

1 Supplementary information for manuscript

2 **Identification and quantification of organic aerosol from**
3 **cooking and other sources in Barcelona using aerosol mass**
4 **spectrometer data**

5
6 **C. Mohr¹, P. F. DeCarlo^{1*}, M. F. Heringa¹, R. Chirico^{1,**}, J. G. Slowik¹, R. Richter¹,**
7 **C. Reche², A. Alastuey², X. Querol², R. Seco³, J. Peñuelas³, J. L. Jiménez^{4,5}, M. Crippa¹, R.**
8 **Zimmermann^{6,7}, U. Baltensperger¹, and A. S. H. Prévôt¹**

9
10 ¹Laboratory of Atmospheric Chemistry, Paul Scherrer Institut (PSI), Villigen, Switzerland

11 ²Institute for Environmental Assessment and Water Research (IDAEA-CSIC), Barcelona, Spain

12 ³Unitat d'Ecologia Global CREA-CEAB-CSIC, Centre de Recerca Ecològica i Aplicacions
13 Forestals, Universitat Autònoma de Barcelona, Barcelona, Spain

14 ⁴Department of Chemistry and Biochemistry, ⁵Cooperative Institute for Research in the
15 Environmental Sciences (CIRES), University of Colorado, Boulder, USA

16 ⁶Helmholtz Zentrum Munchen, Joint Mass Spectrometry Center, Institute of Ecological
17 Chemistry, Neuherberg, Germany

18 ⁷University of Rostock, Rostock, Germany

19 * now at: Department of Civil, Architectural, and Environmental Engineering, Drexel University,
20 Philadelphia, PA, USA

21 ** now at: Italian National Agency for New Technologies, Energy and Sustainable Economic
22 Development (ENEA), UTAPRAD-DIM, Frascati, Italy

23

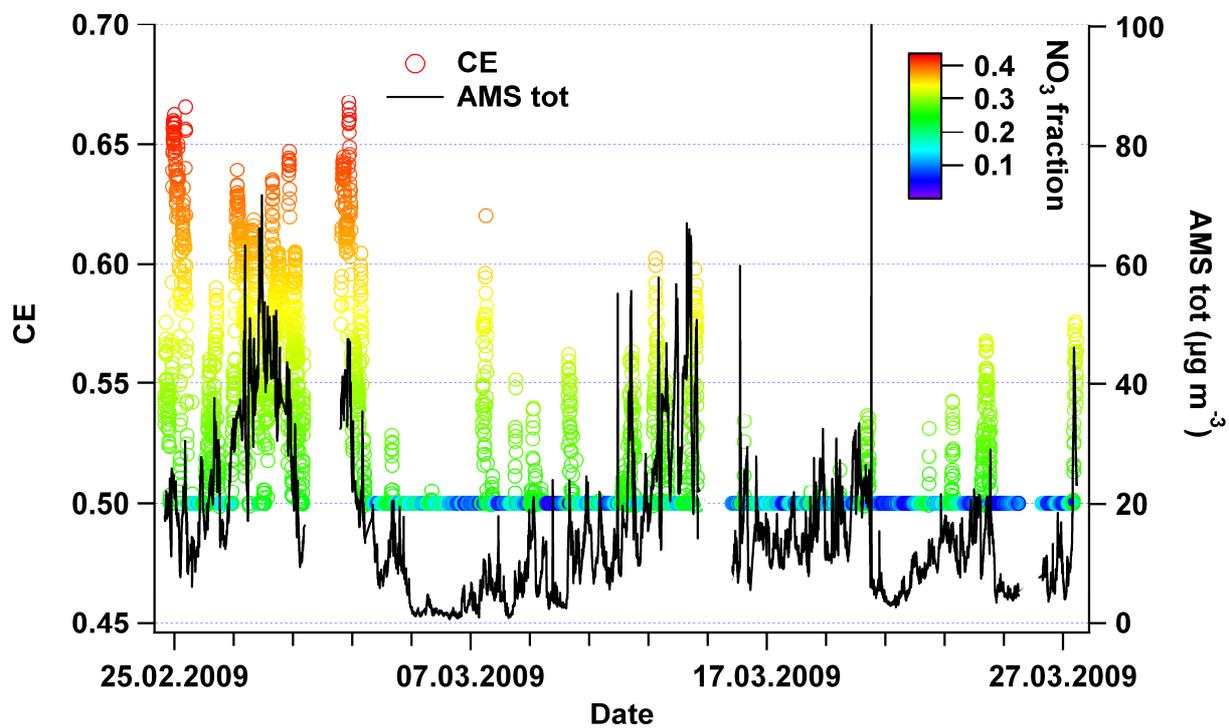
24

25

26 Correspondence to: A. S. H. Prévôt (andre.prevot@psi.ch)

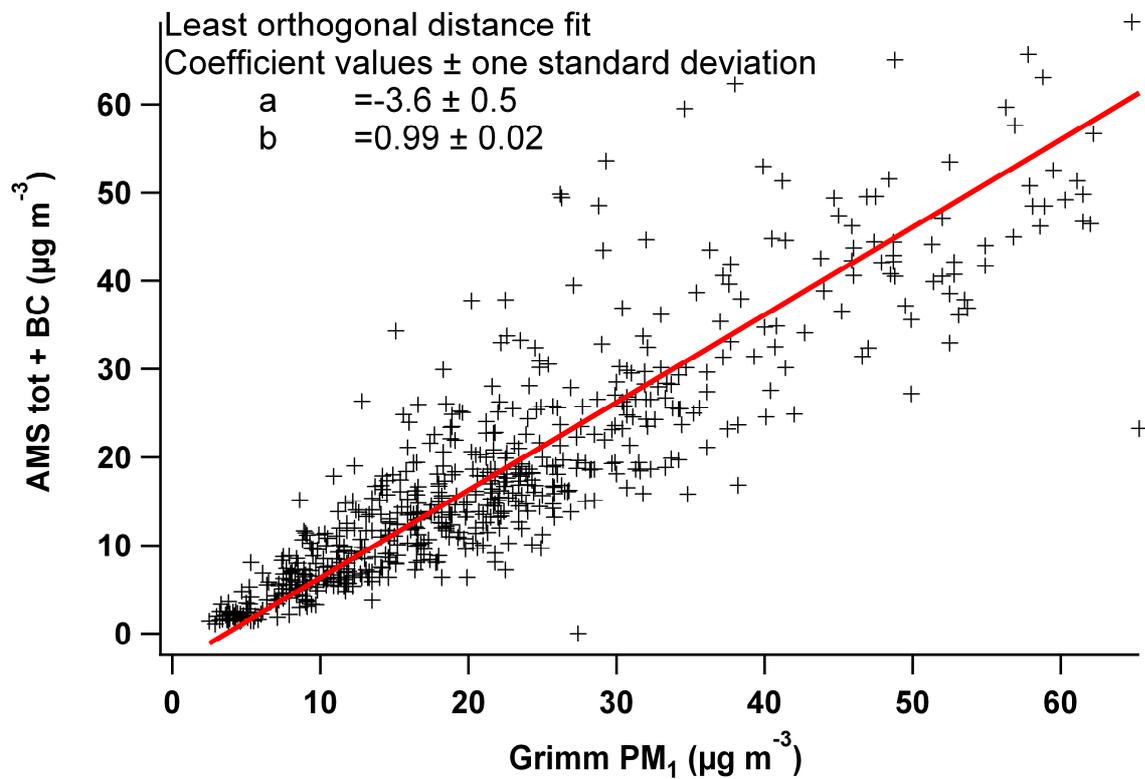
27

28 1 Collection efficiency (CE)



29
30 Figure S1: Time series of the collection efficiency (CE) used for the present dataset (left axis) and total
31 concentration of species measured by AMS (right axis).
32
33
34

