Review of “Organic aerosol source apportionment by offline-AMS over a full year in Marseille” by Bozzetti et al.

This manuscript reported the chemical composition of $PM_{2.5}$ over al full year in an urban site in France, based on offline filter analysis. The focus is to study the sources of organic aerosol by performing PMF analysis on AMS data. The authors did very careful analysis to optimize PMF results. The authors found that BBOA is the dominant OA source in winter and OOA is the main OA source in summer. The authors also compared the offline-AMS results with online-AMS results, but the measurements were not performed simultaneously. The levoglucosan/BBOA ratio was found to evolve over time, which was attributed to different types of biomass burning combustions, instead of the photochemical aging. Overall, the analysis is adequate, the conclusions are generally solid, but none of the results are earth shattering or unexpected given the preexisting literature. I recommend accepting manuscript after major revisions.

Major comments.

1. Instrument inter-comparison.

For the filter analysis, some ions and species (such as SO_4, NO_3, WSOC, etc) are quantified by more than one techniques. It is natural to include the instrument inter-comparison in the manuscript. In one previous study of the authors (Bozzetti et al., 2017), the comparison of SO_4 concentration between offline-AMS and IC has a non-linear relationship. Does the non-linear relationship also exist in current study? Have the authors investigated more about the non-linear relationship since the previous study? Also, I wonder if the AMS_SO_4/IC_SO_4 ratio is similar to AMS_OC / WSOC ratio?

2. The interpretation of recovery ratio and extract the water solubility of OA factors.

If I understand correctly, the recovery ratio is a function of nebulizer efficiency (i.e., species loss during nebulization), AMS collection efficiency (i.e., include the lens transmission efficiency. The collection efficiency at the vaporizer may or may not be included, depending on whether the CE has been applied), and the water solubility of OA factors. Since the species are internally mixed in the solution and in the nebulized particles, the nebulizer efficiency and AMS collection efficiency should be the same for all OA factors. Thus, the recovery ratio only depends on the water solubility of OA factors. Figure S12 shows that OOA has the largest recovery ratio and HOA has the smallest recovery ratio, which is consistent with that OOA is more water soluble than HOA. Thus, this provides a potential opportunity to estimate the water-solubility of OA factors. For example, could the authors make use of the instrument inter-comparison (i.e., SO_4 between offline AMS and IC) to correct for the efficiency of offline AMS system, as similarly done in Xu et al. (2016)? Then, use the model proposed by Psichoudaki and Pandis (2013) to relate the fraction of a compound extracted in WSOC as a function of compound water-solubility at dissolution equilibrium. The model in Psichoudaki and Pandis (2013) is designed for filter extraction analysis. In this way, more useful information about OA factors can be extracted.

3. The selection of PMF solutions is very careful, but some related descriptions require more clarifications.

(1) Did the author constrain all OA factors or only COA and HOA? More importantly, it should be clearly stated that how the anchor profiles are selected? If the authors use the average reference mass spectra of HOA and COA as anchor profiles, would it make the analysis easier?

(2) Did the authors constrain industry-related OA (INDOA) in the offline analysis? The INDOA factor is resolved in whole year offline dataset and 2008 July online dataset, but not in 2011 February online dataset. Have the authors tried to constrain the INDOA for the 2011 February online dataset?

(3) The application of cluster analysis to select PMF solutions is very nice, but the description of cluster analysis is not clear in the main text. For example, what does “PMF solutions” refer to in Page 10 Line 33? Please be more clear that “PMF solution” is a full set of solution (i.e., including both time series and mass spectra of all OA factors). I suggest to remove some descriptions from the SI to the main text.

(4) After all the discussions on the optimization of PMF solutions, it is not clear what is the final PMF results. Did the authors use the average of all retained PMF solutions? It would be useful to comment on how the finally optimized PMF solution is different from that by using PMF2 solver without any constrain, since the PMF2 solver is most widely used method.

4. The comparison between online and offline measurements.

(1) It would be useful to include a table to summarize the sampling periods of online and offline measurements. Page 20 Line 4, it is very misleading to claim this study as the first comparison between HR online AMS and offline AMS, since the online and offline measurements are not simultaneous. Please rephrase.

(2) In Figures 7 and 8, how are winter and summer defined for offline AMS measurements? How many filters are included in winter and summer? The comparison in concentration looks generally good, but there are many disagreements as well. For example, INDOA is resolved in offline dataset, but not in 2011 February online dataset. I suggest to include all factors in the figures, instead of only including the overlapped factors.

