Manuscript 10.5194/amt-2021-272, submitted to Atmospheric Measurement Techniques.
Laboratory evaluation of the scattering matrix of ragweed, ash, birch and pine pollens towards pollen classification
by D. Cholleton, E. Bialic, A. Dumas, P. Kaluzny, P. Rairoux and A. Miffre

Author's detailed response with a list of the changes made to our manuscript

We thank Reviewer #1 and #2 for their comments which add value to our manuscript. We carefully addressed each comment point by point, by first recalling it with italic text, then addressing it in a detailed response. A list of the changes made to our manuscript is then provided in italics and also as red-lined version in the revised manuscript.

Reviewer #1

"The authors studied here a pollen classification method so as to be able to automate the identification process in the future. The topic is novel and important, as the pollen recognition and counting comprises an arduous task and also does not allow for (near-)real-time pollen information. The methods seem sound, the concept novel, the results robust enough. However, I wish to suggest some corrections and additions based on several concerns I have that, if addressed, would further improve the paper."

- We thank Reviewer #1 for identifying the novelty and significance of our manuscript. We carefully addressed each remaining comment, which add value to our manuscript.

Comment 1
"*Pollens: correct throught the manuscript into pollen."*

- **Answer to comment 1 and list of changes made to the manuscript:**

  Thank you for the remark. We agree and did the correction all along the manuscript.

Comment 2
"*lines 10-11: how (and where) are these the most common? You need to further justify here and in the Methods."*

- **Answer to comment 2 and list of changes made to the manuscript:**

  A detailed discussion on how and where the four considered pollen taxa are considered as the most common is provided in the original manuscript in the dedicated paragraphs (2.1 for ragweed, 2.2 for ash, 2.3 for birch and 2.4 for pine). We however agree that we could have been clearer in the writing at lines 10-11 and removed the expression "most common" from the abstract for the sake of clarity. To account for the Reviewer's comment, we also added the following sentence at the end of the introduction to ease the reading: "*We focus on ragweed, ash and birch pollen, which are allergenic, and currently monitored in several countries countries in North America and in Europe. Also, pine pollen is studied in complement as strong pine pollen events have been reported in the literature (Spänkuch et al., 2000)."*

Comment 3
line 24: this is a dramatic generalisation: the referred paper actually cites a projection for Ambrosia pollen, which, being an invader, could be expected. Plese be more neutral here.

- **Answer to comment 3 and list of changes made to the manuscript:**

  Thank you for your comment. The four times increase in concentration reported by (Hamaoui-Laguel et al., 2015) is indeed related to Ambrosia pollen, so we agree to be me more neutral. To our knowledge, there is no similar projection for other taxa. Based on previous observed evolutions (1994-2010), (Zhang et al., 2015) however reported a 42 % increase in the birch pollen concentration in the US. To account for the reviewer's comment, we modified the corresponding sentence in the introduction as follows: "*Indeed, Ambrosia pollen concentrations are expected to increase by 400 % in the following decades (Hamaoui-Laguel et al., 2015) for*

Ambrosia is an invasive plant. To our knowledge, no similar projection exists for other taxa. Based on observed evolutions over the 1994-2010 period however, (Zhang et al., 2015) reported a 42 % increase in the birch pollen concentration in the US."

Comment 4

line 26: citation?

- **Answer to comment 4 and list of changes made to our manuscript:**

To account for the Reviewer's comment, we added the two following references to our revised manuscript together with the following sentences: "*Due to climate change, which increases the global temperature and CO_2 atmospheric concentrations, the length of the pollen season should extend (Bielory and al., 2012). Moreover, the geographical repartition of some pollen plants such as ragweed is also expected to extend (Ziska et al., 2011).*"

Bielory, L., Lyons, K., and Goldberg, R.: Climate Change and Allergic Disease, Curr. Allergy Asthma Rep., 12, 485–494, https://doi.org/10.1007/s11882-012-0314-z, 2012.

