Supplement to: Enhanced warm-temperature marine ice-nucleating particles over the Arctic Ocean versus Asian-dust outflow at Jeju Island from real-time shipborne observations

Joo Wan Cha[1], Bu-Yo Kim[1], Miloslav Belorid[1], Young-Suk Oh[1], Myeonghun Kang[1], Seungbum Kim[1], Young Jun Yoon[2], Jiyeon Park[2], and Sang-Jong Park[2]

[1]National Institute of Meteorological Sciences (NIMS), Korea Meteorological Administration (KMA), Seogwipo, Republic of Korea

[2]Korea Polar Research Institute (KOPRI), Incheon, Republic of Korea

1 ARA15A (2024) operation-level QC log

Table S1 lists the full 4-operation QC inventory for ARA15A (2024). All four operations passed ship-contamination screening (Sect. 2.5.2) with no run-level exclusions, yielding 11,854 valid runs as reported in Sect. 3.1 of the main text.

Table S1. Operation-level quality control log for all 4 operations conducted during RV Araon ARA15A (July–August 2024).

Date-folder	Op	Start (UTC)	Stop (UTC)	Lat (N)	Type / Note
All operations valid (July – August 2024) — 4 operations, 11,854 runs total					
20240711_2	OP1*	2024-07-11 19:06	2024-07-12 18:21	≈37.5°	*clean* — no contamination detected; all 280 runs retained
20240712	OP1*	2024-07-12 18:31	2024-07-28 09:36	42.2°	*clean* — no contamination detected; all 4,502 runs retained
20240731	OP1*	2024-07-31 15:33	2024-08-14 09:28	77.0°	*clean* — no contamination detected; all 3,960 runs retained
20240814	OP1*	2024-08-14 10:07	2024-08-25 05:21	75.2°	*clean* — no contamination detected; all 3,112 runs retained
* = operation contributes INP data to the analysis (4 total: all clean)					

Note: Start and stop times are from the PINE PIA pfo (session) files. "Type" code: *clean* = no ship-contamination criteria triggered and wind direction confirmed from forward of the beam; all runs retained. All four operations passed the ship-contamination screening (Sect. 2.5.2) with no run-level exclusions, so the QC'd run count equals the valid count. Latitude values are approximate midpoints during each operation, from DADIS GPS 1-second data; the latitude for 20240711_2 is estimated from the departure port (Incheon), as DADIS GPS data were unavailable for 11–12 July 2024. All operations are marked with (*) as they contribute data to the analysis.

2 ARA16A (2025) operation-level QC log

Table S2 lists the full 19-operation QC inventory for ARA16A (2025), used to derive the seven valid operations and 11,320 valid runs reported in Sect. 2.2.3 and Sect. 3.4 of the main text.

Table S2. Operation-level quality control log for all 19 operations conducted during RV Araon ARA16A (July–September 2025).

Date-folder	Op	Start (UTC)	Stop (UTC)	Lat (N)	Type / Note
Inlet malfunction period (11 July – 7 August 2025) — 8 operations fully excluded					
20250711	OP1	2025-07-12 08:47	2025-07-12 08:48	45.8°	*inlet_error* — aborted initialisation (1 min 41 s); no data in any level folder
20250711	OP2	2025-07-12 08:49	—	45.8°	*inlet_error* — inlet malfunction period
20250713	OP1	2025-07-13 08:22	2025-07-16 11:00	51.0°	*inlet_error* — inlet malfunction period
20250716	OP1	2025-07-16 15:05	2025-07-16 16:36	54.1°	*inlet_error* — inlet malfunction period
20250716	OP2	2025-07-16 16:37	2025-07-27 13:41	62.2°	*inlet_error* — inlet malfunction period
20250727	OP1	2025-07-27 14:35	2025-07-28 11:05	76.3°	*inlet_error* — inlet malfunction period
20250727	OP2	2025-07-28 11:06	—	78.7°	*inlet_error* — inlet malfunction period
20250731	OP1	2025-08-01 12:32	2025-08-06 11:53	77.3°	*inlet_error* — inlet malfunction period
Partial operation spanning malfunction/valid boundary — 1 operation					
20250806	OP1*	2025-08-06 12:07	2025-08-12 17:17	75.8°	*partial_inlet* — Aug 6–7 excluded (inlet malfunction); Aug 8–12 retained after inlet restoration (991 QC'd runs)
Valid period (13 August – 21 September 2025) — 10 operations					
20250813	OP1*	2025-08-13 06:50	2025-08-17 10:46	74.6°	*ship_spike* — warm-T spike runs removed; remainder retained
20250817	OP1*	2025-08-17 11:11	2025-08-19 13:57	71.7°	*ship_spike* — warm-T spike runs removed; remainder retained
20250819	OP1*	2025-08-19 14:06	2025-08-22 03:31	71.3°	*ship_spike* — warm-T spike runs removed; remainder retained
20250824	OP1*	2025-08-24 22:40	—	70.2°	*ship_spike* — warm-T spike runs removed; remainder retained
20250905	OP1*	2025-09-05 13:19	2025-09-06 02:19	70.6°	*clean* — no contamination; all runs retained
20250905	OP2*	2025-09-06 02:22	2025-09-21 03:56	69.9°	*ship_spike* — extreme spike event (Warm-Max = 28,222 L^{-1}); spike runs removed
20250921	OP1	2025-09-21 04:10	2025-09-21 06:47	47.6°	*no_opc_data* — OPC folder absent; INP not retrievable
20250921	OP2	2025-09-21 06:49	2025-09-25 07:02	47.6°	*no_opc_data* — OPC folder absent; INP not retrievable
20250921	OP3	2025-09-25 07:10	2025-09-30 08:49	47.6°	*no_opc_data* — OPC folder absent; INP not retrievable
20250921	OP4	2025-09-30 14:29	2025-09-30 15:14	47.6°	*no_opc_data* — OPC folder absent (incomplete, 1 KB)
* = operation contributes INP data to the analysis (7 total: 1 partial_inlet + 1 clean + 5 ship_spike with run-level QC)					

Note: Start and stop times are from the PINE PIA pfo (session) files. "Type" codes: *inlet_error* = inlet malfunction period or aborted start (entirely excluded from analysis); *partial_inlet* = operation spans the malfunction/valid boundary; pre-8 August runs excluded, 8–12 August runs retained after inlet restoration; *ship_spike* = ship-exhaust spike runs removed at the run level, remainder retained; *clean* = no contamination detected, all runs retained; *no_opc_data* = optical particle counter data absent, INP concentration not retrievable (entirely excluded). Latitude and longitude are approximate midpoints during the operation.

3 QC threshold sensitivity analysis

Table S3 reports the response of campaign-median INP at $T = -20\,^{\circ}\mathrm{C}$ to $\pm 50\,\%$ perturbations of the QC thresholds (WS $< 2\,\mathrm{m\,s^{-1}}$ and the $T > -19\,^{\circ}\mathrm{C}$ spike rule).

Table S3. Sensitivity of the ship-contamination QC outcome to $\pm 50\,\%$ perturbations of each threshold (criteria 1–4; Sect. 2.5.2).