(3) Page 17 Line 24-25, HOA from offline analysis does not correlate with NO_x. This is consistent with 2008 July online dataset (as noted in the main text), but not consistent with 2011 February online dataset (i.e., R = 0.86 in Page 16 Line 6). Please comment on this discrepancy. Btw, does OOA from offline analysis correlate with NH_4^+?

(4) Were filter samples collected during 2011 February? If not, how are levoglucosan and vanillic acid measured for this period (Page 23 Line 23-24)? If yes, could the authors analyze them using offline-AMS and compare to simultaneous online AMS?

Minor comment.

1. Eqn. (11). In Bozzetti et al. (2017), there is a term TEOC (traffic emission OC) in the equation. Please discuss the rationale to replace TEOC with WSHOA/(R_{HOA}*OM/OC) in current study. Also, WSOC is mentioned in Page 14 Line 17, but it is not clear how the WSOC is used in determining R_k.
2. Figure 1. Please show the detailed time series of stacked $PM_{2.5}$ compositions and the measured total $PM_{2.5}$ concentrations.
3. Figure 10. How and why are the data smoothed?
4. The discussions on PAH sources are confusing. Based on the discussions in Page 16 Line 14-29, it is likely that PAH is mainly from biomass burning for 2011 February. However, in Page 18 Line 10-31, PAH is mainly from industry based on offline analysis on the whole year. Please clarify related discussions. Also, does AMS-PAH agree with filter PAH?
5. In Figure S18, why are there so many negative $CO_2^+{}_{OOA}$ values?
6. In the abstract, please mention that 216 filter were collected over a full year, but only 58 filters were analyzed.
7. Page 9 Line 7. What is the difference in signal between measurement blank and a sample with relatively low loading?
8. Page 14 Line 3, R_{COA} is not defined yet.

9. Page 14 Line 6-7. I believe KOA_i and $WSKOA_i$ should be swapped.

10. Page 15 Line 23. The shallower BLH and less dispersion also likely contribute to higher OM in winter.

11. Page 17 Line 10. What's the correlation between OOA and SO_4 time series?

12. Page 18 Line 15. Replace "INDOA" with "industry". Replace "BBOA" with "biomass burning".

13. Page 20 Line 3. Remove "one".

14. Page 20 Line 19-21. It is a bit surprising that the factor representing 5% of total OA concentration can be resolved from offline analysis, but not online analysis.

15. Page 23 Line 4-5. What's the concentration of nitrate radical and hydroxyl radical for the corresponding lifetimes?

16. Page 23 Line 8-14. It is not clear what the key points are. Are the authors suggesting that BBOA has secondary OA? Or the BBOA/OOA separation is comparable between online and offline measurements, regardless of different temporal resolution? Please rephrase this paragraph as well as Page 26 Line 18 in the conclusions.

17. Page 24 Line 8-10. Please show the correlation between nss-K^+ and syringic acid in summer. In order to be consistent with previous argument (Page 24 Line 5) that nss-K+ is mainly from dust in summer, the correlation has to be poor.

Reference

Bozzetti, C., Sosedova, Y., Xiao, M., Daellenbach, K. R., Ulevicius, V., Dudoitis, V., Mordas, G., Byčenkienė, S., Plauškaitė, K., Vlachou, A., Golly, B., Chazeau, B., Besombes, J. L., Baltensperger, U., Jaffrezo, J. L., Slowik, J. G., El Haddad, I., and Prévôt, A. S. H.: Argon offline-AMS source apportionment of organic aerosol over yearly cycles for an urban, rural, and marine site in northern Europe, Atmos. Chem. Phys., 17, 117-141, 10.5194/acp-17-117-2017, 2017.

Psichoudaki, M., and Pandis, S. N.: Atmospheric Aerosol Water-Soluble Organic Carbon Measurement: A Theoretical Analysis, Environ Sci Technol, 47, 9791-9798, 10.1021/es402270y, 2013.

Xu, L., Guo, H., Weber, R. J., and Ng, N. L.: Chemical Characterization of Water-Soluble Organic Aerosol in Contrasting Rural and Urban Environments in the Southeastern United States, Environ Sci Technol, 10.1021/acs.est.6b05002, 2016.