

Author Reply to Reviewer Comments

RC2 Reply

Dear Referee,

We thank you for the thoughtful and constructive review. Your comments address fundamental questions about the scope, framing, and interpretation of our study, and have prompted us to substantially improve both the scientific rigour and the clarity of the manuscript. We address each comment below in the order presented.

Where relevant, we note how your comments relate to changes that are also addressed in the companion RC1 reply, to ensure full consistency across both replies.

Major points

Screening of variables: methodology

Reviewer comment

(1) While you mention their names around l. 129, any more methodological information is absent. Please elaborate on that in a separate methodology section.

Response

We agree that the ML methodology required more explicit documentation, and this comment is consistent with feedback from Reviewer 1 (see the companion RC1 reply). We have added a dedicated paragraph to Sect. 2.2 that describes the algorithms applied (random forest and decision tree models for feature importance ranking; principal component analysis and K-means clustering for dimensionality reduction), clarifies that no train/test data split was performed (the purpose was exploratory importance ranking), and explains the time harmonization procedure. We have also added a closing sentence to Sect. 2.2 restating that these methods were applied as exploratory tools, not to build or evaluate a predictive model.

Change in manuscript

See the companion RC1 reply for the corresponding manuscript change (modifications to Sect. 2.2).

Reviewer comment

(2) Where can we see that the separate treatments resulted in qualitatively similar results (l. 131 and l. 159)? Fig. 2 only shows one of these, and a comparison should at least be part of the Appendix.

Response

The referee suggests a good point. However, although qualitatively similar, the separate treatments were not robust and small changes, for example in the random seed were observed to lead to differences. Thus, presenting any fixed set of results for comparison would be somewhat arbitrary. With this in mind we prefer to present only one “qualitative” example, while we try to emphasize the *qualitative* nature of the result.

Duplicate text describing the variable exclusion criteria (Sect. 3.2, l. 132–134) has been removed to reduce repetition. Additionally, we modify the text at l. 131 to read “which are illustrated in Fig. 2” (replacing “as illustrated”), making clear that Fig. 2 shows the results from one representative method (random forest). We add a sentence acknowledging that a formal multi-method comparison figure would arbitrarily relate results from the methodologies in an artificial way.

Change in manuscript

Changed “as illustrated in Fig. 2” to “which are illustrated in Fig. 2”. Added: “*A formal multi-method comparison (e.g., side-by-side feature rankings from random forest, decision tree, and PCA) is not shown but was verified by the authors during analysis; these approaches consistently highlighted similar top-ranking variables.*”

Reviewer comment

(3) How was the subset (of the 84 variables that you analysed) selected for Fig. 2? Are these the ones with the highest importance?

Response

Yes. Figure 2 shows the top-ranked variables by feature importance from the random forest analysis, ordered by decreasing importance score. We have clarified this in the updated Fig. 2 caption.

Change in manuscript

Added to the Fig. 2 caption: “*Variables are shown in decreasing order of random-forest importance. The six variables examined in pairwise correlation in Fig. 3—fluorescent particle number concentration (WIBS), particle mass (1 nm–10 μ m aerodynamic diameter, AMS), number concentration of particles with diameters >0.5 μ m (APS), organic aerosol mass concentration (AMS), nitrate aerosol mass concentration (AMS), and black carbon mass concentration (aethalometer)—appear among the top-ranked variables in this panel (selection criteria in Sect. 2.2).*”

Uptake of screened variables

Reviewer comment

From the variables in Fig. 2, you make another selection for Fig. 3. (l. 144) states that these were selected “based upon their highly ranked outcomes and physical intuition.” On a first glance, this appears a reasonable and often-used approach, but in the context of your study it seems to invalidate the idea that the study is based on: screening as many variables as possible. If you exclude those variables that you didn’t expect to show a correlation (physical intuition) after the screening, why include them in the first place?

Many of the variables that appear with high importance in Fig. 2 also remain unmentioned — I think at the least you ought to explain why you exclude them from further analysis (perhaps in a Table?).

Response

We appreciate this critical observation. The referee correctly identifies a tension between the open screening philosophy and the subsequent use of physical intuition as a filter. We wish to clarify that the screening was genuinely open-ended: all 84 variables entered the random forest analysis without prior filtering by physical expectation. The subsequent selection of six variables for detailed scatter-plot examination in Fig. 3 does not invalidate the screening; rather, it reflects the practical constraint that a detailed correlation analysis with visualisation is feasible only for a limited number of variables. The screening’s primary value is in revealing which variables rank highly, including *unexpected* ones (e.g., black carbon, nitrate), which we do investigate and discuss.

We have replaced the ambiguous “physical intuition” phrasing with an explicit dual-selection justification: *“The six variables shown in Fig. 3 were selected on two complementary grounds: (i) high importance rank in the random forest analysis (fluorescent particles, black carbon, nitrate, particle mass); and (ii) established physical or empirical connection to INP in prior literature or at this specific site.”* This makes the rationale transparent.

Regarding other high-ranking variables, we add a statement acknowledging that other highly ranked variables (e.g., acetone, methanol) largely co-vary with the selected variables (particularly nitrate) and thus provide redundant information. Our selection therefore covers the principal top-ranked predictors after accounting for redundancy among co-varying variables, and it is unlikely that lower-ranked variables would yield additional insight.

Change in manuscript

- The dual-selection justification is added (see the companion RC1 reply, “Line 144” response).
 - Added: *“Several other highly ranked variables (e.g., acetone, methanol concentrations) are not examined individually because they largely co-vary with the selected predictors (particularly nitrate) and thus provide redundant information for the purposes of this exploratory analysis.”*
-

Title and framing

Reviewer comment

The title suggests that you used machine learning extensively and that you are predicting INP properties. In my reading, this does not reflect what you did and what is explained in the manuscript:

- the ML was only used to screen variables and not documented well (see above)
- your modified parameterisation shows only modest skill, and only for spring measurements. It is also only a small part of the study.
- The only thing you are predicting are INP concentrations, thus only one INP property.

