Enhancing non-refractory aerosol apportionment from an urban industrial site through receptor modelling of complete high time-resolution aerosol mass spectra

Supplementary Material

M. L. McGuire[1], R. Y. –W. Chang[1,2]*, J. G. Slowik[1,2*], C. –H. Jeong[1], R. M. Healy[1], G. Lu[3], C. Mihele[3], J. P. D. Abbatt[1,3], J. R. Brook[3], G. J. Evans[1]

[1](Southern Ontario Centre for Atmospheric Aerosol Research, University of Toronto, 200 College St., Toronto, Ontario, Canada)

[2](Department of Chemistry, University of Toronto, St. George St., Toronto, Ontario, Canada)

[3](Air Quality and Research Division, Science and Technology Branch, Environment Canada, 4905 Dufferin St., Toronto, Ontario, Canada)

*Now at (School of Engineering and Applied Sciences and Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, United States)

*Now at (Paul Scherrer Institut, Villigen, Switzerland)

Correspondence to: G. J. Evans (greg.evans@utoronto.ca)

S1 PMF$_{\mathrm{Full\ MS}}$ analysis solution justification

The six factor solution was chosen to represent the AMS data as it fulfilled all mathematical diagnostic criteria and provided the most physically meaningful results. The factors were as follows: Amine, Sulphate-OA, Nitrate-OA, Chloride, HOA, and OOA. Comparisons between the Amine factor MS with MS from amine compounds reported in the NIST library showed that the Amine factor's profile was most similar (r = 0.55) to the NIST electron impact ionization (EI) mass spectrum for triethylamine (TEA: $C_6H_{15}N$, 101 g mol^{-1}) (Figure S-1.1) (Stein, 2013). The six factor PMF solution was robust with a global minimum Q value of 227,508 found from 100 runs initiated from random seeds. Furthermore, bootstrapping indicated that all factors demonstrated successful re-mapping to the base run (at least 99/100 runs for each factor). A $Q_{robust}/Q_{exp} = 0.79$ was calculated for this solution, indicating that the uncertainties provided to the PMF model were appropriate for the data. The calculated r^2 for the solution was 0.996, which was a very high degree of reconstruction. Figure S-1.2 shows that the Q value levels off at approximately six factors, indicating that the increase in explained variance on adding factors decreases beyond this point. While a six factor solution was chosen for analysis, solutions ranging from two to ten factors were explored. Figure S-1.2a highlights the $\Delta Q/Q_{exp}$ with an increase in number of factors, where a significant change from the previous factor addition indicates that the newly added factor has captured meaningful variability from the residuals. The amount of variability captured began to decrease with the addition of the seventh factor, although a portion of some large "spikes" were accounted for by the addition of the seventh and eighth factors. These spikes were largely attributed to transient elevated SO_4^{2-}, which was not fully captured by the Sulphate-OA factor in the six factor solution. From Figure S-1.2b, it is evident that solutions beyond eight factors failed to capture any significant additional variability, as compared to PMF solutions of lower order.

The degree of correlation between factor mass spectra and time series was also examined (Figure S-1.3), as examining incremental changes in factor similarity with the addition of new factors can provide indication of where a solution may no longer be deemed acceptable. The five through eight factor solutions all showed reasonable separation in terms of the factor time series, as correlation coefficients were all around 0.6 or lower. In terms of mass spectra, some solutions showed a high degree of correlation in terms of factor mass spectra. For instance, in the five factor solution, the Sulphate-OA and Chloride factors were highly correlated in terms of mass spectra, but not in temporality. Such a high correlation in terms of factor profiles may be explained by these two factors being defined by only a few peaks, with the rest being reasonably low, and thus highly similar. With factor mass spectra defined by so few peaks (particularly for inorganic factors), comparing degree of correlation in terms of temporality was more reasonable. Nonetheless, examination of seemingly split factors, provided useful information. For instance, the Local Sulphate and Sulphate-OA factors were reasonably similar in terms of mass spectra, as expected, yet showed low temporal correlation, indicating they were reasonably different. However, the OOA and Other OA factors in the eight factor solution showed that they were highly correlated, both in terms of mass spectra and time series. When examining the chosen six factor solution, it can be seen that none of the factor combinations show the same high degree of correlation for both mass spectra and time series, providing further justification for this solution.

Before the six factor solution was fully accepted, it was further investigated through forced global rotations using the FPeak parameter. This methodology was used to examine for rotational ambiguity, and also to determine whether another rotation appeared more suitable. Figure S-1.4 shows how Q_{robust}/Q_{exp} varies from an FPeak of -10 to 10. It is evident that the central rotation provides the lowest value, indicating the lowest residuals and best fit for the data. Despite the wide range of tested FPeak values, the range of variation in Q_{robust}/Q_{exp} was reasonably low (0.793-0.805). Rotations were also assessed through the effect on resulting mass spectra according to the following observations: the variation in the fractional contribution of *m/z* 44 (F44) for factors containing significant oxidized organic content (F44 > 0.05 for the base run); the extent of neutralization for factors defined by salts; and the resulting change in mass spectra and associated compositional change. Table S-1.1 lists the effect of FPeak variation on F44. FPeak variation had negligible influence on the Sulphate-OA and Nitrate-OA factors as there was negligible change in the F44 between FPeaks of -10 and +10. A slight effect was observed for the OOA factor, whereby the F44 showed slight variation with changes in FPeak. Overall, variation in the FPeak did not significantly alter the F44 contribution for any of the factors with notable oxidized organic aerosol material, and the variation still fell along the range of values reported for known OOA factors from around the world (~0.10-0.20) (Ng et al., 2011). Table S-1.2 shows the effect of FPeak variation on the extent of neutralization for the salt factors. It is possible that rotational uncertainty is partly associated with uncertainties in the RIEs of inorganic species, as default values were assumed in this study. As such, it was difficult to use $Neut_{ext}$ values in an absolute manner to assess whether an alternate solution was preferred. However, an assessment of values in a relative sense indicated that $\mathrm{Neut}_{\mathrm{ext}}$ values tended to increase away from the central rotation with both positive and negative rotations, towards more basic values , indicating the central rotation appeared the most reasonable. Ultimately, FPeak analysis indicated that between the inorganic factors, the Chloride factor appeared to be the least rotationally fixed, showing the greatest degree of variation in composition among all factors, as most evident in Figure S-1.5.

In addition to the three aforementioned metrics, the resulting mass spectra from FPeak rotations were examined and compared with other known spectra from the AMS mass spectral library (Ulbrich et al., 2010), and are shown in Figure S-1.6. Again, comparisons were only made between factors for which a reliable AMS reference spectrum existed (i.e., NH_4NO_3,[1] $(NH_4)_2SO_4$[2], HOA[3], and OOA[4]), and where correlations with reference spectra were greater than 0.5 (i.e., not the case between the Amine factor and TEA). Only a very minor improvement in correlation was noted for the FPeak of +2 value, although this was not deemed sufficient to justify use of this value.

The variation in degree of correlation between the HOA factor and key external tracers was also examined as a function of FPeak (i.e., NO, BC, Benzene, CO, and particle number concentration) (Figure S-1.7). In each case, the degree of correlation between time series was found to be slighly better with low positive rotations, although the difference between the base case and any of the positive rotations was not sufficient to justify choosing an alternate rotation over the base case. Overall, FPeak rotations revealed some rotational ambiguity in the solution, mostly associated with the Chloride factor. Furthermore, no significant justification could be found for the selection of an alternate solution other than the base run, and as such, the central FPeak rotation (FPeak = 0) was chosen.

To validate the six factor solution, the adjacent five and seven factor solutions were examined. Furthermore, the eight factor solution was also considered due to the presence of an interesting sulphate factor. The solutions were examined for stability, and for physical meaningfulness of the factor mass spectra. The latter was assessed through visual inspection, extent of neutralization, as well as from correlation analysis with reference spectra and key

---

[1] AMS Mass Spectral Database: L_STD_Q_001
[2] AMS Mass Spectral Database: L_STD_Q_002
[3] AMS Mass Spectral Database: A_DEC_Q_012
[4] AMS Mass Spectral Database: A_DEC_Q_013

external time series (Figure S-1.8 and Figure S-1.9). An effective RIE was applied to the Amine factors in these alternate solutions as in the chosen six factor solution. Figure S-1.10 presents the factor mass spectra and time series from the five factor solution. The factors extracted in this solution were largely similar to those from the six factor solution, which was partly reflected in the similarities in extent of neutralization for the salt factors (Table S-1.3). Correlation analysis of factor MS showed that the spectra were similar to those from the six factor solution, except that the five factor HOA factor MS correlated slightly better with the reference HOA spectrum. However, a notable difference was the absence of an OOA factor: instead, the dominant OOA peak, *m/z* 44, was apportioned mainly among the Sulphate-OA, Nitrate-OA and Chloride factors. The lack of an OOA factor appeared to render the solution unstable, as no global minimum Q value could be found. Instead, the 100 random seed runs were split nearly equally between two different global minimum Q values. This instability was taken as a reflection of the robustness of the OOA factor extracted in the six factor solution.

Figure S-1.11 presents the factor mass spectra and time series from the seven factor solution. While there was a slight increase in the correlation between the OOA factor and formadehyde at seven factors (Figure S-1.9), there was a significant deterioration in the mass spectral correlations for the HOA and OOA factors and reference mass spectra (Figure S-1.8). However, as compared to five factors, the seven factor solution showed a greater degree of stability, as a global minimum solution was found. While this stability suggested that seven factors may have been better than five, two undesirable effects were noted. First, the extent of neutralization for the Chloride factor deviated significantly from unity, as the ratio of cations to anions increased from 1.09 to 1.26 (Table S-1.3). Although the RIE value for Chloride was not precisely known for this study (the default value was assumed), this relative increase still suggested that the solution was less ideal; in the six factor solution, all of the inorganic factors appeared neutral, and more importantly, showed the same relative degree of neutrality. In this seven factor solution, this was no longer the case. Second, the additional factor caused the OOA factor to be split into two less physically meaningful sub-factors: the factor appearing most similar to the OOA factor identified in the six factor solution showed a drop in correlation with the reference mass spectrum (Figure S-1.8), and the “Other OA” factor identified showed no significant correlation ($R > 0.5$) with any reference mass spectra from the AMS mass spectra library. The effect of the split OOA factors is further discussed in section 3.1.6.

The eight factor mass spectra and time series are shown in Figure S-1.12. While the eight factor solution did comprise of a global minimum as determined from PMF initiation from 100 random seeds, ultimately it could not be accepted as the splitting of the Sulphate-OA factor led to two Sulphate factors that could not be fully justified. In addition to a sulphate factor that appeared mainly influenced by synoptic flow (named Regional Sulphate), an additional factor, termed Local Sulphate, appeared to better capture the short-lived SO_4^{2-} spikes (Figure S-1.12). An analysis of the molar ratios of the cations to anions (Table S-1.3) showed that the Local Sulphate factor appeared to be highly acidic (ratio = 0.25), consistent with the short-lived, and likely local nature of the SO_4^{2-} spikes that it mainly represented. While the Local Sulphate factor appeared to capture additional variability not captured in the six or seven factor solutions, its extraction still could not be justified for three reasons. First, as shown in Table S-1.3, the Regional Sulphate and Chloride factors in the eight factor solution now appeared to be overneutralized (Regional Sulphate neutralization ratio = 1.27, and Chloride neutralization ratio = 3.05). Again, while the absolute value of the extent of neutralization could not be used due to uncertainties in inorganic species relative ionization efficiencies, there was a notable shift in these values while the reasonably robust Nitrate-OA factor remained stable, and close to neutral. Second, the “Other OA” factor which was also extracted in the seven factor solution, was still present, and still did not appear physically meaningful. Third, correlations mass spectral and temporal correlations for the HOA and OOA factors in terms of reference spectra and time series were not improved as compared to lower order solutions (Figure S-1.8 and Figure S-1.9). As a result, while the eight factor solution appeared interesting in terms of the existence of an acidic Local Sulphate factor, extraction of this factor altered the other factors to an unjustifiable extent. Thus, due to the physical meaningfulness of the factor profiles, as well as the mathematical robustness of the solution, the six factor solution was deemed most acceptable.

