Editor comments response in orange

Thank you for your response to the comments by Reviewer 2. I think all the comments have been addressed but some additional info and explanation regarding the Q/Qexp question would be useful. Specifically, as the reviewer noted, some individual m/z have very high Q/Qexp, up to 300, but the overall Q/Qexp is only around 5. It is unclear why this is the case. It appears that the individual m/z having high Q/Qexp are HOA type ions. Have you checked the input errors for PMF manually, and evaluated if these input errors are too low? If not, it would be useful to double check the errors of these m/z manually.

All these info would be useful for the readers to know to better understand how the PMF solution was obtained. Please explain these observations (very high Q/Qexp for some m/z, but an overall Q/Qexp of around 5), you can keep Figures S10 and S11 in the SI, and add the explanations to the figure captions.

We appreciate the effort from the editor to provide these comments that will increase the quality of the manuscript.

We looked into the input errors for m/z 55 and m/z 57 and they do not have too low values.

We use the Source Finder (SoFi version 4.8) tool (Canonaco et al., 2013) to run PMF. SoFi gives three types of Q/Qexp values to look into detail on the solutions. One overall Q/Qexp, it also gives a row (time series) Q/Qexp and a column (m/z) Q/Qexp. The overall Q/Qexp is calculated as the mean of the time series Q/Qexp.

We looked into detail on the Q/Qexp values for all the PMF analyses and also PMF analyses from previous projects. The overall Q/Qexp and the time series Q/Qexp are correct. However, in all the analysed datasets the m/z Q/Qexp values are similarly high. We identified that dividing the values in figure S11 by the Q/Qexp and by the number of m/z variables we estimate a value similar to the overall Q/Qexp. We consider SoFi 4.8 might have a bug or missed a calculation to estimate the Q/Qexp for m/z. Currently there is a new version of SoFi 6.0, we will see if the Q/Qexp for m/z are reported correctly, otherwise we will contact the developers. We are using the old SoFi 4.8 because we have developed our own code to explore the solution space from runs perfomed in SoFi 4.8.

We took the idea of dividing by Qexp divided by number of species from the EPA PMF 5.0 fundamentals user guide that states: “the Q/Qexp for a species is the sum of the squares of the scaled residuals for that species, divided by the overall Qexpected divided by the number of strong species”. The change we are doing to figure S11 is change the label Q/Qexp for Squares of scaled residual.

We calculate the sum of the values (Squares of scaled residual) plotted in Figure S11 for the specific PMF run. For example, for the optimal solution (PMF_5F_S2) the sum = 11269.56.

We calculate the Qexp = n * m - p*(n + m) = 147127 for the 5-factor solution, where

n = num of samples = 2169, m = num of m/z = 73, p = num of factors = 5

We also divide the Qexp by m, which gives 2015.44.

Finally, Q/Qep = 11269.56/2015.44 = 5.59, which is close to the overall Q/Qexp = 5.2.

We did a similar estimation for other datasets; for example, for the HR-AMS in ND preMonsoon we estimated a new Q/Qexp = 1.70, with n = 40542, m = 77 and p = 5, when the overall Q/Qexp = 1.60.

We have updated the scale of figure S11 for from 300 to 600 and we can see that the highest values are for m/z 15 and m/z 38 (around 550), while m/z55 and m/z57 have values of around 350. Figures S10 and S11 are used to identify particularly high Q/Qexp values for m/z or episodes on time series that might require attention. We are more interested on the relative contribution rather than on the absolute values, hence we consider that this change on figure S11 does not impact on the selection of the optimal solutions. For example, we can see an improvement on both time series and m/z plots when going from 4-factor to 5-factor solutions. However, no remarkable improvement between 5-factor solution was observed. Hence, we can use the overall Q/Qexp to select the optimal solution.

We have kept figures S10 and S11 in the SI as suggested with the following explanation:

We added this paragraph to the SI in line 133

The PMF analysis was performed using the the SoFi tool version 4.8 (Canonaco et al., 2013). In order to select the optimal PMF solution, we analysed the overall Q/Qexp (Figure S9), the row (time series) Q/Qexp (Figure S10) and the Squares of scaled residual (Figure S11).

We added this paragraph to the SI in line 150

The overall Q/Qexp =5.2 is calculated as the mean of the time series Q/Qexp. The overall Q/Qexp can also be estimated from the Squares of scaled residuals. We calculate the sum of the values (Squares of scaled residual) plotted in Figure S11 for the specific PMF run. For example, for the optimal solution (PMF_5F_S2) the sum = 11269.56.

We calculate the Qexp = n * m - p*(n + m) = 147127 for the 5-factor solution, where

n = num of samples = 2169, m = num of m/z = 73, p = num of factors = 5

We also divide the Qexp by m, which gives 2015.44.

Finally, Q/Qep = 11269.56/2015.44 = 5.59, which is close to the overall Q/Qexp = 5.2.

Figures S10 and S11 are used to identify particularly high Q/Qexp values for m/z or episodes with high Q/Qexp on time series that might require further analysis. For example, we can see an episode with high Q/Qexp values on 20/20/2018 (Figure S10), which is related to an episode with high Org concentrations. Also, we can see an improvement on both time series (Figure S10) and m/z (Figure S11) plots when going from 4-factor to 5-factor solutions. However, no remarkable improvement between 5-factor solution was observed. Hence, we can use the overall Q/Qexp to select the optimal solution.

