Supplement of Atmos. Chem. Phys., 17, 13265–13282, 2017
https://doi.org/10.5194/acp-17-13265-2017-supplement

Atmospheric Chemistry and Physics
Open Access
EGU

Supplement of

Long-term chemical analysis and organic aerosol source apportionment at nine sites in central Europe: source identification and uncertainty assessment

Kaspar R. Daellenbach et al.

Correspondence to: André S. H. Prévôt (andre.prevot@psi.ch) and Imad El Haddad (imad.el-haddad@psi.ch)

- Number of factors:

Based on the input data for PMF_{block}, we evaluate the influence of the number of factors, p, on $Q_{i,j}$. For this experiment, both the traffic and cooking signatures were constrained using adapted reference profiles from Crippa et al. (2013b) as described in section III.1. Based on this evaluation, we chose to perform PMF using 6 factors.

$Q_{i,j}$ is computed using the PMF residuals (e_{ij}) and the PMF input errors ($s_{i,j}$):

$$Q_{i,j} = \left(\frac{e_{i,j}}{s_{i,j}}\right)^2 \quad \text{(S1)}$$

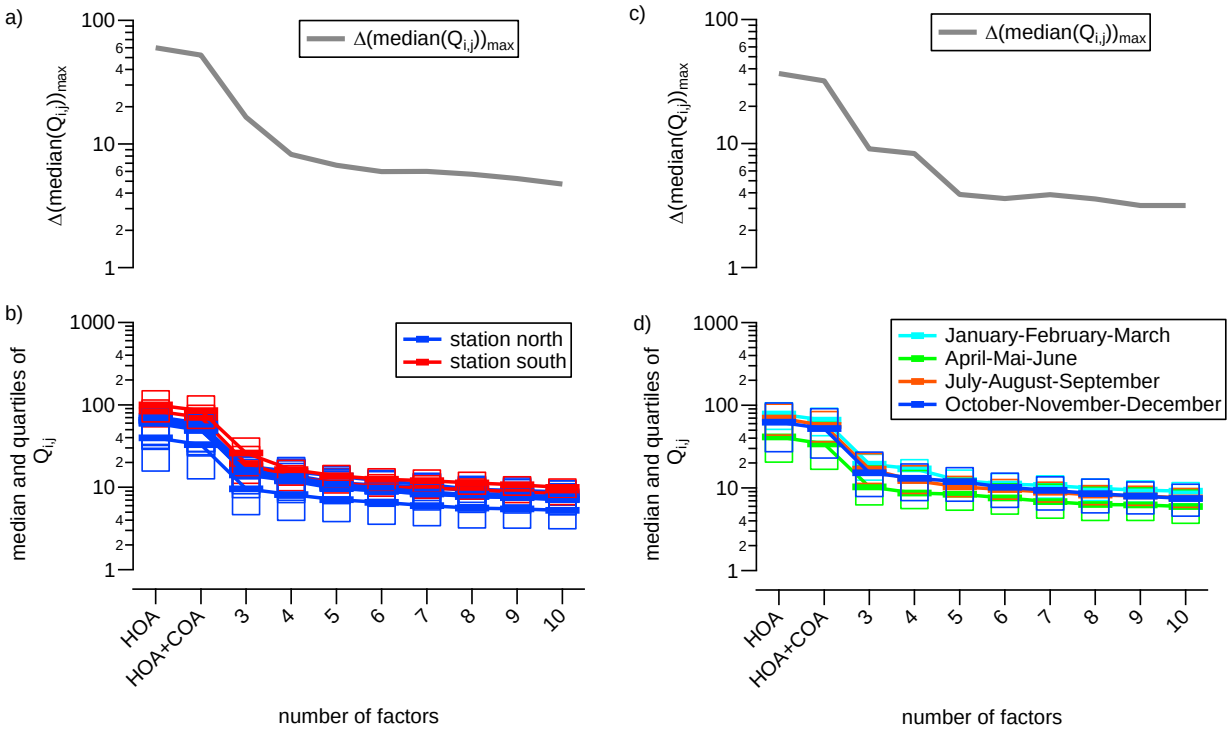

Figure S1: $Q_{i,j}$ as a function of the number of factors for a reference experiment with all data used in PMF (9 sites, full year 2013, HOA and COA constrained with a=0.0 (b and d). $\Delta(\text{median}(Q_{i,j}))_{max}$ is evaluated for the different periods during the year 2013 (January-February-March, April-Mai-June, July-August-September, October-November-December) and for all sites (a and c). The grey line depicts the difference between the category (geographical or season) with the highest and the lowest median $Q_{i,j}$.

