

We thank the referees for their comments, which helped improving the quality of our manuscript. A point by point response (in blue) to the reviewers' comments (in black, italics) will follow. Changes in the text are indicated in in black.

Anonymous Referee #1

The manuscript presents the results obtained from the off-line analysis of filter samples collected at 9 sites in central Europe with different exposure characteristics. The study is mainly focused on the identification of the different sources that contribute to the organic aerosol loadings using PMF analysis. The obtained results indicate that biomass burning is a major contributor to primary organic aerosol with 88% in the alpine valley and 43% north of the alpine crest. On the contrary, the sum of HOA, COA and BBOA contributes less at the sites north of the alpine crest than at the southern alpine valley. Secondary organic aerosol production is enhanced during summer due to the increased biogenic emissions with temperature. Finally, it is estimated that primary biological particles which cannot be resolved by PMF could contribute significantly to PM10 organic aerosol.

The manuscript is very well written, coherent and easy to follow. A lot of effort has been put into the selection of the PMF solutions and the subsequent sensitivity analysis in order to provide the most sane and justified selection of factors, given the statistical nature of the analysis. This paper can be recommended for publication after some minor corrections listed below:

1) As the title reads "Long-term chemical analysis and organic aerosol source apportionment.." I would expect a short description of the trends in the chemical species as well of the 9 sites. It would be helpful as in a subsequent section the correlation with ammonium and nitrates is mentioned.

We have mentioned in the manuscript (page 3, lines 11-14) that two papers regarding the offline analysis of this dataset are planned. This section has been modified as follows:

"... This paper focuses on the identification of the main factors influencing the OA concentrations at the different sites and the assessment of the associated uncertainties. In a second paper, we will investigate the site-to-site differences and general trends in the factor time series and their relationship with external parameters. ..."

In the corrected version of the manuscript we have added a discussion on the fraction of ammonium that can be attributed to nitrate and to sulfate, for different seasons. This section reads as follows (P13 L30-33):

"... Here, we have used ammonium as a proxy for aged aerosols affected by anthropogenic emissions, as WOOA correlates better with ammonium than with nitrate sulfate. We note that in winter, when WOOA is highest, 56% of ammonium can be attributed to nitrate, whereas in summer ammonium sulfate dominates (97% of ammonium can be attributed to sulfate). Therefore, WOOA correlates more with nitrate ($R^2 = 0.64$) than sulfate ($R^2 = 0.48$). ..."

2) The possibility that WOOA could partially originate from the oxidation of BBOA could also be mentioned. During BB events ammonium is in a significant excess compared to sulfate, and this could possibly explain the good correlation between WOOA and NH_4^+ .

As we have mentioned above, the excess of ammonium compared to sulfate in winter is attributed to ammonium nitrate, which is thermodynamically more stable under lower temperatures and

higher relative humidity. As both WOOA and nitrate originate from a similar process – i.e. oxidation of precursors during winter time – both species do correlate. However, we do not think that this correlation is sufficient evidence to support that WOOA could partially originate from the oxidation of BBOA. We indeed think that this is the case based on modelling results we have recently published (Ciarelli et al., 2017) and on the chemical analysis of the same samples on a molecular level, using ultra-high mass resolution spectrometric techniques. The latter results will be presented in an upcoming publication.

Minor corrections:

Abstract L21: ... smaller than 10 μm from 9 stations ...

The text has been adapted to ... smaller than 10 μm at 9 stations...

“... We present offline-AMS measurements for particulate matter smaller than 10 μm at 9 stations in central Europe with different exposure characteristics for the entire year of 2013 (819 samples). ...”

P16, L26: ... closely correlates with NH_4^+ .

The text has been corrected accordingly.

“...WOOA, dominant SOA category during winter, closely correlates with NH_4^+”

References:

Allan, J. D., Jimenez, J. L., Williams, P. I., Alfarra, M. R., Bower, K. N., Jayne, J. T., Coe, H., and Worsnop, D. R.: Quantitative sampling using an Aerodyne aerosol mass spectrometer 1. Techniques of data interpretation and error analysis, *J. Geophys. Res.*, 108, 4090, doi:10.1029/2002JD002358, 2003.

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., 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.

Bozzetti, C., Sosedova, Y., Xiao, M., Daellenbach, K. R., Ulevicius, V., Dudoitis, V., Mordas, G., Byčenkienė, S., Plauškaitė, K., Vlachou, A., Golly, B., Chazeau, B., Besombes, J.-L., Baltensperger, U., Jaffrezo, J.-L., Slowik, J. G., El Haddad, I., and Prévôt, A. S. H.: Argon offline-AMS source apportionment of organic aerosol over yearly cycles for an urban, rural and marine site in Northern Europe, *Atmos. Chem. Phys. Discuss.*, doi:10.5194/acp-2016-413, 2017a.

Bozzetti, C., El Haddad, I., Salameh, D., Daellenbach, K. R., Fermo, P., Gonzalez, R., Minguillón, M. C., Iinuma, Y., Poulain, L., Müller, E., Slowik, J. G., Jaffrezo, J.-L., Baltensperger, U., Marchand, N., and Prévôt, A. S. H.: Organic aerosol source apportionment by offline-AMS over a full year in Marseille, *Atmos. Chem. Phys.*, 17, 8247-8268, <https://doi.org/10.5194/acp-17-8247-2017>, 2017b.

Chen, Q., Miyazaki, Y., Kawamura, K., Matsumoto, K., Coburn, S. C., Volkamer, R., Iwamoto, Y., Kagami, S., Deng, Y., Ogawa, S., Ramasamy, S., Kato, S., Ida, A., Kajii, Y., and Mochida, M.: Characterization of chromophoric water-soluble organic matter in urban, forest, and marine aerosols by HR-ToF-AMS analysis and excitation emission matrix spectroscopy, *Environ. Sci. Technol.*, 50, 10,351–10,360, 2016.

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.

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.

Sun, Y., Zhang, Q., Zheng, M., Ding, X., Edgerton, E. S., and Wang, X.: Characterization and source apportionment of water-soluble organic matter in atmospheric fine particles (PM_{2.5}) with high-resolution aerosol mass spectrometry and GC–MS, *Environ. Sci. Technol.*, 45, 4854–4861, 2011.

We thank the referees for their comments, which helped improving the quality of our manuscript. A point by point response (in blue) to the reviewers' comments (in black, italics) will follow. Changes in the text are indicated in in black.

P. Paatero (Referee)

This work develops first a mathematical machinery whereby they attempt to formulate a small number of factor spectra (2 assumed "known" mass spectra plus 4 fitted mass spectra) whereby a large number (819) of mass spectra, measured in 9 locations in different times of the year, could be successfully fitted. In the second part, they perform a very large number of repetitions of the modeling, so that different details of the model are varied. From variation of the results, they deduce reliability estimates for the obtained spectral profiles and contribution time series.

The mathematical operations are not adequately explained. It is impossible to know what they have done (and why) in different stages of the work. One of their key concepts is a constraining of F factor elements. Unfortunately, the defining equation (3) is so unclear (and possibly contains a typo) that I cannot even guess what they might mean by this equation. This review is limited to the first part of the work, modeling the measured mass spectra by a bilinear model. Although the second part is important, and has required a large amount of work, it cannot be analyzed in the time that would be available for such analysis. Thus I do not comment on the second part. I recommend that this manuscript should be published after the listed problems in mathematical presentation and elsewhere have been corrected.

Problems in the mathematical presentation

The fundamental principle of science is repeatability. Used methods should be defined clearly and in sufficient detail so that colleagues will be able (at least in principle) to repeat what was done by the authors. In the following, I discuss mathematics that have not been described in an understandable manner. In general, equations are the language of mathematics. Mathematical work should be described using equations. Verbal explanations may only help in understanding the equations, they cannot replace equations. In order to use equations, it is necessary to define symbols for various quantities. If necessary, use two-letter symbols.

Consider typographic questions: The prime should be avoided. Complicated notation in subscripts or in superscripts is often very difficult to read.

We have taken into account the reviewer comments, clarified the mathematical operations used and their purpose and added/modified mathematical equations when needed. In the following we respond to the different points raised by the reviewer and indicate the changes in the text we have made.

P5 L9-11 (page 5, lines 9-11)

.. by the estimated organic matter (OM) concentration, calculated as the product of the OC concentrations .. and the OM/OC ratios from the AMS measurements.

Reading this verbal explanation, it is very hard to understand what was done. In fact, I would need to try to write the equations in order to understand. Please do this equation writing for the benefit of your readers!

We have added an additional equation (now equation 3), to clarify the quantification operation. The text reads as follows:

“... ”

Input data and error matrices consisted of 202 organic ions. The organic fragments, $x'_{i,j}$, obtained from offline AMS analyses do not directly represent ambient concentrations. Therefore, the signal of each fragment was converted to such an ambient concentration ($x_{i,j}$ in $\mu\text{g m}^{-3}$), by multiplying the fraction of this signal with the estimated organic matter (OM) concentration. The latter was calculated as the product of the OC concentrations measured by the Sunset OC/EC analyzer and the OM/OC ratios from the offline AMS measurements $(\text{OM/OC})_{\text{oAMS}}$ (Eq. 3). Note that such scaling does not change the outcome of Eq. 2 since both data and error matrices are scaled in the same manner and the fingerprints ($f_{k,j}$) are not changed.

$$x_{i,j} = \frac{x'_{i,j}}{\sum_i x'_{i,j}} * OC * (\text{OM/OC})_{\text{oAMS}} \quad (3)$$

“... ”

P5 EQ(3): I do not understand eq(3) at all.

- what are f_{kn} and f'_{kn} (or is it f_{kn} '), what difference is there between them?

- is there a typo in the equation? As shown, the equation appears impossible.

This is a key detail in the manuscript as it describes constraints that perhaps have never been published in similar work. It must be described so that it is understandable.

We do agree with the reviewer that the equation is misleading. The constraints were performed in the same way as in Canonaco et al. (2013) and the equation in the text was adapted to the one in the mentioned publication. The modified text reads as follows:

“... The PMF algorithm was solved using the multilinear engine-2 (ME-2, Paatero, 1999). Normalization of the PMF solution during the iterative minimization process is disabled as implemented in SoFi (Canonaco et al., 2013). ME-2 enables an efficient exploration of the solution space by a priori constraining the $f_{k,j}$ elements within a certain range defined by the scalar a ($0 \leq a \leq 1$) from a starting value $f'_{k,j}$, such that the modelled $f_{k,j}$ in the solution satisfy Eq. 4:

$$f_{k,j} = f'_{k,j} + a * f'_{k,j} \quad (4)$$

$f'_{k,j}$ is the starting value used as a priori knowledge from previous studies and $f_{k,j}$ is the resulting value in the solution. ...”

What was included in matrix X, to be modeled by PMF. It remains unclear what information was contained in X. Section 2 lists a number of variables whose concentrations were measured. Were all these included in X, or were some of these included, or none of them?

P5 L8-9 say: *Input data and error matrices included 202 organic ions. Do you mean that input data consisted of 202 organic ions? The formulation you used may also be interpreted so that among other data, input data also included 202 organic ions.*

This sentence was indeed misleading. **X** consisted of 202 ions (variables) measured over n time points. The text has been adapted accordingly:

“... Input data and error matrices consisted of 202 organic ions.”

Also regarding matrix X:

Input data ... were rescaled by the estimated organic matter (OM) concentration ... In PMF, one is allowed to rescale data rows in any way, provided that error rows are also scaled in the same manner. Thus your scaling is OK. However, it would help the reader a lot if you state briefly why to scale, what advantage was achieved by scaling. Your scaling (which does not change profile spectra in any way) is useful for plotting figure 9. On the other hand, your scaling influences (improves or worsens) correlations between factor values and marker concentrations. Please explain and/or correct. Note that this scaling does not change the computations performed according to Eq(4) in any way.

The scaling procedure does not influence in any way the PMF calculations. However, as the offline AMS signals do not directly represent ambient concentrations the factors retrieved should be scaled to the real ambient concentrations, before any correlation with external markers and data interpretation is made. Thus the scaling could indirectly affect the result by influencing the solution selection. The scaling could be achieved before or after running the PMF (as it was mentioned correctly by the reviewer the scaling does not change the computations performed according to Eq. 4). We have scaled the data before running PMF for computation reasons as in this way we do not need to scale the results obtained from every PMF solution repeatedly. The text related to this section has been adapted as follows:

“... ”

Input data and error matrices consisted of 202 organic ions. The organic fragments, $x'_{i,j}$, obtained from offline AMS analyses do not directly represent ambient concentrations. Therefore, the signal of each fragment was converted to such an ambient concentration ($x_{i,j}$ in $\mu\text{g m}^{-3}$), by multiplying the fraction of this signal with the estimated organic matter (OM) concentration. The latter was calculated as the product of the OC concentrations measured by the Sunset OC/EC analyzer and the OM/OC ratios from the offline AMS measurements $(OM/OC)_{oAMS}$ (Eq. 3). Note that such scaling does not change the outcome of Eq. 2 since both data and error matrices are scaled in the same manner and the fingerprints ($f_{k,j}$) are not changed.

$$x_{i,j} = \frac{x'_{i,j}}{\sum_i x'_{i,j}} * OC * (OM/OC)_{oAMS} \quad (3)$$

“... ”

P5 L5-7 (input errors):

> *The input errors ... include ... the uncertainty related to ion counting statistics and ion-to-ion signal variability at the detector.*

*I understand that you counted the ions. Then, ion-to-ion signal variability is *not* a source of uncertainty. If ion current is measured, instead of counting ions, then ion signal variability *is* a source of error. Please correct or clarify.*

The AMS does not count ions, but rather an integrated signal (bits) that can be related to ion counts by means of a single ion calibration. Thus single ion signal variability is a source of uncertainty for PMF analysis, in addition to counting statistics. Both sources of errors are included in Allan et al. (2003, alpha and I, respectively).

Allan, J. D., Jimenez, J. L., Williams, P. I., Alfarra, M. R., Bower, K. N., Jayne, J. T., Coe, H., and Worsnop, D. R.: Quantitative sampling using an Aerodyne aerosol mass spectrometer 1. Techniques of data interpretation and error analysis, *J. Geophys. Res.*, 108, 4090, doi:10.1029/2002JD002358, 2003.

P5 L20-24

describe a complicated method for filling knowledge gaps in the known/assumed fixed factor spectra ("reference profiles") of HOA and COA. I am not fully convinced about the performance of this method. The "natural" alternative method is to leave the unknown elements in reference profiles (profiles of HOA and COA) as ordinary factor elements, to be fitted by ME-2 together with all "normal" non-fixed F factor elements, as explained in detail below. Please include this remark in the corrected ms, so that future colleagues are encouraged to follow the safer and simpler method instead of your complicated method.

We added the following statement to the manuscript, based on the reviewer's suggestions:

"... Alternatively, such missing ions can be also treated as ordinary factor elements, to be fitted by ME-2 with all other ordinary factor elements. ..."

Using "constrained factors" based on known profiles of HOA and COA. This topic was very difficult to understand at first. It was not clear what is "constrained" by what. Now I assume that you mean the following: In all PMF runs, two constant F factors were used, i.e. two rows of factor F were defined as a-priori fixed, so that the values of these constant factors were set equal to previously known mass spectra of HOA and COA. Is this what you mean? – Using constant or constrained factors is not familiar to PMF users, not at all. Such unusual methodology must be carefully explained so that all readers have the possibility to understand what you have done. In particular, you should explain that using fixed factors is not the same as using inequalities in order to constrain factors to lie between upper and lower limits, set very close to each other. Also, you should go into technical details here, because it is possible to implement constant factors in two different ways in ME-2. You should guide your readers to the optimal usage. It is possible to use "constant factors" that reside in a different matrix, which is clumsy. The alternative is to keep all F factors in the same matrix but define that the elements in two first (say) rows of F are "locked", not allowed to change during the fitting process. These elements are set equal to the known profile values before initiating the fit. If there are gaps in the knowledge of HOA and COA, then those unknown elements in "locked" F rows should simply not be locked at all, so that they may obtain their best possible values during the fit.

For a better readability we adapted the text to:

“... $f_{k,j}'$ is the starting value used as a priori knowledge from previous studies and $f_{k,j}$ is the resulting value in the solution. In all PMF runs (unless mentioned otherwise), we used the high resolution mass spectra for HOA and COA (cooking OA) from Crippa et al. (2013b) as constraints, i.e. two rows of $f_{k,j}$ were set equal to the mass spectra of HOA and COA. Ions that were present in our datasets but not in the reference profiles for HOA and COA were inferred from published unit mass resolution (UMR) profiles (Ng et al., 2011 and Crippa et al., 2013c). For this purpose, the fraction of signal at a specific m/z in the UMR reference spectrum ($f_{UMR,m/z}$) was compared to the fraction of signal of all ions at this m/z in the HR reference spectrum ($f_{HR,m/z}$). The difference $f_{UMR,m/z} - f_{HR,m/z}$ was used as entries in $f_{k,j}'$ for such missing ions. For these ions, an α -value of unity was set. For the other factors, the factor elements were fitted by ME-2. Alternatively, such missing ions can be also treated as ordinary factor elements, to be fitted by ME-2 with all other ordinary factor elements.

...”

Use of constant factors or constrained factors often causes so-called "normalization conflicts". How did you protect your bilinear model against normalization conflicts? This is another important detail that should be communicated to colleagues who might follow your example.

We believe that the reviewer comment on normalization conflicts is referring to the normalization equations for the factor contribution present in the default ME-2 instruction file. If so, when calling ME-2 from SoFi, these normalization equations are disabled.

We added this technical aspect to the manuscript (P5 L18-23):

“... The Source Finder toolkit (SoFi v.4.9, Canonaco et al., 2013) for Igor Pro software package (Wavemetrics, Inc., Portland, OR, USA) was used to configure the PMF model and for post- analysis. The PMF algorithm was solved using the multilinear engine-2 (ME-2, Paatero, 1999). Normalization of the PMF solution during the iterative minimization process is disabled as implemented in SoFi (Canonaco et al., 2013). ME-2 enables an efficient exploration of the solution space by a priori constraining the $f_{k,j}$ elements within a certain range defined by the scalar α ($0 \leq \alpha \leq 1$) from a starting value $f_{k,j}'$, such that the modelled $f_{k,j}$ in the solution satisfy Eq. 4: ...”

P7 L16-17

> For each of the four PMF datasets, 2420 PMF runs were performed for
> evaluating the sensitivity of the model to the chosen α -value and the seed.

This statement mentions sensitivity of the model to random seed. The random seed determines the pseudorandom initial values of PMF fit. In plain language, this statement says that there were local solutions so that depending on seed, PMF iteration converged to different local solutions. Presumably, these solutions had comparable Q values because otherwise, Q values would be used for selecting between solutions. Now these different solutions are somehow pooled together and their presence is otherwise ignored.

The presence of multiple solutions should be properly reported (e.g. how many of PMF runs had multiple solutions, how many different individual solutions per PMF run were obtained at most and on the average, are the solutions rotationally equivalent having identical residuals of fit, etc.) There are no fixed rules on what to do with multiple solutions. On one extreme, it has been suggested that scientists may at will pick the one solution they like most and ignore the others. At the other extreme,

PMF modeling of such data may be considered failing if there are several local solutions with comparable Q values.