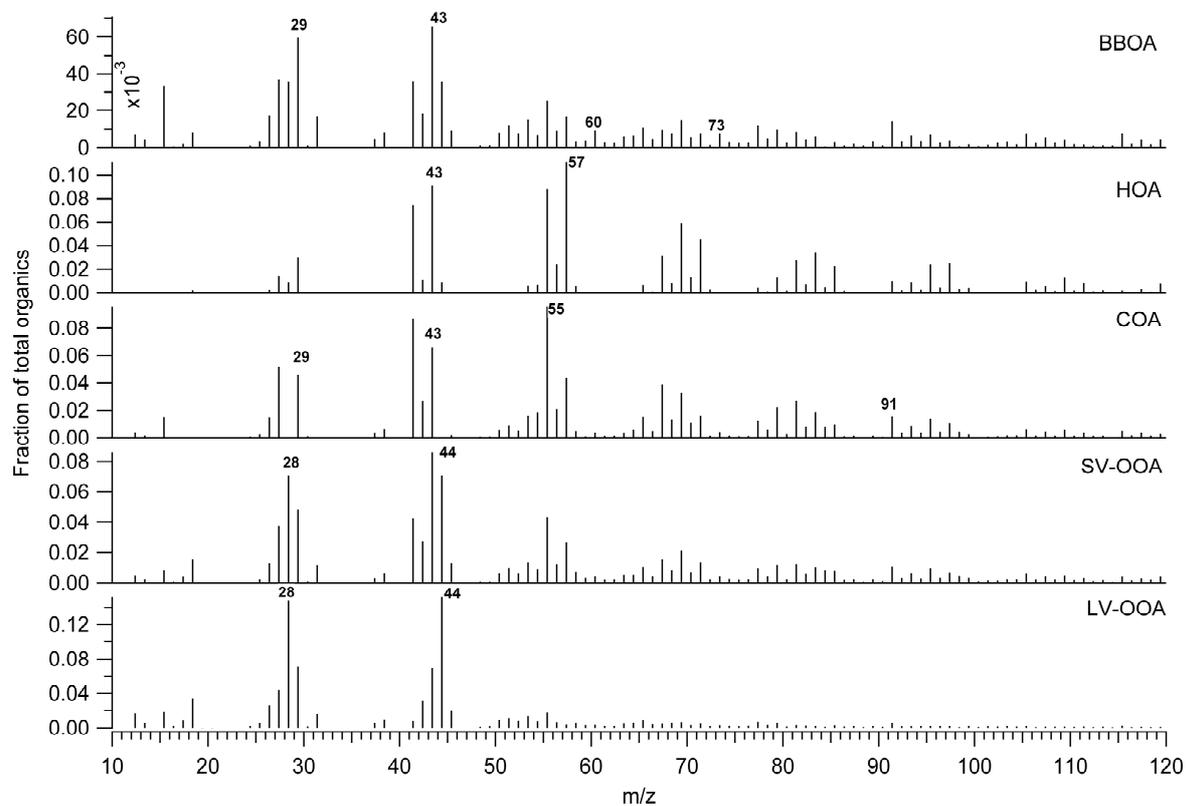
35 **2 PM₁ time series**



36
37 **Figure S2: Scatterplot of combined time series of total AMS species (HR) and BC (y-axis) and Grimm PM₁.**
38 **The data were fitted with a least orthogonal distance fit (red line).**

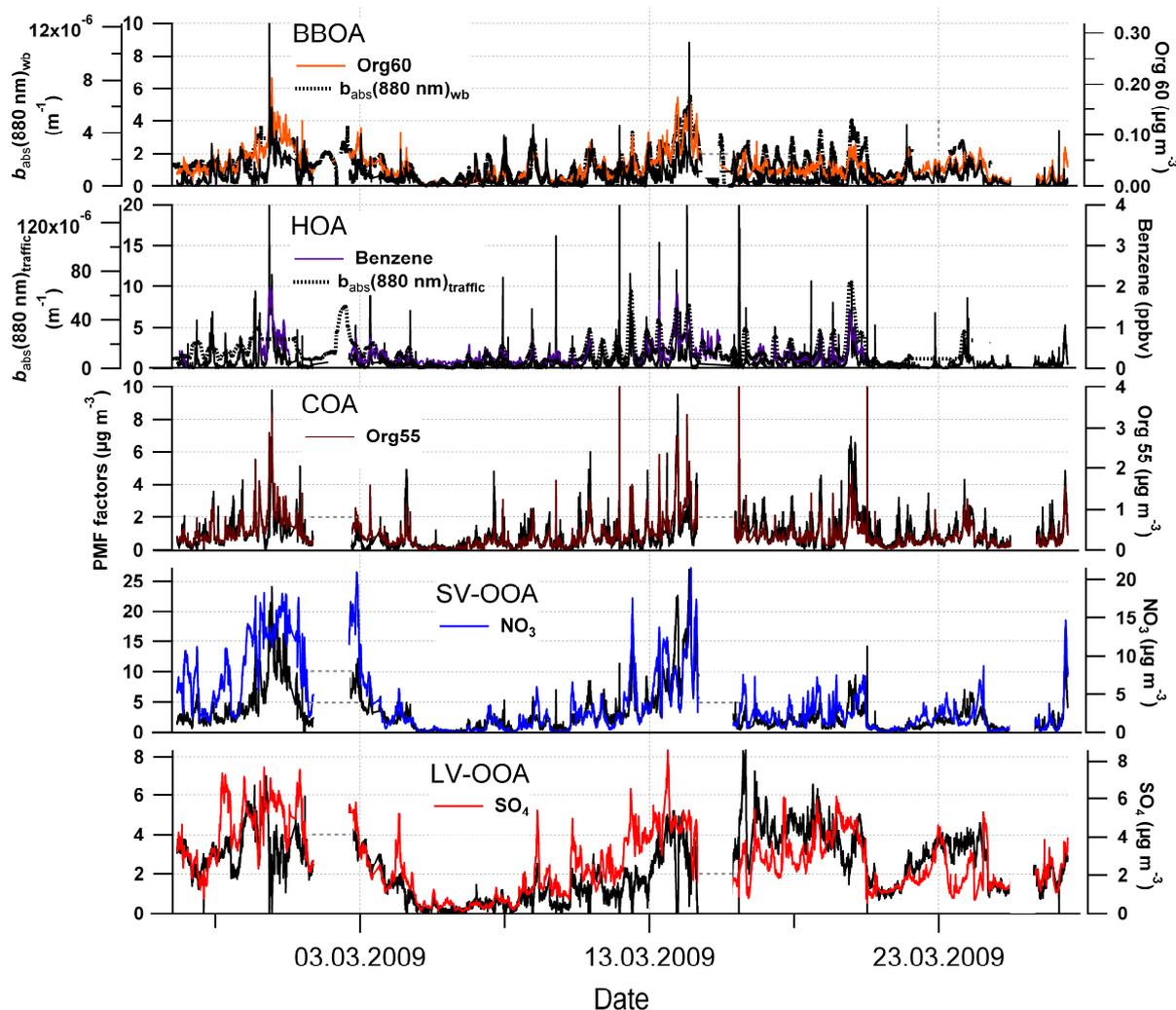
39 **3 PMF**

40 **3.1 UMR solution**



41

42 **Figure S3: Mass spectra of the UMR 5-factor-PMF solution.**



43

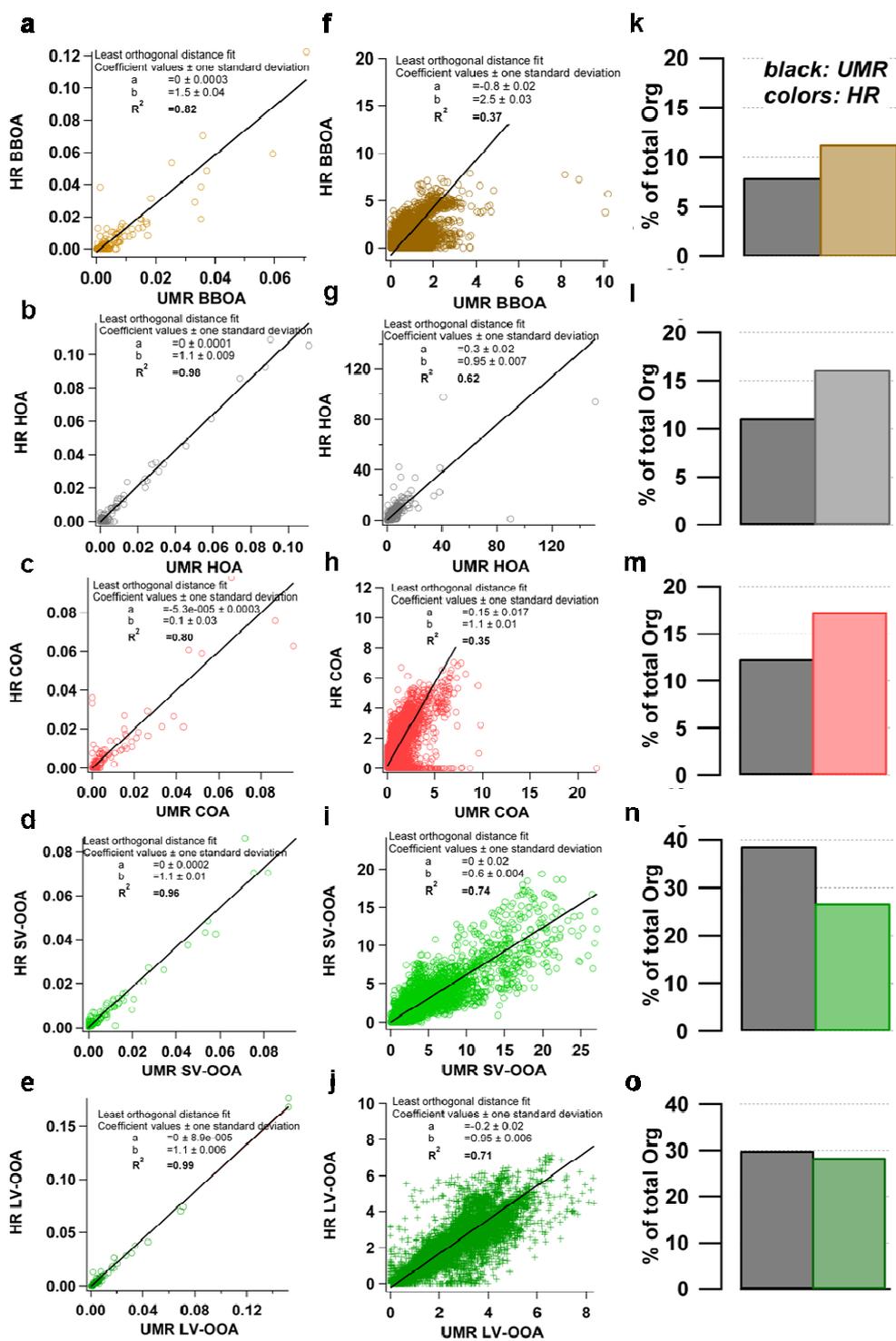
44 **Figure S 4: Time series of the UMR 5-factor-PMF solution and ancillary data.**