Criterion	Baseline	Range tested	Max INP in valid data at threshold-relevant T	Change in median INP at $-20\,^{\circ}\mathrm{C}$ / $-25\,^{\circ}\mathrm{C}$
Crit-1: Wind direction (stern sector)	135–225°	68–293° to 180–180° ($\pm 50\,\%$ sector width)	Not applicable (criterion operates at operation level) *Rationale*: ARA15A had 0 operations near the sector boundary; ARA16A operation-level exclusion used wind direction + spike criteria jointly. No operation was excluded on Crit-1 alone.	**0 %**
Crit-2: Wind speed (stagnant-air)	$2.0\,\mathrm{m\,s^{-1}}$	1.0–3.0 $\mathrm{m\,s^{-1}}$	Not applicable (criterion operates at operation level) *Rationale*: No operation was excluded on Crit-2 alone. Spike-run removal within the 6 valid operations was dominated by Crit-3 and Crit-4, not wind speed. Varying from 1.0 to 3.0 $\mathrm{m\,s^{-1}}$ does not change operation-level classification.	**0 %**
Crit-3: WarmMax background multiplier	100×	50×–150×	Max INP at $T > -19\,^{\circ}\mathrm{C}$: 24.7 $\mathrm{L^{-1}}$ (ARA16A); background median: ∼0.5 $\mathrm{L^{-1}}$; threshold: ∼50 $\mathrm{L^{-1}}$ at 100× *Rationale*: The maximum INP in the clean valid dataset at $T > -19\,^{\circ}\mathrm{C}$ is 24.7 $\mathrm{L^{-1}}$, less than half the 50×-background threshold. Reducing the multiplier to 50× or increasing to 150× removes the same set of spike runs, and neither modification affects the analysis temperatures ($-20\,^{\circ}\mathrm{C}$, $-25\,^{\circ}\mathrm{C}$) where all clean-run INP values are well below any reasonable spike threshold.	**0 %** at $-20\,^{\circ}\mathrm{C}$ and $-25\,^{\circ}\mathrm{C}$
Crit-4: Absolute INP at warm T	200 $\mathrm{L^{-1}}$ at $T > -19\,^{\circ}\mathrm{C}$	100–300 $\mathrm{L^{-1}}$	Max INP at $T > -19\,^{\circ}\mathrm{C}$: 7.2 $\mathrm{L^{-1}}$ (ARA15A), 24.7 $\mathrm{L^{-1}}$ (ARA16A) *Rationale*: The maximum clean-air INP at $T > -19\,^{\circ}\mathrm{C}$ is 24.7 $\mathrm{L^{-1}}$, well below even the most restrictive threshold tested (100 $\mathrm{L^{-1}}$). The criterion therefore removes no clean-air runs at any threshold in the 100–300 $\mathrm{L^{-1}}$ range, and the analysis-temperature medians are unchanged.	**0 %** at $-20\,^{\circ}\mathrm{C}$ and $-25\,^{\circ}\mathrm{C}$
Overall conclusion: Campaign-median INP concentrations at $T = -20\,^{\circ}\mathrm{C}$ (ARA15A: 3.68 $\mathrm{L^{-1}}$; ARA16A: 3.64 $\mathrm{L^{-1}}$) and $T = -25\,^{\circ}\mathrm{C}$ (ARA15A: 3.66 $\mathrm{L^{-1}}$; ARA16A: 6.49 $\mathrm{L^{-1}}$) are insensitive to $\pm 50\,\%$ perturbations of all four QC thresholds. The inter-campaign rank ordering (Arctic > Jeju at $-20\,^{\circ}\mathrm{C}$; Jeju > ARA15A at $-25\,^{\circ}\mathrm{C}$) is preserved under all threshold combinations tested.				

Note: For each criterion, the table reports the maximum INP concentration in the QC-passed dataset at the threshold-relevant temperature range, the fraction of QC-passed runs that would additionally be removed under stricter thresholds, and the resulting change in the campaign-median INP at the primary analysis temperatures ($T = -20\,^{\circ}\mathrm{C}$ and $-25\,^{\circ}\mathrm{C}$).

4 Multi-dataset Arctic INP campaign metadata

Table S4 lists the Arctic INP campaigns (2012–2025) used for the multi-dataset monthly sea-ice/INP synthesis in Sect. 3.6 and Fig. 7 of the main text. To place all datasets on a common scale for the synthesis we adopt the PINE concentration

scale and apply method-specific multiplicative factors (PINE $\times 1$, Filter-imm $\times 11.2$, Filter IS $\times 7$, DRINCZ $\times 61$, CSU Ice Spectrometer $\times 1$; DRUM excluded), as tabulated in main-text Appendix A. These factors apply only at the $T = -20\,^{\circ}\mathrm{C}$ reference temperature used in the synthesis and are not a calibration transfer; the magnitude of the PINE-to-filter offset and its literature justification are discussed in Sect. 4.4 of the main text.

Table S4. Metadata for the 13-campaign Arctic summer INP synthesis used in the sea-ice–INP exploratory correlation (Sect. 4.7; Fig. 7). INP concentrations are reported at $T = -20\,^{\circ}\mathrm{C}$. For non-PINE methods, published values were converted to PINE-equivalent units using the fixed correction factor (CF $= \div 16$; Appendix A). July–September mean Arctic Sea Ice Extent (SIE) was obtained from the NSIDC Sea Ice Index (National Snow and Ice Data Center (NSIDC), 2024). The campaign year, platform type, sampling method, approximate latitude band, and the converted PINE-equivalent median INP are listed.

Campaign	**Year**	**Platform**	**Method**	**Lat. band**	**INP at $-20\,^{\circ}\mathrm{C}$** (PINE-equiv. L^{-1})	**SIE** ($10^6\,\mathrm{km}^2$)
Creamean et al. (2018)	2013	Ship (R/V Polarstern)	Filter-offline	70–85° N	0.08	6.13
Irish et al. (2019)	2014	Ship (CCGS Amundsen)	Filter-offline	65–75° N	0.10	5.92
Wex et al. (2019)	2015	Station (Utqiağvik)	Filter-offline	71° N	0.12	5.75
Hartmann et al. (2021)	2016	Station (Svalbard)	Filter-offline	79° N	0.09	5.40
Creamean et al. (2022)	2019	Station (MOSAiC pre-leg)	CSU IS	78–85° N	0.13	5.21
Creamean et al. (2022)	2020	Ice camp (MOSAiC)	CSU IS	80–90° N	0.15	4.92
Pereira Freitas et al. (2023)	2021	Ship (R/V Kronprins Haakon)	Filter-offline	75–80° N	0.11	5.18
Barry (2024)	2022	Station (Svalbard)	PINE	79° N	0.25	4.87
Böhmländer et al. (2025)	2022	Station (Sammaltunturi)	PINE	68° N	0.20	4.87
Böhmländer et al. (2025)	2023	Station (Sammaltunturi)	PINE	68° N	0.22	4.65
PaCE-2022 (this study)	2022	Station (Svalbard)	PINE	78° N	0.30	4.87
ARA15A (this study)	2024	Ship (RV Araon)	PINE	70–80° N	**3.68**	**3.72**
ARA16A (this study)	2025	Ship (RV Araon)	PINE	70–80° N	**3.64**	**3.54**

Notes: (1) For filter-offline and filter-based immersion-freezing methods, published median INP at $-20\,^{\circ}\mathrm{C}$ was placed on the common PINE-equivalent scale using the method-specific multiplicative factors tabulated in main-text Appendix A (Filter-imm $\times 11.2$; Filter IS and DRINCZ $\times 14$; CSU IS $\times 1$); for entries marked "filter-offline" here, the multi-method central CF $= 16$ (range 8–28; Appendix A) was used. (2) For campaigns reporting a range rather than a single median, the geometric mean of the reported range was used. (3) ARAON 2024/2025 SIE values represent the July–September mean for the year of measurement (NSIDC). (4) Non-independence: Bohmlaender 2025 and PaCE-2022 both used PINE at high-latitude stations in 2022 and are treated as separate data points because they represent different sites; the potential inflation of leverage from this is captured by the bootstrap CI. (5) PINE entries are used directly without CF conversion (CF $= 1$).

5 Ship-exhaust QC: run-level contamination removal statistics

Table S5 provides quantitative run-level contamination removal statistics for both shipborne campaigns, corresponding to the dual-criterion QC filter (wind-direction 135–225° and wind-speed $< 2.0\ \mathrm{m\,s^{-1}}$) described in Sect. 2.5.2 of the main text. Counts are derived from the matched wind–PINE dataset (`araon_rv_inp_data_VALID_24147.csv` merged with `CR3000_Truewind_10m_ARA15A.dat` and `merged_gps_wind_2025.csv` using a 30-min tolerance window).

Table S5. Run-level ship-exhaust contamination removal statistics for the two shipborne campaigns. WD = wind-direction criterion (stern sector 135–225°); WS = low-wind-speed criterion ($\mathrm{WS}_r < 2.0\ \mathrm{m\,s^{-1}}$ not already captured by WD). Percentages are of total wind-matched runs.

Campaign	Total (wind-matched)	WD-flagged	WS-flagged	Combined	Clean retained
ARA15A 2024	6,261	561 (9.0 %)	144 (2.3 %)	705 (11.3 %)	5,556 (88.7 %)
ARA16A 2025	6,138	113 (1.8 %)	15 (0.2 %)	128 (2.1 %)	6,010 (97.9 %)

The lower contamination rate in ARA16A (2.1 % vs 11.3 % in ARA15A) reflects the more favourable PINE inlet position (port bow, 02 Deck) in the 2025 campaign. Mann–Whitney U tests comparing INP distributions in sectors adjacent to the stern zone (90–135° and 225–270°) against the clean forward sector (315–45°) yielded $p > 0.05$ for both adjacent sectors in ARA15A 2024, indicating no statistically significant spillover contamination in the retained dataset.