Table S-1.1: Variation in F44 with FPeak rotation for the $\text{PMF}_{\text{Full MS}}$ six factor solution.

FPeak	Sulphate-OA	Amine	Chloride	Nitrate-OA	HOA	OOA
10	0.15	0.04	0.00	0.13	0.00	0.09
8	0.15	0.04	0.00	0.13	0.00	0.09
6	0.15	0.04	0.00	0.12	0.00	0.09
4	0.15	0.04	0.00	0.12	0.00	0.09
2	0.15	0.04	0.02	0.12	0.00	0.09
0	0.15	0.04	0.03	0.12	0.01	0.09
-2	0.15	0.04	0.02	0.12	0.01	0.07
-4	0.15	0.04	0.02	0.13	0.01	0.07
-6	0.15	0.04	0.01	0.13	0.01	0.07
-8	0.15	0.04	0.01	0.13	0.01	0.07
-10	0.15	0.04	0.01	0.13	0.01	0.07

Table S-1.2: Variation in extent of neutralization with FPeak rotation for the predominantly inorganic factors from the $\text{PMF}_{\text{Full MS}}$ six factor solution.

FPeak	Sulphate-OA	Nitrate-OA	Chloride
10	1.04	1.05	1.13
8	1.04	1.05	1.13
6	1.04	1.05	1.12
4	1.04	1.05	1.11
2	1.04	1.05	1.10
0	0.99	1.04	1.09
-2	1.05	1.06	1.11
-4	1.05	1.06	1.13
-6	1.05	1.06	1.14
-8	1.04	1.06	1.14
-10	1.04	1.06	1.14

Table S-1.3: Extent of neutralization for the five to eight factor PMF solutions for factors dominated by salts, where the extent of neutralization defined as the molar equivalent ratio of cations to anions (i.e., values greater than one indicate a greater concentration of cations to anions).

Factor	Number of Factors			
	5	6	7	8
Nitrate-OA	1.03	1.04	1.02	1.01
Chloride	1.07	1.09	1.26	3.05
Sulphate-OA (Regional Sulphate for 8 factor solution)	0.99	0.99	0.86	1.27
Local Sulphate	n/a	n/a	n/a	0.25

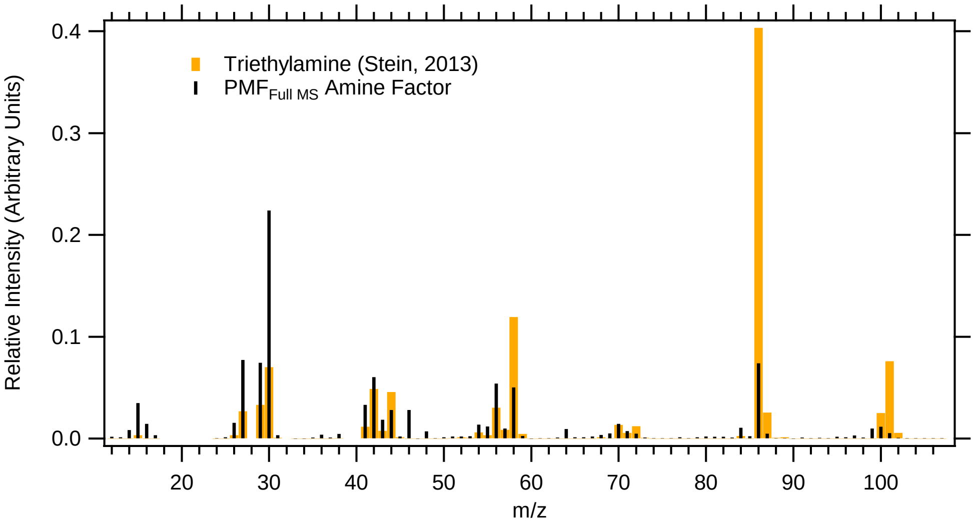


Figure S-1.1: Comparison of the Amine factor from the PMF$_{Full\ MS}$ analysis six factor solution with triethylamine reference electron impact ionization reference spectrum from the NIST database (Stein, 2013).

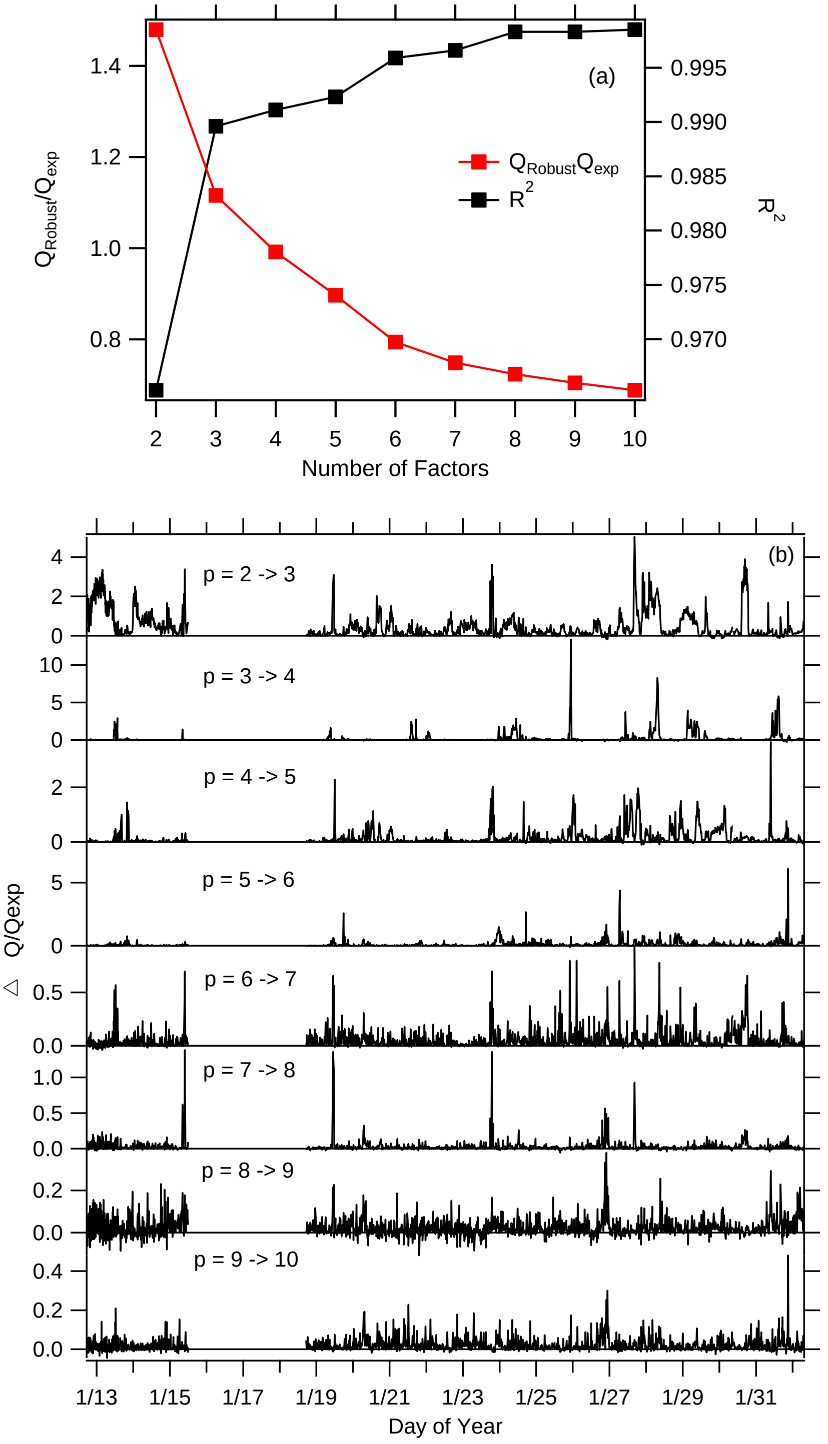


Figure S-1.2: Decrease in the overall Q/Qexp (a), and the Q/Qexp as a function of time (b) as the number of factors (p) is increased in the PMF solution for the PMF$_{\text{Full MS}}$ analysis.

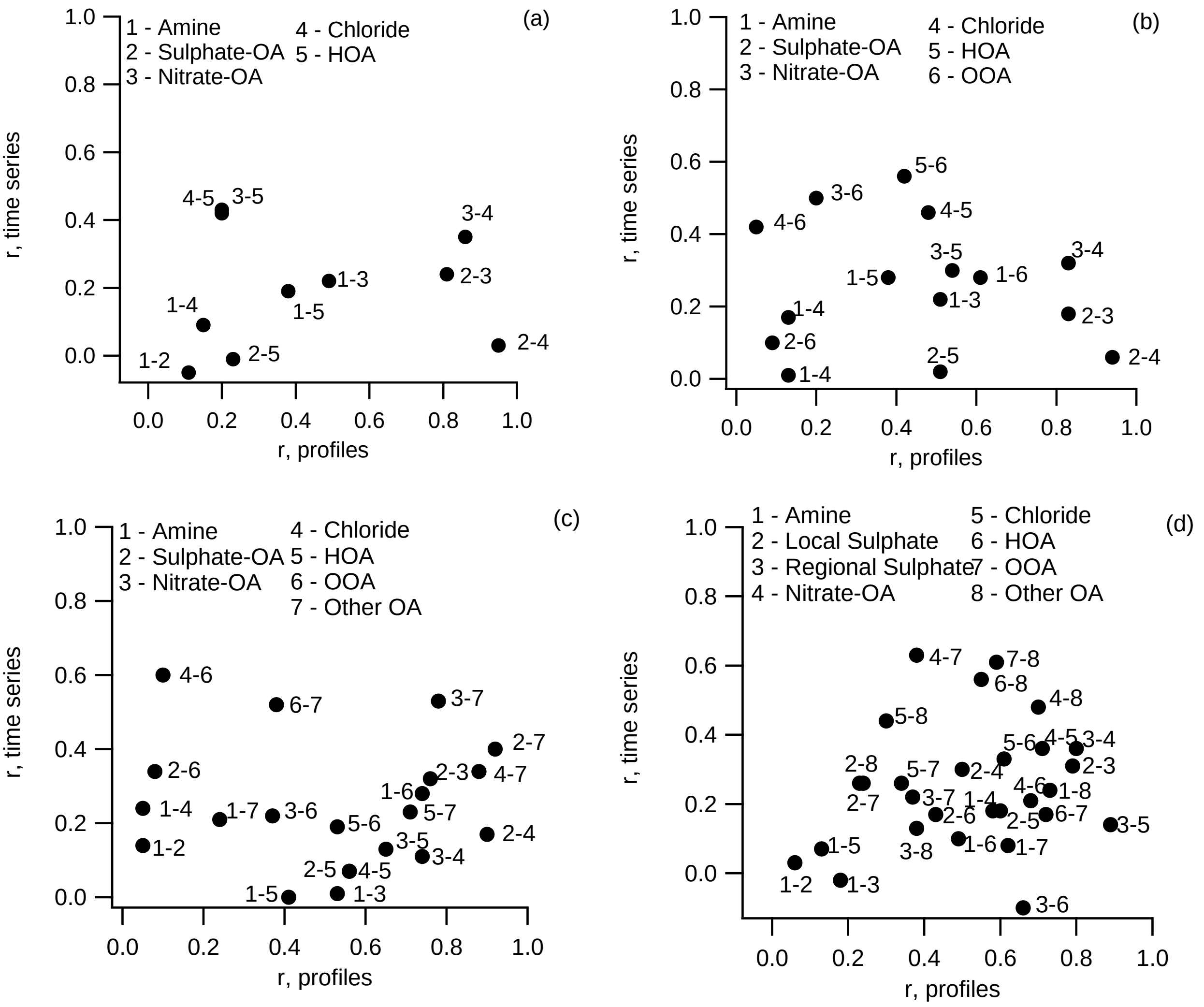


Figure S-1.3: Pearson R correlation coefficients between factor mass spectra and time series for the five (a), six (b), seven (c), and eight (d) factor $\mathrm{PMF_{Full\ MS}}$ solutions.