Fig. S1 shows $Q_{i,j}$ s as a function of the number of factors for different sites (b) and seasons (d) and the difference between the highest (a) and lowest (c) median to evaluate the maximal difference in the mathematical quality of the solutions. As expected, forcing PMF to explain the variability in the dataset only with the 2 constrained factors (p=2), results in very high median $Q_{i,j}$. $\Delta(\text{median}(Q_{i,j}))_{max}$ shows the difference in the median $Q_{i,j}$ between groups of points like sites or season. The

smaller the $\Delta(\text{median}(Q_{i,j}))_{max}$, the smaller are the differences in the mathematical quality of the PMF solution for the different seasons/sites. To explain the temporal and geographical variability at least 5 factors are required. However, the difference between the site that is best explained and the site that is least explained is approximately 6 when using 5 or 6 factors. When increasing to 6 factors, also a factor explaining the variability of sulfur-containing organic ions (especially, $CH_3SO_2^+$) is resolved. Therefore, we opted to perform PMF using 6 factors. Using 6 factors, there is also no difference between the average $Q_{i,j}$ on week-days and weekend (Fig. S2).

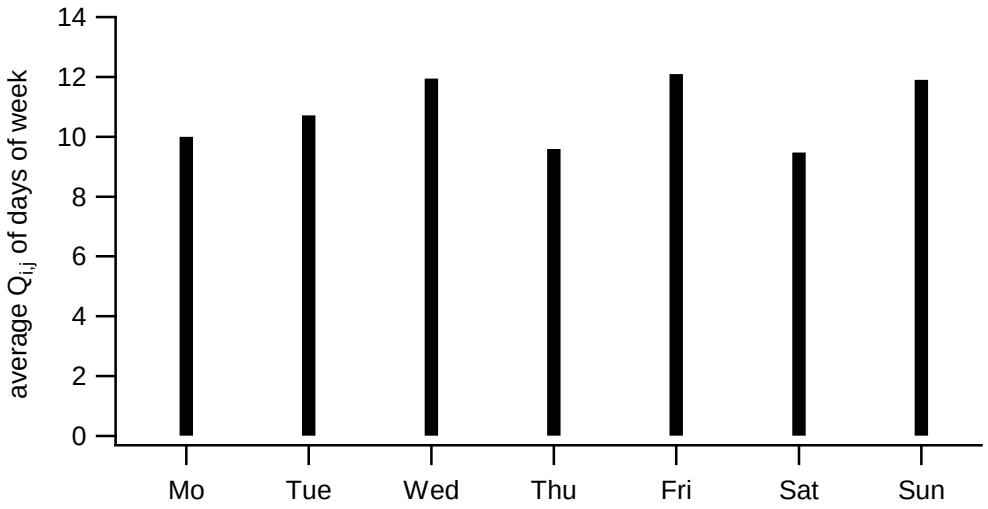

Figure S2: $Q_{i,j}$ as a function of the day of the week.

However, for PMF_{block} also with 6 factors, the average $Q_{i,j}$ is clearly larger (7 only the Zurich data points) than the ideal value of 1, i.e. the PMF residuals are larger than the measurement uncertainties. In comparison to PMF_{block}, the average $Q_{i,j}$ for Zurich is slightly reduced for the same number of factors when only including 1 site in PMF ($PMF_{zue,isol,}$ $PMF_{zue,reps}$, average $Q_{i,j}$ 6). In this study, we analyse yearly cycles and, thereby, assume constant factor profiles throughout the year which can contribute to Q>1.