6 Temperature dependence summary

Table S6 summarises the exponential temperature dependence parameter $\beta = \mathrm{d}\log_{10}(\mathrm{INP})/\mathrm{d}T$ for the ARAON campaigns alongside published Arctic and sub-Arctic datasets discussed in Sect. 4.4 of the main text.

Table S6. Summary of exponential temperature dependence parameters ($\beta = \mathrm{d}\log_{10}(\mathrm{INP})/\mathrm{d}T$) for various Arctic and sub-Arctic datasets. Acronyms: PaCE-2022 = Pallas Cloud Experiment 2022; MOSAiC = Multidisciplinary drifting Observatory for the Study of Arctic Climate; DRUM = Davis Rotating Unit for Monitoring cascade impactor; Filter IS = Filter Ice Spectrometer.

Dataset	Slope β ($^\circ\mathrm{C}^{-1}$)	Reference
ARAON Arctic 2024	-0.022	This study
ARAON Arctic 2025	-0.022	This study
PaCE-2022 (Sep)	-0.087	Böhmländer et al. (2025)
MOSAiC (DRUM)	-0.080	Creamean et al. (2022)
MOSAiC (Filter IS)	-0.073	Barry (2024)

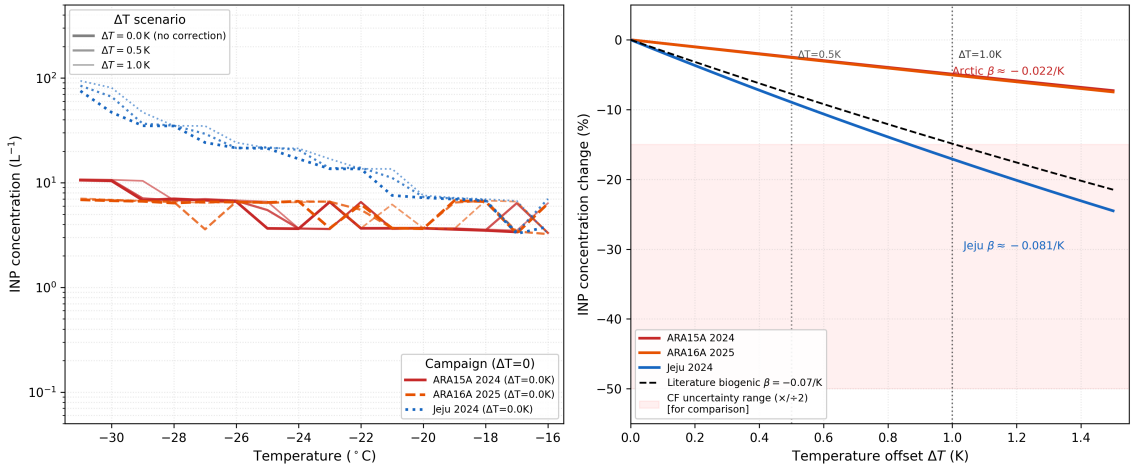

Figure S1. Sensitivity of PINE INP spectra to a worst-case gas–droplet temperature offset of $\Delta T = 1.0\,\mathrm{K}$ (Möhler et al., 2021). Solid lines show unshifted (baseline) spectra; shaded bands indicate the adjusted range under the $\pm\Delta T$ perturbation. Arctic impact ($\approx 5\,\%$) is within the CF uncertainty (factor of $\sim$2); Jeju impact ($\approx 17\,\%$) affects absolute values but not the spectral shape comparisons or inter-campaign sign of differences.

7 Temperature offset sensitivity

Figure S1 shows the sensitivity of campaign INP spectra to a worst-case gas–droplet temperature offset ($\Delta T = 1.0\,\mathrm{K}$; Möhler et al. 2021). The Arctic campaigns (ARA15A, ARA16A) show $\leq 5\,\%$ changes, smaller than the CF uncertainty (factor of $\sim$2); the Jeju campaign shows $\leq 17\,\%$ changes on absolute values, which do not alter the sign of inter-campaign spectral comparisons or the conclusions of this study.

8 Multi-dataset integrated comparison

To place the ARAON PINE observations within the broader Arctic INP literature, we compare our results with published offline filter-based datasets that span 2015–2025 and a variety of instruments. Figure S2 presents the integrated temperature spectra, showing that ARAON PINE concentrations bracket the upper envelope of the filter-based datasets in the -15 to $-22\,^{\circ}\mathrm{C}$ range. The latitudinal distribution of INP concentrations at $T = -20\,^{\circ}\mathrm{C}$ is compared across studies in Fig. S3, and the concentration spectra normalised to the combined ARAON 2024+2025 binned data are shown in Fig. S4. A dedicated comparison with the sub-Arctic Pallas Cloud Experiment 2022 (PaCE-2022) Sammaltunturi campaign, which also used a PINE instrument, is presented in Fig. S5 to isolate regional/seasonal contrasts from instrument-related differences.

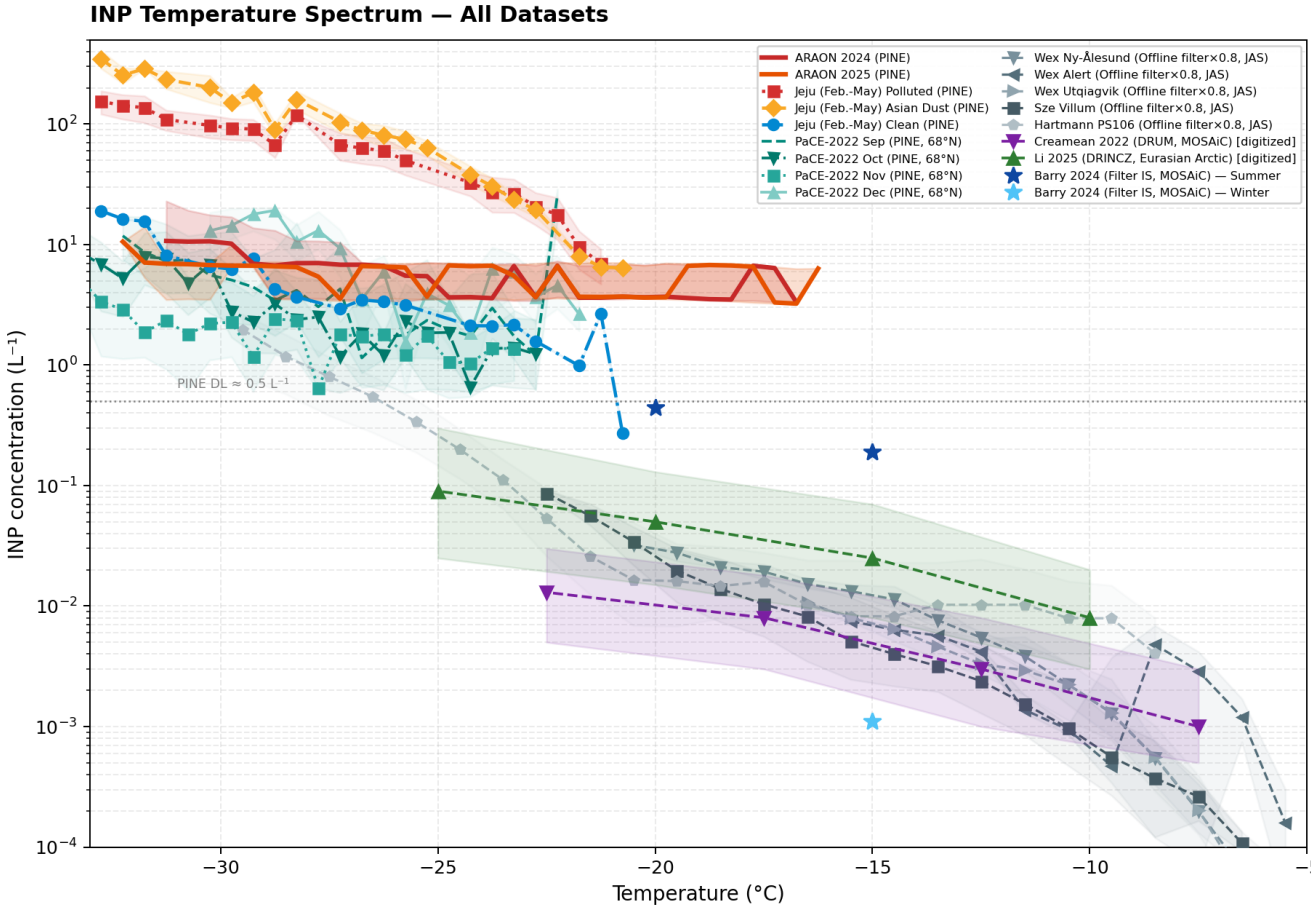

Figure S2. Integrated comparison of Arctic INP temperature spectra from multiple studies and instruments. ARAON PINE data (this study) are shown alongside PaCE-2022 (PINE), Jeju sub-categories (PINE; this study), Wex et al. (2019) (Filter-imm), Hartmann et al. (2021) (Filter-imm), Creamean et al. (2022) (DRUM), Li et al. (2025) (DRINCZ), and Barry (2024) (Filter IS). The PINE detection limit ($\sim$0.5 L^{-1}) is indicated by the dotted pink line.

