

Dear editor and referees,

Thank you for the helpful reviews and suggestions, which we believe helped to make significant improvements to the manuscript. The reviewers' concerns about retaining a critical view to the classification algorithm's validation were especially important to the revisions. We agree that the manuscript would be improved by ensuring thorough, fair discussion on classification by discussing the advantages and disadvantages of using the ML-based methods described in our study, and by offering guidance on ensuring fair use of the resulting classification methods. We have now implemented these changes in the new version, and would like to invite the referees and editor to read the revised manuscript.

We have addressed each referee comment below, with our responses in black and the reviewers' original comments given in grey italics. We refer to the line numbers in the revised manuscript (L__) where changes were made in response to their comments, and include the changes (**blue text** for reformulations/additions in the revised manuscript, and **red text** for the relevant text in the original manuscript). Some further, minor corrections and reformulations for clarity, ease of reading, and consistency were also made to the manuscript; these are specified under the section "Further changes".

Sincerely,

Aiden Jönsson
Dept. of Environmental Science, Stockholm University
Contact: aiden.jonsson@aces.su.se

On behalf of all coauthors: Jinglan Fu, Gabriel Pereira Freitas, Ian Crawford, Pavla Dagsson-Waldhauserová, Radovan Krejci, Yutaka Tobo, Karl Espen Yttri, and Paul Zieger

Response to RC1

1) I am missing one compressed paragraph (for example towards the end of the discussion) that compares the new algorithm with Freitas' in detail: What are the advantages? Is it superior? If not, does it need more validation in order to be? If yes, why? etc. This paragraph should clearly justify why this new algorithm is worth using.

The newly developed algorithm takes more input variables for the particle classification into account compared to the DT algorithm presented by [Freitas et al. \(2022\)](#), which was purely based on fluorescence measurements and earlier performed sea spray simulation experiments with marine bacteria. [Freitas et al. \(2023\)](#) already discussed the potential over- or undercounting of PBAPs, since not all PBAPs are fluorescent with the MBS's excitation/emission wavelengths, or since some non-PBAP particles auto-fluoresce in similar ways to PBAPs. While our algorithm is based on laboratory-sourced input data and then takes more variables (mainly shape parameters) as input for learning, it is also susceptible to misclassification (and thus under- or overcounting), as already discussed in the paper. However, our results clearly illustrate that there are myriad differences among PBAPs from different sources, warranting a flexible algorithm covering a larger range of PBAP types than the DT method can capture. No algorithm is superior to the other, they are rather complementary to the standard decision tree classification (e.g. ABCDEFGH in the MBS and ABC in the WBS) and should always be taken with caution and, ideally, compared to independent tracer and auxiliary aerosol measurement data (as done in this work and in [Freitas et al., 2023](#)) in each measurement context. We also believe that no validation is final, and hence this need, and that each assessment can reveal weaknesses and/or strengths and lead to improvements.

We agree with the reviewer that there is a need to summarize the advantages and recommendations clearly, and that it should be clearly present at the end of the manuscript. We have therefore rewritten the conclusions section to summarize these points around classification and recommendations for using and developing the algorithm going forward. This can be found in lines L598-623 in the new manuscript:

"Using the combined fluorescence and morphology characterization data, we developed a particle classification algorithm using supervised ML trained on known particle types and applied it to one year of field observations from the Zeppelin Observatory, Svalbard. Compared to the previously used fluorescence-only approach, our algorithm yields up to two orders of magnitude higher fluorescent PBAP concentrations during summer, in agreement with previously reported concentrations of PBAPs in Ny-Ålesund during July-August (Johansen and Hafsten, 1988), while reproducing the annual PBAP cycle reported in Freitas et al. (2023a). Although the algorithm identified fungal spores as a major type of PBAP present at Zeppelin Observatory, in line with previous studies (Johansen and Hafsten, 1988), it clearly overestimated pollen fragment concentrations, which may be attributable to similarities shared with plant debris and possibly biomass burning particles. In addition, we trained a dust detection model by exploiting optical morphology differences between dust and SSA particles, which may be used to help determine marine and terrestrial influence in ambient observations. Although laboratory-generated dust and SSA particles were readily distinguished, classification of ambient SSA proved challenging. This suggests significant differences between laboratory-generated and natural aerosols that may stem from environmental processing, motivating improvements in laboratory aerosolization and characterization methods.

Compared to fluorescence-only approaches, the ML framework explicitly incorporates morphology and accounts for a broader range of PBAP classes as well as major non-biological interferents. Domain adaptation further improved performance, although both ML-based and fluorescence-only methods likely remained sensitive to combustion aerosols. It is unlikely that any classification scheme or instrument configuration will fully eliminate the influence of interferents, highlighting the need for parallel observations of e.g., black/brown carbon and source-specific chemical tracers for assessing fPBAP quantifications. Despite site-specific variability in aerobiomes and fluorescence signatures, the framework provides a flexible and expandable basis for estimating PBAP concentrations across environments and reanalyzing existing MBS datasets. However, additional training data are clearly needed to improve both interferent and specific PBAP class identification before it may be suitable for use as a standalone tool. Its continued development will require further validation with field observations and more characterization studies with e.g. biomass burning, pollen emission, fungal sporulation, and atmospheric aging in controlled conditions. The effects of implicating morphology information in fPBAP identification may also be further explored as new characterization data with more realistic aerosolization mechanisms become available. Coordinated collection and open sharing of such reference datasets across research groups will be essential for improving identification methods.”

The first paragraph of the Conclusions section was subsequently changed to more concisely summarize the results of the source experiments (L590-597):

“Although UV-LIF spectroscopy allows for rapid online detection, characterization, and quantification of fluorescent PBAPs, interference from non-biological particle types must be considered. We present an extensive reference dataset of pollen, bacteria, microplastic, cellulose, and dust aerosols made with a multiparameter bioaerosol spectrometer (MBS) with UV-LIF and optical morphology measurements, supplemented with previously published measurements from the same instrument. We show that biological and non-biological particles can share fluorescence signatures, biasing fluorescence-only classifications, and that optical morphology information can improve particle discrimination. While real plant debris was not measured in our characterization experiments, we found broad similarities in the fluorescence properties of cellulose and pollen fragments, suggesting that pollen fragments and plant debris may be easily confused in fluorescence-only classification methods.”

Finally, we have changed all uses of the term “validation” to “assessment” and “validation results” to “auxiliary data” (Sections 2.6 and 3.4 titles, L236, L255, L259, L262, L375, L538) to reflect that the algorithm is not in a “final form”, and explicitly stated that correlations between the classifier results and tracers – especially those used to tune the algorithm – should not be taken as proof of validation but as supporting evidence. This we believe is more honest about the algorithm’s current capabilities and better reflects that it may be a living tool to be further developed and improved. The corresponding statements in the Abstract, Results, and Conclusions were changed. This includes the last paragraphs of the Conclusions section quoted above (L598-623) and the revised abstract’s second paragraph which now does not report correlations (L11-19):

“To explore solutions for these confounding factors, we develop a supervised machine learning classification algorithm integrating fluorescence and morphology information, and evaluate it using MBS and comprehensive chemical tracer observations from Zeppelin Observatory, Svalbard. This evaluation illustrates challenges for inductive

methods in distinguishing biomass burning from biological particles, and dust from SSA, suggesting that important particle classes may be missing and/or 15 laboratory-generated aerosols significantly differ from real-world counterparts. We show that domain adaptation using complementary observations can help address these difficulties. Compared to a fluorescence-only approach, the domain-adapted algorithm reproduces the previously published annual bioaerosol cycle while yielding higher summertime concentrations matching those reported from offline analyses. This open-source algorithm provides a basis for quantifying bioaerosols across diverse environments and can be refined with future field and laboratory efforts.”

As well as the Results section at lines L453-455 and L502-504, respectively:

“This may be reflected in the correlations between fPBAP detected by our algorithm and BB tracers (Fig. 7e), and between cellulose and levoglucosan (Fig. S8).”

→ “This is reflected in improved correlations between fPBAP concentrations with biological tracers arabinol and mannitol (Fig. 7e and Fig. S6); however, these correlations do not constitute fully independent validation of the adapted model and should be interpreted cautiously due to their use in the tuning process.”

“This results in a significant positive correlation with silicon and anti-correlation with SSA tracers sodium, chloride, and magnesium. These correlations are stronger than those from the untuned model’s output (Fig. S7b).”

→ “Although correlations do not independently validate the model, this adaptation results in a significant positive correlation with silicon and anti-correlation with SSA tracers sodium, chloride, and magnesium, which are strengthened relative to the untuned model’s output (Fig. S7b).”

2) One main concern lies in the definitions, interpretations and discussions regarding pollen (3.1: and 480 – 487):

Regarding the definition of SPP (Sub-pollen particles): There is no uniform definition of SPP and people use the term loosely. The term originally comes from the medical community and refers only to the solid starch granules from the inside of the pollen grains (Bacsi et al., 2006; Schleh et al., 2010; Sénéchal et al., 2015; Burkart et al., 2021). These solid starch granules are usually between 0.5 and 2.5 μm in size (e.g. Burkart et al., 2021 for birch; Suphioglu et al., 1992; Schleh et al., 2010 for grasses).

Beside soluble cytoplasmic content, these starch granules are the particles that get expelled during rupture at high RH or when immersed in water. However, at least for birch, this only works for freshly collected pollen grains, and not for purchased ones, even if you put them in water for a longer period (Burkart et al., 2021). Since your measured size distributions of the pollen samples (Fig. 1) are mostly above 2 μm , and you used purchased (not fresh) pollen, I think it is unlikely that the majority of the measured particles are starch granules. This should be clarified.

Some authors also use the term SPP to describe aerosolized and dried aqueous extracts of pollen (Gute and Abbatt, 2020; Gute et al., 2020; Mikhailov et al., 2019; Steiner et al., 2015), but these are typically in the range <500 nm. Since your measured size distribution does not fit any of those two definitions, I would suggest simply calling them pollen fragments, as done, for example by Hughes et al., 2020, or to clearly define what you are talking about. In general, I would switch out “pollen” with “pollen fragments” whenever you discuss your results, since you are not actually measuring pollen grains. This could be confusing to readers.

I would also suggest examining one pollen sample under a microscope to see what these fragments > 2 µm look like. It should be easy to identify if these are starch granules (smooth surface, round to ellipsoid in shape, compare with Fig. 2-4 in Burkart et al., 2021), or really fragments of the pollen exine. You could include microscopy images in the Supplement.

This should also be considered in your discussion in 480-487.

We are thankful to the reviewer for pointing out and clarifying the distinction made for SPPs in the literature, and agree that the manuscript should be made consistent with the body of literature, as well as more correct in terms of what is being observed in the experiments, and in line with the difficulty of identifying which exact components of the pollen were being measured. We have therefore chosen to follow the reviewer's suggestion to replace all uses of the term (SPP) with "pollen fragment", and adjusted the discussion around pollen fragments in the section indicated (L516-520) in the revised manuscript).

