Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC2
https://doi.org/10.5194/amt-2022-100-RC2, 2022

Comment on amt-2022-100

Anonymous Referee #2

Referee comment on "The effect of the averaging period for PMF analysis of aerosol mass spectrometer measurements during offline applications" by Christina Vasilakopoulou et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2022-100-RC2, 2022

The manuscript titled *"The Effect of the Averaging Period for PMF Analysis of Aerosol Mass Spectrometer Measurements during Off-Line Applications"* aims to understand how differently time resolved measurements affect the source apportionment results. This is an important problem to tackle as it will allow for systematic comparisons between online and offline source apportionment results and their effectiveness in interpreting real-world source contributions. As the authors pointed out, developing this understanding is especially important since it is not always possible to make online measurements.

The paper is well written in terms of language and the analysis is nicely structured. I am also happy with their input protocols for SoFi (lines 86-95). I do have some major concerns however, which are primarily rooted in data treatment/processing and explanation of causatives for observations. I think the paper needs to undergo major revisions before it could be considered suitable for publication. My comments are listed below:

1) One of my major concerns is that the paper in its current form reads like a commentary on what percentage variations where observed while running SoFi on differently averaged datasets. However, no insights are provided on likely factors controlling the variations. I see this as a little bit of a drawback since it is not helping the reader in using this paper as a reference in interpreting their own results. For example, in lines 174-177, COA is said to have an opposite behavior to HOA in low temporal resolution analysis. COA is overestimated during high COA concentration periods and underestimated during low concentration periods. However, no informed speculation is made for why this should be happening, which can help the reader.

I am guessing that this may be due to SoFi allocating more fraction of the total signal to

other factors when data-to-error ratio is lower during low concentration periods.

2) Another concern is regarding the data averaging methods employed in this study. The way PMF results would pan out depends on how the data was averaged across different time periods. The authors should discuss whether it was arithmetic or geometric average or weighted average etc. and how it influenced SoFi inputs. SoFi would produce robust results for high data-to-error ratios (e.g. high concentration periods) separating them more clearly in bootstrap tests. However, depending on how the peaks were handled during averaging, the data-to-error ratios could decrease influencing SoFi outputs and making them more uncertain. Given Figure 2, which shows many high concentration events for organic aerosol, a discussion on the averaging technique is very important.

3) Authors show variations in factor contributions broadly in the range of 6-15% across their analysis. Following from (2), I am curious about how significant are these numbers because they do not seem too large. This is encouraging for offline analysis. Could the differences be explained by error propagation in the averaging technique employed? The paper would benefit from a more thorough discussion on this. I recommend running a sensitivity analysis on this matter comparing PMF results from 30-min original data with results obtained for just the 24-h averaged using different averaging techniques, or at least discussing potential effects in the text.

4) In lines 154-159, the authors mention O:C ratios of MO-OOA and LO-OOA to be moving toward each other as the time resolution decreased. This seems to me an evidence of the averaging effect on the data. SoFi could separate the two better for higher time resolution measurements since the data captures temporal dynamics in concentrations resulting from ambient oxidation chemistry more clearly telling apart more from less oxidized. Consequently, SoFi can comfortably separate such contributions into their own factors. Is this reasonable thinking? It would be good if the authors shed more light on the O:C trends they noticed for the secondary components of their measurements.

5) How different would be a 24-hr filter sample data from a 24-hr averaged high time resolution online measurement? The authors should at least discuss the challenges

involved with replicating a 24-hr offline measurement this way. For example, in mathematical treatment of high resolution data, very short term concentration outliers (peaks or drops over few hours) can drag averages up or down to some extent in replicating lower resolution. However, such real-time, short-term concentration peaks or drops may not have the exact same impact on a filter sample being collected over a 24-hr period. Some insights on this issue would be helpful for a reader.

6) I like figures 5 and 6 a lot because they give a very nice overview of the results of this paper. It is very interesting that while % contributions for most individual factors appear to be within 2 standard deviations from average across the different time intervals, the theta angles are considerably different in some cases (e.g. HOA, COA and LO-OOA). Now, the primary to secondary component split is pretty similar across the different averaging intervals. So, looking at HOA, COA and BBOA in Figure 6, does this mean that the aerosol signal is being differently allocated by SoFi between the primary components for different averaging intervals?

7) For 2- and 4-hr averaging intervals, the theta angles for all factors except BBOA are considerably small. Based on figure S8, I assume that a theta angle less than 10 degrees would more or less replicate the factor profile obtained at the 30-min resolution measurement. Hence, I am curious about why the theta angle for BBOA is much higher than others at these averaging intervals even though its % contribution (Figure 5) only changes extremely minimally. Where does this variation in BBOA factor profile come from? Some discussion would be useful for the paper.

8) It is also not clear why: (i) the source apportionment for different time periods was not performed with constraining at least some primary factors, and, (ii) no assessment of uncertainty is made for solutions from differently averaged datasets through a bootstrap analysis. This is important.

9) The way this paper is written seems to suggest high time-resolution to be the truth and then checks for deviations from this "truth" by reducing the time-resolution. I am not sure whether this is the best approach to handling this comparison because both offline and

online techniques provide useful, scientific information in their own right. Also, comparison results could change a lot when looking at offline high spectral-resolution AMS data instead of ACSM. The authors should defend why high time-resolution data should be taken as the baseline for comparisons.

Additional comments:

- lines 15-20: It is important to discuss which percentages are closer to the truth in authors' opinion.

- lines 75-78: It should be mentioned somewhere here that extraction efficiency in offline analysis can play a role in causing differences.

- lines 94-95: The authors should at least briefly describe the process of choosing the optimum Fpeak.

- line 124: There is an older paper to cite for this: Identification of the mass spectra signature of organic aerosols from wood burning emissions, Alfarra et al, 2007 ES&T.

- line 186: The BBOA mass spectrum for the 24-hr average looks more comparable to literature than the 30-min averages where e.g. f43 is nearly the same as f73 which is unusual.

- lines 190-192: It is evident that LO- and MO-OOA will not compare well as the diurnal variability is driving the separation between the OOAs much more for the 30-min solution than for the 24-hr solution which is driven by day-to-day variability. If more seasons were used, then seasonal variability may have separated OOAs to some extent.

- line 253: As I said in (9) above, not necessarily. The dataset in this analysis uses ACSM but AMS used now-a-days provide high spectral-resolution data with more spectral information.

- In section 3.2, discuss how MO-OOA and LO-OOA are separated in terms of their spectral signatures.

-line 178: “underestimate”

-line 214: remove comma after “little”

-line 239: add “of” after “tendency”