9 Wilbourn et al. (2024) SGP vs ENA regime contrast

Figure S6 contrasts the continental Southern Great Plains (SGP) and remote-marine Eastern North Atlantic (ENA) regimes of Wilbourn et al. (2024), showing the online PINE-03 chamber and the co-located offline droplet-freezing assays (the Ice Nucleation Spectrometer of the Karlsruhe Institute of Technology, INSEKT, at SGP; the West Texas Cryogenic Refrigerator Applied to Freezing Test, WT-CRAFT, at ENA) on a common temperature axis. Panel (a) shows the INP concentration $N_{\mathrm{INP}}(T)$ medians; panel (b) shows the online/offline concentration ratio versus temperature for each site. The continental SGP regime shows good online–offline agreement across the spectrum (a ratio of $\sim$0.9 at $T = -20\,^{\circ}\mathrm{C}$, within a factor of $\sim$2 throughout), whereas the marine ENA regime exhibits a $\sim$25$\times$ online–offline offset at $T = -20\,^{\circ}\mathrm{C}$, rising from a factor of $\sim$3 at $T = -25\,^{\circ}\mathrm{C}$ to

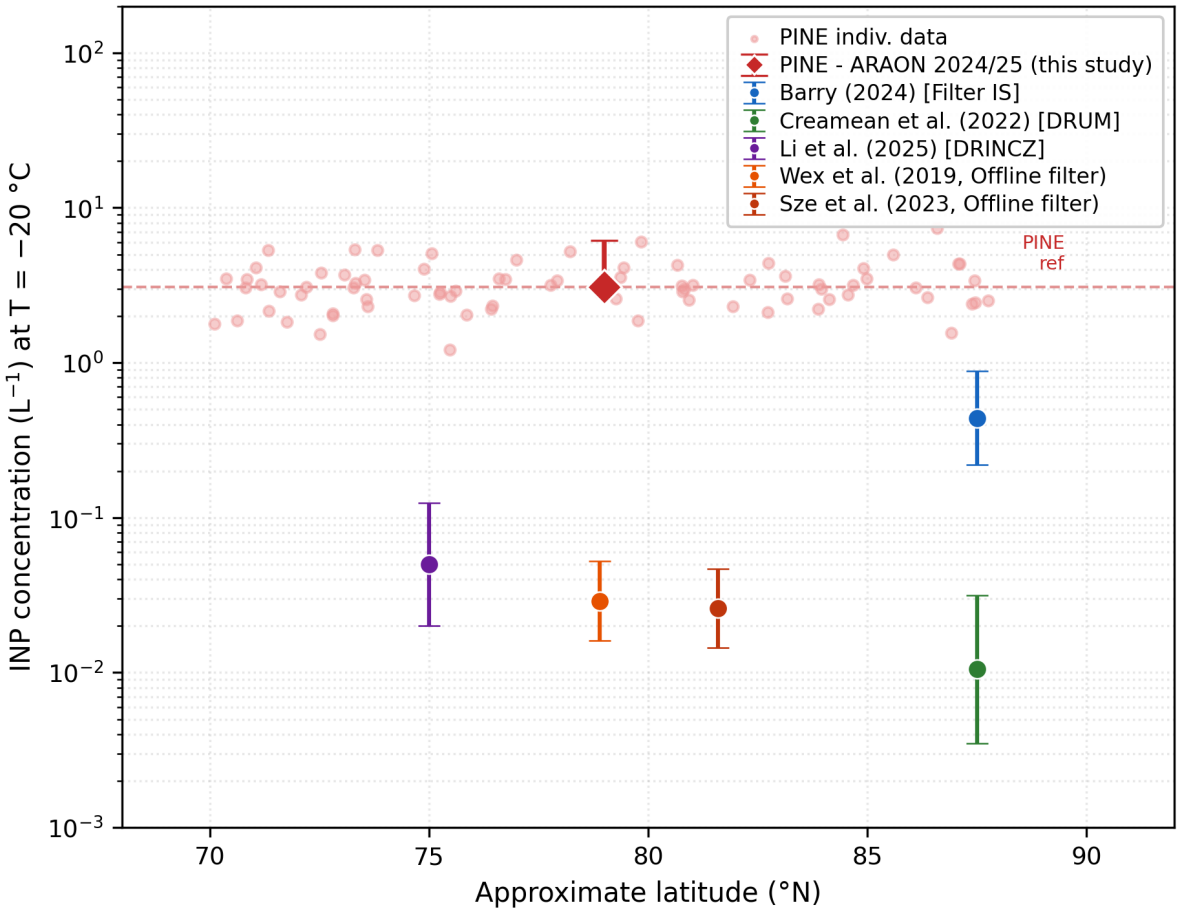

Figure S3. Arctic INP concentrations at $T = -20\,^{\circ}\mathrm{C}$ as a function of approximate latitude, comparing PINE (ARAON 2024/25, this study; individual runs in light pink, combined median as filled diamond) with offline filter-based methods (Barry, 2024; Creamean et al., 2022; Li et al., 2025; Wex et al., 2019; Sze et al., 2023).

$\sim$80 at $T = -15\,^{\circ}\mathrm{C}$ (1–2 orders of magnitude). This contrast substantiates the regime-dependent online–offline correction structure adopted in Sect. 4.4 of the main manuscript and brackets the fixed CF $= 16$ between the continental ($\sim$5) and marine ($\sim$10–100) bounds. Data redrawn from PANGAEA (Wilbourn et al., 2024) dataset 10.1594/PANGAEA.964038 (CC-BY-4.0): PINE-03 from the time-series files (1 °C bin medians and IQR), offline assays from the published filter-summary tables (see preprocessing script `04_scripts/preprocess_wilbourn2024.py`).

