EGUsphere, author comment AC1
https://doi.org/10.5194/egusphere-2022-269-AC1, 2022

Reply on RC1

Marta Via et al.

Author comment on "*Rolling* vs. *seasonal* PMF: real-world multi-site and synthetic dataset comparison" by Marta Via et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-269-AC1, 2022

Reply to Anonymous Referee #1 (Egusphere-2022-269)

Received and published: 23rd June 2022

The authors would like to thank the reviewer for all the comments and suggestions, which helped improving the quality of this work. A new version of the manuscript has been prepared following the suggestions from the reviewers. We provide below detailed replies to each of the comments in a point□by□point manner. Figures and tables cited in the following document can be found in the supplementary annex.

Comments from the reviewer

The manuscript by Marta Via et al. performed a comprehensive comparison between the two methodologies of fine organic aerosol (OA) source apportionment through the Positive Matrix Factorization (PMF) model: rolling and seasonal PMF. They found that the rolling PMF can be considered more accurate and precise, globally, than the seasonal one, although both meet the standards of quality required by the source apportionment protocol. In addition, the results showed that the selection of anchor profiles is highly influencing the OA factors, so local reference profiles are encouraged to minimise this impact. The topic fits well within the scope of Atmospheric Measurement Techniques. Overall, the data analysis is solid and the manuscript is clearly written. Before its publication, the following comments need to be addressed.

- **Line 308: What are the contributions of SOA species to m/z 55? Looking into these datasets would be helpful to evaluate the uncertainty of using m/z 55 as a marker for HOA.**

According to *rolling* PMF results, the m/z 55 is attributed to each source in a percentage depicted in the table R.1.1 (in the supplementary document).

As can be perceived from the table, HOA is always the main contributor to m/z 55 except for those sites in which COA is present, in which cases, HOA is in second place, and in Dublin, where the main contributor to COA is PCOA. These results demonstrate that the posterior assumption that the m/z 55 can be used as a marker for HOA is sufficiently accurate. Besides that, SOA (LO-OOA + MO-OOA) apportionment of m/z 55 does not exceed the 30% of apportioned m/z 55 and the ratio m/z55_{HOA}-to-m/z55_{SOA} is always superior to 1.5 (up to 3.4) except for the Dublin site, in which the combustion sources are significant sources of this ion. This brief analysis proves that m/z55 can indeed be used as a marker for HOA in the multi-site comparison.

Regarding the synthetic dataset (Table R.1.2.), the apportionment of this ion is more equally distributed, but one should take into consideration that SOA_{tr} might have grabbed some HOA markers and therefore its prevalence in HOA must be higher. Hence, not by far, but HOA is yet in the synthetic dataset the highest contributor of the m/z55.

For these reasons, the following sentence (starting at line 356) can be left unchanged:

The adaptability of the models can be assessed from Figure 3 (b), where the 60/55 vs. 44/43 (which are proxies for the BBOA-HOA differentiation and the SOA oxidation, respectively) is plotted for the truth and both methods.

- **Are comparisons of Rolling vs. Seasonal PMF dependent on the type of site (e.g., Urban Background, Suburban) and/or the instruments (i.e., Q-ACSM and ToF-ACSM). Please be specific.**

In both site-dependent and instrument-dependent, the groups are too unbalanced in terms of the number of sites to assure a proper assessment of the group differences.

Regarding the type of site:

- Urban background / Suburban / Peri-urban (7): BCN-PR, DUB, ATOLL, INO, MRS-LCP, SIR, TAR.
- Remote/Rural (2): CAO-AMX, MAG.

Regarding the type of instrument:

- Q-ACSM (8): BCN-PR, CAO-AMX, DUB, ATOLL, INO, MAG, SIR, TAR.
- ToF-ACSM (1): MRS-LCP.

Therefore, the authors believe that this matter should not be addressed as the results would not be sufficiently conclusive.

- **Figure 3: I noticed that there are three distinguished lines in the triangle plot of f44 vs. f43 using seasonal PMF data, while this phenomenon does not appear using the rolling PMF and the truth PMF. Please, describe and explain these differences in detail.**

Figure 3(c) (or in the supplementary document R.1.3.) is a triangle plot (Ng et al., 2010), in which the SOA profile f44 vs. the SOA profile f43 concentrations are shown. This plot aims to compare the time series behaviour of the truth vs. the profile concentrations of the rolling PMF (in red) and seasonal PMF (in blue) results. The big grey dots represent the SOA profile concentrations of the truth, both presenting the three SOA sources (round dots) and the weighted mean of all of them (squared dot).

