Supplement of Atmos. Chem. Phys., 18, 2155–2174, 2018
https://doi.org/10.5194/acp-18-2155-2018-supplement

Atmospheric Chemistry and Physics
Open Access
EGU

Supplement of

Insights into organic-aerosol sources via a novel laser-desorption/ionization mass spectrometry technique applied to one year of PM_{10} samples from nine sites in central Europe

Kaspar R. Daellenbach et al.

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

Field blank analyses

Analyses of field blanks (also spiked with $AgNO_3$) exhibited low signal with 99% of peaks below detection limit (defined as $3 * \sigma_{abs}$, Fig S1).

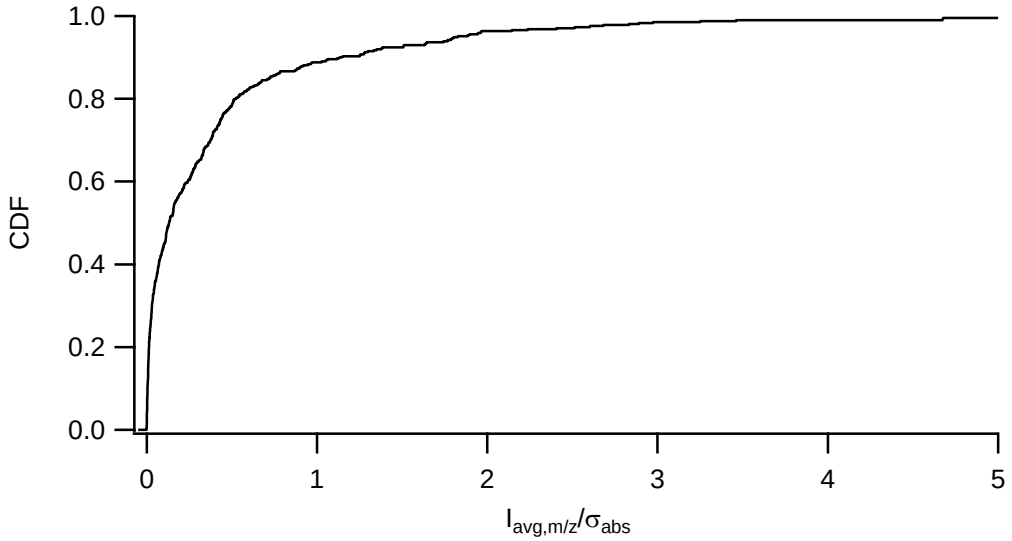

Figure S1: Cumulative density function (CDF) of the peak intensity in field blank analyses ($I_{avg,m/z}$) normalized to the minimal absolute error (σ_{abs}).

Mass spectra from field blanks exhibit low signals throughout the spectrum (Fig S1 and S2). 1% of all peaks (*m/z* 197, 249, 251, 322,324) show signals which are higher than the detection limit (Fig S2a, highlighted in black). These peaks are also subjected to a high variability between the repeated analyses (Figure S2b, highlighted in black).

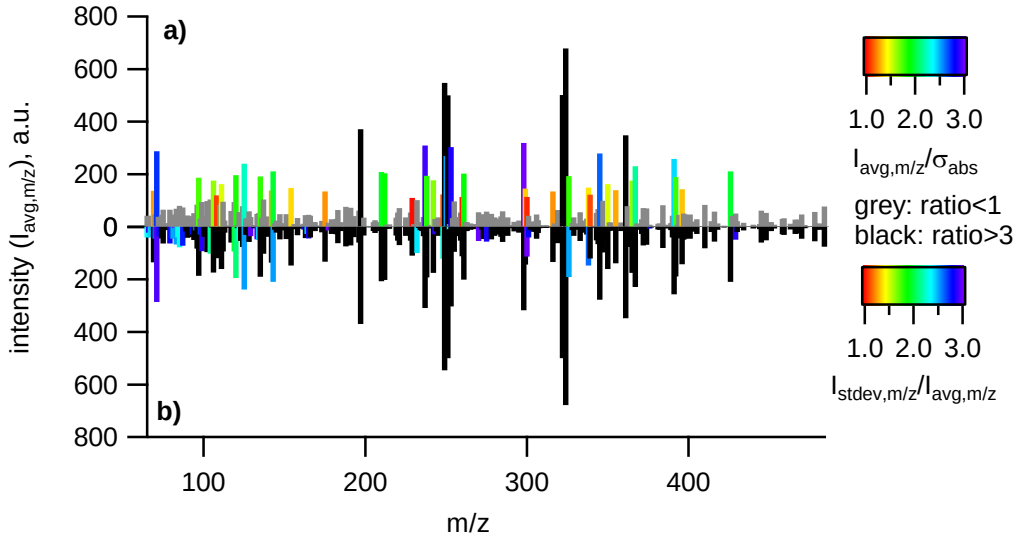

Figure S2: mass spectral signature of field blank analysis: a) colour-coded with the ratio of peak intensity ($I_{avg,m/z}$) divided by the minimal absolute error (σ_{abs}), b) colour-coded with the ratio of the measurement of the peak intensity ($I_{stdev,m/z}$) divided by $I_{avg,m/z}$.

Comparison of total LDI-MS signal and external measurements:

The total LDI-MS-intensity measured for the filter samples does not show a relation to the filter loading in terms of OC, sum of OC and EC, or PM_{10} (Fig S3).