Unfortunately, the pollen samples used in the experiment are now many years old, and visually identifying the fragments observed in our experiments would not be possible. We have therefore elected not to do a visual inspection of the samples. The majority of particles in the dry experiments seem to have highly irregular shape, but as we cannot confirm by sampling the particles measured during the experiments and visually inspecting them, we will retain the term "pollen fragment" terminology and keep a straightforward discussion of their morphology as measured by the MBS in the manuscript. All discussions around the particles seen in the pollen experiments have now been adjusted to ensure that they are open to possibilities of which particles these fragments are (i.e. using "pollen fragments" and "pollen-derived particles" consistently, and removing points where the reader may be confused that we may have been referring to whole pollen).

"However, since the MBS can only detect SPPs and not intact pollen grains, wet-processed pollen samples are useful for understanding SPPs, which are primarily generated in humid conditions. This may be especially useful for detecting cloud-activated pollen fragments sampled via, e.g. counterflow virtual impactor (CVI) inlets (Freitas et al., 2024). As a complicating factor, pollen rupture is stochastic causing high variability in SPP morphologies, reflected in the wide range of scattering signatures we observed."

→ L516-520: "These results may be especially useful for detecting pollen fragments produced by humid/wet atmospheric processes, as well as cloud-activated pollen fragments sampled via e.g. counterflow virtual impactor (CVI) inlets (Freitas et al., 2024) and by whole-air inlets during cloudy conditions. As a complicating factor, pollen rupture is stochastic, leading to high variability in pollen fragment morphologies (as reflected in the wide range of scattering signatures we observed)."

Further instances of "SPP" were changed to "pollen fragments" elsewhere in the document (L278-280, L291, L304, L330, L333-334, L362, L373, L385, L395, L405-407, L521-524, L534-535,).

3) 425-445:

The discussion about pollen (fragments) needs to be even more critical. I highly doubt that 60-80 % of bioaerosols in the arctic are pollen grains, let alone pollen fragments which aerobiologists do not observe in higher concentrations than intact pollen grains in their Hirst traps (they would be colored pink after treatment with fuchsin, as the pollen grains are, and therefore easy to spot). Some fragments, however, were observed by Hughes et al., 2020 (see

Fig. S5 in their Supplement).

You already discuss possible reasons for misclassification. However, this could be better structured. Therefore, I would start with a statement that such high pollen (fragment) concentrations are highly unlikely, and support this with a comparison to earlier studies that you already mention throughout the paragraph (SEM pictures and pollen counts). You state that earlier studies found low concentrations of pollen grains. What is the approximate concentration Johansen and Hafsten found, and how does this compare to your pollen (fragments) concentration?

I would then continue to explain the possible reasons why the particles classified as pollen (fragments) here are too high. You have already discussed several things. For instance, misclassification of combustion particles and fungal spores. However, the correlation of pollen and fungal spores with mannitol and arabitol are not similar, but pollen correlation is significantly higher. This should be clearly emphasized.

Parts of this paragraph already read more as a discussion than results, and lines 514-527 in the discussion section partly repeat the earlier discussion but also introduces new ideas, such how the validation could be improved (for instance through comparison with pollen and fungal spore counts). I recommend merging those parts and put all in the discussion.

We agree with the assessment that it is unlikely that pollen comprise such a portion of the bioaerosols present at Svalbard as discussed in the manuscript, and we agree that the discussion around this part of the results should be clearly critical. We believe the reviewer's suggestion to clearly report on the likely inconsistency in the Results and consolidate the discussion around factors responsible for misclassification in the Discussion section helps to avoid overstatement and confusion about our interpretation, and have chosen to follow this. The changes can be found in the revised manuscript's Results section discussing Figure 7c (L466-477), which we now reformulated to focus on reporting the algorithm results and maintain a critical tone of the pollen identification result:

"Pollen contributes the largest class share to fPBAP by our algorithm throughout the year but more consistently in summer, followed by fungal spores and bacteria (Fig. 7c). A portion of the fPBAP class may be combustion particles falsely identified as pollen due to their similarities described in Sections 3.2 and 3.3. Another portion may be fungal spores misclassified as pollen, as evidenced by their similar behaviors in correlations with fungal tracers. Bacteria makes up a small proportion (~0.1-1%) of fPBAP identified by our classifier. Unfortunately, contributions by specific fPBAP classes are difficult to validate with tracer data available at Zeppelin Observatory for this time. Although arabitol and mannitol are considered fungal tracers, they can be found to lesser degrees in pollen and bacteria (Lau et al., 2006; Di Filippo et al., 2013). PBAP analyzed by Freitas et al. (2023a) and Tobo et al. (2024) with electron microscopy could not be precisely distinguished, although visual comparison suggested that they were likely fungal spores and bacteria. Tobo et al. (2024) also found irregular carbonaceous particles resembling plant debris, a globally abundant PBAP (Sánchez-Ochoa et al., 2007; Winiwarter et al., 2009), or possibly pollen fragments. Considering the similarities between pollen, fungal spores, and pure cellulose seen in Sections 3.2 and 3.3, plant debris not flagged as combustion particles would likely be classified as pollen- or fungal spore-like by our classifier due to their broad fluorescence spectra (as in Fig. 6a). Furthermore, plant debris can be emitted by BB (Holden et al., 2011), with the degree of similarity to smoke particles depending on the degree of influence by combustion. This may be reflected in

the correlations between fPBAP detected by our algorithm and BB tracers (Fig. 7e), and between cellulose and levoglucosan (Fig. S8). Earlier identifications of PBAP classes in Ny-Ålesund in two summers found very low concentrations of intact pollen and found local species to dominate over exotic pollen (Johansen and Hafsten, 1988), although sediment cores from Svalbard indicate considerable variability in the remote transport of exotic pollen (Poliakova et al., 2024). These earlier assessments also found a near constant presence of airborne fungal spores at considerable concentrations in Ny-Ålesund (Johansen and Hafsten, 1988), which is consistent with the prevalence of fungal spores identified by our classifier and the analyses of Freitas et al. (2023a) and Tobo et al. (2024)."

→ "The pollen class contributes the largest share to particles identified as fPBAP by our algorithm throughout the year but more consistently in summer, followed by fungal spores and bacteria (Fig. 7c). Bacteria makes up a small proportion (~0.1- 1%) of fPBAP identified by our classifier. It is unlikely that pollen fragments contribute such a large share of PBAPs in Svalbard; earlier identifications of PBAP classes in Ny-Ålesund in two summers found very low concentrations (~10⁻³-10⁻¹ L⁻¹ in summer) of intact pollen (Johansen and Hafsten, 1988), while the mid-July peak concentration of pollen-class particles identified by our classifier is 3.30 L⁻¹. Considering the similarities between fPBAP classes, cellulose, and combustion particles seen in Sections 3.2 and 3.3, plant debris or combustion particles may be misidentified as pollen- or fungal spore-like by our classifier due to their broad fluorescence spectra (as in Fig. 6a). This is corroborated by the persistent presence of pollen-like and fungal spore-class particles identified throughout winter (Fig. 7c). Fungal spores misclassified as pollen fragments may possibly also contribute to the dominance of the pollen class in fPBAP identifications, as suggested by the pollen-like class correlating more strongly with fungal tracers arabinol and mannitol ($\rho = 0.69$ and 0.65 , respectively) than the fungal spore class does ($\rho = 0.49$ and 0.43 , respectively; Fig. 7e)."

We have now also merged discussion on reasons for misclassification that are extraneous to or not confirmable with our results and moved them to the Discussion (L545-572):

"Our classifier identified pollen fragments and fungal spores as the most common classes of fPBAPs observed at Zeppelin Observatory in 2020. While previous work establishes fungal spores as common in Ny-Ålesund (Johansen and Hafsten, 1988), consistent with the prevalence of fungal spores identified by our classifier and the analyses of Freitas et al. (2023a) and Tobo et al. (2024), whole pollen concentrations were very low (~10⁻³-10⁻¹ L⁻¹ in summer; Johansen and Hafsten 1988). PBAP sampled during the NASCENT campaign analyzed by Freitas et al. (2023a) and Tobo et al. (2024) with electron microscopy could not be precisely distinguished, although visual comparison suggested that they were likely fungal spores and bacteria. Tobo et al. (2024) also found irregular carbonaceous particles resembling plant debris or possibly pollen fragments. Additional tracers could help verify class contributions to fPBAP detected by our classification algorithm, but source overlap complicates tracer interpretation. Although arabinol and mannitol are considered fungal tracers, they can be found to lesser degrees in pollen and bacteria (Lau et al., 2006; Di Filippo et al., 2013). Wildfires have also been found to directly promote PBAP emissions (Holden et al., 2011; Moore et al., 2020; Kobziar et al., 2022; Ellington et al., 2024), thus it is not unexpected that BB particles and PBAP can co-exist in ambient observations. More source-specific tracers, such as sucrose for pollen (Mampage et al., 2022), dipicolinic acid for bacteria (Mampage et al., 2022), and ergosterol for fungal spores (Lau et al., 2006), are necessary

to thoroughly validate PBAP class contributions.

Pollen fragments are difficult to robustly identify and quantify in ambient conditions due to high variability in their emission pathways (Suphioglu et al., 1992; Subba et al., 2021). Our results suggest that BB aerosols may contain particles with strong similarities to pollen fragments, as evidenced by the stronger correlations between BB tracer levoglucosan concentrations and this particle class relative to other fPBAP classes, and by their near-constant presence throughout winter in Svalbard when fluorescent aerosols are likely comprised mostly of combustion-derived particles originating from continental sources (Yttri et al., 2024). Including BB aerosols in future experiments may therefore improve fPBAP and, in particular, pollen fragment identification in our classification algorithm. Based on the apparent similarities in fluorescence characteristics between other fPBAP classes and cellulose assessed in our experiments, it is also possible that many of the particles in the Zeppelin Observatory dataset identified as fPBAP by our classifier were misclassified plant debris. Since real plant debris was missing from the characterization data presented here, representing a potentially important missing class of a globally abundant PBAP (Sánchez-Ochoa et al., 2007; Winiwarter et al., 2009), we suggest that future characterizations measuring e.g., leaf litter be conducted with the MBS to improve identification. Future assessments for specific fPBAP class validations when e.g., pollen trap counts, digital holography-based pollen identification, and aerobiome DNA sequencing are conducted in parallel to MBS observations should be made to provide further validation.”