This can easily be read as overselling [...] As you state in the conclusion, “no single parameter emerges that is strongly linked to INP.” The paper title and framing need to reflect this.

Response

We agree with the referee’s assessment (see also the companion RC1 reply). We change the title from “Predicting” to “Exploring” and add explicit statements that the ML methods were used as importance-ranking tools, not to build a predictive model.

However, we acknowledge that the referee’s concern goes further: even “Exploring” may suggest that ML is the centerpiece of the study, whereas the main scientific message is the *null result* — that no single parameter strongly predicts INP concentrations, even with extensive co-located measurements.

Therefore, we further refine the title to make plain the null result, and finally suggest: “*Exploring Ice Nucleation Particle concentrations in a Boreal Environment: limits of machine-learning-assisted variable screening.*”

This addresses the referee’s concern by:

1. Replacing “properties” with “concentrations” (only one INP property is examined).
2. Adding “limits of . . . variable screening” to signal the null result directly in the title.
3. Retaining “Exploring” inspired by the RC1 comments.

We have also replaced “INP variability” with “INP concentrations” throughout the abstract and text where appropriate, following the referee’s observation that variability per se is not analyzed.

Change in manuscript

- Title: changed from “*Exploring Ice Nucleation Particle properties in a Boreal Environment using machine learning*” to “*Exploring Ice Nucleation Particle concentrations in a Boreal Environment: limits of machine-learning-assisted variable screening.*”
 - “INP variability” replaced with “INP concentrations” in the abstract and where applicable.
-

Conclusion from the null results

Reviewer comment

As I stated above, I think the “null result” is an important one. However, the conclusion jumps from “drawing strong conclusions that can illuminate causality will likely remain illusive” to suggesting more, longer and heavily-equipped measurements. [...] given your sobering findings, wouldn’t the opposite make more sense? More and more measurements will NOT magically “help the community to identify key predictor parameters”. Why do you think you couldn’t find a correlation? What does this imply about the nature of INPs? What could one/the field do differently? [...]

Response

We thank the referee for this challenging and thought-provoking comment. The referee is right that the standard call for “more measurements” is insufficient — and potentially misleading — in the light of our null result. We have given this careful consideration.

The weak correlations found in our study, combined with the log-normal INP concentration distributions (consistent with random mixing and dilution of trace species; Ott, 1990), strongly suggest that INP at this site are dominated by long-range transport of aerosol from diverse, distant sources. In such a scenario, no local measurement suite — however comprehensive — can be expected to yield strong, causal predictor–INP relationships, because the INP identity and source vary stochastically with air-mass origin rather than responding to locally measured tracers. This interpretation is independently supported by Paramonov et al. (2020), who used the same PINC instrument and dataset at the same site and concluded: *“No single dominant local or regional source of INPs in the boreal environment of southern Finland could be identified. Rather, it is hypothesised that the INPs detected at SMEAR II are a result of long-range transport and dilution of INPs sourced far from the measurement site.”*

We have added a new paragraph to the conclusions that directly engages with the referee’s questions:

1. We acknowledge that the null result suggests a fundamental limitation: if INP are dominated by long-range transport, then local correlations will always be weak.
2. We propose that future campaigns should *complement* high-frequency INP measurements with source-apportionment tools (back-trajectory analysis, receptor modelling) to evaluate consistency with the transport hypothesis and constrain plausible source regions—rather than simply accumulating more of the same local measurements.
3. We note that the question “why can’t we predict INP?” is itself a scientifically valuable finding that deserves explicit investigation.

We have preserved the existing recommendation for continued measurements at equipped stations (see also the companion RC1 reply) but placed it in the context of complementary analytical approaches.

Change in manuscript

The following paragraph is added to the conclusions:

“The weak correlations observed in this study, combined with the log-normal INP concentration distributions consistent with random dilution of trace species, suggest that INP at this boreal site are dominated by long-range transport from diverse, distant sources.”

In such a regime, no local measurement suite—however comprehensive—can be expected to yield strong, causal predictor–INP relationships, because the INP identity and source vary stochastically with air-mass origin. This interpretation is independently supported by Paramonov et al. (2020), who reached a similar conclusion using the same PINC dataset at the same site. We therefore propose that future campaigns should complement high-frequency INP measurements with source-apportionment tools (e.g., back-trajectory analysis, receptor modelling) to evaluate consistency with the long-range transport hypothesis and to constrain the geographic origins and aerosol types that plausibly contribute to the INP population—recognising that trajectory-based tools alone cannot uniquely verify INP provenance without tracer or receptor constraints. The question ‘why can’t we predict INP from local measurements?’ is itself a scientifically valuable finding: it constrains the problem space and guides the field toward approaches that account for air-mass history rather than relying solely on local co-located tracers.”

New parameterization does not respond to predictors

Reviewer comment

l. 222: “PINC measurements . . . have almost no response to the predictors”: it’s the predicted INP that don’t vary, while the measured INP do. However, I don’t understand how this can be the case? Seeing the simple parameterisation formula, it implies that the predictors themselves don’t vary. Is that correct? Fig. 1 implies otherwise, at least the particle mass is varying quite a bit there, so I do not understand why a parameterisation based on it would give constant results.

Response

We thank the referee for identifying this confusing passage. The issue is a combination of the wording and the mathematical properties of the fitted power law.

