Review: Vertical distribution, optical characterization and automated classification of airborne pollen from in-situ measurements

To the Authors

Summary: The manuscript entitled "Vertical distribution, optical characterization and automated classification of airborne pollen from in-situ measurements" presents an interesting and valuable investigation of airborne pollen using lightweight optical sensors installed at multiple heights, together with Hirst-type measurements, a Cloud Droplet Analyzer, SILAM simulations and machine-learning classification. The measurements extending from 4 to 272 m a.g.l. are useful and provide an opportunity to examine the vertical distribution of airborne pollen. It is evident that a substantial amount of experimental, analytical and methodological work has been invested in this study. The manuscript is generally well organized, the objectives are scientifically relevant and the results have the potential to make a meaningful contribution to the field of aerobiology and pollen monitoring. However, it could benefit from revision and further clarification of several methodological and interpretative aspects, which are described in detail in the comments below.

General Comments:

1) In Lines 36-52, please clarify the relationship between sampling height and transport distance. Long-range transported pollen can be mixed downward and detected at heights of 10-30 m or at canopy level. However, measurements at these heights alone cannot show whether the pollen originates from local, regional or distant sources. Rojo et al. (2019) mainly examined the effect of sampling height, while Damialis et al. (2017) used aircraft measurements over a wider altitude range. Therefore, I suggest revising the final sentence so that these studies are not presented as direct evidence of medium and long-range transport. In addition, in Line 38, Sofiev et al. (2013) is a valid reference and discusses long-range transport, but it is mainly a description of a birch-pollen emission module. It may not directly support the precise statement that transport over 100-1000 km becomes the dominant contribution before local flowering begins. Please provide more direct references for this statement or modify that sentence.

Response:

We have revised the Introduction to better distinguish between the concepts of vertical distribution and transport origin. Specifically, we modified the description of the studies by Rojo et al. (2019) and Damialis et al. (2017) so that they are presented as investigations of the vertical distribution of airborne pollen rather than as direct evidence of medium- or long-range transport. The concluding sentence has also been revised to clarify that measurements at different heights provide information on the vertical distribution and atmospheric mixing of pollen, while the identification of local, regional, and long-range transport requires supporting meteorological information or atmospheric transport modelling. In addition, we revised the sentence citing Sofiev et al. (2013) to remove the unsupported statement that transport over 100–1000 km becomes the dominant contribution before local flowering, and instead state more generally that long-range atmospheric transport can contribute to airborne pollen concentrations under favorable meteorological conditions.

2) The abstract states that the optical pollen sensors showed good agreement with the CDA measurements ($R^2 \geq 0.91$). However, the reported R^2 values appear to describe the comparison among the three PS2 sensors rather than a PS2-CDA regression. Could the authors please clarify this point? If a PS2-CDA regression was performed, it would be helpful to report the corresponding statistics. Otherwise, please revise abstract and conclusions accordingly.

Response:

The reported R^2 values (≥ 0.91) in the original abstract referred to the intercomparison among the three PS2 sensors rather than to the comparison with the Cloud Droplet Analyzer (CDA). To avoid this ambiguity, we have revised the abstract and the corresponding statement in the conclusions so that they now refer only to the strong agreement among the three PS2 sensors, without implying a direct PS2–CDA regression.

3) The correction of PS22 and PS23 requires clarification in Lines 264-268. The original slopes are reported as 0.93 and 0.76, while the manuscript states that the data were corrected by multiplying them by the respective slope factors. If the dependent sensor has a slope below one relative to PS21, multiplication by that slope would normally increase the disagreement rather than correct it. Please provide the exact correction equations and clarify which variable was treated as dependent.

Response:

Our wording was misleading. The regression slopes (0.93 for PS22 and 0.77 for PS23 relative to PS21) were not applied directly as multiplicative correction factors. Instead, $PS2_1$ was used as the reference instrument and the raw $PS2_2$ and $PS2_3$ concentration measurements were rescaled to the $PS2_1$ response by dividing by their respective regression slopes:

This calibration transforms the raw relationships to:

$$C_{\mathrm{PS22}} = \frac{C_{\mathrm{PS21}}}{0.93} = 1.065\, C_{\mathrm{PS21}}$$

and

$$C_{\mathrm{PS23}} = \frac{C_{\mathrm{PS21}}}{0.77} = 1.304\, C_{\mathrm{PS21}}$$

We have revised the manuscript , so our wording is not misleading.

4) Figure 3 shows that birch and pine pollen have different optical signatures and may therefore provide a basis for identifying pollen at higher levels. However, the procedure used in the vertical profile analysis is not fully described. Could the authors please clarify whether the PS2 measurements at 115 and 272 m were identified as birch by comparing their Ppol and DoP values with the laboratory clusters shown in Figure 3 or whether they were considered birch mainly because the Burkard measurements at 4 m showed a birch dominated period? If the laboratory values were used, please describe the classification criteria or exact boundaries applied.