4) *The last thing that is unclear for me is the following: You used soft labels in order to tune the algorithm. Since prior to tuning, the results seemed unreasonable, do you expect that you need these labels and therefore tracer measurements in every future campaign, or was this sufficient as a one-time adjustment to get meaningful classes for future measurements? Please comment on this.*

This is a good comment to discuss, both here and in the manuscript, as it would help provide suggestions and guidance for its use and clarify our intentions. Future adjustments by further tuning with new ambient measurements might only be useful if the measurements have been sufficiently studied and documented; the Zeppelin dataset was chosen because of the extent of validation and supporting observations that exist for it, which is rather rare. Our intention is to use supporting measurements in future campaigns, such as of eBC, and assess the performance of the algorithm in each setting; this was suggested in the concluding paragraph of the original manuscript. Supporting measurements may be used to choose thresholds for identification, such as in the pollution model; this was also suggested in the original manuscript. The future development of the algorithm and its models is discussed in the Discussions and Conclusions sections, but with a focus on including new data from future laboratory characterization experiments (specifically, including biomass burning/woodsmoke and plant litter in future iterations) where particle type is constrained, rather than future tuning. This we believe is preferable to tuning, which our study explores as an option when facing a lack of training data. We have stressed this in the revised Discussion paragraph on this (L499-512), which better reflects that this is an exploration of the domain adaptation method rather than a final product, and that future development is recommended primarily through making more characterization data from controlled settings available.

“The ease of tuning the LRM allowed us to implement domain adaptation using field validation data, improving correlations between our algorithm’s fPBAP estimates and

PBAP tracer concentrations. The significant correlation between the wood pyrolysis tracer levoglucosan and HFP concentrations indicates that particles originating from BB were likely misidentified as fPBAP by the untuned classifier. These may differ from diesel engine-sourced combustion particles such as those included in our training data and could have been modified by aging processes. However, wildfires have been found to directly promote PBAP emissions (Moore et al., 2020; Kobziar et al., 2022; Ellington et al., 2024). Determining the degree to which biological tracer concentrations are influenced by, e.g., plant debris and PBAPs emitted via BB is ultimately difficult. Our classification algorithm may be improved in the future by including source characterization data from wood and biomass burning experiments.”

→ “The ease of tuning the LRM allowed us to implement domain adaptation using field validation data, improving correlations between our algorithm’s fPBAP estimates and PBAP tracer concentrations. Significant correlations between fPBAP concentrations derived with the untuned classifier and wood pyrolysis tracer (levoglucosan) concentrations indicates that particles originating from BB were likely misidentified as fPBAPs. These may differ from diesel engine-sourced combustion particles like those included in our training data and could have been modified by aging processes. Our classification algorithm may be improved in the future by including source characterization data from BB experiments; we suggest that the models comprising it be trained from scratch and re-assessed when such data are available before further tuning is applied.”

Specific comments:

The Supplement could be better structured according to their appearance in the main text.

The supplement has now been restructured to be in line with the order of appearance in text, and references to supplementary figures and tables were subsequently changed to reflect their new order in the SI.

Titel: Wouldn't "anthropogenic" be a better word than "human"?

We agree that this is a more fitting wording! The title has now been changed as suggested.

6: I suggest to add the instrument here.

This is a good suggestion; we have added the name of the instrument in the new formulation of the abstract (L5-8):

“Using laboratory measurements of single-particle ultraviolet light-induced fluorescence (UV-LIF) spectroscopy and morphology, we provide a new reference dataset for coarse-mode aerosols from common sources, including pollen, dust, bacteria, and microplastics.”

→ “Using laboratory-based single-particle ultraviolet light-induced fluorescence spectroscopy and morphology measurements made with a Multiparameter Bioaerosol Spectrometer (MBS), we provide a new reference dataset for coarse-mode aerosols from common sources, including pollen, dust, bacteria, and microplastics.”

30 – 31: *Can the concentration range you give for bacteria also be found in Huang et al., 2021?*

If yes, make that clear, if not, find a reference supporting this range.

We agree that this is a confusing citation; [Huang et al. \(2021\)](#) focuses on the PBAP-INP link and is thus an indirect source for PBAP variability. We have now reformulated the sentence (L31-33) to instead refer to [Depres et al. \(2012\)](#) where it is easily found:

“Although PBAPs are far less abundant and highly variable, typically between 10^{-3} - 10^1 L⁻¹, even low concentrations in mixed-phase clouds may substantially affect cloud properties due to highly efficient glaciation feedbacks (Huang et al., 2021).”

→ “Although PBAPs are far less abundant and highly variable, typically between 10^{-3} - 10^1 L⁻¹ (Deprés et al., 2012), even low concentrations may substantially affect mixed-phase cloud properties due to highly efficient glaciation feedbacks (Huang et al., 2021).”

33 & 35: Since “online aerosol mass spectrometry” includes “Single-particle mass spectrometry”, I suggest rephrasing the former, or explain both under the parent phrase “online aerosol mass spectrometry”.

These sentences have now been reformulated to name both bulk and single-particle mass spectrometers under the parent phrase (L35-39):

“Online aerosol mass spectrometry offers higher time resolution for chemical composition, but it can typically resolve only one size range per instrument (Nash et al., 2006; Pratt and Prather, 2012).”

→ “Online aerosol mass spectrometry offers higher temporal resolution for chemical composition; these instruments are designed to determine bulk aerosol composition in only a narrow size range (Nash et al., 2006; Pratt and Prather, 2012) or at single-particle resolution.”

The sentence following this was also reformulated due to the changes above (L35-39):

“Single-particle mass spectrometry can offer valuable, highly resolved particle composition information, but has highly size-dependent, low ($\ll 50\%$) detection efficiencies (e.g. Jacquot et al. 2024) and is costly and difficult to miniaturize, making its employment quite rare (Lai et al., 2025).”

→ “Although single-particle mass spectrometers can offer valuable information, they typically have very low ($\ll 50\%$) detection efficiencies (Jacquot et al., 2024) and are costly and difficult to miniaturize, making their employment quite rare (Lai et al., 2025).”

34: change “only one size range” to “only a narrow size range”.

This change was included in the reformulation (L39; see previous response for revised sentence).

60: While it may be true, that microplastics contain PAHs, they are not produced during production. The main reason for their fluorescence is either because they contain aromatic structures that are expected to fluoresce (e.g. PET, PS) (Gratzl et al., 2024), the production of trace amounts of α - β -unsaturated carbonyls during synthesis or storage (e.g. PE, PP) (Grabmayer et al., 2014; Htun and Klein, 2010), or due to additives (Morgana et al., 2024). I also suggest including Beres et al., 2024 in the references, which also used autofluorescence to detect airborne microplastics.

We are grateful for the correction! This sentence has now been split in two and reformulated (L67-70) to name that both base and additive materials are expected to fluoresce given that

they contain aromatics, and that they may arise from aging/storage. These references on storage and Beres et al. (2024) have also been included.

“Plastic polymers may contain trace amounts of PAHs from production and thus fluoresce, a property used to detect airborne and marine microplastics (Morgana et al., 2024; Gratzl et al., 2024; Pandey et al., 2022).”

→ “Plastic polymers typically fluoresce due to the presence of aromatic groups in base materials or trace additives used in their synthesis (Gratzl et al., 2024), or by the production of fluorophores during aging (e.g. Htun and Klein 2010; Grabmayer et al. 2014). This property has been used to detect airborne and marine microplastics in the environment (Pandey et al., 2022; Morgana et al., 2024; Beres et al., 2024).”

89: *Add the names of the channels (A - H). I think it would also make it easier for the readers if the central wavelengths were written here.*

The channel labels and central wavelengths have now been added in the reformulated sentence (L97-100). Note also that we have included a detail on the spectral range of the MBS; there is a long-pass filter that limits the lower wavelength limit of detected light, which we have now specified in the revised sentence.

“Fluorescence emitted after excitation is collected by a hemispheric mirror, dispersed using a diffraction grating, and measured in eight acceptance intervals (bands, channels) with roughly equidistant central wavelengths spanning approximately 315-640 nm (Fig. S3).”

→ “Fluorescence emitted after excitation is collected by a hemispheric mirror, limited to exclude <305 nm light with a long-pass filter, dispersed using a diffraction grating, and measured in eight acceptance intervals (channels, labeled A-H) with roughly equidistant central wavelengths (315, 364, 414, 461, 508, 552, 595, and 640 nm, respectively) spanning a detection region of approximately 305-655 nm (Fig. S1).”

95: *Is the background only measured with one single light pulse?*

The background is measured with 42 pulses, from which the mean and standard deviations of channel detector background intensities are determined; this detail has now been added to the sentence, with a reference to [Freitas et al. \(2022\)](#) where it is also explained (L106-108).

“The MBS periodically measures the background fluorescence, i.e. without sample flow, once when beginning the measurement and then typically after 30,000 samples.”

→ “The MBS periodically measures background intensities sensed by each detector, i.e. that of the optical stage after a pulse of excitation light without sample flow, with a series of 42 xenon lamp flashes once when beginning the measurement and then typically after 30,000 samples (Freitas et al., 2022).”

We have also reformulated the following sentences (L108-111) to be more clear about the data treatment using the background detection:

“This provides an estimate of the instrument’s baseline, which is subtracted from each particle’s measured spectrum.”

→ “This provides an estimate of the instrument’s baseline detector intensity spectrum, which is subtracted from each subsequently measured particle’s measured spectrum. Fluorescence intensity can then be expressed in terms of standard deviations (σ ,

determined by the spread in background measurements) above the background mean.”

124: *How was it ground into particles?*

The cellulose used was in crystalline powder form; this phrase was left from a previous formulation of the sentence when written together with PE, and has been removed L(135). We thank the reviewer for helping to catch this.

“We used pure, synthesized crystalline cellulose (Sigma Aldrich, cat. no. 435236) ground into particles.”

→ “We used pure, synthesized crystalline cellulose (Sigma Aldrich, cat. no. 435236).”

125-132: *How old were the pollen samples? How were they stored?*

The pollen samples were collected the last pollen season (spring/summer 2021) and stored in HDPE sample tubes in a cool room (<8° C) at the supplier. After receipt in Sweden (January 2022) they were stored in the lab freezer; the experiments were carried out between January-March 2022. This information has now been added to the manuscript (L143-144).

“Pollen was collected by hand, with the exception of juniper, which was collected with vacuum. All samples were processed dry (sifting and drying), with the exception of willow, which was defatted with acetone.”

→ “Pollen were collected by hand, with the exception of juniper, collected with vacuum, the spring/summer before the experiments took place (January 2022). All samples were processed dry (sifting and drying), with the exception of willow, which was defatted with acetone. Samples were stored cold (<8° C) in centrifuge tubes at the supplier prior to dispatch and in a freezer (~-20° C) thereafter; experiments were conducted over the next two months.”

137: *How was the PE ground? How was the UV-aging exactly achieved? What light source / what wavelength was chosen for the aging? How long was the exposure?*

The PE samples were provided by a colleague in both normal and UV aged format. The process of production, aging and shredding were performed as with current standards in the microplastics field (such as those used by our colleague, e.g. [Reichelt & Gorokhova, 2023](#)). It is however not known the process of aging involved in this specific sample. Thus, readers should be cautious when directly comparing their results to ours when analyzing different stages of aging. We have confirmed that the response of the PE particles' fluorescence properties to aging matches those expected by other more sensitive methods such as [Ho et al. \(2026\)](#). Thus, the main results stem from the ability of the MBS to detect their fluorescence (i.e., that the particles small enough to be aerosolized yield enough photons to be detected) and that even normal shredding practices can produce them. We have now added a sentence urging readers to keep this in mind in the methods section (L152-155).