45

46 **3.2 Comparison of UMR and HR PMF solution**

47 The R^2 of the correlation of the mass spectra of the UMR and HR PMF solution range from 0.80
 48 (COA) to 0.99 (LV-OOA), confirming their similarities. Bigger differences can be seen in the
 49 time series of the corresponding factors. The COA time series show discrepancies in the total
 50 mass especially in the beginning of the campaign (until 02 March 2009), visualized in the data
 51 points with a much lower slope in Fig. S5 h. For the BBOA, the UMR time series features peaks
 52 not inherent to the HR time series. Concerning the mass attribution to each factor, HR generally
 53 assigns more mass to the primary OA factors and less to the OOA factors. Here the higher

54 resolution and, relied to that, the signal on an individual ion basis of the HR data matrix adds
55 additional information to the HR data matrix and thus allows for a better quantification of
56 primary and secondary OA.



57
58
59
60

Figure S5: Scatter plots of UMR and HR PMF spectra (a-e), time series (f-j) and a comparison of the mass attributed to each factor relative to OA (k-o).

61 3.3 UMR solution criteria

62 The PMF solution for a chosen number of factors p is a weighted iterative least squares fit
63 minimizing Q as in Eq. (1), with m and n denoting the rows and columns of the input matrices,
64 respectively. The known standard deviations σ_{ij} of the measured input values x_{ij} are used to
65 determine the weights of the residuals e_{ij} .

$$66 \quad Q = \sum_{i=1}^m \sum_{j=1}^n (e_{ij} / \sigma_{ij})^2 \quad (1)$$

67 If the model is appropriate and the data uncertainties estimations are accurate, $(e_{ij} / \sigma_{ij})^2$ is ~ 1 and
68 the expected Q ($Q_{expected}$) = $mn - p(m+n) \approx mn$, the degrees of freedom of the fitted data. The Q -
69 value is one mathematical criterion for the quality of the PMF solution: $Q/Q_{expected} \gg 1$
70 indicates an underestimation, $Q/Q_{expected} \ll 1$ an overestimation of errors in the input data
71 (Paatero et al., 2002). The mathematically correct value of p in PMF would be where the line
72 changes the slope in the plot of a series of p values versus their respective minimized Q (Fig. S6
73 a). However, a PMF solution has to be feasible in an ambient context and thus does not
74 necessarily correspond to the mathematically correct value of p .

75 Rotational ambiguity is a significant problem in the use of factor analysis (Paatero et al., 2002).
76 PMF solutions are not unique since linear transformation (still conserving the non-negativity
77 constraint) are possible ($\mathbf{GF} = \mathbf{GTT}^{-1}\mathbf{F}$). The rotational freedom of the chosen solution can be
78 explored through a non-zero valued user-specified rotational parameter f_{peak} . $f_{peak} > 0$ tries to
79 impose rotations on the emerging solutions using positive coefficients r in \mathbf{T} , $f_{peak} < 0$ vice
80 versa. $f_{peak} = 0$ produces the most central solution (Fig. S6 b).

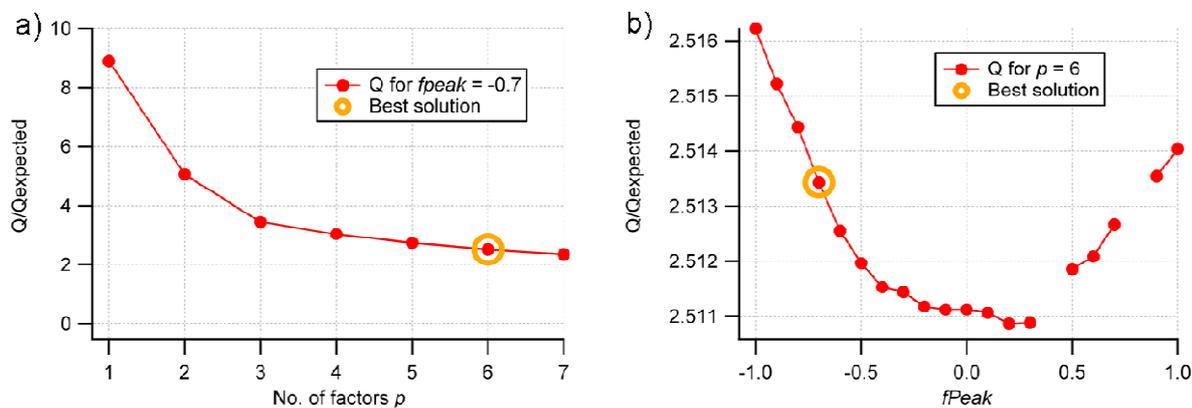
81 The number of factors p was chosen to be 6 for the UMR dataset (Fig. S7). In the solution with
82 $p = 5$ (Fig. S8 a), the spectra of BBOA, HOA, and COA are less clearly separated (e. g. high
83 signal at m/z 57 in the top factor resembling BBOA, but very little signal at m/z 57 in the red
84 factor resembling COA). Figure S8 b shows the time series of the 5-factor solution – they are
85 less clearly distinct than those of the 6-factor solution. The 7-factor solution (Fig. S9, a) features
86 a factor consisting mostly of signal at m/z 43 and a factor (orange) with single, isolated peaks
87 inconsistent with regular ion series. The time series show a more similar evolution (Fig. S9 b),
88 indicating a split of factors.

89 For the PMF solution presented in the manuscript, the 6-factor solution was chosen and the two
90 factors assigned to SV-OOA (black and purple) regrouped to one SV-OOA, using the sum for
91 the time series and the loadings-weighted average of the spectra.

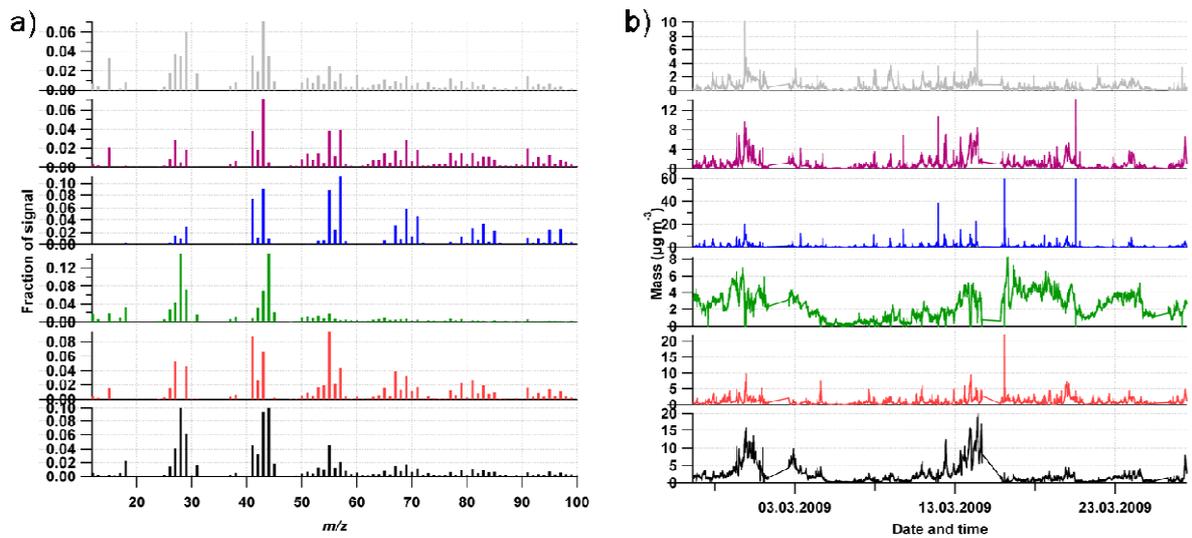
92 Figure S10 presents the explained variance of the organics as a function of f_{peak} for the chosen
93 6-factor solution. f_{peak} was chosen to be -0.7 based on correlations of the corresponding factors
94 with reference spectra.