For each of the four PMF datasets, 2420 PMF runs were performed for evaluating the sensitivity of the model to the chosen α -value and the seed. The quality of each of the 2420 PMF runs was individually assessed using the criteria lined out in Sec. 3.6. We assessed the environmental interpretability based on the correlation of factor time series and markers of the respective source. All solutions fulfilling these criteria are accepted and considered plausible. Using these criteria 331 PMF runs were selected for PMF_{block} (for PMF_{zue,isol} 230, for PMF_{zue,reprs} 99, and for PMF_{1filter/month} 269). This information was added to the text. We found that 80% of the accepted solutions have an α -value \leq 0.3 for HOA and an α -value \leq 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 now 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. This information was added to the supplementary. In a last step, all accepted solutions are pooled together. In Sec. 3.3, we added a remark linking to Sec. 3.6:

The modified Sec. 3.6 reads:

“... Monte Carlo simulations were performed and simulations for which res-OC_i distributions were significantly different from 0 ($Q_{25}<0<Q_{75}$, details in SI) were discarded until 500 acceptable simulations were found. Thereby, 331 PMF runs were selected for PMF_{block} (230 for PMF_{zue,isol}, 99 for PMF_{zue,reprs}, and 269 for PMF_{1filter/month}). Median factor time series and recovery parameters from all retained simulations were then determined and the interquartile range (IQR) represents our best estimate of the uncertainties for the single PMF datasets. The Monte Carlo process was repeated for the four different PMF datasets described above and the resulting median time series of their estimated uncertainties were compared. The resulting uncertainty estimates and the method are described in Sec. 4.2.1. and in the SI. ...”

The added information in the SI reads:

“... ”

Cumulative density functions for the α -values of HOA and COA are presented for the accepted solutions in Fig. S4. We found that 80% of the accepted solutions have an α -value \leq 0.3 for HOA and an α -value \leq 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.

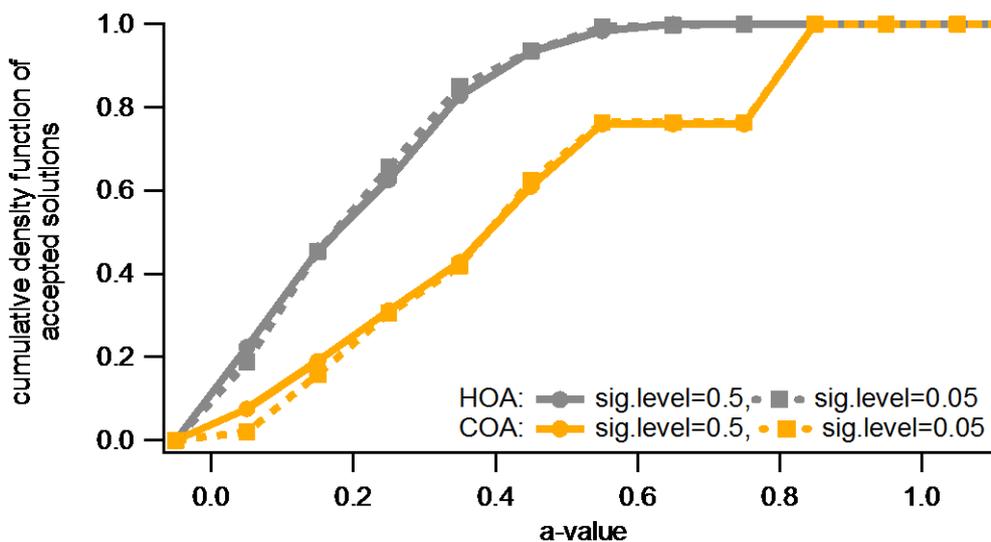


Figure S4: Cumulative density functions of a-values for HOA and COA for the accepted solutions.

The yearly average factor concentrations of all selected PMF_{block} solutions after R_k correction are shown 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.

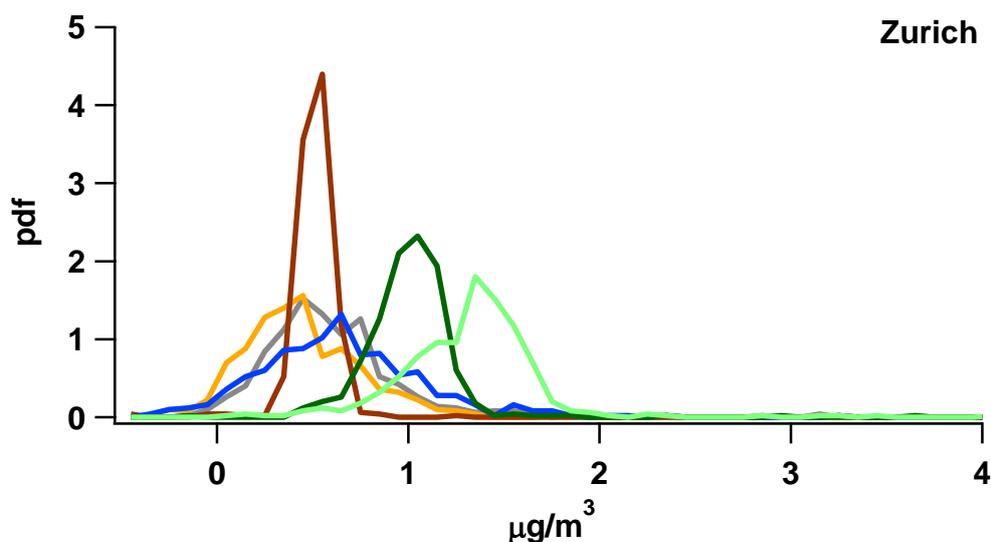


Figure S5: Histograms of yearly average factor concentrations of all selected PMF_{block} solutions (after R_k correction)....”

If DISP is used for uncertainty estimation of a case with several local solutions, one often obtains the outcome that the model is "Not Well Defined" or "NWD". I would not suggest what the authors should do with their many-solution cases in addition to discussing them. Whatever they opt to do, they should describe it: what was done and why.

The procedure that we have adopted aims at assessing the sensitivity of the PMF to the constrained profiles (a-value), to the pseudo-random starting point (seed), to the measurement uncertainties (by

including repeated measurements from one of the sites) and to the input data/rotational ambiguity (bootstrap). We have discarded a number of solutions based on a set of criteria. The PMF runs satisfying these criteria were all considered as plausible and discussed in the manuscript. Differences amongst these solutions are used as an estimate of the combined uncertainties mentioned above. The process is described in Sec. 4.2.1. and in the Supplementary Information. We added the following sentence at the end of Sec. 3.6 linking to Sec. 4.2.1.

"... The resulting uncertainty estimates and the method are described in Sec. 4.2.1. and in the SI. ..."

Section 3.3, sensitivity analysis

I cannot comment more on this analysis because I do not understand what the α -values are and how they were used.

We do agree with the reviewer that the equation is misleading. The constraints were performed in the same way as in Canonaco et al. (2013) and the equation in the text was adapted to the one in the mentioned publication. The modified text reads as follows:

"...

The PMF algorithm was solved using the multilinear engine-2 (ME-2, Paatero, 1999). Normalization of the PMF solution during the iterative minimization process is disabled as implemented in SoFi (Canonaco et al., 2013). ME-2 enables an efficient exploration of the solution space by a priori constraining the $f_{k,j}$ elements within a certain range defined by the scalar α ($0 \leq \alpha \leq 1$) from a starting value $f'_{k,j}$, such that the modelled $f_{k,j}$ in the solution satisfy Eq. 4:

$$f_{k,j} = f'_{k,j} + \alpha * f'_{k,j} \quad (4)$$

$f'_{k,j}$ is the starting value used as a priori knowledge from previous studies and $f_{k,j}$ is the resulting value in the solution.

"..."

P6 L26-28 say:

> Paatero et al. (2014) compare the effectiveness in estimating modelling errors using two different approaches: the displacement (DISP) and bootstrap analysis (BS), respectively.

Here seems to be a terminological problem: in the quoted paper, Paatero et al. estimated the uncertainties of estimated F factor elements in the situation when no modelling errors are present. These F uncertainties depend mainly on random error in X and on rotational freedom of factor matrices G and F. It was specifically emphasized that the obtained uncertainties do not cover effects of modelling errors in the results. (Examples of modelling errors: non-constant factor profiles, wrong uncertainties assigned to data values.) Thus modelling errors were not estimated in the quoted paper.

P6 L29:

DISP involves running PMF several times using randomly perturbed factor profile elements of a reference solution

In fact, DISP estimation does not involve any randomness at all. F factor elements are perturbed in a systematic fashion by DISP. Perhaps here is confusion with Monte Carlo methods where random perturbations may be applied. DISP is not Monte Carlo.

We apologize for the terminological mistake. The respective part of the manuscript has been corrected.

“... While this approach has been proven very effective in selecting a range of environmentally relevant solutions (Elser et al., 2016a, 2016b and Daellenbach et al., 2016), the resulting uncertainties may be underestimated. Paatero et al. (2014) compared the effectiveness in estimating uncertainties of factor elements using two different approaches: the displacement (DISP) and bootstrap analysis (BS). BS involves applying the model to input matrices consisting of a subset of the entire dataset. DISP involves running PMF several times using systematically perturbed factor profile elements of a reference solution, but allowing a defined difference in Q from the reference solution. Both approaches are computationally intensive, especially DISP. Because of such computational limitations the combination of BS and DISP was not feasible for the dataset presented here, especially in combination with α -value sensitivity tests. Therefore, we chose to perform 4 sensitivity tests performing PMF runs using 4 different input datasets, presented in the following. These sensitivity tests allow conclusions on the stability of PMF analysis when reducing the temporal or spatial resolution as well as the influence of the measurement repeatability.

..”

In table (3), uncertainty estimates of percentage concentrations are not correctly computed.

The overall uncertainties on the mass concentrations were estimated as the sum in quadrature of two error terms. The first relates to the variability in the yearly average mass concentrations for the different factors at the different sites calculated by changing the COA and HOA α -values using the PMF_{block} setting (where all data points are considered, σ_a). The second error term relates to the additional variability in the yearly average mass concentrations for the different factors at the different sites due to the change in the input matrix (PMF_{block} , $PMF_{zue,isol}$, $PMF_{zue,repr}$). As the latter was determined only for the case of Zurich, then the same additional relative error (σ_b) was assumed for the other sites. This is indeed only a best estimate of the errors. In the corrected version of the manuscript we clarify the way these errors were determined (in Caption to Tab. 3) and we do not show errors on the percentage concentrations.

“...Table 1: Yearly average contribution and uncertainty of resolved factors for PMF_{block} run for the different sites and the average for all sites. The uncertainty is calculated based on the variability of the yearly averages from PMF_{block} and the variability between the sensitivity tests. ...”

Notation:

In supplement, subscript "i" is used as a subscript of Q. It is not defined what "i" means here. Does it mean number of factors? If yes, then the symbol used for number of factors should be used. If i does not have a definable meaning, then it might be clearer to omit the subscript in this case. In general, systematic use of subscripts would be a help for the reader. E.g. use i only as the index of sample (time), j as index of column of X, k index of factor. For other quantities, select other symbols and define what they mean.

The index i in this figure and the text was removed.

In different places, factor elements are denoted as G_{ik} and as g_{ik} . This may confuse readers. Eventually they will recognize that this difference does not mean anything, but first they will waste time trying to understand. Either, select one notation (preferred), or, in the section "Notation" (to be written), specify that G_{ik} and g_{ik} mean the same, and also F_{kj} and f_{kj} mean the same.

G_{ik} was replaced by g_{ik} .

Eq (2) is incorrect. If you wish to use matrix element notation, then summation over k must be indicated. If you wish to use vector-matrix notation, then its use should be defined, especially because many of your readers are not familiar with such notation. In vector-matrix notation, index k would not be visible.

Equation 2 and the adjacent text have been corrected:

“...

Source apportionment of the organic aerosol is performed using positive matrix factorization (PMF, Paatero, 1994). PMF is a statistical un-mixing model explaining the variability of the organic mass spectral data ($x_{i,j}$), as linear combinations of static factor profiles ($f_{j,k}$) and their time-dependent contributions ($g_{i,k}$), see Eq. 2 (where p is the number of factors). The index i represents a specific point in time, j an ion, and k a factor. The elements of the model residual matrix are termed $e_{i,j}$.

$$x_{i,j} = \sum_{k=1}^p g_{i,k} f_{k,j} + e_{i,j} \quad (2)$$

...”

P5 L10: sunset → Sunset OC/EC analyzer

Has been corrected.

P5 L21

> Fitted ions in our datasets missing in the reference ...

What do you mean by "fitted ions"? I do not understand this sentence.

The sentence was rephrased to:

“...Ions that were present in our datasets but not in the reference profiles for HOA and COA were inferred from published unit mass resolution (UMR) profiles (Ng et al., 2011 and Crippa et al., 2013c). ...”

P7 L21:

> The identity of HOA and COA were identified first as their mass spectra were initially constrained.

Why do you need to identify HOA and COA? I would assume that they are on preselected rows of F , such as rows 1 and 2. No identification is needed for factors that are in known positions. What is wrong here? Am I understanding all this completely wrong?

Yes, indeed HOA and COA are on preselected rows of F. The text has been adapted to:

“... Since HOA and COA were initially constrained on preselected rows of F, they did not need to be identified. ...”

Supplement

Figure SI.1

shows ratios of obtained Q vs. expected Q_{exp}. How were Q_{exp} computed? Did you take into account that downweighted columns of X contribute very little to Q_{exp}? How many downweighted columns were present?

The results displayed in FigS1 refer to the distributions of $Q_{i,j}$ (median and quartiles and not to Q/Q_{exp} . $Q_{i,j}$ is computed as:

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

We corrected the text and axis labels accordingly and refer only to $Q_{i,j}$.

In PMF_{block}, only 7 of 202 ions were downweighted by a factor 3. Therefore, the bias in Q between using the non-downweighted or the downweighted $s_{i,j}$ is smaller than the variability in Q for the same set of solutions. Thus we do not think that this bias has an influence on the results, all the more since the absolute Q-values are not used in the solution selection.

Obtained ratios Q/Q_{exp} are of the order of 10 for 6 factors. This indicates that there are one or several significant modeling errors. It has been assumed among atmospheric scientists that a ratio of 4, say, would not be significant, that it could be caused by random variations from the expected Q. This assumption is totally wrong. Such ratios (>1.5, say) always have a cause that preferably should be understood in a project where careful mathematical analysis is attempted.

Possible causes of modeling errors are: underestimation of random errors in mass spectra, variation of factor profiles with time and between sites, systematic errors in preprocessing m/z spectra, and spurious sporadic local sources that cannot be modeled by PMF. An attempt should be made at understanding and discussing those errors, even if the effect to obtained results cannot be eliminated any more at this final stage. One useful diagnostic is to examine contributions to Q from different m/z values, from different times of day and days of week, and so on.

Variation of factor profiles between sites:

We have explicitly represented the average $Q_{i,j}$ per site and per season to examine the performance of the model and reasons behind the observed $Q_{i,j}$ values. As we have already pointed out in the manuscript the model does get the data at sites in the north better than at sites in the South during winter. We did attribute this difference to the variability in the wood burning source profiles, here represented by only one factor. Indeed, we do agree that additional reasons may result in the high $Q_{i,j}$ observed. These are discussed below.

First, as pointed out above we do represent sources at all sites with a single profile and the examination of the site-to-site median $Q_{i,j}$ values do indicate a discrepancy in the model performance between sites in the north and south of the Alps. 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,reprs}, average $Q_{i,j}$ 6). The difference between the site that is best explained and

the site that is least explained is also approximately 6 (Fig. SI.1). Therefore, we do indeed attribute a significant part of the unexplained variability to variation in source profiles at different sites.

Variation of factor profiles with season:

We note that these differences can also occur because of representing SOA at different seasons with only two factors: For example the difference in Q values between the best explained season (April-Mai-June) and the least well explained season (January-February-March) is approximately 4 (Fig. SI.1). A consistent difference between Q for week-days and weekends cannot be found (Fig. S2).

Underestimation of random errors in mass spectra

As mentioned in the manuscript, the samples from Zurich were measured at two instances, 5 months apart. When performing PMF on either batch separately (PMF_{zue,isol} and PMF_{zue,rep}, respectively) their average $Q_{i,j}$ is comparable (Q=6). Therefore, there does not seem to be a difference in the representation of random errors between these measurements.

Systematic errors in preprocessing m/z spectra

As we have mentioned in the manuscript the error matrix considers the detector counting statistics and the variation in the background. However, additional uncertainties can result from peak fitting of the high resolution AMS data (attribution of signal at a nominal mass to several ions). This procedure is affected by errors in the m/z calibration accuracy and precision and by the peak width. These errors are not currently taken into account, but may result in a significant underestimation of the measurement uncertainties especially for overlapping peaks, which may explain (at least partially) the high average $Q_{i,j}$.

In Fig. S3, we present the average $Q_{i,j}$ of individual ions. There is no clear dependence on the ion molecular weight and therefore, m/z. However, the average $Q_{i,j}$ for ions with a small mass defect (nominal mass – exact ion mass) is higher (~10) than for the other ions (~3). There are a lot of ions with small mass defects close to 0.03 a.m.u. in our dataset. This makes these peaks prone to overlap with other ions and thus their error subject to underestimation because this effect is not considered in the s_{ij} calculation.

Sporadic local sources that cannot be modeled by PMF:

The ion families CHN and especially CS show higher average Q than CH, CHO1, and CHOgt1. Since CH₃SO₂⁺ shows an event-driven time series, at least the high Q related to CS ions could be related to the inability of PMF to resolve these events accurately.

The corresponding section in the SI has been thoroughly reworked and reads now:

“... ”

Based on the input data for PMF_{block}, we evaluate the influence of the number of factors, p, on Q_i . 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 ($s_{i,j}$):

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

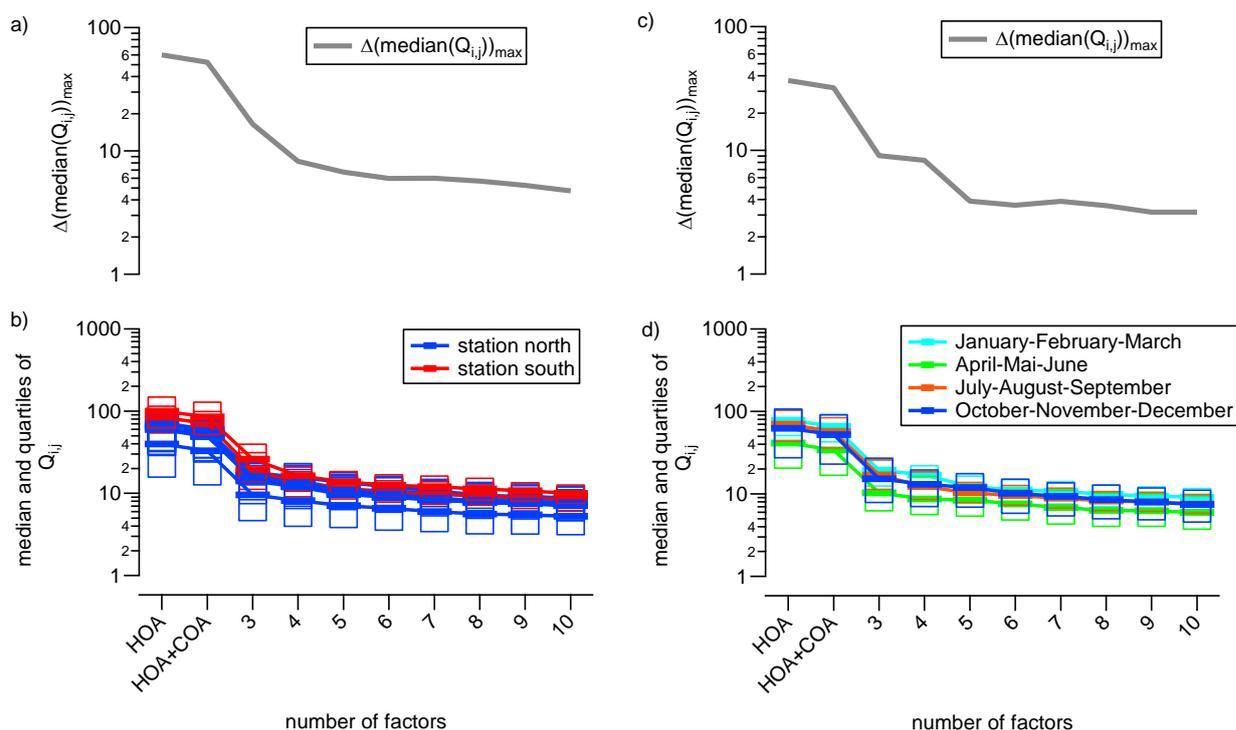


Figure S1: $Q_{i,j}$ as a function of the number of factors for a reference experiment with all data used in PMF (9 sites, full year 2013, HOA and COA constrained with $a=0.0$ (b and d). $\Delta(\text{median}(Q_{i,j}))_{\text{max}}$ is evaluated for the different periods during the year 2013 (January-February-March, 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 $Q_{i,j}$. $\Delta(\text{median}(Q_{i,j}))_{\text{max}}$ shows the difference in the median $Q_{i,j}$ between groups of points like sites or season. The smaller the $\Delta(\text{median}(Q_{i,j}))_{\text{max}}$, 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).

