Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC1
https://doi.org/10.5194/amt-2022-81-RC1, 2022

Comment on amt-2022-81

Anonymous Referee #1

Referee comment on "Combining Mie–Raman and fluorescence observations: a step forward in aerosol classification with lidar technology" by Igor Veselovskii et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2022-81-RC1, 2022

The manuscript describes several case studies of lidar observations where fluorescence observations combined with lidar depolarization shows significantly different properties for pollen, smoke, dust and anthropogenic aerosol. I'm excited to see the potential of these new measurements, which give completely independent and orthogonal information about aerosol particles, at single bin resolutions, significantly increasing the information available for aerosol typing. The case studies are a nice selection of different types and mixtures and interesting to see.

The manuscript seems to suffer from an identity problem, however. Mostly it is an illustrative set of cases studies that demonstrate differences in the two-dimensional space of fluorescence capacity and particle depolarization. It includes nice analysis of some mixtures of types as well. However, the paper claims to be an algorithm description paper, and for that purpose, analysis of a few hand-selected case studies really isn't sufficient, and the mixture analysis doesn't exactly fit, because it is not part of the algorithm. Apparently in consequence of this uncertainty about the desired focus of the paper, some aspects of the paper seem superficial, or rather, inconsistent in depth. The inferences in the paper about the types seem very reasonable, but many are not backed up by any independent information or compared with other methods of classification, which should be done to demonstrate the validity of the new algorithm, particularly if this is the algorithm description paper. Also there's insufficient information about how the thresholds in the algorithm were chosen. In the analysis of the case studies, there should be a consistent effort to include complimentary information to validate the case identifications using other measurements (in situ or other lidar measurements that reveal type) and backtrajectories. And if a major focus of the paper is to showcase the performance of a new (and better) classification algorithm, then the results should be shown on a bulk of data in addition to the case studies, and comparisons with other classification methods should be made and discussed.

Specific comments:
L24. "and their mixtures". The mixture analysis is an interesting part of the paper, and apparently new compared to the authors' other papers, but it appears it's not really part

of the classification algorithm, in the sense that mixture analysis can only be done on a case-by-case basis. Any discussion about that? This could be clarified in the abstract. Also, the mixture analysis is not even mentioned in the introduction. Discussing it there would help to clarify the novel aspects of the paper.

L73-75. I very much agree that adding independent aerosol information will improve classification, but this specific point is unconvincing. Yes, the variables used for classification so far have variability within types but there's nothing to suggest that this won't also be true for fluorescence capacity, is there? So, I'm not sure this is exactly the right motivation.

L105. Good point that the resolution is higher since fluorescence capacity can be calculated using data at a single bin, unlike extinction or other quantities related to extinction. This seems particularly useful for Raman measurements.

L105-107. Veselovskii et al. 2021a is referenced extensively in the introduction, including to say that it already demonstrates the ability of the 2-d measurement space to separate all the aerosol types. I couldn't follow how the purpose and scope of this paper is different from 2021a.

L183-193. Calculation of the backscatter coefficient using a calibration constant sounds so straightforward, that I didn't realize that it hadn't been done before. This is great. It's good to see a relatively straightforward innovation discovered and put into practice that will produce a significant amount of additional retrievals, in profiles when the reference height is not accessible to the lidar.

L231-232. Add an earlier reference for spectral dependence of the depolarization ratio, Burton et al. 2015.

L240-241. Since line 223 just said that Veselovskii et al. 2021a already demonstrated that the two dimensional diagram can separate types, is the part about mixtures the main purpose of this manuscript? If so, the abstract and intro should make that clearer and the examples should be chosen to align with that purpose.

L248-249. Burton et al. (2012) or Burton et al. (2013), referenced elsewhere in the manuscript, is an earlier lidar aerosol classification methodology with depolarization ratio ranges listed for common types.

L247. "The ranges are based on results obtained in LOA". The algorithm is a simple thresholding method in two dimensions, so the ranges are the single most important aspect of the algorithm description. This statement is much too vague to support and explain how the ranges were derived, and I'm eager to know more. What results? From cases published in other publications? From a completely independent subset of cases than the results shown in this manuscript? Are the results only inferences from the lidar measurements of depolarization and fluorescence capacity, or do they include other coincident measurements that provide stronger evidence for the type identifications? Is there a set of training cases that are classified using other external measurements and/or source information? Are the cases shown in this paper the training cases or are they independent cases that demonstrate the validation of the algorithm? All this should be part of the methodology discussion.