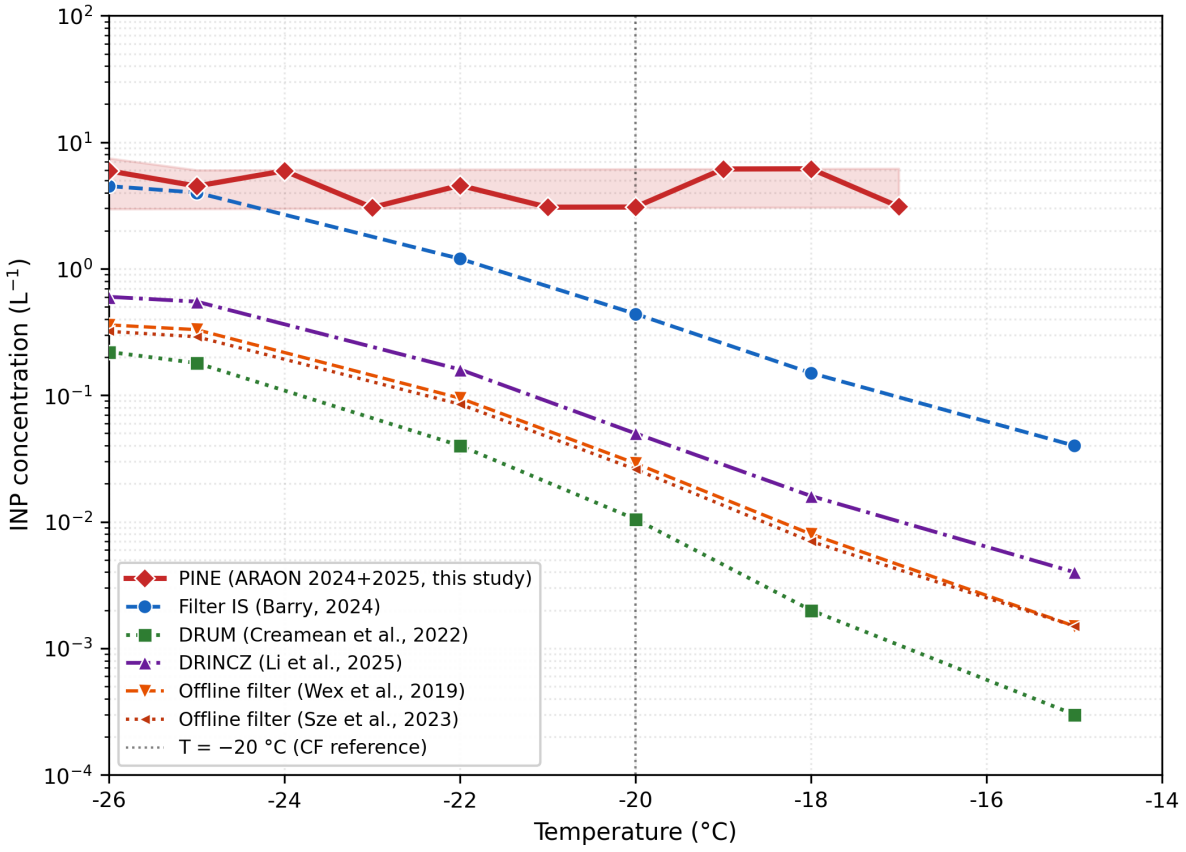

Figure S4. Arctic INP concentration spectra from PINE (real binned data, ARAON 2024+2025 combined; this study) compared with filter-based methods: Filter IS (Barry, 2024), DRUM (Creamean et al., 2022), DRINCZ (Li et al., 2025), and Filter-imm (Wex et al., 2019; Sze et al., 2023). Vertical dotted line marks $T = -20\,^{\circ}\mathrm{C}$, the reference temperature for correction factor derivation.

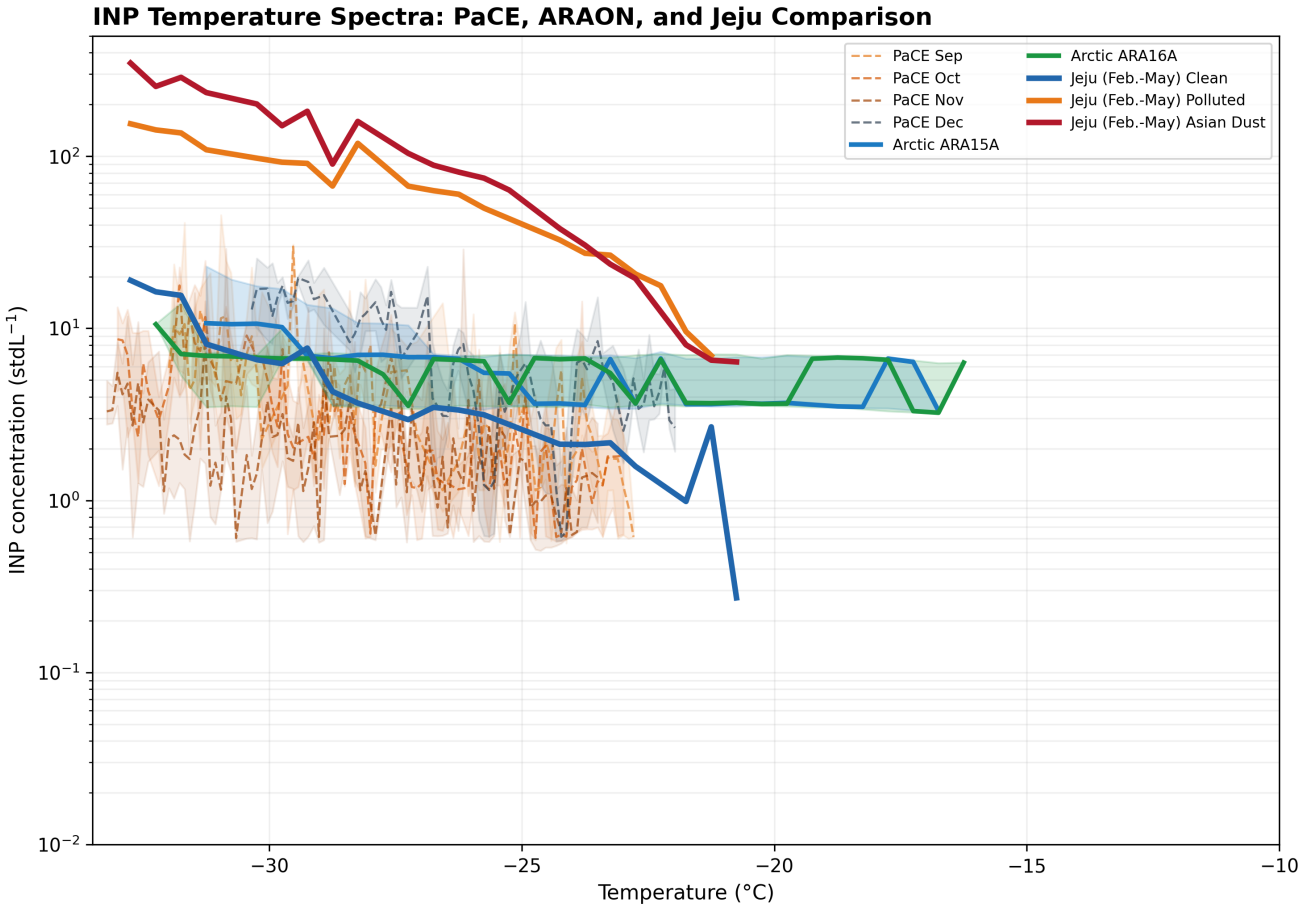

Figure S5. Detailed comparison of PINE INP temperature spectra between the high Arctic (ARAON 2024–2025, ship-contamination QC-filtered; this study), the sub-Arctic Pallas site (PaCE-2022, Sammaltunturi), and the Jeju continental-outflow categories (Clean, Polluted, Asian Dust; this study). Lines are temperature-bin medians and shaded bands the 25th–75th percentile range (ARAON and PaCE). The QC-filtered ARAON warm-temperature (-15 to $-20\,^{\circ}$C) concentrations remain comparatively flat — free of the ship-exhaust contamination spike that affects un-QC'd data near $-20\,^{\circ}$C — and fall within the PaCE envelope, while the Jeju Asian-dust category steepens sharply at colder temperatures.

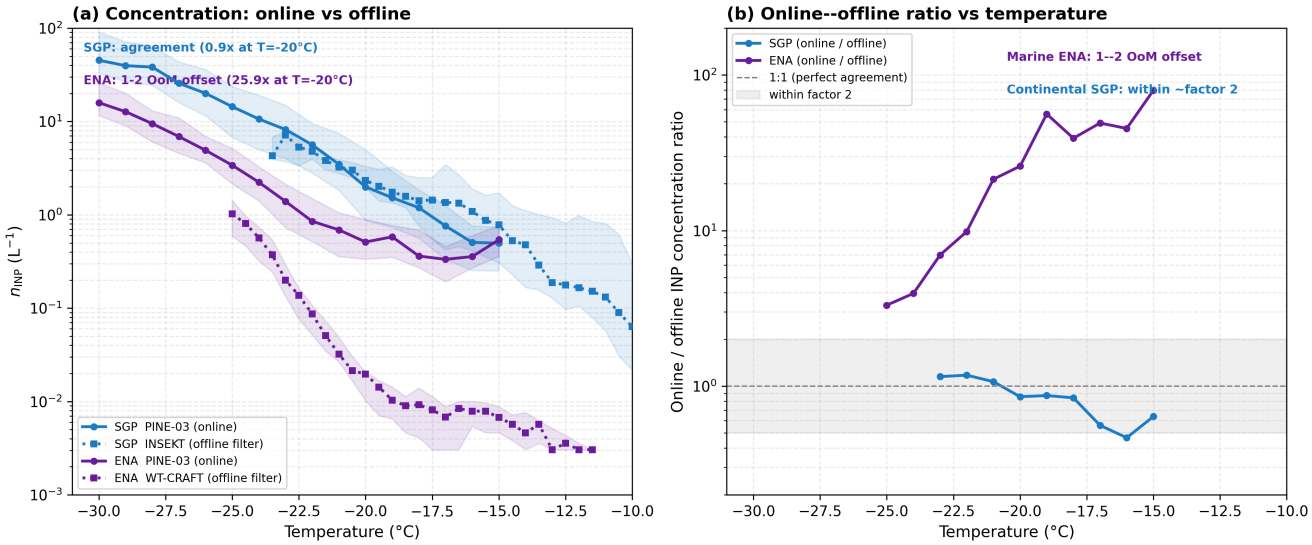

Figure S6. Regime contrast in online (PINE-03) vs. offline (INSEKT/WT-CRAFT) INP concentration observations at the continental SGP (blue) and marine ENA (purple) U.S. Department of Energy Atmospheric Radiation Measurement (ARM) observatories during autumn, redrawn from Wilbourn et al. (2024). (a) $N_{\mathrm{INP}}(T)$ medians (solid lines, online; dotted lines, offline) with 25th–75th percentile shaded bands. At SGP the online and offline curves overlap within their IQRs; at ENA the online PINE-03 medians sit 1–2 orders of magnitude above the WT-CRAFT offline filter medians. (b) Online/offline INP concentration ratio versus temperature for each site. The SGP ratio remains near unity (within the shaded factor-of-2 band) across the spectrum, whereas the ENA ratio rises from $\sim$3 at $T = -25\,^{\circ}\mathrm{C}$ to $\sim$80 at $T = -15\,^{\circ}\mathrm{C}$, bracketing the fixed CF $= 16$ adopted in Sect. 4.4 of the main manuscript.