→ “While we compare the relative effects of aging in our experiments, we do not attempt to quantify physico-chemical changes resulting from UV aging as the MBS is not suitable to this end. Measured quantities reported here are thus not expected to be directly comparable with those found in other experiments using more exact methods.”

140-145: *Are these commercial dust samples or did you collect them?*

These were field samples collected by coauthors Pavla Dagsson-Waldhauserová and Yutaka Tobo, and by Ingrid Zieger; this has now been specified (L157).

“Natural dust samples from four sources, ...”

→ “Natural dust samples collected in the field from four sources, ...”

156: *Were background measurements taken between the samples?*

We are unsure of which background measurements are being referred to; if this is meant to mean the fluorescence background (or force trigger/FT mode), these are automatically performed at the beginning of every measurement cycle (after stopping measurements) and after every 30,000 particles, and the processing routine removes the most recent FT measurement, i.e. the background changes with time. Therefore, this is done at the beginning of every sample measurement as the MBS was turned off in between. If this is meant to mean blank measurements of empty containers with the instrumental setup, these were performed periodically throughout the experiments. This was not after every sample, but typically between workdays and after cleaning the setup.

Table 1: Include the size of the PSL.

The PSL size has now been added to Table 1.

174: *here you call them objectives, but later on you refer to task 1, task 2, ... Maybe stick to task.*

“Objectives” has been changed to “tasks” as suggested (L194).

215: *eBC data: Since there is a minutely resolution available, why did the author only compare 24 h data, and not for example 1 h data?*

The daily eBC data as published in [Freitas et al. \(2023\)](#) are processed and cleaned; we elected to use the published data to reduce the work of re-processing data and maintain comparability with their results. The MBS samples relatively small volumes and requires longer sampling times for reliable concentration estimates, especially for rare particles such as highly fluorescent particles in the Arctic, making daily average concentrations more reliable and less stochastic. This consideration also went into the use of daily averages in [Freitas et al. \(2023\)](#).

216: *You should include for what aerosol classes the tracers are for in Table S3 and give some examples here in the main text.*

We agree that it is useful to include the aerosol types traced by each species in Table S3; this has now been added to the supplementary Table S3. As examples and references for tracers are given in the next paragraph (L243-253), we have elected not to add more examples to the main text.

264: *You could add the central emission wavelength of the B channel here.*

We agree this would be helpful, it has now been added following the words “B channel peak” as suggested (L286).

“Figures 2a-c show that fluorescence spectra recorded for bacteria are dominated by a B channel peak (~315 nm), ...”

265: *The authors may correct me if I'm wrong, but I did not find evidence that bacteria fluorescence peaks in this channel based on the two references provided. While I agree that tryptophan, the main fluorophore responsible for bacterial fluorescence peaks in the B channel, whole cells are not discussed in these publications. I suggest to reference Kinahan et al., 2019;*

Savage et al., 2017 and/or Hernandez et al., 2016). They do not use the exact same excitation/emission wavelength as the MBS but measure actual cells.

This could stand to be clarified; the sentence has been adjusted to specify that the region of fluorescence emission of tryptophan is noted in [Pöhlker et al. \(2012\)](#) and [Huffman et al. \(2020\)](#), and that the B channel peak has been linked especially to bacteria by [Freitas et al. \(2022\)](#). The suggested references are good additions to the finding, and have been added in the reformulated sentence.

“..., consistent with Pöhlker et al. (2012) and Huffman et al. (2020), and the decision tree criterion of Freitas et al. (2022).”

→ “... , consistent with the known region of fluorescence emission by tryptophan when excited at this wavelength (Pöhlker et al., 2012; Huffman et al., 2020), the findings of Freitas et al. (2022), and characterizations of bacteria with other UV-LIF instruments (e.g. Hernandez et al. 2016; Savage et al. 2017).”

293: *What is the approximate size range of the combustion particles?*

The combustion particles were quite fine and clearly separated by size from most biologically sourced particles, with median optical diameters <2 μm (Fig. 4b).

296: *regarding discrimination: Studies found a high correlation of eBC concentrations with a WIBS channel using a different excitation wavelength than the MBS (Gratzl et al., 2025; Gao et al., 2024; Beck et al., 2024).*

This is a good way to illustrate this! We have now reformulated and split up the sentence to show that this has been useful in the field, and added references (L303-306).

“Using instruments with multiple excitation wavelengths, such as the WIBS (Markey et al., 2024), can add discriminating power; despite higher spectral resolution in fluorescence emission detection, the MBS is limited in this regard, motivating a combined fluorescence-morphology classification approach.”

→ “Using multiple excitation wavelengths can add discriminating power to UV-LIF instruments; in studies employing the WIBS, particles excited at a different wavelength than that of the MBS were shown to correlate with combustion tracers (e.g., Gao et al. 2024; Beck et al. 2024; Gratzl et al. 2025). Despite higher spectral resolution in fluorescence emission detection, the MBS is limited in this regard, motivating a combined fluorescence-morphology classification approach.”

301: *Why was 3 μm used?*

This was an arbitrary choice since morphology signals become clearer with larger particle sizes. This sentence has been reformulated for clarity (L308-309).

“For clarity, we show morphology parameter distributions for <3 μm and $\geq 3 \mu\text{m}$ particles.”

→ “Since optical morphology signals are stronger with larger particles, we visualize morphology parameter distributions for <3 and $\geq 3 \mu\text{m}$ particles.”

310: *The explanation of what PSL particles are should come in the methods, not here.*

We agree that this should be in the methods section, and this has now been added (L185-188) – we are thankful for pointing this out.

Added: “In addition, we include data from polystyrene latex spheres (PSLs) measured for

calibration in Beck et al. (2024); these represent particles with well-constrained optical and morphological properties (spherical and opaque) and provide another sample to the microplastic class.”

The sentence pointed out by the reviewer was also adjusted accordingly (L322).

“Figure 4c-f and Fig. S5 compare dust, nascent SSA (Beck et al., 2024), and PSLs (Beck et al., 2024) (a spherical, opaque reference particle).”

→ “Figures 4c-f and Fig. S5 compare dust, nascent SSA (Beck et al., 2024), and PSLs (Beck et al., 2024).”

310-328. The dust and SSA results should also be compared to the other particle types and not just to each other (e.g. low to zero HFAP contribution).

We agree that this is a helpful addition to the results, and have now added two sentences with this comparison (L344-346):

“Optically, pollen fragments and fungal spores have irregular scattering signatures resembling dust with low peak counts and high asymmetries, mean signals, and PTMRs. Meanwhile, bacteria optically resemble SSA and PSLs, consistent with smooth surfaces.”

We have also added a small comparison on their fluorescence fractions (L309-310):

“Natural dust measured in our experiments and the filtered seawater measurements in Beck et al. (2024) also contained low fractions of HFPs (<0.1%), indicating some organics or potential PBAPs present in these samples.”

4: A color legend for the different classes would be nice.

This is a good suggestion, and has now been implemented in the revised Fig. 4.

340: Why was 2.5 μm chosen as the threshold?

2.5 μm was chosen as the scattering signal, and differences between classes, increases with particle size; differences become difficult to distinguish with smaller particles, as can be seen in Figures S5-S6 We chose this threshold as an arbitrary threshold based on inspecting differences in morphologies by size, as in the aforementioned supplementary figures, and by training the algorithm with different lower cutoff sizes to manually search where performance affected. Another method used for choosing particles for a sufficiently measured optical scattering signal is to use the sum, or mean (Figures S5d and S6d), although this introduces problems since these are size-dependent. This has now been explicitly stated in the text (L369-371):

“Following the discussion around Fig. 4, we trained the dust model only on dust and SSA particles $\geq 2.5 \mu\text{m}$ to ensure more distinct separation between their scattering signatures, as significant overlap in morphology signal distributions among particles of different classes are seen below this size range.”

The later mention of this was removed (where it would have been L415 in the revised MS; L383-384 in the original MS) following this change.

341: The terms “Pollution model” and “pollution-likelihood” should appear already in Section 2.5. Same for “dust model” at line 348.

This has now been defined in Section 2.5 (L202-205) in the following revised sentences:

“This approach is well suited to task 1, flagging observations as interferents or non-interferents, and to the dust vs. non-dust decision in task 3. The model output is a probability estimate between 0 and 1, obtained by averaging logistic function outputs across input features.”

→ “This approach is well suited to task 1, flagging observations as interferents or non-interferents in a pollution model, and to the dust vs. non-dust decision in a dust model for task 3. The model output is a probability estimate between 0 and 1, obtained by averaging logistic function outputs across input features, which we name pollution-likelihood and dust-likelihood for the pollution and dust models, respectively.”

343, 348: *What is the confidence threshold?*

This threshold is meant to be flexible and can be adjusted for confidence needs or as a result of validation/correlations with any available tracers in each measurement setting, but in the initial assessment we used a simple majority threshold (0.5) which is the typical “default” for logistic regression/binary classification. The last sentence of this paragraph has been adjusted to more clearly define these (L349-350).

“In our assessment, we used the most likely class (the rounded probability for LRMs and the class of highest 350 probability for the kNN) for identification.”

→ “In this assessment, we used the most likely class (the rounded probability for LRMs, i.e. α_{pol} , $\alpha_{dust} = 0.5$, and the class of highest probability for the kNN) for identification.”

396: *Since you reference to the Supplement figure here that includes “tuned” and “untuned” data, you should introduce these two terms already earlier in the section.*

The paragraph has been reworded to refer to the model trained only on lab data as the untuned model for clarity (L415-420).

“The algorithm, trained only on source characterization data, struggled to recognize the types of combustion particles seen at Zeppelin Observatory, as evidenced by the high correlation between fPBAP concentrations and levoglucosan and eBC (Fig. S7b). The fact that fPBAP concentrations correlated more strongly with levoglucosan than with eBC may suggest an influence by specific combustion particle types, i.e., BB- rather than fossil fuel-derived. Biological tracer and DT-based (Freitas et al., 2023a) fPBAP concentrations indicate a clear seasonal cycle of biological aerosol influence, which our algorithm initially failed to capture (Fig. 7b). This resulted in likely overestimated fPBAP concentrations, especially evident between January April when eBC concentrations were high.”

→ “The algorithm in its raw, untuned form, trained only on source characterization data, struggled to recognize the types of interferents seen at Zeppelin Observatory, as evidenced by the high correlations between fPBAP concentrations and levoglucosan and eBC (Fig. S6b). That fPBAP concentrations derived from the untuned model correlated more strongly with levoglucosan than with eBC may suggest an influence by specific combustion particle types, i.e., BB- rather than fossil fuel-derived. Biological tracer and DT-based (Freitas et al., 2023a) fPBAP concentrations indicate a clear seasonal cycle of biological aerosol influence (Fig. 7a), which our algorithm initially failed to capture in its untuned form (Fig. 7b). This resulted in likely overestimated fPBAP concentrations, especially evident between January-April when eBC concentrations were high.”