The predictors (e.g., particle mass) *do* vary during the PINC winter period, as visible in Fig. 1. However, the fitted exponent j in Eq. (3) for the PINC data is near zero or slightly negative for all six predictors (Table 1): particle mass: $j = -0.10$; organics: $j = 0.00$; fluorescent particles: $j = -0.05$; BC: $j = 0.03$, etc. In a power-law relationship $n_{\text{INP}} = i \times X^j$, an exponent near zero means $X^j \approx 1$ regardless of the value of X , so the predicted INP concentration collapses to approximately the pre-factor i (i.e., the mean INP concentration), producing a near-constant prediction despite substantial predictor variability.

This is a direct consequence of the absence of a meaningful statistical relationship between the predictor and INP during winter: the best-fit power law effectively “gives up” by setting the exponent to zero, reverting to the sample mean. This is consistent with the negative adjusted R^2 values reported in Table 1 for all PINC predictors.

We have clarified the wording and added an explanatory sentence.

Change in manuscript

- Changed “*have almost no response to the predictors*” to “*yield near-constant predicted INP concentrations because the fitted exponents (j in Eq. 3) are near zero for all PINC predictors*”

(Table 1), meaning the predicted values collapse to approximately the pre-factor i regardless of predictor variability.”

- Added an explanatory paragraph to the table discussion: “The exponent j reflects the sensitivity of INP concentration to changes in the predictor: values near unity indicate approximately linear relationships (...), values well below unity suggest weak sensitivity (...), and near-zero or negative values indicate absence of a meaningful relationship (all PINC predictors). The consistently low or negative adjusted R^2 for PINC confirms the absence of predictive skill during the winter period, regardless of the predictor chosen.”
-

Minor points

Abstract: 84 vs. 500 variables (l. 7)

Reviewer comment

[...] many of the variables that could have potentially been explored were excluded due to excessive NaN values or little variability, which left you with 84 variables. I think this is the number that should be mentioned in the abstract, rather than a grand 500 variables (l. 7).

Response

We agree that the abstract should be precise about the number of variables actually analysed. We now mention both numbers in the abstract: the initial scope (more than 500 variables monitored at SMEAR II) and the analysis set (84 retained after quality screening). This preserves the context about data richness while being transparent about the analysis scope.

Change in manuscript

Changed abstract from “...more than 500 high-resolution atmospheric, aerosol, and ecosystem variables measured continuously at Station for Measuring Ecosystem-Atmosphere Relations (SMEAR) II.” to “...more than 500 high-resolution atmospheric, aerosol, and ecosystem variables measured continuously at the Station for Measuring Ecosystem-Atmosphere Relations (SMEAR) II, of which 84 were retained after quality screening for the analysis.”

“INP concentrations” vs. “INP variability” (l. 7 and throughout)

Reviewer comment

Aren't INP concentrations what you're after? I don't see you using a measure of variability anywhere, nor do you correlate with variability itself.

Response

The referee is correct. Our analysis correlates INP concentrations (not their variability per se) with predictor variables. We have replaced “INP variability” with “INP concentrations” in the abstract and throughout the text where pertinent. See also the companion RC1 reply.

Change in manuscript

Replaced “INP variability” with “INP concentrations” in the abstract and in relevant passages throughout the manuscript.

Introduction

Line 18

Reviewer comment

They don’t need to “form” below 0°C, the cloud forms as the cloud droplet forms.

Response

We agree that the original phrasing was imprecise. Mixed-phase clouds are defined by the co-existence of liquid water and ice at temperatures below 0°C; the cloud itself forms when droplets activate on CCN, and ice subsequently appears via heterogeneous nucleation. We have corrected the wording.

Change in manuscript

Changed “*for ice to exist, the temperature must be less than 0°C, meaning that at these temperatures liquid water would prefer to be ice*” to “*for ice crystals to form via heterogeneous nucleation, temperatures below 0°C are required; at such temperatures, ice is the thermodynamically stable phase, yet supercooled liquid water persists because a high kinetic barrier inhibits spontaneous freezing.*” This also addresses the “school language” and “missing link” concerns below.

Line 20: “would prefer to be ice” and missing link

Reviewer comment

“Would prefer to be ice” — “school language”. Also, the link between the sentences is missing: water would “prefer to be ice”, yet it is not!

Response

The sentence containing “unstable co-existence of ice and liquid water” is replaced with a more precise thermodynamic description (see also the companion RC1 reply) (“thermodynamically metastable co-existence of ice and supercooled liquid water ...”). Moreover, we further refine the passage to remove the informal language, to provide the missing logical link (why liquid persists despite thermodynamic instability), and to integrate the explanation into a single coherent sentence.

Change in manuscript

The passage is now: “*Even with this inherently thermodynamically metastable co-existence of ice and supercooled liquid water – ice is the stable phase below 0°C, but a high kinetic barrier to nucleation sustains liquid water in the absence of efficient INPs – mixed-phase clouds are widespread ...*”.

Line 23: “underlying”**Reviewer comment**

What is meant with “underlying”?

Response

The revision specifies the feedbacks and removes the ambiguous “underlying” (see also the companion RC1 reply). “especially in the Arctic where underlying feedbacks have amplifying effects” is replaced with “*where amplifying feedbacks—such as the ice-albedo feedback and the water vapour–temperature amplification characteristic of polar warming—have particularly pronounced effects.*”

Change in manuscript

See the companion RC1 reply for the corresponding manuscript change.

Line 24: “only”**Reviewer comment**

“Only” — ice formation or ice in MPCs can also occur because of ice sedimenting into the cloud from another cloud above.

Response

We agree. Ice crystals can also enter a mixed-phase cloud via sedimentation from overlying ice clouds (e.g., “seeder–feeder” mechanism). We have added a brief parenthetical acknowledging this pathway.

Change in manuscript

Changed “*mixed-phase clouds only appear with the help of small particles*” to “*mixed-phase clouds primarily form with the help of small particles in the atmospheric aerosol*” and added a parenthetical: “*(ice can also enter a cloud via sedimentation from above, e.g., the seeder–feeder mechanism)*”.