95 A boxplot of the scaled residuals (boxes are +/- 25% of points) per m/z is shown in Fig. S11,
96 time series of the residuals and $Q/Q_{expected}$ are shown in Fig. S12. On 16 March 2009, a power
97 failure led to a breakdown of the instrument and a subsequent pumping down effect (Fig. S12).
98 Downweighting this period in the input for PMF did not alter the solution.

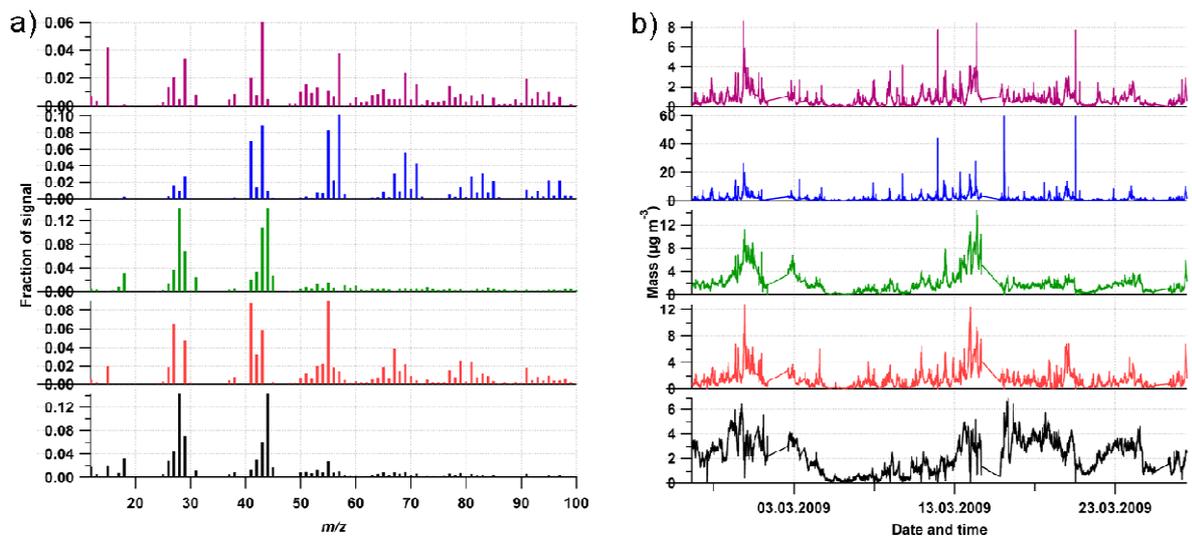
99 The solution space for the chosen $p = 6$ (central rotation) was explored by running PMF with 50
100 random initial values ($SEED$) at iteration start (Figs. S13 – 14). Roughly three solution groups
101 can be identified (numbers in Fig. S14). Groups 1 and 2 feature a factor spectrum predominantly
102 consisting of m/z 43 and two spectra that are basically identical. The spectrum with BBOA-like
103 features shows no contributions at m/z 44, which is inconsistent with previous studies. For group
104 3, all spectra not assigned to OOA show very high similarities. The solution with a central
105 rotation ($f_{peak} = 0$) was thus discarded regardless of $SEED$ values. Similar information was also
106 published in the supplementary information of Mohr et al. (2011).



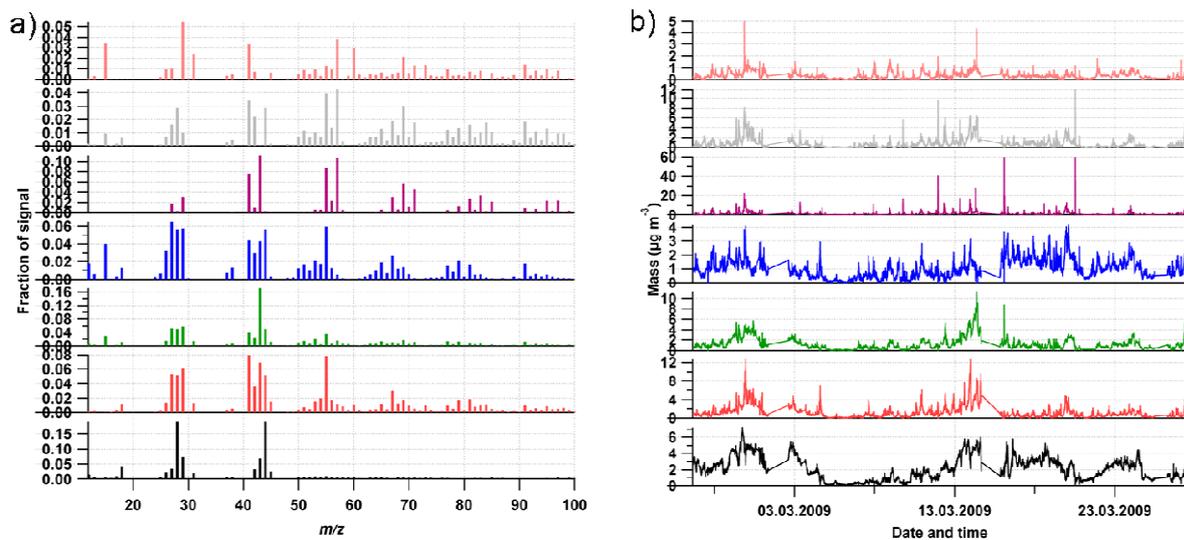
107
108 **Figure S 6: $Q/Q_{expected}$ versus the number of factors p (a) or f_{peak} (b). The orange circle denotes the chosen**
109 **UMR solution.**



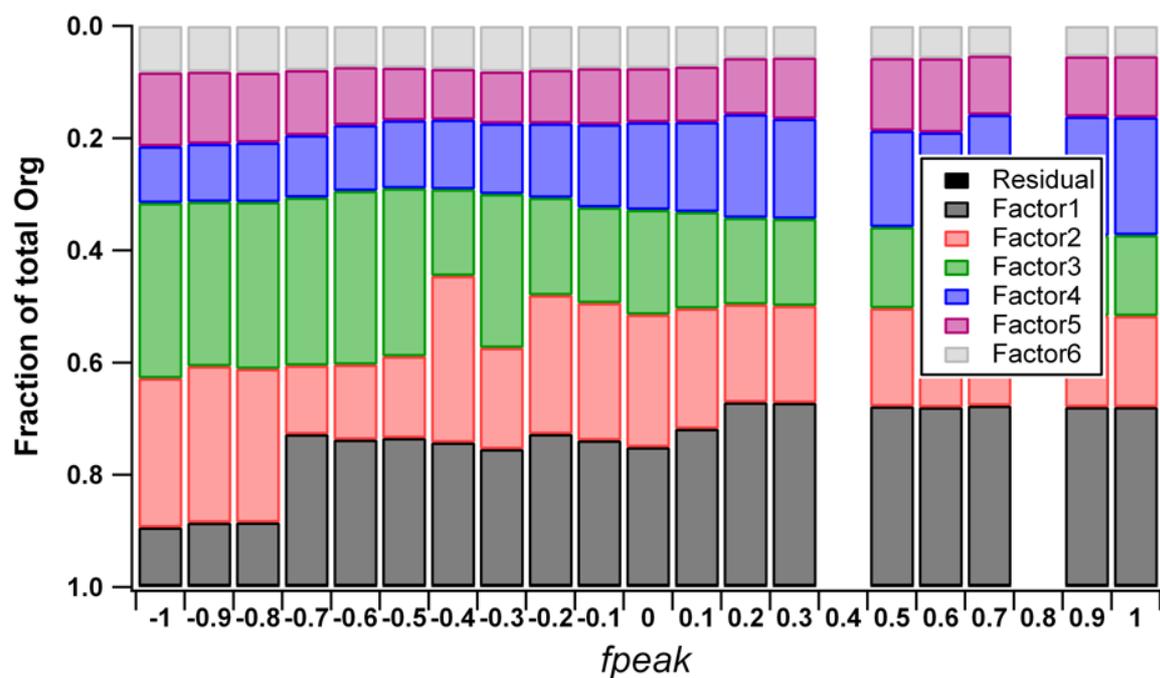
110
 111 **Figure S7: 6-factor UMR solution chosen, mass spectra (a) and time series (b). The black and the purple**
 112 **factor (SV-OOA 1 and 2) were regrouped to SV-OOA.**
 113