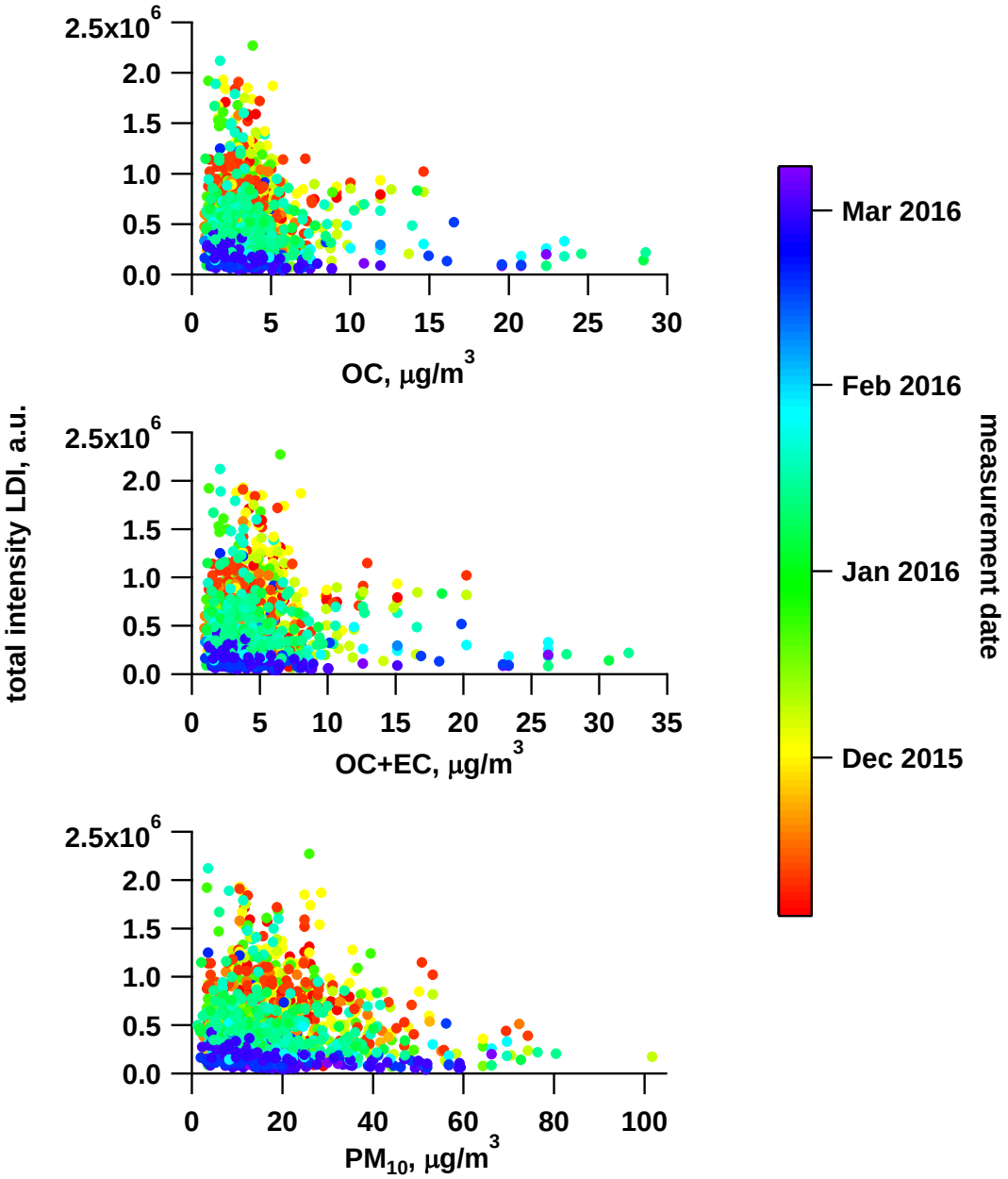

Figure S3: measured total intensity plotted against filter loadings in OC, OC+EC, and PM_{10} (color-coded with measurement date).

Influence of measurement time delays on total LDI-MS signal recorded:

In order to assess the repeatability of the total intensity, all repeats are normalized to the first measurement of the respective sample (Fig. S4). The intra-day variability (without the ones used to parametrize the error model) ranges from -27% (first quartile) to +33% (third quartile) of the average measurement. The inferred instrumental drift has an effect on the measured intensity (Kruskal-Wallis-test, p-value<0.05, Kendall-Tau-Test, p-value<0.05). Therefore, for quantification purposes we use external data like OA.

