## Supplement of

# Long-term chemical analysis and organic aerosol source apportionment at 9 sites in Central Europe: Source identification and uncertainty assessment

5

## 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<sub>block</sub>, we evaluate the influence of the number of factors, p, on the  $Q_i/Q_{i,exp}$  contribution. 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.

5 Fig. SI.1 shows the  $Q_i/Q_{i,exp}$  contribution 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  $Q/Q_{exp}$  contributions.  $\Delta(Q_i/Q_{i,exp})$  shows the difference in  $Q_i/Q_{i,exp}$  between groups of points like sites or season. The smaller the  $\Delta(Q_i/Q_{i,exp})$ , the smaller are the differences in the mathematical quality of the PMF solution for the



10 different seasons/sites. To explain the temporal and geographical variability at least 5 factors are required.

Figure SI.1:  $Q_i/Q_{i,exp}$  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).  $\Delta Q_i/Q_{i,exp}$  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 median  $Q_i/Q_{i,exp}$  and the lowest.

• Quality assessment of solutions:

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

Table 51.1. Set of criteria used to diseard bad solutions
---

criteria on profile					
	$f(\mathrm{CO}_2^+)$	$f(C_2H_4C)$	<b>D</b> <sub>2</sub> <sup>+</sup> )		
НОА	<0.4	< 0.004			
COA	<0.4	<0.01			
Criteria on time series					
НОА	r(HOA,NOx)>0 & r(HOA,NOx)>r(COA,NOx)				
BBOA	r(BBOA,levo)>0				
SC-OA	r(SC-OA,CH <sub>3</sub> SO <sub>2</sub> <sup>+</sup> )>0				
Mass closure criteria					
OC <sub>res</sub> total		$Q_{25}(OC_{res}) < 0 \& Q_{75}(OC_{res}) > 0$			
Magadino winter, Magadino summer		$Q_{25}(OC_{res}) < 0 \& Q_{75}(OC_{res}) > 0$			
Zurich winter, Zurich summer			$Q_{25}(OC_{res}) < 0 \& Q_{75}(OC_{res}) > 0$		
Magadino, Zurich			$Q_{25}(OC_{res}) < 0 \& Q_{75}(OC_{res}) > 0$		
HOC <median, hoc="">median</median,>		$Q_{25}(OC_{res}) < 0 \& Q_{75}(OC_{res}) > 0$			
COC <median, coc="">median</median,>		$Q_{25}(OC_{res}) < 0 \& Q_{75}(OC_{res}) > 0$			
BBOC <median, bboc="">median</median,>		$Q_{25}(OC_{res}) < 0 \& Q_{75}(OC_{res}) > 0$			
SC-OC <median, sc-oc="">median</median,>		$Q_{25}(OC_{res}) < 0 \& Q_{75}(OC_{res}) > 0$			
WOOC <median, wooc="">median</median,>		$Q_{25}(OC_{res}) < 0 \& Q_{75}(OC_{res}) > 0$			
SOOC <median, sooc="">median</median,>		$Q_{25}(OC_{res}) < 0 \& Q_{75}(OC_{res}) > 0$			
for PMF with 12 filters per site summer Magadino and			$Q_{25}(OC_{res}) < 0 \& Q_{75}(OC_{res}) > 0$		
Zurich					
for PMF with 12 filters per site winter Magadino and			$Q_{25}(OC_{res}) < 0 \& Q_{75}(OC_{res}) > 0$		
Zurich					

5

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

Figure SI.2 demonstrated the high similarity between the retrieved BBOA signature and the mass spectrum of nebulized levoglucosan.



10 Figure SI.2: mass spectral fingerprints of BBOA (PMF<sub>block</sub>) and nebulized levoglucosan.

#### • <u>Uncertainty estimation and propagation:</u>

10

The uncertainty described by the interquartile range from the *a*-value sensitivity assessment ( $\sigma_a$ ) 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

5 uncertainty is estimated by propagating the variability between the median concentrations for the 4 sensitivity tests ( $\sigma_b$ ) and half the interquartile range of PMF<sub>block</sub> ( $\sigma_a$ , Eq. SI.1):

$$err_{i,k,tot} = \sqrt{\sigma_a^2 + \sigma_b^2} = \sqrt{\left(\frac{interquartile\ ragne\ PMF\ block}{2}\right)_{i,k}^2 + \left(\sigma(median\ bootstrap\ solutions)\right)_{i,k}^2}$$
SI.1

In absence of  $\sigma_b$  for all other points, we parametrize  $\sigma_b$ . We express  $\sigma_b$  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. SI.2):

$$\frac{\sigma_b}{|conc_{i,k}|} = \frac{\sqrt{\sigma_{minimal}^2 + k^2 * conc_{i,k}^2}}{|conc_{i,k}|}$$
SI.2

The uncertainty ( $\sigma_a$  and  $\sigma_b$ ) is propagated for all points using the parameters from Eq. SI.2 in order to obtain the total uncertainty for all points in the dataset (Eq. SI.3):

$$err'_{i,k,tot} = \sqrt{\sigma_a^2 + \left(\sigma_{minimal}^2 + k^2 * conc_{i,k}^2\right)}$$
SI.3

15 The resulting coefficients of the error model are presented in table SI.2:

factor	$\sigma_{minimal}$	k
НОА	0.14±0.02	$0.40\pm0.47$
COA	0.04±0.01	0.76±0.15
BBOA	0.05±0.01	0.50±0.07
SC-OA	0.26±0.01	0.26±0.38
WOOA	0.04±0.02	0.76±0.30
SOOA	0.03±0.01	$0.27 \pm 0.05$

### Table SI.2: $\sigma_{minimal}$ and k for the different factors including their uncertainty.

References:

Bozzetti, C., Daellenbach, K., R., Hueglin, C., Fermo, P., Sciare, J., Kasper-Giebl, A., Mazar, Y., Abbaszade, G., El Kazzi, M., Gonzalez, R., Shuster Meiseles, T., Flasch, M., Wolf, R., Křepelová, A., Canonaco, F., Schnelle-Kreis, J., Slowik, J. G.,

5 Zimmermann, R., Rudich, Y., Baltensperger, U., El Haddad, I., and Prévôt, A. S. H.: Size-resolved identification, characterization, and quantification of primary biological organic aerosol at a European rural site, Environ. Sci. Technol., 50, 3425-3434, doi:10.1021/acs.est.5b05960, 2016.