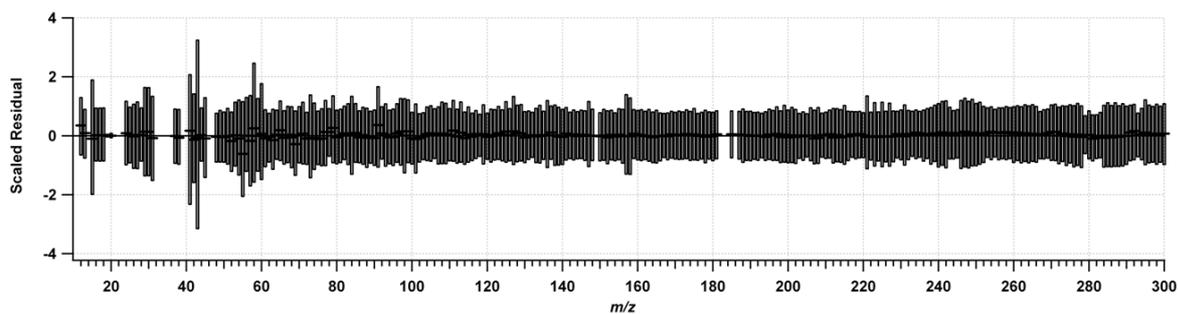
114
 115 **Figure S8: 5-factor UMR solution, mass spectra (a) and time series (b).**



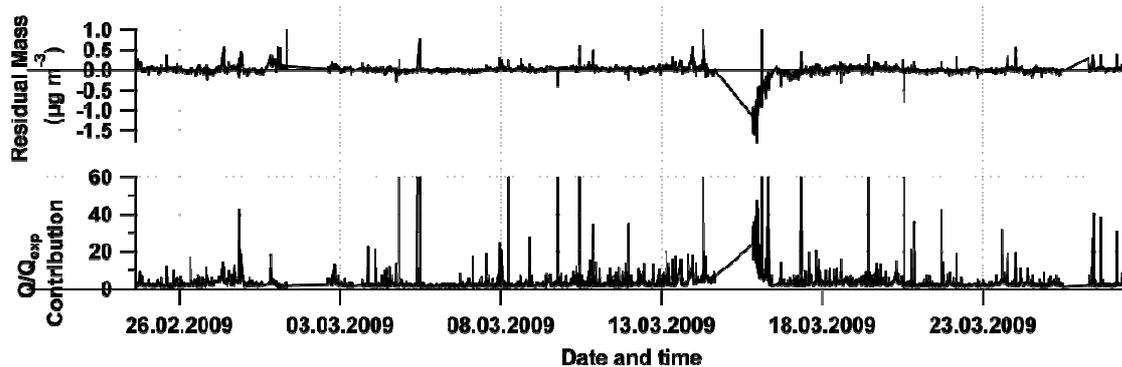
116
 117 **Figure S9: 7-factor UMR solution, mass spectra (a) and time series (b).**
 118



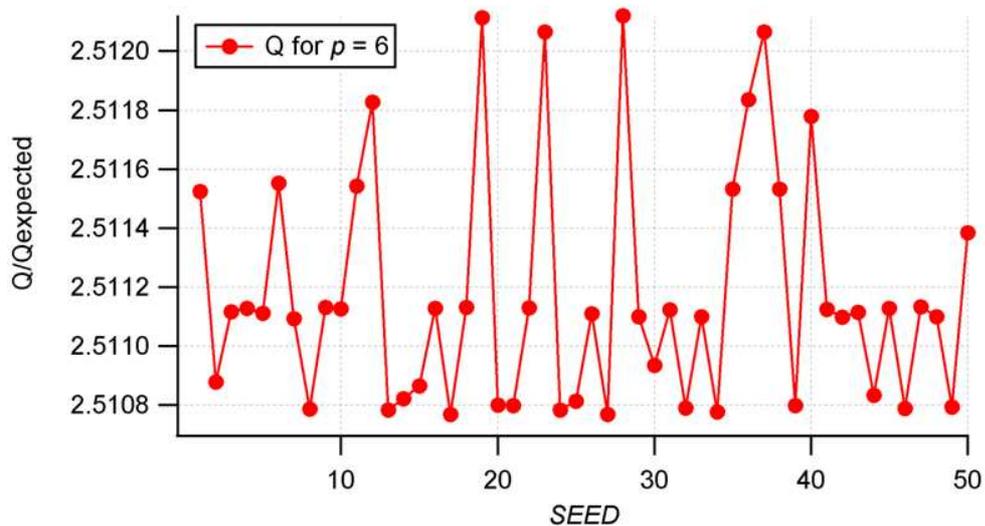
119
 120 **Figure S10: Variance explained by PMF due to the 6-factor UMR solution as a function of f_{peak} . For the**
 121 **solution presented, $f_{peak} = -0.7$.**
 122
 123
 124



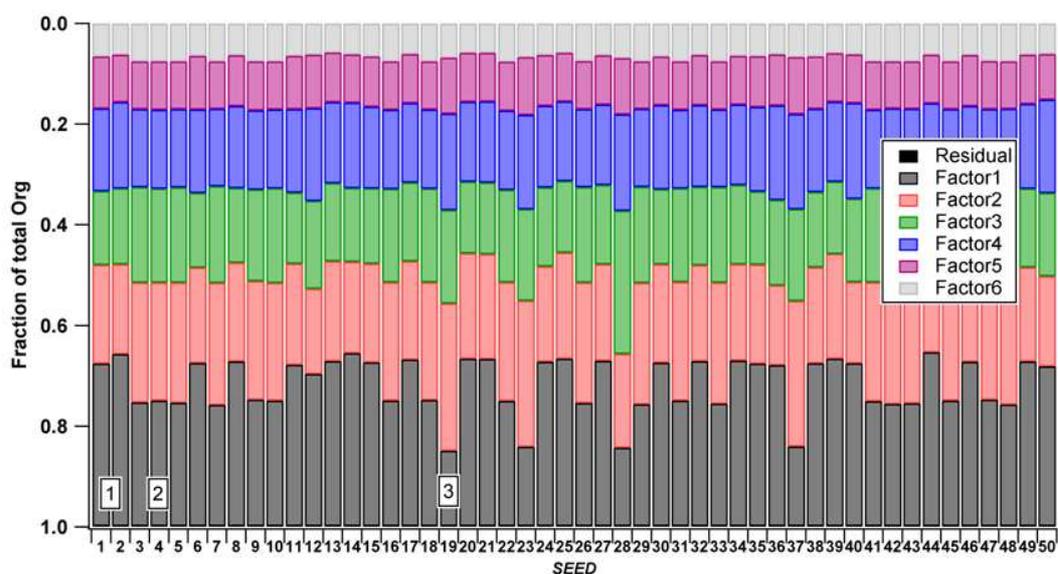
125
 126 Figure S11: Median black strokes) and lower/upper quartiles (boxes) of the scaled residuals per m/z .
 127



128
 129 Figure S12: Time series of scaled residuals (top panel) and $Q/Q_{expected}$ (lower panel).
 130



131
 132 Figure S13: $Q/Q_{expected}$ as a function of different $SEED$ values.
 133



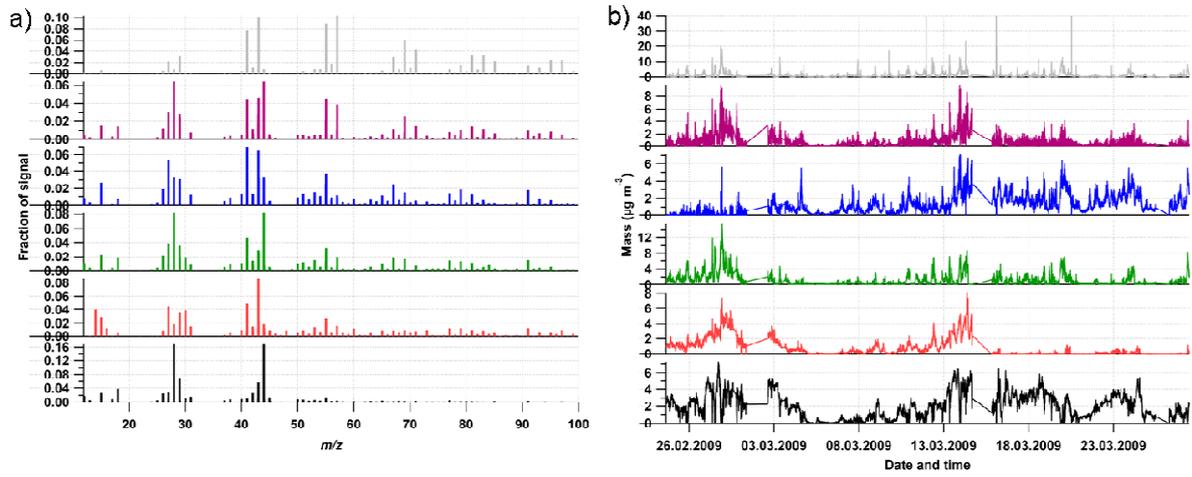
134
 135 **Figure S14: Variance explained by PMF due to the 6-factor UMR solution as a function of *SEED*. The**
 136 **numbers 1, 2 and 3 denote the three solution groups identified (see text).**
 137

138 3.4 HR solution criteria

139 Figure S15 shows the chosen 6-factor solution for the HR dataset. The SV-OOA factors (green
 140 and purple) were regrouped to a single SV-OOA factor as described in the previous UMR
 141 section. The 5-factor solution (Fig. S16) was discarded due to the high similarity of two factors
 142 (green and blue, spectra and time series). The 7-factor solution (Fig. S17) features three LV-
 143 OOA spectra: PMF seems to assign an individual LV-OOA factor to three different sections in
 144 the time series.