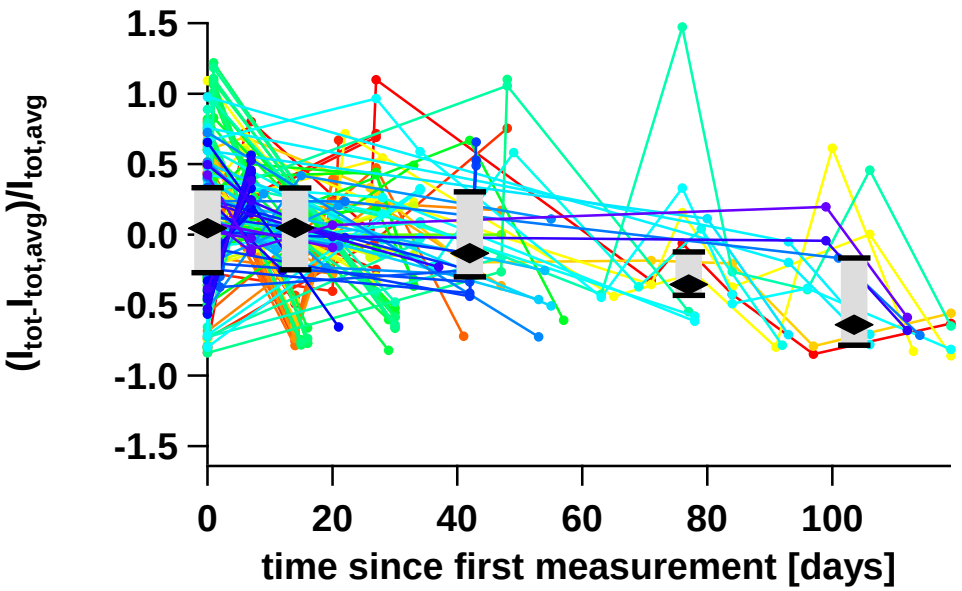

Figure S4: Difference of total measured signal intensity (I_{tot}) of samples repeated on different occasions to the average of I_{tot} normalized to the average of I_{tot} (lines with dots with varying colors for the different samples). The data is summarized using the median (black diamond) and quartiles (black horizontal lines), binned as intra-day repeats, repeats within 30, 30-60,60-90,90-120 days, respectively.

Comparison of selected *m/z*s and external measurements:

For visual comparison certain *m/z*s are plotted as a function of external measurements (Fig. S5).

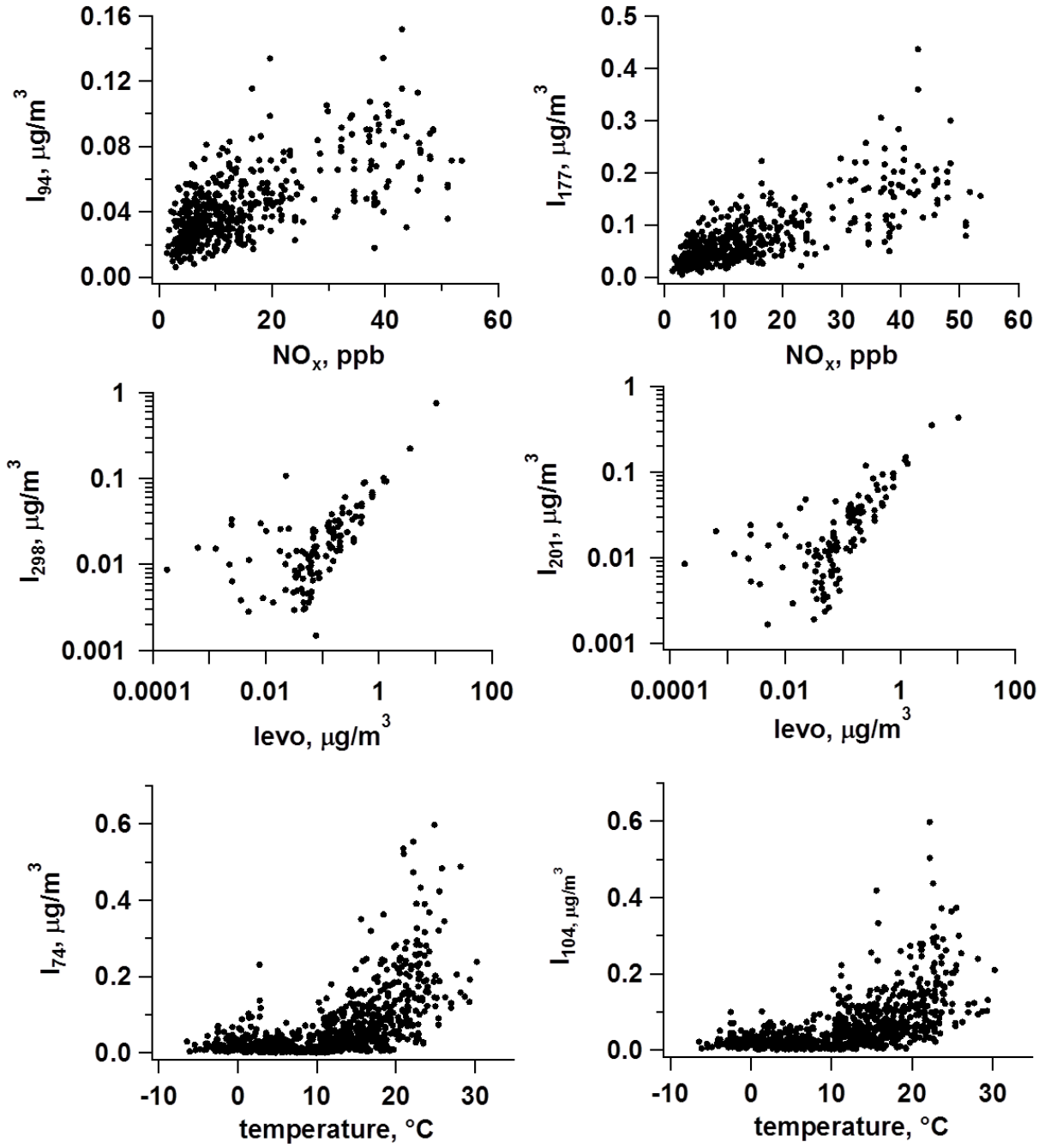

Figure S5: Scatterplots of selected *m/z*s (scaled to OA, µg/m^3) and NO_x (ppm), levoglucosan (µg/m^3), and temperature (°C).

