RC3: Anonymous Referee #3, 24 Feb 2026

General comment

This preprint describes an innovative method for measuring and identifying pollen grains using a standard laboratory flow cytometer and machine learning tools. The authors made considerable efforts to collect and prepare a comprehensive pollen reference database, which they then used to test their method. The results are promising, showing that liquid flow cytometry coupled with machine learning algorithms could become an effective alternative to manual pollen counting and identification. I only have a few questions related to the content and some technical corrections. In my opinion, no major changes are needed.

We would like to thank the reviewer for the very favorable assessment of our manuscript, for the constructive comments and suggestions, and for all the minor technical corrections; they were very helpful.

Specific comments

- l.88-91 “Their inclusion enabled the model to learn to discriminate tree pollen from other common airborne particle types, as real-world environmental samples typically comprise a heterogeneous mix of tree, grass, and weed pollen, along with various non-pollen particulates.” Are non-pollen particles a majority in your samples? What proportion did you observe?

We thank the reviewer for the comment. The proportion of debris varies considerably from one sample to another, as it depends on how easy (or difficult) it is to separate the pollen from the rest of the floral unit, as well as on the filtration stage. The objective during model training is to include a limited number of non-pollen particles so that the model has an option for defining an “other” category, but not to achieve an exhaustive representation of the debris that might be found in the samples; we therefore do not consider it relevant to include the proportions in the manuscript.

When balancing the data, we kept only 1,000 for the species-level model and 10,000 for the genus-level model, like the other categories.

- l.104-106 “Depending on their peptide composition, these proteins absorb light at a certain wavelength and emit light radiation at a different wavelength in return producing a characteristic fluorescence signature that varies among species.” Is the fluorescence signal expected to differ from airflow cytometry due to the suspension in PBS?

The PBS we used is Dulbecco’s phosphate-buffered saline (without calcium and without magnesium), as described by Dunker et al. and Aloisi et al. [1][2]. This is a standard solution used in flow cytometry; it is a neutral, isotonic buffer solution that has no effect on fluorescence.

[1] Aloisi, Iris; Cai, Giampiero; Tumiatti, Vincenzo; Minarini, Anna; Del Duca, Stefano (2015): Natural polyamines and synthetic analogs modify the growth and the morphology of Pyrus communis pollen tubes affecting ROS levels and causing cell death. Plant Sci 239, pp. 92–105. DOI: https://doi.org/10.1016/j.plantsci.2015. 07.008.

[2] Dunker S, Motivans E, Rakosy D, Boho D, Mäder P, Hornick T, et al. Pollen analysis using multispectral imaging flow cytometry and deep learning. New Phytol. 2021;229(1):593–606. https://doi.org/10. 1111/nph.16882

To address this issue, on L98-101, we added « The filtrate was then suspended in Dulbecco's phosphate-buffered saline (PBS) solution, a standard neutral isotonic buffer commonly used in flow cytometry to minimize aggregation (Aloisi et al., 2015; Dunker et al., 2021b) (see detailed protocol in the supplementary material). »

- l.106-108 "For each laser, avalanche photodiode (APD) detectors measure the intensity of light emitted at different wavelengths using ten filters: 450/45, 525/40, 610/20 (violet laser), 525/40, 585/42, 690/50, 780/60 (blue laser), 660/10, 712/25, 780/60 nm (red laser)." It would be nice to have an explanation on these values, e.g., for 450/45 what does 450 and 45 represent? Most people in the aerobiology community are not familiar with these measurements.

We thank the reviewer for the comment; this has been clarified in the revised manuscript: L114-118: «Each filter value, such as 450/45, follows a simple convention: the first value corresponds to the central wavelength (in nanometers, nm), which is the midpoint of the light allowed to pass through the filter ; the second corresponds to the bandwidth, i.e., the width of this "window" of light. Thus, a 450/45 filter transmits light between 427.5 nm and 472.5 nm (i.e., 450 ± 22.5 nm). »

- l.108-109 "In addition to fluorescence, two scatter parameters were recorded to describe particle morphology: grain size and granularity." Give a precise definition of what size and granularity represent in this study as they can be misinterpreted. Is the size value the maximum axis of the pollen grain? What values do we expect for granularity?

We have clarified this in the revised manuscript.

L119-126: We changed for "The forward scatter (FSC) measures light diffracted by the pollen grain at a flat angle, reflecting the approximate diameter of the grain. The sidewards scatter (SSC) measures light diffracted by the pollen grain at a right angle, reflecting its granularity." with "The forward scatter (FSC) detects light scattered at low angles in the forward direction, which correlates with the cross-sectional area of the particle, equivalent to a spherical diameter. For non-spherical particles like pollen (prolate, oblate, tricolporate, etc.), FSC reflects an average optical cross-section as the particle passes through the laser in a random orientation. The sidewards scatter (SSC) detects light scattered at ~90° (orthogonal) to the laser beam, which is sensitive to internal complexity and surface irregularities. In pollen, this captures internal granularity, wall sculpturing, apertures, vacuoles/ pollen sacs. The more complex the structure and texture of the pollen grain, the higher the granularity values will be."

- l.125-126 "This resulted in a total of 25 parameter values per particle." Since the previous sentence is quite long I would summarise here the number of parameters: "This resulted in three values for size, two for granularity and 20 for fluorescence, with a total of 25 parameter values per particle."

We made the suggested changes, see L139-140.

- l.127-129 “For each species, pollen grains were manually separated from debris using scatter density plots (size vs. granularity) and histograms of all fluorescence features.” This is a very interesting approach that needs to be shared in the aerobiology community!

We thank the reviewer for this positive feedback.

- l.140-141 “Among these, the Random Forest algorithm showed the best performance and was therefore selected for subsequent analysis.” Which criterion was used to determine that Random Forest showed the best performance? F1 score? Please state it briefly in the text again.

We used the F1 score; this is mentioned in the revised manuscript on L156.

- l.160-173 It is very interesting to compare the species level with the genus level. In these two paragraphs, you describe the taxa that performed with reduced accuracy. What about those that performed best?

L179-180: We added “Most species perform very well, with 75% of species achieving a F1-score above 0.70.”

- l.212-214 “The models achieved high classification performance (F1=0.76 at the species level and 0.90 at the genus level) highlighting the potential of this approach as a scalable alternative to traditional microscopy for pollen identification.” For clarity, it would be good to state that this performance was achieved under lab conditions.

We have added this in the revised manuscript.

L240-242: added « While these findings are promising, they were obtained using reference pollen grains collected directly on trees; further validation using atmospheric samples will be necessary before implementation in an airborne pollen monitoring network. »

Technical corrections

Technical corrections are highlighted in the attached PDF file.

We thank the reviewer very much for taking the time to provide this feedback; we really appreciate it, and we have made all suggested changes in the revised manuscript.