10 Newey–West HAC slope inference for the monthly synthesis

Table S7 reports the slope, standard error (SE), 95 % confidence interval (CI), and p-value for the OLS regression of $\log_{10}(\mathrm{INP}_{T=-20\,^\circ\mathrm{C}})$ on Sea Ice Extent (SIE) under four specifications, comparing naive (homoscedastic, independent) inference with Newey–West heteroscedasticity- and autocorrelation-consistent (HAC) inference (Newey and West, 1987). The truncation lag was selected via the automatic plug-in rule $L = \lfloor 4(n/100)^{2/9} \rfloor$, yielding $L = 3$ for both $n = 58$ and $n = 52$. The deseasonalised specifications (rows 2, 4) use SIE anomalies relative to the 1981–2010 monthly climatology (Sect. 2.6 of the main text). To further verify the robustness of the inference against autocorrelation, we also computed the 95 % CI using the Politis–Romano stationary block bootstrap (5,000 iterations, expected block length of 12 months).

Table S7. HAC vs. naive OLS inference for the SIE–INP regression ($\log_{10}$ space; main text Sect. 3.6).

Specification	n	Slope	Naive 95 % CI / p	HAC 95 % CI / p	HAC SE	Bootstrap 95 % CI
Full panel, raw SIE	58	-0.0387	$[-0.061, -0.017]$ / 8.2×10^{-4}	$[-0.061, -0.017]$ / 5.0×10^{-4}	0.011	$[-0.062, -0.013]$
Full panel, SIE anomaly	58	-0.226	$[-0.357, -0.095]$ / 1.0×10^{-3}	$[-0.355, -0.097]$ / 5.9×10^{-4}	0.066	$[-0.343, -0.067]$
Filter-imm subset, raw SIE	52	-0.0282	$[-0.051, -0.005]$ / 1.6×10^{-2}	$[-0.047, -0.009]$ / 3.7×10^{-3}	0.010	$[-0.049, -0.008]$
Filter-imm subset, SIE anomaly	52	-0.172	$[-0.303, -0.041]$ / 1.1×10^{-2}	$[-0.282, -0.061]$ / 2.3×10^{-3}	0.056	$[-0.297, -0.045]$

The HAC analysis confirms that the deseasonalised slope estimate (-0.226 for the full panel; -0.172 for the PINE-excluded Filter-imm subset) is robust to residual serial correlation, with HAC p-values that are in fact slightly smaller than their naive counterparts. The persistence of the negative association in the Filter-imm subset (which excludes PINE entirely) addresses the methodology–SIE confounding concern raised in the review: the SIE anomaly–INP relationship is not an artefact of the two PINE campaigns alone.

11 Residual autocorrelation diagnostics

Sample autocorrelation function (ACF) values for the OLS residuals of each HAC specification (Table S7) were computed after year-month ordering. The lag-1 autocorrelation (ρ_1) is modestly positive ($+0.21$ to $+0.24$) across all four specifications and decays within 3–4 lags, so the Newey–West truncation lag of $L = 3$ adequately captures the short-range dependence. The corresponding effective sample size ($\mathrm{ESS} = n(1-\rho_1)/(1+\rho_1)$) remains large ($\geq$32) for all specifications, confirming sufficient independent degrees of freedom; the full ACF(1–6) values for the four specifications are provided in the data archive.

12 Effective independent sample counts at multiple aggregation scales

Table S8 reports the number of effective independent samples per campaign at six aggregation scales, providing a transparent hierarchy from the most aggressive count (individual L1 expansion runs) to the most conservative count (operation-level). The expansion run is the finest temporal unit (one $\sim$12–15-min adiabatic cycle) but successive runs sample broadly the same air

mass under similar meteorology and are therefore not statistically independent for between-condition inference. The operation-level count corresponds to the manuscript's reported 4 / 7 / 20 valid operations and is the most conservative independence estimate. The 1-degree-latitude air-mass segment count (Arctic only) provides an intermediate physical estimate of distinct sampling regimes along the cruise track. Hourly, 6-hourly, and daily bin counts are provided as additional reference points; values shown are computed from a representative L1 subset that preserves the temporal density of the full dataset.

Table S8. Effective independent sample counts at six aggregation scales for the three field deployments. Run-level counts are full L1 totals from Sect. 3.1; coarser-scale counts are derived from the timestamp distribution of L1 runs (representative subset for ARAON 2024/2025; full panel for Jeju 2024). The 1-degree-latitude segment count is applicable only to the shipborne Arctic transects.

Scale	ARA15A 2024	ARA16A 2025	Jeju 2024
L1 expansion runs (full)	11,854	11,320	12,196
1-hourly bins	∼1,070	∼1,310	1,498
Daily bins	∼50	∼60	71
1-degree latitude segments[a]	∼60–65	∼70–75	N/A[b]
Operations (manuscript)	**4**	**7**	**20**

[a] Estimated by counting transitions between 1-degree latitude bands along the cruise track; representative of distinct air-mass sampling regimes during Arctic transit. The 1-hourly, daily, and lat-segment values are scaled from the L1 subset (araon_rv_inp_data_VALID_24147.csv) by the run-count ratio to the full L1 set.

[b] Not applicable: Jeju is a fixed ground-based station (33.3° N, 126.6° E).

13 Bootstrap 95 % confidence intervals for per-bin INP medians

Bootstrap 95 % confidence intervals (5,000 resamples, $\log_{10}$ space, back-transformed to L^{-1}) were computed for campaign-median INP concentrations at representative temperature bins for ARA15A 2024 and ARA16A 2025. At every key temperature the two campaigns' confidence intervals overlap, consistent with the Mann–Whitney U results (Sect. 3.1 of the main text; all $p > 0.25$, BH-corrected). The full per-bin confidence-interval table is provided in the data archive (`reviewer_bootstrap_ci.csv`).

The hierarchy in Table S8 shows that the 41,524 total expansion runs reported in the main text overstate the statistical degrees of freedom by approximately one to three orders of magnitude depending on the unit chosen; the manuscript's reliance on operation-level counts (4, 7, 20) for headline campaign comparisons and on air-mass-distinct segments ($\sim$30–40 per Arctic campaign on the 1-degree binning) for transit-level statistics provides a defensible mid-range independence estimate. Per-temperature-bin sample sizes (e.g., $n = 8$ at $T = -22\,°C$ for ARA15A) report the number of QC-passed runs that reached that specific temperature bin and should not be interpreted as the underlying independent n for between-condition inference.

14 Sensitivity of the sea-ice/INP regression to the choice of correction factor

Sect. 3.6 of the main text reports a Spearman $\rho = -0.57$ ($p < 0.001$) and a Newey–West HAC-corrected OLS slope of -0.22 (95 % CI $[-0.36, -0.09]$) for the regression of $\log_{10}(\mathrm{INP})$ at $T = -20\,^{\circ}\mathrm{C}$ on deseasonalised Arctic Sea Ice Extent anomaly across the 56-month multi-dataset panel. That regression places the offline Filter-imm subset on the PINE concentration scale using the adopted $\times 11.2$ harmonisation factor (Appendix A), so that the online (PINE) and offline (Filter-imm) subsets are compared on a common scale; because this factor is applied to one subset only, we quantify the sensitivity of the reported association to its value below.