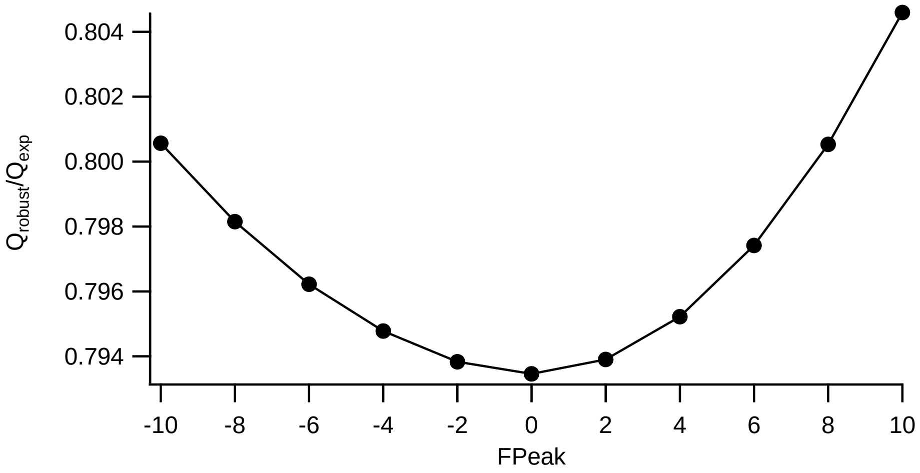
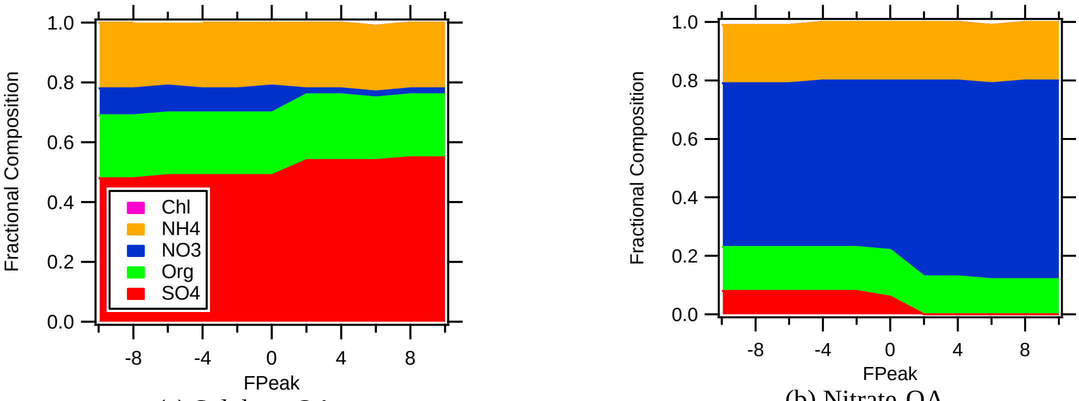
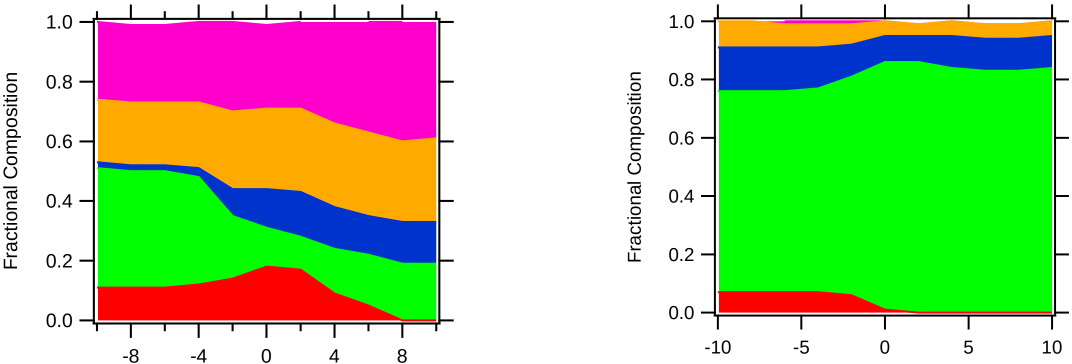


Figure S-1.4: Variation in Q_{robust}/Q_{exp} for the range of tested FPeak values for the $PMF_{Full\ MS}$ six factor solution.



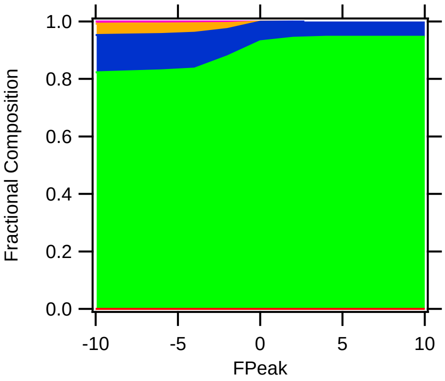
(a) Sulphate-OA

(b) Nitrate-OA



(c) Chloride

(d) HOA



(e) OOA

Figure S-1.5: Variation in fractional factor composition for the six factor solution for the $PMF_{Full\ MS}$ analysis by FPeak value.

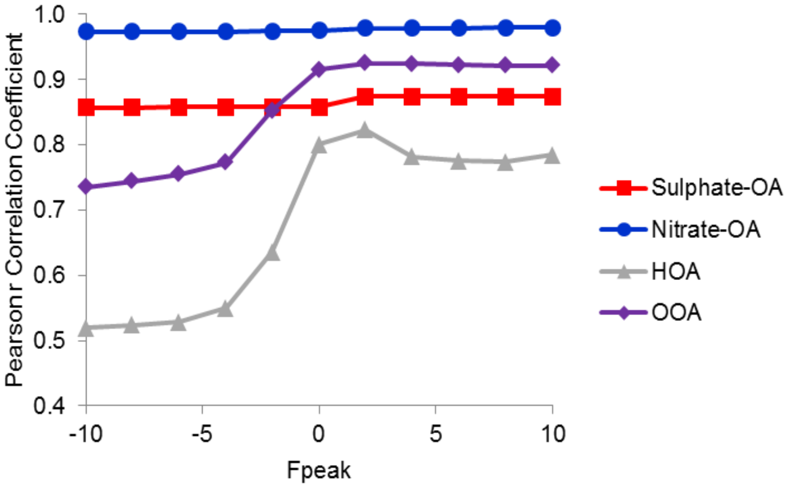


Figure S-1.6: Pearson r correlation coefficients for factor mass spectra from the six factor solution of the PMF$_{\text{Full MS}}$ analysis between the various FPeak values (FPeak = 0 refers to the base case), and selected reference mass spectra from the AMS mass spectral library.

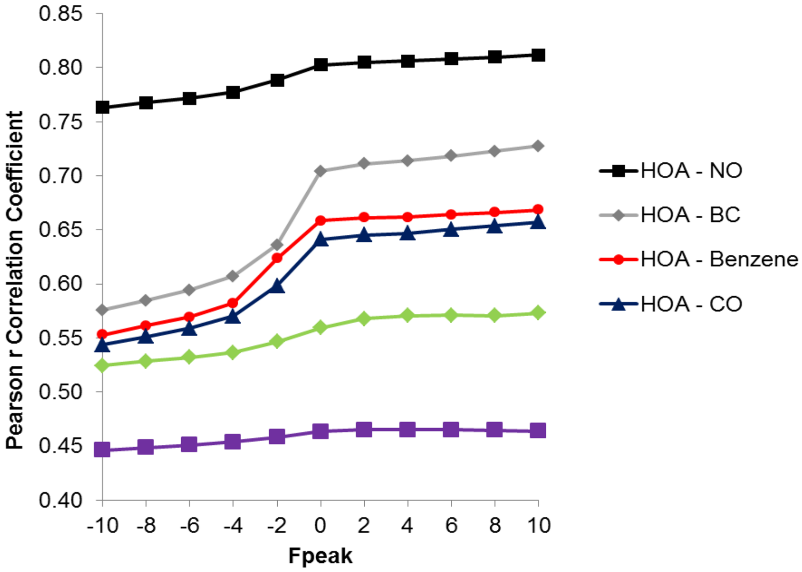


Figure S-1.7: Pearson r correlation coefficients for factor time series from the six factor solution of the PMF$_{\text{Full MS}}$ analysis between the various FPeak values (FPeak = 0 refers to the base case), and selected external time series (PNC refers to CPC-measured particle number concentration).

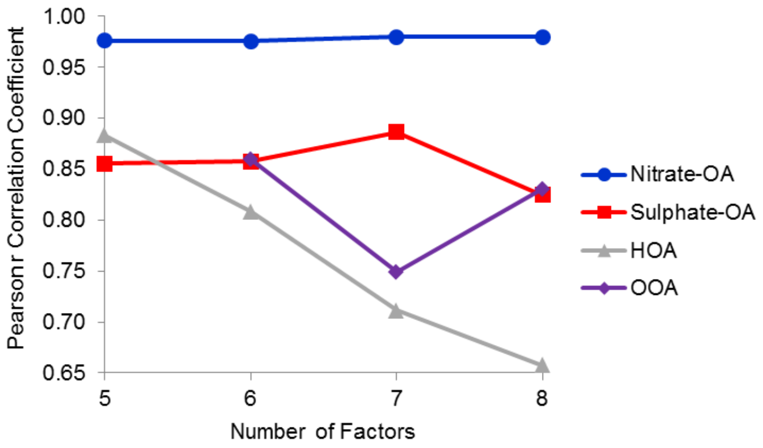


Figure S-1.8: Correlation coefficients for factor mass spectra between the five and eight factor PMF solutions and key reference mass spectra from the AMS mass spectral library.

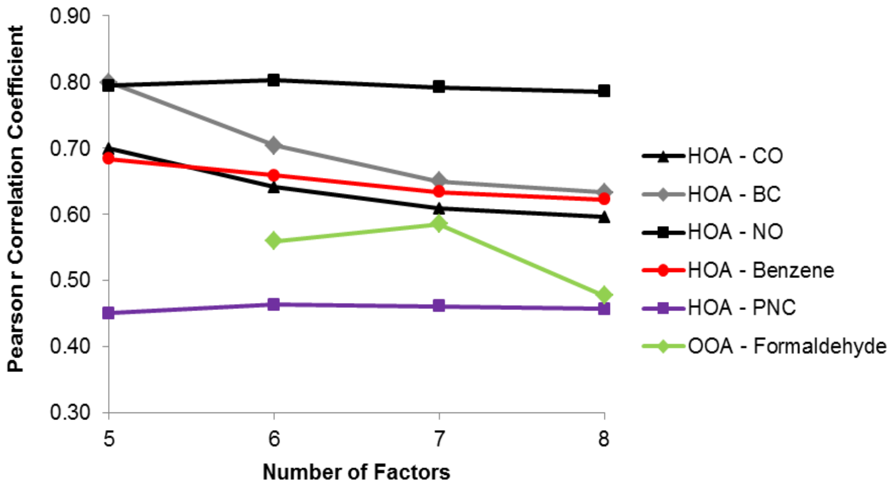


Figure S-1.9: Correlation coefficients for factor time series between the five and eight factor PMF solutions and key external time series.