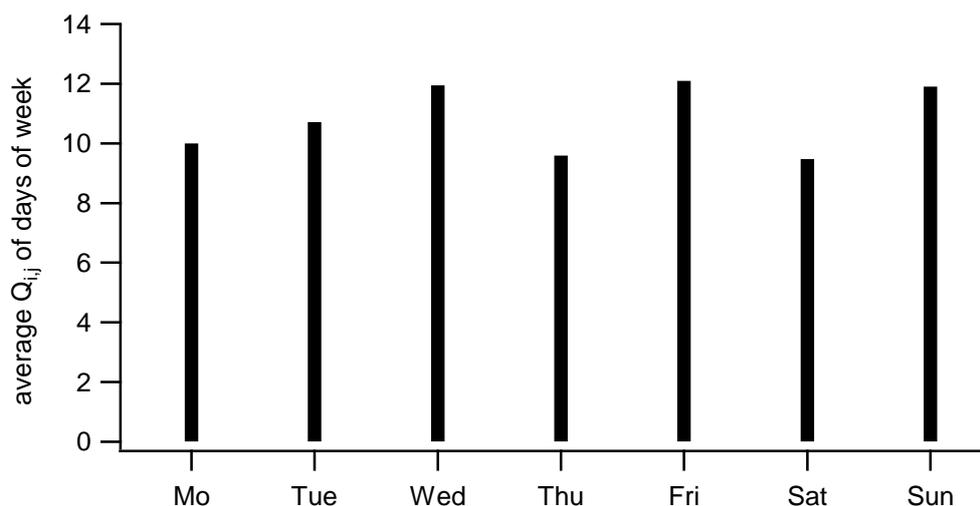


Figure S2: Q_{ij} as a function of the day of the week.

However, for PMF_{block} also with 6 factors, the average Q_{ij} 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_{ij} for Zurich is slightly reduced for the same number of factors when only including 1 site in PMF ($PMF_{zue,isol}$, $PMF_{zue,reprs}$, average Q_{ij} 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_{ij} 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_{ij} than hydrocarbon ions (CH, only C and H) and oxygenated ions ($CHO_{z=1}$ with 1 oxygen and $CHO_{z>1}$ with more than 1 oxygen). Since the time series of $CH_3SO_2^+$ is event-driven, the high Q_{ij} of this ion hints to the fact that PMF is unable to accurately resolve all of these events.

The average Q_{ij} 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).

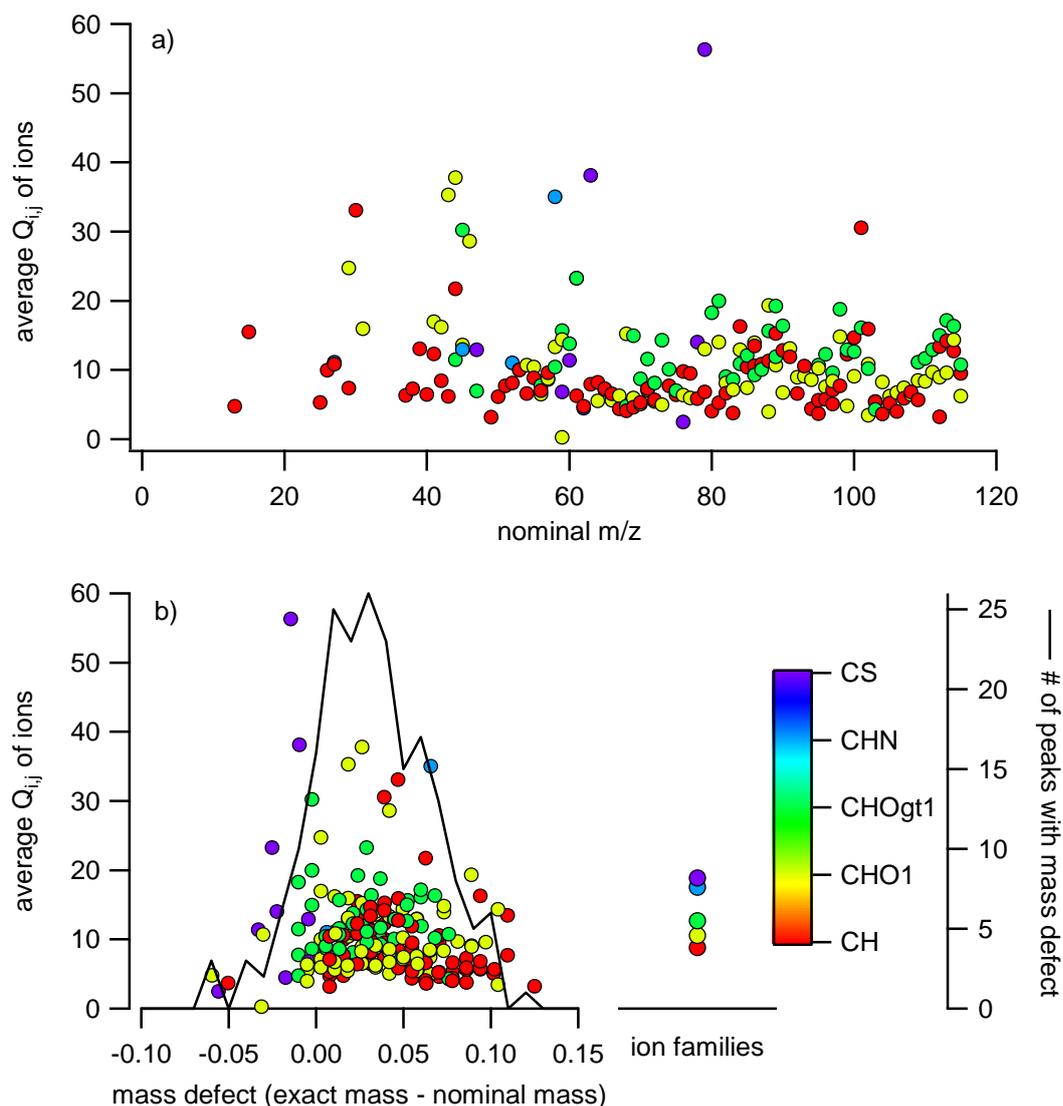


Figure S3: a) Average $Q_{i,j}$ of ions in PMF_{block} as a function of their mass-to-charge ratio (m/z). The ions are color-coded with their composition (CH: ions consisting only of C and H; CHO1: ions consisting of C, H, and 1 O; CHOgt1: ions consisting of C, H, and more than 1 O; CHN: ions consisting 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.

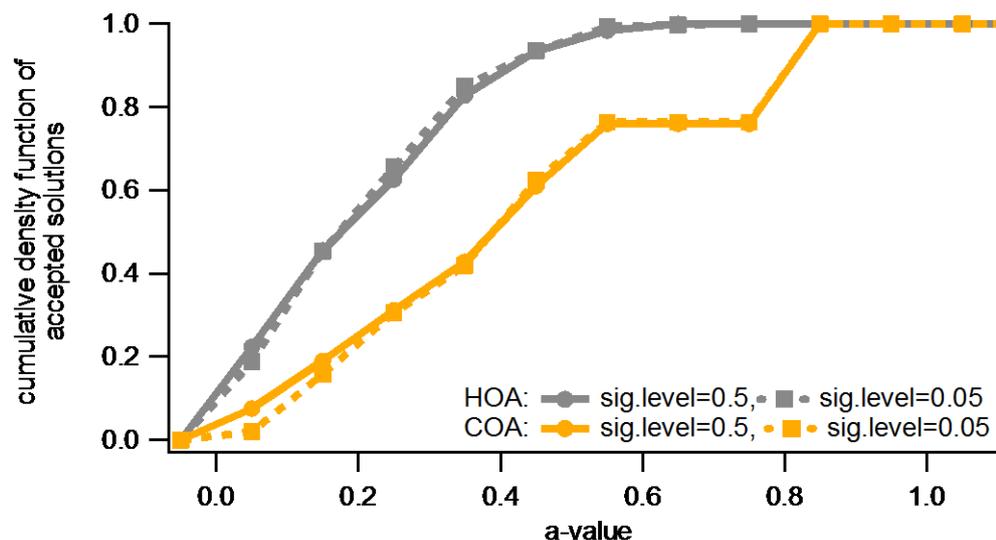


Figure S4: Cumulative density functions of a-values for HOA and COA for the accepted solutions.

The yearly average factor concentrations of all selected PMF_{block} solutions after R_k correction are shown 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.

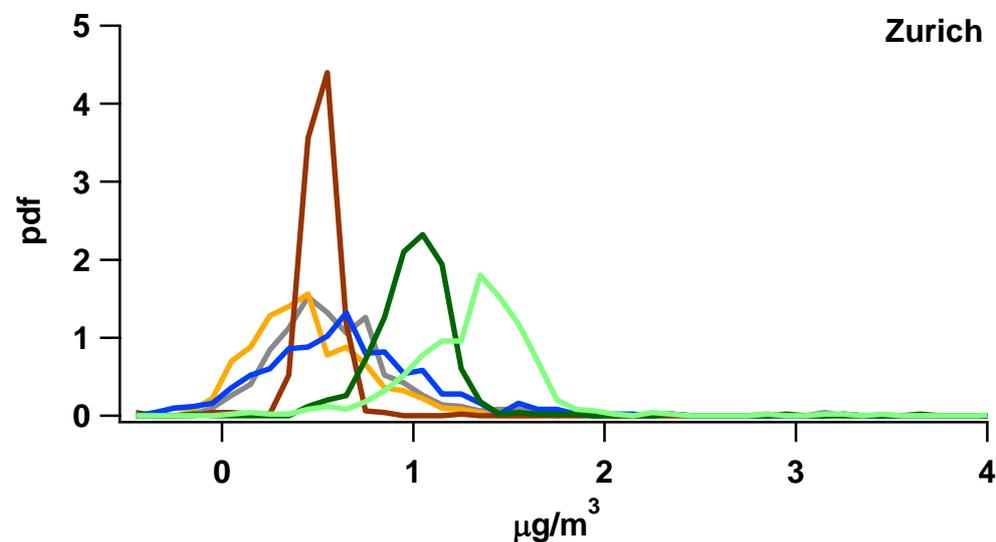


Figure S5: Histograms of yearly average factor concentrations of all selected PMF_{block} solutions (after R_k correction).

... “

Table SI.1

Why is there a table for mass closure criteria (used for rejecting bad solutions) when all entries in this table are identical? The criteria seem to concern distribution of residuals of OC fitting. Is it really so

that the fit is rejected if 1st quartile point is negative and 3rd quartile point is positive? In other words, the fit is rejected if residuals are symmetrical around zero. Usually, such residuals would be considered desirable.

Symmetrical residuals around 0 were considered desirable and solutions were accepted if the first quartile was smaller than 0 and the third quartile larger than 0. The title of the Table was changed to:

“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.”

References:

Allan, J. D., Jimenez, J. L., Williams, P. I., Alfarra, M. R., Bower, K. N., Jayne, J. T., Coe, H., and Worsnop, D. R.: Quantitative sampling using an Aerodyne aerosol mass spectrometer 1. Techniques of data interpretation and error analysis, *J. Geophys. Res.*, 108, 4090, doi:10.1029/2002JD002358, 2003.

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., 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.

Bozzetti, C., Sosedova, Y., Xiao, M., Daellenbach, K. R., Ulevicius, V., Dudoitis, V., Mordas, G., Byčenkienė, S., Plauškaitė, K., Vlachou, A., Golly, B., Chazeau, B., Besombes, J.-L., Baltensperger, U., Jaffrezo, J.-L., Slowik, J. G., El Haddad, I., and Prévôt, A. S. H.: Argon offline-AMS source apportionment of organic aerosol over yearly cycles for an urban, rural and marine site in Northern Europe, *Atmos. Chem. Phys. Discuss.*, doi:10.5194/acp-2016-413, 2017a.

Bozzetti, C., El Haddad, I., Salameh, D., Daellenbach, K. R., Fermo, P., Gonzalez, R., Minguillón, M. C., Iinuma, Y., Poulain, L., Müller, E., Slowik, J. G., Jaffrezo, J.-L., Baltensperger, U., Marchand, N., and Prévôt, A. S. H.: Organic aerosol source apportionment by offline-AMS over a full year in Marseille, *Atmos. Chem. Phys. Discuss.*, doi:10.5194/acp-2017-54, in review, 2017b.

Chen, Q., Miyazaki, Y., Kawamura, K., Matsumoto, K., Coburn, S. C., Volkamer, R., Iwamoto, Y., Kagami, S., Deng, Y., Ogawa, S., Ramasamy, S., Kato, S., Ida, A., Kajii, Y., and Mochida, M.: Characterization of chromophoric water-soluble organic matter in urban, forest, and marine aerosols by HR-ToF-AMS analysis and excitation emission matrix spectroscopy, *Environ. Sci. Technol.*, 50, 10,351–10,360, 2016.

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.

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.

Sun, Y., Zhang, Q., Zheng, M., Ding, X., Edgerton, E. S., and Wang, X.: Characterization and source apportionment of water-soluble organic matter in atmospheric fine particles (PM_{2.5}) with high-resolution aerosol mass spectrometry and GC-MS, *Environ. Sci. Technol.*, 45, 4854–4861, 2011.

We thank the referees for their comments, which helped improving the quality of our manuscript. A point by point response (in blue) to the reviewers' comments (in black, italics) will follow. Changes in the text are indicated in in black.

Anonymous Referee #3

Comments on "Long-term chemical analysis and organic aerosol source apportionment at 9 sites in Central Europe: Source identification and uncertainty assessment" by Daellenbach et al. The manuscript presents new research which clearly fits within the scope of the journal. The text is well-written and fairly easy to follow. Some of figures, however, compile several information and are not as straightforward to interpret (e.g. Figure 8) – please make sure to modify them (color axis, split into subplots, etc.) to improve readability.

We removed the second axis (PBOA) in Figure 8 since the difference is only a scaling factor. We improved the readability of several figures.

The technique described here is a follow-up of the characterization of OA measurements based on filter collections followed by water extraction and analysis by HR-AMS, previously published, being the novelty a large statistics from 9 sampling sites and, most importantly, PMF analysis of the OA spectrum from filters. Although the former is unquestionably of scientific interest, it is the latter that will allow others to apply the technique and indeed reach its goals as described in the introduction. At its current stage, the manuscript doesn't fully achieve it.

Major comments:

* The description depth of the PMF applied to this very specific dataset doesn't seem to be proportional to its level of development in regard to the widely used techniques. Please detail it more carefully.

Based on the comments of reviewer 2 and 3 we have significantly adapted the parts related to the description of the PMF. We also highlighted more clearly the part of the model that is typically applied and the other parts that are developed within this work. For example, PMF is widely used, but the ME-2 implementation is not. Also, in the revised version of the manuscript, we do describe more clearly the steps we have adopted for the PMF solution selection.

"...

Source apportionment of the organic aerosol is performed using positive matrix factorization (PMF, Paatero, 1994). PMF is a statistical un-mixing model explaining the variability of the organic mass spectral data ($x_{i,j}$), as linear combinations of static factor profiles ($f_{j,k}$) and their time-dependent contributions ($g_{i,k}$), see Eq. 2 (where p is the number of factors). The index i represents a specific point in time, j an ion, and k a factor. The elements of the model residual matrix are termed $e_{i,j}$.

$$x_{i,j} = \sum_{k=1}^p g_{i,k} f_{k,j} + e_{i,j} \quad (2)$$

atrix, each filter sample was represented on average by 11 mass spectral repetitions to examine the influence of the AMS measurement repeatability on the PMF outputs. A preceding blank from nebulized ultrapure water was subtracted from each mass spectrum. The input errors $s_{i,j}$ required for the weighted least-squares minimization by the model consist of the blank variability ($\sigma_{i,j}$) and the uncertainty related to ion counting statistics and ion-to-ion signal variability at the detector ($\delta_{i,j}$; Allan et al., 2003; Ulbrich et al., 2009). We applied a minimum error according to Ulbrich et al. (2009), and a down-weighting factor of 3 to all fragments with an average signal to noise lower than 2 (Ulbrich et al., 2009). Input data and error matrices consisted of 202 organic ions. The organic fragments, $x'_{i,j}$, obtained from offline AMS analyses do not directly represent ambient concentrations. Therefore, the signal of each fragment was converted to such an ambient concentration ($x_{i,j}$ in $\mu\text{g m}^{-3}$), by multiplying the fraction of this signal with the estimated organic matter (OM) concentration. The latter was calculated as the product of the OC concentrations measured by the Sunset OC/EC analyzer and the OM/OC ratios from the offline AMS measurements $(\text{OM/OC})_{\text{oAMS}}$ (Eq. 3). Note that such scaling does not change the outcome of Eq. 2 since both data and error matrices are scaled in the same manner and the fingerprints ($f_{k,j}$) are not changed.

$$x_{i,j} = \frac{x'_{i,j}}{\sum_i x'_{i,j}} * OC * (\text{OM/OC})_{\text{oAMS}} \quad (3)$$

The Source Finder toolkit (SoFi v.4.9, Canonaco et al., 2013) for Igor Pro software package (Wavemetrics, Inc., Portland, OR, USA) was used to configure the PMF model and for post-analysis. The PMF algorithm was solved using the multilinear engine-2 (ME-2, Paatero, 1999). Normalization of the PMF solution during the iterative minimization process is disabled as implemented in SoFi (Canonaco et al., 2013). ME-2 enables an efficient exploration of the solution space by a priori constraining the $f_{k,j}$ elements within a certain range defined by the scalar a ($0 \leq a \leq 1$) from a starting value $f'_{k,j}$, such that the modelled $f_{k,j}$ in the solution satisfy Eq. 4:

$$f_{k,j} = f'_{k,j} + a * f'_{k,j} \quad (4)$$

$f'_{k,j}$ is the starting value used as a priori knowledge from previous studies and $f_{k,j}$ is the resulting value in the solution. In all PMF runs (unless mentioned otherwise), we used the high resolution mass spectra for HOA and COA (cooking OA) from Crippa et al. (2013b) as constraints, i.e. two rows of $f_{k,j}$ were set equal to the mass spectra of HOA and COA. Ions that were present in our datasets but not in the reference profiles for HOA and COA were inferred from published unit mass resolution (UMR) profiles (Ng et al., 2011 and Crippa et al., 2013c). For this purpose, the fraction of signal at a specific m/z in the UMR reference spectrum ($f_{\text{UMR},m/z}$) was compared to the fraction of signal of all ions at this m/z in the HR reference spectrum ($f_{\text{HR},m/z}$). The difference $f_{\text{UMR},m/z} - f_{\text{HR},m/z}$ was used as entries in $f'_{k,j}$ for such missing ions. For these ions, an a -value of unity was set. For the other factors, the factor

elements were fitted by ME-2. Alternatively, such missing ions can be also treated as ordinary factor elements, to be fitted by ME-2 with all other ordinary factor elements.

...”

