phase clouds (Ansmann et al., 2009; Burrows et al., 2022).

Significant research has been done to understand what constitutes a highly active INP (Hoose and Möhler, 2012; Murray et al., 2012; Burrows et al., 2022), but much uncertainty remains. Mineral dust (Kumai, 1961; Niemand et al., 2012; Murray et al., 2012; Boose et al., 2016) and fertile soil dust with an organic component (O'Sullivan et al., 2014; Tobo et al., 2014; DeMott et al., 2025) nucleate ice at temperatures below -20 °C. Biological material may explain the nucleation at warmer temperatures (Schnell and Vali, 1972; Wilson et al., 2015; O'Sullivan et al., 2018). Biological aerosols include particles such as bacteria, fungal spores, non-proteinaceous polymers, and pollen (Després et al., 2012; Kinney et al., 2024). Primary marine organic aerosols (PMOA) are emitted to the atmosphere via sea spray and are an important source of INPs over oceans and in remote regions (DeMott et al., 2016). Biomass-burning aerosol, comprising black carbon, organic aerosol, mineral dust, and plant debris, has been identified as a complex INP source (Burrows et al., 2022). Recent literature indicates that black carbon has a low INP impact in the mixed-phase temperature range in the immersion mode (Vergara-Temprado et al., 2018; Schill et al., 2020), but the importance of organic matter (OM) remains unclear (Chen et al., 2025; Gohil et al., 2025). Finally, ice nucleating capabilities have been found in volcanic ash (Hoyle et al., 2011; Umo et al., 2021), metal oxides (Worringen et al.,

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

2015), as well as microplastics (Tatsii et al., 2025). These species tend to be found at colder nucleation temperatures and are less common in the atmosphere.

Aerosol-aware models have enabled global modelling of INPs. Using ice nucleation active site (INAS) density parameterisations for immersion freezing, these models can represent the primary drivers of nucleation: temperature, number concentration and surface area concentrations of ice nucleation active aerosol species. A great effort has been put into the development of INAS density parameterisations for dust (Niemand et al., 2012; Ullrich et al., 2017; Harrison et al., 2019). Recent modelling studies by Herbert et al. (2025) (H25), and Chatziparaschos et al. (2025) (CH25), have produced global distributions of INPs based on (mainly) the INAS density for dust and PMOA. The representation of dust is parameterised using fertile soil dust for H25 (O'Sullivan et al., 2014) and K-Feldspar (mineral dust) (Harrison et al., 2019) for CH25. While sea salt (the inorganic component of sea spray) is not an efficient INP, McCluskey et al. (2018) developed an INAS density parameterisation for sea spray aerosols (SSA) of unknown PMOA content.

Many global and regional models, such as numerical weather prediction (NWP) and climate models, operate with simplified immersion freezing schemes (Fletcher, 1962; Cotton et al., 1986; Cooper, 1986; Meyers et al., 1992) that are not coupled to aerosols due to computational restrictions. Yet, studies show that if included, explicit aerosol-ice interactions improve weather forecasting (Hermes et al., 2024). The simplified immersion freezing parameterisations depend solely on temperature and are insensitive to changes in aerosol concentrations driven by seasonality or climate change. Motivated by this shortcoming and the possibility of bypassing assumption-based INAS density parameterisations, this study shows how machine learning (ML) methods can be used to predict a global INP distribution. Recent advancements in the field have brought an increase in atmospheric applications of ML methods. The eXtreme Gradient Boosting algorithm (XGBoost) (Chen and Guestrin, 2016) has been used to understand the Southern Ocean radiation bias (Fiddes et al., 2024), improving the prediction of low-cloud fraction (Zhang et al., 2024), as well as for predictor selection for tropical rainfall forecast (Satheesh et al., 2025). XGBoost is a supervised learning algorithm that builds predictions by combining many simple decision-tree models. Each decision tree makes a series of basic yes-or-no splits based on the input features (variables used to train the ML model), grouping similar data points together to produce a prediction. In XGBoost, these trees are constructed sequentially using a process called boosting: each new tree is trained to focus on the mistakes made by the previous trees. By repeatedly correcting earlier errors, the full collection of trees produces increasingly accurate predictions.

In this study, we train an XGBoost model to predict a global distribution of INP concentration. The input data include field measurements of INP concentration at surface sites from 5 continents and from ship-based campaigns spanning 49 datasets amounting to 416,469 data points (Table A1, Fig. A1) in immersion freezing mode at activation temperatures between -36 °C and -3 °C. The features include: activation temperature of the immersion freezing INPs, and collocated Copernicus Atmosphere Monitoring Service (CAMS) aerosol mass concentration reanalysis. This includes aerosol mass concentrations from three modes of dust and sea salt, and hydrophilic OM from natural and anthropogenic sources. These species contribute or act as a proxy to the INP concentration in the mixed-phase temperature range. The XGBoost model is evaluated by assessing its: 1) predictive skill using unseen data, 2) ability to capture the temperature dependence in regional climatological INP predictions, and 3) ability to produce realistic global INP distributions.

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

2 Data and Methods

2.1 INP Datasets

The INP datasets used for this proof-of-concept study are outlined in Table A1 and shown spatially in Fig. A1. Some limitations were imposed on the selection of datasets. Due to the temporal collocation with CAMS data, measurements must fall within the years 2003-2025. The selected INP datasets were not chosen for a complete inclusion of all INP measurements. Rather, a set of measurements was selected for a broad distribution across the globe. Aircraft samples were excluded to homogenise the sampling altitude at the surface. However, data collected from multiple measurement techniques, i.e. online and filter measurements, were used to enable a broader global distribution as well as to cover the entire temperature range for immersion freezing. The focus is on immersion freezing experiments due to the importance of this process for mixed-phase clouds, and the limited availability of deposition nucleation data. The measurement techniques are also indicated in Table A1. For discussion on differences between instruments, we refer the interested reader to Lacher et al. (2024). The data was cleaned for outliers above $10^4\,\mathrm{L}^{-1}$, and below $10^{-6}\,\mathrm{L}^{-1}$. This removes zeros (no detected INPs) in some datasets. Further outliers were removed after visual inspection of the data. The large dataset from Beijing was randomly subsampled to 60,000 data points to reduce domain dominance. Care was taken to ensure the subsampled dataset represents the full range of INP concentrations measured. For the Southern Ocean dataset, the ACE campaign measurements were not corrected for background INP levels, which may have resulted in reported INP concentrations that were higher than those actually present. The split between land and ocean data points is approximately 85% land and 15% ocean (including coastal Arctic sites). The number of data points per location is tabulated in Table A1. The temporal and size resolutions are heavily dependent on the measurement technique. For specifics on each campaign, we refer the reader to the corresponding articles listed in Table A1.

2.2 CAMS Data

The observed INPs were collocated through nearest neighbour interpolation with a daily mean aerosol mass concentration reanalysis product from CAMS (Inness et al., 2019). Specifically, the ECMWF Atmospheric Composition Reanalysis 4 (EAC4) global reanalysis between 2012 and 2025 (Copernicus Atmosphere Monitoring Service (CAMS), 2020a). The reanalysis contains assimilated aerosol optical depth (AOD) and is continuously evaluated against observations (Langerock et al., 2024). A daily mean (6-hourly data) was calculated for the closest model level to the altitude at which measurements were performed. For measurement altitudes below 50 m, the surface model level was used to simplify the collocation. For the observations performed at higher elevation, the model level used for collocation was approximated from the calculated model level height during 01-01-2025 (an arbitrary date), as the model levels are not constant in time. The use of another date (01-08-2025) shifts the model level by one, a maximum of 300 m.

EAC4 has a horizontal grid spacing of 0.75 °, which limits our spatial resolution. Included are surface aerosol mass mixing ratios of three modes of dust (0.03 µm -0.55 µm (small dust), 0.55 µm - 0.9 µm (mid dust), and 0.9 µm - 20 µm (large dust)), three modes of sea salt (0.03 µm -0.5 µm (small sea salt), 0.5 µm - 5 µm (mid sea salt), 5 µm - 20 µm (large sea salt)), and hydrophilic OM generated from wildfires, anthropogenic sources such as biomass burning, fossil fuels, and secondary organic

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

aerosols (Dentener et al., 2006; Inness et al., 2019) aged by sulfate. SHAP values (see below) were used to examine the relative importance of each species. Across 4 Arctic stations, sporadic dates when the daily mean of all aerosol fields is zero were removed (1600 data points, primarily from Villum, Greenland). For evaluation against modelled global INP distributions from H25 and CH25 (Herbert et al., 2025; Chatziparaschos et al., 2025), we calculate a four-year climatology between 2020-2024 based on day-count weighted monthly mean aerosol fields (EAC4) (Copernicus Atmosphere Monitoring Service (CAMS), 2020b) at 600 hPa. For the regional comparison, we calculate a four-year climatology between 2007 and 2011 (Copernicus Atmosphere Monitoring Service (CAMS), 2020b) at the surface (model level 60) to ensure no temporal overlap with the training data.

2.3 Machine Learning Methods

The prediction of INPs is a regression task and is performed by XGBoost (Friedman et al., 2000; Chen and Guestrin, 2016), package version 3.2.0 (Chen et al., 2016). Furthermore, linear regression, using the package from scikit-learn version 1.8.0 (Pedregosa et al., 2011), is used as a benchmark for the XGBoost performance.

2.3.1 Features and Target

The features for the INP prediction were chosen based on the results of DeMott et al. (2010) who found a correlation between INPs and the number concentration of aerosols larger than 0.5 μm in diameter. From a mechanistic standpoint, INP concentration scales with surface area. This parameter is, however, not readily available. We hypothesise a correlation, but not necessarily linear, between the mass in each size mode of aerosols and the presence of INPs following previous heterogeneous immersion freezing parameterisations (Phillips et al., 2013). Thus, the features used for the prediction include aerosol mass concentrations for the established dust INP. Furthermore, sea salt can be used as a tracer for PMOA as it is co-emitted in sea spray (DeMott et al., 2016). To what level OM is acting as an INP in the mixed-phase temperature range is still debated (Chen et al., 2025). However, the inclusion of this species allows for particles from biomass burning to act as INP. The mass concentrations were log-transformed with a base of 10. Further features include INP activation temperature. The activation temperature is the temperature at which the INP concentration was measured. This introduces a temperature dependence for the INP prediction to capture the INP spectra (Fig. 2). The observed INP data (Table A1, Fig. 2a,b) used as our target variable were log-transformed with a base of 10.

Feature importance was explored using the SHapley Additive exPlanations (SHAP) values (Lundberg and Lee, 2017) and the TreeExplainer (Lundberg et al., 2020). SHAP values are a type of feature importance that is calculated using game theory and enables machine learning models to be explainable. Temperature dominates the feature importance. This is an expected outcome as the INP spectra are highly dependent on temperature (Fig. 2). The SHAP values can be understood through the example of the impacts of temperature; low temperatures tend to increase the INP concentration, thus in Fig. B3, low values of temperature have positive SHAP values. This indicates an increase in the prediction of INP concentration compared to the mean. For warmer temperatures (red colours), the SHAP values are negative, which indicates a reduction in INP concentration compared to the mean. Mid-size dust particles have a high importance, while the largest dust particles have the third lowest

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

importance. This may be due to both the lack of large dust particles in the training data and the correlation to OM (not shown), which has high importance. Whether this is due to OM acting as an INP or rather that it is used as a tracer for other aerosol species is impossible to disentangle, and we refrain from stating any physical importance of OM apart from its use in predicting the INP concentration.