398: Refer to Fig. 7A after "biological aerosol influence".

Added in the reformulated sentence (L406):

"Biological tracer and DT-based (Freitas et al., 2023a) fPBAP concentrations indicate a clear seasonal cycle of biological aerosol influence, which our algorithm initially failed to capture (Fig. 7b)."

→ "Biological tracer and DT-based (Freitas et al., 2023a) fPBAP concentrations indicate a clear seasonal cycle of biological aerosol influence (Fig. 7a), which our algorithm initially failed to capture in its untuned form (Fig. 7b)."

7a: fPBAP (this study) is the tuned one, right? This should be in the legend.

Agreed, this figure (Fig. 7) has been revised and the labels adjusted accordingly.

404: Which tracers were used as soft labels?

Here eBC and arabitol were used for soft labels; "used for soft labeling" was now added following "pollution influence" to specify that the influence metric is used in soft labeling (L432-436). The same was done for the dust/SSA metric (L491-494).

"This is accomplished by choosing a number of sample periods with sufficient tracer data and labeling HFPs observed during these periods as either pollution or not based on the soft label, and further training the LRM on these data (the tuning block) with lower weights than those of the initial training data to limit its influence on model fit; details of this step are explained the SI."

→ L432-436: "This is accomplished by choosing a number of sample periods with sufficient tracer data and labeling HFPs observed during these periods as either pollution or not based on the soft label, and further training the LRM on these data (the tuning block) with lower weights than those of the initial training data to limit its influence on model fit; details of this step are explained the SI (Section S1). The pollution influence metric used for soft labeling was calculated using normalized eBC and arabitol concentrations."

"To address this weakness, we applied domain adaptation on the dust LRM, this time using mineral constituents silicon, aluminum, iron, manganese, titanium, nickel, and chromium for mineral dust mass and the sum of major SSA ions (sodium, chloride, magnesium, and potassium) for sea spray; details of our dust influence metric's construction are further explained in the SI."

→ L491-494: "To address this weakness, we applied domain adaptation on the dust LRM, this time constructing soft labels using mineral constituents silicon, aluminum, iron, manganese, titanium, nickel, and chromium for mineral dust mass and the sum of major SSA ions (sodium, chloride, magnesium, and potassium) for SSA mass; details of our dust influence metric's construction are further explained in the SI (Section S1)."

407: Why 20%?

There is no hard-and-fast rule about how much weight to assign samples used in domain adaptation, but it is generally agreed that a conservative weighting is preferable to avoid overfitting on unknown samples ([Kouw and Loog, 2021](#)). In testing the sensitivity to this, we found that 10% influence was not enough to significantly change the behavior of the classifier, but that large differences began to occur at 30%. We therefore continued to use 80% training/20% soft-labeled influence split as a benchmark with which to assess the

effects of domain adaptation on it. We have now reformulated the text around this and stated these considerations (L438-434).

“The pollution influence metric was calculated using normalized eBC and arabitol concentrations, and the combined weighting of the tuning block on the LRM’s fitting was set to 20% of the training data’s. Individual weights were calculated for tuning data based on this scaling and our confidence in the soft labels, determined by the pollution influence metric.”

→ “To reduce the risk of overfitting to uncertain pseudo-labeled particles while preserving the classifier’s fit to the laboratory reference data, the total contribution of the soft-labeled tuning set was limited to 20% of the weight assigned to the laboratory440 derived source training data. This follows the common domain adaptation strategy of down-weighting uncertain samples and treating their influence as a regularization term rather than as fully trusted labels (e.g., Kouw and Loog 2021). A sensitivity test determined that 20% influence by the tuning samples was conservative enough to preserve accuracy in in-sample testing with the training data while leading to significant changes in classification behavior; hence, we chose this degree of tuning to investigate the effects of domain adaptation on the classification algorithm.”

411: *Interferents in HFPs are not directly visible in Fig. 7A. You could include the interferents concentration in Fig 7A if it does not become too messy. Alternatively, explain again that interferents are HFPs minus fPBAPs.*

We have elected to reword the text for clarity (L445-453) and to clearly point out that the main finding is the annual cycle of identified fPBAPs in order to avoid confusion about the interferent concentrations, which we feel would complicate panel Fig. 7a. However, we have revised Fig. 7b to include the interferents class as in the contributions to HFPs, to make it more clear that HFPs are divided into interferents and the fPBAP classes. We have also revised the discussion around this figure, which we believe better reflects, including a reference to Freitas et al. (2023) where a second peak can be seen in INP concentrations not captured by the DT method, which the presented algorithm captured (L445-461):

“The tuned fluorescent particle classifier reproduces the DT-based fPBAP concentration time series (Freitas et al., 2023a), with few ($\sim 10^{-3}$ - 10^{-1} L⁻¹) fPBAP during winter and early spring, a peak in July, and a subsequent decrease during winter onset (Fig. 7a-b). The high proportion of HFPs labeled as interferents between January and April (Fig. 7a), along with significant correlations between the pollution class and lead eBC, and levoglucosan (Fig. 7e), shows that the classification algorithm robustly reduces the influence of combustion particles. In August, when biological tracer concentrations were highest, our algorithm detected fPBAP concentrations nearly two orders of magnitude higher than the DT method did. This is expected because the DT-based fPBAP criterion is validated for one type of fPBAP, while our classification algorithm targets multiple fPBAP types with diverse fluorescence properties. Using scanning electron microscopy, Tobo et al. (2024) determined that most carbonaceous coarse-mode particles sampled at Zeppelin Observatory in July 2020 were likely PBAPs (such as microorganisms and plant debris), while none were detected in March 2020. The low (high) proportion of HFPs classified as fPBAPs during March (July) reflects a similar seasonal fPBAP cycle as that described in Freitas et al. (2023a). Both our algorithm’s and DT based fPBAP concentrations correlate significantly with BB tracers (Fig. 7e), in agreement with Freitas et al. (2023a).”

→ “The tuned fluorescent particle classifier reproduces the DT-based fPBAP

concentration time series (Freitas et al., 2023a), with few ($\sim 10^{-3}$ - 10^{-2} L⁻¹) fPBAP during winter and early spring, a peak in July ($\sim 10^1$ L⁻¹), and a subsequent decrease during winter onset (Fig. 7a-b). This cycle is also in agreement with Tobo et al. (2024), who determined that most carbonaceous coarse-mode particles sampled at Zeppelin Observatory in July 2020 and analyzed with scanning electron microscopy were likely PBAPs (such as microorganisms and plant debris), while none were detected in March 2020. Compared to the DT-based method, the ML-based classifier also better captures a second peak in PBAP tracers during October that coincides with a significant increase in warm-activating INPs (Fig. 1c of Freitas et al. 2023a). The low proportion of HFPs identified as fPBAP between January-April (Fig. 7a) illustrates that the tuned classification algorithm robustly reduces the influence of combustion particles on fPBAP concentrations.”

412: There is no “pollution class” in Fig. 7e. Do you mean combustion? Be consistent! What does “lead” mean in lead eBC?

We have now revised the manuscript to use the “pollution/interferents” class in figures (Figs. 5-7) and throughout the text, and we’re thankful for pointing out the inconsistencies. Lead was included as a tracer for pollution in an earlier analysis; this was dropped to streamline the analysis, and the word somehow slipped through. However, the phrase pointed out has been removed to focus on the PBAP annual cycle (see previous comment).

413: I would add “on fPBAP” after “reduces the influence of combustion particles”.

Added as suggested (L453).

415: For what type is it validated?

The decision tree method was validated primarily for marine bacteria; this has now been specified in the sentence (L456-57).

“This is expected because the DT-based fPBAP criterion is validated for one type of fPBAP, while our classification algorithm targets multiple fPBAP types with diverse fluorescence properties.”

→ “This is expected because the DT-based fPBAP criterion is validated for one type of fPBAP (marine bacteria; Freitas et al., 2022), while our classification algorithm targets multiple fPBAP types with diverse fluorescence properties.”

424: State the correlation coefficients here.

Added (L463-465):

“This has implications for the interpretation of fructose and glucose as PBAP tracers, as illustrated by their correlations with levoglucosan (Fig. S8).”

→ “This has implications for the interpretation of fructose and glucose as PBAP tracers, as illustrated by their correlations with levoglucosan (Spearman $\rho = 0.56$ and 0.59 , 465 respectively; Fig. S7).”

446: Name the SSA and dust tracers.

These have now been specified (L453-454).

“Figure 7d reveals that the untuned dust LRM is also insensitive to variability in SSA and dust tracers, indicated by the nearconstant, high fraction ($\sim 80\%$) of particles labeled as

dust.”

→ “Figure 7d reveals that the untuned dust LRM is also insensitive to variability in SSA and dust tracers (sodium and silicon, respectively), indicated by the near-constant, high fraction (~80%) of particles labeled as dust.”

Technical comments:

24: < - 20 °C (the minus sign is missing)

Added (L27).

161: “of” missing: inlet OF the MBS.

This was grammatically correct, but this sentence has now been adjusted to avoid confusion (L166-168).

“The pollution events were observed during an Arctic Ocean research cruise, when the ship’s plume was sampled directly by the inlet housing the MBS.”

→ “The pollution events were observed during an Arctic Ocean research cruise when the ship’s exhaust plume was transported directly to the inlet from which the MBS was sampling.”

220: “and the” is written twice.

Fixed (L242).

240: you just refer to the SI here. Which part?

This has now been specified (Section S1) (L262):

“The details of the domain adaptation implementation are further explained alongside our assessment of the initial classification algorithm in Section 3.4 and in the SI (Section S1).”

326: Citation should be in brackets.

Fixed (L353).

364: Missing full stop after “classes”.

This sentence has been adjusted and the full stop was included (L370-372).

“Figure 6a shows UMAP projections for fluorescent particles using only pollen, bacteria, fungal spores, and combustion particles for training, illustrating clear separation between combustion particles and all fPBAP classes; the clearer the separation between clusters, the more certainty can be obtained 365 in discriminating between them.”

→ “Figure 6a shows UMAP projections for fluorescent particles using only pollen fragments, bacteria, fungal spores, and combustion particles for training, illustrating clear separation between the combustion particles and all fPBAP classes.”

405: Where in the SI?

This has now been specified (L412).

“This is accomplished by choosing a number of sample periods with sufficient tracer

data and labeling HFPs observed during these periods as either pollution or not based on the soft label, and further training the LRM on these data (the tuning block) with lower weights than those of the initial training data to limit its influence on model fit; details of this step are explained the SI.”

→ “This is accomplished by choosing a number of sample periods with sufficient tracer data and labeling HFPs observed during these periods as either pollution or not based on the soft label, and further training the LRM on these data (the tuning block) with lower weights than those of the initial training 435 data to limit its influence on model fit; details of this step are explained in the SI (Section S1).”