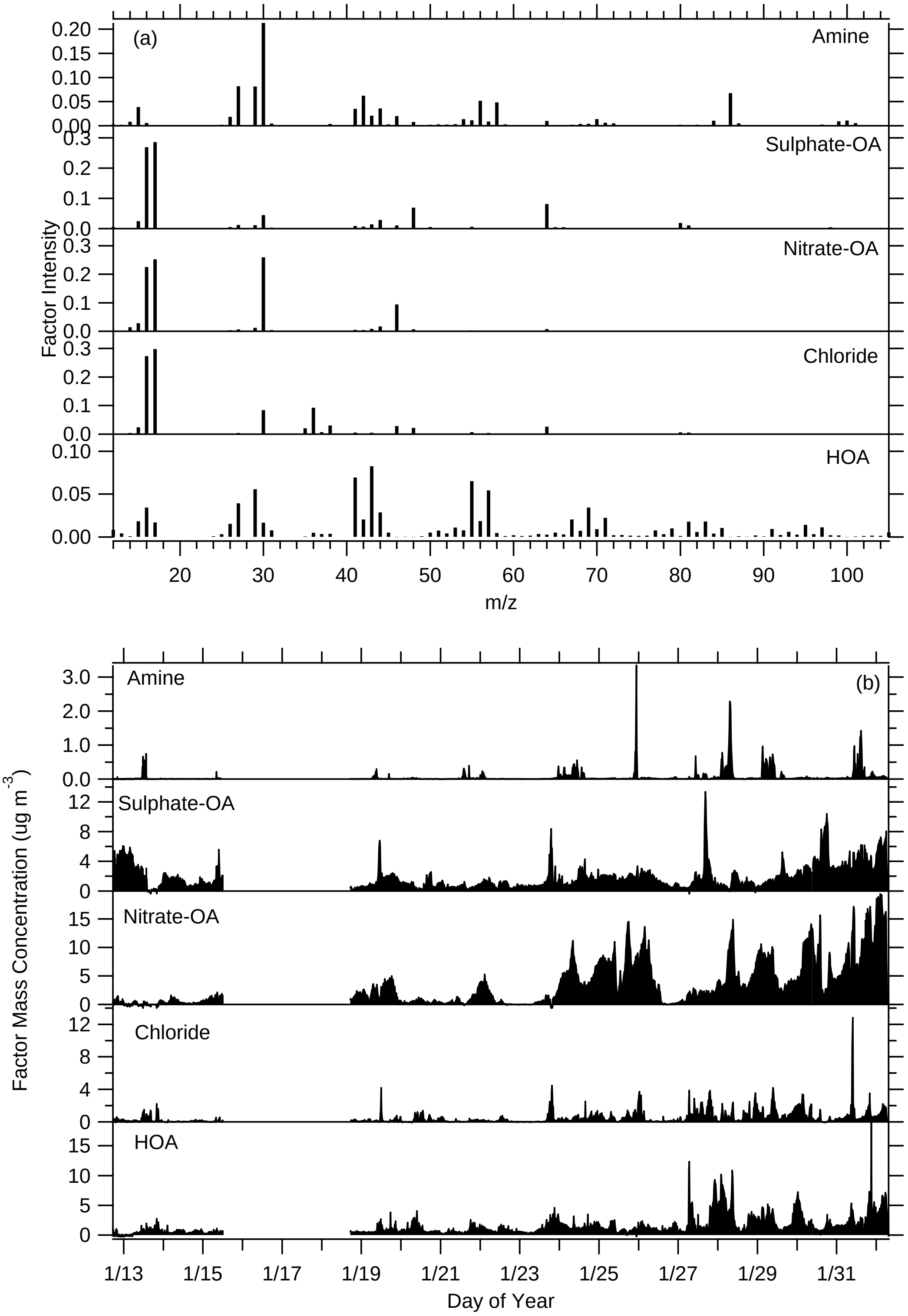
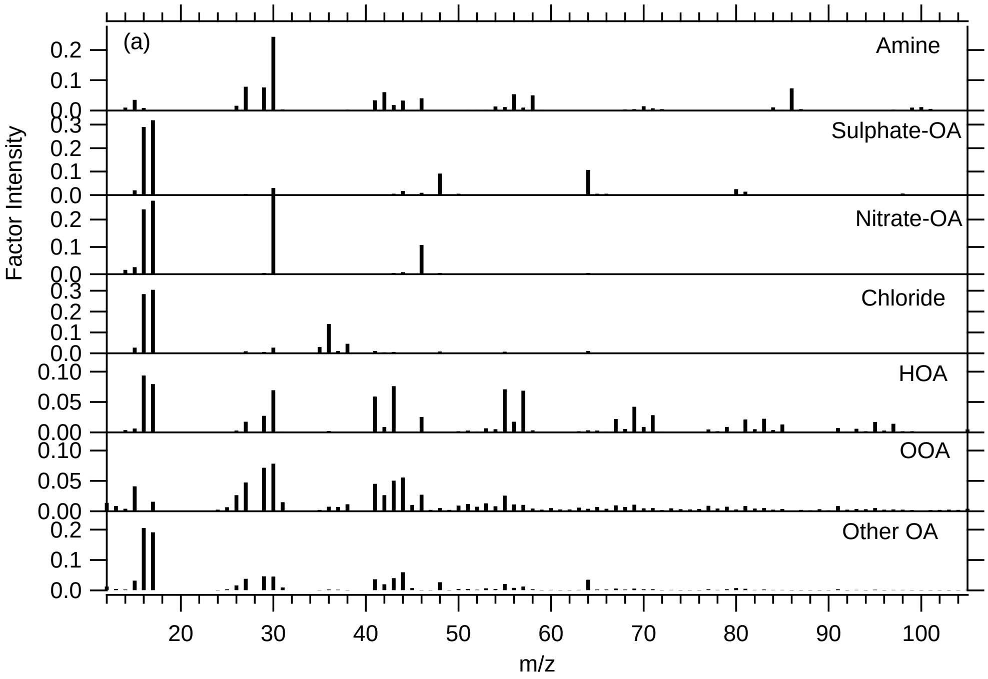


Figure S-1.10: Mass spectra (a) and time series (b) from the five factor $PMF_{Full\ MS}$ analysis solution.



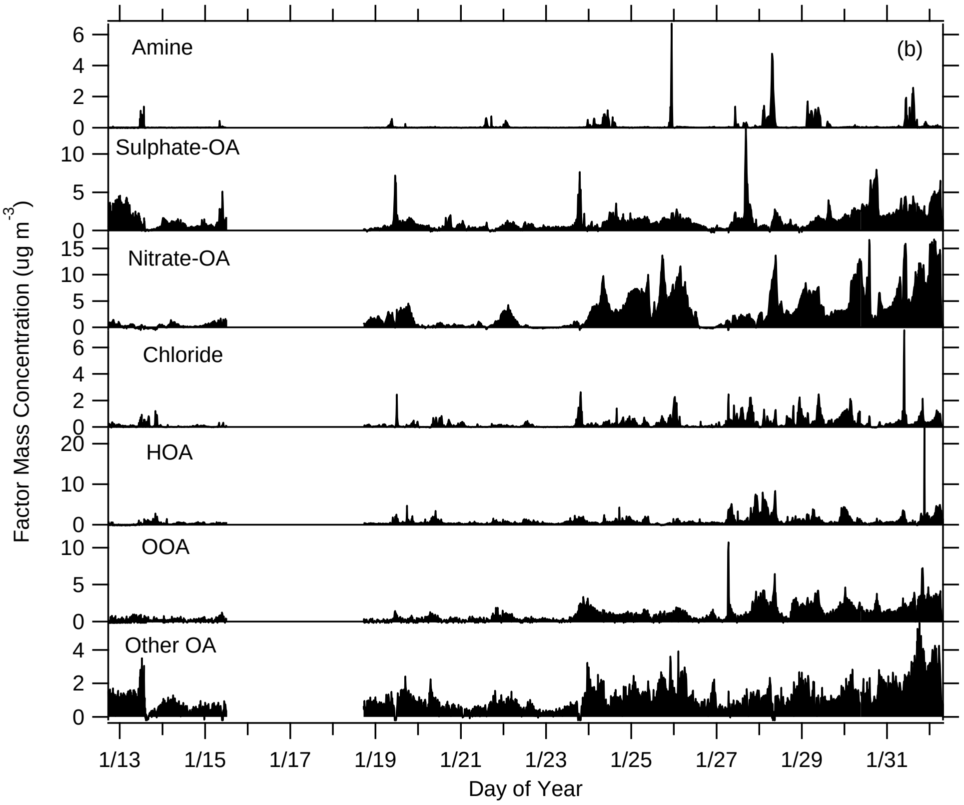


Figure S-1.11: Mass spectra (a) and time series (b) from the seven factor $PMF_{Full\ MS}$ analysis solution.

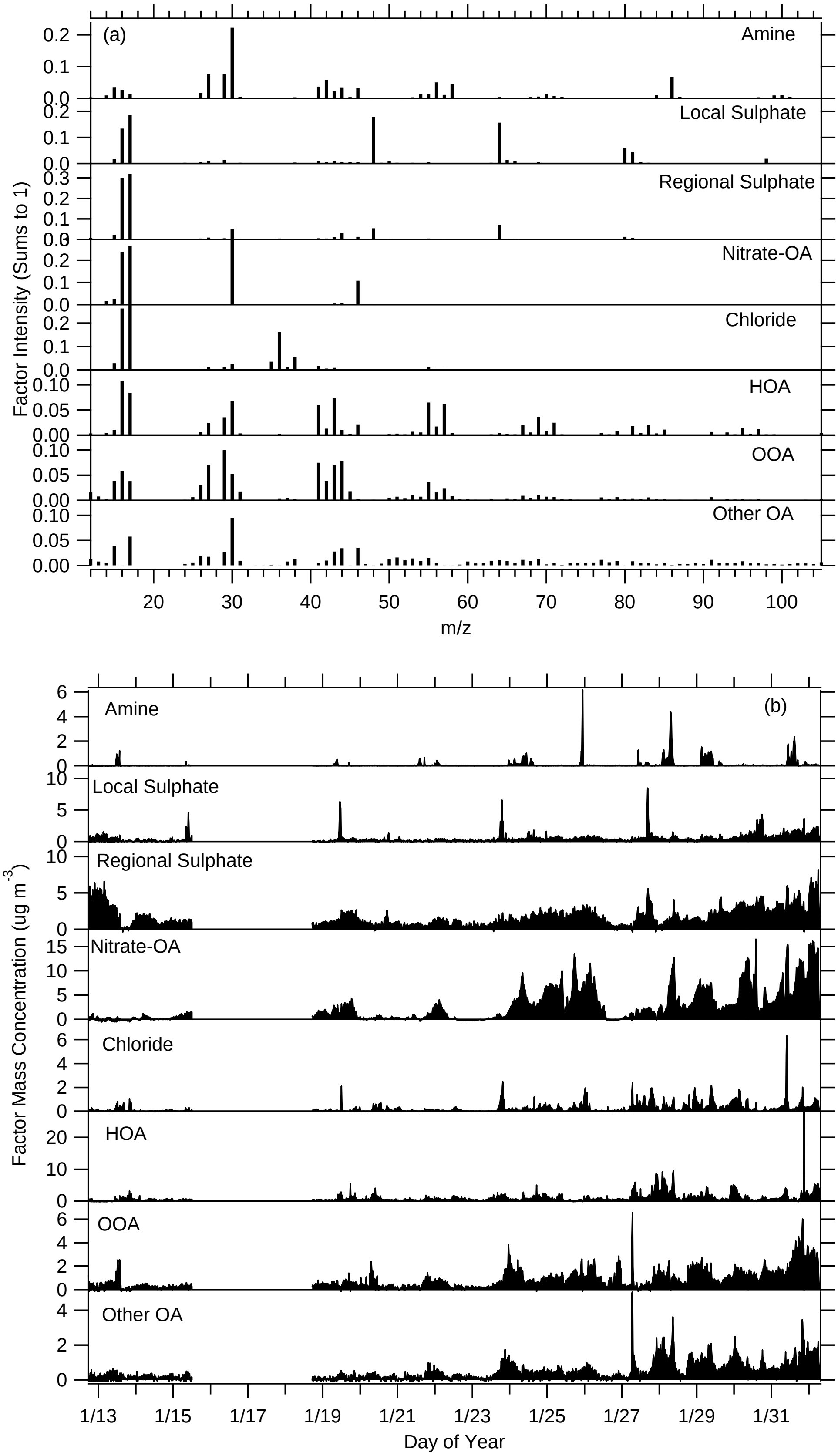


Figure S-1.12: Mass spectra (a), and time series (b) from the eight factor PMF$_{\text{Full MS}}$ analysis solution.