However, if the editor considers these details could confuse the reader, we can remove Figure S11 from the supplement.

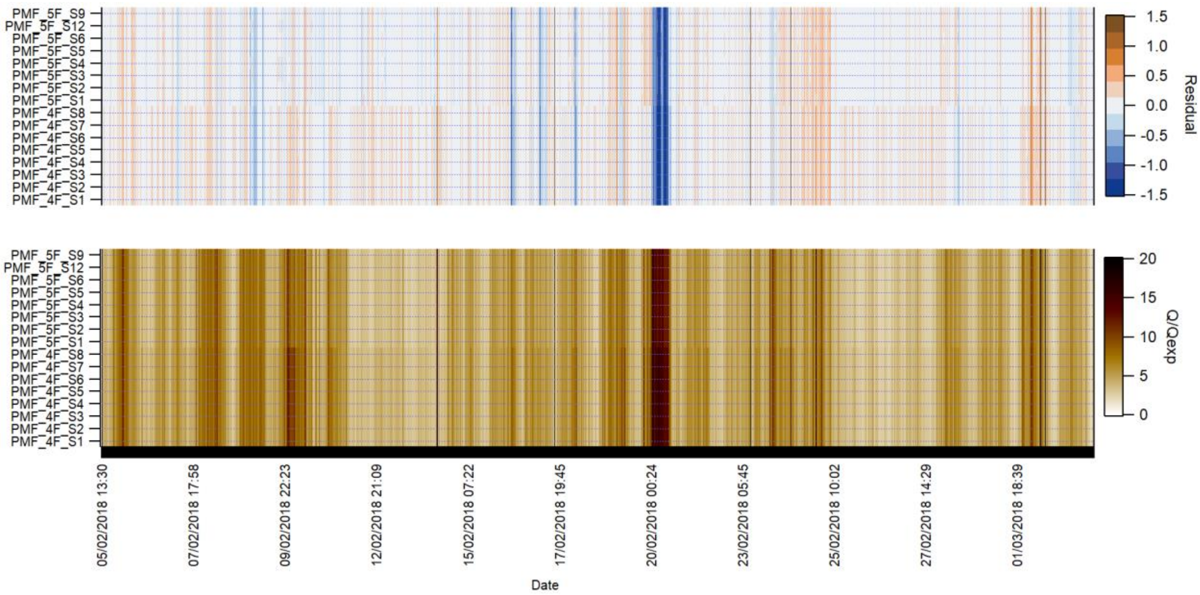

Figure S1. Time series of residuals and Q/Qexp values.

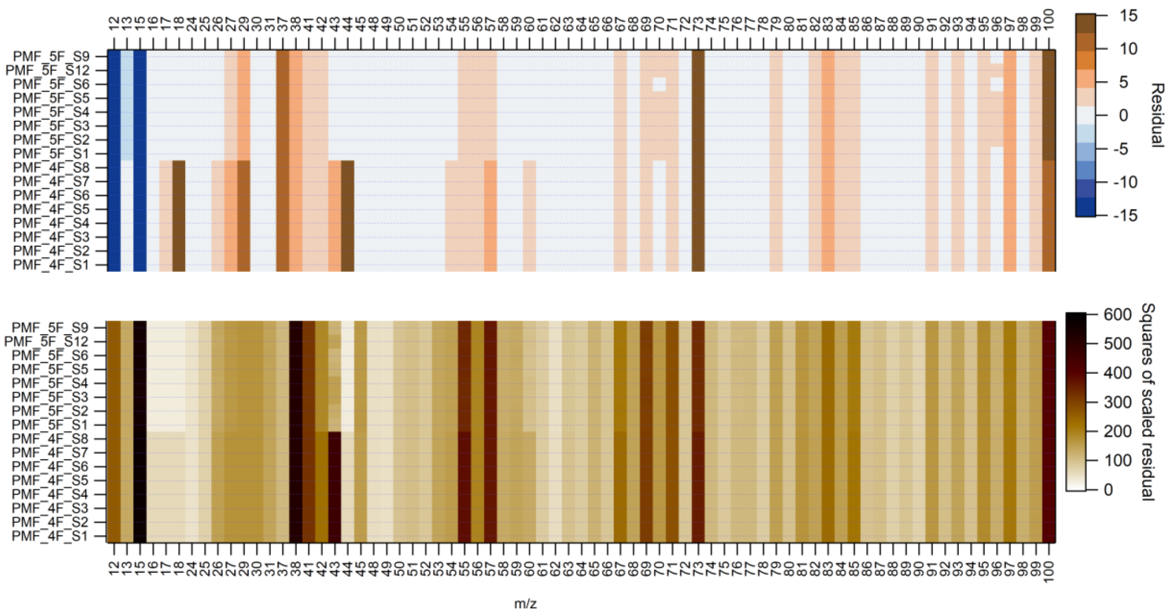

Figure S2. Residuals and Squares of scaled residual for m/z.

Canonaco, F., Crippa, M., Slowik, J. G., Baltensperger, U., and Prevot, A. S. H.: SoFi, an IGOR-based interface for the efficient use of the generalized multilinear engine (ME-2) for the source apportionment: ME-2 application to aerosol mass spectrometer data, Atmos Meas Tech, 6, 3649-3661, DOI 10.5194/amt-6-3649-2013, 2013.