Another possible reason for Q>1 is an underestimation of the measurement uncertainty. A main contributor in high-resolution AMS data treatment (attribution of the signal at a nominal mass to several ions) stems from errors in the *m/z* calibration which could not be incorporated in the current data analysis. Recent studies demonstrate that for overlapping peaks (ions) the measurement uncertainties are strongly underestimated (Cubison et al., 2015; Corbin et al., 2015). For PMF_{block} using 6 factors, average $Q_{i,j}$ do not depend on *m/z* but rather on the ion family (Fig. S3): ions consisting of C, H, S, (and O) summarized under the name (CS) and ions consisting of C, H, N, (and O) summarized under the name CHN have a higher $Q_{i,j}$ than hydrocarbon ions (CH, only C and H) and oxygenated ions ($CHO_{z=1}$ with 1 oxygen and $CHO_{z>1}1$ with more than 1 oxygen). Since the time series of $CH_3SO_2^+$ is event-driven, the high $Q_{i,j}$ of this ion hints to the fact that PMF is unable to accurately resolve all of these events.

The average $Q_{i,j}$ for ions with a mass defect (nominal mass – exact ion mass) around 0.03 a.m.u. is higher than for the other ions (Fig S3). Mass defects in this range are most common in our dataset. This makes these peaks prone to overlap with

other ions and thus their error prone to an underestimation because this effect is not considered in the s_{ij} calculation (described above).