S2 $PMF_{Org\ MS}$ analysis solution justification

The three factor solution was deemed most acceptable for the $PMF_{Org\ MS}$ analysis. A global minimum Q_{robust} value of 221,281 was found for this solution. Figure S-2.1a shows the decrease of the overall Q/Q_{exp} value with increasing factors. It can be seen that the gains in explained variance begin levelling off at three factors. This is also apparent in Figure S-2.1b where the magnitude of the difference between residuals as a function of time drops dramatically after the addition of the fourth factor. These diagnostics show that at least three factors are required to capture a significant amount of variance in the data. Similar to the six factor solution from the $PMF_{Full\ MS}$ analysis, the three factor solution from the $PMF_{Org\ MS}$ analysis resulted in a very good degree of reconstruction, with an r^2 of 0.994. A Q_{robust}/Q_{exp} = 0.78 was calculated, again very similar to the value calculated for the $PMF_{Full\ MS}$ analysis, and indicating that the uncertainties were appropriate for the data.

The rotational robustness of the three factor solution was examined through variation of the FPeak. Figure S-2.2 shows how Q_{robust}/Q_{exp} varies from an FPeak of -10 to 10. The central rotation provides the lowest value, indicating the best fit for the data. Furthermore, despite the wide range of tested FPeaks, the range of variation in Q_{robust}/Q_{exp} is reasonably low (0.776-0.780). The variation in the fraction of *m/z* 44 to the factor mass spectrum (F44) was also examined (Table S-2.1). Variation was low for both the HOA and Amine factors, where values were low. For the OOA factor, it varied more significantly from the central rotation value of 0.15, ranging from 0.12 to 0.16. These values still fell within the range of values reported for OOA factors from around the world (~0.10-0.20) (Ng et al., 2011). Furthermore, factor profile correlations across FPeaks with the reference HOA and OOA mass spectra showed negligible variation in correlation coefficients across rotations, indicating the central rotation was appropriate. Correlations between external time series and factor time series across FPeaks showed slightly more variation, but any improvement in correlation was still insufficient to justify choosing another solution over of the central rotation (Figure S-2.3).

The degree of correlation between factor mass spectra and time series was also examined (Figure S-2.4), as examining incremental changes in factor similarity with the addition of new factors can provide indication of where a solution may no longer be deemed acceptable. Both the three and four factor solutions showed reasonable separation in terms of factor time series, as correlation coefficients were all lower than 0.6. However, in the five factor solution and higher, several factors showed a higher degree of correlation in terms of their mass spectra.

Examination of the mass spectra from the three to six factor solutions shows that the three factor solution provides the most physically meaningful results. At four factors (Figure S-2.5), an Other OA 1 factor is produced, which shows a different mass spectrum and time series than the OOA factor. However, the mass spectrum for this factor cannot be physically justified, as it does not compare sufficiently well with any other known mass spectra, for instance any known HOA or oxidized organic aerosol type spectra. Instead, it displays very high peaks at *m/z* 15 and 29, with *m/z* 43 being the next highest fragment. Some minor features at higher *m/z* appeared in this factor at *m/z* 91 and 115, suggestive of biomass burning emissions. While the correlation with the average BBOA mass spectrum from the AMS mass spectral library shows an r^2 of 0.69 (Table S-2.2), it did not show distinctive contributions from key levoglucosan marker species. Instead, contributions from *m/z* 60 and 73 appeared on par with species of similar mass, resulting in no discernable mass spectral signature. Ultimately, this factor could not be accepted as it was not sufficiently similar to any known source profiles, and the solution was not significantly improved from the previous one by its inclusion. This unknown factor also appeared in the five factor solution, as the Other OA 2 factor (Figure S-2.6). In this solution, it shows a higher r^2 value with BBOA (0.79), however, the time series for the Other OA 1 and Other OA 2 factors were very highly correlated (r = 0.8) (as shown in Figure S-2.4c), and appear to have resulted from factor splitting. This effect continues into the six factor solution, as the Other OA 2 and Other OA 3 factors appear highly similar in terms of their time series (Figure S-2.7). The six factor solution produces an Other OA 1 factor, which displays a very high signal at *m/z* 29, and moderate signal at *m/z* 41 and 43, and low signal elsewhere. A comparison between this factor's MS and others from the AMS mass spectral database shows no significant matches with any known mass spectra, further preventing justification of the six factor solution.

The time series of the four factor OOA and Other OA factors were most correlated with the Nitrate-OA and OOA factors respectively from the $PMF_{Full\ MS}$ analysis (r of 0.77 and 0.93) (Table S-2.3). Although the four factor $PMF_{Org\ MS}$ solution alluded to the presence of other types of organics that were somewhat correlated with inorganics, mass spectral analysis of the Other OA factor from the four factor organic solution showed that this factor did not sufficiently resemble any other known mass spectra to accept the solution. Furthermore, the correlation between the $PMF_{Org\ MS}$ four factor OOA factor and the $PMF_{Full\ MS}$ OOA factor (r of 0.66) was not that much different with that of the $PMF_{Full\ MS}$ Nitrate-OA factor.

The time series of the factors from the three to six factor solutions were compared with those from key external tracers (Table S-2.4). The addition of more OA factors in the four and five factor solutions leads to higher correlations between the Other OA 1 and 2 factors in the 4 and 5 factor solutions respectively, with black carbon, NO, CO, and PNC. However, at the same time, the correlation between HOA and these key combustion species also decreases, indicating that the HOA factor is being split to some degree as well. The impact on the HOA factor, along with the difficulty in justifying the MS of the additional OA factors in the higher order solutions led to the three factor solution being chosen.

The mass spectra of the factors from the three to six factor solutions were also compared to the organic fraction mass spectra from the five organic containing factors from the $\mathrm{PMF}_{\mathrm{Full\ MS}}$ analysis to determine whether higher order organic PMF solutions contained the same information as the organic fraction as determined from the $\mathrm{PMF}_{\mathrm{Full\ MS}}$ analysis (Table S-2.5). The assessment criterion for a good match was considered to be where there was a correlation of $r > 0.95$ (i.e., HOA factors). These cross-correlations showed that the higher order solutions did not resolve the organics in the same way as the $\mathrm{PMF}_{\mathrm{Full\ MS}}$ analysis: not every factor from the $\mathrm{PMF}_{\mathrm{Full\ MS}}$ analysis displayed a significant correlation with corresponding factors from the $\mathrm{PMF}_{\mathrm{Org\ MS}}$ analysis. Thus, the $\mathrm{PMF}_{\mathrm{Full\ MS}}$ analysis extracted factors differently, and led to enhanced apportionment of the organic fraction, as higher order $\mathrm{PMF}_{\mathrm{Org\ MS}}$ solutions were less physically meaningful than those extracted from the six factor $\mathrm{PMF}_{\mathrm{Full\ MS}}$ analysis.

In terms of solution stability, all solutions from three to six factors all showed a global minimum Q value from 100 random seed runs. Bootstrapping runs showed that the three, four, and five factor solutions were all robust, as all solutions were fully remapped to the base run; this robustness declined somewhat at the six factor solution, where all factors were fully remapped to the base run, except for one, where 94 out of 100 runs were remapped.

Table S-2.1: Variation in F44 with FPeak rotation for the $\mathrm{PMF}_{\mathrm{Org\ MS}}$ three factor solution.

FPeak	HOA	OOA	Amine
10	0.00	0.16	0.04
8	0.00	0.16	0.04
6	0.00	0.16	0.04
4	0.00	0.16	0.04
2	0.00	0.16	0.04
0	0.02	0.15	0.05
-2	0.02	0.12	0.05
-4	0.02	0.12	0.05
-6	0.02	0.12	0.05
-8	0.02	0.12	0.05
-10	0.02	0.12	0.05

Table S-2.2: Pearson r correlation comparisons between OOA factor mass spectra from the three, four, five, and six factor $\mathrm{PMF}_{\mathrm{Org\ MS}}$ solutions with reference mass spectra from the AMS mass spectral library.

	OOA Average[5]	**LV-OOA Average**[6]	**SV-OOA Average**[7]	**BBOA Average**[8]
3F OOA	0.93	0.94	0.65	0.60
4F OOA	0.93	0.91	0.72	0.65
4F Other OA 1	0.39	0.41	0.61	0.69
5F OOA	0.94	0.93	0.66	0.57
5F Other OA 1	0.54	0.47	0.49	0.35
5F Other OA 2	0.37	0.38	0.63	0.79
6F OOA	0.89	0.87	0.47	0.31
6F Other OA 1	0.23	0.20	0.43	0.52
6F Other OA 2	0.36	0.38	0.54	0.60
6F Other OA 3	0.50	0.49	0.62	0.77

---

[5] AMS Mass Spectral Database: A_DEC_Q_016
[6] AMS Mass Spectral Database: A_DEC_Q_017
[7] AMS Mass Spectral Database: A_DEC_Q_018
[8] AMS Mass Spectral Database: A_DEC_Q_019

Table S-2.3: Pearson r correlation comparisons between factor time series from the three, four, five, and six factor PMF$_{\text{Org MS}}$ solutions with factor time series from the PMF$_{\text{Full MS}}$ 6 factor solution.