461: *Where in the SI?*

This has now been specified (L494).

“... details of our dust influence metric’s construction are further explained in the SI (Section S1).”

465: *should be “repeated”.*

This was fixed in the reformulated sentence (L472).

“The model’s overall precision in identifying dust and SSA in laboratory data in r epeated in-sample testing was not significantly affected by the tuning (Fig. S9).”

→ “The model’s overall precision in identifying dust and SSA in repeated in-sample testing with laboratory data was not significantly affected by the tuning (Fig. S9).”

Supplement:

Page 8, Fig. S3: It would make more sense to add the starting wavelength of each channel instead of random 100 nm steps on the lower x-axis. Add the excitation wavelength in the caption.

In figure form, using the starting/ending wavelengths does not make the figure more readable. However, the initial version of the figure was more rough and did not take into account the effects of the diffraction grating, which causes each channel’s span to be ± 15 nm rather than neat transitions to neighboring channels. The revised Fig. S1 has now been updated to be exactly as the instrument is described in the manual, and the channel spans have now been clarified in the caption, as well as the excitation wavelength.

S4: a and b are switched!

Fixed in the revised Fig. S2, we are thankful for the catch.

Table S3: You should include what these substances are tracers for in the table.

This is a good idea! The aerosol classes traced were now added to the revised Table S3.

S5: One bar is always orange instead of red.

Fixed in the revised Fig. S5.

Response to RC2

I do not understand how chemical tracers were used in domain adaptation and how it helped testing possibility that model trained in lab can be used in real life field observation. More detailed discussion on expected performance of the MBS coupled with machine learning algorithm for identification listed bioaerosols in real life field measurements is needed.

We are thankful for the reviewer's comment that the expected performance and use must be clearly discussed in the manuscript. This discussion was present in the original manuscript, but we see a need to consolidate and clearly summarize these takeaways between the Discussion and Conclusions sections. Taken together with the conservative performance of the decision tree (DT) method and the assessed performance of the machine learning (ML) method, each method has its strengths and limitations, and researchers must consider them when choosing a method for identification in their applications. This discussion is an important one to be had in the field of online bioaerosol detection, and a suggestion to focus the discussion regarding this in Referee Comment 1 was also made. We have adjusted the Discussion section to more clearly discuss the expected performance (L539-583) in the revised manuscript, and rewritten the Conclusions section to include recommendations and considerations (L598-623) (see responses to comments 1 for Conclusions passages, and 3 for Discussion passages regarding classification).

If the purpose of the measurements are to enable automatic quantification of different airborne coarse particles then I think the methodological approach limits it notably. Sampling only 0.33 l/min (and analyzing half of it) is hardly enough to representatively detect and quantify rare airborne particles such as pollen and fungal spores. If you check the aerobiological studies, atmosphere only in rare events contains more than 5000 pollen grains/m³ and more than 50000 fungal spores/m³ (most often it is even a magnitude less). This way

We understand that there are limitations to single-particle characterization methods such as UV-LIF when it comes to identifying and quantifying extremely rare atmospheric particles, constrained by e.g. measurement and sample flow rates. However, UV-LIF instrumentation with comparable sampling rates (such as the WIBS; e.g., [Hughes et al., 2020](#)) has provided robust quantifications of PBAP concentrations in many situations (e.g. and agreed well with offline analyses using techniques analyzing greater air volumes and particle numbers (e.g. [Negron et al., 2020](#)). We are also not sure if there was a misunderstanding on the sampling rate; over 100 particles per second can be fully measured with all fluorescence and optical scattering parameters, and at an example flow rate of 0.33 L min⁻¹, this corresponds to a maximum number concentration of particles able to be measured by the MBS (i.e. in the correct size range) of several 10⁴ L⁻¹. This level of coarse-mode concentration is in the upper range of ambient concentrations in many settings, such as in Saharan dust events that were observed in a recent campaign with the MBS at a mountaintop observatory. At the example flow rate given, 0.33 L min⁻¹ analyzes roughly 20 L of air per hour; this may not be enough to observe and quantify rapid fluctuations or events on timescales <1 h, but is useful for measuring variability on hourly or greater timescales, with limits of detection of 50 m⁻³ for 1 h, and less the longer the averaging period. Finally, the concentrations of total fluorescent PBAPs identified by our algorithm agree on the orders of magnitudes seen in previous work on quantifying PBAPs in Ny-Ålesund using offline methods (10¹ L⁻¹ during July-August; [Johansen & Hafsten, 1986](#)), indicating that the MBS can sufficiently analyze enough particles to come close to methods sampling larger volumes of air.

We also found a discrepancy with the maximum rate of measurement reported, which was

that of a different version of the instrument; this error has now been corrected and it is now explicitly stated that this is enough to characterize most particles in most ambient conditions, which we find helps reduce confusion about the instrument's capabilities. We are thankful for the reviewer pointing out a need for the instrument capabilities to be concretely stated. The changes can be found at lines L102-104:

"The maximum particle measurement rate is also limited by the xenon lamp's recharge time, yielding an upper rate of about 100 particles s⁻¹."

→ "The maximum particle measurement rate is also limited by the xenon lamp's recharge time, yielding an upper rate of about 160 particles s⁻¹ (Freitas et al., 2022), enough to characterize most particles in even high ambient coarse-mode particle concentrations."

The authors claim pollen rupture is the reason for detecting large variety of particles shapes and sizes. It should be noted that pollen grains exine is not prone to rupturing in normal conditions. In the atmosphere it actually rarely happens and seen variability should rather be attributed either to irregular shapes of dry pollen grains or their non-homogenous structure (cytoplasm with different water and chemical content. sometimes even air trapped in exine) as it is suggested in a study that tested Rapid-E+ air flow cytometer (Sikoparija et al. 2024). In fact those authors argued that due to wrinkly and inhomogeneous pollen cytoplasm function for size estimation usually derived from measurements PSL does not fit.

I do not agree if the authors with "wet nebulized pollen show more spherical morphologies, that many particles are dried residues of dissolved constituents rather than physical fragments" claim that dry pollen is fragmented in air. Dry pollen is wrinkly which might make it appear as irregular fragment in the scattering signal, but it is still an intact particle that should be discriminated from other coarse particles of similar size.

We do not argue that dry pollen is fragmented in ambient air in the real world, nor does the experimental setup used for the characterization experiments allow for the potential real-world mechanisms of this to be tested. That some degree of fragmentation must have occurred during the experiments is highly likely given that the size range of particles able to be measured by the MBS is well below that of most whole pollen grains (the smallest size denoted by the company who prepared our samples was 18 μm), as is stated in the description of results for the pollen experiments (L253-255). Most particles we observed in the experiments were in the 3-6 μm range. While pollen grains are known for their elasticity and foldability, enabling them to fold into smaller volumes upon dehydration as an evolutionary strategy for surviving moisture fluctuations ([Katifori et al., 2010](#)), a reduction in mode optical diameter from the >20 μm sizes reported for our samples to 3-6 μm is unlikely, based off of pollen size reductions observed in electron microscopy studies ([Halbritter & Hesse, 2003](#); [Ermolaev et al., 2024](#)). Fragmentation during our experiments may have occurred due to acceleration forces/impact with the sample holding tube walls during dry nebulization, or in the water-pollen solution prepared for wet nebulization (as the resulting particles would first be emitted as droplets and then evaporated into particle phase).

Given this, pollen rupture in the atmosphere is still a potentially important process, which has potentially important impacts on health and climate, and has been observed previously in ambient conditions (e.g., [Hughes et al., 2020](#); [Mampage et al., 2025](#); [Zhang et al., 2025](#)) and reproduced with reasonable ease in laboratory conditions (e.g., [Stone et al., 2021](#); [Taylor et al., 2004](#)). However, the extent to which this occurs in the atmosphere is unknown, further motivating studying how they may look in ambient observations with online methods.

To help clarify these points, we have now reformulated much of the discussion around pollen in the revised manuscript to maintain a critical tone towards pollen fragment identification capabilities (L466-477 in the Results and L545-567 in the Discussion; see also our response to comment 3 of RC1 for the changed passages). We have also replaced the term “sub-pollen particle” with “pollen fragment” throughout the manuscript, better reflecting that it is not known which exact components were nebulized and measured in our experiments. The revised Discussion section now clearly and critically discusses pollen classification difficulties (L545-572):

“Our classifier identified pollen fragments and fungal spores as the most common classes of fPBAPs observed at Zeppelin Observatory in 2020. While previous work establishes fungal spores as common in Ny-Ålesund (Johansen and Hafsten, 1988), consistent with the prevalence of fungal spores identified by our classifier and the analyses of Freitas et al. (2023a) and Tobo et al. (2024), whole pollen concentrations were very low ($\sim 10^{-3}$ – 10^{-1} L⁻¹ in summer; Johansen and Hafsten 1988). PBAP sampled during the NASCENT campaign analyzed by Freitas et al. (2023a) and Tobo et al. (2024) with electron microscopy could not be precisely distinguished, although visual comparison suggested that they were likely fungal spores and bacteria. Tobo et al. (2024) also found irregular carbonaceous particles resembling plant debris or possibly pollen fragments. Additional tracers could help verify class contributions to fPBAP detected by our classification algorithm, but source overlap complicates tracer interpretation. Although arabinol and mannitol are considered fungal tracers, they can be found to lesser degrees in pollen and bacteria (Lau et al., 2006; Di Filippo et al., 2013). Wildfires have also been found to directly promote PBAP emissions (Holden et al., 2011; Moore et al., 2020; Kobziar et al., 2022; Ellington et al., 2024), thus it is not unexpected that BB particles and PBAP can co-exist in ambient observations. More source-specific tracers, such as sucrose for pollen (Mampage et al., 2022), dipicolinic acid for bacteria (Mampage et al., 2022), and ergosterol for fungal spores (Lau et al., 2006), are necessary to thoroughly validate PBAP class contributions.

Pollen fragments are difficult to robustly identify and quantify in ambient conditions due to high variability in their emission pathways (Suphioglu et al., 1992; Subba et al., 2021). Our results suggest that BB aerosols may contain particles with strong similarities to pollen fragments, as evidenced by the stronger correlations between BB tracer levoglucosan concentrations and this particle class relative to other fPBAP classes, and by their near-constant presence throughout winter in Svalbard when fluorescent aerosols are likely comprised mostly of combustion-derived particles originating from continental sources (Yttri et al., 2024). Including BB aerosols in future experiments may therefore improve fPBAP and, in particular, pollen fragment identification in our classification algorithm. Based on the apparent similarities in fluorescence characteristics between other fPBAP classes and cellulose assessed in our experiments, it is also possible that many of the particles in the Zeppelin Observatory dataset identified as fPBAP by our classifier were misclassified plant debris. Since real plant debris was missing from the characterization data presented here, representing a potentially important missing class of a globally abundant PBAP (Sánchez-Ochoa et al., 2007; Winiwarter et al., 2009), we suggest that future characterizations measuring e.g., leaf litter be conducted with the MBS to improve identification. Future assessments for specific fPBAP class validations when e.g., pollen trap counts, digital holography-based pollen identification, and aerobiome DNA sequencing are conducted in parallel to MBS observations should be made to provide further validation.”