Ziska, L., Knowlton, K., Rogers, C., Dalan, D., Tierney, N., Elder, M. A., Filley, W., Shropshire, J., Ford, L. B., Hedberg, C., Fleetwood, P., Hovanky, K. T., Kavanaugh, T., Fulford, G., Vrtis, R. F., Patz, J. A., Portnoy, J., Coates, F., Bielory, L., and Frenz, D.: Recent warming by latitude associated with increased length of ragweed pollen season in central North America, Proc. Natl. Acad. Sci., 108, 4248–4251, https://doi.org/10.1073/pnas.1014107108, 2011.

Comment 5

line 27: change to allergic reactions.
line 30: change to 'consisting of sampling the...'.

- **Answer to comment 5 and list of changes made to the manuscript:**

Thank you for your comment, we agree and did the correction.

Comment 6

Lines 35-36: it is not as recent. Check also Sauliene et al. 2021, and Schaefer et al 2021, and references therein, in both (https://doi.org/10.1016/j.scitotenv.2021.148932).

- **Answer to comment 6 and list of changes made to the manuscript:**

Thank you for your remark. We agree and quoted a previous paper from (Oteros et al., 2015) as a reference for the classification performed with the BAA500.

Comment 7

line 56: pleae provide, here and in Methods, a full justification. for instance, pine pollen is not considered as allergenic because of its size. And based on their geographical distribution, there are other pollen types that are definitely more allergenic and abundant. You need to explicitly justify the selection criteria, at least for the locality, if not in an international context, as it might be considered otherwise a convenience sample.

- **Answer to comment 7 and list of changes made to the manuscript:**

Thank you for the remark. As above explained (see answer to comment 2), our selection criteria is indeed not only based on the allergenic character of the considered taxon as we wrote in the original manuscript (see lines 106-109): pine pollen, though rarely causing allergies, "*however remains an interesting study case, as it may impact the Earth's climate locally: (Spänkuch et al., 2000) reported that a high pine pollen concentration increased the down welling infrared flux up to eight times the monthly means.*" We however agree that we could have been clearer in the writing and to account for the Reviewer's comment, we moved the sentence at lines 106-109 related to the pine pollen to the introduction, to ease the reading.

Moreover, please note that, as we wrote in the outlook section of our original manuscript, our methodology is "*applicable to other sets of pollen taxa, provided that the same accuracy is experimentally achieved*." In agreement with the scope of the AMT journal, the novelty of our manuscript is here to develop a new and precise measurement technique and apply it to four existing case studies, rather than studying the light scattering properties of all taxa in a given geographical region.

Therefore, no convenience sample exists and each sample can be studied using our new methodology, it is a question of precision and our work exhibits the required precision. For the sake of clarity, we added the following sentences to our revised manuscript, at the end of the introduction: "*Our main selection criteria for choosing our pollen samples was the geographical location (Europe, North America) and the allergenic character or / and the climatic impact of the pollen taxon. Still, other pollen taxa could be likewise studied by applying our new measurement technique, provided that the same accuracy is experimentally achieved. Extension of this work to all pollen taxa in a given geographical region is however beyond the scope of our contribution".*

Comment 8

lines 66-71: unless the journal formatting and writing style requires it, it is not necessary to state here the structure of the paper. As currently is, it rather resembles a thesis. Please rephrase/omit.

- **Answer to comment 8 and list of changes made to the manuscript:**

  These sentences were aimed to ease the reading for future readers. To account for the reviewer's comment, we rephrased these sentences by removing the explicit structure of the paper: "*The paper first presents the studied pollen samples, then details our laboratory methodology to precisely evaluate the scattering matrix of these pollen. From that, the evaluation of their scattering matrix is presented, and a principal component analysis is proposed as an outlook to help classifying among these pollen.*"

Comment 9

lines 82-83: Thibaudon et al. studied the pollen distribution; there are other studies and reviews that provide such information. Please change.

- **Answer to comment 9 and list of changes made to the manuscript:**

  We agree and quote the paper by (Dahl et al., 1999) to modify our manuscript as follows: "*Ragweed pollen induces particular allergic reactions such as asthma twice more often as other pollen (Dahl et al., 1999) with an annual economic cost of 7 billion euros in Europe (Schaffner et al., 2020).*"

Comment 10

line 92: Why Fraxinus americana, since you refer to France as the study area? Why not Fraxinus excelsior or similar?

