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

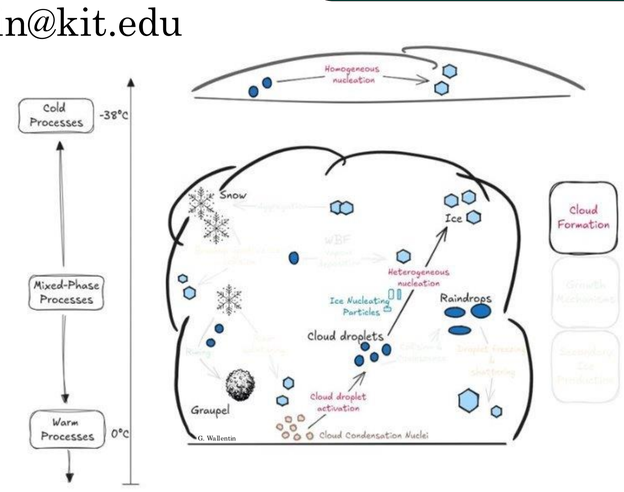
Karlsruhe Institute of Technology

Gabriella Wallentin, Alexander Böhmländer, Ross Herbert, Heike Wex, Charlotte Beall, Pia Bogert, Zoé Brasseur, Jie Chen, Jessie M. Creamean, Paul J. DeMott, Oliver Eckermann, Xianda Gong, Silvia Henning, Naruki Hiranuma, Christina S. McCluskey, Ottmar Möhler, Mark D. Tarn, Christian Tatzelt, Ping Tian, Yutaka Tobo, Franziska Vogel, André Welti, Elise Wilbourn, and Corinna Hoose

gabriella.wallentin@kit.edu

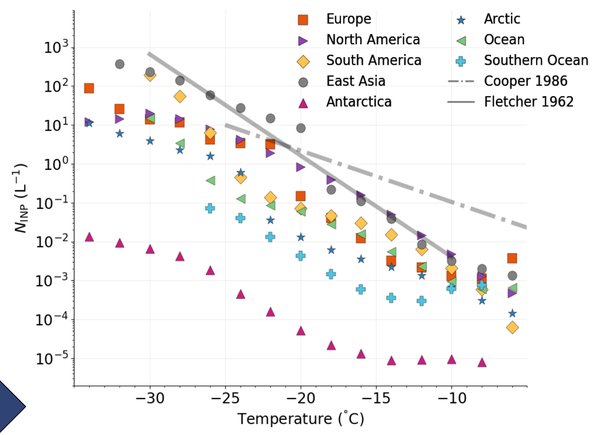
The Problem

Ice nucleating particles (INPs) are crucial for cloud ice formation in the mixed-phase temperature range



Measurements of INPs are scarce in time and space

Note the strong temperature dependence!

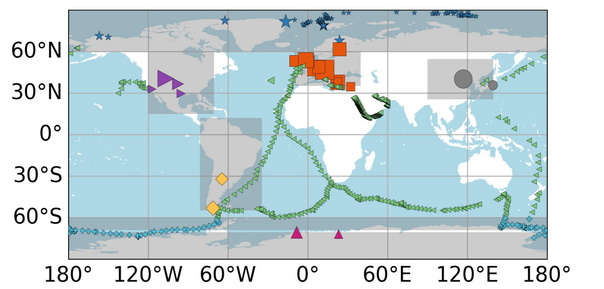


Their variability across space and time is difficult to capture with aerosol-independent models. Instead simplistic, temperature-dependent equations are defined to model this process globally

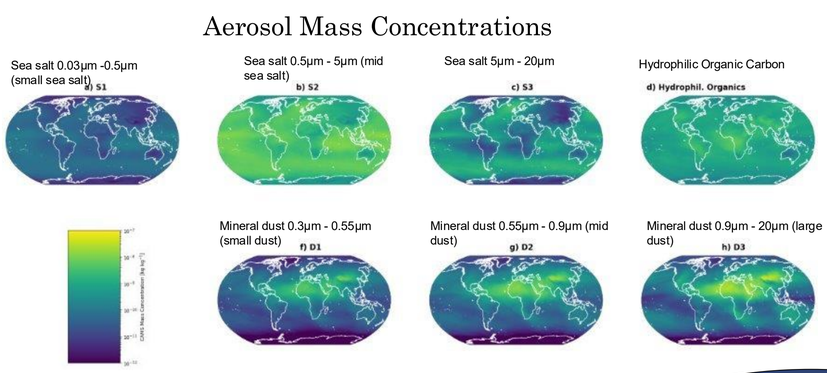
Solution: Predict INPs based on machine learning