** Section 4.2.2 seems quite weak, three methods to estimate PBOA are presented, but no clear conclusion is given other than it underestimates based on previous literature. From my perspective this section doesn't add too much to the manuscript and could easily be removed, however if the authors wish to keep it, please make sure to better constrain the methods into a valid scientific output.*

We believe that removing this section would be misleading and not sufficiently transparent and would hide the fact that SOOA might be overestimated due to the contribution of PBOA, which could not be separated by PMF. Our main objective is not quantifying PBOA, but rather its potential contribution to SOOA. Therefore, with the three approaches presented we attempt to estimate this PBOA contribution. These approaches suggest that PBOA would contribute between 0.3 and 1.0 give units, or 18-58% of SOOA during the warm season. In the revised version of the manuscript, we have clarified the aim of the section related to PBOA contribution estimation. This section reads as follows:

“... Unresolved sources in PMF are an inherent uncertainty of source apportionment analyses. As Bozzetti et al. (2016) show, PBOA can present considerable contributions to OA in PM10 (constituting a large part of coarse OA). In the present analysis, PBOA could not be separated by PMF (neither unconstrained nor using the mass spectral signature from Bozzetti et al., 2016). This inability might be caused by the low water-solubility and the absence of PM2.5 filters in the dataset. Since these coarse particles are only abundant in PM10 and not in PM2.5/PM1, the presence of both PM10 and PM2.5 samples, exhibiting a large gradient in PBOA, might allow an unambiguous separation of PBOA. The aim of this section is to estimate the influence of PBOA on the source apportionment results. A quantification of this fraction is, however, beyond the scope of this paper. In the following, we estimate the influence of PBOA in three alternative ways: ...

...”

Minor comments:

** Abstract. L.21: add the word “from” between μm and 9.*

As suggested by another reviewer, we added “at”.

** Abstract. L.24: remove “which is” and add a comma before related.*

The text has been corrected.

** P.2, L.31: Please remove “restricted to WSOA in”.*

The respective part has been removed.

** P.12, L.10: Remove the word “here”.*

The word “here” was removed.

* External gas-phase tracers (besides the use of NOx just to separate HOA, COA) could also add some information of the surrounding chemistry – for example, what is the ozone (over 24h, or just afternoon) in regard to SOOA and WOOA? And Ox?

For Zurich, we added a comparison of SOOA concentrations to ozone and Ox (O₃+NO₂):

“...

In Figure S7, we compare the SOOA concentrations to ozone and Ox (O₃+NO₂) 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 $R_{s,SOOA,O_3}=0.33$, Fig. S7.b) and Ox ($R_{s,SOOA,Ox}=0.38$, Fig. S7.c).

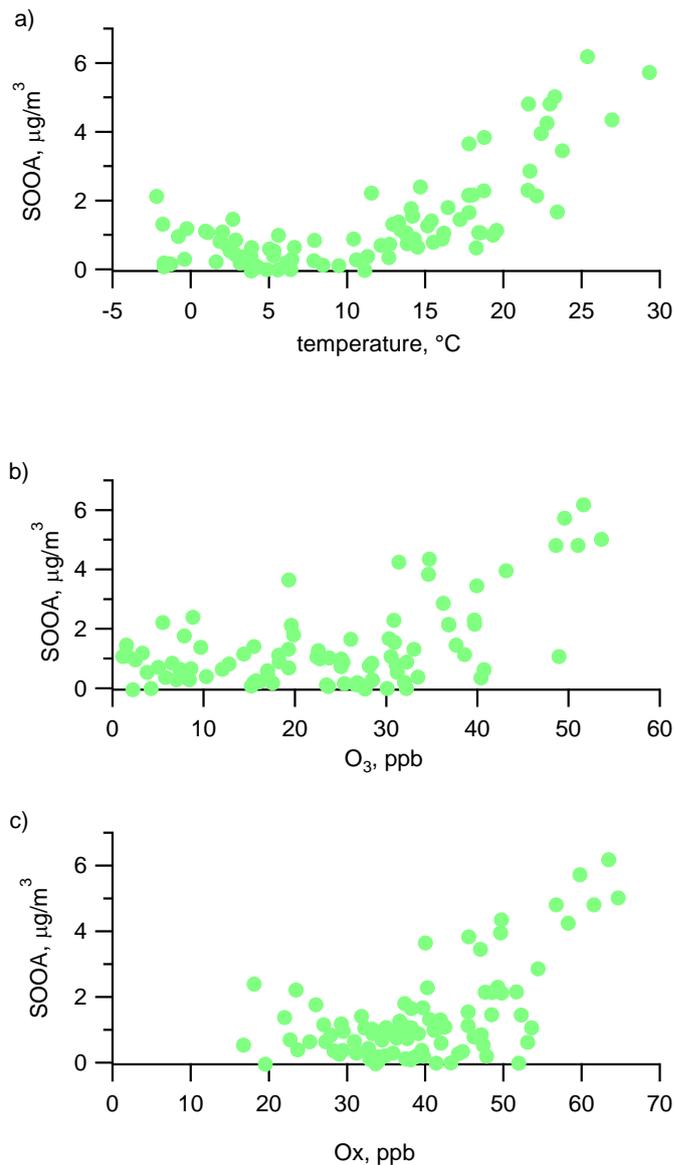


Figure S7: SOOA concentrations compared to temperature, ozone, and Ox (O₃+NO₂) for Zurich.

“...”

References:

Allan, J. D., Jimenez, J. L., Williams, P. I., Alfarra, M. R., Bower, K. N., Jayne, J. T., Coe, H., and Worsnop, D. R.: Quantitative sampling using an Aerodyne aerosol mass spectrometer 1. Techniques of data interpretation and error analysis, *J. Geophys. Res.*, 108, 4090, doi:10.1029/2002JD002358, 2003.

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., 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.

Bozzetti, C., Sosedova, Y., Xiao, M., Daellenbach, K. R., Ulevicius, V., Dudoitis, V., Mordas, G., Byčenkienė, S., Plauškaitė, K., Vlachou, A., Golly, B., Chazeau, B., Besombes, J.-L., Baltensperger, U., Jaffrezo, J.-L., Slowik, J. G., El Haddad, I., and Prévôt, A. S. H.: Argon offline-AMS source apportionment of organic aerosol over yearly cycles for an urban, rural and marine site in Northern Europe, *Atmos. Chem. Phys. Discuss.*, doi:10.5194/acp-2016-413, 2017a.

Bozzetti, C., El Haddad, I., Salameh, D., Daellenbach, K. R., Fermo, P., Gonzalez, R., Minguillón, M. C., Iinuma, Y., Poulain, L., Müller, E., Slowik, J. G., Jaffrezo, J.-L., Baltensperger, U., Marchand, N., and Prévôt, A. S. H.: Organic aerosol source apportionment by offline-AMS over a full year in Marseille, *Atmos. Chem. Phys. Discuss.*, doi:10.5194/acp-2017-54, in review, 2017b.

Chen, Q., Miyazaki, Y., Kawamura, K., Matsumoto, K., Coburn, S. C., Volkamer, R., Iwamoto, Y., Kagami, S., Deng, Y., Ogawa, S., Ramasamy, S., Kato, S., Ida, A., Kajii, Y., and Mochida, M.: Characterization of chromophoric water-soluble organic matter in urban, forest, and marine aerosols by HR-ToF-AMS analysis and excitation emission matrix spectroscopy, *Environ. Sci. Technol.*, 50, 10,351–10,360, 2016.

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.

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.

Sun, Y., Zhang, Q., Zheng, M., Ding, X., Edgerton, E. S., and Wang, X.: Characterization and source apportionment of water-soluble organic matter in atmospheric fine particles (PM_{2.5}) with high-resolution aerosol mass spectrometry and GC-MS, *Environ. Sci. Technol.*, 45, 4854–4861, 2011.

We thank the referees for their comments, which helped improving the quality of our manuscript. A point by point response (in blue) to the reviewers' comments (in black, italics) will follow. Changes in the text are indicated in in black.

Anonymous Referee #4

General comments:

In this paper the concentrations of the six types of organic aerosol (OA) components (HOA, COA, BBOA, WOOA, SOOA, and SC-OA) over Switzerland are reported based on the off-line analysis of the water-soluble aerosol components in aerosol samples using an aerosol mass spectrometer (AMS). The characteristics of the retrieved OA components, e.g., the relative abundances and seasonality, are presented. Further, the uncertainty of the concentrations of the retrieved OA is discussed. The source identification of OA components based on long-term samplings at multiple locations is important, and the application of the aerosol mass spectrometry for the chemical analysis of aerosol samples collected on filters made it possible in this study. The contributions of the major sources of OA to the atmospheric concentrations in the studied area have been characterized well in view of location and seasonality. Although the results presented in this paper are highly valuable, this paper needs substantial improvement in terms of the presentation quality. The explanations for the statistical analyses are not fully comprehensive, and a part of them would be flawed. Further, the point of this study is not very clear because both the methodology of the analysis itself and the results based on it are presented and discussed. To make the point clearer, it may be better to move the discussion on the uncertainty based on the results in Figures 6 and 7 to the experimental section or the supplement. Other minor issues regarding the presentation quality include inadequate explanations, undefined abbreviations/symbols, and grammatical errors. For the reasons above, substantial improvement is required for the publication of this paper in its final form. More specific comments are listed below.

Specific comments:

Page 3, 1st paragraph: It may be better to explain more about previous source apportionment studies for organic aerosols using off-line AMS measurement techniques. The group of the first and corresponding authors reported two more studies, both of which were also for European sites (Bozzetti et al., 2017a, 2017b). There are also other source apportionment studies based on statistical analysis for the mass spectra obtained using off-line AMS techniques (Sun et al., 2011; Chen et al., 2016). Emphasis should be on which characteristics of atmospheric aerosols have not been studied tentatively even by the use the off-line AMS techniques.

In the revised version of the manuscript, we have mentioned previous work that used a similar methodology to here. While Chen et al. (2016) have used factor analysis for the mass spectra obtained using off-line AMS techniques; their focus was on the identification of chromophores. Indeed, the studies by Bozzetti et al. (2017a, b) from our group use the same methodology for a similar purpose that is the determination of spatially resolved trends of OA sources. However, the sites studied and the challenges faced in the aforementioned studies are different from here. Here, as the sites are not homogeneous the main aim of the paper is providing a methodology to satisfactorily represent the OA by few factors, with a systematic and objective assessment of the results and the underlying uncertainties.

This paragraph reads now:

“... This approach allows the retroactive investigation of specific events, e.g. haze events in China (Huang et al., 2014) as well as AMS measurements of coarse mode aerosol (Bozzetti et al., 2016) and long-term source apportionment studies (Bozzetti, 2017b, 2017a). Such an approach was also used in recent studies for identifying the different types of water-soluble chromophores (Chen et al., 2016). Additionally, such filters are routinely collected and are already available over multi-year periods at many air quality monitoring stations around the world for years/decades. ...”

Page 3, 2nd paragraph: The chemical analysis using the AMS was limited to the watersoluble component of organics in PM10, although the water-insoluble organic component was also taken into consideration in the source apportionment. This point should be addressed more explicitly.

Indeed, the chemical analysis using the AMS was limited to the water soluble fraction. However, using factor-specific recoveries determined in Daellenbach et al. (2016), the contribution of the different factors to WSOC could be scaled to OC. This is described in detail in Section 2.5.

Page 3, line 13: The site-to-site differences and time series are not explained in a specific part of this paper.

This sentence was misleading. Detailed analysis of the site-to-site differences and time series will be presented in a second paper. The sentence has been adapted to:

“...In a second paper, we will investigate the site-to-site differences and general trends in the factor time series and their relationship with external parameters. ...”

Page 4, lines 1-3: How were the mass spectra of the extracts from aerosol samples corrected for field blanks? Because the sensitivity of an AMS to aerosol components depends on the particle size, the signal intensity of organics should not be proportional to the organic mass flux from the nebulizer. For this reason, the assessment of the blank level is not straightforward. More explanation to this point is necessary.

Between two samples we measured ultrapure water. The recorded signal was subtracted from the sample spectra. In a previous study, we showed that the organic blank measurements collected by ultrapure water nebulization provide a comparable blank estimate to the organic blanks determined from the nebulization of NH_4NO_3 (Bozzetti et al., 2017a). However, we also analyzed field blanks which were extracted and measured in the same way as the exposed samples.

..

We added accordingly a statement in the manuscript on P3 L30 –P4 L7:

“... The measurement blank was determined before and after every filter sample. Each sample was recorded for 480 seconds (AMS V-mode, m/z 12-447), with a collection time for each spectrum of 30 seconds. Ultrapure water was measured for 720 seconds. Once per day, ultrapure milliQ water was nebulized with a particle filter interposed between the nebulizer and the AMS, for the determination of the gas-phase contribution to the measured mass spectrum, which was then subtracted during analysis from both blanks and filter samples. The filters from Zurich were analysed twice with a time difference of approximately 5 months to assess the measurement repeatability. High resolution mass spectral analysis was performed for each m/z (mass to charge) in the range of 12- 115. The measurement blank was subtracted from the sample spectra. In a previous study, it has been shown that the measurement blank is comparable to the organic blanks obtained from the nebulization of NH_4NO_3 (Bozzetti et al. (2017a). The interference of NH_4NO_3 on the CO_2^+ signal described by Pieber et al. (2016) was corrected as follows (Eq. 1): ...”

We have previously shown that the organic signal from the nebulization of MQ water is not statistically significantly different from the organic signals from the nebulization of NH_4NO_3 (Bozzetti et al. (2017a)), which might potentially act as a carrier seed of contaminants. Therefore, we considered the MQ water to be an adequate representation of the measurement background. In addition to the measurement blanks, we have measured field blanks following the same procedure. These samples showed WSOC and OC concentrations higher than instruments detection limits. As this contamination can contribute to different extents to different factors, data have been corrected post PMF as described in Section 2.5.

In order to account for the effect of the field blanks on the source apportionment, we subtracted the blank concentrations factor after the PMF analysis. To that purpose we performed PMF runs using $\text{PMF}_{\text{block}}$ while also the field blank measurements were included in the PMF run. Thereby, we found how much the different factors contributed to the field blanks. Finally, we subtracted this effect from the factor time series. In addition, the OC blank levels used in the previous version of the manuscript were overestimated and have now been updated. Therefore, all numbers in the manuscript related to the source apportionment analysis slightly changed (yearly average factor concentrations changed by around 15%). However, the main conclusions remain the same.

The respective text was adapted for a better readability:

“... For a limited number of PMF runs ($\text{PMF}_{\text{block}}$) also the field blank analyses were included in the PMF input data. This provides the contribution of different factors to the field blanks which were used to correct the output factor time series. Uncertainties induced by the blank subtraction were propagated. ...”

Page 4, lines 9-10: The expression in the parenthesis is unclear and needs to be reworded.

The text has been adapted in the revised version of the manuscript:

“... The correction factor $\left(\frac{\text{CO}_2, \text{meas}}{\text{NO}_3, \text{meas}}\right)_{\text{NH}_4\text{NO}_3, \text{pure}}$ was determined based on measurements of aqueous NH_4NO_3 conducted regularly during the entire measurement period and varied between ~1% and ~5% (Pieber et al., 2016). ...”

Page 5, lines 9-11: The method for rescaling here and that explained in the 2nd paragraph of page 9 does not seem identical.

While OC concentrations are available for all samples, WSOC concentrations are only available for a subset of all samples (Magadino and Zurich). Therefore, for the samples from Magadino and Zurich it was possible to evaluate mass closure, i.e. whether the sum of WSOC factor concentrations after R_k correction matched the measured OC ($\text{OC}_{i, \text{res}} = \text{OC}_{i, \text{meas}} - \sum_k \text{WSOC}_{i, k} / R_k$). For all other samples this was not possible because of the unavailability of WSOC concentrations and, therefore, these samples needed to be scaled to OC.

The text has been adapted in the revised manuscript:

“... The last criterion relates to OC mass closure. A Monte Carlo approach was applied to evaluate whether a combination of water soluble factor time series and recovery parameters would achieve OC mass closure, as described in the following. For the samples from Zurich and Magadino, for which WSOC concentrations were available (in contrast to the other samples), offline AMS measurements were scaled to the water soluble organic matter (WSOM), calculated using the WSOC measurements and OM/OC ratios from the AMS HR analysis. The water-soluble contributions from an identified aerosol source in a sample i were rescaled to their total organic matter concentrations ($OA_{i,k}$), where k represents a given factor, using combinations of factor recoveries as determined by Daellenbach et al. (2016, medians of the used combinations being: R_{HOA} : 0.11, R_{COA} : 0.54, R_{BBOA} : 0.65, and R_{OOA} : 0.89 used for WOOA and SOOA). ...

Page 5, equation 3: The constraint represented by equation 3 seems erroneous because the left and the right parts of the equation are identical.

There was indeed a typo in the equation. The equation was adapted to the presentation in Canonaco et al. (2013), since the interface presented therein was used and the same type of constraints was applied.

Page 5, lines 21-22: Were the inferred fitted ions also for constraint? Does this sentence mean all the factors other than HOA and COA were inferred from published UMR profiles?

Ions that were present in our dataset but not in the reference profiles for HOA and COA were inferred and constrained. However, such ions were given an α -value of unity. For the other factors besides HOA and COA, factor elements were not constrained but fitted by ME-2. The section was adapted to:

“... $f_{k,j}'$ is the starting value used as a priori knowledge from previous studies and $f_{k,j}$ is the resulting value in the solution. In all PMF runs (unless mentioned otherwise), we used the high resolution mass spectra for HOA and COA (cooking OA) from Crippa et al. (2013b) as constraints, i.e. two rows of $f_{k,j}$ were set equal to the mass spectra of HOA and COA. Ions that were present in our datasets but not in the reference profiles for HOA and COA were inferred from published unit mass resolution (UMR) profiles (Ng et al., 2011 and Crippa et al., 2013c). For this purpose, the fraction of signal at a specific m/z in the UMR reference spectrum ($f_{UMR,m/z}$) was compared to the fraction of signal of all ions at this m/z in the HR reference spectrum ($f_{HR,m/z}$). The difference $f_{UMR,m/z} - f_{HR,m/z}$ was used as entries in $f_{k,j}'$ for such missing ions. For these ions, an α -value of unity was set. For the other factors, the factor elements were fitted by ME-2. Alternatively, such missing ions can be also treated as ordinary factor elements, to be fitted by ME-2 with all other ordinary factor elements....”

Page 5, lines 22-24: The explanation in this sentence is not clear. This sentence should be reworded.

The respective sentence has been reworded to:

“ ... For this purpose, the fraction of signal at a specific m/z in the UMR reference spectrum ($f_{UMR,m/z}$) was compared to the fraction of signal of all ions at this m/z in the HR reference spectrum ($f_{HR,m/z}$). The difference $f_{UMR,m/z} - f_{HR,m/z}$ was used as entries in $f_{k,j}'$ for such missing ions ...”

Page 8, line 1: The values of the recoveries used in this study should be presented.

Recoveries are presented in Fig. 6 and the text was adapted to:

The water-soluble contributions from an identified aerosol source in a sample i were rescaled to its total organic matter concentration ($OA_{i,k}$), where k represents a given factor, using combinations of factor recoveries as determined by Daellenbach et al. (2016, medians of the combinations being: R_{HOA} : 0.11, R_{COA} : 0.54, R_{BBOA} : 0.65, and R_{OOA} : 0.89 used for WOOA and SOOA).

Page 8, line 2: The meaning of “the contributions of different factors to the field blank samples” is not clear. What was done is not clear, either.

The text was adapted for a better readability:

“... For a limited number of PMF runs (PMF_{block}) also the field blank analyses were included in the PMF input data. This provides the contributions of different factors to the field blanks which were used to correct the output factor time series. Uncertainties induced by the blank subtraction were propagated. ...”

Page 8, line 27: Is “ $\alpha=0.5$ ” the significance level? Fifty percent is too high.

For PMF_{block} , we performed an additional sensitivity test with $\alpha=0.05$ instead of $\alpha=0.5$. The results are described in a new section in the supplementary material and the comparison mentioned in the main text.

The sentence in the main text reads:

“... A t-test is then used to verify the significance ($\alpha=0.5$) of the average correlation coefficient between factor and marker time series, r_{avg} (Eq. 7):

$$t_{avg} = \frac{r_{avg}}{\sqrt{\frac{1-r_{avg}^2}{N-2}}} \quad (7)$$

Here, r_{avg} is the correlation coefficient averaged over the different stations, derived from the average z value, t_{avg} is the corresponding t-value and N is the average number of samples at the different stations. Results with significance level $\alpha=0.05$ are summarized in Fig. S8.