145 As the variance in the solution space $p = 6$ as a function of f_{peak} is negligible (Figs. S18 – 19),
 146 the most central solution ($f_{peak} = 0$) was chosen and different *SEED* values were explored (Figs.
 147 S20 – 21). *SEED* = 64 was chosen as the best solution due to correlations with reference spectra.
 148 A boxplot of the scaled residuals (boxes are +/- 25% of points) per m/z is shown in Fig. S22,
 149 time series of the residuals and Q/Q_{exp} are shown in Fig. S23. The same irregularities as for the
 150 UMR data can be observed as well in the HR data.

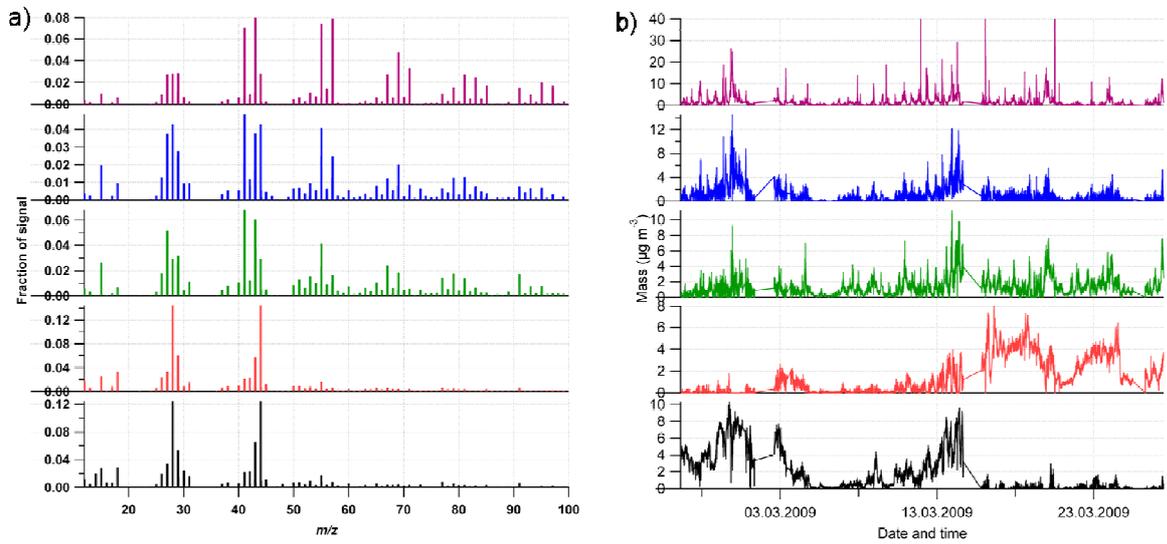
151
 152



153

154 **Figure S15: Chosen 6-factor solution of the HR dataset, mass spectra (a) and time series (b).**

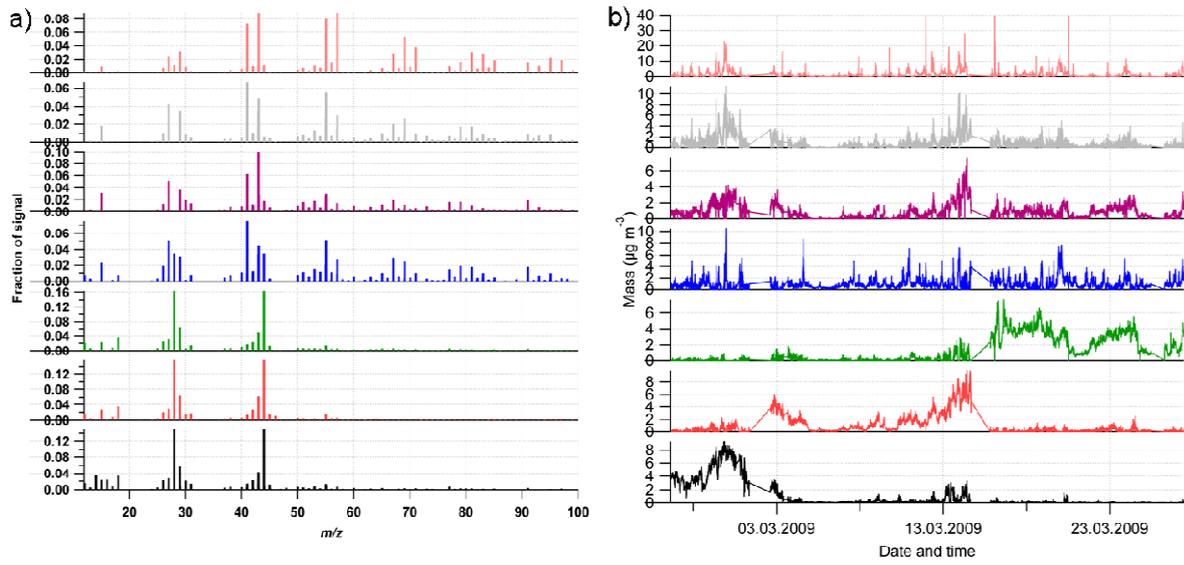
155



156

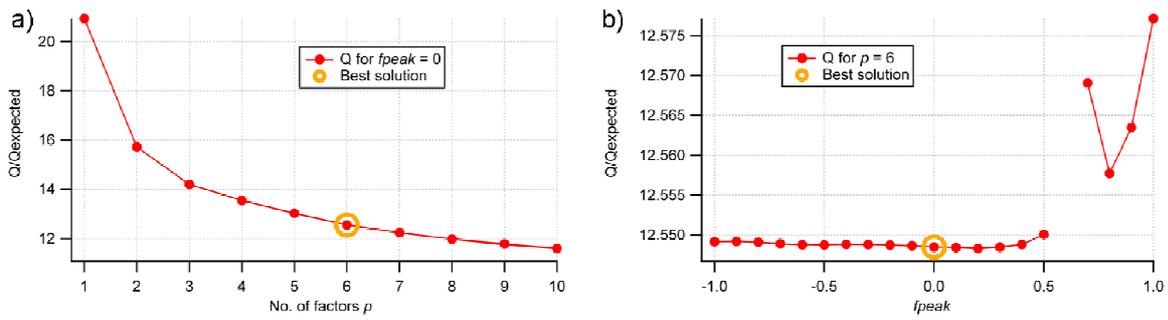
157 **Figure S16: 5-factor solution for the HR dataset, mass spectra (a) and time series (b).**

158



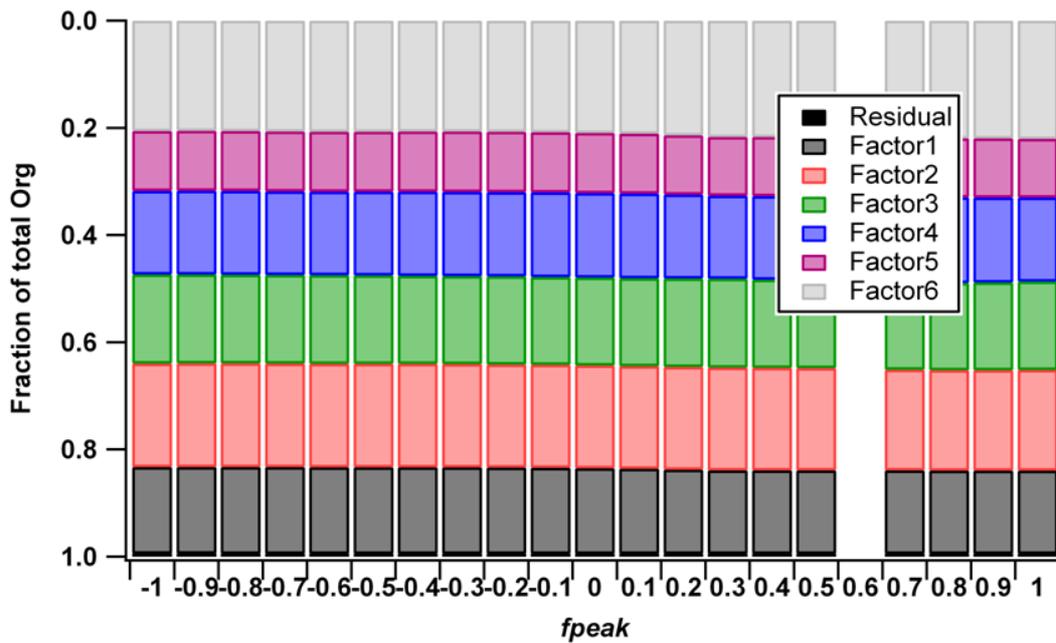
159
160
161

Figure S17: 7-factor solution for the HR dataset, mass spectra (a) and time series (b).