2.3.2 Model Setup and Evaluation

The training, test, and validation data were split with a 70%-20%-10% distribution. The data were split according to dates to reduce the correlation between the training, test, and validation datasets. No identical data points exist in all three categories, but we note that data points from the same location were present in training, validation, and test data. This method was chosen as the INP distributions differ substantially between each dataset, and the spatial coverage is too sparse to exclude locations for testing. The final product is a prediction of the spatial distribution, not the individual stations, where the presence of similar samples in training and test data may inflate the scoring. The prediction was evaluated on a previously unseen test data set from different dates (Fig. 1) using root mean square error (RMSE) and the coefficient of determination (R^2) score. We define these as

$$\mathrm{R}^2 = 1 - \frac{\mathrm{res}}{\mathrm{tot}}, \quad \text{where} \quad \mathrm{res} = \sum_i (y_i - f_i)^2, \quad \text{and} \quad \mathrm{tot} = \sum_i (y_i - \bar{y})^2 \tag{1}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_i^n (y_i - f_i)^2} \tag{2}$$

where y_i is the observed value, f_i the predicted, $\bar{y}$ the observed mean, and n the sample size. When calculating the RMSE in log-space, it becomes non-dimensional, and the RMSE is represented by a factor calculated as 10^{RMSE}. For the RMSE for XGBoost, an RMSE of 0.71 corresponds to approximately a factor of 5 error, while the linear regression RMSE of 0.86 corresponds to a factor of 7.

The model was trained by minimising the mean square error, with RMSE used for evaluation. The hyperparameter optimisation package Optuna (Akiba et al., 2019) was used to optimise the number and depth of trees, and learning rate using a 5-fold cross-validation on the training dataset. The depth of the trees and the learning rate were identified as the main hyperparameters impacting the results. A range of 4-7 for the maximum depth of the trees, and 0.01-0.3 for the learning rate, was used to produce a range of XGBoost models. Using a 10-fold cross-validation on the training data, 20 of these models obtained RMSE values within 5% of the minimum. These models were chosen to quantify the uncertainty in the spatial variability of the predicted INPs as an XGBoost ensemble. This is shown in Fig. 4. The final XGBoost model (used in Fig. 1, 2, 3 and as a scatter in Fig. 4) was decided for a maximum depth of 5 and a learning rate of 0.05, as this model produced the lowest RMSE (10-fold cross-validation on training data). Early stopping was further used to limit the trees (final number of trees (n estimators) is 175). XGBoost can predict an INP concentration at any temperature. Outside of the observed INP temperature range,

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

extrapolation occurs and becomes unreliable. Evaluation of XGBoost is thus only performed within the temperature range of the observations.

There is a clear domain dominance (85% data points over land) in the dataset. Ways to mitigate this were explored by 1) selecting an equal number of data points from each region, and 2) boosting the Oceanic and Antarctic datasets by weights. Both these methods subsequently failed to capture a realistic spatial distribution (not shown).

3 Results

3.1 Predicting INPs with Machine Learning

INP concentrations are predicted using an XGBoost regression with input features and parameters as described in Sect. 2.3. As a benchmark for the performance of XGBoost, a linear regression model is used with the same features. Figure 1 shows the observed INP concentration, Test $\mathrm{INP_{OBS}}$, (from dates of samples not included in training), vs predicted INP concentration for the linear regression model (Test $\mathrm{INP_{LIN}}$) and XGBoost (Test $\mathrm{INP_{XGB}}$). To evaluate the ML models, the point-based metrics, coefficient of determination, R^2, and root mean square error, RMSE, in log space (see Sect. 2.3) are used. Metrics show better performance for XGBoost, with a log-based RMSE of 0.71 compared to 0.86 for the linear regression model. R^2 values reach 0.83 (0.75) for XGBoost (linear regression), where a perfect fit would render 1.0. Most of the overprediction by both ML models is due to the very low observed INP concentrations over Antarctica (Wex et al., 2025). This is discussed further below.

3.2 Regional INPs as a Function of Temperature

Regionally averaged datasets of observed INPs, $\mathrm{INP_{OBS}}$, summarised in Table A1, are shown in Fig. 2a,b as INP spectra (INP concentration as a function of temperature). The corresponding predicted regional climatological surface spectra for XGBoost ($\mathrm{INP_{XGB}}$ (SFC, T)) are shown in Fig. 2c,d, where the spread in the 5th to 95th percentile indicates the variability across the latitudes and longitudes selected for each region. The prediction over continental regions follows the Cooper parameterisation (Cooper, 1986) at low temperatures (below -20 °C), and the Fletcher parameterisation (Fletcher, 1962) at higher temperatures (above -20 °C), similar to the observations. The INP prediction is driven by this temperature dependence (Sect. 2.3, Fig. B3) together with mid-sized dust particles. Discrepancies between the observations at low temperatures (Fig. 2 and Fig. B1) may be due to instrument limitations, such as detection limits of the different instruments, as well as differences in the presence of low-temperature activating INPs, such as dust. The normalised occurrence of $\mathrm{INP_{OBS}}$ for each temperature bin is shown in the histograms above Fig. 2a,b and indicates a comparatively lower frequency of INPs measured at temperatures below -25 °C for the oceanic regions. A bias ratio is defined as the mean XGBoost values divided by the mean observational values at the same temperature. For Europe, Asia, North and South America, the bias ratio varies between 0.1 and 10 (Fig. 2e), indicating a good fit to the observations by XGBoost. The mean absolute bias (over temperature bins) reaches 2.4 for Europe, while East Asia has the overall smallest bias (0.98). The Arctic and Southern Ocean have a mean bias of about 4, while oceanic regions are better predicted at a mean bias of 1.5, with a tendency of over-prediction by XGBoost. For Antarctica, a bias of up to

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

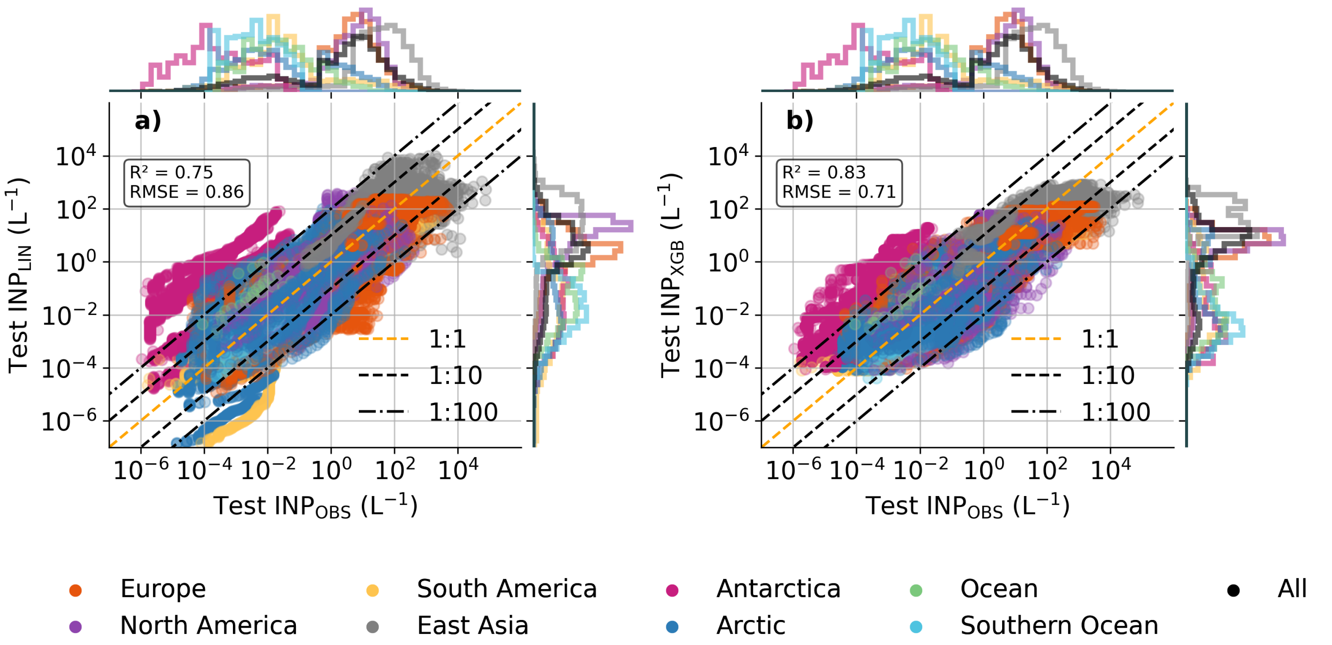

Figure 1. Observed vs predicted INP concentration for the linear regression model (a) and the XGBoost regression (b). The orange line shows the 1:1 relationship. The 10x and 100x biases are outlined in black. The box shows the log-based RMSE and R^2 values. Colours signify the region of the test data. Normalised distributions for each region are shown in the histograms, where the combined dataset is shown in black.

3 orders of magnitude is obtained (mean bias of 152) (Fig. 2f). This large discrepancy may be due to less well-constrained aerosol concentrations over this region and the small number of data points, focused on West Antarctica, limiting its presence in the training dataset. The use of annual mean aerosol mass concentrations for INP prediction does not support the presence of a strong seasonal cycle in dust and OM over Antarctica (not shown) and thus leads to overprediction. In general, XGBoost tends to have a smaller spread between the regions, indicating that the prediction of both high and low values is limited. This is likely due to the usage of an annual mean aerosol mass concentration, which smooths out variability, as well as an inherent characteristic of the XGBoost setup, as it is developed to reduce the mean error and not to capture extremes.