Response:

We thank the reviewer for this important question. The PS2 measurements at 115 and 272 m were not classified as birch based on the laboratory polarization clusters shown in Fig. 3. Instead, the analysis was restricted to intensive pollen periods (IPPs) identified from the Burkard Hirst-type sampler at 4 m. Specifically, we selected periods during which birch accounted for more than 90% of the total pollen concentration according to the Burkard observations, as described in Sect. 2.5.3. The simultaneous PS2 measurements at 4, 115 and 272 m obtained during these birch-dominated periods were then analyzed. Thus, the interpretation that the PS2 observations predominantly represent birch pollen is based on the Burkard reference measurements rather than on a direct classification of the PS2 optical properties. We added The PS2 measurements at 115 and 272 m were not individually classified using their polarization properties. Instead, the same IPPs identified from the Burkard Hirst-type sampler at 4 m were used for all three PS2 instruments, and the simultaneous measurements were assumed to be representative of the dominant pollen type during each period.

5) In Line 220, please explain the basis for the threshold of 142 particles m^{-3} used to filter the PS2 measurements.

Response:

This concentration corresponds approximately to the detection of two particles within a 2h Burkard sample. Concentrations below this threshold are associated with very low particle counts and therefore high counting uncertainty. The same concentration threshold was applied to the PS2 measurements to ensure consistency with the Burkard data.

6) The Random Forest approach appears suitable for this classification task. However, the validation method may require some clarification. The stratified shuffle-split procedure preserves the proportions of the classes, but the 2 h observations are still divided randomly. As a result, neighboring observations from the same pollen episode may be included in both the training and testing datasets, which could lead to an optimistic estimate of model performance. Could the authors please clarify whether the temporal structure of the data was considered during validation? A temporally separated validation, for example using different years for training and testing, could provide a more independent assessment of the model.

Response:

We thank the reviewer for this important comment. We agree that randomly assigning individual 2 h observations to the training and test datasets may result in temporally neighboring observations from the same pollen episode occurring in both datasets, possibly leading to an optimistic estimate of classification performance.

Temporally separated validation is a good idea. However, a leave-one-year-out validation is unfortunately not well suited to our dataset because the occurrence of the different pollen species is strongly uneven between years, and some species occur only, or almost exclusively, during individual years. Consequently, a year-wise separation would result in some classes being absent from either the training or test dataset and would therefore not provide a meaningful evaluation of the multi-class classifier.

To explicitly assess the potential influence of temporal dependence, we repeated the classification using a grouped stratified validation. Instead of treating individual 2 h observations independently, we divided the time series into consecutive 3-day blocks and all observations belonging to the same 3-day block were assigned exclusively to either the training or validation data. Therefore, temporally neighboring observations within a 3-day period could not occur in both datasets. Stratification was retained as far as possible to account for the strongly unequal representation of the pollen classes.

The temporally grouped validation resulted in an overall accuracy of 74.5%, compared with 76.9% using the original stratified shuffle-split procedure. Therefore, imposing a substantially stronger temporal separation reduced the overall accuracy by only 2.4 percentage points. The classification rates of several of the better-represented classes also remained similar, for example, PINU changed from 79.7% to 75.5%, POAC from 90.2% to 90.5%, MIX from 31.9% to 30.0%, and NONE from 84.5% to 83.0%. A somewhat larger decrease was observed for BETU (88.2% to 75.4%), while the results for the least represented classes are more variable because of their small validation sample sizes.

We conclude that temporal correlation between neighboring 2 h observations has some influence on the classification results, but does not substantially change the overall performance or the conclusions drawn from this analysis. We have clarified this issue in the manuscript.

7) Some references may not directly support the statements with which they are associated. For example, Schneider et al. (2015) is a diagnostic study of patients with suspected asthma and does not appear to support the reported prevalence of allergic diseases in industrialized areas. Kuparinen et al. (2009) focuses on the influence of temperature and atmospheric conditions on wind driven plant dispersal, rather than on changes in the water content of individual pollen grains. In addition, the cited Oteros et al. (2019) paper concerns the spatial interpolation of pollen concentrations and may not be the intended reference for the description of the BAA500 instrument. Cariñanos et al. (2000) compares two methods for counting pollen on Hirst-trap slides and does not appear to support the statement regarding the complementary use of the PS2 and Hirst sampler. Finally, Sofiev (2019) discusses pollen data assimilation and the adjustment of seasonal emissions, but it does not directly demonstrate a general tendency of SILAM to overestimate pollen concentrations. Could the authors please check these references or/and either provide more appropriate citations?