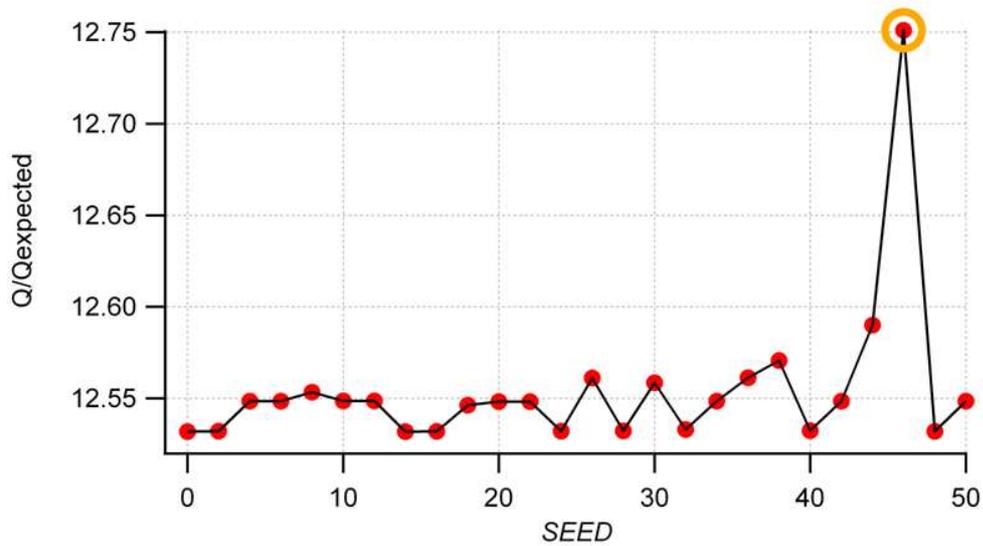


162
163
164
165

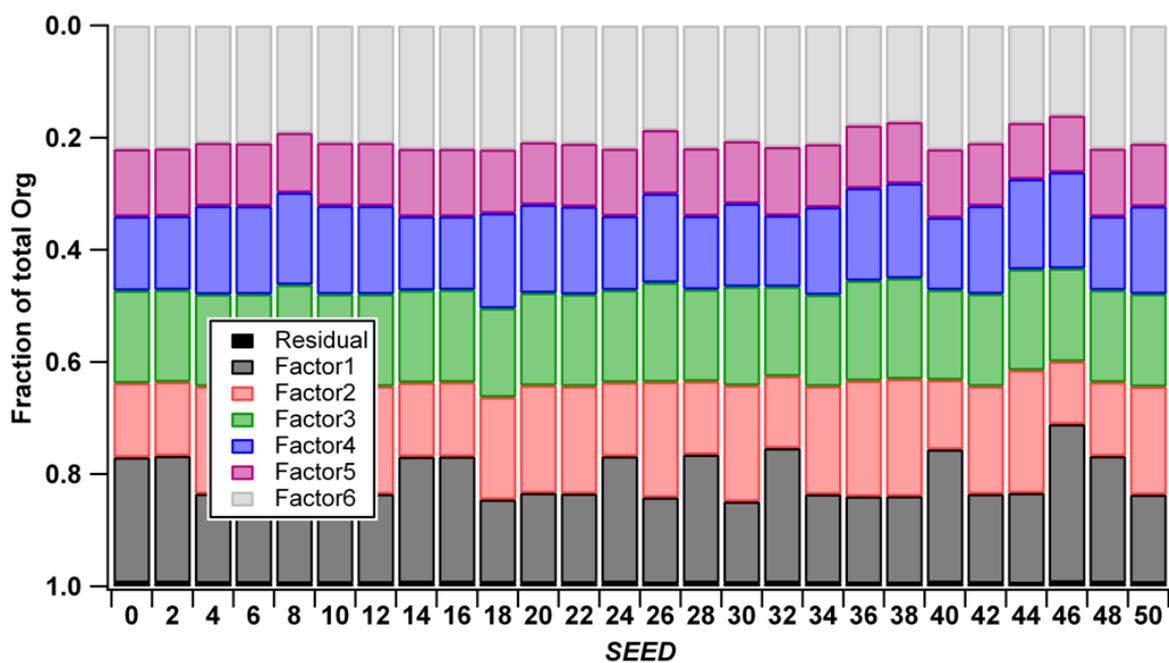
Figure S18: Q/Q_{expected} versus the number of factors p (a) or f_{peak} (b), HR PMF. The orange circle denotes the chosen solution.



166
 167 **Figure S19: Variance explained by PMF due to the 6-factor HR solution as a function of f_{peak} . For the**
 168 **solution presented, $f_{peak} = 0$.**
 169



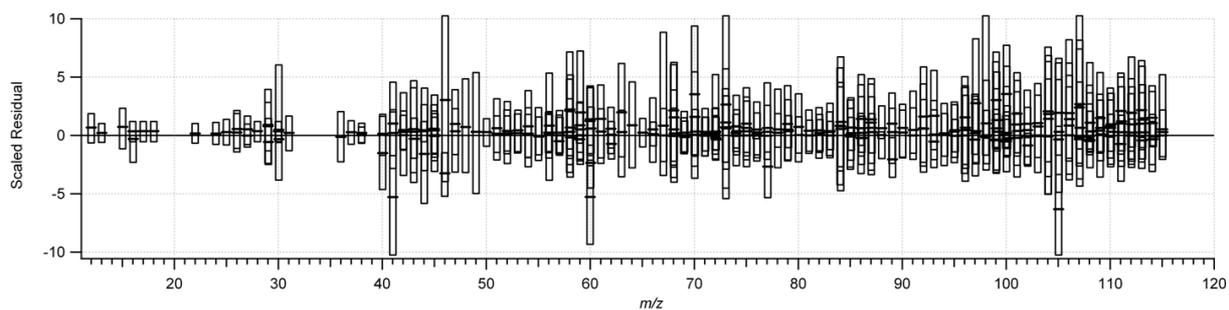
170
 171 **Figure S20: $Q/Q_{expected}$ versus $SEED$ for the HR solution.**
 172



173

174 **Figure S21: Variance explained by PMF due to the 6-factor HR solution as a function of SEED.**

175

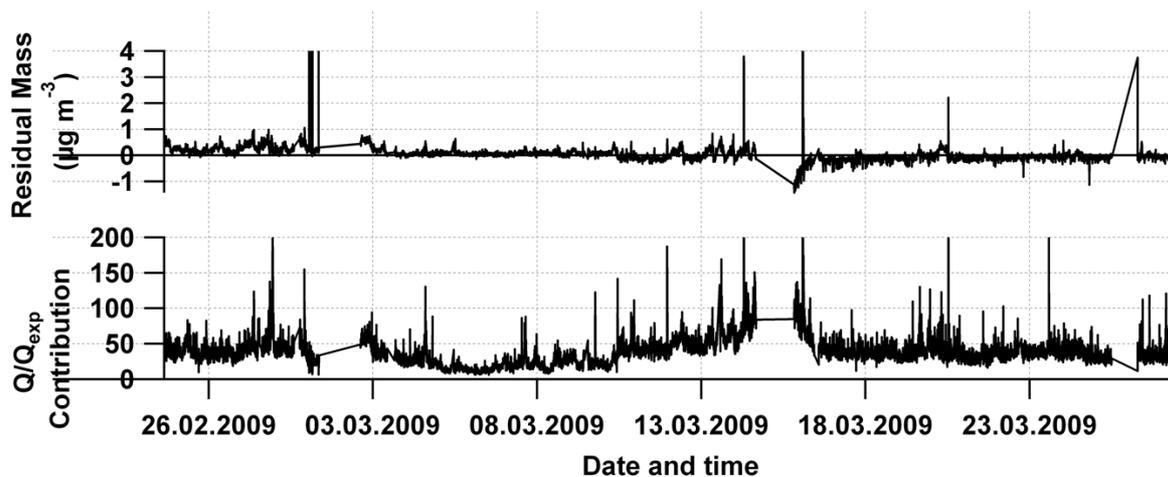


176

177 **Figure S22: Median (black strokes) and lower/upper quartiles (boxes) of the scaled residuals per m/z (HR**