3.3 A Global Spatial Distribution of INPs across Models

To further test XGBoost, a spatial distribution of INPs is predicted at approximately 600 hPa and at -20 °C, INP_{XGB} (600 hPa -20 °C), shown in Fig. 3a, together with the modelled INPs from CH25 (Chatziparaschos et al., 2025) (Fig. 3b) and H25 (Herbert et al., 2025) (Fig. 3c). Across the two aerosol-aware global models, CH25 and H25, large differences can be seen, owing to differences in their aerosol and INP assumptions, but the main features of the INP distribution are consistent. The peaks in CH25 and H25 coincide with the large dust sources: the Sahara and the Gobi desert, and across the south-western part of Asia, also seen in XGBoost. However, the INP plume extending from the Sahara over the Atlantic is weaker in XGBoost

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

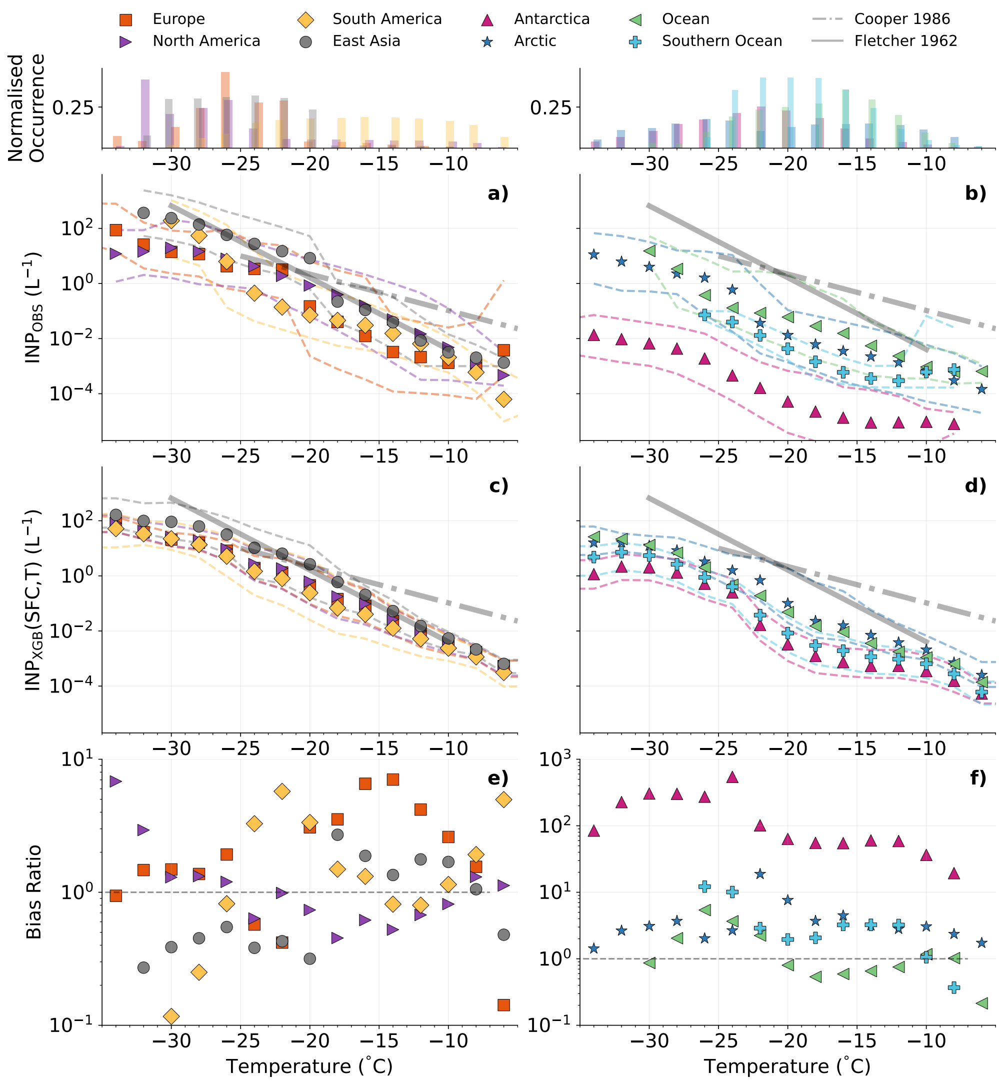

Figure 2. Mean (geometric) continental (a) and remote/oceanic (b) $\mathrm{INP_{OBS}}$ spectra with temperature for each region, the regions are separated for better readability. The temperature is binned in bins of 2 K for temperature ranges containing more than 10 data points. A histogram showing the normalised occurrence frequency of $\mathrm{INP_{OBS}}$ for each temperature bin is included on top of the INP spectra. c) and d) Predicted $\mathrm{INP_{XGB}}$ (SFC, T) spectra with temperature for each region using a 2007-2011 surface aerosol climatology based on CAMS (Copernicus Atmosphere Monitoring Service (CAMS), 2020a). Common parameterisations are shown in a) -d) from Cooper (1986) (Cooper, 1986) and Fletcher et al. (1962) (Fletcher, 1962) for their respective temperature range of validity. e) and f) Bias ratio, defined as the mean XGBoost values divided by the mean observed values for each temperature. Below (above) 1, XGBoost is under- (over-)predicting. The region separation is shown in Fig. A1. These region definitions do not further impact XGBoost and are shown for a general overview of the data. Dashed lines outline the predicted and observed 5th and 95th quantiles.

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

compared to CH25 and H25. In H25, this peak is unreasonably large (Herbert et al., 2025), attributed to the increase in INP concentration with the fertile soil dust parameterisation (O'Sullivan et al., 2014). The prediction over Antarctica has very low concentrations, comparable to CH25. XGBoost produces a larger peak over Central North America in comparison to the other models. INP prediction over Australia is also high, as seen in H25. The oceans, including the Arctic region, have a larger variability across the two aerosol-aware global models, CH25 and H25, with a lack of consensus in any region apart from the most remote parts of the Pacific. XGBoost predicts between H25 and CH25 for most of the regions with observational data. The regional geometric mean for the regions explored in Fig. 2 is shown in Fig. 4. The INP concentration over the oceans in XGBoost is greater than H25 and CH25 by a factor of 2.5 (Fig. 4). To explore the XGBoost model uncertainty, in terms of how XGBoost is tuned with hyperparameters (parameters that tune the model, see Sect. 2.3.2), 20 XGBoost models with similar RMSE are chosen to quantify the variance across the spatial distribution. The violin plots in Fig. 4 show the spread of the predicted INP concentration and can be thought of as an uncertainty in the prediction. Overall, all regions have less than an order of magnitude spread. Antarctica has the comparatively largest spread, reflected in its contribution to the dataset with only approximately 13700 data points, corresponding to 0.03% of the full dataset. Overall, XGBoost predicts a realistic global INP distribution. The spatial distributions at -10° C and -30° C are shown in Fig. B2.

4 Discussion

Using the gradient boosting decision tree algorithm XGBoost, a global INP distribution is obtained. CAMS surface aerosol mass concentrations from three modes of dust and sea salt, and natural and anthropogenic hydrophilic organic matter, are used together with the activation temperature of observed INPs to produce a five-dimensional (space, time, and temperature) INP product.

Spatially, a realistic distribution is obtained, with peaks over the deserts and lower values over the oceans and remote regions (Fig. 3). A comparison with two aerosol-aware global models, using INAS densities for INP prediction, shows a decent agreement over the oceans and continents, within the uncertainty between the aerosol-aware global models (CH25 and H25) themselves. INP concentration varies strongly with temperature, and XGBoost captures this inherent quality (Fig. 2). As global INPs cannot be accurately simplified to a single parameterisation, XGBoost can better represent regional nuances in aerosol availability (Fig. 2, 3). Regional INP spectra can be extracted from XGBoost and can serve as a better alternative to global parameterisations, reducing biases introduced by aerosol-cloud interactions. To put our result into perspective, observational intercomparisons at the same location often vary within an order of magnitude (Lacher et al., 2024; DeMott et al., 2025); thus, our predictions are within observational measurement uncertainty.

As with any machine learning approach, the results presented here remain influenced by the characteristics of the underlying training data. In particular, the predicted spatial distributions depend on the quality and coverage of the aerosol reanalysis used as input. The framework is also constrained by the availability and representativeness of existing INP observations. Although the current dataset captures a broad range of environments, some regions remain comparatively undersampled (Fig. A1).

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

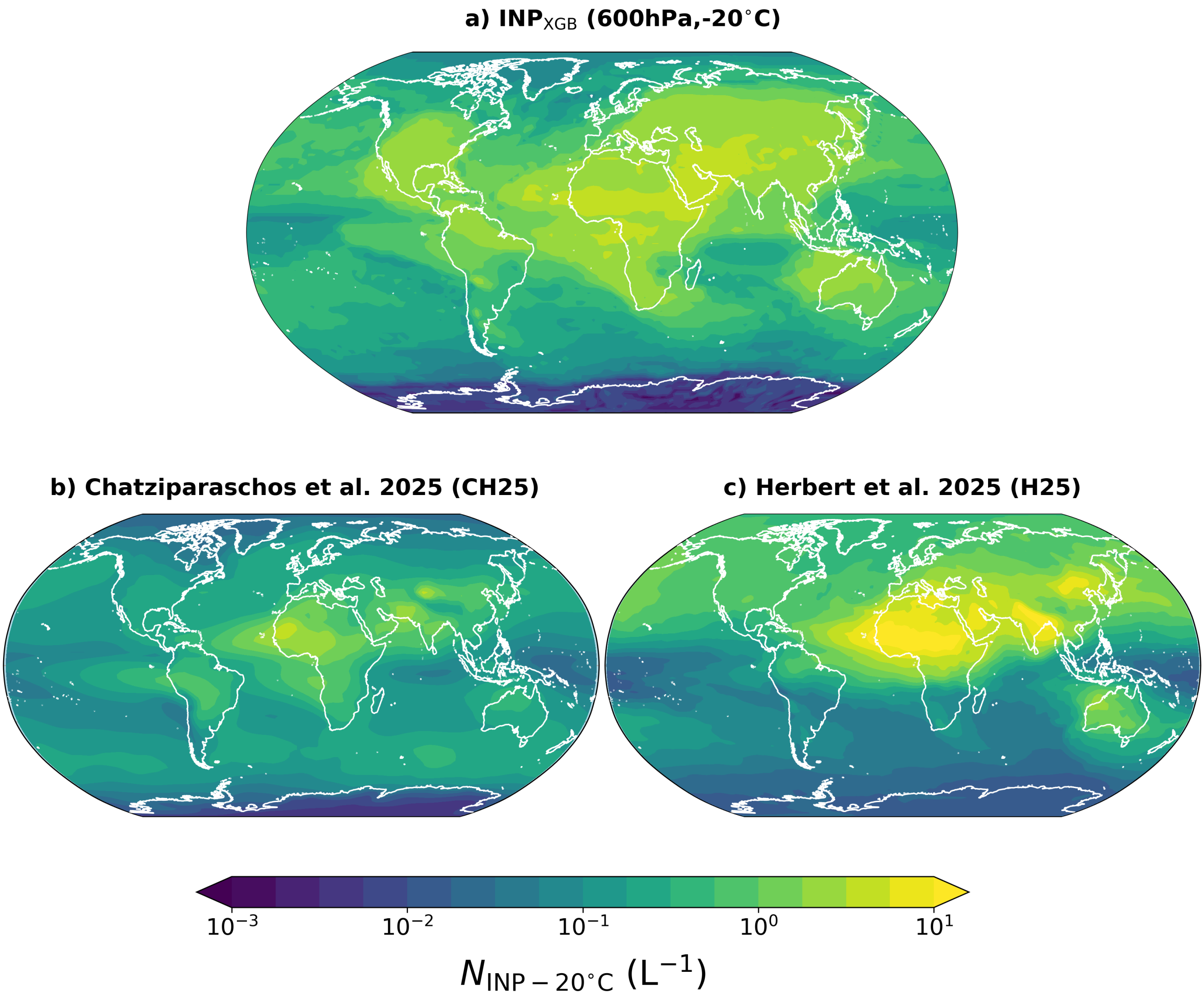

Figure 3. Global distributions of predicted INP concentration, N_{INP}, at approximately 600 hPa and at -20 °C. In a) INP$_{\mathrm{XGB}}$ (600 hPa - 20 °C), a global INP distribution is predicted at an activation temperature of -20 °C using CAMS aerosol mass concentrations at 600 hPa (Copernicus Atmosphere Monitoring Service (CAMS), 2020b). It is linearly interpolated to the same grid as CH25 (b). The horizontal grid spacing is 2 ° x 3 ° for CH25 and 1.25 ° x 1.875 ° for H25 (c). H25 is an annual mean from a 20-year simulation with aerosol emissions from 2010, while CH25 represents the annual mean with simulations spanning 2009-2016.