...”

The section in the supplementary information reads:

“ ...

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.

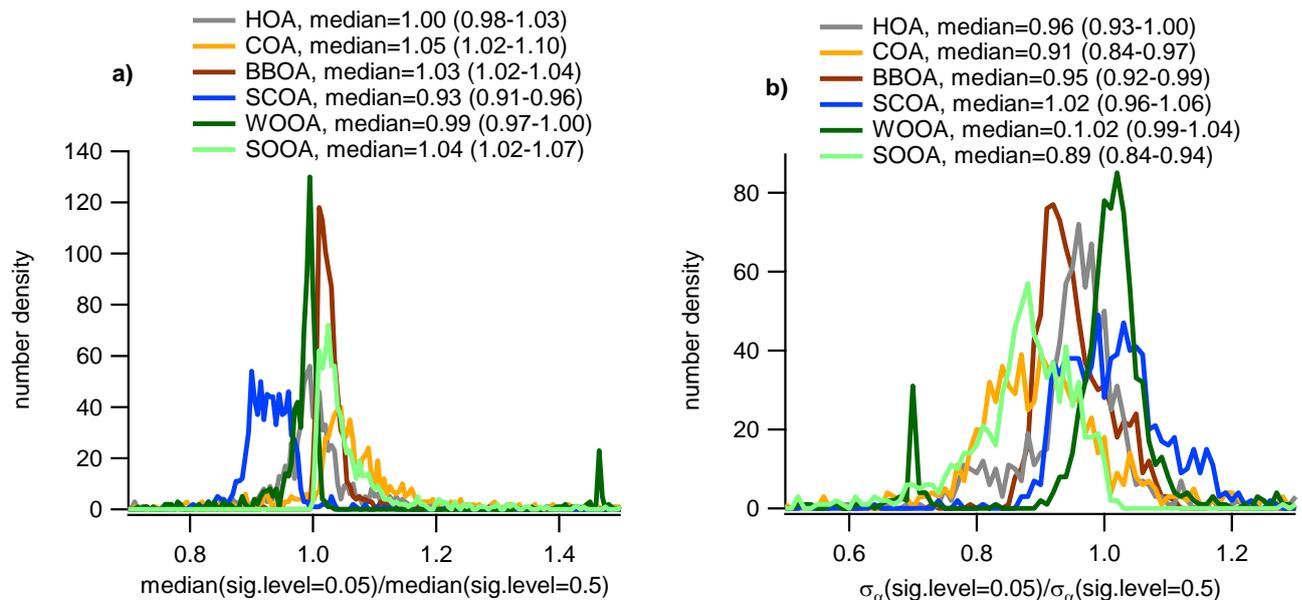


Figure S8: number density functions of source apportionment results obtained using a significance level of 0.05 normalized to results obtained using a significance level of 0.5: a) Comparison of factor concentrations b) Comparison of uncertainty estimate (σ_a).

“ ...”

Page 8, line 26-28: How the statistical analysis using the average values from different stations can be justified? The validity of this method is not obvious.

The ratios of factor concentrations to marker concentrations cannot be assumed to be the same at all sites. Therefore, correlation coefficients need to be calculated at different sites. However, to achieve an optimized system for the entire dataset the average R should be considered.

Page 9, lines 2-4: More details in the calculation should be given so that the readers can assess its validity.

We have adapted the text and added further information:

“ ...

The first two criteria (1-2) ensure an appropriate separation of HOA and COA from OOA and BBOA, respectively. Criteria 3-6 relate to the evaluation of the correlation between factor and marker time

series. This was achieved by computing the Fisher-transformed correlation coefficient z at different stations (Eq. 6):

$$z = 0.5 * \ln\left(\frac{1+r}{1-r}\right) = \arctan(r) \quad (6)$$

where r is the correlation coefficient between factor and marker at a given station. The obtained z values at the different stations are subsequently averaged and transformed back to r_{avg} before further analysis. A t-test is then used to verify the significance ($\alpha=0.5$) of the average correlation coefficient between factor and marker time series, r_{avg} (Eq. 7):

$$t_{avg} = \frac{r_{avg}}{\sqrt{\frac{1-r_{avg}^2}{N-2}}} \quad (7)$$

Here, r_{avg} is the correlation coefficient averaged over the different stations, derived from the average z value, t_{avg} is the corresponding t-value and N is the average number of samples at the different stations. Results with significance level $\alpha=0.05$ are summarized in Fig. S8.

To evaluate whether HOA correlated significantly better with NOx than COA did, the average z values obtained between HOA and NOx and between COA and NOx (Eq. 6) were compared, using a standard error on the z distribution of $1/\sqrt{N-3}$ (Zar, 1999).

...”

Page 9, lines 14-16: Is the issue really explained in the supplement?

This paragraph has been reworded and further information has been added. Now the paragraph reads:

“...The sum of $OC_{i,k}$ from all factors k (mod- OC_i) was then evaluated against the measured OC (meas- OC_i). For this, the residual OC mass (res- OC_i) for each sample was calculated (meas- $OC_i - \text{mod-}OC_i$), and the residual distributions were examined for different conditions that are specified in the Supplement. In summary, a solution was only accepted if res- OC_i were normally distributed around 0 considering all points and subsets of points: a) summer, b) winter c) Magadino, d) Zurich, e) low and high concentrations of the single factors (see Table S1). ...”

Page 10, line 2: What are the percentages of the accepted data?

We added this information to the manuscript at P10 L5:

“...Thereby Thereby, 331 PMF runs were selected for PMF_{block} (230 for PMF_{zue,isol}, 99 for PMF_{zue,repr}, and 269 for PMF_{1filter/month}). ...”

Page 11, line33 – page 12, line 1: This sentence is not clear. Does COA relate to the discussion here?

Besides the correlation between the yearly average concentrations of SC-OA and NO_x, we also present the same values for HOA vs. NO_x and COA vs. NO_x. This comparison allows the conclusion that not all anthropogenically influenced factors show a relation to NO_x.

“... SC-OA instead exhibits low background levels episodically intercepted by remarkable ten-fold enhancements, especially at urban sites affected by traffic emissions (e.g. the SC-OA contribution is significantly higher at sites with higher yearly NO_x average levels). The hypothesis of an influence of traffic activity on SC-OA is provided by the correlation of the yearly average concentrations with NO_x ($R_{s,SC-OA,NO_x} = 0.65$, $n=9$, $p<0.06$) which is, however, comparable to the correlation of HOA and COA (e.g., $R_{s,HOA,NO_x} = 0.68$, $n=9$, $p<0.05$, $R_{s,COA,NO_x} = 0.68$, $n=9$, $p<0.05$).. ...”

Page 13, line 1: The “uncertainties” here should be relative uncertainties. This should be addressed explicitly.

In Figure 7, we present the uncertainties relative to the factor concentrations. We corrected the mistake in the manuscript:

“... We note that relative uncertainties related with SOOA increase with decreasing concentrations (Fig. 7). A small error in modelling sources with high contributions (BBOA, WOOA) in winter can result in a large error of SOOA with its small contribution during winter. Furthermore, some other sources like primary biological OA (PBOA, see Sec. 4.2.2) might also mix into SOOA. ...”

Page 13, lines 2-3: The meaning of “contribution from other more significant wintertime sources” is not clear. Further, justification of the explanation in this sentence should be provided.

This sentence has been adapted and reads now:

“... A small error in modelling sources with high contributions (BBOA, WOOA) in winter can result in a large error of SOOA with its small contribution during winter. ...”

Page 13, lines 3-4: It is not clear why the mixing of some winter-time SOA into SOOA results in a larger uncertainty.

SOOA concentrations are relatively small in winter compared to BBOA or WOOA concentrations. Therefore, a small error in modelling BBOA or WOOA can result in a rather big error in SOOA. The corresponding section in the manuscript was adapted:

“... We note that relative uncertainties related with SOOA increase with decreasing concentrations (Fig. 7). A small error in modelling sources with high contributions (BBOA, WOOA) in winter can result in a large error of SOOA with its small contribution during winter. Furthermore, some other sources like primary biological OA (PBOA, see Sec. 4.2.2) might also mix into SOOA. ...”

Page 14, lines 1-2: How was σ_b calculated?

12 samples are present in all PMF datasets (PMF_{block}, PMF_{zue,iso}, PMF_{1filt/month}, PMF_{zue,reprs}). For these 12 samples we determine the median concentration for the independently treated PMF datasets. σ_b is the variability of the median concentrations of these 12 samples. This information can be found in the supplementary information (section uncertainty estimation and propagation).

We adapted the main text to the following:

“... The variability of the factor time series for the single PMF sensitivity tests (PMF_{block}, PMF_{zue,isol}, PMF_{1filter/month}, PMF_{zue,eps}) is used as an uncertainty estimate (shaded area in Fig. 4). This estimate (σ_a) depends on the measurement repeatability (10 single mass spectra included for each sample) and on the selected PMF solution/ R_k combinations and, therefore, also on the α -value. However, the variability depending (1) on the choice of input points (time and site; PMF_{block}, PMF_{zue,isol}, PMF_{1filter/month}) and (2) on the instrumental reproducibility (PMF_{zue,reprs}) of the offline AMS measurements is not accounted for. The contribution of (1) and (2) to the uncertainty is assessed through the sensitivity tests by examining the variability of the median factor time-series (σ_b). σ_b is the variability of the median factor concentrations from the PMF sensitivity tests using PMF_{block}, PMF_{zue,isol}, PMF_{1filter/month}, PMF_{zue,reprs} for the 12 samples common to all 4 PMF datasets. For the 12 filters common in all PMF datasets (PMF_{block}, PMF_{zue,isol}, PMF_{1filter/month}, PMF_{zue,reprs}), we calculate a best estimate of the overall uncertainty (err_{tot}), by propagating both error terms: σ_a and σ_b”

Page 14, line 8: The meaning of “ σ_b – rotational ambiguity” is not clear.

The part “- rotational ambiguity” has been removed and more information has been added to the text:

“... It is worthwhile to note that for major factors exhibiting a similar seasonality, i.e. WOOA and BBOA, a great part of the uncertainty arises from σ_b . Thus the variability between the PMF solutions using PMF_{block}, PMF_{zue,isol}, PMF_{1filter/month}, PMF_{zue,reprs} (σ_b) and, therefore, the sensitivity of the factor concentrations on the chosen PMF dataset significantly contribute to the uncertainty. ...”

Page 14, lines 12-15: This sentence is not very organized and needs to be reworded.

We reworded the sentence, now it reads:

“... In the present analysis, PBOA could not be separated by PMF (neither unconstrained nor using the mass spectral signature from Bozzetti et al., 2016). This inability might be caused by the low water-solubility and the absence of PM2.5 filters in the dataset. ...”

Page 14, lines 24-26: What is the definition of the site-to-site variability? Was standard deviation calculated for the average values at respective sites?

The site-to-site variability is the standard deviation of average concentrations at the different sites. This information has been added to the text and the sentence has been reworded for easier readability:

“...Using this approach, we estimate that PBOA contributes $0.30 \mu\text{g}/\text{m}^3$ during the warm months (site-to-site variability computed as standard deviation of the average concentration of all sites of $0.03 \mu\text{g}/\text{m}^3$). During the same period, SOOA concentrations are $1.78 \mu\text{g}/\text{m}^3$ (site-to-site variability of $0.18 \mu\text{g}/\text{m}^3$) and OA concentrations $4.32 \mu\text{g}/\text{m}^3$ (site-to-site variability of $0.44 \mu\text{g}/\text{m}^3$). ...”

Page 15, line 2: The use of the word “however” does not seem appropriate.

The word “however” has been removed and the sentence reworded. Now it reads:

“...The ion $C_2H_5O_2^+$ (indicator for PBOA) shows higher concentrations with increasing OC_{coarse} concentrations. ...”

Page 16, line 5: Is POA here the sum of HOA, COA and BBOA? Shouldn't it be defined here instead of line 9?

We moved the definition of POA to line 9. Now the paragraph reads:

“...In general, the seasonality of the factor time series is consistent for all the 9 sites in the entire study area (Fig. 9). In summer, SOOA is the main contributor to OA, while in winter POA (HOA+COA+BBOA) becomes more important although WOOA still contributes significantly. In comparison to the sites in northern Switzerland, OA in the southern alpine valleys is dominated by BBOA in winter, while in the north WOOA also plays a role. The different factors contribute 0.47 ± 0.12 (HOA, average and site-to-site variability), 0.31 ± 0.13 (COA), 1.37 ± 1.77 (BBOA), 0.67 ± 0.31 (SC-OA), 1.11 ± 0.23 (WOOA), 1.31 ± 0.13 (SOOA) $\mu g/m^3$ for all sites during the entire year (Table 3). In northern Switzerland, POA contributes less to OA (POA/OA=0.3) than in the southern alpine valleys where POA/OA is equal to 0.6. ...”

Figure 2: The aHOA and aCOA are not defined explicitly.

The figure caption was adapted to:

“...**Figure 1: Step-by-step outline of adopted source apportionment approach (factor recoveries R_k). a_{HOA} and a_{COA} represent the a-value applied for HOA and COA, respectively. ...”**

Figure 3: The definition of fm/z should be given.

The definition of fm/z has been added to the figure caption:

“...**Figure 2: PMF factor profiles of HOA, COA, BBOA, SOOA, WOOA, SC-OA, color-coded with ion family of PMF_{block} (average). fm/z is the relative intensity at a specific mass-to-charge ratio (m/z). ...”**

Figure 9: The definition of OAexpl is not given explicitly.

The figure caption has been adapted to:

“...**Figure 3: Map of Switzerland with yearly cycles. Negative concentrations were set to 0 prior to normalization for display. The OA mass explained by the source apportionment analysis is termed OA_{expl}”**

Page 2 (supplement): The relationship among “ $Q_i/Q_{i;exp}$ ”, “ $\Delta (Q_i/Q_{i;exp})$ ”, “ $\Delta Q_i/Q_{i;exp}$ ”, “ $(Q_i/Q_{i;exp}$ contribution)”, and “ $\Delta (Q_i/Q_{i;exp}$ contribution)” is not clear.

The nomenclature has been unified. The results refer to the distributions of $Q_{i,j}$ –(median and quartiles or average) and not to Q/Q_{exp} . $Q_{i,j}$ (referred to as Q-contribution) is computed as:

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

We corrected the text and axis labels accordingly and refer only to Q-contribution.

“... $\Delta(Q\text{-contribution})$ shows the difference in the median Q-contributions between groups of points like sites or season. The smaller the $\Delta(Q\text{-contribution})$, the smaller are the differences in the mathematical quality of the PMF solution for the different seasons/sites. ...”

Page 3 (supplement): The definitions of “r(. . .)”, “Q25(OCres)” and “Q75(OCres)” are not given.

The table caption has been modified as:

“...**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. ...**”

Page 4 (supplement): The definition of “fion” is not given.

The figure caption has been updated:

“...**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.**”

Technical corrections:

Page 1, line 21: Should “at” be added between “10_m” and “9 stations”?

This mistake has been corrected.

Page 3, line 20: “HiVol” should be spelled out.

Spelled out as High-Volume samplers.

Page 5, lines 29-30: The subscripts of “PMF” are not written consistently in the paper.

The PMF datasets are now consistently called:

PMF_{block}

PMF_{zue,isol}

PMF_{1filter/month}

PMF_{zue,reprs}

Page 10, line 18: Should “from” be added after “profile”?

The missing word has been added as well as the number corrected:

“...COA profile elements were constrained using the COA profile from Crippa et al. (2013b) and the obtained factor profile maintains the same features (OM/OC of 1.32, IQR 1.30-1.33, Fig. 3). ...”

Page 13, line 29: Should “Fig. 5” be “Fig. 4”?

Yes it should be Fig. 4. The mistake has been corrected.

Page 14, lines 30 and 33: Should “is in summer” be “in summer is”?

The mistake has been corrected.

Page 14, line 33: “OCcoarse” should be defined in line 30..

OCcoarse is now defined on line 30.

“... Bozzetti et al. (2016) showed that coarse OC ($OC_{\text{coarse}} = OC_{\text{PM}_{10}} - OC_{\text{PM}_{2.5}}$) in summer is dominated by PBOA for samples collected at a rural site in Switzerland (Payerne). ...”

Page 19, lines 9-10: The list of authors are incorrect.

The citation has been corrected:

“Daellenbach, K. R., Bozzetti, C., Křepelová, A., Canonaco, F., Wolf, R., Zotter, P., Fermo, P., Crippa, M., Slowik, J. G., Sosedova, Y., Zhang, Y., Huang, R.-J., Poulain, L., Szidat, S., Baltensperger, U., El Haddad, I., and Prévôt, A. S. H.: Characterization and source apportionment of organic aerosol using offline aerosol mass spectrometry, *Atmos. Meas. Tech.*, 9, 23-39, doi:10.5194/amt-9-23-2016, 2016.”

Table 1: The commas after “St. Gallen” and “San Vittore” in the column “Site (station code)”, and the periods after “m” in the column “altitude” should be omitted. The initial letter of “altitude” should be capitalized.

In the column further information on the station location have been added. Altitude has been capitalized and the periods are omitted.

Site (station code)	Classification	General location	Altitude
Basel, St. Johann (bas)	Urban/background	North of Alps/Swiss plateau	308 m
Bern, Bollwerk (ber)	Urban/traffic	North of Alps/Swiss plateau	506 m
Frauenfeld, Bahnhofstr. (fra)	Suburban/backgroun d	North of Alps/Swiss plateau	403 m
Payerne (pay)	Rural/background	North of Alps/Swiss plateau	539 m
St. Gallen, Rorschacherstr. (gal)	Urban/traffic	North of Alps/Swiss plateau	457 m
Zurich, Kaserne (zue)	Urban/background	North of Alps/Swiss plateau	457 m
Vaduz, Austrasse (vad)	Urban/traffic	North of Alps/alpine valley	706 m
Magadino, Cadenazzo (mag)	Rural/background	South of Alps	254 m
San Vittore, Zentrum (vi)	Rural/traffic	South of Alps/alpine valley	330 m

Figure 4 caption: It is better to write “HOA, COA, . . .” in the order of the corresponding panels.

The figure caption has been adapted and reads now:

“...Figure 4: HOA, COA, BBOA, WOOA, SOOA, and SC-OA and respective marker concentrations as a function of time for Zurich in 2013. Depicted are the median factor time series results for the different PMF datasets (median) including the uncertainties for PMF_{block} (first and third quartile) (green: PMF_{block} , black: $PMF_{zue,isol}$, red: $PMF_{zue,repr}$, pink bullets: $PMF_{1filter/month}$). ...”

Figure 5 caption: Should “NH4” be “NH4+”?

The figure caption has been corrected:

“...Figure 5: Scatter-plots for the different extreme sensitivity tests for Zurich and for all sites for PMF_{block} median concentrations): a) HOA vs NO_x , b) BBOA vs levoglucosan, c) SOOA vs temperature, d) WOOA vs NH_4^+”

Figure 7: Should “[” after “concentration” be “]”?

In the y-axis |concentration| refers to the absolute concentration.

Page 4 (supplement): The “interquartile range PMF block” should be represented by a symbol because it is in a mathematical formula. It may be better to write “median bootstrap solutions” as the subscript of σ .

Since σ_a and σ_b are explained in detail in the text we remove this part of the mathematical expression. The equation reads now:

“ ...

$$\text{err}_{i,k,\text{tot}} = \sqrt{\sigma_a^2 + \sigma_b^2}$$

...”

References:

Allan, J. D., Jimenez, J. L., Williams, P. I., Alfarra, M. R., Bower, K. N., Jayne, J. T., Coe, H., and Worsnop, D. R.: Quantitative sampling using an Aerodyne aerosol mass spectrometer 1. Techniques of data interpretation and error analysis, *J. Geophys. Res.*, 108, 4090, doi:10.1029/2002JD002358, 2003.