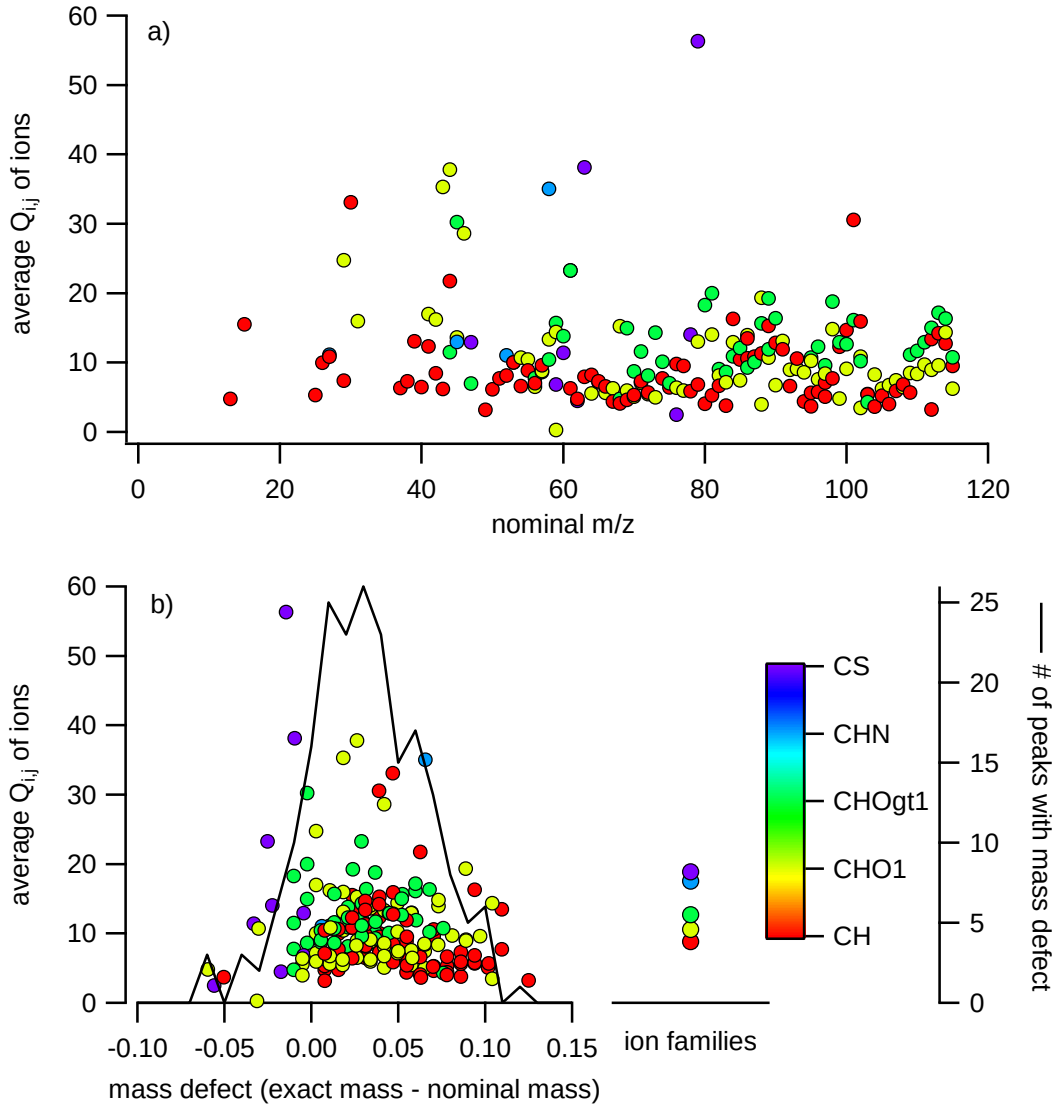

Figure S3: a) Average $Q_{i,j}$ of ions in PMF_{block} as a function of their mass-to-charge ratio (m/z). The ions are color-coded with their composition (CH: ions consisting only of C and H; CHO1: ions consisting of C, H, and 1 O; CHOgt1: ions consisting of C, H, and more than 1 O; CHN: ions consisting of C, H, N, (and O); CS: ions consisting of C, H, S, (and O)). b) Average $Q_{i,j}$ of the ions in PMF_{block} as a function of their mass defect (exact mass – nominal mass) as well as a histogram of the number of ions with a certain mass defect. The mean $Q_{i,j}$ of the ion families is displayed separately.

Cumulative density functions for the a-values of HOA and COA are presented for the accepted solutions in Fig. S4. We found that 80% of the accepted solutions have an a-value≤0.3 for HOA and an a-value≤0.5 for COA. The output HOA and COA factor profiles are therefore not significantly variable and very similar to the input profiles, indicating that similar solutions were selected. Furthermore, the yearly average factor concentrations of all selected PMF_{block} solutions after R_k correction are shown for the case of Zurich as an illustration in Fig. S5. The distributions of each of the different factors do not show more than 1 distinct mode, indicating that we do not have several populations of solutions.

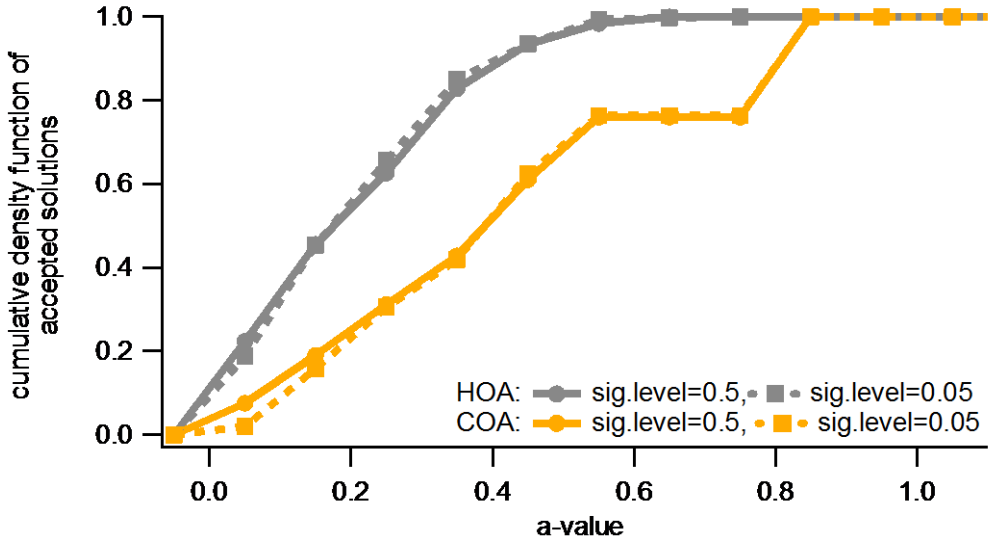

Figure S4: Cumulative density functions of a-values for HOA and COA for the accepted solutions.