Additional observations from Africa, Asia, Australia, Antarctica, and especially remote oceanic regions would help further constrain, evaluate, and improve the model.

Finally, this study serves as a proof-of-concept, producing a global INP distribution with fewer INP assumptions than aerosol-aware global models while retaining a similar accuracy. INP prediction can thus be performed at ambient tempera-

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

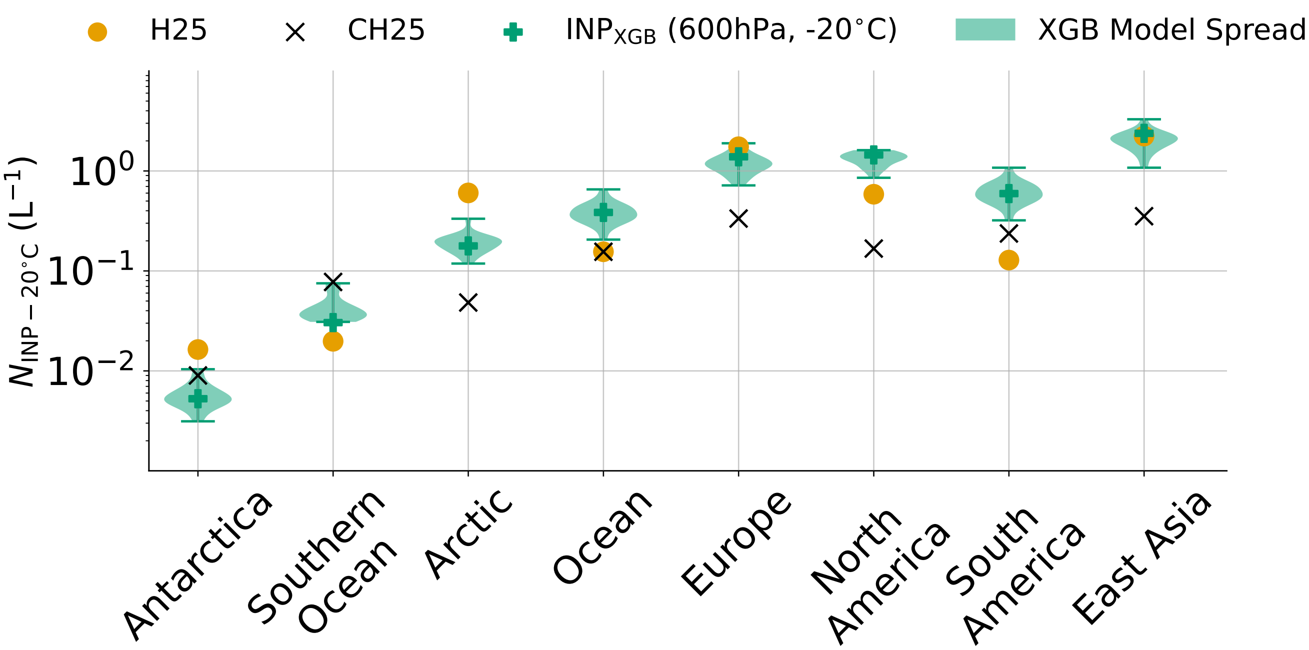

Figure 4. Geometric mean INP concentration for H25, CH25, and XGBoost for data shown in Fig. 3 averaged over specific regions defined in Fig. A1. Violin plots show the spread of the predicted INP concentration using 20 XGBoost models with varying hyperparameters.

tures across the globe and allows for a more variable INP concentration, as it is not constrained by a species-specific INAS density. The XGBoost model can serve as a cheap and easy-to-implement immersion freezing parameterisation in regional and global weather and climate models, potentially reducing long-standing biases in cloud representation.

Data availability. RADAR4KIT repositories (Wallentin, 2026a, b) will contain the pre-processing scripts for the aerosol collocation, Python scripts for the machine learning models, and the final machine learning models upon manuscript acceptance. XGBoost INP predictions for the surface and 600 hPa at the three activation temperatures are included. The INP data are available under the respective references in Table A1. The CAMS data used for daily collocation have been obtained from Copernicus Atmosphere Monitoring Service (CAMS) (2020a), while the monthly means used for prediction are available from (Copernicus Atmosphere Monitoring Service (CAMS), 2020b). Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains. The CH25 data used for model comparison in Fig. 3b can be located in the open research section of Chatziparaschos et al. (2025). The data from H25 Herbert et al. (2025) has been extracted from their model simulation to match the same pressure levels as presented here.

Appendix A: INP Datasets

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

Table A1: Ice nucleating particle datasets used in this study. Measurement type is indicated in the third column, where PINE refers to the online measurements and INSEKT, CRAFT, SIOAIS, IS/INS, INDA, and LINA are offline measurements using filters and droplet freezing assays. When only the word "Filter" is given, the measurement technique is unknown. References are given for the dataset itself and the related articles when both exist. The TAN dataset (Welti et al., 2020) has been updated following personal correspondence with P. DeMott. The data points column refers to the number of individual data points considered in the study from each campaign, and the dates column shows the time range of the measurements.

Name	Region	Measurement Type	Data points	Dates	Reference
Karlsruhe	Europe (Germany)	PINE	14304	2021-03-25, 2021-05-02	Vogel (2025c); Lacher (2025a)
Sonnblick Observatory	Europe (Germany)	PINE	69809	2021-07-28, 2022-10-14	Bogert (2025a, b)
Puy de Dôme	Europe (France)	PINE	3201	2018-09-10, 2018-10-25	Lacher et al. (2024)
Mount Helmos	Europe (Greece)	PINE & INSEKT	23377	2021-06-13, 2022-05-08	Vogel (2025a, b) and unpublished KIT
Chopin	Europe (Greece)	PINE	4773	2024-10-10, 2025-03-25	Unpublished EPFL
Aigio	Europe (Greece)	PINE	2714	2023-03-27, 2023-04-24	Vogel (2025d)
Hyytiälä	Europe (Finland)	PINE & INSEKT	16290	2018-03-13, 2024-04-29	Brasseur et al. (2022) and unpublished KIT
Cyprus	Europe	INSEKT & LINA	2887	2017-04-02, 2017-04-28	Gong et al. (2019); Gong et al. (2019)
Mace Head	Europe (Ireland)	IS/INS	1941	2015-08-05, 2015-08-28	McCluskey et al. (2018)
Mace Head	Europe (Ireland)	INSEKT	6275	2017-11-21, 2019-05-24	Unpublished KIT
WAO	Europe (UK)	PINE	9619	2023-01-12, 2023-12-31	Herbert et al. (2026)
CountIce1	Europe (UK)	PINE	15960	2021-01-09, 2022-12-01	Tarn and Murray (2025a)
CountIce2	Europe (UK)	PINE	18794	2023-01-04, 2023-12-08	Tarn and Murray (2025b)
Beijing	Asia	PINE	124967	2022-01-04, 2023-12-07	Herbert et al. (2026)
Beijing	Asia	INDA & LINA	960	2016-11-28, 2016-12-22	Chen et al. (2018)
Tokyo	Asia	CRAFT	4155	2016-08-01, 2017-07-30	Tobo et al. (2020); Tobo (2020)
Punta Arenas	South America	LINA & INDA	6037	2019-05-12, 2020-02-16	Gong et al. (2022); Gong et al. (2022)
CACTI	South America	IS/INS	3633	2018-10-05, 2019-04-28	DeMott (2021); Testa et al. (2021)
Southern Great Plains	North America	PINE	9226	2019-10-01, 2019-11-14	Vepuri et al. (2021)
Southern Great Plains	North America	IS/INS	1580	2020-10-19, 2024-05-09	Creamean et al. (a)
Colorado	North America	PINE	41282	2021-01-11, 2022-12-07	Lacher (2025b)
Gunnison	North America	IS/INS	3915	2021-09-09, 2023-06-12	Creamean et al. (b)
La Jolla	North America	IS/INS	2288	2023-02-16, 2024-02-13	Creamean et al. (d)
Houston	North America	IS/INS	2819	2022-06-04, 2022-09-29	Creamean et al. (c)
Oliktok	Arctic	IS/INS	1538	2020-08-20, 2021-06-02	Creamean et al. (e)
Utqiavik	Arctic	INDA	4734	2012-06-29, 2013-05-31	Wex et al. (2019); Wex et al. (2019)
Ny Ålesund	Arctic (Svalbard)	INDA	2271	2012-03-23, 2012-09-04	Wex et al. (2019); Wex et al. (2019)
Alert	Arctic (Canada)	INDA	5357	2015-04-22, 2016-04-25	Wex et al. (2019); Wex et al. (2019)
Villum	Arctic (Greenland)	LINA & INDA	25408	2018-07-13, 2020-09-11	Sze et al. (2023); Sze et al. (2023)

Continued on next page...

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

Table A1 – continued from previous page

Name	Region	Measurement Type	Data points	Dates	Reference
Zeppelin	Arctic (Svalbard)	PINE	12615	2023-04-26, 2023-10-23	Böhmländer (2025b)
Gruvebadet	Arctic (Svalbard)	PINE	4489	2023-03-15, 2023-04-10	Böhmländer (2025a)
MOSAiC	Arctic	IS/INS	4354	2019-10-27, 2020-09-19	Creamean et al. (2022)
Pallas	Arctic (Finland)	PINE	9262	2022-09-23, 2022-12-23	Böhmländer et al. (2025)
East North Atlantic	Oceanic (North Atlantic)	PINE	1473	2020-10-02, 2020-12-01	Wilbourn and Hiranuma (2023); Wilbourn et al. (2024)
ACAPEX	Oceanic (East Pacific)	IS/INS	407	2015-01-17, 2015-02-10	DeMott and Hill (2016); Welti et al. (2020)
NM	Antarctica	LINA & INDA	10383	2019-06-12, 2021-11-06	Wex et al. (2025); Eckermann et al. (2025b)
PES	Antarctica	LINA & INDA	3144	2020-10-12, 2022-02-01	Wex et al. (2025); Eckermann et al. (2025a)
TAN	Southern Ocean	Filter	236	2015-01-02, 2015-12-02	Welti et al. (2020)
ACE	Southern Ocean	INDA	2717	2016-12-21, 2017-03-18	Tatzelt et al. (2022, 2020)
SHIPPO	Oceanic (West Pacific)	IS	256	2012-07-16, 2014-04-15	Welti et al. (2020); DeMott et al. (2016)
PS95	Oceanic (Atlantic)	Filter	440	2015-11-01, 2015-11-30	Welti et al. (2020)
PS79	Oceanic (Atlantic)	Filter	7800	2012-01-05, 2012-12-05	Welti et al. (2020)
AQABA	Oceanic (Arabian Peninsula)	SIOAIS	6646	2017-07-05, 2017-08-31	Beall et al. (2022)