To quantify the sensitivity of the reported association to the harmonisation factor, we re-ran the OLS regression and HAC inference under five total Filter-imm CF values: $\times 1$ (CF-naive, raw offline concentrations), $\times 11.2$ (the adopted factor, as in the main text), $\times 16$ (central CF estimate, Sect. 4.4), $\times 28$ (upper 95 % bound), and $\times 47$ (an extreme upper stress-test bound; the corrected Filter IS MC median is ≈ 8 and the multi-method offsets range up to $\sim$120, Appendix A Fig. A1c, so the $\times 47$ case probes robustness well beyond the plausible Filter IS offset). PINE rows are held at $\times 1$ in all cases. Results are tabulated in Table S9 and shown in Fig. S7.

Three points stand out. (i) The adopted $\times 11.2$ harmonisation yields a well-constrained negative slope (-0.22, HAC 95 % CI $[-0.35, -0.09]$). (ii) Across the full range from $\times 1$ (CF-naive) to $\times 47$ the OLS slope varies monotonically from -0.355 to -0.144, reflecting the mechanical effect of the one-sided rescaling on the cross-method linear fit; the CF-naive endpoint ($\times 1$) has the steepest slope but the widest HAC interval ($[-0.67, -0.04]$). (iii) Crucially, the Spearman rank correlation ρ remains in the narrow range -0.46 to -0.58 across all five CF values (Spearman $p \leq 3 \times 10^{-4}$ in every case), confirming that the monotonic SIE/INP association is robust to the choice of harmonisation factor; only the magnitude of the linear slope is CF-sensitive.

Table S9. Sensitivity of the SIE–INP regression results (Sect. 3.6) to the choice of Filter-imm multiplicative CF. PINE rows are held at $\times 1$. $n = 56$ in all rows. Slope inference uses Newey–West HAC ($L = 3$) 95 % confidence intervals (a CI excluding zero indicates significance at the 5 % level); the Spearman rank correlation is reported as the harmonisation-robust association measure.

CF	OLS slope	HAC 95 % CI	Spearman ρ (p)
$\times 1$ (CF-naive)	-0.355	$[-0.67, -0.04]$	-0.58 (2.4×10^{-6})
$\times 11.2$ (main text)	-0.223	$[-0.36, -0.09]$	-0.57 (4.7×10^{-6})
$\times 16$	-0.203	$[-0.32, -0.09]$	-0.57 ($< 10^{-4}$)
$\times 28$	-0.173	$[-0.28, -0.07]$	-0.54 ($< 10^{-4}$)
$\times 47$	-0.144	$[-0.25, -0.03]$	-0.46 (3×10^{-4})

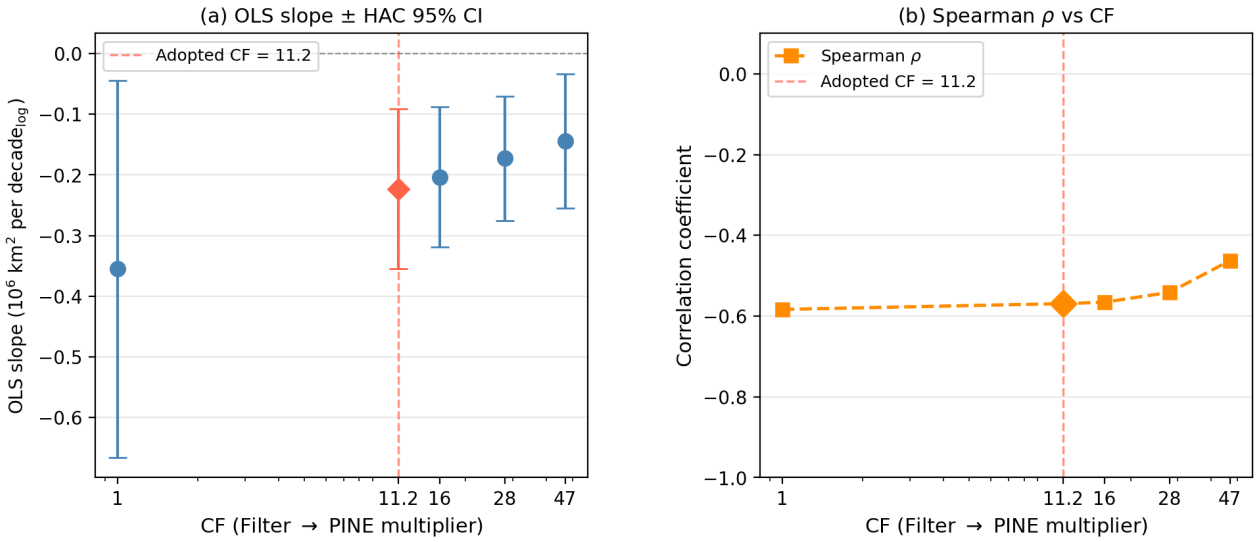

Figure S7. Sensitivity of the SIE–INP regression (Sect. 3.6) to the choice of Filter-imm multiplicative CF. (a) OLS slope ± Newey–West HAC 95 % CI as a function of CF. (b) Spearman ρ as a function of CF. The dashed vertical line marks the adopted main-text harmonisation factor ($\times 11.2$). The OLS slope is sensitive to CF (because applying CF to the offline subset shifts the cross-method intercept), whereas the Spearman rank correlation ρ remains in the narrow band -0.46 to -0.58 across all tested CF values, indicating that the monotonic SIE/INP association is robust to harmonisation-factor choice.

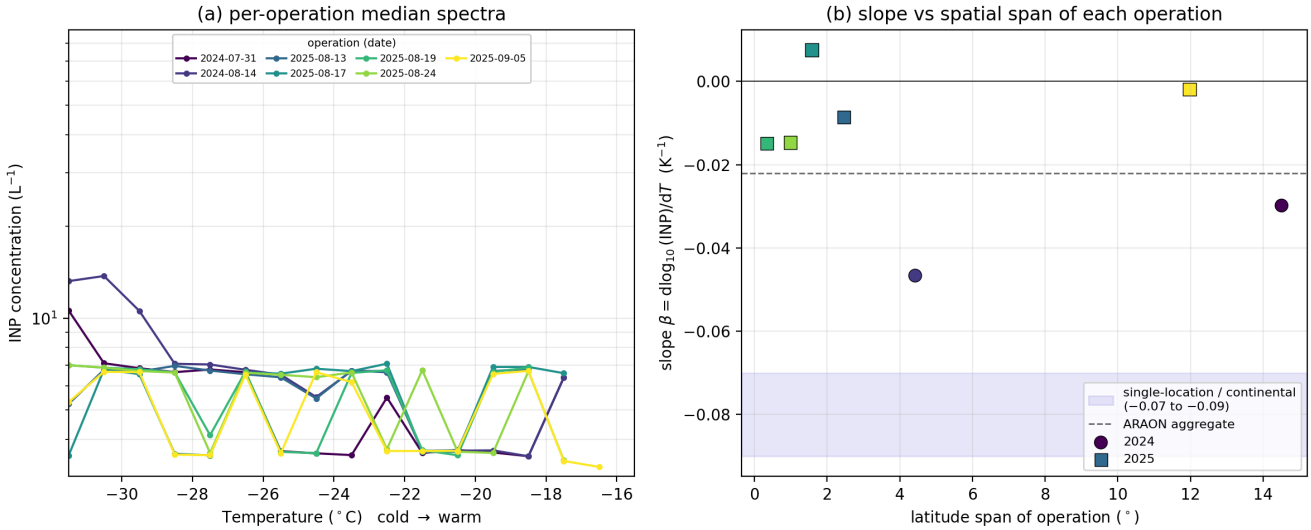

Figure S8. Per-operation breakdown of the Arctic INP temperature slope, supporting the intrinsic-flatness interpretation of Sect. 4.1. (a) Per-operation bin-median temperature spectra (1 °C bins, $n \geq 5$); each line is one shipborne operation. (b) Per-operation slope $\beta = \mathrm{d}\log_{10}(\mathrm{INP})/\mathrm{d}T$ (common window -31 to $-18\,^{\circ}\mathrm{C}$) versus the latitude span of each operation, coloured by year (circles: 2024; squares: 2025). The shaded band marks the -0.07 to $-0.09\,\mathrm{K}^{-1}$ slope of single-location continental and Arctic records; the dashed line is the ARAON campaign-aggregate slope ($-0.022\,\mathrm{K}^{-1}$). Every operation — including those spanning less than 1° of latitude, which sample an approximately single air and water mass — is flatter than the single-location band, and the slope shows no tendency to steepen as the spatial span narrows. This indicates that the flat aggregate slope is intrinsic to the marine Arctic boundary layer rather than an artefact of aggregating, along the cruise track, populations with differing onset temperatures.