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., 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.

Bozzetti, C., Sosedova, Y., Xiao, M., Daellenbach, K. R., Ulevicius, V., Dudoitis, V., Mordas, G., Byčenkienė, S., Plauškaitė, K., Vlachou, A., Golly, B., Chazeau, B., Besombes, J.-L., Baltensperger, U., Jaffrezo, J.-L., Slowik, J. G., El Haddad, I., and Prévôt, A. S. H.: Argon offline-AMS source apportionment of organic aerosol over yearly cycles for an urban, rural and marine site in Northern Europe, *Atmos. Chem. Phys. Discuss.*, doi:10.5194/acp-2016-413, 2017a.

Bozzetti, C., El Haddad, I., Salameh, D., Daellenbach, K. R., Fermo, P., Gonzalez, R., Minguillón, M. C., Iinuma, Y., Poulain, L., Müller, E., Slowik, J. G., Jaffrezo, J.-L., Baltensperger, U., Marchand, N., and Prévôt, A. S. H.: Organic aerosol source apportionment by offline-AMS over a full year in Marseille, *Atmos. Chem. Phys.*, 17, 8247–8268, <https://doi.org/10.5194/acp-17-8247-2017>, 2017b.

Chen, Q., Miyazaki, Y., Kawamura, K., Matsumoto, K., Coburn, S. C., Volkamer, R., Iwamoto, Y., Kagami, S., Deng, Y., Ogawa, S., Ramasamy, S., Kato, S., Ida, A., Kajii, Y., and Mochida, M.: Characterization of chromophoric water-soluble organic matter in urban, forest, and marine aerosols by HR-ToF-AMS analysis and excitation emission matrix spectroscopy, *Environ. Sci. Technol.*, 50, 351–360, 2016.

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.

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.

Sun, Y., Zhang, Q., Zheng, M., Ding, X., Edgerton, E. S., and Wang, X.: Characterization and source apportionment of water-soluble organic matter in atmospheric fine particles (PM_{2.5}) with high-resolution aerosol mass spectrometry and GC-MS, *Environ. Sci. Technol.*, 45, 4854–4861, 2011.

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

Kaspar R. Daellenbach¹, Giulia Stefenelli¹, Carlo Bozzetti¹, Athanasia Vlachou¹, Paola Fermo², Raquel Gonzalez², Andrea Piazzalunga^{3,*}, Cristina Colombi⁴, Francesco Canonaco¹, Christoph Hueglin⁵, Anne Kasper-Giebl⁶, Jean-Luc Jaffrezo⁷, Federico Bianchi^{1,**}, Jay G. Slowik¹, Urs Baltensperger¹, Imad El-Haddad¹, and André S. H. Prévôt¹

¹Paul Scherrer Institute (PSI), 5232 Villigen-PSI, Switzerland

²Università degli Studi di Milano, 20133 Milano, Italy

10 ³Università degli Studi di Milano-Bicocca, 20126 Milano, Italy

⁴ARPA Lombardia, Regional Centre for Air Quality Monitoring, 20122 Milan, Italy

⁵Swiss Federal Laboratories for Materials Science and Technology, 8600 Dübendorf, Switzerland.

⁶Institute of Chemical Technologies and Analytics, Vienna University of Technology, 1060 Wien, Austria

⁷Université Grenoble Alpes, CNRS, IGE, 38000 Grenoble, France

15 ^{*}now at: Water and Soil Lab, 24060 Entratico, Italy

^{**}now at: Department of Physics, University of Helsinki, 00014 Helsinki, Finland

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

Abstract. Long-term monitoring of the organic aerosol is important for epidemiological studies, validation of atmospheric models, and air quality management. In this study, we apply a recently developed filter-based offline methodology of the aerosol mass spectrometer to investigate the regional and seasonal differences of contributing organic aerosol sources. We present offline-AMS measurements for particulate matter smaller than 10 μm at 9 stations in central Europe with different exposure characteristics for the entire year of 2013 (819 samples). The focus of this study is a detailed source apportionment analysis (using PMF) including in-depth assessment of the related uncertainties. Primary organic aerosol (POA) is separated in three components: hydrocarbon-like OA-~~which is~~, related to traffic emissions (HOA), cooking OA (COA), and biomass-burning OA (BBOA). We observe enhanced production of secondary organic aerosol (SOA) in summer, following the increase in biogenic emissions with temperature (summer oxygenated OA, SOOA). In addition, a SOA component was extracted that correlated with anthropogenic secondary inorganic species which is dominant in winter (winter oxygenated OA, WOOA). A factor (SC-OA) explaining sulfur-containing fragments (CH_3SO_2^+), which has an event-driven temporal behavior, was also identified. The relative yearly average factor contributions range for HOA from ~~34~~ to ~~45~~14%, for COA from 3 to ~~34~~11%, for BBOA from 11 to ~~64~~59%, for SC-OA from 5 to 23%, for WOOA from 14 to ~~28~~27%, and for SOOA from ~~44~~15 to ~~40~~38%. The uncertainty of the relative average factor contribution lies between ~~5~~2 and ~~9~~12% of OA. At the sites north of the alpine crest, the sum of HOA, COA, and BBOA (POA) contributes less to OA (POA/OA=0.3) than at the southern alpine valley sites (0.6). BBOA is the main contributor to POA with ~~88~~87% in alpine valleys and ~~43~~42% north of

the alpine crest. Furthermore, the influence of primary biological particles (PBOA), not resolved by PMF, is estimated and could contribute significantly to OA in PM10.

Formatted: Font color: Auto

1 Introduction

The development and field deployment of the Aerodyne aerosol mass spectrometer (AMS, Canagaratna et al., 2007) have greatly improved air quality monitoring by providing real-time measurements of the non-refractory (NR) submicron aerosol (PM1) components. The application of factor analysis on the collected organic aerosol (OA) mass spectra enabled the efficient disentanglement of aerosol factors, which could be subsequently related to specific aerosol sources and processes (Lanz et al., 2007, 2010; Jimenez et al., 2009; Ulbrich et al., 2009, Zhang et al., 2011; Ng et al., 2010; Crippa et al., 2014). Factors typically extracted include directly emitted primary OA (POA) from biomass burning (BBOA) or traffic (HOA), oxygenated OA (OOA) that is typically associated with secondary OA (SOA), formed through the oxidation of organic vapor precursors or heterogeneous processes. The model is not capable of identifying the main SOA precursors, but often differentiates OOA based on its volatility and degree of oxygenation (semi-volatile fraction: SV-OOA and low-volatility fraction: LV-OOA) due to the available highly time-resolved data.

However, the cost and operational requirements of the AMS make its deployment impractical throughout a dense monitoring network and over longer time periods. As a result, most available datasets are often limited to few weeks of measurements and factors are extracted mainly based on diurnal variations in POA emission strength and SOA oxygen content (Zhang et al., 2011; El Haddad et al., 2013). Highly mobile measurements on platforms as aircrafts (e.g. DeCarlo et al., 2008) or vehicles (e.g. Mohr et al., 2011) are designed for regional studies, but are even more limited by cost, availability and time than stationary studies. This hinders the determination of the aerosol regional and seasonal characteristics and evaluation of long-term emission trends, limiting the information required for model validation and development of efficient mitigation strategies. Furthermore, the negligible transmission efficiency of the AMS inlet for coarse particles, prevents the characterization of their chemical nature and contributing sources.

The recent development of the aerosol chemical speciation monitor (ACSM, Ng et al., 2011, Fröhlich et al., 2013) has enabled the establishment of dense networks of long-term AMS-type measurements and source apportionment of the organic aerosol (e.g. Crippa et al., 2014 using AMS for shorter campaigns within the EUCAARI project or EMEP/ACTRIS projects for longer multi-season campaigns using ACSM). However, the mass spectrometers used by the ACSMs have far lower mass resolution than the AMS, reducing their performance for OA characterization and source apportionment. An alternate monitoring strategy involves extending AMS spatial and temporal coverage by measuring the nebulized water extracts of filter samples (Daellenbach et al., 2016, ~~restricted to WSOA in~~ Mihara et al., 2011). This approach allows the retroactive investigation of specific events, e.g. haze events in China (Huang et al., 2014) ~~and as well as~~ AMS measurements of coarse

mode aerosol (Bozzetti et al., 2016a, b, 2016) and long-term source apportionment studies (Bozzetti, 2017a, 2017b). Such an approach was also used in ~~prep~~-recent studies for identifying the different types of water-soluble chromophores (Chen et al., 2016). Additionally, such ~~offline~~ filters are routinely collected and are already available over multi-year periods at many air quality monitoring stations around the world for years/decades. Unlike single-season online AMS studies, the offline AMS analysis of filter samples may reveal seasonal and long-term variations in the emissions of POA and SOA precursors, required for model validations and the establishment of efficient mitigation strategies.

Here, we present offline AMS measurements of PM₁₀ (particulate matter with an aerodynamic diameter smaller than 10 μm) at 9 stations in central Europe with different exposure characteristics for the entire year of 2013 (819 samples). The sites cover rural and urban locations, including urban background and traffic and wood-burning influenced stations. Such long-term multi-site analyses allow the quantitative description of the temporal and spatial variability of the main OA sources and may provide further insights into SOA precursors and formation pathways. This paper focuses on the identification of the main factors influencing the OA concentrations at the different sites and the assessment of the associated uncertainties without analysing the other aerosol components in detail. In the second ~~part~~paper, we will investigate the site-to-site differences and general trends in the factor time series and their relationship with external parameters.

2 Methods

2.1 Study area and aerosol sampling

PM₁₀ samples were collected at 9 sites in Switzerland and Liechtenstein (Tab. 1 and Fig. 1). 7 of the sites (Basel, Bern, Payerne, Zurich, Frauenfeld, St. Gallen, Vaduz) are located in northern Switzerland and Liechtenstein and 2 (Magadino and San Vittore) in southern Switzerland. Aerosol was sampled at the selected sites every 4th day for 24h throughout the year 2013 onto quartz fibre filters (14.7 cm) using HiVol/High-Volume samplers (500 l min⁻¹). Filters were then wrapped in aluminium foil or lint-free paper and stored at -20°C. Field blanks were collected following the same approach.

2.2 Offline AMS analysis

The offline AMS analysis summarized below was carried out following the methodology developed by Daellenbach et al. (2016). For each analyzed filter sample, 4 16 mm diameter filter punches were sonicated together in 10 mL ultrapure water (18.2MΩcm, total organic carbon TOC/ < 5 ppb, 25 °C) for 20 min at 30 °C. Liquid extracts were then filtered (0.45 μm) and nebulized in synthetic air (80% Vol N₂, 20% Vol O₂, Carbagas, Gümligen CH-3073 Switzerland) using a customized Apex Q nebulizer (Elemental Scientific Inc., Omaha 24 NE 68131 USA) operating at 60°C. The resulting droplets were dried using a Nafion® dryer, and then injected and analyzed by the HR-ToF-AMS. Three types of measurements were performed: (i) filter samples, (ii) field blanks (collected and treated in the same way as the exposed filters), and (iii) measurement blanks (nebulized ultrapure water without filter extract). The measurement blank was determined before and

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

after every filter sample ~~or field blank~~. Each sample was recorded for 480 seconds (AMS V-mode, m/z 12-447), with a collection time for each spectrum of 30 seconds. Ultrapure water was measured for 720 seconds. Once per day, ultrapure milliQ water was nebulized with a particle filter interposed between the nebulizer and the AMS, ~~providing for the determination of~~ the gas-phase contribution to the measured mass spectrum, which was then subtracted during analysis from both blanks and filter samples. The filters from Zurich were analysed twice with a time difference of approximately 5 months to assess the measurement repeatability. High resolution mass spectral analysis was performed for each m/z (mass to charge) in the range of 12- 115. ~~The measurement blank was subtracted from the sample spectra. In a previous study, it has been shown that the measurement blank is comparable to the organic blanks obtained from the nebulization of NH_4NO_3 (Bozzetti et al. (2017a)).~~ The interference of NH_4NO_3 on the CO_2^+ signal described by Pieber et al. (2016) was corrected as follows (Eq. 1):

$$CO_{2,real} = CO_{2,meas} - \left(\frac{CO_{2,meas}}{NO_{3,meas}} \right)_{NH_4NO_3,pure} * NO_{3,meas} \quad (1)$$

The correction factor $\left(\frac{CO_{2,meas}}{NO_{3,meas}} \right)_{NH_4NO_3,pure}$ was determined based on measurements of aqueous NH_4NO_3 conducted regularly during the entire measurement period ~~and varied between ~1% and ~5% (Pieber et al., 2016, from~~ $\left(\frac{CO_{2,meas}}{NO_{3,meas}} \right)_{NH_4NO_3,pure} < 1\%$ ~~and up to approximately 5%);~~.

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

2.3 Other chemical analysis

Organic and elemental carbon (OC, EC) content were measured by a thermo-optical transmission method with a Sunset OC/EC analyzer (Birch and Cary, 1996), following the EUSAAR-2 thermal-optical transmission protocol (Cavalli et al., 2010). Water-soluble carbon was measured by water-extraction followed by catalytic oxidation, non-dispersive infrared detection of CO_2 using a total organic carbon analyser, only for the samples from Magadino and Zurich. Water-soluble ions (K^+ , Na^+ , Mg^{2+} , Ca^{2+} , NH_4^+ ; and SO_4^{2-} , NO_3^- , Cl^-) and methane sulfonic acid were analyzed using ion chromatography (Piazzalunga et al., 2013 and Jaffrezo et al., 1998). Levoglucosan measurements were performed by a high-performance anion exchange chromatography (HPAEC) with pulsed amperometric detection (PAD) using an ion chromatograph (Dionex ICS1000) following Piazzalunga et al. (2010 and 2013a). Free cellulose was determined by an enzymatic conversion to D-glucose (Kunit and Puxbaum, 1996) and subsequent determination of glucose with HPAEC (Iinuma et al., 2009). Online measurements of gas-phase compounds and meteorology were also performed at selected sites.

3 Source apportionment

3.1 General principle

Source apportionment of the organic aerosol is performed using positive matrix factorization (PMF, Paatero, 1994). PMF is a statistical un-mixing model explaining the variability of the organic mass spectral data ($x_{i,j}$), as linear combinations of static factor profiles ($f_{j,k}$) and their time-dependent contributions ($g_{i,k}$), see Eq. 2 (where p is the number of factors). The index i represents a specific point in time, j an ion, and k a factor. The elements of the model residual matrix are termed $e_{i,j}$.

$$x_{i,j} = \sum_{k=1}^p g_{i,k} f_{k,j} + e_{i,j} \quad (2)$$

In the input data matrix, each filter sample was represented on average by 11 mass spectral repetitions to examine the influence of the AMS measurement repeatability on the PMF outputs. A corresponding preceding measured blank from nebulized ultrapure water was subtracted from each mass spectrum. The input errors $s_{i,j}$ required for the weighted least-squares minimization by the model include consist of the blank variability ($\sigma_{i,j}$) and the uncertainty related to ion counting statistics and ion-to-ion signal variability at the detector ($\delta_{i,j}$ Allan et al., 2003; Ulbrich et al., 2009). We applied a minimum error according to Ulbrich et al. (2009), and a down-weighting factor of 3 to all fragments with an average signal to noise lower than 2 (Ulbrich et al., 2009). Input data and error matrices included consisted of 202 organic ions and were rescaled by the organic fragments, $x'_{i,j}$, obtained from offline AMS analyses do not directly represent ambient concentrations. Therefore, the signal of each fragment was converted to such an ambient concentration ($x_{i,j}$ in $\mu\text{g m}^{-3}$), by multiplying the fraction of this signal with the estimated organic matter (OM) concentration. The latter was calculated as the product of the OC concentrations measured by the sunsetSunset OC/EC analyzer and the OM/OC ratios from the offline AMS measurements— $(\text{OM}/\text{OC})_{\text{AMS}}$ (Eq. 3). Note that such scaling does not change the outcome of Eq. 2 since both data and error matrices are scaled in the same manner and the fingerprints ($f_{k,j}$) are not changed.

$$x_{i,j} = \frac{x'_{i,j}}{\sum_i x'_{i,j}} * \text{OC} * (\text{OM}/\text{OC})_{\text{AMS}} \quad (3)$$

The Source Finder toolkit (SoFi v.4.9, Canonaco et al., 2013) for Igor Pro software package (Wavemetrics, Inc., Portland, OR, USA) was used to configure the PMF model and for post-analysis. The PMF algorithm was solved using the multilinear engine-2 (ME-2, Paatero, 1999), which Normalization of the PMF solution during the iterative minimization process is disabled as implemented in SoFi (Canonaco et al., 2013). ME-2 enables an efficient exploration of the solution space by a priori constraining the $f_{k,j}$ elements (n and m being any two realizations of j) within a certain range defined by the scalar a ($0 \leq a \leq 1$) from a starting value $f_{k,j}'$, such that the modelled $f_{k,j}$ in the solution satisfy Eq. 34:

$$\frac{(1-a)f_{k,m}}{(1+a)f_{k,m}} \leq \frac{f_{k,m}'}{f_{k,m}'} \leq \frac{(1-a)f_{k,m}}{(1+a)f_{k,m}} \quad (3)$$

$$f_{k,j} = f'_{k,j} + a * f_{k,j} \quad (4)$$

$f'_{k,j}$ is the starting value used as a priori knowledge from previous studies and $f_{k,j}$ is the resulting value in the solution. In all PMF runs (unless mentioned otherwise), we used the high resolution mass spectra for HOA and COA (cooking OA) from Crippa et al., (2013b) as constraints. ~~Fitted ions, i.e. two rows of $f_{k,j}$, were set equal to the mass spectra of HOA and COA. Ions that were present~~ in our datasets ~~missing but not~~ in the reference profiles ~~for HOA and COA~~ were inferred from published unit mass resolution (UMR) profiles (Ng et al., 2011 and Crippa et al., 2013c). ~~The relative intensity allocated to a missing ion was~~ For this purpose, the ~~intensity fraction of signal at a specific m/z in the corresponding UMR peak minus reference spectrum ($f_{UMR,m/z}$) was compared to the intensities fraction of signal~~ of all ~~HR ions present at this m/z~~ in the HR reference ~~profile spectrum ($f_{HR,m/z}$). The difference $f_{UMR,m/z} - f_{HR,m/z}$ was used as entries in $f'_{k,j}$ for such missing ions. For these ions, an a -value of unity was set. For the other factors, the factor elements were fitted by ME-2. Alternatively, such missing ions can be also treated as ordinary factor elements, to be fitted by ME-2 with all other ordinary factor elements.~~

Formatted: Font: +Headings (Times New Roman)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

15 Source apportionment analysis was performed following the scheme shown in Fig. 2 and discussed below. Exploratory unconstrained and constrained PMF runs provided information on the number of interpretable factors (Sec. 3.2). Multiple constrained PMF runs were then performed, to assess the model sensitivity to the chosen a -value, the model starting point and input matrix (entire dataset: PMF_{block} , only Zurich: $PMF_{zue,isol}$, 1 filter per site and month: $PMF_{filter,month}$, repeated measurements for Zurich: $PMF_{zue,repr}$) and repeated measurements (Sec. 3.3). The obtained factors were then classified and corrected for their recovery (Sec 3.4 and 3.5). Finally, the different solutions were evaluated and only the solutions that satisfied a set of predefined criteria (Sec. 3.6 and supplement) were considered.