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

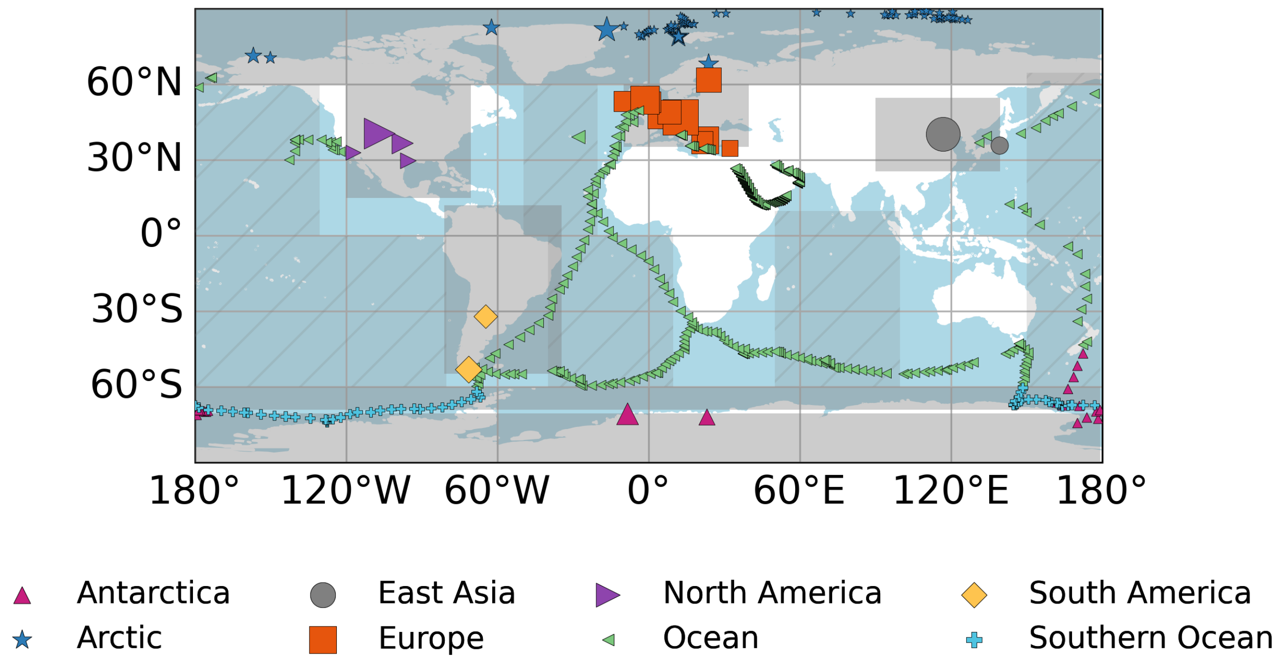

Figure A1. Locations of observational datasets used with region specification as indicated in the legend. The size of the scatter indicates the size of the dataset. Grey patches indicate the region selected for XGBoost comparison in Figs. 2 and 4. The ocean region is roughly defined to cover most of the oceanic regions.

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

Appendix B

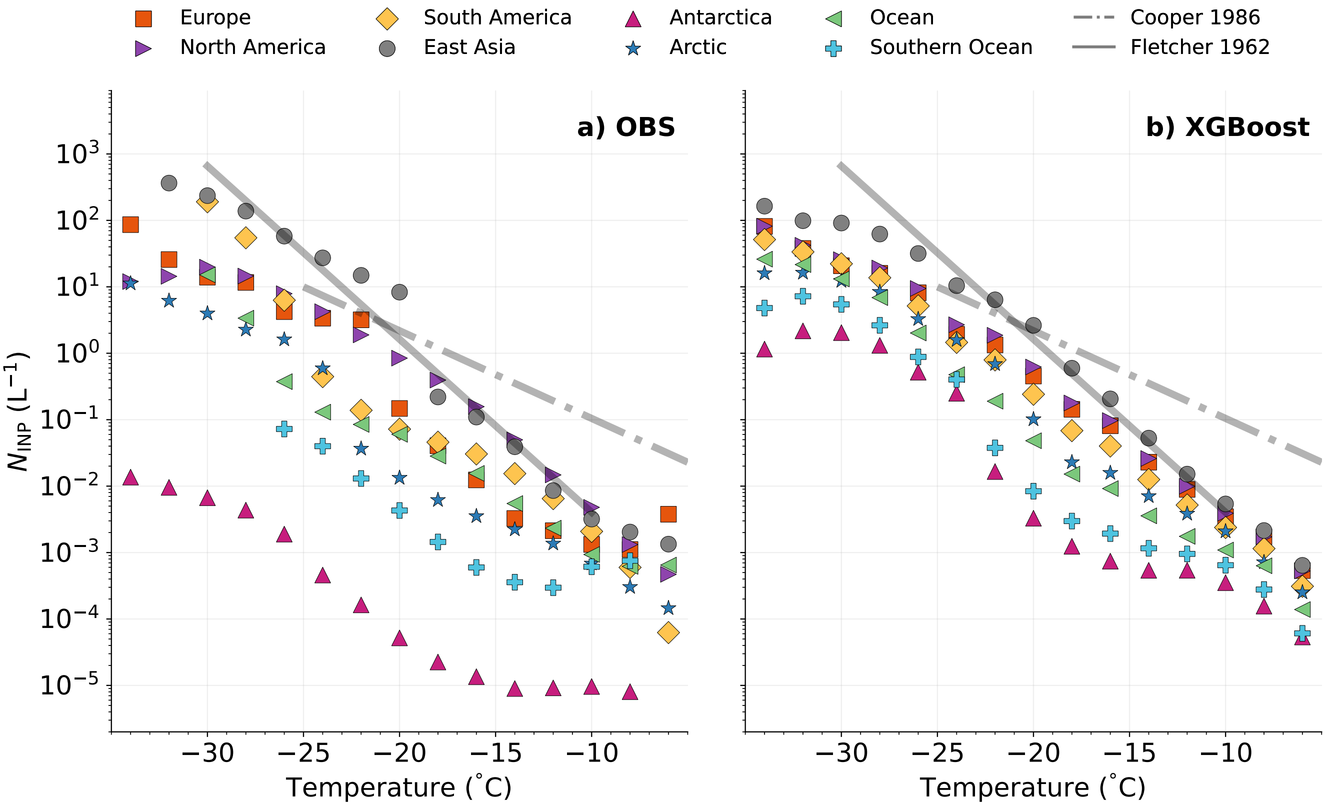

Figure B1. INP number concentration (N_{INP}) as in Fig. 2 but combined for all regions for a) observations, b) XGBoost, $\mathrm{INP_{XGB}}$ (SFC, T), with INP prediction performed at the surface.

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

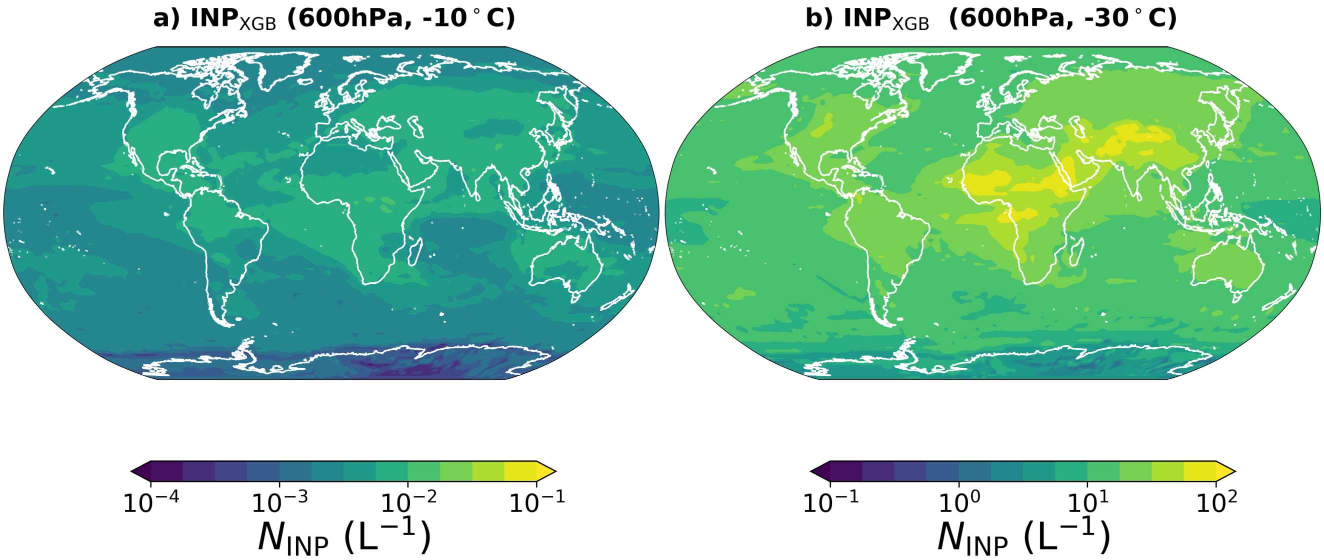

Figure B2. As in Fig. 3 at an activation temperature for INP prediction of a) -10° C and b) -30° C.

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

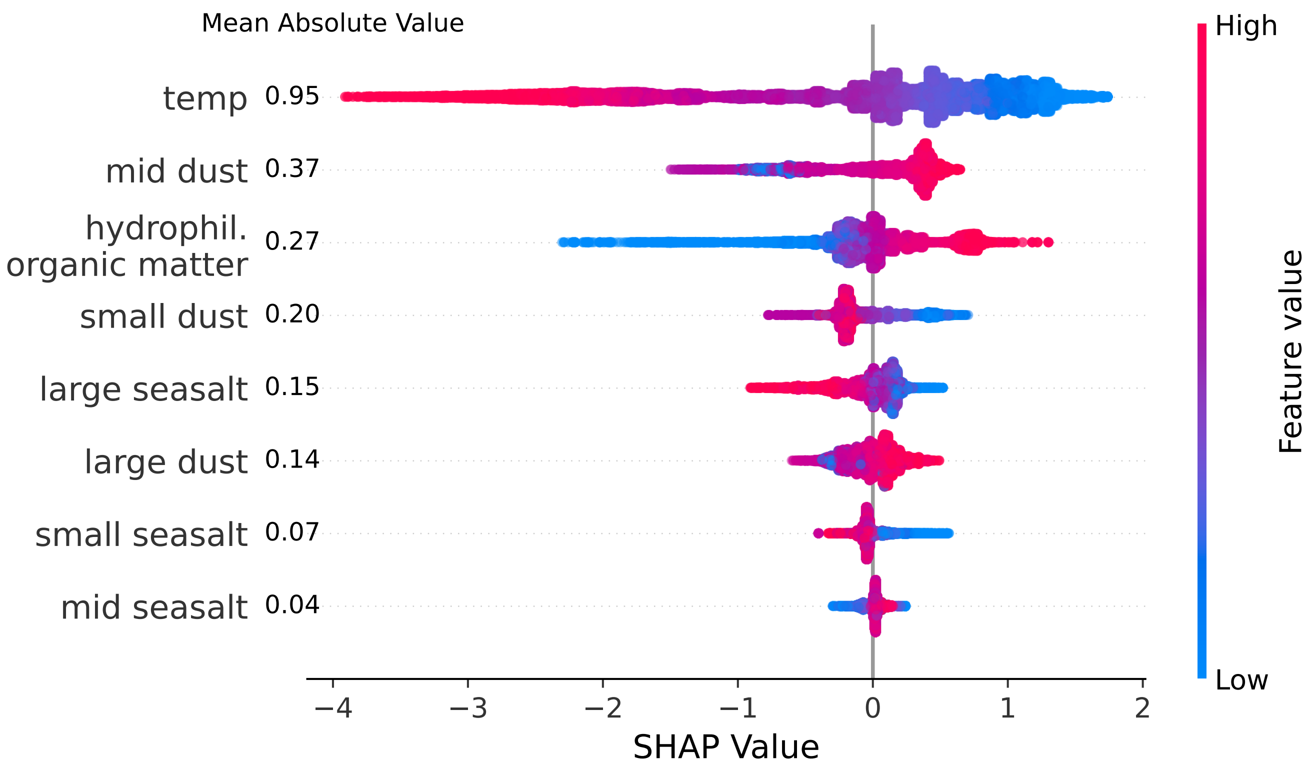

Figure B3. Feature importance using SHAP (Lundberg and Lee, 2017) values. Features are sorted according to their mean absolute SHAP value, as shown on the right of each feature label. A SHAP value of 0 indicates the mean prediction, while negative (positive) values indicate a prediction that is lower (higher) than the mean due to the specific feature. The feature values are indicated by the colour. Temperature shows an expected feature importance where high temperatures, as indicated by the red colours, decrease the INP concentration (negative SHAP value) while low temperatures tend to increase the concentration.