Preliminary source apportionment analysis and factor identification:

A preliminary source apportionment analysis is performed by unconstrained PMF (no a priori information on source fingerprints used) in order to determine the number of factors for the source apportionment analysis. Based on Q_{avg}, defined as $\frac{\Sigma_{i=1}^{m}\Sigma_{j=1}^{n}\left(\frac{e_{i,j}}{\sigma_{PMF,i,j}}\right)^2}{n*m}$ (n: number of spectra, m:number of m/zs), the model explains the sites north and south of the alpine crest equally well when allowing for at least 4 factors (Fig. S6). Furthermore, we assess the change in time-dependent $Q_{avg,i}$, $\frac{\Sigma_{j=1}^{n}\left(\frac{e_{i,j}}{\sigma_{PMF,i,j}}\right)^2}{n}$, when increasing the number of factors, $\Delta Q_{avg,i}$ (Fig. S7). A significant $\Delta Q_{avg,i}$ signifies that structure in the residuals disappeared when adding an additional factor. Up to 5 factors, removed structure is evident but not when adding further factors. When increasing the number of factors up to 7, a factor with a similar signature as stable phase wood burning and a second traffic factor appear. A further increase leads to the separation of another possibly traffic related factor which contributes less than 5%. Therefore, we consider 7 factors to be optimal for this dataset.

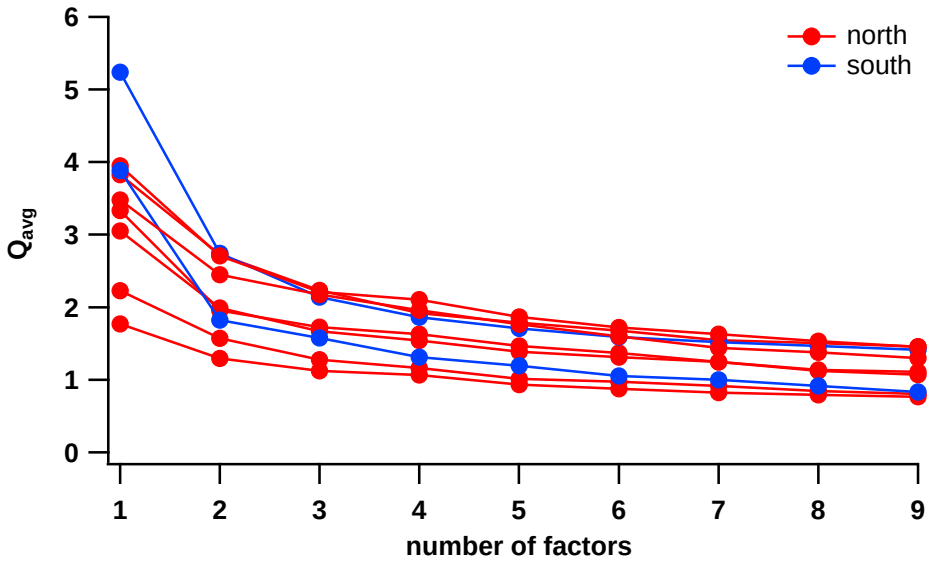

Figure S6: Q_{avg} as a function of the number of factors for the single sites (color-coded with geographical region).