Figure 3. The mixing lines all go through the box that's marked "pollen". This highlights the unavoidable weakness of typing with just two dimensions. Presumably, anything that falls within this box needs context to distinguish between pollen, a pollen mixture, or a smoke-dust mixture that has nothing to do with pollen. Identification by context (particularly where supporting measurements are available) is fine for the purpose of case studies, but there must be significant potential for misidentification in the automated algorithm, I suppose. It would be good to discuss weaknesses as well as strengths of the approach.

L268. Clouds are also shown in the aerosol typing masks and line 308 mentions both ice and water droplets, so the thresholds values for ice and water droplets should also be included in Table 1.

Figs 4,5. It's confusing that the ice cloud is only partially included in this example. It's shown in the type mask, but not discussed, and it's not shown in the scatter plot in Fig 5a. It's included in Fig 4, but apparently off-scale. The authors should decide whether they want to include the cloud in their analysis and discussion or not. If not, cut off the plots at an altitude below the cloud. If so, rescale Figure 4, include it in Fig 5 and add discussion about cloud.

Figure 5 and similar figures. What's the purpose of the boxes and cross-hairs in the fluorescence vs. depolarization diagrams? The boxes would probably be more useful to readers if they were all the same, and used the values from Table 1. That way, we can see visually how the identified types fall into the broad category already established. I can guess that the crosshairs represent the mean and (probably) standard deviation of identified pure types, but those aren't discussed anywhere in the paper.

L315-321. The explanation of the smoothing procedure is missing something. Z is a number, but the classification IDs are not numbers that can be added and weighted, but just labels. How are the classifications convolved with Z? Just guessing, I suppose the fluorescence capacity and depolarization ratio are what's averaged using the Z-weightings, and then the classification is done on these smoothed measurements instead? Please clarify in the text.

L339 and 341 and elsewhere. I'd suggest avoiding describing values as "typical" and expand the description to be more specific. For instance, perhaps this is within the ranges seen in your previous publications and/or other publications for cases that have been identified as smoke and urban based on independent data? "Typical" is a bit dangerous, in that it implies a generality that is not established after only a few handfuls of case studies, particularly since the case study identifications seem to mostly be rather dependent on expectations about the typical values. Statements like this unfortunately seem to be quoted and referenced repeatedly so that they become ingrained without becoming better supported. After all, we now know that it is quite common for smoke (in the upper troposphere and stratosphere) to have depolarization values that are much larger than this, and previously published ranges of depolarization for urban aerosol also include significantly larger depolarization values than this.

L347. Says that the fluorescence capacity can decrease as a function of relative humidity, explaining a range of variables. Why does it produce variability rather than reducing the fluorescence capacity uniformly?

L361-367 and Figure 6-7. I agree that the shape of the curve in Figure 7a is very striking and reminiscent of a mixing line. However, I also just read in the previous section that fluorescence capacity is strongly impacted by relative humidity, making me wonder quantitatively how much impact RH has, compared to the impact of mixing. Is there a model (theoretical or empirical) of G_F dependence on relative humidity? The RH profile should be added to Figure 8 (and all the other profile figures). Another aspect that puzzles and surprises me is the increased G_F specifically in parts of the curtain where the backscatter is lower. This hints that the variation in G_F might be quite strongly related to RH; alternately that the pollen is more diffuse and widespread than the urban aerosol, which I think would be unusual. A curtain of RH (perhaps from MERRA-2 since there is insufficient sonde data to produce a curtain) and/or backtrajectories might help make the scenario more clear.

L368-369. It's good that 1064 nm depolarization is included here, because in general, the more data shown, the better the patterns can be understood. However, the text highlights larger values of 1064 nm depolarization to support the inference of pollen, but that's also true for urban aerosol (e.g. Burton et al. 2012). Then "both depolarization ratios decrease with height" as the pollen concentration decreases (L372), but 1064 continues to be larger than 532, so again this is not definitive. Any further comment about this?