Author contributions. GW designed the study, developed the methods, and analysed the results with input from CH. All co-authors provided datasets used in the study and input on the manuscript.

Competing interests. The authors declare no competing interests.

Acknowledgements. The authors thank Thomas Hill for providing data and feedback, and Baptiste Testa for providing data. The authors would like to express their gratitude to all co-authors on the data publications used in this publication. Furthermore, thanks are extended to Athul Rasheeda Satheesh for valuable discussions and feedback. GW, CH, AN, and CM acknowledge support for the research of this work from the Horizon Europe project CleanCloud (Grant Agreement no. 101137639). JW acknowledges funding from the ICON-SmART project (Grant number 4823DWDP4) funded by the German Federal Ministry of Transport within the Hans Ertel Centre for Weather Research. RH acknowledges funding provided by the Natural Environment Research Council under grant NE/T00648X/1 (M-Phase). XG acknowledges funding provided by the National Natural Science Foundation of China, NSFC (Nos. 42505070, 22576164). YT acknowledges funding provided by the Environment Research and Technology Development Fund (JPMEERF20232001) of the Environmental Restoration and Conservation Agency of Japan.

This work was performed with the help of the Large Scale Data Facility at the Karlsruhe Institute of Technology, funded by the Ministry of Science, Research and the Arts Baden-Württemberg and by the Federal Ministry of Education and Research. The authors gratefully acknowledge the computing time provided on the high-performance computer HoreKa by the National High-Performance Computing Center at KIT (NHR@KIT). This center is jointly supported by the Federal Ministry of Education and Research and the Ministry of Science, Research and the Arts of Baden-Württemberg, as part of the National High-Performance Computing (NHR) joint funding program (https://www.nhr-verein.de/en/our-partners). HoreKa is partly funded by the German Research Foundation (DFG). Grammarly has been used for the stylistic and grammatical improvement of the manuscript, and ChatGPT has been used in the development of the machine learning workflow.

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

References

Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M.: Optuna: A Next-generation Hyperparameter Optimization Framework, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2019.

Ansmann, A., Tesche, M., Seifert, P., Althausen, D., Engelmann, R., Fruntke, J., Wandinger, U., Mattis, I., and Müller, D.: Evolution of the ice phase in tropical altocumulus: SAMUM lidar observations over Cape Verde, Journal of Geophysical Research Atmospheres, 114, https://doi.org/10.1029/2008JD011659, 2009.

Beall, C. M., Hill, T. C., Demott, P. J., Köneman, T., Pikridas, M., Drewnick, F., Harder, H., Pöhlker, C., Lelieveld, J., Weber, B., Iakovides, M., Prokeš, R., Sciare, J., Andreae, M. O., Stokes, M. D., and Prather, K. A.: Ice-nucleating particles near two major dust source regions, Atmospheric Chemistry and Physics, 22, 12 607–12 627, https://doi.org/10.5194/acp-22-12607-2022, 2022.

Bogert, P.: PINE-04-02 SBO21 2021, Karlsruhe Institute of Technology (KIT) RADAR4KIT, https://doi.org/10.35097/g828m8z2det94691, [Dataset], 2025a.

Bogert, P.: PINE-04-02 SBO21_new 2021–2022, Karlsruhe Institute of Technology (KIT) RADAR4KIT, 10.35097/dxcsmnvfunkx76vx, [Dataset], 2025b.

Böhmländer, A., Lacher, L., Fösig, R., Büttner, N., Nadolny, J., Brus, D., Doulgeris, K.-M., and Möhler, O.: Measurement of the ice-nucleating particle concentration with the Portable Ice Nucleation Experiment during the Pallas Cloud Experiment 2022, https://doi.org/10.5194/essd-2025-89, 2025.

Boose, Y., Welti, A., Atkinson, J., Ramelli, F., Danielczok, A., Bingemer, H. G., Plötze, M., Sierau, B., Kanji, Z. A., and Lohmann, U.: Heterogeneous ice nucleation on dust particles sourced from nine deserts worldwide - Part 1: Immersion freezing, Atmospheric Chemistry and Physics, 16, 15 075–15 095, https://doi.org/10.5194/acp-16-15075-2016, 2016.

Brasseur, Z., Castarède, D., Thomson, E. S., Adams, M. P., Drossaart Van Dusseldorp, S., Heikkilä, P., Korhonen, K., Lampilahti, J., Paramonov, M., Schneider, J., Vogel, F., Wu, Y., Abbatt, J. P., Atanasova, N. S., Bamford, D. H., Bertozzi, B., Boyer, M., Brus, D., Daily, M. I., Fösig, R., Gute, E., Harrison, A. D., Hietala, P., Höhler, K., Kanji, Z. A., Keskinen, J., Lacher, L., Lampimäki, M., Levula, J., Manninen, A., Nadolny, J., Peltola, M., Porter, G. C., Poutanen, P., Proske, U., Schorr, T., Silas Umo, N., Stenszky, J., Virtanen, A., Moisseev, D., Kulmala, M., Murray, B. J., Petäjä, T., Möhler, O., and Duplissy, J.: Measurement report: Introduction to the HyICE-2018 campaign for measurements of ice-nucleating particles and instrument inter-comparison in the Hyytiälä boreal forest, Atmospheric Chemistry and Physics, 22, 5117–5145, https://doi.org/10.5194/acp-22-5117-2022, 2022.

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, https://doi.org/10.1029/2021RG000745, 2022.

Böhmländer, A.: PINE-04-02 Exinp_GVB 2023, Karlsruhe Institute of Technology (KIT) RADAR4KIT, https://doi.org/10.35097/dc0vhg052pzx95wx, [Dataset], 2025a.

Böhmländer, A.: PINE-04-02 Exinp_ZEP 2023–2024, Karlsruhe Institute of Technology (KIT) RADAR4KIT, https://doi.org/10.35097/tdwrwzd5z7nakajv, [Dataset], 2025b.

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.

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

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.

Chen, J., Jakob, F. M. O., Voliotis, A., Wu, H., Syafira, S. A., Oghama, O., Shardt, N., Fauré, N., Kong, X., Mcfiggans, G., and Kanji, Z. A.: Ice Nucleation Abilities and Chemical Characteristics of Laboratory-Generated and Aged Biomass Burning Aerosols, Environmental Science and Technology, 59, 2575–2586, https://doi.org/10.1021/acs.est.4c04941, 2025.

Chen, T. and Guestrin, C.: XGBoost: A scalable tree boosting system, in: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, vol. 13-17-August-2016, pp. 785–794, Association for Computing Machinery, ISBN 9781450342322, https://doi.org/10.1145/2939672.2939785, 2016.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., Li, M., Xie, J., Lin, M., Geng, Y., Li, Y., Li, Y., and Guestrin, C.: XGBoost: Extreme Gradient Boosting, https://xgboost.readthedocs.io/, 2016.

Cooper, W. A.: Ice Initiation in Natural Clouds, Meteorological Monographs, 21, 29 – 32, https://doi.org/10.1175/0065-9401-21.43.29, 1986.

Copernicus Atmosphere Monitoring Service (CAMS): CAMS global reanalysis (EAC4), Copernicus Atmosphere Monitoring Service (CAMS) Atmosphere Data Store, https://doi.org/10.24381/d58bbf47, accessed on 10-09-2025, 2020a.

Copernicus Atmosphere Monitoring Service (CAMS): CAMS global reanalysis (EAC4) monthly averaged fields., Copernicus Atmosphere Monitoring Service (CAMS) Atmosphere Data Store, https://doi.org/10.24381/fd75fff2, accessed on 10-09-2025, 2020b.

Cotton, W. R., Tripoli, G. J., and Rauber, R. M.: Numerical Simulation of the Effects of Varying Ice Crystal Nucleation Rates and Aggregation Processes on Orographic Snowfall, Journal of Climate and Applied Meteorology, 25, 1658–1680, 1986.

Creamean, J., Hill, T., Hume, C., Vazquez, M., and Shi, Y.: Ice Nucleation Spectrometer for INP measurement (INP), 2020-10-19 to 2024-05-09, Southern Great Plains (SGP), Central Facility, Lamont, OK (C1), Atmospheric Radiation Measurement (ARM) user facility, https://doi.org/10.5439/1770816, [Dataset][Last Access 2025-06-01], a.

Creamean, J., Hill, T., Hume, C., Vazquez, M., and Shi, Y.: Ice Nucleation Spectrometer for INP measurement (INP), 2021-09-09 to 2023-06-12, ARM Mobile Facility (GUC), Gunnison, CO; Supplemental Facility 2 (S2), Atmospheric Radiation Measurement (ARM) user facility, https://doi.org/10.5439/1770816, [Dataset][Last Access 2025-06-01], b.

Creamean, J., Hill, T., Hume, C., Vazquez, M., and Shi, Y.: Ice Nucleation Spectrometer for INP measurement (INP), 2022-06-04 to 2022-09-29, ARM Mobile Facility (HOU), Houston, TX; AMF1 (main site for TRACER) (M1), Atmospheric Radiation Measurement (ARM) user facility, https://doi.org/10.5439/1770816, [Dataset][Last Access 2025-06-01], c.

Creamean, J., Hill, T., Hume, C., Vazquez, M., and Shi, Y.: Ice Nucleation Spectrometer for INP measurement (INP), 2023-02-16 to 2024-02-13, ARM Mobile Facility (EPC), La Jolla, CA; AMF1 (main site for EPCAPE on Scripps Pier) (M1), Atmospheric Radiation Measurement (ARM) user facility, https://doi.org/10.5439/1770816, [Dataset][Last Access 2025-06-01], d.

Creamean, J., Hill, T., Hume, C., Vazquez, M., and Shi, Y.: Ice Nucleation Spectrometer for INP measurement (INP), 2020-08-20 to 2021-06-02, Oliktok Point, AK, USA; Long-term Mobile Facility (OLI), Oliktok Point, Alaska; AMF3 (M1), Atmospheric Radiation Measurement (ARM) user facility, https://doi.org/10.5439/1770816, [Dataset][Last Access 2025-06-01], e.

Creamean, J. M., Barry, K., Hill, T. C., Hume, C., DeMott, P. J., Shupe, M. D., Dahlke, S., Willmes, S., Schmale, J., Beck, I., Hoppe, C. J., Fong, A., Chamberlain, E., Bowman, J., Scharien, R., and Persson, O.: Annual cycle observations of aerosols capable of ice formation in central Arctic clouds, Nature Communications, 13, https://doi.org/10.1038/s41467-022-31182-x, 2022.