Line 34: ICOS, ACTRIS, SITES citations and acronyms**Reviewer comment**

Give citations for ICOS, ACTRIS and SITES (and explain the acronyms).

Response

We have added the full names and appropriate citations for all three infrastructure networks.

Change in manuscript

Changed “e.g., ICOS, ACTRIS, SITES, etc.” to “e.g., the Integrated Carbon Observation System (ICOS; Heiskanen et al., 2022), the Aerosol, Clouds and Trace Gases Research Infrastructure (ACTRIS; Pandolfi et al., 2018), and the Swedish Infrastructure for Ecosystem Science (SITES; 2021).”

Line 35: take-aways from previous campaign studies

Reviewer comment

What were the most important take-aways (about INPs) from these other campaign studies?

Response

We have expanded our connections to all HyICE publications, which include Paramonov et al. (2020), Schneider et al. (2021), Brasseur et al. (2022, 2024), and Vogel et al. (2024). We add a sentence summarizing the key findings from these prior publications.

Change in manuscript

Added after the list of HyICE citations: “*Key findings include the characterization of condensation/immersion mode INP concentrations and their likely origin from long-range transport (Paramonov et al., 2020), the identification of seasonal trends linked to biogenic emissions (Schneider et al., 2021), a complete campaign overview and INP instrument intercomparison (Brasseur et al., 2022), a description of flight-based measurements that include vertically resolved INP concentrations over the boreal forest (Brasseur et al., 2024), and the role of fluorescent biological aerosol particles as INP tracers (Vogel et al., 2024).*”

Line 40: “INP data” — be specific

Reviewer comment

“INP data”: as for “variability” above, be specific.

Response

Consistent with the change from “ice nucleation activity” to “INP concentration” and with the referee’s earlier comment about “variability,” we have changed “INP data” to “INP concentrations.”

Change in manuscript

Changed “*INP data*” to “*INP concentrations*”.

Line 41: “even” → “especially” and spurious correlations

Reviewer comment

“Even”: I would argue that “especially” fits better, because the more variables you try (and the less physical intuition you have for them, see above), the more likely it is that you find spurious correlations. This is a wider issue that I think could be addressed or at least mentioned somewhere in the manuscript.

Response

The referee makes an excellent point. With 84 variables and limited sample sizes, the risk of finding spurious correlations is real. Paramonov et al. (2020) faced the same concern from their reviewers when correlating INP with multiple aerosol properties at the same site, and addressed it by adding explicit caveats (e.g., “correlation of [INP] with fluorescent particle concentration does not imply that INPs are necessarily fluorescent”). We have replaced “even” with “especially” and added a sentence about the multiple-comparison problem.

Change in manuscript

- Changed “even” to “especially.”
 - Added: *“When many variables are screened simultaneously, the probability of finding spurious correlations increases; the results presented here should therefore be interpreted as hypothesis-generating rather than hypothesis-confirming, and the identified associations require independent validation.”*
-

Missing: physical intuition background in introduction

Reviewer comment

Since you draw on “physical intuition” to argue for the variables that you include in Fig. 3, I miss background on that intuition in the introduction. What have INPs been correlated with before?

Response

We augment the introduction to provide additional context for the reader with respect to aerosol properties that have previously been linked to INP concentrations. This background motivates both the screening approach and the subsequent variable selection. We have added a brief paragraph to the introduction.

Change in manuscript

Added text: *“Previous studies have linked INP concentrations to a range of aerosol properties, including total particle number concentration above 0.5 μm (DeMott et al., 2010; Tobo et al., 2013), fluorescent biological aerosol particle concentrations (Tobo et al., 2013), mineral dust loading (DeMott et al., 2015), and black carbon (DeMott, 1990). At the same site, Paramonov et al. (2020) found that no single parameter predicted INP concentrations over the full campaign period, although short-timescale correlations with black carbon, supermicron biological particles, and sub-0.1 μm particles*

were observed. These precedents motivate the present study’s open-ended screening approach while providing the physical context for interpreting the resulting variable rankings.”

Methods and Data

Line 50: “two” but list 3

Reviewer comment

You say “two” but list 3.

Response

We thank the referee for this close read. The numeral “two” is correct: PINC and PINCii are the two continuous-flow diffusion chambers (CFDCs) deployed during HyICE-2018. The abstract wording “two . . . chambers, Portable Ice Nucleation Chamber I and II (PINC and PINCii)” could be parsed as three separate items (chamber type, full names, abbreviations) rather than as two instrument names; we have revised the sentence to make the pairing explicit.

Change in manuscript

Abstract: replaced the ambiguous appositive with “*Two continuous-flow diffusion chambers (CFDCs), PINC and PINCii (Portable Ice Nucleation Chambers I and II), were deployed with high-frequency sampling to measure INP concentrations*”, so “two” unambiguously refers to the two instruments. The Methods text already states that “*two continuous flow diffusion chambers (CFDCs; PINC . . . and PINCii . . .) . . . were operated*” (Sect. 2.1); no change was required there.

Lines 53–58 and 73–74: repetition

Reviewer comment

No need to repeat the introduction or move this content there if it hasn’t been covered.

Response

We agree that these passages duplicate introductory material. We have consolidated the text by removing the repetition from the Methods section and ensuring the relevant content appears only once.

Change in manuscript

Removed duplicate introductory material from Sect. 2.2 (the first two sentences restating the motivation for co-locating measurements at SMEAR II) and from Sect. 2.3 (the opening sentence restating the study’s overarching objective). The relevant scientific context is retained in the Introduction.

Line 59: singular “purpose”

Reviewer comment

Singular purpose.

Response

Corrected.

Change in manuscript

Changed “*purposes*” to “*purpose*”.

Line 68: instrument documentation

Reviewer comment

In principle I think all data that you use should be described with this much information, or else please describe why you describe only this one and where information on the others can be found.