The rolling PMF dots move around the triangle plot, nevertheless, the seasonal dots are shown as three different blue lines. The seasonal PMF has been applied in this analysis in three different seasons, therefore we get three different profiles, one per season. Each of these profiles has a concrete f44-to-f43 ratio, which could be represent in three dots in this plot. Nevertheless, here we show the time series adaptation (or lack of adaptation in the seasonal case), this is why the plot presents the profile multiplied by the time series concentrations. This way of plotting the data, in the seasonal PMF case, gets represented into three different lines whose slopes correspond to the three different f44-to-f43 ratios of the three seasons. The points in each of the lines represent the i-th time point concentration f44-to-f43, which all concur in the slope established for each season. In contrast, the rolling PMF dots show unconstrained mobility inside the triangle.

- **The sampling period that appears in Figure 2 is not in Table 1 (Participant sites). Please have a check.**

The synthetic dataset sampling period has not been added in the participant sites (Table1) because it did not use real data in its time series (coming from the CAMx model) but it did use the Zurich ToF-ACSM data to generate the error time series mimicking the procedure of error calculation of a real-world dataset. The authors consider that for the sake of clarity, this Zurich dataset should not be included in the participant site table, in order to differentiate the synthetic dataset as an standalone approach. However, we acknowledge some more information of the sampling period should be provided by the authors.

The following sentence has been added at section 2.2.

For this purpose, the dataset used is that from the Zurich site which ranges from February 2011 until December 2011. Hence, the same CAMx outcoming time series period was used to generate the concentration matrix.

- **It would be better to change the name of "truth PMF", because no one knows what the truth is like, and our goal is to pursue infinite access to the truth.**

The label of "truth" has been used only in the synthetic analysis to refer the synthetically created dataset, consisting on the time series and the profiles of five OA sources. This dataset has been arranged as a matrix, in order to be the starting point to launch the OA source apportionment in both rolling and seasonal PMF modelling. The outcomes of each methodologies are compared to the initial data in this study, which is actually a proxy for the truth, as the matrix is directly the multiplication of the synthetically designed.

Hence, the use of "truth" in this article does not imply that the synthetic dataset is a representation of the atmosphere. The labelling of the synthetic dataset as "truth" intends to remark what the results would be if the model could represent with perfect effectivity the input matrix. This is why we the authors would like to maintain the "truth" labelling.

- **The figure captions for each panel should be clearly stated. Take Figure S4 for example, what is SHINDOA representing? In addition, '"58-OOA" must be defined.**

Some figure and table captions have been rephrased in order to promote a quicker understanding of the plot content.

Figure 1. OA apportionment results for rolling and seasonal methods and truth output.

Figure 2. Rolling, seasonal and truth (synthetic dataset original values) (a) time series (in hourly averages for the sake of clarity), (b) diel profiles and (c) scatter plots.

Figure 3. Synthetic dataset solution (a) Profiles; (b) Time-dependent profile variability of ratios 60/55 vs. 44/43; (c) Triangle plot of f44 vs. f43; for rolling PMF (red), seasonal PMF (blue) and truth (black).

Figure 4. Pie charts of the mean concentrations of the main factors for the ensemble of all sites.

Figure 8. Kernel density estimation of the histograms of the subtraction of the m/z 44-to-43 ratio from the raw (from input matrices) time series data minus the apportioned quantity profiles. These plots only contain those time lapses among the change of season (transition periods).

Table S1. Multi-site assessment dataset characteristics.

Figure S54. Pie plots for rolling and seasonal source apportionment solution for each site. The factor acronyms correspond to: Hydrocarbon-like OA (HOA), Biomass Burning OA (BBOA), Less Oxidized Oxygenated OA (LO-OOA), More Oxidised Oxygenated OA (MO-OOA), Cooking-like OA (COA), Peat Combustion OA (PCOA), Coal Combustion OA (CCOA), Wood Combustion OA (WCOA), 58-related OA (58-OA) and Shipping + Industry OA (SHINDOA).

Figure 6 was changed slightly for the sake of clarity and its caption was modified as follows:

Figure 6. Rolling and seasonal boxplots of the Pearson-squared correlation coefficient of OA sources with their respective markers for all sites.

References

Ng, N. L., Canagaratna, M. R., Zhang, Q., Jimenez, J. L., Tian, J., Ulbrich, I. M., Kroll, J. H., Docherty, K. S., Chhabra, P. S., Bahreini, R., Murphy, S. M., Seinfeld, J. H., Hildebrandt, L., Donahue, N. M., Decarlo, P. F., Lanz, V. A., Prévôt, A. S. H., Dinar, E., Rudich, Y. and Worsnop, D. R.: Organic aerosol components observed in Northern

Hemispheric datasets from Aerosol Mass Spectrometry, Atmos. Chem. Phys., 10(10), 4625–4641, doi:10.5194/acp-10-4625-2010, 2010.

Please also note the supplement to this comment:
https://egusphere.copernicus.org/preprints/2022/egusphere-2022-269/egusphere-2022-269-AC1-supplement.pdf