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

DeMott, P.: Processed data associated with the manuscript titled "Ice nucleating particles in the boundary layer of the Sierras de Córdoba, Argentina, during the Cloud, Aerosol, and Complex Terrain Interactions experiment.", Colorado State University. Libraries, http://dx.doi.org/10.25675/10217/229310, [Dataset], 2021.

DeMott, P. and Hill, T.: ACAPEX-Ship-Based Ice Nuclei Collections Field Campaign Report, Tech. rep., 2016.

DeMott, P., 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., Hwang, C. Y., Rhee, T. S., Snider, J. R., McMeeking, G. R., Dhaniyala, S., Lewis, E. R., Wentzell, J. J., Abbatt, J., Lee, C., Sultana, C. M., Ault, A. P., Axson, J. L., Martinez, M. D., Venero, I., Santos-Figueroa, G., Stokes, M. D., Deane, G. B., Mayol-Bracero, O. L., Grassian, V. H., Bertram, T. H., Bertram, A. K., Moffett, B. F., and Franc, G. D.: Sea spray aerosol as a unique source of ice nucleating particles, Proceedings of the National Academy of Sciences of the United States of America, 113, 5797–5803, https://doi.org/10.1073/pnas.1514034112, 2016.

DeMott, P. J., Prenni, A. J., Liu, X., Kreidenweis, S. M., Petters, M. D., Twohy, C. H., Richardson, M. S., Eidhammer, T., Rogers, D. C., and Tolbert, M. A.: 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/-/DCSupplemental, 2010.

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

Dentener, F., Kinne, S., Bond, T., Boucher, O., Cofala, J., Generoso, S., Ginoux, P., Gong, S., Hoelzemann, J. J., Ito, A., Marelli, L., Penner, J. E., Putaud, J.-P., Textor, C., Schulz, M., Van Der Werf, G. R., and Wilson, J.: Atmospheric Chemistry and Physics Emissions of primary aerosol and precursor gases in the years 2000 and 1750 prescribed data-sets for AeroCom, Atmos. Chem. Phys, 6, 4321–4344, www.atmos-chem-phys.net/6/4321/2006/, 2006.

Després, V. R., Alex Huffman, J., Burrows, S. M., Hoose, C., Safatov, A. S., Buryak, G., Fröhlich-Nowoisky, J., Elbert, W., Andreae, M. O., Pöschl, U., and Jaenicke, R.: Primary biological aerosol particles in the atmosphere: A review, Tellus, Series B: Chemical and Physical Meteorology, 64, https://doi.org/10.3402/tellusb.v64i0.15598, 2012.

Eckermann, O., Mangold, A., Van Overmeiren, P., Zeppenfeld, S., van Pinxteren, M., Dall'Osto, M., Henning, S., and Wex, H.: INP data from Antarctica, Princess Elisabeth station and west of and on the northern Antarctic Peninsula, PANGAEA, https://doi.org/10.1594/PANGAEA.974235, [Dataset], 2025a.

Eckermann, O., Wex, H., Jurányi, Z., Weller, R., and Henning, S.: INP data from Antarctica, Neumayer Station III, during two years, PANGAEA, https://doi.org/10.1594/PANGAEA.974187, [Dataset], 2025b.

Fiddes, S. L., Mallet, M. D., Protat, A., Woodhouse, M. T., Alexander, S. P., and Furtado, K.: A machine learning approach for evaluating Southern Ocean cloud radiative biases in a global atmosphere model, Geoscientific Model Development, 17, 2641–2662, https://doi.org/10.5194/gmd-17-2641-2024, 2024.

Fletcher, N.: Physics of Rain Clouds, Cambridge University Press, 1962.

Friedman, J., Hastie, T., and Tibshirani, R.: Additive Logistic Regression: A Statistical View of Boosting, The Annals of Statistics, 28, 337–407, 2000.

Gohil, K., Asch, N., Sullivan, R., DeMott, P. J., and Gordon, H.: Global Simulations Suggest Biomass Burning Aerosol Emissions From Grassland Fires Could Be Important Ice Nucleating Particles, Journal of Geophysical Research: Atmospheres, 130, https://doi.org/10.1029/2025JD043647, 2025.

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

Gong, X., Wex, H., Müller, T., Wiedensohler, A., Höhler, K., Kandler, K., Ma, N., Dietel, B., Schiebel, T., Möhler, O., and Stratmann, F.: Characterization of aerosol properties at Cyprus, focusing on cloud condensation nuclei and ice-nucleating particles, Atmospheric Chemistry and Physics, 19, 10 883–10 900, https://doi.org/10.5194/acp-19-10883-2019, 2019.

Gong, X., Wex, H., Müller, T., Wiedensohler, A., Höhler, K., Kandler, K., Ma, N., Dietel, B., Schiebel, T., Möhler, O., and Stratmann, F.: Ground-based measurements on aerosol particles at Paphos, Cyprus, in March-April 2017, PANGAEA, https://doi.org/10.1594/PANGAEA.904758, [Dataset], 2019.

Gong, X., Radenz, M., Wex, H., Seifert, P., Ataei, F., Henning, S., Baars, H., Barja, B., Ansmann, A., and Stratmann, F.: Significant continental source of ice-nucleating particles at the tip of Chile's southernmost Patagonia region, Atmospheric Chemistry and Physics, 22, 10 505–10 525, https://doi.org/10.5194/acp-22-10505-2022, 2022.

Gong, X., Wex, H., Henning, S., and Stratmann, F.: Ice-nucleating particles at the tip of Chile's southernmost Patagonia region from May 2019 to February 2020, PANGAEA, https://doi.org/10.1594/PANGAEA.944440, [Dataset], 2022.

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.

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., Böhmländer, A., Tarn, M. D., Canzi, A., Pantoya, A., Freney, E., Höhler, K., Bogert, P., Planche, C., Tian, P., Adams, M., Barr, S., Brus, D., Büttner, N., Daily, M., Doulgeris, K., Eleftheridadis, K., Forster, G., Fösig, R., Georgakopoulos, D., Gini, M., Hallar, A. G., Krejci, R., Ludewig, E., Mazzola, M., McCubbin, I. B., Petäjä, T., Robinson, J., Vogel, F., Zieger, P., Arnold, S. R., Carslaw, K. S., Hiranuma, N., Möhler, O., and Murray, B. J.: GloPINE dataset: model-ready measurements of INP concentrations using PINE instruments, Earth System Science Data Discussions, 2026, 1–24, https://doi.org/10.5194/essd-2026-41, 2026.

Hermes, K., Quinting, J., Grams, C. M., Hoose, C., and Hoshyaripour, G. A.: Impact of Saharan dust outbreaks on short-range weather forecast errors in Europe, Quarterly Journal of the Royal Meteorological Society, 150, 1704–1723, https://doi.org/10.1002/qj.4666, 2024.

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.

Hoyle, C. R., Pinti, V., Welti, A., Zobrist, B., Marcolli, C., Luo, B., oskuldsson, , Mattsson, H. B., Stetzer, O., Thorsteinsson, T., Larsen, G., and Peter, T.: Ice nucleation properties of volcanic ash from Eyjafjallajökull, Atmospheric Chemistry and Physics, 11, 9911–9926, https://doi.org/10.5194/acp-11-9911-2011, 2011.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, Atmospheric Chemistry and Physics, 19, 3515–3556, https://doi.org/10.5194/acp-19-3515-2019, 2019.

IPCC: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, vol. In Press, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/9781009157896, 2021.

Kinney, N. L., Hepburn, C. A., Gibson, M. I., Ballesteros, D., and Whale, T. F.: High interspecific variability in ice nucleation activity suggests pollen ice nucleators are incidental, Biogeosciences, 21, 3201–3214, https://doi.org/10.5194/bg-21-3201-2024, 2024.

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

Kumai, M.: Snow Crystals and the Identification of the Nuclei in the Northern United States of America, Journal of Meteorology, 18, 139–150, 1961.

Lacher, L.: PINE-01-A CORONA 2021, Karlsruhe Institute of Technology (KIT) RADAR4KIT, https://doi.org/10.35097/m61r3c6tn0t3qk1y, [Dataset], 2025a.

Lacher, L.: PINE-01-A SPL01 2021–2022, Karlsruhe Institute of Technology (KIT) RADAR4KIT, https://doi.org/10.35097/rmfxfbr6jegua5cv, [Dataset], 2025b.

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

Langerock, B., Arola, A., Benedictow, A., Bennouna, Y., Blake, L., Bouarar, I., Cuevas, E., Errera, Q., Eskes, H. J., Griesfeller, J., Ilic, L., Kapsomenakis, J., Mortier, A., Pison, I., Pitkänen, M. R. A., Richter, A., Schoenhardt, A., Schulz, M., Thouret, V., Tsikerdekis, A., Warneke, T., and Zerefos, C.: Validation report of the CAMS global reanalysis of aerosols and reactive trace gases, years 2003-2023, Documentation, Copernicus Atmosphere Monitoring Service, https://doi.org/10.24380/g8h7-kd21, 2024.

Lundberg, S. M. and Lee, S.-I.: A Unified Approach to Interpreting Model Predictions, Curran Associates, Inc., 2017.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., and Lee, S.-I.: From local explanations to global understanding with explainable AI for trees, Nature Machine Intelligence, 2, 2522–5839, 2020.

McCluskey, C. S., Ovadnevaite, J., Rinaldi, M., Atkinson, J., Belosi, F., Ceburnis, D., Marullo, S., Hill, T. C., Lohmann, U., Kanji, Z. A., O'Dowd, C., Kreidenweis, S. M., and DeMott, P. J.: Marine and Terrestrial Organic Ice-Nucleating Particles in Pristine Marine to Continentally Influenced Northeast Atlantic Air Masses, Journal of Geophysical Research: Atmospheres, 123, 6196–6212, https://doi.org/10.1029/2017JD028033, 2018.

Meyers, M. P., DeMott, P. J., and Cotton, W. R.: New Primary Ice-Nucleation Parameterizations in an Explicit Cloud Model, Journal of Applied Meteorology, 31, 708–721, 1992.

Murray, B. J., O'Sullivan, D., Atkinson, J. D., and Webb, M. E.: 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.

Niemand, M., Möhler, O., Vogel, B., Vogel, H., Hoose, C., Connolly, P., Klein, H., Bingemer, H., Demott, P., Skrotzki, J., and Leisner, T.: A particle-surface-area-based parameterization of immersion freezing on desert dust particles, Journal of the Atmospheric Sciences, 69, 3077–3092, https://doi.org/10.1175/JAS-D-11-0249.1, 2012.

O'Sullivan, D., Murray, B. J., Malkin, T. L., Whale, T. F., Umo, N. S., Atkinson, J. D., Price, H. C., Baustian, K. J., Browse, J., and Webb, M. E.: 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.

O'Sullivan, D., Adams, M. P., Tarn, M. D., Harrison, A. D., Vergara-Temprado, J., Porter, G. C., Holden, M. A., Sanchez-Marroquin, A., Carotenuto, F., Whale, T. F., McQuaid, J. B., Walshaw, R., Hedges, D. H., Burke, I. T., Cui, Z., and Murray, B. J.: Contributions of biogenic material to the atmospheric ice-nucleating particle population in North Western Europe, Scientific Reports, 8, https://doi.org/10.1038/s41598-018-31981-7, 2018.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, J. Mach. Learn. Res., 12, 2825–2830, 2011.