3.2 Preliminary PMF

We explored constrained PMF solutions, ranging from 1-10 factors. This investigation is performed on the entire dataset, including all stations and seasons (details in SI). The impact of the number of factors on the residuals is examined in the supplement. The introduction of 2 factors, in addition to HOA and COA, resulted in a significant reduction in the residuals and the separation of BBOA and OOA contribution. BBOA exhibited a prominent seasonal variation with a significant increase during winter and contributed most to the explained variation of the fragment $C_2H_4O_2^+$, originating from the decomposition of anhydrous sugars, i.e., from cellulose pyrolysis. OOA was identified based on its mass spectral fingerprint, with high contribution from oxygenated ions at m/z 43 and 44. A further increase in the number of factors did not significantly contribute to the reduction in the residuals. However, the introduction of a 5th factor allowed the separation of the OOA into two different factors, with distinct seasonal variability and different relative contributions from oxygenated

fragments at m/z 43 and 44. The two OOA factors will be referred to as winter and summer OOA (WOOA and SOOA), according to their seasonality. The introduction of a 6th factor allowed resolving a factor with a distinct time series explaining the variability of sulfur-containing fragments (e.g. CH_3SO_2^+). This factor will be referred to as sulfur containing organic aerosol (SC-OA). We explored higher order solutions, but could not interpret the resulting factor separations. Therefore, we further consider a 6-factor solution below.

3.3 Sensitivity analysis

We assessed the model sensitivity to the chosen a -value for HOA and COA, the model starting point (independently for all four PMF inputs, as described below). The a -values were independently varied for HOA and COA (a -value from 0 to 1 with increments of 0.1, giving 121 a -value combinations). For every a -value combination, the model was initiated from five different pseudo-random starting points (seeds), yielding 605 total runs. As the selection of the a -value combination was randomized, the process was repeated four times in order to ensure that every a -value combination was represented at least once (2420 runs), which in turn provided an assessment of the seed effect on the results.

While this approach has been proven very effective in selecting a range of environmentally relevant solutions (Elser et al., 2016a, 2016b and Daellenbach et al., 2016), the resulting ~~modelling errors~~~~uncertainties~~ may be underestimated. Paatero et al. (2014) ~~compare~~~~compared~~ the effectiveness in estimating ~~modelling errors~~~~uncertainties of factor elements~~ using two different approaches: the displacement (DISP) and bootstrap analysis (BS) ~~, respectively~~. BS involves applying the model to input matrices consisting of a subset of the entire dataset. DISP involves running PMF several times using ~~randomly~~~~systematically~~ perturbed factor profile elements of a reference solution, but allowing a defined difference in Q from the reference solution. Both approaches are computationally intensive, especially DISP. Because of such computational limitations the combination of BS and DISP was not feasible for the dataset presented here, especially in combination with a -value sensitivity tests. Therefore, we chose to perform 4 sensitivity tests performing PMF runs using 4 different input datasets, presented in the following. These sensitivity tests allow conclusions on the stability of PMF analysis when reducing the temporal or spatial resolution as well as the influence of the measurement repeatability.

1. PMF_{block}: PMF was performed on data from all seasons and all sites combined (all measured in October 2014). The corresponding data and error matrices involved 819 samples from 9 sites with 202 ions and on average 11 spectra per sample. This represents the base case.
2. PMF_{zue,isol}: PMF was performed on data from Zurich alone (isolated from PMF_{block} input). The corresponding data and error matrices involved 91 samples with 202 ions and on average 11 spectra per sample.

3. ~~PMF_{fin}~~PMF_{filter/month}: PMF was performed on data from all sites but only considering the 1st filter collected for every month (12 filters/site), as for these samples levoglucosan and cellulose data was available. The corresponding data and error matrices involved 108 samples with 202 ions and on average 11 spectra per sample.
4. PMF_{zurc.reps}: PMF was performed on data from the repeated measurements of Zurich samples. The corresponding data and error matrices comprised 91 samples with 196 ions and on average 14 spectra per sample.

For each of the four PMF datasets, 2420 PMF runs were performed for evaluating the sensitivity of the model to the chosen α -value and the seed. The quality of each of the 2420 PMF runs was individually assessed using criteria lined out in Sec. 3.6.

3.4 Factor classification

From the sensitivity analysis, a large number of solutions were generated. Systematic analysis of these solutions required automatic identification/classification of the retrieved factors within each solution. We applied a sequential classification algorithm as follows. ~~The identity of~~ Since HOA and COA were ~~identified first as their mass spectra were~~ initially constrained on preselected rows of F, they did not need to be identified. In a second step, the factor showing the highest explained variation for C₂H₄O₂⁺ among the 4 remaining factors was identified as BBOA. In a third step, the factor with the highest explained variation for CH₃SO₂⁺ among the 3 remaining factors was identified as SC-OA. From the last two factors, the one with the highest explained variation of CO₂⁺ was identified as WOOA and the other as SOOA.

3.5 Recovery and blank corrections

After factor identification, factor time series are corrected using factor-specific recoveries (Eq. 45, resulting in OA_{*i,k*}) determined in Daellenbach et al. (2016) for HOA, COA, BBOA, and OOA.

$$OA_{i,k} = \frac{\frac{c_{ijk}}{R_k}}{\sum_k \frac{c_{ijk}}{R_k}} \frac{g_{i,k}}{\sum_k g_{i,k}} * OA_i \quad (45)$$

where $g_{i,k}$ are the concentrations of factor k at the timepoint i , R_k the recoveries of the respective factor and OA_{*i*} the OA concentration. ~~The~~ For a limited number of PMF runs (PMF_{block}) also the field blank analyses were included in the PMF input data. This provides the contributions of different factors to the field ~~blank samples were estimated by inserting in the input matrix the field blank mass spectra. This was performed only for a limited number of runs, using the PMF_{block} dataset.~~ The factor contributions to the field blanks which were used to correct the output factor time series, ~~and the uncertainty.~~ Uncertainties induced by the blank subtraction ~~was/were~~ propagated.

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

3.6 Solution selection

Each of the 2420 PMF solutions were per PMF dataset (PMF_{block}, PMF_{zue.isol}, PMF_{zue.reps}, PMF_{ifilter/month}) was evaluated based on their factor profiles, time series, and the OC mass closure. Solutions were selected if they satisfied the following set of criteria:

1. $f_{\text{CO}_2^+} < 0.04$ in HOA and COA factor profiles (HOA based on Aiken et al., 2009, Mohr et al., 2012, Crippa et al., 2013b, 2014 and COA based on Crippa et al., 2013b, 2013c, Mohr et al., 2012);
2. $f_{\text{C}_2\text{H}_4\text{O}_2^+} < 0.004$ and 0.01 in HOA and COA, respectively (HOA based on Aiken et al., 2009, Mohr et al., 2012, Crippa et al., 2013b, 2014 and COA based on Crippa et al., 2013b, 2013c, Mohr et al., 2012);
3. HOA correlates significantly with NOx being the sum of NO and NO₂ (defined below);
4. HOA correlates significantly better with NOx than COA; BBOA correlates significantly with levoglucosan (defined below);
5. SC-OA correlates significantly with CH₃SO₂⁺ (defined below);
6. for samples from Zurich and Magadino, where WSOC data are available, modelled and measured OC mass are comparable for a different set of conditions (see below and in SI).

The first two criteria (1-2) insure an appropriate separation of HOA and COA from OOA and BBOA, respectively. Criteria 3-6 relate to the evaluation of the correlation between factor and marker time series. This was achieved by computing the Fisher-transformed correlation coefficient z at different stations (Eq. 56):

$$z = 0.5 * \ln\left(\frac{1+r}{1-r}\right) = \arctan(r) \quad (6)$$

where r is the correlation coefficient between factor and marker at a given station. The obtained z values at the different stations are subsequently averaged and transformed back to r_{avg} before further analysis. A t-test is then used to verify the significance ($\alpha=0.5$) of the average correlation coefficient between factor and marker time series, t_{avg} (Eq. 67):

$$t_{\text{avg}} = \frac{r_{\text{avg}}}{\sqrt{\frac{1-r_{\text{avg}}^2}{N-2}}} \quad (67)$$

Here, r_{avg} is the correlation coefficient averaged over the different stations, derived from the average z value, t_{avg} is the corresponding t-value and N is the average number of samples at the different stations. Results with significance level $\alpha=0.05$ are summarized in Fig. S8.

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: Font: Not Italic

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: Font: Not Italic, English (U.S.)

Formatted: English (U.S.)

Formatted: Font: Not Italic, English (U.S.)

Formatted: English (U.S.)

Formatted: Font: Not Italic, English (U.S.)

Formatted: English (U.S.)

To evaluate whether HOA correlated significantly better with NO_x than COA did, the average z values obtained between HOA and NO_x and between COA and NO_x (Eq. 6) were compared, using a standard error on the z distribution of $\frac{1}{\sqrt{N}} - \frac{3}{1/\sqrt{N} - 3}$ (Zar, 1999).

The last criterion relates to OC mass closure. A Monte Carlo approach was applied to evaluate whether a combination of water soluble factor time series and recovery parameters would achieve OC mass closure, as described in the following. For the samples from Zurich and Magadino, for which WSOC concentrations were available (in contrast to the other samples), offline AMS measurements were scaled to the water soluble organic matter (WSOM), calculated using the WSOC measurements and OM/OC ratios from the AMS HR analysis. The water-soluble contributions from an identified aerosol source in a sample i were rescaled to ~~its~~their total organic matter ~~concentration~~concentrations ($OA_{i,k}$), where k represents a given factor, using combinations of factor recoveries as determined by Daellenbach et al. (2016). (2016, medians of the used combinations being: R_{HOA} : 0.11, R_{COA} : 0.54, R_{BBOA} : 0.65, and R_{OOA} : 0.89 used for WOOA and SOOA). For SC-OA, whose recovery was not previously determined, a recovery value was stochastically generated between 0 and 1. The obtained $OA_{i,k}$ concentrations were then converted to organic carbon concentrations $OC_{i,k}$, using factor specific OM/OC ratios determined from the factor profiles. The sum of $OC_{i,k}$ from all factors k (mod-OC _{i}) was then evaluated against the measured OC (meas-OC _{i}). For this, the residual OC mass (res-OC _{i}) for each sample ~~is~~was calculated (meas-OC _{i} – mod-OC _{i}), and the residual distributions were examined for different conditions that are ~~thoroughly described in the Supplement~~specified in the Supplement. In summary, a solution was only accepted if res-OC _{i} were normally distributed around 0 considering all points and subsets of points: a) summer, b) winter c) Magadino, d) Zurich, e) low and high concentrations of the single factors (see Table S1).

For each of the Monte Carlo simulations, the water soluble factor time series satisfying criteria 1-6 were used together with a combination of factor recoveries from Daellenbach et al. (2016) as input data. The water soluble OC used for scaling $G_{i,k}$ matrix and the meas-OC _{i} used for residual calculation were varied within their uncertainties (5%) and biases (5%) assuming a normal distribution of the errors. Likewise, constant biases were also introduced into the initial recovery distributions from Daellenbach et al. (2016). Monte Carlo simulations were performed and simulations for which res-OC _{i} distributions were significantly different from 0 ($Q_{25} < 0 < Q_{75}$, details in SI) were discarded until 500 acceptable simulations were found. Thereby, 331 PMF runs were selected for PMF_{block} (230 for PMF_{ue,isol}, 99 for PMF_{ue,repr}, and 269 for PMF_{filter/month}). Median factor time series and recovery parameters from all retained simulations were then averaged and the standard deviation of these averages represent interquartile range (IOR) represents our best estimate of the uncertainties for the single PMF datasets. The Monte Carlo process was repeated for the four different PMF datasets described above and the resulting average median time series of their estimated uncertainties were compared. The resulting uncertainty estimates and the method are described in Sec. 4.2.1. and in the SI.

Formatted: English (U.S.)

Formatted: Font: Not Italic, English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: Font: Not Italic, English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

4 Results and discussions

In this section, the final source apportionment results are presented and validated. The source signatures are presented in Fig. 3 for PMF_{block} color-coded with the ion family. Fig. 4 shows the time series for Zurich obtained from all PMF approaches and Tab. 2 summarizes the correlation coefficients between factor and marker time series for Zurich (all PMF runs) and the other sites in the study area (PMF_{block}) while the relation between factor and marker time series is displayed in Fig. 4 and 5. Presented are median (and quartile) results for all PMF runs accepted following the criteria described above.

4.1 Interpretation of PMF factors

HOA: HOA profile elements were constrained using the reference profile from Crippa et al. (2013b). The final factor profile (Fig. 3) maintains the same features, characterized by high contributions from hydrocarbon fragments. The fraction of oxygenated organic fragments that were missing in the initial reference profile, that were added based on UMR spectra, show an increased contribution to the ions above m/z 100 (see Sec. 3.1). While this indicates a possible overestimation of the contribution of these fragments, using this methodology, this increase does not substantially affect the results: e.g. HOA OM/OC ratio remains low (1.32, IQR 1.30-1.33). The HOA time series follows an expected pattern that matches the NO_x yearly cycle (Fig. 4.a) except for San Vittore which is very likely due to the extremely high contribution of biomass burning at this site during winter, which may result in additional NO_x inputs and/or may affect the separation of HOA by PMF. The HOA/NO_x (Fig. 5.a) ratio at the different sites ($0.014015 \pm 0.014011 \mu\text{g m}^{-3} \text{ppb}^{-1}$) lies within the range of literature values (0.001 to 0.028 ($\mu\text{g}/\text{m}^3$)/ppb, Lanz et al., 2007 and Kirchstetter et al., 1999). A similar average ratio was obtained for Zurich from the different sensitivity tests, but with high variability ($0.012013 \pm 0.009 \mu\text{g m}^{-3} \text{ppb}^{-1}$) similar to that obtained between the different sites. This implies that the observed site-to-site differences are not statistically significant given our uncertainty in extracting HOA contributions.

COA: COA profile elements were constrained using the COA profile from Crippa et al. (2013b) and the obtained factor profile maintains the same features (OM/OC of 1.32, IQR 1.30-1.33, Fig. 3). For COA, no molecular marker is available for validation purposes. Daellenbach et al. (2016) demonstrated that COA concentrations can be estimated with offline AMS (in Zurich at the same site) by constraining its signatures, but only with a high uncertainty. This was performed by comparing offline AMS results to those from a collocated ACSM, which owing to its higher time resolution enabled the identification of cooking emissions based on their diurnal cycles (Canonaco et al., 2013). Here, while no ACSM data were available, we followed the same methodology used in Daellenbach et al. (2016) to estimate the contribution of COA. The average COA contributions estimated here and their yearly variability are similar to those from previous studies at the same sites, but as expected have high uncertainties (Fig. 4.b).

BBOA: BBOA is identified based on its spectral fingerprint (OM/OC 1.74, IQR 1.74-1.75, Fig. 3), which, similar to previously extracted BBOA factors at other locations (Daellenbach et al., 2016, Lanz et al., 2007; Crippa et al., 2014), exhibits high contributions from oxygenated fragments (CHO^+ , $\text{C}_2\text{H}_4\text{O}_2^+$, $\text{C}_3\text{H}_5\text{O}_2^+$) from anhydrous sugars fragmentation (see comparison to nebulized levoglucosan in Supplement Fig. S4S6). Similar to levoglucosan, the BBOA time series shows an expected seasonal variation with high concentrations in winter, supporting the identification of this factor (Fig. 4.c). Except for Bern and Magadino (6.87.5 and 40.911.2), a similar ratio of BBOA to levoglucosan is found at all other sites (3.59 to 5.7), despite apparent site-to-site differences in the model residuals during winter due to significantly higher contributions of BBOA at the southern stations (Fig. 5.b). The obtained ratios are within the range of values reported in literature (between 4 and 18 assuming OM/OC ratios between 1.6 and 1.8 for the non-AMS analyses, Zotter et al., 2014, Herich et al., 2014, Minguillon et al., 2011, Crippa et al., 2013a, and Favez et al., 2010). We note that a similar ratio is also found for the different PMF datasets performed for the case of Zurich (BBOA/levoglucosan between 4.8 and 11.03.9 and 12.1). Taken together, the high (for most sites) correlation ($R^2=0.78$ for all sites, single sites in Tab. 2) between levoglucosan and BBOA and their consistent ratios at different sites and between the different PMF datasets indicates that BBOA is well resolved by PMF at all sites, despite potential site-to-site differences in BBOA composition.

SC-OA: Sulfur-containing fragments (e.g., CH_3SO_2^+) are predominantly apportioned to this factor, which also has a high OM/OC ratio (1.82, IQR 1.80-1.9293, Fig. 3). As mentioned in Sec. 3.6, the recovery of SC-OA was unknown and had to be determined by mass closure, while the recoveries of the other factors were determined by comparison to their online counterparts (albeit for a different dataset, Daellenbach et al., 2016). In the lack of specific constraints (like an online counterpart), the recovery of SC-OA is highly uncertain and thus also the factor time series. A similar factor profile had been extracted from previous online AMS datasets, and was related to the fragmentation of methane sulfonic acid (MSA) present in PM1 particles, a secondary product of marine origin (Crippa et al 2013b, Zorn et al., 2008). However, the SC-OA factor extracted here did not seem to be related to marine emissions because neither its variability nor its levels matched those of MSA (Fig. 4.d). First we compared the MSA levels measured in Zurich using ion chromatography to those estimated based on the concentration of sulfur containing fragments from offline AMS measurements in SC-OA (Eq. 78), based on Crippa et al. (2013b):

$$MSA_{i,est} = SC - OA_i * \frac{f_{SC-OA}(\text{CH}_2\text{SO}_2^+) + f_{SC-OA}(\text{CH}_3\text{SO}_2^+) + f_{SC-OA}(\text{CH}_4\text{SO}_3^+)}{0.147} \quad (78)$$

Here, $MSA_{i,est}$ is the estimated MSA concentration, $SC-OA_i$ the factor concentration of the sulfur-containing factor, $f_{SC-OA}(\text{CH}_2\text{SO}_2^+)$ and the following summands the fractional contributions of the respective organic fragment to SC-OA, while 0.147 is a scaling factor from Crippa et al. (2013b). The estimated MSA levels are 5.6 times higher than the measured MSA, indicating the presence of another source of sulfur-containing species. Second, unlike marine OA factors from previous

online datasets (lower size cut-off, typically PM1), the SC-OA time series does not correlate with MSA ($R^2=0.0302$). While MSA concentrations show a clear enhancement during summer, the SC-OA time series exhibit a very weak seasonal variability with slightly higher concentrations in winter. ~~SC-OA instead exhibits instead~~ low background levels episodically intercepted by remarkable ten-fold enhancements, especially at urban sites affected by traffic emissions (e.g. the SC-OA contribution is significantly higher at sites with higher yearly NO_x average levels). The ~~hypothesis of an influence of traffic activity on SC-OA is provided by the correlation of the~~ yearly average concentrations ~~of SC-OA with NO_x ($R_{s,SC-OA,NO_x}$~~ $=0.865$, $n=9$, $p<0.0106$) which is, however, comparable to the correlation of HOA and also of different POA factors ~~correlate with NO_x ($R_{s,COA}$ (e.g., $R_{s,HOA,NO_x}=0.468$, $n=9$, $p<0.27$), COA (05, $R_{s,COA,NO_x}=0.4368$, $n=9$, $p<0.2305$)).~~ Besides that, the SC-OA time series also correlates with that of NO_x (overall $R^2=0.432$, for sites in Tab. 2). While HOA and BBOA also correlate with NO_x , both of the secondary factors, WOOA and SOOA, do not, supporting the hypothesis that SC-OA consists of locally emitted anthropogenic (primary) OA. The site-to-site differences in SC-OA concentrations and temporal behaviour suggest that this factor, which to the best of our knowledge is reported here for the first time, is influenced by primary sources.

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

15

Oxygenated OA factors: Unlike oxygenated OA factors from limited-duration intensive online campaigns characterized by a high temporal resolution in which factor variability is thought to be primarily driven by volatility and/or local oxidation reactions, OOA factors are ~~here~~-resolved based on differences in their seasonal behavior: SOOA (in summer) and WOOA (in winter). The SOOA (summer) and WOOA (winter) mass spectral signature (Fig. 3) show similarities with OOA from earlier measurements (Ng et al., 2011, Canonaco et al., 2013, 2015), with high contributions of $\text{C}_2\text{H}_3\text{O}^+$, CO_2^+ and high OM/OC ratios though SOOA (OM/OC = 1.89, IQR 1.88-1.89) is less oxidized than WOOA (OM/OC = 2.12, IQR 2.11-2.14). The mass spectral fingerprints (Fig. 3), the temporal behaviour (Fig. 4e and f) and the relation to markers (Fig. 5c and d) of the two factors are in agreement with those from earlier work at other locations, including Zurich (Daellenbach et al., 2016), Payerre (Bozzetti et al., 2016a2016), and Lithuania (Bozzetti et al., 2016b2017a). This OOA separation appears to be typical for PMF analysis of long-term, low time-resolution OA mass spectra of filter samples.