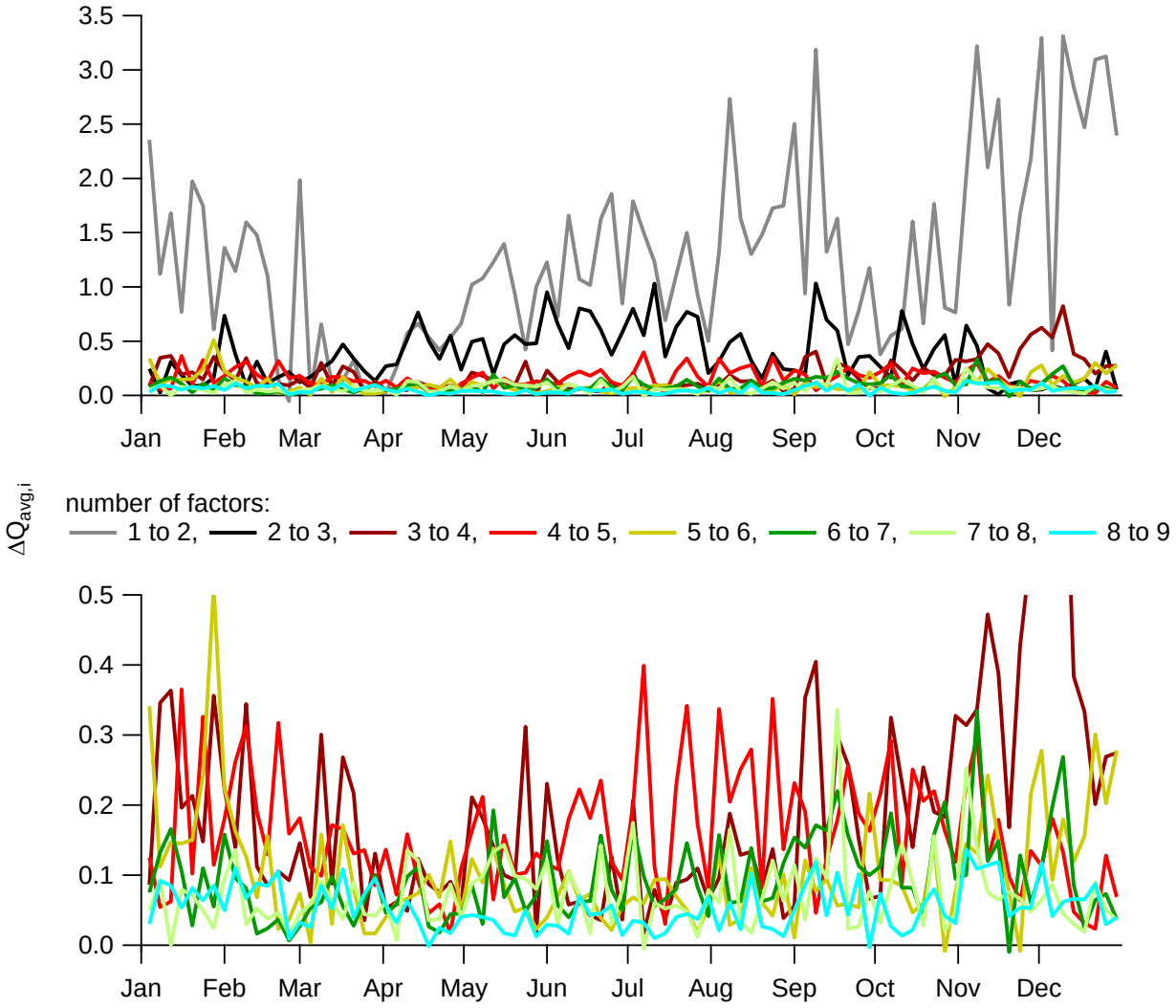

Figure S7: change in time-dependent contribution $Q_{avg,i}$ as a function of the number of factors, $\Delta Q_{avg,i}$.

For the 7 factor solution, we assess how well the different *m/z*s are explained by PMF using the quantity $Q_{avg,j}$, $\frac{\sum_{i=1}^{m}\left(\frac{e_{i,j}}{\sigma_{PMF,i,j}}\right)^2}{m}$ (Fig. S8). $Q_{avg,j}$ shows that even with 7 factors not all *m/z*s are explained within their measurement uncertainty. This might be linked to an underestimation of the measurement uncertainty itself. Additionally, factor profiles, which are assumed to be constant, might vary with changing seasons contributing to increased $Q_{avg,j}$ of certain *m/z*s. The *m/z*s 322, 324, 326 might be affected by the signal from silver-trimers, even though these peaks were removed prior to PMF.

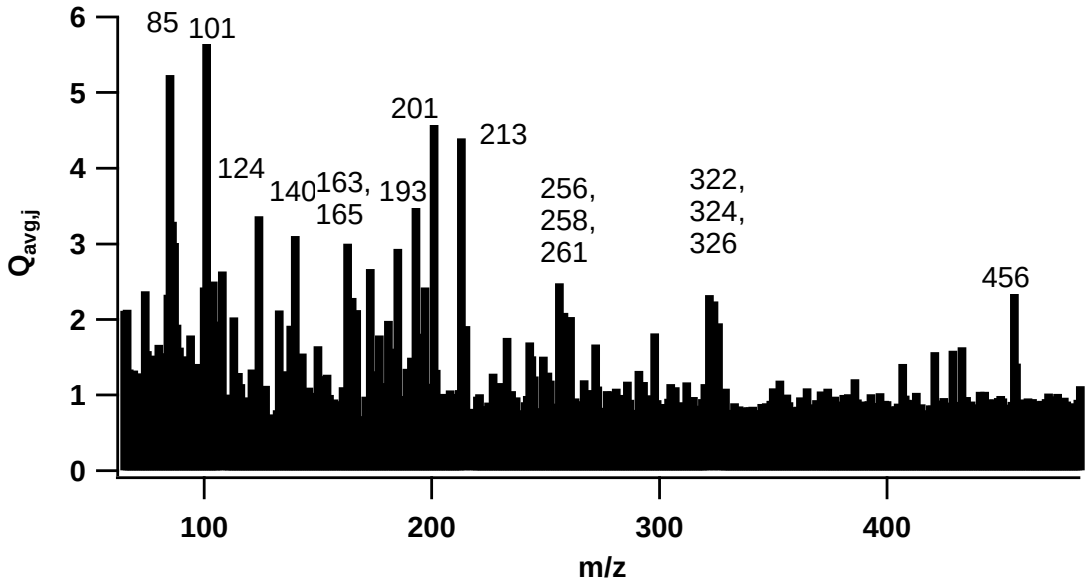

Figure S8: Q_{avg} **as a function of** *m/z* **when allowing for 7 factors.**

The explained variation (EV, Paatero et al., 2004, Canonaco et al., 2013) describes how much of the measured variation (time or variable) is explained by each PMF-factor (Eq. S1, for variable *j*):

$$EV_{j,k} = \frac{\sum_{i=1}^{n}(|g_{i,k}f_{k,j}|/\sigma'_{PMF,i,j})}{\sum_{i=1}^{n}\left((\sum_{k=1}^{p}|g_{i,k}f_{k,j}|+|e_{i,j}|)/\sigma'_{PMF,i,j}\right)} \quad \text{for k=1,…,p} \qquad \text{(S1)}$$