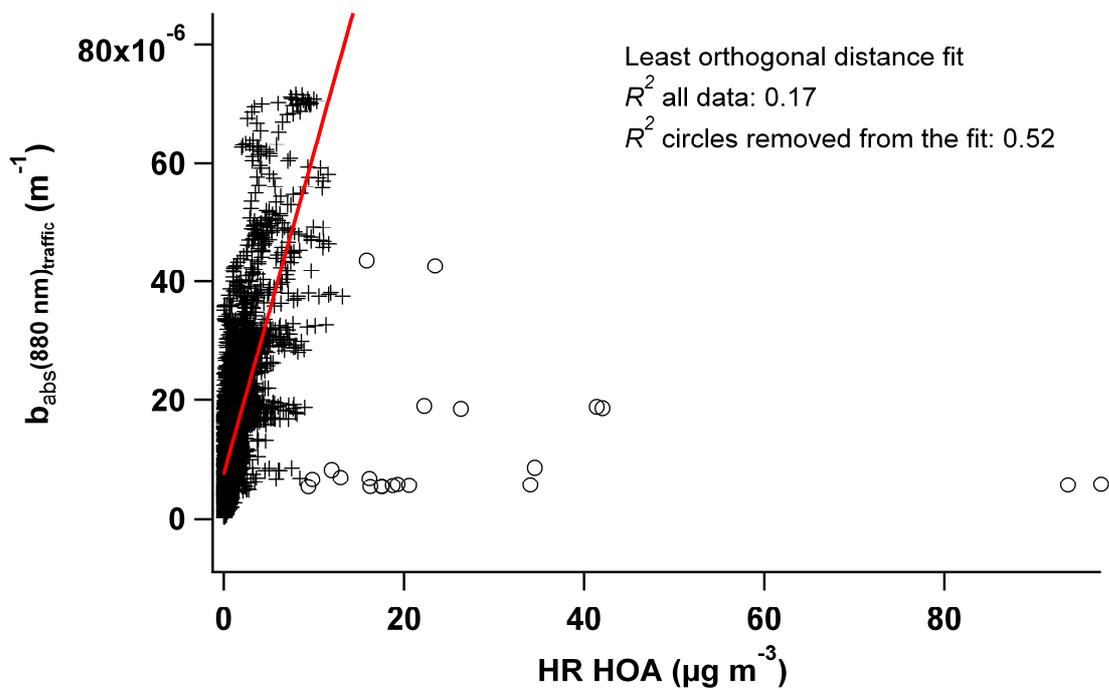
178 **solution).**

179



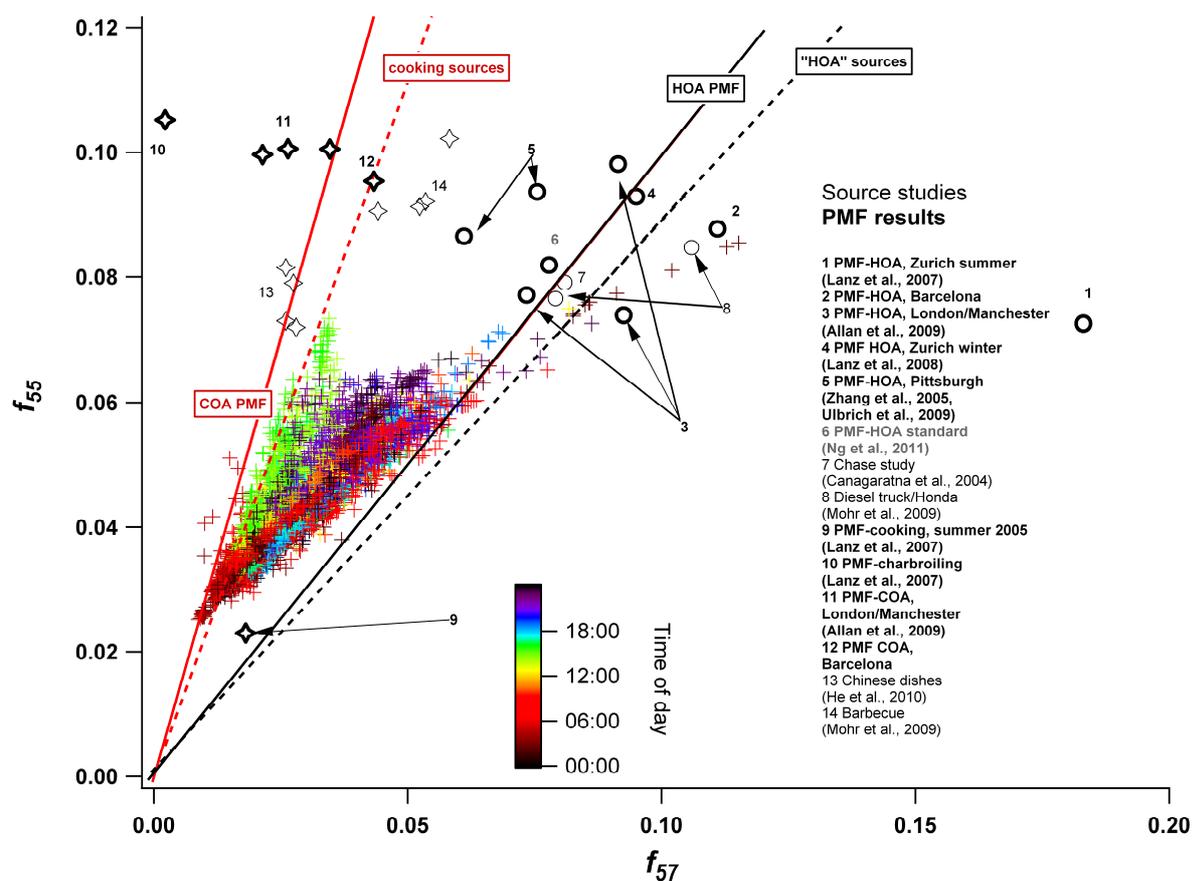
180
181
182

Figure S23: Time series of scaled residuals (top panel) and $Q/Q_{expected}$ (lower panel) for the HR solution.



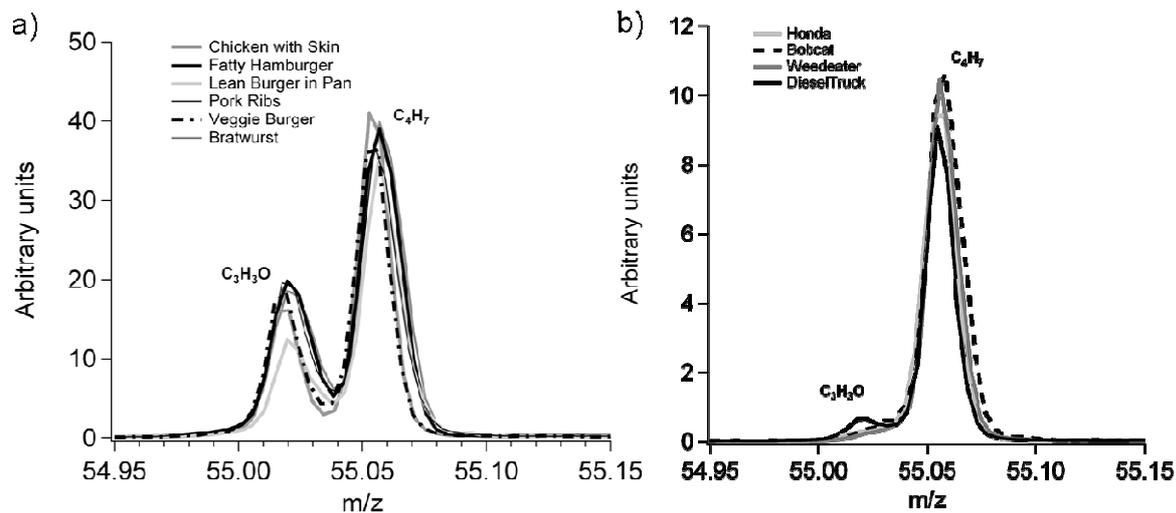
183
184
185
186

Figure S24: Scatter plot of the time series of $b_{abs}(880 \text{ nm})_{traffic}$ and HOA. The red line is the least orthogonal distance fit where the circle data points were removed.



187
188 **Figure S 25: m/z 55/Org (f_{55}) plotted against m/z 57/Org (f_{57}).**

189



190
191 **Figure S26. Signal at m/z 55 in the HR spectra of meat cooking sources (a) and vehicle engine sources (b). In**
192 **the engine exhaust spectra, the signal is almost entirely due to the reduced hydrocarbon ion $C_4H_7^+$, whereas in**

193 the cooking spectra there is also substantial contribution from the oxygen-containing ion $C_3H_3O^+$. Reprinted
194 from Mohr et al. (2009).

195

196 4 References

197 Mohr, C., Huffman, J. A., Cubison, M. J., Aiken, A. C., Docherty, K. S., Kimmel, J. R., Ulbrich,
198 I. M., Hannigan, M., and Jimenez, J. L.: Characterization of primary organic aerosol emissions
199 from meat cooking, trash burning, and motor vehicles with high-resolution aerosol mass
200 spectrometry and comparison with ambient and chamber observations, *Environ. Sci. Technol.*,
201 43, 2443-2449, 2009.

202 Mohr, C., Richter, R., DeCarlo, P. F., Prévôt, A. S. H., and Baltensperger, U.: Spatial variation of
203 chemical composition and sources of submicron aerosol in Zurich during wintertime using
204 mobile aerosol mass spectrometer data, *Atmos. Chem. Phys.*, 11, 7465-7482, 2011.

205 Paatero, P., Hopke, P. K., Song, X. H., and Ramadan, Z.: Understanding and controlling rotations
206 in factor analytic models, *Chemometr. Intell. Lab.*, 60, 253-264, 2002.

207

208