The Method

Collect INP data



Collocate CAMS aerosols

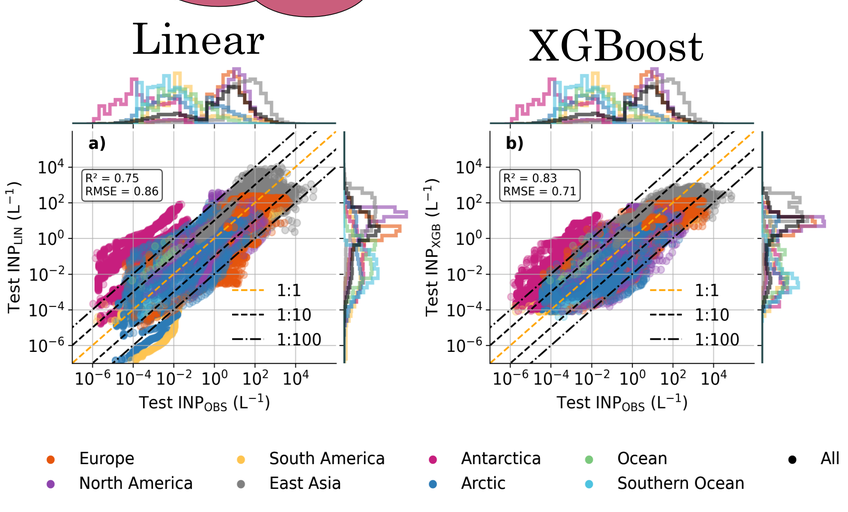


Train an XGBoost Regression Model

XGBoost is an ensemble of decision trees, based on simple yes-or-no questions where the trees are trained sequentially

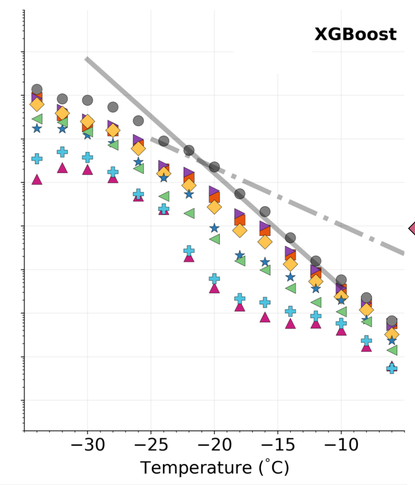
The Results

1: Does it work?

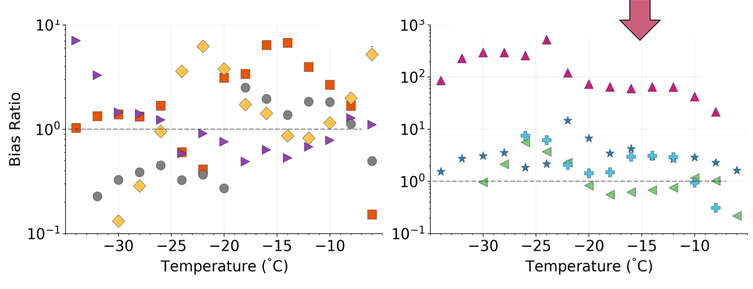


- Observed vs predicted INP concentrations separated by region.
- The linear regression benchmark performs well but XGBoost has better scores (R^2 and RMSE)
- Antarctica is difficult to predict due its small dataset used for training, less well-constrained aerosol reanalysis and difficulties predicting extreme outliers.

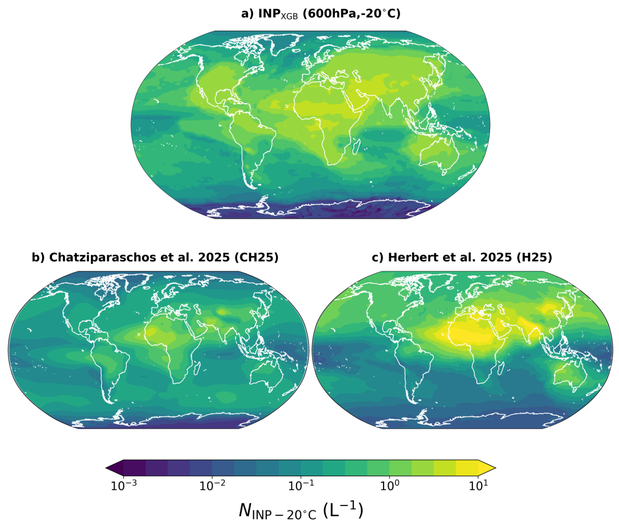
2: Is there a strong temperature dependence?



- INP concentrations is predicted at surface level
- The strong temperature dependence is captured
- Most regions compare well with observations (above)
- Mean bias (below), model mean/obs. mean shows a good agreement within mostly 1 order of magnitude



3: Realistic spatial distribution?



- INP concentration predicted at 600hPa at an activation temperature of -20°C to explore the global spatial distribution
- Comparison with aerosol-aware models with assumption based INP concentrations calculated from prognostic aerosols show a decent agreement with peaks over deserts and lower values over Antarctica and remote regions

The Conclusions

- Even with a small dataset it is possible to predict a realistic global INP distribution
- Regional climatological INP spectra (with temperature) can be extracted and agree within mostly one order of magnitude of the observations
- Spatially varying immersion freezing is possible to model without prognostic aerosols