- **Answer to comment 11 and list of changes made to the manuscript:**

  We never refer to France in our manuscript as can be easily checked (the word France is not used in our original manuscript). As above explained (see answers to comments 2 and 7), we here focus on ragweed, ash and birch pollen, which are allergenic, and currently monitored in several countries in North America and in Europe. Hence, for ash pollen, Fraxinus Americana was selected for North America, though Fraxinus excelsior may also have been considered for Europe. Considering both is however beyond the scope of our manuscript: our contribution is not aimed at studying the light scattering properties of all taxa in a given geographical region. We recall (see answer to comment 7) that, in line with the scope of the AMT-journal, the novelty of our manuscript is to develop a new atmospheric measurement technique and that each pollen sample can be studied by using our new methodology: it is there a question of precision and our work exhibits the required precision. To account for the reviewer's comment, we modified our manuscript as follows "*We here consider ash (fraxinus americana) pollen which is a relevant source of allergenic reactions in North America. In Europe, ash pollen has been underestimated for a long time as its bloom season overlaps with that of birch (Imhof et al., 2014). Hence, Fraxinus excelsior may also be considered as an outlook of this work, provided that the same experimental accuracy is achieved."*

Comment 11

line 94: please be more specific, which ones?

- **Answer to comment 11 and list of changes made to the manuscript:**

The asked specific information is detailed in the publication by (Niederberger et al., 2002) which we quoted in our original manuscript. To account for the Reviewer's comment, we added the example of birch pollen for cross-reactivity with its antigens and hence modified our manuscript as follows: "*Moreover, it presents a high cross-reactivity from allergens from other plant species such as birch pollen, as underscored by (Niederberger et al., 2002).*"

Comment 12
line 96: 3-4 colpi
line 99: in central and north Europe.

- **Answer to comment 12 and list of changes made to the manuscript:**

Thank you for the remark, we did the correction.

Comment 13
line 106: Again, why Pinus strobus? Please justify the selection of species.

- **Answer to comment 13 and list of changes made to the manuscript:**

Please refer to our response to comments 2, 7 and 10 where the selection of species is justified. We recall that, in line of the AMT-journal, the goal of our manuscript is to develop and present a new atmospheric measurement technique, which can be applied to any pollen taxon: it is a question of experimental precision and our work exhibits the required precision.

Comment 14
line 121: why were commercial pollen grains were used? Was it considered that there might a difference compared to the real-life, fresher pollen? It is known from other laboratory experiments that pollen may deteriorate and lose physical and chemical (and optical?) properties when not fresh and manipulated from commercial samples. Plese justify and discuss fully.

- **Answer to comment 14 and list of changes made to the manuscript:**

Thank you for your question. Actually, we did not have fresher pollen grains, but we carefully checked that the size and the shape of our studied pollen grains (evaluated from the microscopic images taken at iLM to be seen in Figure 1), were identical to that referenced in the literature (see Paldat pollen database for example). To account for the Reviewer's comment, we also carefully looked at the literature on the comparison of optical properties of pollen in laboratory / field. This literature is extremely rare with however a very recent contribution by (Miki and Kawashima, 2021), which was quoted in our original manuscript. We however note that these authors did not observe any temporal changes in light scattering by alnus pollen over a ten days period. We then followed their conclusion. To account for the Reviewer's comment, we modified our manuscript by adding the following sentences to Section 2.5: "*These commercial pollen grains may differ from that of atmospheric pollen grains. However, our Figure 1 microscopic images did not exhibit differences in size nor shape compared with fresher pollens. Otherwise, the most recent literature does not report any changes in the pollen light scattering properties over a ten days period (Miki and Kawashima, 2021)*".

Comment 15
Fig. 2 and relevant text: while i am in agreement with the method used before and here, i wonder whether specifically the pine pollen would exhibit a different behaviour that might have affected the scattering results too. Despite irregularities and peculiarities of all pollen grains per species, they all have in common that they are rather ovoid. Exception is the pine pollen, which is not symmetrical (only along it axis), it is much heavier and with distinct texture and shape between the main body and the sacci. Please discuss more in depth on this and provide further justification and comparisons.