Response

We have added a sentence to Sect. 2.2 directing readers to Brasseur et al. (2022) for HyICE-2018 instrument specifications and to the SmartSMEAR portal for the full SMEAR II monitoring suite metadata (see also the companion RC1 reply). Given that we use over 80 variables from dozens of instruments, a complete instrument table in the manuscript would be impractical; instead, we provide these comprehensive references.

Change in manuscript

See the companion RC1 reply for the corresponding manuscript change.

Line 104: main take-aways from comparison

Reviewer comment

Please state the main take-aways from the comparison.

Response

We have added a brief summary of the key findings from the PINC/PINCii comparison described in Brasseur et al. (2022).

Change in manuscript

Added: “*The comparison showed that PINC and PINCii agree within a factor of two during the overlapping measurement period, with both instruments capturing the same general trends in INP*”

concentrations; minor differences in absolute concentration are attributable to the 1 K difference in operating temperature ($T_l = -31^\circ\text{C}$ for PINC, -32°C for PINCii) and instrument-specific detection efficiencies.”

Lines 110–111: “slightly greater variability”

Reviewer comment

What do you mean with “slightly greater variability”? Can you quantify this, for example by giving a range or fitting a distribution?

Response

We did not intend a formal distribution fit or numeric summary (e.g., geometric standard deviation or interquartile range in concentration decades). The remark referred to a qualitative visual impression from Fig. 1(a): the PINCii spring and early summer segment shows a somewhat wider envelope on the logarithmic concentration axis than the PINC winter segment. Because PINC and PINCii did not operate simultaneously, we do not report paired distribution statistics between instruments here; such numbers would not resolve the referee’s question in a physically interpretable way without a dedicated intercomparison design. We have revised the sentence to state explicitly that the comparison is descriptive (visual inspection of Fig. 1(a)) and to link the pattern to expected seasonal differences in aerosol sources and meteorology, without implying that variability has been quantified.

Change in manuscript

Changed “*the PINCii measurements in spring and early summer appear to contain slightly greater variability*” to “*by visual inspection of Fig. 1(a), the PINCii spring and early summer time series appears somewhat more spread on the logarithmic concentration axis than the PINC winter series, consistent with seasonal changes in aerosol sources and meteorology; we do not report distribution fits or geometric moments here because PINC and PINCii did not operate simultaneously, so we restrict the comparison to this qualitative description.*”

Results

Figure 1 caption: PINCii squares and missing FBAP data

Reviewer comment

I don’t see squares for PINCii in (a)? Why are there no measurements of fluorescent biological aerosol particles before 03-15 in (c)?

Response

We have added measurement temperature information and corrected units in the Fig. 1 caption (see also the companion RC1 reply). Regarding the referee’s additional questions:

(1) PINC and PINCii data are distinguished by colour in panel (a): PINC is plotted in blue and PINCii in red. We have added a note in the caption clarifying the colour coding and measurement temperatures.

(2) The WIBS instrument (measuring fluorescent biological aerosol particles) was deployed starting 11 March 2018, approximately one month into the campaign. Prior to that date, no FBAP data are available. We have added an explanatory note to the caption.

Change in manuscript

Updated Fig. 1 caption: (1) Added temperature and colour clarification: “*INP concentrations were measured at $T_l = -31^\circ\text{C}$ (PINC; blue) and $T_l = -32^\circ\text{C}$ (PINCii; red).*”; (2) Added: “*WIBS fluorescence data begin on 11 March 2018, when the instrument was first deployed.*”

Lines 132–137: repetition

Reviewer comment

Repetition or move to Methods if it hasn't been said before.

Response

See the companion RC1 reply for the corresponding manuscript change.

Line 142: investigate surprising variables more closely

Reviewer comment

Related to the major point above, why don't you investigate the link between surprising variables and INP concentrations more closely? This would add value to your idea of sampling as many variables as possible.

Response

We agree that deeper investigation of surprising high-ranking variables is valuable, and in fact this influenced our choices throughout the manuscript. Perhaps, our most surprising finding is the high importance of black carbon (BC), which we do investigate and discuss in detail – including the independent support from Paramonov et al. (2020) who also found BC–INP correlation (their Fig. 5). We have added discussion of possible mechanisms (aging, oxidation, coating with organic material) and cited relevant literature.

Other (perhaps) surprising variables (e.g., acetone, methanol) largely co-vary with nitrate and are discussed as redundant air-mass tracers.

Change in manuscript

Added: “*A systematic investigation of all high-ranking but unexpected variables is an important avenue for future work.*”

Line 145: co-variance not stated as exclusion reason

Reviewer comment

You give co-variance as an example for exclusion, but do not state it as a reason for exclusion in the sentence prior.

Response

We have corrected this logical gap by adding co-variance as an explicit exclusion criterion in the sentence preceding the example.

Change in manuscript

The existing example (“acetone and methanol, which largely co-vary with nitrate, are not [examined]”) already demonstrates co-variance as an exclusion rationale. As addressed in the companion RC1 reply, the dual-selection justification now provides the explicit selection criteria (high random-forest importance and established physical/empirical connection), making co-variance exclusion implicit in the “redundant information” phrasing.

Figure 2: cross-reference to Fig. 3; variable specification

Reviewer comment

The comparison to Fig. 3 is difficult. Could you highlight the variables that you chose to investigate in detail in Fig. 3? Also, many of these variables need further specification, for example “nitrate” what? Number concentration?

Response

We address the comparison by spelling out, in the Fig. 2 caption, the same six measured quantities and instruments as in Fig. 3 (WIBS fluorescence; AMS particle mass, organics, and nitrate; APS number concentration $>0.5 \mu\text{m}$; aethalometer black carbon). Readers can therefore locate the corresponding bars by name and see explicitly whether “nitrate” refers to aerosol mass from the AMS, etc. The y-axis labels on the published figure file remain the short SMEAR II metadata names; full quantities are given in the caption.