The yearly average factor concentrations of all selected PMF_{block} solutions after R_k correction areshown for the case of Zurich as an illustration (Fig. S5). The distributions of each of the different factors do not show more than 1 distinct mode.

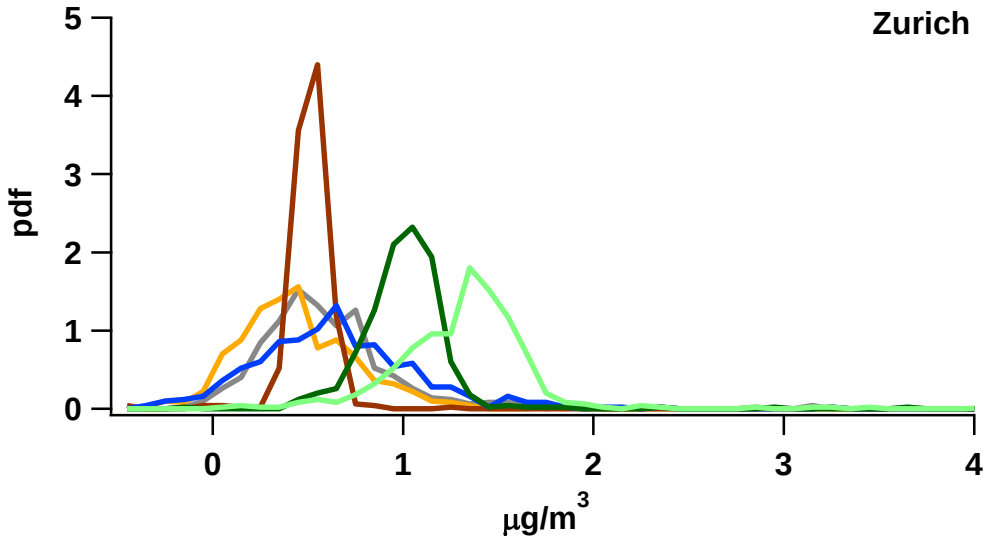

Figure S5: Histograms of yearly average factor concentrations of all selected PMF_{block} solutions (after R_k correction).

- Quality assessment of solutions:

Set of criteria used when assessing quality of a single PMF run:

Table S1: set of acceptance criteria used. r is the correlation coefficient between a factor time series and the respective marker. Q25 is the 1st quartile and Q75 the 3rd quartile.

criteria on profile		
	$f(CO_2^+)$	$f(C_2H_4O_2^+)$
HOA	<0.4	<0.004
COA	<0.4	<0.01
Criteria on time series		
HOA	r(HOA,NOx)> 0 & r(HOA,NOx)> r(COA,NOx)	
BBOA	r(BBOA,levo)> 0	
SC-OA	r(SC-OA,$CH_3SO_2^+$)>0	
Mass closure criteria		
OC_{res} total		Q_{25}(res-OC)<0 & Q_{75}(res-OC)>0
Magadino winter, Magadino summer		Q_{25}(res-OC)<0 & Q_{75}(res-OC)>0
Zurich winter, Zurich summer		Q_{25}(res-OC)<0 & Q_{75}(res-OC)>0
Magadino, Zurich		Q_{25}(res-OC)<0 & Q_{75}(res-OC)>0
HOC<median, HOC>median		Q_{25}(res-OC)<0 & Q_{75}(res-OC)>0
COC<median, COC>median		Q_{25}(res-OC)<0 & Q_{75}(res-OC)>0
BBOC<median, BBOC>median		Q_{25}(res-OC)<0 & Q_{75}(res-OC)>0
SC-OC<median, SC-OC>median		Q_{25}(res-OC)<0 & Q_{75}(res-OC)>0
WOOC<median, WOOC>median		Q_{25}(res-OC)<0 & Q_{75}(res-OC)>0
SOOC<median, SOOC>median		Q_{25}(res-OC)<0 & Q_{75}(res-OC)>0
for PMF with 12 filters per site summer Magadino and Zurich		Q_{25}(res-OC_i)<0 & Q_{75}(res-OC_i)>0
for PMF with 12 filters per site winter Magadino and Zurich		Q_{25}(res-OC_i)<0 & Q_{75}(res-OC_i)>0

- Comparison of mass spectral signature of BBOA and nebulized levoglucosan:

Figure S6 demonstrated the high similarity between the retrieved BBOA signature and the mass spectrum of nebulized levoglucosan.

