Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., author comment AC2
https://doi.org/10.5194/amt-2021-272-AC2, 2021

Reply on RC2

Danaël Cholleton et al.

Author comment on "Laboratory evaluation of the scattering matrix of ragweed, ash, birch and pine pollen towards pollen classification" by Danaël Cholleton et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2021-272-AC2, 2021

Reviewer #2

In the submitted work a laboratory evaluation of light scattering by four of the most common pollen taxa, namely ragweed, birch, pine, and ash, is provided for the purpose of pollen classification. The authors conduct the experiment of light scattering by grains of mentioned pollen taxa and represent it through scattering matrix formalism at two different wavelengths (532 and 1064nm) of incident radiation. Elements of a scattering matrix describe how the polarization state of the incident radiation has changed by light scattering of the studied pollen grains. A Principal component analysis (PCA) is applied on estimated ten scattering elements (five per wavelength) to reduce the dimensionality of the feature space to two by explaining 99% of the variance in the data. In the transformed domain, defined by the range of PC1 and PC2 components, a pollen identification is performed based on the area size of cluster regions of projected scattering matrix elements.The methodology is well explained within the manuscript, and the results are clearly represented. However, there is space for the manuscript improvement if the following remarks are addressed:

- We thank Reviewer #2 for the time she / he spent to carefully analyse our manuscript and to identify our methodology and its novelty. We here provide point by point answers to her / his remaining comments.

Comment 1

Line 35 -40 A recent work of Sauliene at el. 2019 https://doi.org/10.5194/amt-12-3435-2019 , considers different modalities of data (among them light scattering data) for real-time pollen identification.

- **Answer to comment 1 and list of changes made to the manuscript:**
Thank you for the remark. We agree and added the paper by Sauliene et al., 2019 to our references. These authors indeed analysed pollen side scattering patterns, which is a complementary method to our work (they focus on a range of scattering angles, we focus on light polarization at a given scattering angle). We added this reference to our manuscript: "Likewise, image recognition on the scattering pattern of pollen grains have been investigated, as described by (Šaulienė et al., 2019), and holographic

images are also used (Giri et al., 2019; Sauvageat et al., 2020; Kemppinen et al., 2020) as an identification methodology"

Comment 2
Line 225-230 Normalization of the detected intensity by that of a photodetector placed at 170 degrees scattering angle is motivated by its dependency on the pollen grains number concentration. Can the authors provide a more informative explanation on this, concretely how pollen grains number influence the measured intensity of the photodetector at 170 degrees?

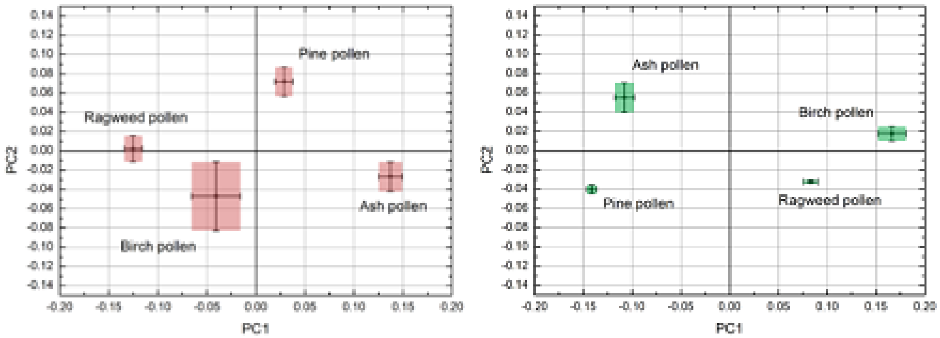
- **Answer to comment 2 and list of changes made to the manuscript**
  Thank you for your comment. As detailed in light scattering textbooks (Mishchenko et al., 2002), light scattering is proportional to the particles number concentration . More precisely, the scattered light intensity at 170° scattering angle is proportional to as this detector is polarization insensitive. To account for the reviewer's comment, we modified our manuscript by adding the following sentences: "Indeed, the scattered light intensity at 170° scattering angle is proportional to the pollen grain concentration as this detector is polarization insensitive. As a result, statistical errors due to potential fluctuations in the pollen grains number concentration are removed by considering the ratio of the two intensities at these two scattering angles."

Comment 3
Line 245 -250 In the (45+)-polarization curve two successive local minima are not equal at wavelength… Some annotation on Figure 4 will be helpful for understanding which exactly two. (optional)

- **Answer to comment 3 and list of changes made to the manuscript:**
  Thank you for this remark, we precised what we intended directly in the body of the revised manuscript: "In the (45+)-polarization curve, the two successive local minima are not equal at wavelength (see, for example, the first and second minima of ragweed pollen at wavelength)".

Comment 4
Line 290-325 The Figure of PC components obtained from five scattering elements (separately for VIS and IR) would clarify the influence of wavelength selection on pollen identification.

- **Answer to comment 4 and list of changes to the manuscript**
  To answer to your comment, we carried out a PCA by considering only five matrix elements, as provided in the figures below for wavelength (left graph, in red), then for wavelength (right graph, in green). The influence of each wavelength is to be seen on each graph and shows the applicability of our methodology already at one wavelength: in both cases, the PCA allows identifying a simple light-scattering criterion to differentiate each taxon. A larger area is however obtained for birch pollen at wavelength as the size of the obtained error bars are related to the achieved precision on each scattering matrix element. Since our experimental error bars are precise, considering the scattering matrix at each wavelength adds value to our understanding of the light scattering characteristics of each pollen taxa. We hence chose to only present the PCA graph by considering both wavelengths. To include the reviewer's comment, we added the following sentence to our revised manuscript: 'When considering each wavelength (,) separately, the PCA still allows identifying a simple light-scattering criterion to differentiate each taxon, with a precision depending on the achieved accuracy in the retrieved scattering matrix elements at the considered wavelength'.
-



Comment 5

Line 290-325 It is well known that PCA is the standard method mostly used for compact data representation while Linear discriminant analysis (LDA) is the standard method used when the discriminant features are needed for classification purposes when the class labels are known. From that perspective, an LDA is more suitable for the considered problem. Therefore I strongly advise authors to consider LDA in the part of data analysis.

- **Answer to comment 5 and list of changes to the manuscript**
  Thank you for your comment. We followed the approach published by Martinez et al. who published a paper dedicated to that topic, entitled "PCA versus LDA". There, theses authors concluded that "PCA might outperform LDA when the number of samples per class is small". In our methodology, each class (pollen taxon) is represented by a single point as being representative of the distinct size and shape of each pollen taxon. The uncertainties associated with this single point correspond to our experimental uncertainties, but we only have one element per class. Moreover, the LDA cannot be applied to our methodology as more than one element per class would then be required. To account for the reviewer's comment, we modified our manuscript as follows: "A Linear Discriminant Analysis (LDA) may appear more suited for labelled classes. However, as published by Martinez et al. (2001), PCA might outperform LDA when the number of samples per class is small, and in our methodology, each class (pollen taxon) is represented by a single point as being representative of the distinct size and shape of each pollen taxon. Applying a LDA would require having more than one single point per class".
- Martinez et al., PCA versus LDA, IEEE Transactions on pattern analysis and machine intelligence, Vol. 23, No. 2, 2001.