		PMF$_{\text{Full MS}}$ factors					
PMF$_{\text{Org MS}}$ Factors		Amine	Sulphate-OA	Nitrate-OA	Chloride	HOA	OOA
3F	Amine	1.00	-0.02	0.27	0.12	0.22	0.29
	HOA	0.21	0.07	0.46	0.49	0.99	0.67
	OOA	0.14	0.43	0.79	0.41	0.36	0.85
4F	Amine	1.00	-0.02	0.25	0.10	0.18	0.26
	HOA	0.18	-0.01	0.33	0.42	1.00	0.51
	OOA	0.10	0.50	0.77	0.29	0.15	0.66
	Other OA 1	0.31	0.18	0.62	0.51	0.66	0.93
5F	Amine	1.00	-0.03	0.24	0.09	0.17	0.24
	HOA	0.16	-0.01	0.30	0.39	0.99	0.44
	OOA	0.06	0.54	0.75	0.28	0.10	0.59
	Other OA 1	0.28	0.15	0.58	0.53	0.73	0.90
	Other OA 2	0.47	0.18	0.68	0.42	0.59	0.89
6F	Amine	1.00	-0.03	0.24	0.09	0.17	0.24
	HOA	0.17	-0.01	0.31	0.41	0.99	0.47
	OOA	0.05	0.55	0.74	0.31	0.15	0.61
	Other OA 1	0.27	0.15	0.58	0.53	0.71	0.91
	Other OA 2	0.49	0.21	0.67	0.39	0.58	0.82
	Other OA3	0.16	0.33	0.68	0.33	0.33	0.71

Table S-2.4: Pearson r correlation coefficients between the time series of key external tracers, and factor time series from the three, four, five and six factor $\mathrm{PMF}_{\mathrm{Org\ MS}}$ solutions.

	External TS	**3F**	**4F**	**5F**	**6F**
HOA	BC	0.78	0.68	0.64	0.66
	NO	0.83	0.80	0.79	0.79
	CO	0.70	0.63	0.60	0.62
	PNC	0.45	0.46	0.45	0.45
OOA	BC	0.64	0.50	0.47	0.53
	NO	0.41	0.26	0.23	0.28
	CO	0.51	0.37	0.34	0.40
	PNC	0.12	0.02	-0.01	-0.03
	Formaldehyde	0.58	0.46	0.45	0.48
Other OA 1	BC		0.78	0.79	0.78
	NO		0.59	0.64	0.63
	CO		0.65	0.67	0.67
	PNC		0.28	0.31	0.32
	Formaldehyde		0.59	0.59	0.59
Other OA 2	BC			0.71	0.71
	NO			0.56	0.56
	CO			0.57	0.56
	PNC			0.30	0.25
	Formaldehyde			0.51	0.49
Other OA 3	BC				0.51
	NO				0.37
	CO				0.40
	PNC				0.22
	Formaldehyde				0.45

Table S-2.5: Pearson r correlation comparisons between factor mass spectra from the three, four, five, and six factor PMF$_{\text{Org MS}}$ solutions with the organic fraction of factor mass spectra from the PMF$_{\text{Full MS}}$ 6 factor solution.

		PMF$_{\text{Full MS}}$ factors				
PMF$_{\text{Org MS}}$ Factors		**Sulphate-OA**	**Nitrate-OA**	**Chloride**	**HOA**	**OOA**
3F	HOA	0.57	0.64	0.68	0.99	0.55
	Amine	0.56	0.65	0.67	0.37	0.76
	OOA	0.93	0.97	0.69	0.40	0.95
4F	HOA	0.61	0.67	0.69	0.99	0.56
	Other OA 1	0.49	0.72	0.53	0.36	0.86
	Amine	0.58	0.65	0.68	0.38	0.75
	OOA	0.96	0.98	0.74	0.55	0.92
5F	Other OA 1	0.78	0.70	0.78	0.49	0.66
	HOA	0.57	0.63	0.66	0.99	0.52
	OOA	0.98	0.96	0.73	0.51	0.89
	Amine	0.55	0.56	0.66	0.31	0.65
	Other OA 2	0.48	0.73	0.50	0.53	0.83
6F	OOA	0.95	0.88	0.59	0.35	0.75
	Other OA 2	0.41	0.69	0.31	0.36	0.76
	Amine	0.56	0.57	0.68	0.33	0.66
	Other OA 3	0.72	0.81	0.75	0.65	0.86
	HOA	0.56	0.61	0.64	0.99	0.49
	Other OA 1	0.52	0.56	0.85	0.49	0.72

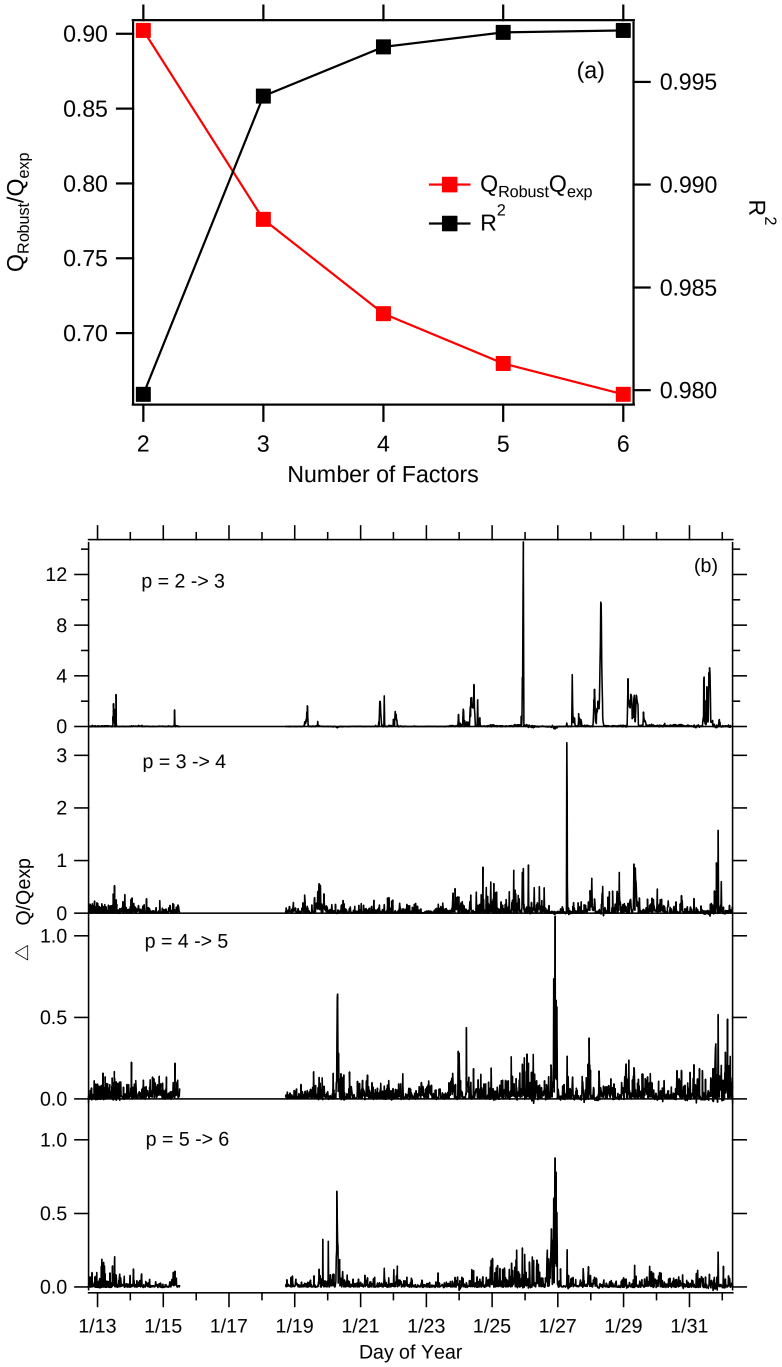


Figure S-2.1: Decrease in the overall Q/Q_{exp} (a), and the Q/Q_{exp} as a function of time (b) as the number of factors (p) is increased in the PMF solution for the $PMF_{Org\ MS}$ analysis.

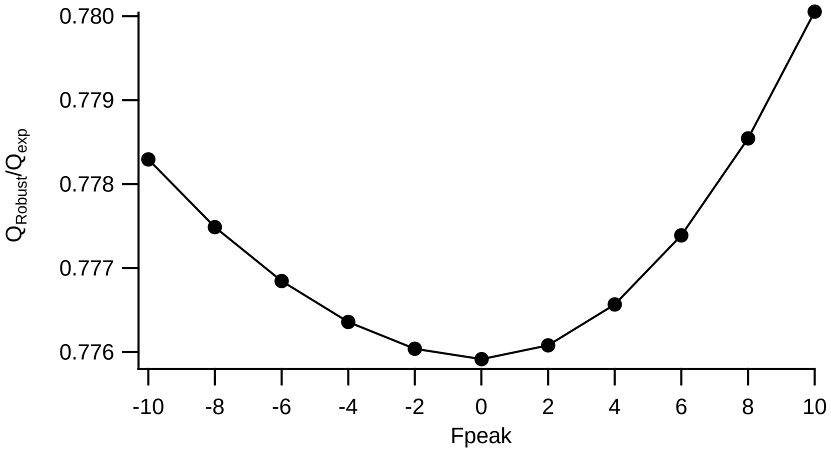


Figure S-2.2: Variation in Q_{robust}/Q_{exp} for the range of tested FPeak values for the $PMF_{Org\ MS}$ three factor solution.

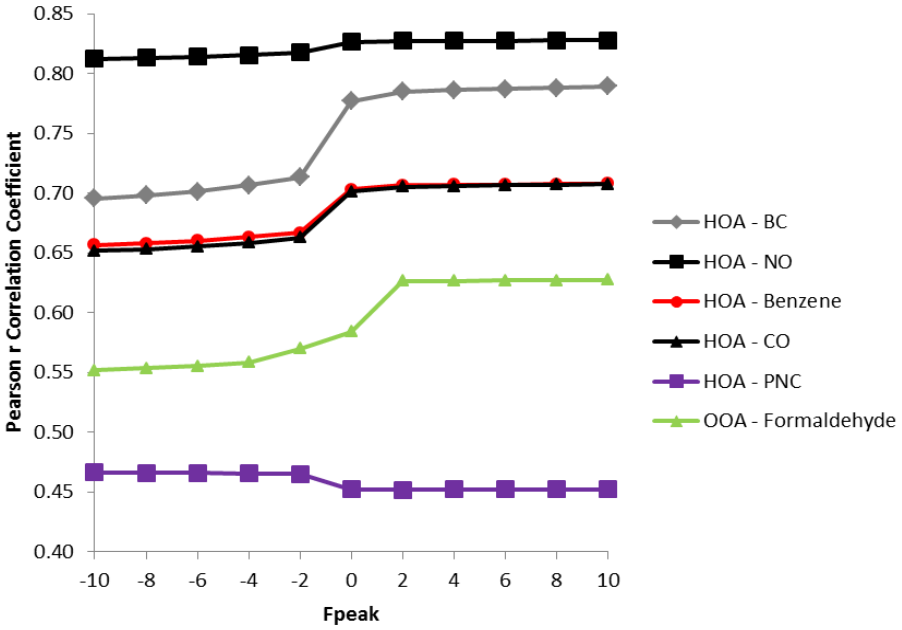


Figure S-2.3: Pearson r correlation coefficients for factor time series from the three factor solution of the PMF$_{\mathrm{Org\,MS}}$ analysis between the various FPeak values (FPeak = 0 refers to the base case), and selected external time series (PNC refers to CPC-measured particle number concentration).