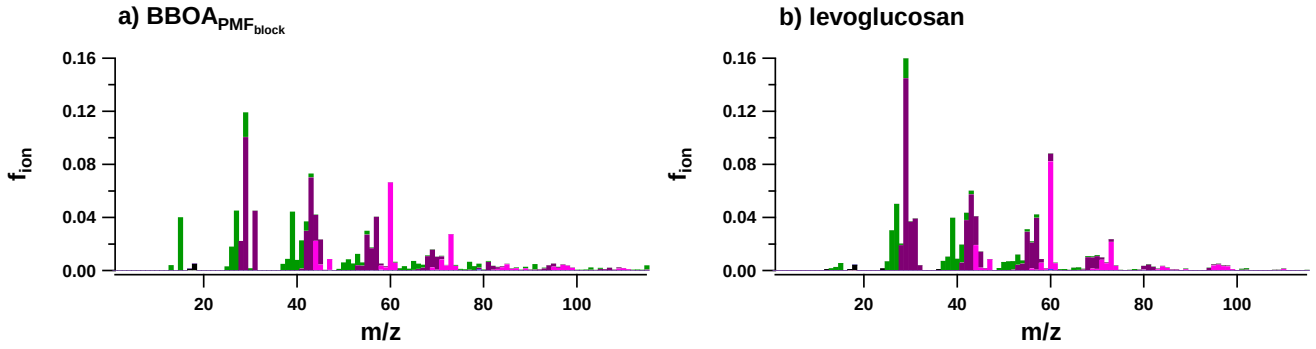

Figure S6: mass spectral fingerprints of BBOA (PMF_{block}) and nebulized levoglucosan. f_{ion} is the fraction of signal of a respective ion to the sum of the total signal.

- **Comparison of SOOA to ozone and Ox**

In Figure S7, we compare the SOOA concentrations to ozone and Ox (O_3+NO_2) for Zurich. The SOOA concentrations follow best the temperature ($R_{s,SOOA,temp}$=0.65, Fig. S7.a) but show also some correlation to ozone $R_{s,SOOA,O3}$=0.33, Fig. S7.b) and Ox ($R_{s,SOOA,Ox}$=0.38, Fig. S7.c).