In my opinion the presented results indicated that MBS measurements have a very limited applicative value for discriminating between tested bioaerosols. From comparison of fluorescence signals between dry and wet nebulized pollens one could see notable differences in signals (and large variability). This indicates that only change in humidity during real life sampling could change signal characteristics to the degree it confuses between different pollen classes. And possibly even between different aerosols. Although model validation results presented in Table 2 indicate models in laboratory conditions sees some differences between classes, in real life field conditions when facing numerous unseen particles the accuracy notably drops.

We acknowledge the reviewer's concerns about applicability in real-world conditions, which we believe became apparent during the validation exercise using the Zeppelin observations. This is the motivation for applying a domain adaptation step to the models, which we wanted to test as a potential solution for the issue of generalizing results from laboratory measurements to real-world conditions. Because of the inherent difficulty of exactly duplicating real-world conditions in the laboratory, we believe that a combined approach is motivated. However, we do not believe that it solves all problems of identification using UV-LIF or the MBS in observations, and introduces new considerations. This concern about classification performance is shared with the first reviewer, and based on their comments as well as those of the second reviewer, we recognize the need to ensure that the manuscript contains a more critical discussion of limitations and guidance for the readers.

To this end, we have now reformulated parts of the paper to maintain a critical tone towards class-specific bioaerosol identification, and to clearly communicate caveats and weaknesses for the classification tools presented in this study. This includes the Discussion section, which we have revised to critically discuss reasons for misclassification and potential future solutions to this end (L532-544):

"We demonstrated the training of ML models on the combined body of MBS source experiment data using UMAP dimensionality reduction and k -nearest neighbors. In in-sample testing, the classifier displayed confusion between bacteria and fungal spores, and between pollen fragments and fungal spores. Problematically, the multiclass fluorescent particle classifier misidentified ~12% of combustion particles as pollen fragments. For more robust filtering, we trained a separate binary classifier using logistic regression for flagging particles as combustion-derived interferents with the same training data, which performed better at distinguishing interferents from fPBAPs. Combining these two methods provides flexibility and builds redundancy into the classifier. The ease of tuning the LRM allowed us to implement domain adaptation using field validation data, improving correlations between our algorithm's fPBAP estimates and PBAP tracer concentrations. Significant correlations between fPBAP concentrations derived with the untuned classifier and wood pyrolysis tracer (levoglucosan) concentrations indicates that particles originating from BB were likely misidentified as fPBAPs. These may differ from diesel engine-sourced combustion particles like those included in our training data and could have been modified by aging processes. Our classification algorithm may be improved in the future by including source characterization data from BB experiments; we suggest that the models comprising it be trained from scratch and re-assessed when such data are available before further tuning is applied."

As well as the concluding paragraphs, which report the weaknesses we found in the assessment and point to future directions (L598-623):

“Using the combined fluorescence and morphology characterization data, we developed a particle classification algorithm using supervised ML trained on known particle types and applied it to one year of field observations from the Zeppelin Observatory, Svalbard. Compared to the previously used fluorescence-only approach, our algorithm yields up to two orders of magnitude higher fluorescent PBAP concentrations during summer, in agreement with previously reported concentrations of PBAPs in Ny-Ålesund during July-August (Johansen and Hafsten, 1988), while reproducing the annual PBAP cycle reported in Freitas et al. (2023a). Although the algorithm identified fungal spores as a major type of PBAP present at Zeppelin Observatory, in line with previous studies (Johansen and Hafsten, 1988), it clearly overestimated pollen fragment concentrations, which may be attributable to similarities shared with plant debris and possibly biomass burning particles. In addition, we trained a dust detection model by exploiting optical morphology differences between dust and SSA particles, which may be used to help determine marine and terrestrial influence in ambient observations. Although laboratory-generated dust and SSA particles were readily distinguished, classification of ambient SSA proved challenging. This suggests significant differences between laboratory-generated and natural aerosols that may stem from environmental processing, motivating improvements in laboratory aerosolization and characterization methods.

Compared to fluorescence-only approaches, the ML framework explicitly incorporates morphology and accounts for a broader range of PBAP classes as well as major non-biological interferents. Domain adaptation further improved performance, although both ML-based and fluorescence-only methods likely remained sensitive to combustion aerosols. It is unlikely that any classification scheme or instrument configuration will fully eliminate the influence of interferents, highlighting the need for parallel observations of e.g., black/brown carbon and source-specific chemical tracers for assessing fPBAP quantifications. Despite site-specific variability in aerobiomes and fluorescence signatures, the framework provides a flexible and expandable basis for estimating PBAP concentrations across environments and reanalyzing existing MBS datasets. However, additional training data are clearly needed to improve both interferent and specific PBAP class identification before it may be suitable for use as a standalone tool. Its continued development will require further validation with field observations and more characterization studies with e.g. biomass burning, pollen emission, fungal sporulation, and atmospheric aging in controlled conditions. The effects of implicating morphology information in fPBAP identification may also be further explored as new characterization data with more realistic aerosolization mechanisms become available. Coordinated collection and open sharing of such reference datasets across research groups will be essential for improving identification methods.”

These changes also includes the abstract, particularly the second paragraph, which we have now revised in order to convey a realistic and critical tone towards the results of the algorithm assessment (L11-19):

“To explore solutions for these confounding factors, we develop a supervised machine learning classification algorithm integrating fluorescence and morphology information, and evaluate it using MBS and comprehensive chemical tracer observations from Zeppelin Observatory, Svalbard. This evaluation illustrates challenges for inductive methods in distinguishing biomass burning from biological particles, and dust from SSA, suggesting that important particle classes may be missing and/or laboratory-generated aerosols significantly differ from real-world counterparts. We show that domain

adaptation using complementary observations can help address these difficulties. Compared to a fluorescence-only approach, the domain-adapted algorithm reproduces the previously published annual bioaerosol cycle while yielding higher summertime concentrations matching those reported from offline analyses. This open-source algorithm provides a basis for quantifying bioaerosols across diverse environments and can be refined with future field and laboratory efforts.”

We also changed our naming of the exercise applying the algorithm to the field observations at Zeppelin from “validation” to “assessment”, and replaced “validation data” with “auxiliary data” (Sections 2.6 and 3.4 titles, L236, L255, L259, L262, L375, L538). This reflects that it is not a final product and that the algorithm may be improved in the future, and is honest to its limitations. In addition, we explicitly stated that care must be taken when interpreting correlations with tracer data, since they were used in tuning; however, we do find the correlations important to report as a means of judging the algorithm. The changes can be found at L453-455 and L502-504, respectively:

“This may be reflected in the correlations between fPBAP detected by our algorithm and BB tracers (Fig. 7e), and between cellulose and levoglucosan (Fig. S8).”

→ “This is reflected in improved correlations between fPBAP concentrations with biological tracers arabitol and mannitol (Fig. 7e and Fig. S6); however, these correlations do not constitute fully independent validation of the adapted model and should be interpreted cautiously due to their use in the tuning process.”

“This results in a significant positive correlation with silicon and anti-correlation with SSA tracers sodium, chloride, and magnesium. These correlations are stronger than those from the untuned model’s output (Fig. S7b).”

→ “Although correlations do not independently validate the model, this adaptation results in a significant positive correlation with silicon and anti-correlation with SSA tracers sodium, chloride, and magnesium, which are strengthened relative to the untuned model’s output (Fig. S7b).”

Further changes

- 1) The first paragraph of the Abstract was slightly reworded for simplicity and clarity:

“Coarse-mode aerosol particles influence the environment, climate, and human health in diverse ways depending on their type. While mineral dust and sea spray aerosol (SSA) dominate this size range, rarer biological particles can have outsized impacts, for example by initiating ice formation at relatively warm temperatures. Hence, accurate, type-specific characterization of coarse-mode aerosol is essential for understanding their roles in climate and the environment. Using laboratory measurements of single-particle ultraviolet light-induced fluorescence (UV-LIF) spectroscopy and morphology, we provide a new reference dataset for coarse-mode aerosols from common sources, including pollen, dust, bacteria, and microplastics. Comparisons with previously published datasets reveal consistent source-dependent fluorescence features, but also highlights similarities between biological and non-biological particles that can bias classifications based on fluorescence alone.”

→ “Coarse-mode aerosol particles influence the environment, climate, and human health in diverse ways depending on their type. While mineral dust and sea spray aerosol (SSA) dominate this size range, rarer biological particles can have outsized impacts, such as initiating hydrometeor freezing at relatively warm temperatures. Accurate, type-specific characterization of coarse-mode aerosol is therefore essential for investigating their roles in climate and the environment. Using laboratory-based single-particle ultraviolet light-induced fluorescence spectroscopy and morphology measurements made with a Multiparameter Bioaerosol Spectrometer (MBS), we provide a new reference dataset for coarse-mode aerosols from common sources, including pollen, dust, bacteria, and microplastics. Including published datasets revealed consistent source-dependent fluorescence features, but also highlighted similarities between biological and non-biological particles that can bias fluorescence-based classifications.”

- 2) A discrepancy was found in the SI; Fig. S2 of the original draft was a figure from a previous version of the dust influence tuning step using an alternate dust mass concentration estimate, and was incorrect. The correct version with weekly mineral dust concentrations using the formula in the supplementary text (eq. 6) is now implemented in the SI (Fig. S4). We subsequently ensured consistency between figures using representative days with low dust influence based on the metric to be the same, as the original draft used different days chosen with different metrics. It is simpler to choose the representative day as the period of wintertime minimum in dust influence. This has now been specified in the supplementary figures with representative coarse-mode particles from this day (Figures S8-S9) and specified in the text (L485-486):

“Inspecting non-fluorescent particles measured at Zeppelin Observatory during a sampling period with minimal dust influence (the minimum in silicon concentrations over 2020) reveals that they resembled dust more than SSA in laboratory characterization data (Fig. S6), suggesting that natural processes (e.g., aging of SSA) may have modified particle morphology.”

→ “Inspecting non-fluorescent particles measured at Zeppelin Observatory during a sampling period with minimal dust influence (the wintertime minimum of dust influence relative to SSA tracers; see Section S1) reveals that they resembled dust more than SSA when compared to laboratory characterization data (Fig. S8), suggesting that natural processes (e.g., aging of SSA) may have modified particle morphology.”