$f_{k,j}$ are the constant factor profiles and $g_{i,k}$ their time-dependent contributions. The index *i* represents a specific point in time (up to the number of points in time *n*, *j* the signal at a specific *m/z*, and *k* a factor (up to the number of factors p). $e_{i,j}$ are the PMF-residuals and $\sigma'_{PMF,i,j}$ the measurement uncertainties. Using the factor identifications based on the preliminary source apportionment analysis, the factors of the bootstrap runs are identified sequentially based on the explained variation as follows:

- traffic1: factor with maximal explained variation of *m/z* 177.
- traffic2: factor with maximal explained variation of *m/z* 163 of the remaining factors.
- BBeff: factor with maximal average explained variation of the *m/z*s 85, 124, 140, 197, and 213 of the remaining factors.
- BBineff: factor with maximal average explained variation of the *m/z*s 284 and 298 of the remaining factors.
- BBineff2: factor with maximal average explained variation of ions bigger than 300 a.m.u. of the remaining factors.
- bio-OA: factor with maximal average explained variation of *m/z* 74 and 104 of the remaining factors.
- LMW-OA: factor with maximal average explained variation of ions smaller than 150 a.m.u. of the remaining factors.

Bootstrap runs that showed mixing between factors described above were not further considered for the analysis. Criteria used for minimizing the effect of mixing based on the explained variation of groups of *m/z*s are lined out in the following:

1.	EV(traffic2,*m/z*s>300)	<	EV(BBineff 2,*m/z*s>300)
2.	EV(traffic2, *m/z*s 85, 124, 140, 197, 213)	<	EV(BBeff, *m/z*s 85, 124, 140, 197, 213)
3.	EV(traffic2, *m/z*s 284, 298)	<	EV(BBineff1, *m/z*s 284, 298)
4.	EV(LMW-OA, *m/z*s 284, 298)	<	EV(BBeff, *m/z*s 284, 298)

- Factor time series vs marker time series

Tab. S1 summarizes the correlation coefficients of all the factor time series and respective markers for all sites and as an average.

Table S1: correlation coefficients between factor and marker time series (R_p: pearson correlation coefficient, R_s: Spearman correlation coefficient).

Correlation			bas*	ber*	fra*	gal*	mag*	pay*	vad*	vi*	zue*	avg
Rp	traffic1	NO_x	-0.18	-0.31	-0.08	0.13	-0.23	-0.26	-0.1	-0.29	-0.19	-0.17
Rp	traffic2	NO_x	0.10	-0.06	0.29	0.35	0.54	0.5	0.24	-0.23	0.32	0.24
Rp	traffic1	eBC_{tr}					-0.31	0.15			-0.01	-0.08
Rp	traffic2	eBC_{tr}					0.73	0.68			0.63	0.69
Rp	BBeff	Levo	0.71	0.65	0.78	0.44	0.83	0.88	0.42	0.18	0.9	0.70
Rp	BBineff1	levo	0.88	0.11	0.64	0.91	0.59	0.93	0.93	0.98	0.58	0.83
Rp	BBineff2	levo	0.9	0.85	0.77	0.44	0.96	0.81	0.65	0.9	0.79	0.83
Rp	BBeff	K^+	0.29	0.61	0.78	0.32	0.67	0.1	0.35	-0.02	0.77	0.48
Rp	BBineff1	K^+	0.08	0.08	0.85	0.26	0.75	0.11	0.5	0.9	0.4	0.52
Rp	BBineff2	K^+	0.19	0.59	0.75	0.4	0.76	0.09	0.48	0.81	0.68	0.57
Rp	BBeff	BC_{wb}					0.76	0.42			0.75	0.67
Rp	BBineff1	BC_{wb}					0.91	0.49			0.55	0.71
Rp	BBineff2	BC_{wb}					0.69	0.4			0.82	0.67
Rs	bio-OA	temp	0.73	0.69	0.74	0.76	0.69	0.78	0.78	0.72	0.75	0.74
Rp	LMW-OA	NH_4^+	0.7	0.46	0.75	0.78	0.78	0.31	0.75	0.76	0.75	0.69

*ber: Bern, bas: Basel, fra: Frauenfeld, gal: St. Gallen, mag: Magadino, Pay: Payerne: vi: S. Vittore, zue, Zurich

Fig. S9 illustrates the relation between traffic1, traffic2 and NOx and eBCtr when only using the summer points.