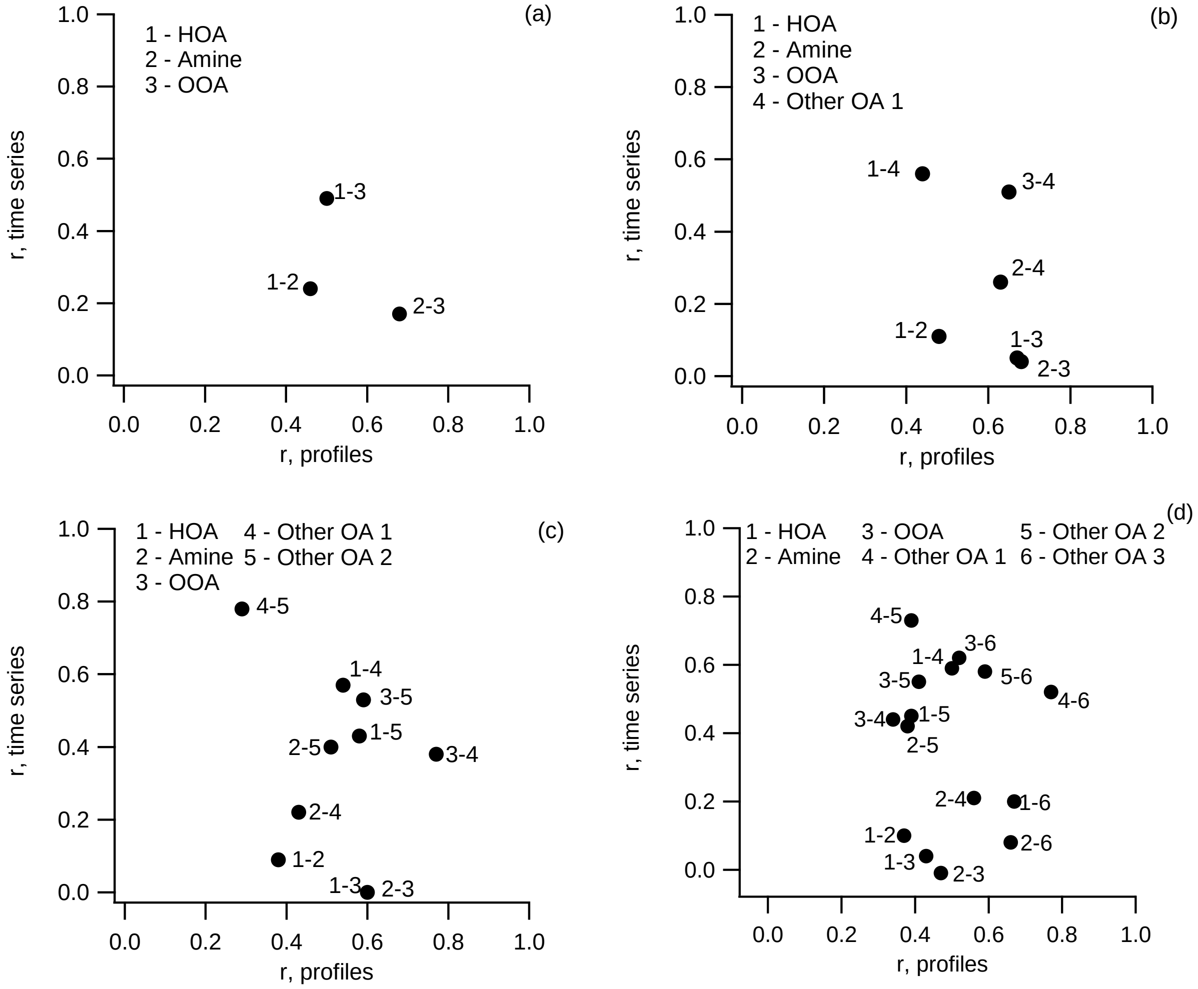


Figure S-2.4: Pearson r correlation coefficients between factor mass spectra and time series for the three (a), four (b), five (c), and six (d) factor $PMF_{Org\ MS}$ solutions.

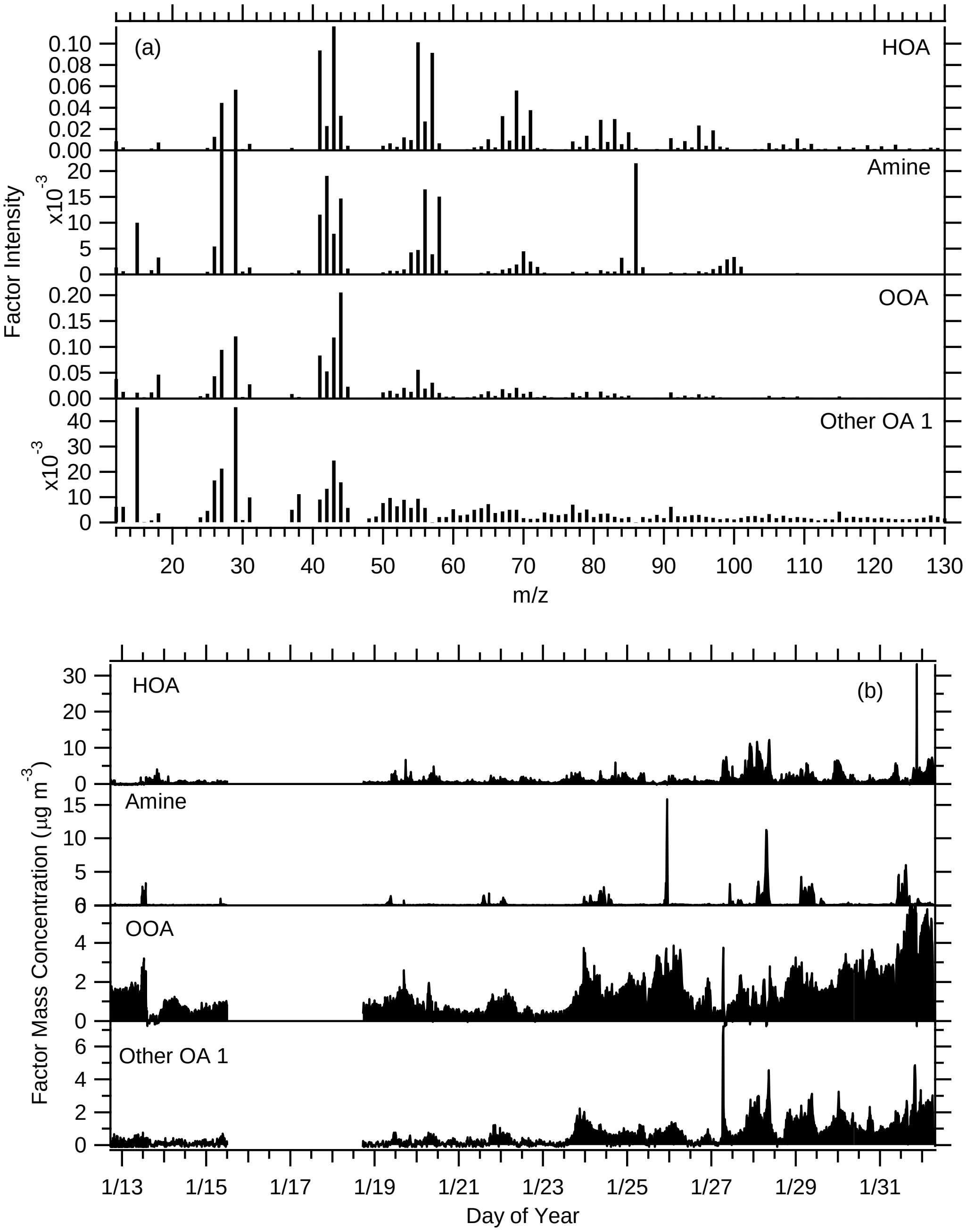


Figure S-2.5: Mass spectra (a), and time series (b) from the four factor PMF$_{\mathrm{Org\ MS}}$ analysis solution.

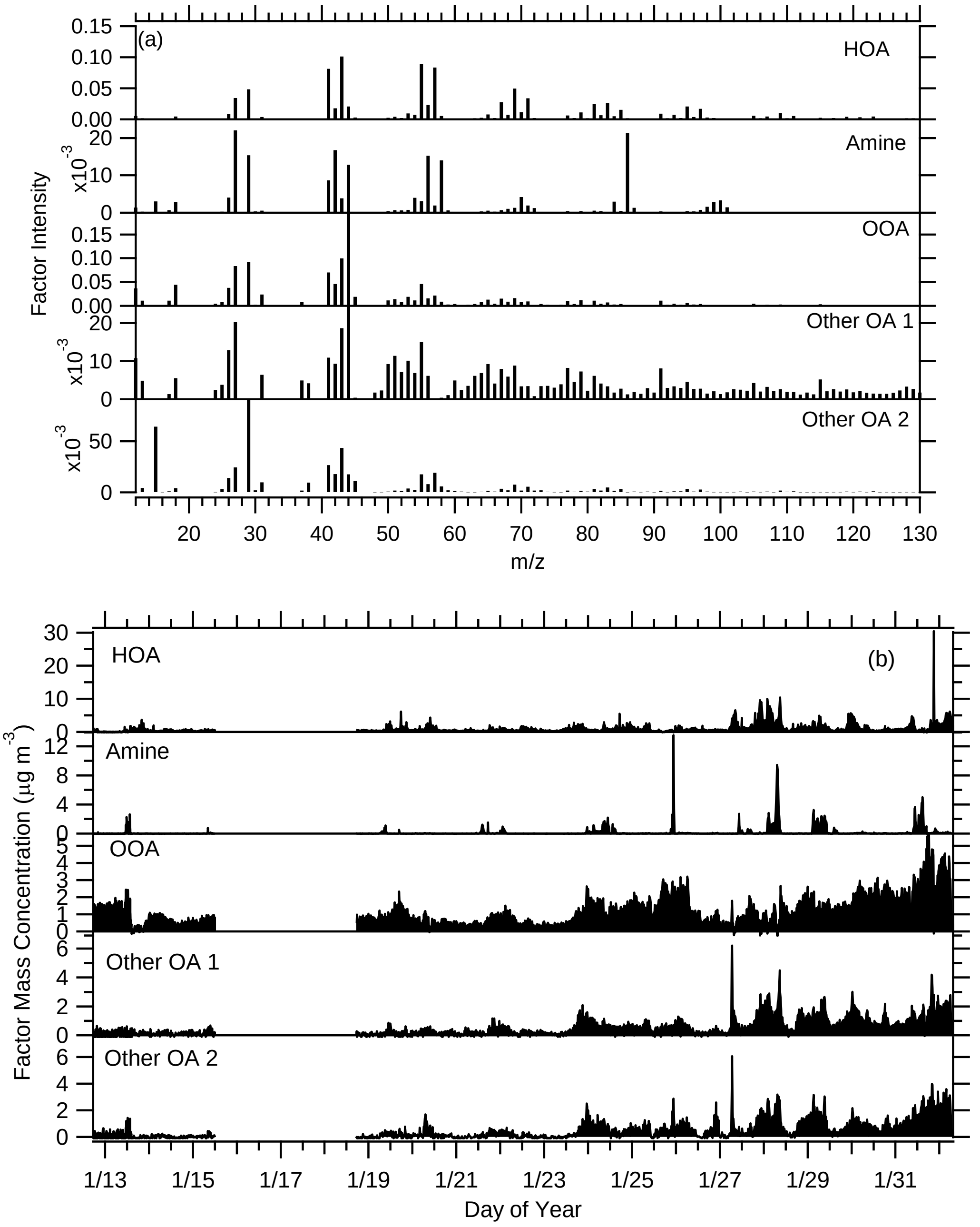


Figure S-2.6: Mass spectra (a), and time series (b) from the five factor PMF$_{\mathrm{Org\ MS}}$ analysis solution.