15 Input dataset reproducibility metadata

Table S10 compiles the input datasets contributing to the multi-dataset sea-ice/INP synthesis (Sect. 3.6) and the Appendix A correction-factor analysis. For each dataset, the table lists the source reference, observation period, location, instrument/method, temperature range, aggregation level, monthly n, censoring rule, and the processing scripts that produced the merged synthesis panel.

Table S10. Input dataset metadata for reproducibility of the Appendix A correction-factor analysis and the Sect. 3.6 sea-ice/INP synthesis. The Filter-imm harmonisation factor $\times 11.2$ is recommended in Appendix A but is *not* applied to the Sect. 3.6 regression; sensitivity to CF choice is documented in Fig. S7 and Table S9 above.

Dataset	Period & Location	Instrument / Method / Notes
ARAON_2024	Aug 2024; ARAON cruise track (65–90° N)	PINE chamber, online condensation-freezing; T = -10 to $-30\,^{\circ}$C; monthly median, $n_{\mathrm{months}} = 1$; LOD applied; ship-contamination QC applied; This study
ARAON_2025	Jul–Sep 2025; ARAON cruise track (65–90° N)	PINE chamber; monthly median, $n_{\mathrm{months}} = 3$; LOD applied; ship-contamination QC applied; This study
Hartmann2020 (Filter-imm)	May–Jul 2017; Ny-Ålesund, Svalbard (78.9° N)	Filter-based immersion-freezing droplet assay; T = -7 to $-25\,^{\circ}$C; monthly aggregated, $n_{\mathrm{months}} = 3$; Hartmann et al. (2021)
Sze2023 (Filter-imm)	Jul 2018–Sep 2020; Villum, Greenland (81.6° N)	Filter-based immersion-freezing droplet assay; T = -5 to $-25\,^{\circ}$C; monthly aggregated, $n_{\mathrm{months}} = 27$; Sze et al. (2023)
Wex2019 (Alert)	Apr 2015–Apr 2016; Alert, Canadian Arctic (82.5° N)	Filter-based immersion-freezing droplet assay; monthly aggregated, $n_{\mathrm{months}} = 7$; Wex et al. (2019)
Wex2019 (Ny-Ålesund)	Mar–Aug 2012; Ny-Ålesund, Svalbard (78.9° N)	Filter-based immersion-freezing droplet assay; monthly aggregated, $n_{\mathrm{months}} = 5$; Wex et al. (2019)
Wex2019 (Utqiagvik)	Nov 2012–Mar 2013; Utqiagvik, Alaska (71.3° N)	Filter-based immersion-freezing droplet assay; monthly aggregated, $n_{\mathrm{months}} = 5$; Wex et al. (2019)
Wex2019 (Villum)	Dec 2013–Dec 2015; Villum, Greenland (81.6° N)	Filter-based immersion-freezing droplet assay; monthly aggregated, $n_{\mathrm{months}} = 5$; Wex et al. (2019)
NSIDC SIE	Oct 1978–present (matched to INP months above)	Daily Sea Ice Index (NSIDC), monthly means and 1981–2010 climatology; $n_{\mathrm{matched\,months}} = 56$; National Snow and Ice Data Center (NSIDC) (2024, 2025)

Processing pipeline. The merged monthly panel was assembled by (i) LOD- and contamination-filtering each PINE campaign at the run level, (ii) aggregating to monthly medians, (iii) harvesting published Filter-imm monthly time series at the $T = -20\,^{\circ}$C reference temperature, (iv) pairing each INP month with the matching NSIDC monthly SIE and deseasonalised SIE anomaly (1981–2010 climatology), and (v) regressing $\log_{10}(\mathrm{INP})$ on SIE anomaly with Newey–West HAC inference at lag

$L = \lfloor 4(n/100)^{2/9} \rfloor = 3$. Processing scripts are archived alongside the data deposit cited under "Data availability" in the main text.

References

Barry, K. R.: Ice nucleating particles in the Arctic: measurement and source tracking, PhD dissertation, Colorado State University, Fort Collins, CO, USA, https://hdl.handle.net/10217/237248, available at Mountain Scholar (Repository ID: islandora:1010323), Ch. 2. Defence: 2023-11-06; advisors: S. Kreidenweis, P. J. DeMott., 2024.

Böhmländer, A., Lacher, L., Fösig, R., Büttner, N., Nadolny, J., Brus, D., Doulgeris, K.-M., and Möhler, O.: PaCE-2022 PINE ice-nucleating particle measurements at Sammaltunturi, Finland, Earth Syst. Sci. Data, 17, 6165–6185, https://doi.org/10.5194/essd-17-6165-2025, underlying dataset archived at Zenodo, doi:10.5281/zenodo.16882069, 2025.

Creamean, J. M., Kirpes, R. M., Pratt, K. A., Spada, N. J., Maahn, M., de Boer, G., Schnell, R. C., and China, S.: Marine and terrestrial influences on ice nucleating particles during continuous springtime measurements in an Arctic oilfield location, Atmos. Chem. Phys., 18, 18 023–18 042, https://doi.org/10.5194/acp-18-18023-2018, key retained as "Creamean2019" for citation continuity; the paper is the 2018 ACP Oliktok-Point campaign., 2018.

Creamean, J. M., Barry, K., Hill, T. C. J., Hume, C., DeMott, P. J., Shupe, M. D., Dahlke, S., Willmes, S., Schmale, J., Beck, I., Hoppe, C. J. M., Fong, A., Chamberlain, E., Bowman, J., Scharien, R., and Persson, O.: Annual cycle observations of aerosols capable of ice formation in central Arctic clouds, Nat. Commun., 13, 3383, https://doi.org/10.1038/s41467-022-31182-x, 2022.

Hartmann, M., Gong, X., Kecorius, S., van Pinxteren, M., Vogl, T., Welti, A., Wex, H., Zeppenfeld, S., Herrmann, H., Wiedensohler, A., and Stratmann, F.: Terrestrial or marine – indications towards the origin of ice-nucleating particles during melt season in the European Arctic up to 83.7°N, Atmos. Chem. Phys., 21, 11 613–11 636, https://doi.org/10.5194/acp-21-11613-2021, 2021.

Irish, V. E., Hanna, S. J., Yu, X., Boyer, M., Polishchuk, E., Ahmed, M., Chen, J., Abbatt, J. P. D., Gosselin, M., and Chang, R. Y.-W.: Revisiting properties and concentrations of ice-nucleating particles in the sea surface microlayer and bulk seawater in the Canadian Arctic during summer, Atmos. Chem. Phys., 19, 7775–7787, https://doi.org/10.5194/acp-19-7775-2019, 2019.

Möhler, O., Adams, M., Lacher, L., Vogel, F., Nadolny, J., Ullrich, R., Boffo, C., Pfeuffer, T., Hobl, A., Weiß, M., Vepuri, H. S. K., Hiranuma, N., and Murray, B. J.: The portable ice nucleation experiment (PINE): a new online instrument for laboratory studies and automated long-term field observations of ice-nucleating particles, Atmos. Meas. Tech., 14, 1143–1166, https://doi.org/10.5194/amt-14-1143-2021, 2021.

National Snow and Ice Data Center (NSIDC): Sea Ice Index, Version 3, boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. https://doi.org/10.7265/N5K072F8, 2024.

National Snow and Ice Data Center (NSIDC): Sea Ice Index, Version 3, boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. https://doi.org/10.7265/N5K072F8, 2025.

Newey, W. K. and West, K. D.: A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica, 55, 703–708, https://doi.org/10.2307/1913610, 1987.

Pereira Freitas, G., Adachi, K., Conen, F., Heslin-Rees, D., Krejci, R., Tobo, Y., Yttri, K. E., and Zieger, P.: Regionally sourced bioaerosols drive high-temperature ice nucleating particles in the Arctic, Nat. Commun., 14, 5997, https://doi.org/10.1038/s41467-023-41696-7, 2023.

Sze, K. C. H., 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, Atmos. Chem. Phys., 23, 4741–4761, https://doi.org/10.5194/acp-23-4741-2023, 2023.

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 concentrations at different Arctic locations, Atmos. Chem. Phys., 19, 5293–5311, https://doi.org/10.5194/acp-19-5293-2019, 2019.

Wilbourn, E. K., Lacher, L., Guerrero, C., Vepuri, H. S. K., 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, Atmos. Chem. Phys., 24, 5433–5456, https://doi.org/10.5194/acp-24-5433-2024, 2024.