- 3) We decided to be more internally consistent with the defined algorithm class “pollution” in the UMAP + kNN multi-class classifier and replaced the “combustion” class with “pollution/interferents”. This is to remain flexible for possible future changes to the algorithm where other interferents or classes not implemented in this version may be binned into this class, e.g., biomass burning or SOA. Figures 5-7 and Table 2 were adjusted with this change, and instances in the text with “combustion class” were replaced accordingly (L394, L405, L536).
- 4) All instances of the mathematical formulation for “order of” (e.g. $O(10^{-1})$) were changed to approximate orders of exponents (e.g. “ $\sim 10^{-1}$ ”) for simplicity (L172, L306, L446). Units were reformatted in LaTeX for typesetting neatness, hence changes seen after numbers (e.g. $L \text{ min}^{-1}$ at L96 in the diff file).
- 5) The data sets have been officially published on the Bolin Centre Database and have been cited in the Code and data availability section (L624-626); these can be found at <https://bolin.su.se/data/jonsson-2026-aerosol-mbs-raw> for the raw MBS output, and <https://bolin.su.se/data/jonsson-2026-aerosol-mbs> for processed data.

Further specific grammatical, formatting, and typo correction changes include:

- L49: Changed for readability: “coenzymes such as riboflavin” → “the coenzyme riboflavin”
- L90: “Xenon” → “xenon”
- L90: Specified: “280 nm light to excite ...”
- L100-101: We did not use arbitrary units for detected fluorescence, so the following sentence was changed:

“Fluorescence intensity is not measured in physical units; we report absolute fluorescence as a detector signal in arbitrary units (au).”
 → “Fluorescence intensity is not measured in physical units; we report fluorescence intensities as fractions normalized by detector signals.”

Fluorescence intensity is not measured in physical units; we report fluorescence intensities as fractions normalized by detector signals.

- Italicized all references to MBS channels (L126, L131-132,
- L133-134: For more general application, reformulated the following sentence:

“The samples characterized with the MBS in this study are selected to represent aerosols relevant to the Northern Hemisphere, including boreal/tree pollen, marine bacteria, dust, and synthetic contaminants (Table 1).”
 → “The samples characterized with the MBS in this study are selected to represent common sources of aerosols potentially relevant in many environments, including tree pollen, marine bacteria, dust, and synthetic contaminants (Table 1).”

- L135: Removed “debris” from “plant matter debris” for generalization and less redundancy
- L163: Fixed units: “5 ml” → “5 mL”
- L196: Changed for consistency: “classify fluorescent particles into broad bioaerosol subgroups, and;” → “classify fluorescent particles into broad PBAP subgroups, and;”
- L202: Fixed spelling for “accomodate”
- L206: For specificity: “Our LRMs” → “These LRMs”
- L213: For fairness, as more examples exist: “(Crawford et al. 2020, 2023)” → “(e.g. Crawford et al. 2020, 2023)”

- L256: Fixed the order of words for correctness: “Domain adaptation is a ML technique for transfer learning” → “Domain adaptation is a transfer learning technique for ML”
- L265: For specificity: “Emission spectra and size distributions for dry and wet pollen samples” → “Emission spectra and size distributions measured in dry and wet pollen experiments”
- L282: To refer to the results specifically: “Such particles would tend to show more droplet-like scattering signatures ...” → “These particles show more regular, symmetric, droplet-like scattering signatures ...”
- L283-285: Fixed for grammar:

“The especially strong wet-dry contrast in willow pollen fluorescence (Fig. 1g and 1n) may be due to acetone de-fatting (Table S1), as removing surface lipids likely increase water uptake and solubility, promoting rupture and altering fluorescence.”

→ “The especially strong wet-dry contrast in fluorescence seen with willow pollen (Fig. 1g and 1n) may be due to acetone de-fatting 285 (Table S1), as removing surface lipids likely increases water uptake and solubility, promoting rupture and altering fluorescence.”

- L297: Specified: “... contribute to pollen fluorescence signals ...”
- L300: Fixed: “contrast” → “contrasts”
- L306: Used defined abbreviation for HFPs
- L311: Specified “fluorescence intensities”
- L313: Fixed the logical inconsistency: “... and often peak in both channels B and C ...” → “... and peak in either channels B or C ...”
- L317: Specified: “complications” → “interference by combustion particles”
- L333: For more physical interpretation: “bacterial B channel peaks” → “bacterial emission spectra”
- L329-332: Reformulated for correctness and easier reading:

“Accordingly, pollen, and to some extent fungal spores, tend to have more irregular morphologies, whereas bacteria and intact spores exhibit more regular scattering signatures (Fig. 4c-f).”

→ “Accordingly, pollen fragments, and to some extent fungal spores, may be expected to have more irregular morphologies whereas bacteria and intact spores may be expected to exhibit more regular scattering signatures; these behaviors are reflected in their measured properties (Fig. 4c-f).”

- L335: Removed “rather than physical fragments” due to logical inconsistency
- L337: Fixed grammar: “differ clearly” → “clearly differ”
- L373-375: Reformulated for specificity on the point of using the threshold:

“A fluorescent particle’s p_{pol} must not surpass a confidence threshold (α_{pol}) for it to be classified as a FPBAP.”

→ “A fluorescent particle’s p_{pol} must not exceed a confidence threshold (α_{pol}), which may be chosen depending on confidence requirements and auxiliary data in each application setting, for it to be classified as a FPBAP.”

- L382: Added reference to Table S2.
- L387: For specificity: “microplastics and cellulose” → “microplastic and cellulose particles”
- L396: Removed “the clearer the separation between clusters, the more certainty can be obtained in discriminating between them” to avoid repetition
- L408: Added for specificity: “fluorescence curves and irregular morphologies”
- Table 2 caption: Grammar: “predicted as” → “predicted to be in”

- L432: For better wording: "biological dominance" → "biological influence"
- L498: For clarity: "laboratory data in repeated in-sample testing" → "in repeated in-sample testing with laboratory data"
- L529-530: Simplified:

"Notably, combustion particles showed broad fluorescence spectra similar to those of pollen and fungal spores with generally substantially stronger intensities, posing a major challenge in UV-LIF-based PBAP detection."

→ "Notably, combustion particles showed broad fluorescence 530 spectra similar to those of many fPBAP types, posing challenges for UV-LIF-based PBAP detection."

- L530-531: Simplified by removing "fPBAP", and specified "bacterial emission spectra". Removed to avoid repetition: "Dust particles exhibited distinct scattering properties, indicating higher asymmetry and irregular, rough surfaces than SSA."
- L537: Replaced "fluorescent PBAP" with the defined abbreviation.

References (in order of appearance)

- Freitas, Gabriel P., et al. "Emission of primary bioaerosol particles from Baltic seawater." *Environmental Science: Atmospheres* 2.5 (2022): 1170-1182. DOI: <https://doi.org/10.1039/D2EA00047D>
- Pereira Freitas, Gabriel, et al. "Regionally sourced bioaerosols drive high-temperature ice nucleating particles in the Arctic." *Nature Communications* 14.1 (2023): 5997. DOI: <https://doi.org/10.1038/s41467-023-41696-7>
- Huang, Shu, et al. "Overview of biological ice nucleating particles in the atmosphere." *Environment International* 146 (2021): 106197. DOI: <https://doi.org/10.1016/j.envint.2020.106197>
- Després, Viviane R., et al. "Primary biological aerosol particles in the atmosphere: a review." *Tellus B: Chemical and Physical Meteorology* 64.1 (2012): 15598. DOI: <https://doi.org/10.3402/tellusb.v64i0.15598>
- Reichelt, Sophia, and Elena Gorokhova. "Aggregation in experimental studies with microparticles: Experimental settings change particle size distribution during exposure." *Environmental Pollution* 336 (2023): 122369. DOI: <https://doi.org/10.1016/j.envpol.2023.122369>
- Ho, Derek, et al. "Glow and behold: How weathering alters the surface, chemical and fluorescence properties of microplastics." *Microplastics and Nanoplastics* (2026). DOI: <https://doi.org/10.1186/s43591-026-00198-1>
- Pöhlker, Christopher, J. Alex Huffman, and Ulrich Pöschl. "Autofluorescence of atmospheric bioaerosols—fluorescent biomolecules and potential interferences." *Atmospheric Measurement Techniques Discussions* 4.5 (2011): 5857-5933. DOI: <https://doi.org/10.5194/amt-5-37-2012>
- Huffman, J. Alex, et al. "Real-time sensing of bioaerosols: Review and current perspectives." *Aerosol Science and Technology* 54.5 (2020): 465-495. DOI: <https://doi.org/10.1080/02786826.2019.1664724>
- Kouw, Wouter M., and Marco Loog. "A review of domain adaptation without target labels." *IEEE transactions on pattern analysis and machine intelligence* 43.3 (2019): 766-785. DOI: <https://doi.org/10.1109/TPAMI.2019.2945942>
- Hughes, Dagen D., et al. "Characterization of atmospheric pollen fragments during springtime thunderstorms." *Environmental science & technology letters* 7.6 (2020): 409-414. DOI: <https://doi.org/10.1021/acs.estlett.0c00213>
- Negron, Arnaldo, et al. "Using flow cytometry and light-induced fluorescence to characterize the variability and characteristics of bioaerosols in springtime in Metro Atlanta, Georgia." *Atmospheric Chemistry and Physics* 20.3 (2020): 1817-1838. DOI: <https://doi.org/10.5194/acp-20-1817-2020>
- Johansen, Stein, and Ulf Hafsten. "Airborne pollen and spore registrations at Ny-Ålesund, Svalbard, summer 1986." *Polar Research* 6.1 (1988): 11-17. DOI: <https://doi.org/10.3402/polar.v6i1.6842>
- Katifori, Eleni, et al. "Foldable structures and the natural design of pollen grains." *Proceedings of the National Academy of Sciences* 107.17 (2010): 7635-7639. DOI: <https://doi.org/10.1073/pnas.0911223107>
- Halbritter, Heidemarie, and Michael Hesse. "Principal modes of infoldings in tricolp (or) ate Angiosperm pollen." *Grana* 43.1 (2004): 1-14. DOI: <https://doi.org/10.1080/00173130310017625>
- Ermolaev, Aleksey, et al. "A simple and user-friendly method for high-quality preparation of pollen grains for scanning electron microscopy (SEM)." *Plants* 13.15 (2024): 2140. DOI: <https://doi.org/10.3390/plants13152140>
- Mampage, C. B. A., et al. "Hot and dry conditions elevate grass pollen and sub-pollen particle concentrations in Melbourne, Australia." *Environmental science: atmospheres* 5.10 (2025): 1081-1098. DOI: <https://doi.org/10.1039/D5EA00024F>
- Zhang, Hao, et al. "Data-Driven Detection of Nocturnal Pollen Fragmentation Triggered by High Humidity in an Urban Environment." *Environmental Science & Technology* 59.25 (2025): 12763-12774. DOI: <https://doi.org/10.1021/acs.est.4c13905>
- Stone, Elizabeth A., et al. "Airborne sub-pollen particles from rupturing giant ragweed pollen." *Aerobiologia* 37.3 (2021): 625-632. DOI: <https://doi.org/10.1007/s10453-021-09702-x>
- Taylor, Philip E., et al. "Birch pollen rupture and the release of aerosols of respirable allergens." *Clinical & Experimental Allergy* 34.10 (2004): 1591-1596. DOI: <https://doi.org/10.1111/j.1365-2222.2004.02078.x>