Phillips, V. T., Demott, P. J., Andronache, C., Pratt, K. A., Prather, K. A., Subramanian, R., and Twohy, C.: Improvements to an empirical parameterization of heterogeneous ice nucleation and its comparison with observations, Journal of the Atmospheric Sciences, 70, 378–409, https://doi.org/10.1175/JAS-D-12-080.1, 2013.

Satheesh, A. R., Knippertz, P., and Fink, A. H.: Machine Learning Models for Daily Rainfall Forecasting in Northern Tropical Africa Using Tropical Wave Predictors, Weather and Forecasting, 40, 1895–1916, https://doi.org/10.1175/WAF-D-24-0192.1, 2025.

Schill, G. P., DeMott, P. J., Emerson, E. W., Marie Rauker, A. C., Kodros, J. K., Suski, K. J., J Hill, T. C., T Levin, E. J., Pierce, J. R., Farmer, D. K., Kreidenweis, S. M., contributions, A., designed research, S., and performed research, E.: The contribution of black carbon to global ice nucleating particle concentrations relevant to mixed-phase clouds, Proceedings of the National Academy of Sciences, 117, 22 705–22 711, https://doi.org/10.1073/pnas.2001674117/-/DCSupplemental.y, 2020.

Schnell, R. and Vali, G.: Athmospheric Ice Nuclei from Decomposing Vegetation, Nature, 236, 163–165, 1972.

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.

Sze, K. C. H., Wex, H., Hartmann, M., Skov, H., Massling, A., and Stratmann, F.: Two years of data on Ice Nucleating Particles (INP) from Villum Research Station (VRS) in Northern Greenland, PANGAEA, https://doi.org/10.1594/PANGAEA.953838, [Dataset], 2023.

Tarn, M. and Murray, B.: Data from the Portable Ice Nucleation Experiment (PINE) during the CountIce (part 1) 2021-2022 campaign, https://doi.org/10.5281/zenodo.17451020, [Dataset], 2025a.

Tarn, M. and Murray, B.: Data from the Portable Ice Nucleation Experiment (PINE) during the CountIce (part 2) 2023 campaign, https://doi.org/10.5281/zenodo.17451024, [Dataset], 2025b.

Tatsii, D., Gasparini, B., Evangelou, I., Bucci, S., and Stohl, A.: Do Microplastics Contribute to the Total Number Concentration of Ice Nucleating Particles?, Journal of Geophysical Research: Atmospheres, 130, https://doi.org/10.1029/2024JD042827, 2025.

Tatzelt, C., Henning, S., Tummon, F., Hartmann, M., Baccarini, A., Welti, A., Lehtipalo, K., and Schmale, J.: Ice Nucleating Particle number concentration from low-volume sampling over the Southern Ocean during the austral summer of 2016/2017 on board the Antarctic Circumnavigation Expedition (ACE)., Zenodo, https://doi.org/10.5281/zenodo.4311665, [Dataset], 2020.

Tatzelt, C., Henning, S., Welti, A., Baccarini, A., Hartmann, M., Gysel-Beer, M., Van Pinxteren, M., Modini, R. L., Schmale, J., and Stratmann, F.: Circum-Antarctic abundance and properties of CCN and INPs, Atmospheric Chemistry and Physics, 22, 9721–9745, https://doi.org/10.5194/acp-22-9721-2022, 2022.

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., Kreidenweis, S. M., and DeMott, P. J.: Ice Nucleating Particle Connections to Regional Argentinian Land Surface Emissions and Weather During the Cloud, Aerosol, and Complex Terrain Interactions Experiment, Journal of Geophysical Research: Atmospheres, 126, https://doi.org/10.1029/2021JD035186, 2021.

Tobo, Y.: Monitoring of atmospheric ice nucleating particle (INP) number concentrations on Tokyo Skytree, Arctic Data archive System (ADS), Japan, https://ads.nipr.ac.jp/dataset/A20200727-004, [Dataset] [Last Acces 2025-03-17], 2020.

Tobo, Y., Demott, P. J., Hill, T. C., Prenni, A. J., Swoboda-Colberg, N. G., Franc, G. D., and Kreidenweis, S. M.: Organic matter matters for ice nuclei of agricultural soil origin, Atmospheric Chemistry and Physics, 14, 8521–8531, https://doi.org/10.5194/acp-14-8521-2014, 2014.

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, https://doi.org/10.1029/2020JD033658, 2020.

Ullrich, R., Hoose, C., Möhler, O., Niemand, M., Wagner, R., Höhler, K., Hiranuma, N., Saathoff, H., and Leisner, T.: A new ice nucleation active site parameterization for desert dust and soot, Journal of the Atmospheric Sciences, 74, 699–717, https://doi.org/10.1175/JAS-D-16-0074.1, 2017.

Umo, N. S., Ullrich, R., Maters, E. C., Steinke, I., Benker, N., Höhler, K., Wagner, R., Weidler, P. G., Hoshyaripour, G. A., Kiselev, A., Kueppers, U., Kandler, K., Dingwell, D. B., Leisner, T., and Möhler, O.: The Influence of Chemical and Mineral Compositions on the Parameterization of Immersion Freezing by Volcanic Ash Particles, Journal of Geophysical Research: Atmospheres, 126, https://doi.org/10.1029/2020JD033356, 2021.

Vepuri, H. S. K., Lacher, L., Nadolny, J., Möhler, O., and Hiranuma, N.: Online Ice-Nucleating Particle Measurements in the Southern Great Plains (SGP) Using the Portable Ice Nucleation Experiment (PINE) Chamber, p. 25, MDPI AG, https://doi.org/10.3390/ecas2020-08469, 2021.

Vergara-Temprado, J., Holden, M. A., Orton, T. R., O'Sullivan, D., Umo, N. S., Browse, J., Reddington, C., Baeza-Romero, M. T., Jones, J. M., Lea-Langton, A., Williams, A., Carslaw, K. S., and Murray, B. J.: Is Black Carbon an Unimportant Ice-Nucleating Particle in Mixed-Phase Clouds?, Journal of Geophysical Research: Atmospheres, 123, 4273–4283, https://doi.org/10.1002/2017JD027831, 2018.

Vogel, F.: PINE-04-01 CALISTHO21_new 2021, Karlsruhe Institute of Technology (KIT) RADAR4KIT, https://doi.org/10.35097/4XAWJAQWTEYVX4FR, [Dataset], 2025a.

Vogel, F.: PINE-04-01 CALISTHO21 2021, Karlsruhe Institute of Technology (KIT) RADAR4KIT, https://doi.org/10.35097/jxwbwycbyu64n2p0, [Dataset], 2025b.

Vogel, F.: PINE-04-01 CORONA 2021, Karlsruhe Institute of Technology (KIT) RADAR4KIT, https://doi.org/10.35097/j823gv823jbwy7b0, [Dataset], 2025c.

Vogel, F.: PINE-05-02 FROSTDEFEND_23 2023, Karlsruhe Institute of Technology (KIT) RADAR4KIT, https://doi.org/10.35097/mdxcqd7p646cwc0u, [Dataset], 2025d.

Wallentin, G.: Models and Scripts for "A Data-Driven Approach to Modelling a Global Distribution of Ice Nucleating Particles", RADAR4KIT, https://doi.org/10.35097/7ajnkmf1wbajeefa, [Dataset], 2026a.

Wallentin, G.: Data for "A Data-Driven Approach to Modelling a Global Distribution of Ice Nucleating Particles", RADAR4KIT, https://doi.org/10.35097/ytm7xb43na2kwh72, [Dataset], 2026b.

Welti, A., Keith Bigg, E., J. Demott, P., Gong, X., Hartmann, M., Harvey, M., Henning, S., Herenz, P., C. J. Hill, T., Hornblow, B., Leck, C., Löffler, M., S. McCluskey, C., Marie Rauker, A., Schmale, J., Tatzelt, C., Van Pinxteren, M., and Stratmann, F.: 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., Sheesley, R., Bossi, R., and Traversi, R.: Annual concentrations of ice nucleating particles at different Arctic stations, PANGAEA, https://doi.org/10.1594/PANGAEA.899701, [Dataset], 2019.

Wex, H., Huang, L., Zhang, W., Hung, H., Traversi, R., Becagli, S., Sheesley, R. J., Moffett, C. E., Barrett, T. E., Bossi, R., Skov, H., Hünerbein, A., Lubitz, J., Löffler, M., Linke, O., Hartmann, M., Herenz, P., and Stratmann, F.: Annual variability of ice-nucleating particle

https://doi.org/10.5194/egusphere-2026-4134
Preprint. Discussion started: 14 July 2026

concentrations at different Arctic locations, Atmospheric Chemistry and Physics, 19, 5293–5311, https://doi.org/10.5194/acp-19-5293-2019, 2019.

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, https://doi.org/10.1029/2024GL112583, 2025.

Wilbourn, E. and Hiranuma, N.: Abundance of ice-nucleating particles and cloud condensation nuclei measured at the Southern Great Plains and Eastern North Atlantic observatories in autumn 2019 and 2020, PANGAEA, https://doi.org/10.1594/PANGAEA.964038, [Dataset], 2023.

Wilbourn, E. K., Lacher, L., Guerrero, C., Vepuri, H. S., Höhler, K., Nadolny, J., Pantoya, A. D., Möhler, O., and Hiranuma, N.: Measurement report: A comparison of ground-level ice-nucleating-particle abundance and aerosol properties during autumn at contrasting marine and terrestrial locations, Atmospheric Chemistry and Physics, 24, 5433–5456, https://doi.org/10.5194/acp-24-5433-2024, 2024.

Wilson, T. W., Ladino, L. A., Alpert, P. A., Breckels, M. N., Brooks, I. M., Browse, J., Burrows, S. M., Carslaw, K. S., Huffman, J. A., Judd, C., Kilthau, W. P., Mason, R. H., McFiggans, G., Miller, L. A., Najera, J. J., Polishchuk, E., Rae, S., Schiller, C. L., Si, M., Temprado, J. V., Whale, T. F., Wong, J. P., Wurl, O., Yakobi-Hancock, J. D., Abbatt, J. P., Aller, J. Y., Bertram, A. K., Knopf, D. A., and Murray, B. J.: A marine biogenic source of atmospheric ice-nucleating particles, Nature, 525, 234–238, https://doi.org/10.1038/nature14986, 2015.

Worringen, A., Kandler, K., Benker, N., Dirsch, T., Mertes, S., Schenk, L., Kästner, U., Frank, F., Nillius, B., Bundke, U., Rose, D., Curtius, J., Kupiszewski, P., Weingartner, E., Vochezer, P., Schneider, J., Schmidt, S., Weinbruch, S., and Ebert, M.: Single-particle characterization of ice-nucleating particles and ice particle residuals sampled by three different techniques, Atmospheric Chemistry and Physics, 15, 4161–4178, https://doi.org/10.5194/acp-15-4161-2015, 2015.

Zhang, H., Zheng, Y., and Li, Z.: Improving Low-Cloud Fraction Prediction Through Machine Learning, Geophysical Research Letters, 51, https://doi.org/10.1029/2024GL109735, 2024.