This case and the first case were also included in earlier publications by the same authors. The papers make different analyses of them, so that's fine, but does this mean they also contributed information relevant to producing the ranges used in the algorithm? If so, they are not such good examples to illustrate the performance of the typing algorithm.

Figure 8 L 715. Why were the profiles created for 21:00-23:00 instead of a later time where the curtain shows pollen at lower altitudes and mixing is discussed? Is this a mistake?

L376-377. I'm not finding the explanation for the lack of variability in the backscatter angstrom exponent to be very convincing. It appears to be saying that the urban particles are growing due to humidification exactly in balance with the effective dry particle size decreasing due to less pollen? (if so, this needs support). Perhaps some quantitative modeling would help. How small of a backscatter Angstrom exponent would be expected for high concentration of pollen, and just how much contribution to the backscatter is there (based on the mixing model) and how much change in Angstrom would you therefore expect? What confuses me is that the fluorescence capacity also mixes linearly according to the backscatter partition, so if there was really too little backscatter contribution to be noticeable, wouldn't that also mean there would be little variation in G_F as well?

L392-393. Unfortunately, the SILAM website only provides current forecast data, so please make the relevant data available as a supplement or shown in a figure. Also, what kind of pollen was it?

L418-419. I'm not quite clear on what the author's intent is here. Is this saying that the algorithm misclassified a mixture as pure urban, or that the mixture only occurs where the classification puts it, but that the two urban layers have quite a lot of difference between them? It would be very helpful (in this case and others) to mark the points in the scatterplots according to the classification result or altitude. I would like to see exactly where the two layers classified as "urban" fall on the apparent mixing line. I think it's interesting that the two layers marked urban have different spectral dependence of depolarization. Backtrajectories would be helpful for this case too, to help understand why the two layers of urban aerosol might have different properties.

L420. "typical for urban-pollen mixture". Actually the mixing curve is significantly to the left of the curve in Figure 3, suggesting that the pure pollen in this mixture is not "typical" compared to the ranges given in the table, but is more of an edge case with relatively low fluorescence capacity.

L430-436. This is a very nice case to demonstrate contrast in fluorescence between different types. But the type identification is entirely made by inference using the two classification dimensions without any other support such as in situ measurements, backtrajectories, or other lidar-measured quantities like 1064 nm depolarization, lidar ratio or angstrom exponents. It's great that two measurements used for the classification appear to give the ability to make these separations, but for such a key demonstration I think the case studies need to be very well supported. In general I suggest bolstering the verification of the identifications for all the cases (not just this one) by including all relevant data. I mean specifically, first of all, other lidar quantities that have been used in previous classification methodologies, including especially lidar ratios, and also 1064 nm depolarization and angstrom exponents for all cases. Also include RH, backtrajectories and any coincident in situ measurements (especially pollen) for all cases.

L445 and Figs 15 and 16. The suggested mixing between layers doesn't look convincing. On the fluorescence vs. depolarization diagram, these intermediate points don't follow a nice mixing line like the other mixing cases, and the boundaries in the measurement curtains appear quite crisp. Could these points be artifacts of the smoothing instead?

Fig 15. The depolarization especially and perhaps also the fluorescence capacity (outside of the smoke plume) seems to be anti-correlated with backscatter, including in regions that seem unlikely to be pollen-dominated (such as the minimum between the smoke and urban layers). Particulate depolarization is especially susceptible to systematic error, particularly overestimation, at low values of backscatter (Freudenthaler et al. 2009, Burton et al. 2015). Have you done a systematic uncertainty calculation? (Also this is another case where color coding of the scatterplot by altitude would be useful).

Fig 15-18. Include the data for depolarization and angstrom exponent (and RH) for these cases also.

L455 "G_F increased ... probably due to the mixing with local pollution". Does this make sense? Nothing prior to this in the manuscript suggests that urban pollution has significant fluorescence capacity. Also, on the scatter plot on Figure 18, there's no suggestion that the higher values of G_F in the dust cluster are correlated with depolarization in any way; that is, they are not following any mixing line. What evidence is there that this is not simply normal variability within dust? Table 1 shows dust can have G_F up to 0.5. Why not 0.6? Also, could some of this variability be correlated with RH?