Response:

Schneider et al. (2015): We agree that this reference does not directly support the statement that 15–30% of the population in industrialized areas suffers from allergic diseases. We have removed this citation and replaced it with a more appropriate reference supporting the reported prevalence (Linneberg, A. (2011), The increase in allergy and extended challenges. Allergy, 66: 1-3. https://doi.org/10.1111/j.1398-9995.2011.02619.x).

Kuparinen et al. (2009): We agree that this study primarily addresses temperature effects on wind-driven plant dispersal and does not directly demonstrate changes in the water content of individual pollen grains. We have therefore rewrote this statement and retained references as below.
Numerous studies have examined how meteorological conditions, including relative humidity (RH), temperature (T), wind speed, sunshine duration, and rainfall, affect pollen release and transport in the atmosphere. Relative humidity and temperature can affect pollen release and phenology, while temperature can also influence the atmospheric instability and turbulent transport that control wind-driven pollen dispersal (Chico-Fernández and Ayuga-Téllez, 2025; Tomczyk et al., 2025; Linkosalo et al., 2017; Kuparinen et al., 2009). For example, low relative humidity combined with high temperature tends to be associated with increased airborne pollen concentrations (Ščevková et al., 2015).

Oteros et al. (2019): We have checked this citation and clarified the associated statement. The correct citation is : Jose Oteros, Gudrun Pusch, Ingrid Weichenmeier, Ulrich Heimann, Rouven Möller, Stefani Röseler, Claudia Traidl-Hoffmann, Carsten Schmidt-Weber, Jeroen T.M. Buters; Automatic and Online Pollen Monitoring. *Int Arch Allergy Immunol* 1 September 2015; 167 (3): 158–166. https://doi.org/10.1159/000436968

Cariñanos et al. (2000): Cariñanos et al. (2000) is now cited in the description of the Hirst methodology, where it appropriately supports the discussion of microscopic pollen counting strategies and the comparability of counting methods used for Hirst-type volumetric traps. The statement on the complementary roles of the PS2 and the Hirst-type sampler has been revised accordingly.

Sofiev (2019): We have revised the manuscript to state that previous evaluations have shown that SILAM generally reproduces the overall temporal evolution and magnitude of airborne pollen concentrations reasonably well, while model performance remains sensitive to factors such as pollen emission parameterization and meteorological input, leading to biases that may vary depending on the study region and season. This revised wording is consistent with the scope and conclusions of the cited reference. Sofiev M, Siljamo P, Ranta H, et al. A numerical model of birch pollen emission and dispersion in the atmosphere. Description of the emission module. *Int J Biometeorol*. 2013;57(1):45-58. doi:10.1007/s00484-012-0532-z

We have carefully checked the remaining citations in the manuscript and revised the relevant statements to ensure that the references directly support the claims being made.

Specific Comments:

1) "Most PBAP cause allergic diseases" is too general in Line 18. Many primary biological aerosol particles are not allergenic. Please refer specifically to allergenic pollen and fungal spores.

Response: The text has been revised to refer specifically to allergenic pollen and fungal spores, rather than PBAPs in general.

2) Please explain how the PS2 thresholds Ppol>1.3 and -0.1<DoP<0.5 were determined and report their sensitivity (Lines 152-153).

Response: We selected the thresholds based on the Figure 3, which illustrates the PS2 count patterns in the Ppol-DoP plane for selected pollen grains, measured in the laboratory, and fog droplets, measured in the field. As we can see from Figure 3, there is a wide gap in the DoP value between them, so we consider our choice for the thresholds safe and reliable for distinguishing pollen grains from fog droplets. We have modified the text accordingly.

3) Figure 5 and the caption use inconsistent periods: 1 May-1 July versus 1 May-31 July.

Response: The figure caption has been corrected so that the time period is consistent throughout the manuscript (1 May–1 July).

4) Please correct "3d" to "3-D" (Line 179), "versions 5.9" to "version 5.9" (Line 185) and the typographical phrase "analysis of the birch and IPPs" (Lines 277-279).

Response: The suggested corrections have been implemented throughout the revised manuscript, including changing "3d" to "3-D", "versions 5.9" to "version 5.9", and correcting the wording of the phrase referring to the analysis of the IPPs.

5) Throughout the manuscript figures, the terminology should be revised to "pollen grains/m^3" rather than "m^3", when referring to units based on the paper of Galán et al. (2017) [1].

Response: We have revised the terminology throughout the manuscript and figures to report pollen concentrations as "Pollen grains m^{-3}", following the recommendation of Galán et al. (2017).

Bibliography:

1. Galán, C.; Ariatti, A.; Bonini, M.; Clot, B.; Crouzy, B.; Dahl, A.; Fernandez-González, D.; Frenguelli, G.; Gehrig, R.; Isard, S.; et al. Recommended Terminology for Aerobiological Studies. *Aerobiologia (Bologna).* **2017**, *33*, 293–295, doi:10.1007/S10453-017-9496-0/METRICS.