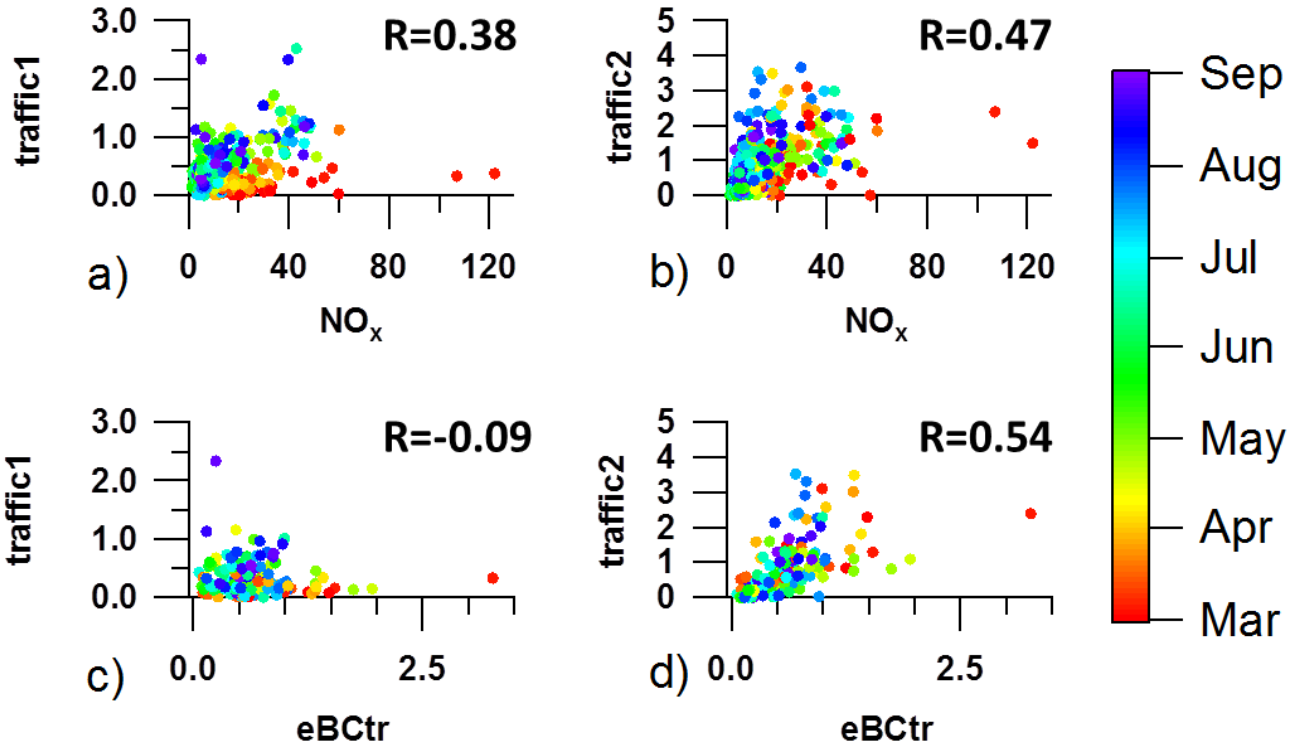

Figure S9: Scatterplots of selected traffic 1 / 2 (µg/m^3) and NOx (ppm), and temperature (°C). Correlation coefficients are computed on all summer points together.

Uncertainty estimation of PMF results:

The variability (standard deviation) of the apportioned factor concentration for the same measurement in the bootstrap runs is the base for the uncertainty estimate (σ_{bs}). Besides σ_{bs}, we also account for the effect of the repeatability on the source apportionment results (assessed by the standard deviation of the apportioned factor concentration, $\sigma_{intraday}$). For the latter purpose, 3 filters were measured on 3 instances throughout the measurement campaign (10 times during 1 day). For these filters, both types of uncertainty can be readily propagated to σ_{tot} (Eq. S2):

$$\sigma_{tot,i,k} = \sqrt{\sigma_{bs,i,k}^2 + \sigma_{intraday,i,k}^2} \quad \text{(S2)}$$

For the other filters, σ_{tot} cannot be estimated directly. Therefore, we parameterized $\sigma_{intraday}$ as a function of two error terms, one absolute and one relative, similarly to the PMF input uncertainty described in Section 2.3.2: (Eq. S3):

$$\sigma_{intraday,i,k} = \sqrt{e_{intraday,k}^2 + \left(r_{intraday,k} * conc_{i,k}\right)^2} \quad \text{(S3)}$$

Eq. S3 was fitted using the repeatability tests for the 3 filters, to obtain the parameters $e_{intraday}$ and $r_{intraday}$ for each of the PMF factors (Tab. S2) and the factor concentration (conc). These parameters were then extrapolated to the other filter samples, such that $\sigma_{i,k,tot}$ could be applied to all filter samples.

Table S2: error model coefficients for parameterized σ_{intrad}.

	$e_{intraday}$, µg/m^3	$r_{intraday}$, -
traffic1	0.01±0.01	0.08±0.02
traffic2	0.09±0.13	0.00±0.21
BBeff	0.01±0.0	0.10±0.02
BBineff1	0.04±0.04	0.13±0.06
BBineff2	0.11±0.05	0.08±0.07
LMW-OA	0.02±0.01	0.12±0.04
bio-OA	0±0.01	0.12±0.04

Influence of instrumental drifts on source apportionment results:

Fig S10 illustrates the difference in the apportioned factor concentrations as a function of the measurement time. All factors are affected by instrumental changes occurring during the measurement campaign (Kruskal-Wallis-test, p-value<0.05) and for all factors but traffic1 (p-value=0.58) and BBineff1 (p-value=0.23) a trend is identified consistent with an impact of an instrumental drift (ranked Mann-Kendall test, p-value<0.05). The intra-day variance explains largely the total variance, which is defined as the sum of intra-day (σ_{intrad}^2) and inter-day (σ_{inter}^2) variance (for traffic1 97%, traffic2 94%, BBeff 85%, BBineff1 89%, BBineff2 82%, LMW-OA 79%, bio-OA 97%).