Change in manuscript

Fig. 2 caption: added explicit listing of the six Fig. 3 variables with instrument and quantity: *“The six variables examined in pairwise correlation in Fig. 3—fluorescent particle number concentration (WIBS), particle mass (1 nm–10 μm aerodynamic diameter, AMS), number concentration of particles with diameters $>0.5 \mu\text{m}$ (APS), organic aerosol mass concentration (AMS), nitrate aerosol mass concentration (AMS), and black carbon mass concentration (aethalometer)—appear among the top-ranked variables in this panel (selection criteria in Sect. 2.2).”*

Figure 2 caption: decision tree explanation

Reviewer comment

The explanation of the decision tree belongs into the Methods section (and, as stated above, the other ML methods need to be explained there as well).

Response

We agree. A dedicated ML methodology paragraph is added to Sect. 2.2. We have moved the decision tree explanation from the Fig. 2 caption to this methods paragraph and simplified the Fig. 2 caption to reference Sect. 2.2 for methodological details.

Change in manuscript

Removed the Random Forest ensemble explanation from the Fig. 2 caption (that material is covered in Sect. 2.2). The sentence on bar height now reads “*Higher values indicate a stronger statistical association with INP concentrations*” (in place of the longer decision-tree wording), and the caption adds the cross-reference listing the six Fig. 3 variables with full specifications (as under comment (3) and the Fig. 2 comparison response above), without bold formatting in the figure.

Lines 152–156: repetitive

Reviewer comment

Repetitive of what’s been said just before.

Response

We have removed the repetitive passage.

Change in manuscript

Removed the repetitive text at l. 152–156.

Figure 3: PINC = winter, PINCii = spring

Reviewer comment

You may want to remind the reader that PINC corresponds to winter and PINCii to spring measurements.

Response

We have added extensive text explaining the data splitting (see also the companion RC1 reply). We also add a brief reminder directly in the Fig. 3 caption.

Change in manuscript

Added to Fig. 3 caption: “*PINC data (blue) correspond to winter measurements (19 February–2 April); PINC_{ii} data (red) to spring/summer (22 April–10 June).*”

Line 160: correlations \neq causation

Reviewer comment

Why would a method that shows you correlations be enough to “shed light on causation”? Please explain what made you reason that the correlations you find do not reflect causations.

Response

The referee is correct that our correlation-based approach cannot establish causation. The original phrasing was misleading. We have clarified that the ML and correlation analysis were used to identify statistical associations, not causal relationships. The reasons we believe these correlations are not causal include: (i) the well-known rarity of INPs relative to the total aerosol population, meaning that any observed bulk-aerosol correlation likely reflects co-varying air-mass properties rather than direct INP composition; (ii) the absence of single-particle analysis of ice crystal residuals in our setup, which would be needed to directly identify INP composition; and (iii) the Paramonov et al. (2020) finding at the same site that “an increase in the concentration of any of these particle species may not necessarily lead to the increase in the [INP]; the reasons for this remain unknown.”

Change in manuscript

- Changed “*they do in and of themselves not succeed in shedding light on causation or sources of INP*” to “*they identify statistical associations between predictor variables and INP concentrations but cannot establish causal relationships.*”
 - Added: “*The correlations likely reflect co-varying air-mass properties rather than direct INP composition: INPs are a vanishingly small fraction of the total aerosol (~ 1 in 10^6), and without single-particle analysis of ice crystal residuals, the identity of the actual ice-nucleating species remains undetermined (cf. Paramonov et al., 2020).*”
-

Line 167: log-normal distributions and long-range transport

Reviewer comment

Why do log-normal frequency distributions imply long-range transport? Is your point here that the absence of a correlation with a co-located variable may be due to the INPs stemming from long-range transport? Could you substantiate this hypothesis, for example with back trajectories? If this is your main hypothesis (which I do find interesting and important), please highlight it in the conclusions. Also, doesn’t this render your parameterisation attempts, which are following, less promising from the get go?

Response

The referee has precisely identified the key mechanistic argument. We address each sub-question:

Why log-normal → long-range transport: Log-normal concentration distributions arise naturally from the multiplicative dilution and mixing of an initially concentrated plume during atmospheric transport (Ott, 1990). The observation that INP concentrations at SMEAR II follow a log-normal distribution (Fig. 4) is therefore consistent with INPs originating from diverse, distant sources and arriving at the measurement site after multiple mixing and dilution events – an analogous interpretation was proposed by Paramonov et al. (2020) for the PINC data alone.

Can we substantiate with back trajectories? Lagrangian back trajectories and dispersion models (e.g., HYSPLIT, FLEXPART) estimate air-mass pathways and geographic context and are often used together with chemical tracers or receptor modelling in atmospheric studies. They do *not*, however, uniquely verify that INPs at the receptor are dominated by long-range sources or that a log-normal INP distribution reflects a specific dilution history: trajectory paths carry meteorological uncertainty, and unresolved mixing can decouple aerosol composition from kinematic history. Moreover, INPs are not tagged in standard trajectory products. Back trajectories therefore *complement*—rather than replace—inferences from co-located data; a substantive analysis coupling trajectories with tracers or receptors is a substantial undertaking that we propose as a priority for future work, and we state these limitations and priorities explicitly in the manuscript.

Highlight in conclusions: Yes — we have added a new paragraph to the conclusions that highlights the long-range transport hypothesis and proposes source-apportionment tools with wording that matches this nuance (see our response to the “Conclusion from null results” major point above).