- **Answer to comment 15 and list of changes made to the manuscript:**

Thank you for your comment. The fact that the pine pollen is not symmetrical along its axis and has a distinct size, texture and shape is indeed responsible for the observed differences in the retrieved scattering matrix elements for pine compared with other taxa. Our spectral and polarimetric light scattering methodology is indeed sensitive to the size and to the shape of each

pollen taxon and the achieved precision in the scattering matrix retrieval allows accounting for these specific size and shape features, which allow identifying each pollen separately. To account for the reviewer's comment, we modified our manuscript in Section 4 as follows: "*The fact that the pine pollen is not symmetrical along its axis and has a distinct size, texture and shape is indeed responsible for the observed differences in the retrieved scattering matrix elements for pine compared with other taxa. Our spectral and polarimetric light scattering methodology is indeed sensitive to the size and to the shape of each pollen taxon and the achieved precision in the scattering matrix retrieval allows accounting for these specific size and shape features, which allow identifying each pollen separately."*

Comment 16

Figure 6 and relevant text: while the results are robust and the method successful for the selected species, my main concern is whether this or a similar technique would be actually operational in a real-life study design, with more species, fresh pollen, and if then there would be indeed a possibility to identify among a larger range of pollen species. Please discuss and set as a study limitation.

- **Answer to comment 16 and list of changes made to the manuscript:**

Thank you for identifying the robustness and the success of our methodology, which is in line of the scope of AMT. The raised issues are important but far beyond the scope of this manuscript which is dedicated to the development of a new measurement technique, in agreement with the scope of the AMT journal. For that reason, in our original manuscript, we considered these aspects in the outlook section, as explained in our answer to comment 7. There and in the outlook section of our original manuscript, we explained that our methodology can be applied to identify among a larger range of pollen species "*provided that the same accuracy is experimentally achieved*." Hence, any pollen taxon can in principle be studied and identified by applying our new methodology: it is a question of precision and our work exhibits the required precision. Our experimental error bars are indeed very low, so does the probability for two different pollen taxa to exhibit the same ten scattering matrix elements (five per wavelength). Laboratory intense work is however required for extending this work to other species, which is far beyond the scope of this contribution, dedicated to the development of this new measurement technique. To account for the reviewer's comment, we added the following sentences to the outlook section: "*Any pollen taxon can in principle be studied and identified by applying our new methodology: our experimental error bars are indeed very low, so does the probability for two different pollen taxa to exhibit the same ten scattering matrix elements (five per wavelength). Laboratory intense work is however required for extending this work to other species, which is far beyond the scope of this contribution, aimed at introducing this new measurement technique.*"

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 N. More precisely, the scattered light intensity at 170° scattering angle is proportional to N 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 λ_{IR} (see, for example, the first and second minima of ragweed pollen at wavelength λ_{IR})".*

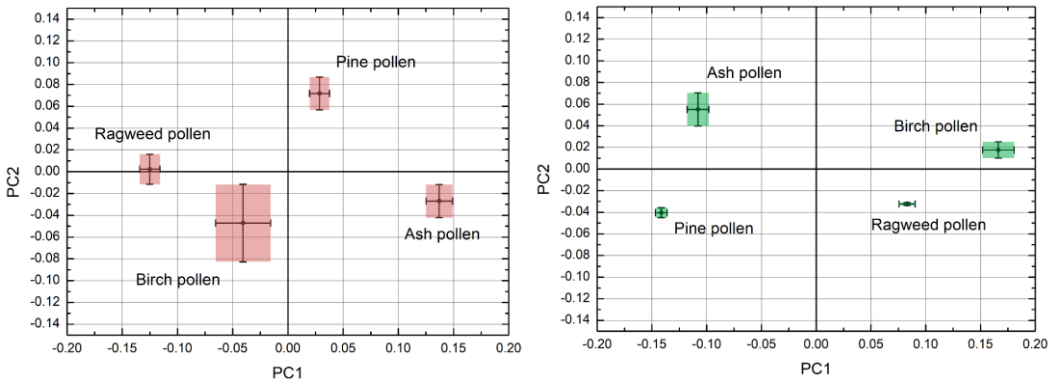
Comment 4

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 λ_{IR} (left graph, in red), then for wavelength λ_{VIS} (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 λ_{IR} 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 (λ_{VIS}, λ_{IR}) 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.