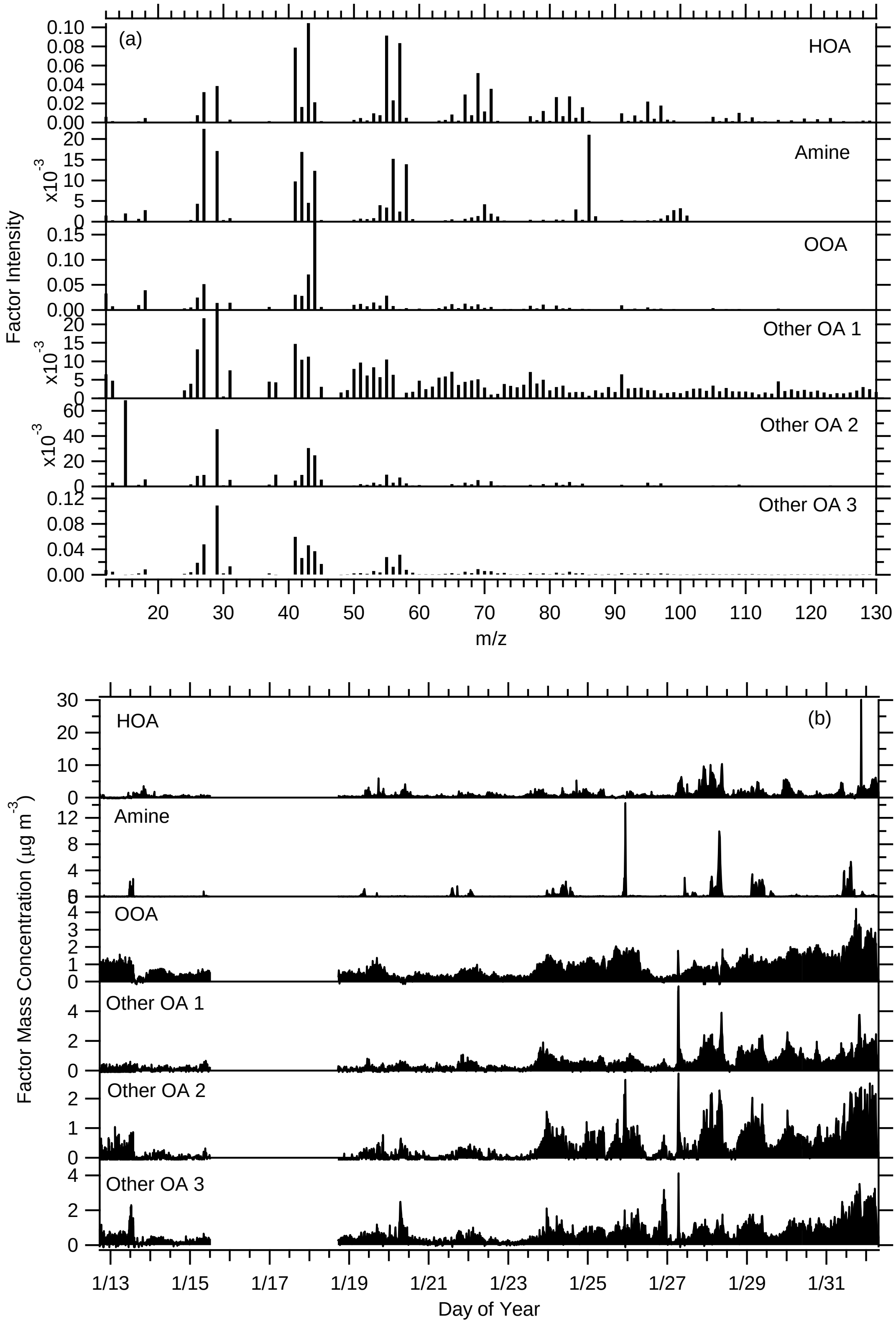


Figure S-2.7: Mass spectra (a), and time series (b) from the six factor $\mathrm{PMF}_{\mathrm{Org\ MS}}$ analysis solution.

S3 Comparison of PMF solutions between the PMF$_{\mathrm{Full\ MS}}$ and PMF$_{\mathrm{Org\ MS}}$ analyses

To understand the effect of introducing the inorganics to organics in PMF analysis of the full mass spectrum, the overall fit and residuals between the PMF$_{\mathrm{Full\ MS}}$ and PMF$_{\mathrm{Org\ MS}}$ analyses were compared. In comparing the overall Q_{robust}/Q_{exp} values for the chosen solutions from the PMF$_{\mathrm{Full\ MS}}$ and PMF$_{\mathrm{Org\ MS}}$ analyses, it can be seen that they are on the whole, very similar (Q_{robust}/Q_{exp} for PMF$_{\mathrm{Full\ MS}}$ is 0.794, while it is 0.776 for PMF$_{\mathrm{Org\ MS}}$). This indicates that the scaled residuals for both solutions are highly similar.

Figure S-3.1 provides a closer examination of this metric with the Q/Q_{exp} for each point in time (a), and across the mass spectrum (b). Figure S-3.1a shows that the temporal distribution of scaled residuals relative to the Q_{exp} is highly similar between the PMF$_{\mathrm{Full\ MS}}$ and PMF$_{\mathrm{Org\ MS}}$ analyses, with the only major instances where the PMF$_{\mathrm{Full\ MS}}$ contributions noticeably exceeded those of the PMF$_{\mathrm{Org\ MS}}$ analysis occurring during the large SO_4^{2-} spikes. As these spikes are mainly attributable to inorganics that were not accounted for in the PMF$_{\mathrm{Org\ MS}}$ analysis, it can be seen that the temporal analysis of scaled residuals to Q_{exp} indicates that the data were modeled to a very similar degree in both analyses. Figure S-3.1b shows the contribution of scaled residuals to Q_{exp} by *m/z*. Only fragments used in both analyses are presented for this comparison. The same effect is observed in this case as in the temporal residuals comparison, as it is mainly the fragments containing inorganics (namely ammonium and sulphate), that display higher residuals. As discussed earlier, these residuals are likely caused by an acidic local sulphate signal not being captured by the PMF model in the six factor solution.

In considering the organic fragments, only *m/z* 44 was modeled somewhat better by the PMF$_{\mathrm{Org\ MS}}$ analysis. A possible explanation for this may be that in the PMF$_{\mathrm{Org\ MS}}$ analysis, *m/z* 44, and ratios between this and other key fragments, is driving factor separation. However, in the PMF$_{\mathrm{Full\ MS}}$ analysis, factor separation is also being driven by the inorganic fragments. With *m/z* 44 co-varying with a series of other inorganic fragments in the PMF$_{\mathrm{Full\ MS}}$ analysis, apportioning this fragment to inorganic factors, and ultimately a greater number of factors, may result in some unapportioned signal. Another notable difference between the two analyses was for *m/z* 15, whereby it showed higher residuals in the PMF$_{\mathrm{Org\ MS}}$ analysis. In examining solutions of higher order, it was noted that while there were modest and more significant improvements to *m/z* 15 residuals in the four and five factor solutions respectively.

While the four and five factor solutions managed to better represent this *m/z*, these solutions were not entirely physically meaningful, and could not be accepted. In considering *m/z* 15 representation in the $\mathrm{PMF}_{\mathrm{Org\ MS}}$ and $\mathrm{PMF}_{\mathrm{Full\ MS}}$ analyses, it is possible that it was modeled better in the $\mathrm{PMF}_{\mathrm{Full\ MS}}$ analysis due to the presence of inorganic signal at *m/z* 15, 16, and 17 from ammonium; in that analysis, the strong co-variance between these fragments may have resulted in a stronger pull towards those fragments. However, higher residuals for *m/z* 15 (as well as a few other *m/z*s such as 55 and 57) as compared to other organic fragments in the $\mathrm{PMF}_{\mathrm{Org\ MS}}$ analysis is most likely attributed the strong signal at this *m/z*, the number of complex sources and processes contributing to it, and overall limitations in the data (unit mass resolution, and 15min time resolution).

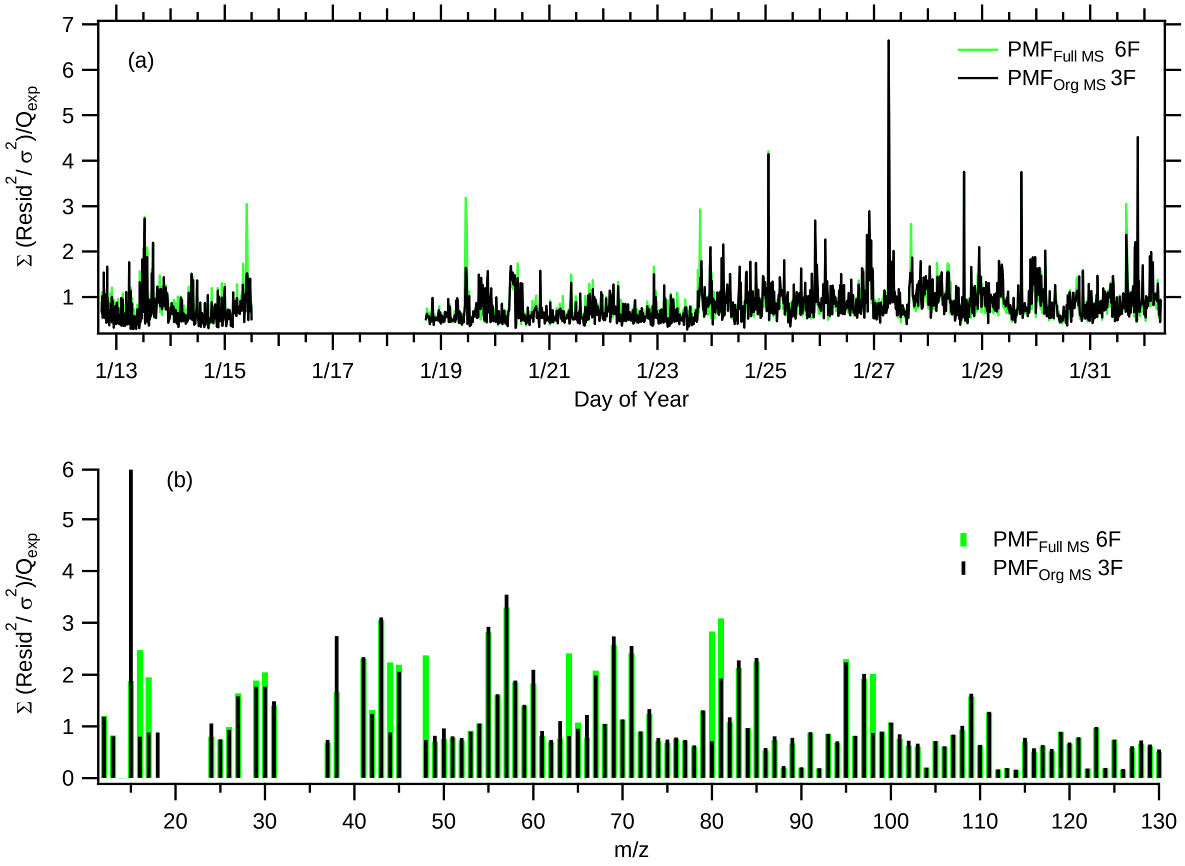


Figure S-3.1: Contribution of scaled residuals to Q_{exp} for both the $PMF_{Full\ MS}$ six factor and the $PMF_{Org\ MS}$ three factor solutions for each point in time (a), and for each *m/z* (b).

References

Ng, N. L., Canagaratna, M. R., Jimenez, J. L., Zhang, Q., Ulbrich, I. M., and Worsnop, D. R.: Real-time methods for estimating organic component mass concentrations from aerosol mass spectrometer data, Environmental Science and Technology, 45, 910-916, 2011.

Stein, S. E.: "Mass Spectra", NIST Chemistry WebBook, NIST Standard Reference Database Number 69, National Institute of Standards and Technology, Gaithersburg MD, 20899, 2013.

Ulbrich, I.M., Lechner, M., and Jimenez, J.L. AMS Spectral Database. URL: http://cires.colorado.edu/jimenez-group/AMSsd/.