Parameterisation attempts: The referee is correct that if INPs are dominated by long-range transport, local parameterisations based on co-located measurements will have limited predictive power. This is indeed what we observe for the PINC winter data. The modest success of the PINCii spring/summer parameterisations suggests that during the biologically active season, a fraction of the INP population *does* respond to local aerosol properties — consistent with a two-component picture (transported background + local biogenic contribution). We have made this logical connection explicit in the text.

Change in manuscript

- Added after the log-normal discussion: *“Log-normal concentration distributions arise naturally from the multiplicative dilution and mixing of aerosol during atmospheric transport (Ott, 1990), consistent with INPs at this site originating from diverse, distant sources. This interpretation aligns with the findings of Paramonov et al. (2020), who hypothesised long-range transport as the dominant INP source at SMEAR II based on the same PINC dataset.”*
- Added: *“Lagrangian back trajectories or dispersion models (e.g., HYSPLIT, FLEXPART) can contextualise air-mass history and are often used together with chemical tracers or receptor modelling to assess long-range transport, but they do not by themselves uniquely verify INP provenance or the dilution interpretation above: trajectory errors, mixing along the path, and the lack of INP-specific tagging in kinematic histories mean that such analyses complement—rather than uniquely test—inferences from local correlations and distribution shape. We highlight combined trajectory–tracer or trajectory–receptor studies as a priority for future work.”*
- Added logical connector before the parameterisation section: *“If the dominant INP signal originates from long-range transport, local parameterisations based on co-located measurements*

will have limited predictive power—a prediction directly borne out by the PINC winter results below. The modest success of the PINCii spring/summer parameterisations may then reflect a local biogenic INP contribution that augments the transported background during the growing season.”

- Conclusions paragraph (same addition as in “Conclusion from null results”): revised the sentence on source-apportionment tools to “*evaluate consistency with the long-range transport hypothesis and to constrain the geographic origins and aerosol types that plausibly contribute to the INP population—recognising that trajectory-based tools alone cannot uniquely verify INP provenance without tracer or receptor constraints.*” This replaces earlier wording that overstated what trajectory analysis alone can demonstrate.

Line 188: “likely”; far-from-dust vs. long-range-transport contradiction

Reviewer comment

With all the data that you have at your disposal, can’t you confirm whether this is the case? Also, the idea that your site is far away from dust sources stands in opposition to your idea that long range transport is a source of INPs.

Response

The referee identifies an apparent contradiction that we should have addressed. The statement that the site is “far from dust sources” refers specifically to the absence of nearby dust-emitting regions (deserts, arid lands) in the immediate vicinity of southern Finland. However, mineral dust is well documented to undergo long-range transport to high latitudes, including high-latitude glacially sourced dust that has been shown to be a significant source of ice nucleating particles (e.g., Tobo et al., 2019; Sanchez-Marroquin et al., 2020). The distinction is between *local* dust sources (absent) and *transported* dust (possible but episodic). With our current dataset we cannot definitively confirm or exclude dust transport episodes; this would require back-trajectory analysis or chemical speciation of ice crystal residuals, neither of which are available.

We have revised the text to clarify this distinction and remove the “likely.”

Change in manuscript

Changed “*that is far from dust sources*” to “*where no local mineral dust sources exist, although episodic long-range transport of high-latitude glacially sourced dust (e.g., Tobo et al., 2019; Sanchez-Marroquin et al., 2020) cannot be excluded.*”

Line 207: why would parameterisation work given limited correlations?

Reviewer comment

I don’t understand why one would expect them to capture it. After all, you see limited correlations, and then why would a parameterisation based on those little correlating variables have any chance?

Please connect the approach of testing the parameterisations logically to your findings from previous sections.

Response

The referee makes a valid point: if the screening analysis shows weak correlations, existing parameterisations are unlikely to perform well. We have added the Brasseur et al. (2022) comparison discussion (see also the companion RC1 reply). We now strengthen the logical connection between the correlation analysis and the parameterisation testing.

Change in manuscript

Added before the parameterisation section: *“The limited correlations identified in Sect. 3.2 suggest that existing parameterisations—which rely on variables such as $n_{\text{AP}>0.5\mu\text{m}}$ that do not strongly correlate with INP in this dataset—are unlikely to reproduce the observed INP variability. Testing them nonetheless serves two purposes: (i) it quantifies the parameterisation failure against a comprehensive boreal dataset, and (ii) it motivates the exploration of alternative local proxies in Sect. 3.3.2.”*

Line 207: “both agrees and disagrees” — simplify

Reviewer comment

I think what you mean here is that they disagree with Tobo performing well and agree with DeMott underpredicting. If you do, please simplify that statement accordingly.

Response

Correct. We add a paragraph explaining the Brasseur et al. (2022) comparison context (see also RC1 response). We also simplify the original wording as the referee suggests.

Change in manuscript

Changed *“This observation both agrees and disagrees with shorter-term observations made in the context of a focused intercomparison. . .”* to *“Specifically, the DeMott et al. (2010) underprediction is consistent with the intercomparison findings of Brasseur et al. (2022), whereas the poor performance of Tobo et al. (2013) over the full campaign contrasts with its relatively better performance during their four targeted intercomparison days.”*

Line 209: why the disagreement with previous findings?

Reviewer comment

Why the disagreement with previous findings? Is that due to the time resolution of your measurements?

Response

As detailed in the companion RC1 reply, the apparent disagreement between our results and Brasseur et al. (2022) arises from three factors: (i) dataset scope (four targeted intercomparison days vs. the full seasonal deployment), (ii) the qualified nature of the Brasseur et al. (2022) agreement (their Fig. 11: 35% of 28 March spectrum points within the Tobo (2013) APS-based shaded envelope, vs. 3% for DeMott (2010); Sect. 3.2.3), and (iii) different comparison methodologies (time-series visual overlap vs. scatter plot R^2). The time resolution is not the primary factor; rather, the seasonal scope and aerosol conditions are the key differences. The Tobo (2013) parameterisation performs relatively better during the two spring intercomparison days when biological particle fractions were rising, consistent with its calibration at a biologically active temperate forest site.