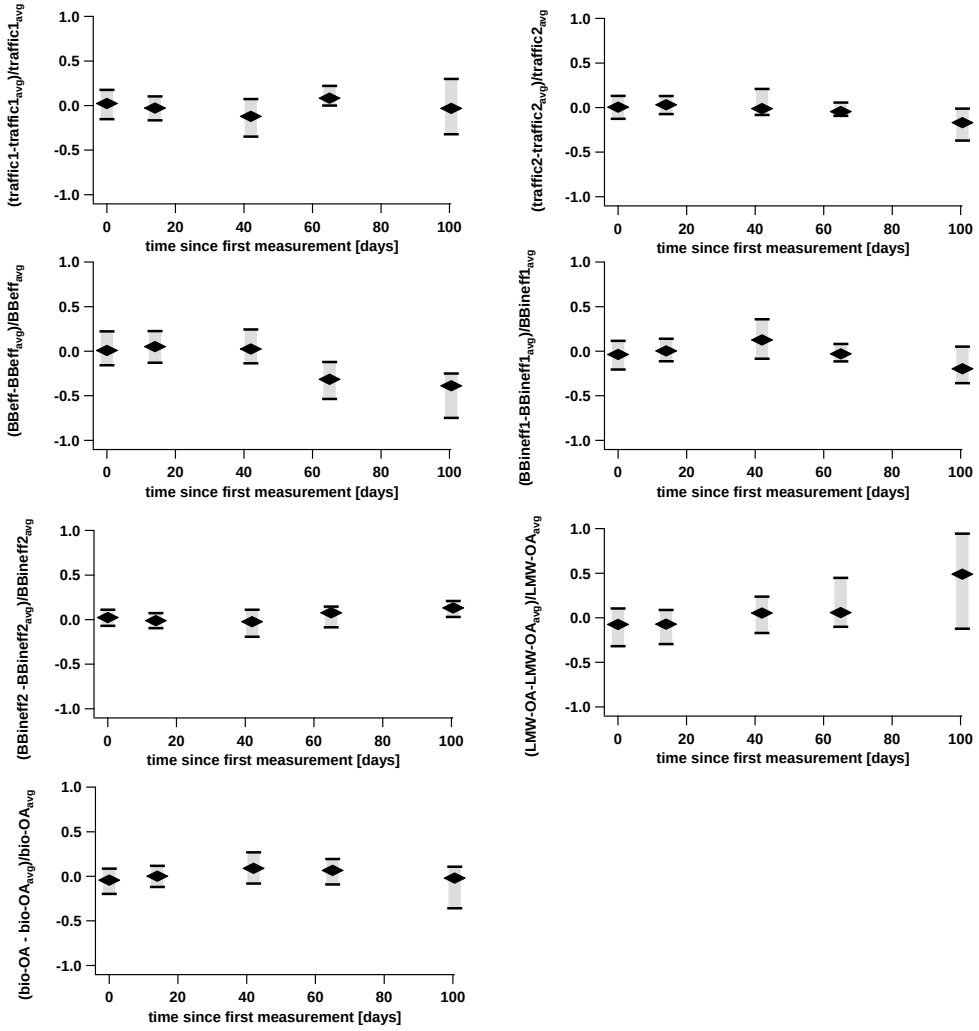

Figure S10: Difference in factor concentration of samples measured on different instances to the average factor concentration normalized to the average concentration as a function of measurement time delay. The data is summarized using the median (diamond) and quartiles (horizontal lines) of the samples, binned as intra-day repeats, repeats within 30, 30-60,60-90,90-120 days, respectively (195, 240, 46, 53, 28 points per bin, respectively).

Relative yearly average contributions of PMF factors for all nine sites:

Fig. S11 presents the relative yearly average contributions of the PMF factors for all nine sites as presented in Tab. 1 in the manuscript.

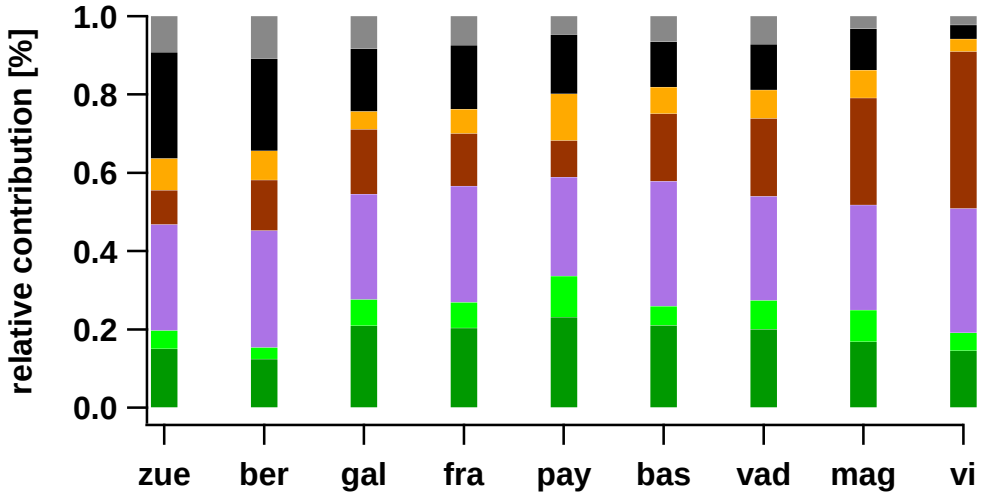

Figure S11: relative yearly average factor contributions for the different sites (grey: traffic1, black: traffic2, orange: BBeff, brown: BBineff1, violet: BBineff2, light green: bio-OA, dark green: LMW-OA.

References:

Paatero, P.: User's guide for positive matrix factorization programs PMF2 and PMF3, 2004.

Canonaco, F., Crippa, M., Slowik, J. G., Baltensperger, U., and Prévôt, 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.