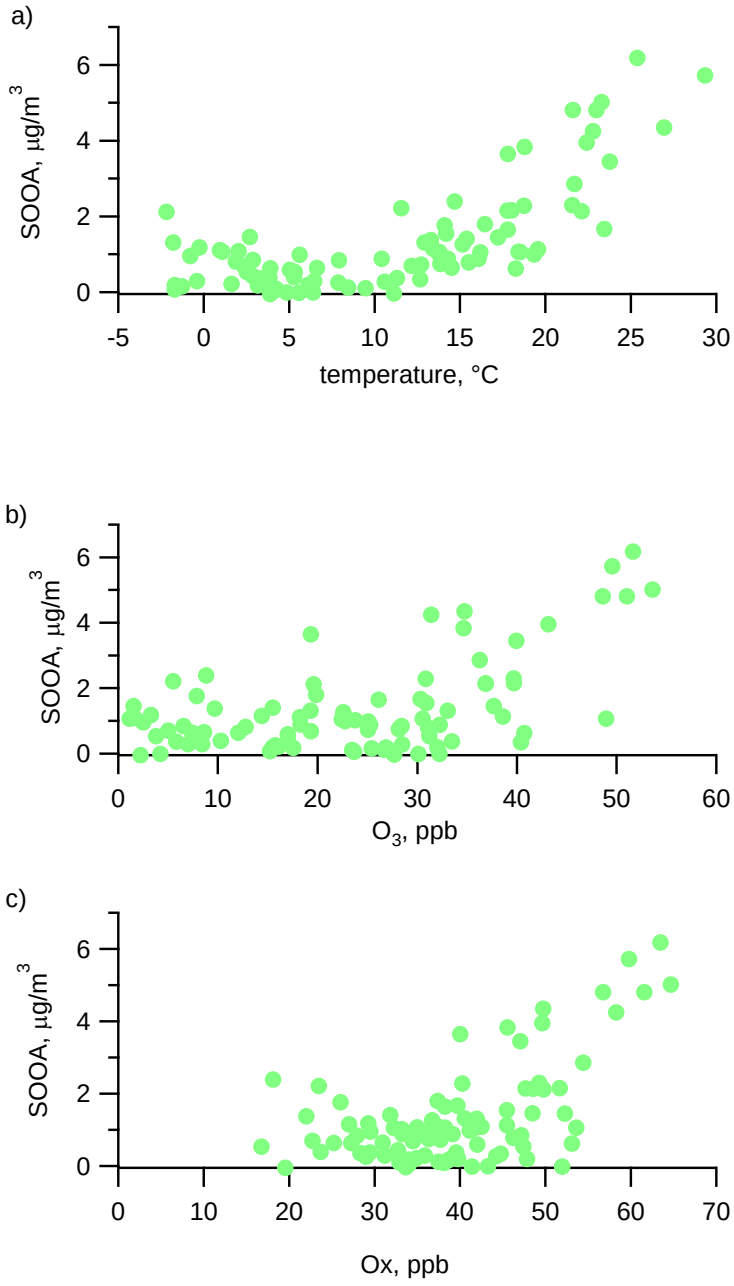

Figure S7: SOOA concentrations compared to temperature, ozone, and Ox (O_3+NO_2) for Zurich.

- Uncertainty estimation and propagation:

The uncertainty described by the interquartile range from the *a*-value sensitivity assessment (σ_a) does not fully explain the variability between the 4 sensitivity tests. In the following, we use the source apportionment results of the 12 filters common to all 4 sensitivity tests for achieving a better estimate of the uncertainty of the factor concentrations. For these 12 filters the uncertainty is estimated by propagating the variability between the median concentrations for the 4 sensitivity tests (σ_b) and half the interquartile range of PMF_{block} (σ_a, Eq. S2):

$$\boldsymbol{err_{i,k,tot}} = \sqrt{\boldsymbol{\sigma_a^2} + \boldsymbol{\sigma_b^2}} \qquad \textbf{S2}$$

In absence of σ_b for all other points, we parametrize σ_b. We express σ_b as a function of a minimal uncertainty ($\sigma_{minimal}$) and an uncertainty proportional (*k*) to the factor concentration and fit the equation using the 12 points in common to all datasets (Eq. S3):

$$\frac{\boldsymbol{\sigma_b}}{|\boldsymbol{conc_{i,k}}|} = \frac{\sqrt{\boldsymbol{\sigma_{minimal}^2} + \boldsymbol{k^2} * \boldsymbol{conc}_{i,k}^2}}{|\boldsymbol{conc_{i,k}}|} \qquad \textbf{S3}$$

The uncertainty (σ_a and σ_b) is propagated for all points using the parameters from Eq. S3 in order to obtain the total uncertainty for all points in the dataset (Eq. S4):

$$\boldsymbol{err'_{i,k,tot}} = \sqrt{\boldsymbol{\sigma_a^2} + \left(\sigma_{minimal}^2 + k^2 * conc_{i,k}^2\right)} \qquad \textbf{S4}$$

The resulting coefficients of the error model are presented in table S2:

Table S2: $\sigma_{minimal}$ and *k* for the different factors including their uncertainty.

factor	$\sigma_{minimal}$	*k*
HOA	0.16±0.06	0.39±0.24
COA	0.09±0.01	0.52±0.09
BBOA	0.06±0.01	0.48±0.05
SC-OA	0.30±0.00	0.32±0.27
WOOA	0.28±0.08	0.42±0.27
SOOA	0.05±0.01	0.24±0.05

- Sensitivity to significance level of statistical tests in PMFblock:

For PMF_{block}, a sensitivity test with significance level of 0.05 instead of 0.5 as in the base case was performed. The factor concentrations and their corresponding uncertainties (σ_a) are compared and displayed as number density functions (Fig. S8). Changes in the estimated factor concentrations are within 10% of the factor concentrations for SCOA and smaller for all other factors. The uncertainty related to COA is decreased when lowering the significance level to 0.05, while the other factors remain largely unaffected.