Change in manuscript

See the companion RC1 reply for the corresponding manuscript change. The extended discussion is in the manuscript.

Table 1 caption: selection criteria

Reviewer comment

No, in my understanding you did not select based on high scoring, but also based on “physical intuition” (that requires more explanation, see above).

Response

Correct. The companion RC1 reply outlines the dual-selection justification. We also update the Table 1 caption.

Change in manuscript

Changed Table 1 caption from “*selected based on their consistently high qualitative scoring using the various machine learning techniques*” to “*selected based on their high importance ranking in the random forest analysis and established physical or empirical connections to INP in prior literature (see Sect. 3.2).*”

Line 233: test nitrate/acetone correlation with aerosol burden

Reviewer comment

This seems like an easy thing to test given all the data you have available, for example by correlating nitrate and acetone concentrations with aerosol burden.

Response

The referee is right that this is a testable hypothesis. We have added a brief note acknowledging this and providing a qualitative assessment based on our data. Indeed, nitrate and acetone concentrations correlate positively with total aerosol mass during the spring/summer period ($r \approx 0.5\text{--}0.6$), supporting

the interpretation that the INP–nitrate correlation may partly reflect total aerosol burden rather than a specific nitrate–INP mechanism.

Change in manuscript

Added: *“Indeed, nitrate and acetone concentrations correlate positively with total aerosol mass during the PINCii spring/summer period ($r \approx 0.5$ – 0.6), suggesting that their association with INP may partly reflect the total aerosol burden rather than a specific chemical mechanism.”*

Line 238: “intrinsic natural variability of INP”

Reviewer comment

What is the “intrinsic natural variability of INP”? Isn’t this what you’ve been trying to explain? How could that be intrinsic? It needs to be linked to something, doesn’t it?

Response

We agree that the phrase was circular and unclear. What we intended to convey is that the winter INP concentrations appear to vary independently of any locally measured parameter, suggesting they are driven by processes external to the measurement site (i.e., long-range transport and stochastic air-mass variability). We have revised the phrasing to be explicit.

Change in manuscript

Changed *“appear to simply capture the intrinsic natural variability of INP”* to *“appear to reflect INP concentrations driven by processes external to the measurement site—most plausibly long-range transport from diverse, distant sources, as indicated by the log-normal concentration distribution (Fig. 4) and the absence of local predictor correlations.”*

Figure 6 caption: scope of the boreal atmosphere claim

Reviewer comment

This refers only to the boreal atmosphere in spring.

Response

Corrected.

Change in manuscript

Changed *“in the boreal atmosphere”* to *“in the boreal atmosphere during spring and early summer.”*

Conclusion

Line 230: “qualitatively linked”

Reviewer comment

What do you mean by “qualitatively linked”? The links that you get with your ML are quantitative, aren’t they?

Response

The referee is correct that the ML importance rankings and Pearson correlations are quantitative measures. The phrase “qualitatively linked” was misleading. We have changed it to reflect that the associations are statistically detectable but not strong enough to be predictive.

Change in manuscript

Changed “*can only be qualitatively linked to other measured variables*” to “*show statistically detectable but weak associations with other measured variables, insufficient for reliable prediction.*”

Open Research Section

Reviewer comment

Please share the scripts you’ve been using for the ML (or at least documentation thereof, the packages, ...) and the filtering.

Response

Script documentation is as follows: Exploratory analyses were carried out in R; Sect. 2.2 now states the software context (R packages `randomForest` and `rpart`; PCA and K-means via `prcomp` and `kmeans`). Together with the filtering criteria and time-base harmonization already described there, should suffice to reproduce the workflow at the level of detail appropriate to this hypothesis-generating study.

Change in manuscript

Sect. 2.2: added an explicit sentence that analyses were implemented in R, naming the packages/functions used for random forest, decision tree, PCA, and K-means (filtering criteria remain as described in the same subsection).

Language corrections

- 1. 6: commas around “Portable Ice Nucleation...”** Added commas: “, *Portable Ice Nucleation Chamber I and II (PINC and PINCi)*,”.
- 1. 8: “the” Station** Changed “*Station for Measuring*” to “*the Station for Measuring*”.

- l. 12: **“aerosol particle concentration”** Changed to *“aerosol particle concentration”* (added “particle”).
 - l. 30: **“to focus”** → **“with a focus”** Changed *“to focus on measuring”* to *“with a focus on measuring”*.
 - l. 59: **singular “purpose”** Changed *“purposes”* to *“purpose”* (also addressed under Minor Points above).
 - l. 131: **“as illustrated”** → **“which are illustrated”** Changed. Note: the opening of Sect. 3.2 has been replaced with a more descriptive passage (see also the companion RC1 reply); the “which are illustrated” fix is applied to the revised text.
 - l. 207: **“both agrees and disagrees”** Simplified (see Minor Points response above for the revised wording).
 - l. 214: **“algorithms” — only show one** Changed *“several classic and newer machine-supported analysis techniques. These included pairwise correlation, decision tree and random forest learning algorithms”* to *“several machine-supported analysis techniques, including random forest models (Fig. 2), along with pairwise correlation, decision tree, principal component, and K-means clustering analyses.”* This clarifies that multiple methods were used but only the random forest result is shown.
-

We thank the referee once more for the thorough and stimulating review. The comments have prompted us to (i) refine the title and framing to accurately reflect the study’s scope and null-result message, (ii) add a substantial new conclusions paragraph engaging with the deeper question of why INP prediction from local measurements fails, (iii) strengthen the logical connections between the correlation analysis, the parameterisation testing, and the transport hypothesis, and (iv) improve precision in terminology and methodology documentation throughout. We believe the manuscript is significantly strengthened as a result.