L485. "during Spring-Autumn seasons". It would be helpful to show a timeseries demonstrating that the pollen signature (elevated depolarization and fluorescence capacity) does NOT occur in winter.

Typographical or wording:

L19. What is meant by "single" in "first single version of the algorithm". I suggest delete "single" or reword.

L18 and L24. Change particle's to particle.

L92. Be specific about which wavelength here.

L247. Define LOA.

L270-281. There should be some discussion or at least references to other analyses of mixtures of aerosols that derive similar equations (especially Eq. 7), e.g. Sugimoto and

Lee 2006, Gross et al. 2011, Gasteiger et al. 2011, Tesche et al. 2009, Burton et al. 2014.

L280. Eq. 8. It probably would be good to remind the reader that fluorescence capacity and backscatter in this equation refer to particular wavelengths.

L282. "We assume". I think this is meant to refer only to the demonstration in Figure 3, not a general assertion. If true, perhaps swap the first two sentences of the paragraph to make it less likely to be misread. As mentioned in the introduction, the quantities have a lot of variability even within types, so assuming single values wouldn't be well supported.

L300. "the height resolution is 7.5 m". Is that really the resolution or only the grid spacing? That is, taking the detectors into account, are measurements at adjacent vertical grid points independent?

L342. spell out FBC

L344. Add a reference to the reminder. (I think it is Veselovskii 2020?)

L445. Typo in "0.2-0.3"

L663-664. It would be helpful to add "using the reference height as Ansmann et al. 1992 (green) or the calibration constant as in Eq 5. (magenta)". (I read figure captions before the text, so having a bit more detail in the captions is vey helpful)

L708. Please add clarification to the caption whether the scatter plot shows data for the entire time period shown in the curtain or only the subset that's included in the profile plots of Figure 8.

Figure 2 and 4. There is a lot of red in these plots hinting that the scales might be cutting off the data. Perhaps the scales should be expanded.

Figure 3. Also show the smoke + pollen mixing line, since one of the selected cases references mixing of those two types.

Figures 4, 6, 912, 15, 17. It would be helpful if the curtains of intensive properties (depolarization and fluorescence capacity) had consistent scales across each of these plots, making it easier to compare one case to another.

References:

Burton, S. P., et al. (2014) Separating mixtures of aerosol types in airborne High Spectral Resolution Lidar data. Atmos. Meas. Tech. 7, 419-436 DOI: 10.5194/amt-7-419-2014

Burton, S. P., et al. (2015) Observations of the spectral dependence of linear particle depolarization ratio of aerosols using NASA Langley airborne High Spectral Resolution Lidar. Atmos. Chem. Phys. 15, 13453-13473 DOI: 10.5194/acp-15-13453-2015

Freudenthaler, V., et al. (2009) Depolarization ratio profiling at several wavelengths in pure Saharan dust during SAMUM 2006. Tellus B 61, 165-179 DOI: 10.1111/j.1600-0889.2008.00396.x

Gasteiger, J. and V. Freudenthaler (2014) Benefit of depolarization ratio at λ = 1064 nm for the retrieval of the aerosol microphysics from lidar measurements. Atmos. Meas. Tech. 7, 3773-3781 DOI: 10.5194/amt-7-3773-2014

Groß, S., et al. (2011) Characterization of Saharan dust, marine aerosols and mixtures of biomass-burning aerosols and dust by means of multi-wavelength depolarization and Raman lidar measurements during SAMUM 2. Tellus B 63, 706-724 DOI: 10.1111/j.1600-0889.2011.00556.x

Sugimoto, N. and C. H. Lee (2006) Characteristics of dust aerosols inferred from lidar depolarization measurements at two wavelengths. Applied Optics 45, 7468-7474 DOI: 10.1364/AO.45.007468

Tesche, M., et al. (2009). "Vertically resolved separation of dust and smoke over Cape Verde using multiwavelength Raman and polarization lidars during Saharan Mineral Dust Experiment 2008." J. Geophys. Res. 114(D13): D13202.