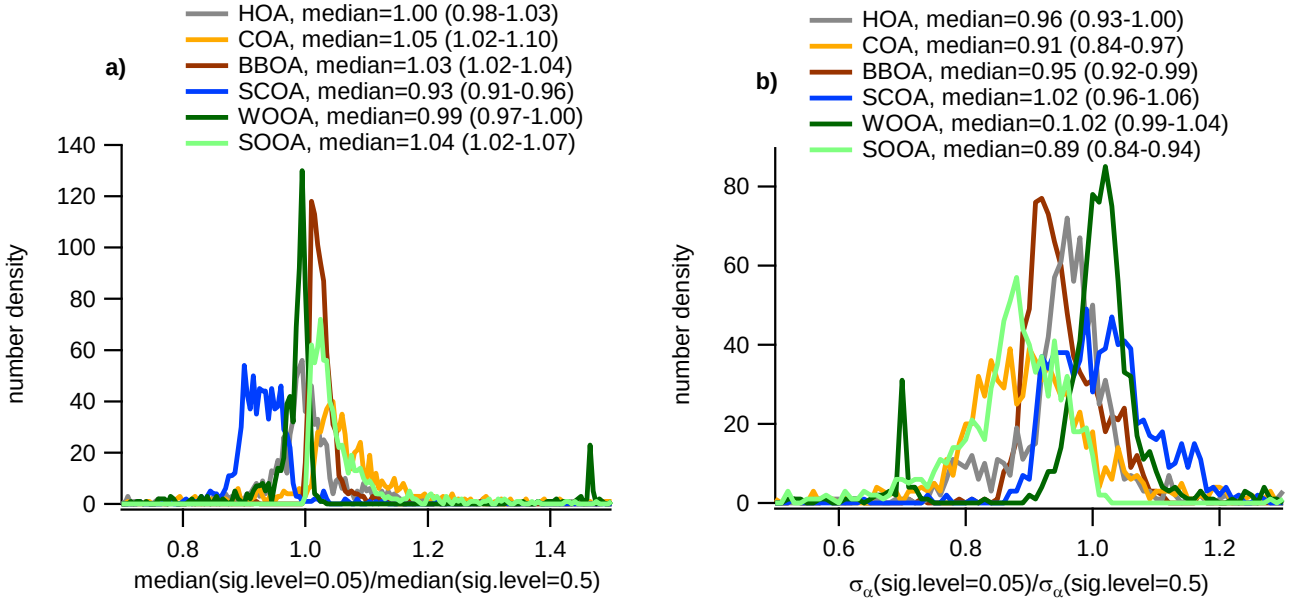

Figure S8: number density functions of source apportionment results obtained using a significance level of 0.05 normalized to results obtained using a significance level of 0.5: a) Comparison of factor concentrations b) Comparison of uncertainty estimate (σ_a).

References:

Cubison, M. J. and Jimenez, J. L.: Statistical precision of the intensities retrieved from constrained fitting of overlapping peaks in high-resolution mass spectra, Atmos. Meas. Tech., 8, 2333–2345, doi:10.5194/amt-8-2333-2015, 2015.

Corbin, J. C., Othman, A., Allan, J. D., Worsnop, D. R., Haskins, J. D., Sierau, B., Lohmann, U., and Mensah, A. A.: Peak-fitting and integration imprecision in the Aerodyne aerosol mass spectrometer: effects of mass accuracy on location-constrained fits, Atmos. Meas. Tech., 8, 4615-4636, doi:10.5194/amt-8-4615-2015, 2015.