20

25

SOOA correlates significantly among the different sites (also south and north of the alpine crest) and with local temperature (Fig. 5c). The SOOA exponential increase with average daily temperatures from 5-30 °C is consistent with the exponential increase in terpenes emissions, which are dominant biogenic SOA precursors (Guenther et al., 2006). This is also consistent with the mass spectral fingerprint of this factor, characterized by an $f_{\text{C}_2\text{H}_3\text{O}^+}$ of 0.10 and an $f_{\text{CO}_2^+}$ of 0.13, which are similar to values reported for chamber SOA from terpenes or at an urban location (Zurich) during summer (Canonaco et al., 2015). A similar temperature dependence of biogenic SOA concentrations has been observed for a terpene-dominated Canadian forest (Leaich et al., 2011) and for the case of Switzerland, using a similar source apportionment model (Daellenbach et al., 2016; Bozzetti et al., 2016a2016). Taken together, these observations suggest that SOOA principally derives from the

30

oxidation of biogenic precursors during summer. Site-to-site SOOA concentrations were not statistically different within our model errors, assessed from the different sensitivity tests for the case of Zurich. Therefore, even though the behavior of SOOA at the different sites studied here might be controlled by various parameters, including tree cover, available OA mass, air mass photochemical age and oxidation conditions (e.g. NO_x concentrations), temperature seems to be the main driver of the SOOA concentrations. Indeed, the aforementioned parameters may contribute, together with model and measurement uncertainties, to the observed scatter in the data. Biogenic VOC emissions might even in winter be non-negligible (Oderbolz et al., 2013; Schurgers et al., 2009; Holzke et al., 2006) and, therefore, significant winter-time SOOA concentrations are not in disagreement with the hypothesized biogenic origin. The lower SOOA concentrations in the temperature range between 7 and 12 °C might be explained by often occurring precipitation in this temperature range. We note that relative uncertainties related with ~~this factor~~SOOA increase with decreasing concentrations (Fig. 7), ~~due to the considerable contribution from other more significant wintertime~~7). A small error in modelling sources ~~(e.g. with high contributions (BBOA, WOOA and BBOA)) in winter can result in a large error of SOOA with its small contribution during winter.~~ Furthermore, some ~~winter-time SOA or~~ other sources like primary biological OA (PBOA, see Sec. 4.2.2) might also mix into SOOA.

Compared to SOOA, the WOOA profile can be distinguished by a higher contribution from CO₂⁺ and a lower C₂H₃O⁺ (Fig. 3), similar to OOA factors previously extracted in this region during winter based on ACSM measurements. This fingerprint is characteristic of highly oxidized SOA from non-biogenic precursors with low H/C ratios (e.g. aromatic compounds from wood combustion emissions, Bruns et al., 2016). WOOA is well correlated with NH₄⁺ (Fig. 5.d, overall R² 0.65 for all sites, overall R² 0.6781 for all PMF runs for Zurich in Tab. 2) which is in agreement with earlier studies (e.g., Zurich in Lanz et al., 2008). This is probably explained by its correlation with other inorganic secondary ions NO₃⁻ and SO₄²⁻ (driven like WOOA by meteorological factors including boundary layer height and temperature), which govern the NH₄⁺ concentration in the aerosol. Here, we have used ammonium as a proxy for aged aerosols affected by anthropogenic emissions, as WOOA correlates better with ammonium than with nitrate sulfate. We note that in winter, when WOOA is highest, 56% of ammonium can be attributed to nitrate, whereas in summer ammonium sulfate dominates (97% of ammonium can be attributed to sulfate). Therefore, WOOA correlates more with nitrate (R²=0.64) than sulfate (R²=0.48). WOOA exhibits a regional behaviour and its concentrations are correlated at all sites in the Swiss plateau. The WOOA mass spectral fingerprint, its seasonal variability and its high correlation with long-range transported anthropogenic inorganic secondary ions suggest that this factor is characteristic of a highly aged OOA influenced by wintertime anthropogenic emissions (e.g., biomass burning).

4.2 Uncertainty analysis

4.2.1 Model uncertainties

PMF uncertainties depend on the factor contribution. According to Ulbrich et al. (2009), reliable interpretation of factors with a low relative contribution is challenging. However, also the specificity of the time series and factor profile (caused by rotational ambiguity) and in this sense also solution acceptance criteria influence the uncertainty. In our analysis, we correct our results from WSOM to OM using R_k and, thereby, introduce additional uncertainties (caused by the uncertainty of R_k or an unknown R_k). The more uncertain R_k is, the higher is the additional uncertainty in the extrapolation (Eq. 45). As mentioned in Sec. 3.5, R_k constraints (recovery combinations for different factors) are available for R_{HOA} , R_{COA} , R_{BBOA} , and R_{OOA} but not for R_{SC-OA} and not for individual OOA factors (Daellenbach et al., 2016). With the available constraints of mass closure (for Magadino and Zurich), R_{SC-OA} can only be determined with a high uncertainty (Fig. 6).

The variability of the factor time series for the single PMF sensitivity tests (PMF_{block} , $PMF_{zue,iso}$, $PMF_{filter/month}$, $PMF_{zue,repr}$) is used as an uncertainty estimate (shaded area in Fig. 54). This estimate (σ_a) depends on the measurement repeatability (10 single mass spectra included for each sample) and on the selected PMF solution/ R_k combinations and, therefore, also on the a -value. However, the variability depending (1) on the choice of input points (time and site; PMF_{block} , $PMF_{zue,iso}$, $PMF_{filter/month}$) and (2) on the instrumental reproducibility ($PMF_{zue,repr}$) of the offline AMS measurements is not accounted for. The contribution of (1) and (2) to the uncertainty is assessed through the sensitivity tests by examining the variability of the median factor time-series (σ_b). σ_b is the variability of the median factor concentrations from the PMF sensitivity tests using PMF_{block} , $PMF_{zue,iso}$, $PMF_{filter/month}$, $PMF_{zue,repr}$ for the 12 samples common to all 4 PMF datasets. For the 12 filters common in all PMF datasets (PMF_{block} , $PMF_{zue,iso}$, $PMF_{filter/month}$, $PMF_{zue,repr}$), we calculate a best estimate of the overall uncertainty (err_{tot}), by propagating both error terms: σ_a and σ_b . As σ_b is not available for all datapoints, we parametrized σ_b as a function of the factor concentration (details in SI) and subsequently used this parameterized σ_b , σ_b' , to calculate an approximated overall error, err'_{tot} . err'_{tot} is displayed in Fig. 7.b in comparison with σ_a (Fig. 7.a). For all factors, err'_{tot} are in general high (~ a factor of 2), especially for low factor concentrations. (~ a factor of 2). It is worthwhile to note that for major factors exhibiting a similar seasonality, i.e. WOOA and BBOA, a great part of the uncertainty arises from σ_b the rotational ambiguity. Thus the variability between the PMF solutions using PMF_{block} , $PMF_{zue,iso}$, $PMF_{filter/month}$, $PMF_{zue,repr}$ (σ_b) and, therefore, the sensitivity of the factor concentrations on the chosen PMF dataset significantly contribute to the uncertainty. By contrast, moderately soluble fractions, COA and HOA, constrained in the PMF the major part of err'_{tot} is related to σ_a .

4.2.2 Influence of unresolved primary biological OA.

Unresolved sources in PMF are an inherent uncertainty of source apportionment analyses. As Bozzetti et al. (2016a,2016) show, PBOA can present considerable contributions to OA in PM10 (constituting a large part of coarse OA). In the present

analysis, ~~this primary source PBOA~~ could not be separated ~~statistically using by~~ PMF ~~(also not by neither unconstrained nor~~ using the mass spectral signature from Bozzetti et al., 2016) ~~without mixing with other factors possibly because of)~~. ~~This inability might be caused by~~ the low water-solubility ~~in combination with and~~ the absence of PM2.5 filters in the dataset. Since these coarse particles are only abundant in PM10 and not in PM2.5/PM1, the presence of both PM10 and PM2.5 samples, exhibiting a large gradient in PBOA, might allow an unambiguous separation of PBOA. ~~The aim of this section is to estimate the influence of PBOA on the source apportionment results. A quantification of this fraction is, however, beyond the scope of this paper.~~ In the following, we estimate the influence of PBOA in three alternative ways:

- Based on factor profiles: Bozzetti et al. (2016a,2016) identified the AMS fragment $C_2H_5O_2^+$ as a possible tracer ion for PBOA. Based on the seasonality of SOOA (high in summer and low in winter), one can assume that SOOA in this study is a linear combination of PBOA and SOOA identified in PM2.5 and PM1. Based on the relative contribution of the ion $C_2H_5O_2^+$ to the factor profiles of SOOA from this analysis and literature profiles of PBOA and SOOA from Bozzetti et al. (2016a,2016, study site: Payerne), we estimate that 17% of the water-soluble SOOA is in fact PBOA (between ~~52~~ and ~~2223~~% for the different sensitivity tests). Using this approach, ~~we estimate that PBOA contributes $0.30 \mu g/m^3$ during the warm months on average for the different sites $0.29 \mu g/m^3$ with a (site-to-site variability computed as standard deviation of the average concentration of all sites of $0.03 \mu g/m^3$ (-). During the same period, SOOA concentrations are $1.7478 \mu g/m^3$ with a (site-to-site variability of $0.1618 \mu g/m^3$), and OA $3.44 \mu g/m^3$ with a (site-to-site variability of $0.3744 \mu g/m^3$)). This approach is very uncertain, mainly due to the uncertainty in PBOA and SOOA profiles, the assumption of a constant PBOA contribution to SOOA, and also the uncertainty of R_{PBOA} .~~

- Based on coarse OC: Bozzetti et al. (2016a,2016) showed that coarse OC ($OC_{coarse} = OC_{PM10} - OC_{PM2.5}$) ~~is~~ in summer ~~is~~ dominated by PBOA for samples collected at a rural site in Switzerland (Payerne). For a subset of the samples used in the present work, OC in the PM2.5 fraction was also analysed (Basel, Bern, Magadino, Payerne, Zurich, accounting for 149 samples in total). For these samples, the OC_{coarse} contribution to OC in the PM10 fraction is in summer 16% higher than in winter (site-to-site variability of 4%). This part of OC might be related to resuspension caused by traffic or emissions of primary biological particles. ~~However, the indicator~~ The ion $C_2H_5O_2^+$ (~~indicator for PBOA~~) shows higher concentrations with increasing OC_{coarse} concentrations. Therefore, this increment can tentatively be ascribed to PBOA, which leads to a contribution of $0.55 \mu g/m^3$ to OC in summer (site-to-site variability $0.16 \mu g/m^3$). This results in an average summer PBOA concentration of $1.21 \mu g/m^3$ with a site-to-site variability of $0.39 \mu g/m^3$ when assuming an OM/OC of 2.2 (or $0.66 \pm 0.21 \mu g/m^3$, for OM/OC 1.2, OM/OC range according to Bozzetti et al., 2016). For Magadino (2014, Vlachou et al., in prep.), OC_{coarse} represents 8% of OC in PM10 in winter while this ratio is 25% in summer. It can be assumed that the difference of 17% in summer can be attributed to PBOA. Extrapolating this estimate to the overall data set from 2013 considered in this study and

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

Formatted: English (U.S.)

assuming an OM/OC of 2.2, PBOA contributes in summer on average $0.97 \mu\text{g}/\text{m}^3$ to OA in PM10, with a site-to-site variability of $0.13 \mu\text{g}/\text{m}^3$ (or $0.63 \pm 0.07 \mu\text{g}/\text{m}^3$ OA with for OM/OC of 1.2).

- **Based on cellulose:** It has previously been shown that free cellulose contributes strongly to PBOA (25% of PBOA mass, for measurements made in Payerne during summer 2012/winter 2013; Bozzetti et al., [2016a2016](#)). Therefore, we can attempt to use cellulose analyses on a subset of samples (the same one as for levoglucosan but Bern, see Sec. 2.3) to estimate PBOA concentrations (Fig. 8). As seen for the case of $\text{OC}_{\text{coarse}}$, cellulose concentrations also increase with higher $\text{C}_2\text{H}_5\text{O}_2^+$ concentrations. For the sites with cellulose measurements available (all sites in the study but Bern), cellulose average concentrations of $0.17 \mu\text{g}/\text{m}^3$ (site-to-site variability $0.08 \mu\text{g}/\text{m}^3$, in the warm season $0.18 \pm 0.07 \mu\text{g}/\text{m}^3$) are observed which corresponds to $0.69 \mu\text{g}/\text{m}^3$ PBOA with a site-to-site variability of $0.34 \mu\text{g}/\text{m}^3$ (, in the warm season $0.77 \pm 0.29 \mu\text{g}/\text{m}^3$), using the ratio cellulose / PBOA from Bozzetti et al. ([2016a2016](#)). In this last study conducted during summer (15 days in June/July 2012), PBOA concentrations of $3 \mu\text{g}/\text{m}^3$ on average (with cellulose concentrations of $0.8 \mu\text{g}/\text{m}^3$) were estimated, which is clearly above the observation made here. However, Bozzetti et al. ([2016a2016](#)) assessed a shorter time period with diurnal resolution, instead of 1 sample per month as in the present work. Cellulose concentrations from other European sites during other years are consistent with the results in this study (Sanchez-Ochoa et al., 2007; Yttri et al., 2011). In general, the background cellulose concentrations at the southern alpine sites are higher and also the temporal behaviour deviates from the one observed at the northern sites: the maximal concentrations are not reached in July/August but rather in May or October/November. The different seasonality might be caused by different agricultural procedures. The higher background concentrations of cellulose for the southern Alpine sites might be caused by interferences from wood burning, which in the absence of glucose analyses cannot be excluded.

All these PBOA estimates (between 0.3 to $1.0 \mu\text{g}/\text{m}^3$ during the warm season) are consistently lower than reported in Bozzetti et al. ([2016a2016](#)), with a factor of 3 to 10 times lower depending on the site. One should keep in mind that these estimates are based on limited data sets in both studies (30 samples in Bozzetti et al. (2016) while 12 samples from the same site in this study).

4.3 Factor relative contribution at different sites

In general, the seasonality of the factor time series is consistent for all the 9 sites in the entire study area (Fig. 9). In summer, SOOA is the main contributor to OA, while in winter POA ($\text{HOA} + \text{COA} + \text{BBOA}$) becomes more important ~~though~~ ~~although~~ WOOA still contributes significantly. In comparison to the sites in northern Switzerland, OA in the southern alpine valleys is dominated by BBOA in winter, while in the north WOOA also plays a role. The different factors contribute 0.3947 ± 0.4412 (HOA, average and site-to-site variability), 0.2431 ± 0.4213 (COA), 1.2537 ± 1.7077 (BBOA), 0.5367 ± 0.2731 (SC-OA), 0.9311 ± 0.23 (WOOA), 1.4231 ± 0.4213 (SOOA) $\mu\text{g}/\text{m}^3$ for all sites during the entire year ([Tab-Table 3](#)). In northern

Switzerland, POA ($\text{HOA} + \text{COA} + \text{BBOA}$) contributes less to OA ($\text{POA}/\text{OA} = 0.3$) than in the southern alpine valleys where POA/OA is equal to 0.6. Among POA, BBOA is the most important, with 88.87% of POA in the south and 43.42% in the north. The higher relative contribution of BBOA to POA in the southern alpine valleys than at the northern sites supports the conclusion that the high BBOA concentrations (e.g. $2.2645 \mu\text{g}/\text{m}^3$ Magadino compared to $0.5362 \mu\text{g}/\text{m}^3$ Vaduz) are not only a consequence of the meteorological situation in the valleys (strong thermal inversion close to the valley ground) but mainly reflect the emission strength. SC-OA, which is possibly linked to a local source of rather primary origin, shows clear site-to-site differences, with e.g. high concentrations at a traffic site in Bern ($1.0825 \mu\text{g}/\text{m}^3$) and low concentrations at a rural site in Payerne ($0.4626 \mu\text{g}/\text{m}^3$). SOOA, believed to have strong influences from biogenic SOA, shows consistently low concentrations at all sites for low temperatures ($0.6576 \pm 0.6067 \mu\text{g}/\text{m}^3$ at $5\text{--}15^\circ\text{C}$) and clearly increased concentrations under warmer conditions ($4.6085 \pm 1.3551 \mu\text{g}/\text{m}^3$ at $25\text{--}35^\circ\text{C}$).

5 Conclusion

Sources contributing to OA are quantitatively separated and their uncertainty estimated statistically at 9 sites in central Europe throughout the entire year 2013 (819 samples). Thereby, 3 primary (HOA, COA, BBOA) OA sources are separated from 2 secondary (WOOA, SOOA) categories and a yet unknown source explaining sulfur-containing fragments (SC-OA). BBOA exhibits clearly higher concentrations at the alpine valley sites in southern Switzerland than at the sites in northern Switzerland. SOOA, characterized by high concentrations in summer, shows a more than linear increase with rising temperatures as is observed from biogenic VOC emissions and biogenic SOA concentrations. WOOA, dominant SOA category during winter, closely correlates with NH_4^+ . The influence of PBOA, not resolved by PMF, is estimated using among others cellulose analyses and could be an important contributor. Cellulose's temporal behaviour suggests maximal PBOA contributions in northern Switzerland during summer while at the southern alpine sites maximal concentrations are reached in spring/autumn.

6 References

Aiken, A. C., Salcedo, D., Cubison, M. J., Huffman, J. A., DeCarlo, P. F., Ulbrich, I. M., Docherty, K. S., Sueper, D., Kimmel, J. R., Worsnop, D. R., Trimborn, A., Northway, M., Stone, E. A., Schauer, J. J., Volkamer, R. M., Fortner, E., de Foy, B., Wang, J., Laskin, A., Shutthanandan, V., Zheng, J., Zhang, R., Gaffney, J., Marley, N. A., Paredes-Miranda, G., Arnott, W. P., Molina, L. T., Sosa, G., and Jimenez, J. L.: Mexico City aerosol analysis during MILAGRO using high resolution aerosol mass spectrometry at the urban supersite (T0) – Part 1: Fine particle composition and organic source apportionment, *Atmos. Chem. Phys.*, 9, 6633–6653, doi:10.5194/acp-9-6633-2009, 2009.

Formatted: Font color: Auto

Formatted: Font color: Auto

- Allan, J. D., Jimenez, J. L., Williams, P. I., Alfarra, M. R., Bower, K. N., Jayne, J. T., Coe, H., and Worsnop, D. R.: Quantitative sampling using an Aerodyne aerosol mass spectrometer 1. Techniques of data interpretation and error analysis, *J. Geophys. Res.*, 108, 4090, doi:10.1029/2002JD002358, 2003.
- 5 Birch, M. E. and Cary, R. A.: Elemental carbon-based method for monitoring occupational exposures to particulate diesel exhaust, *Aerosol Sci. and Tech.*, 25, 221–241, doi:10.1080/02786829608965393, 1996.
- 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.,
10 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, ~~2016a~~2016.
- Bozzetti, C., Sosedova, Y., Xiao, M., Daellenbach, K. R., Ulevicius, V., Dudoitis, V., Mordas, G., Byčenkienė, S.,
15 Plauškaitė, K., Vlachou, A., Golly, B., Chazeau, B., Besombes, J.-L., Baltensperger, U., Jaffrezo, J.-L., Slowik, J. G., El Haddad, I., and Prévôt, A. S. H.: Argon offline-AMS source apportionment of organic aerosol over yearly cycles for an urban, rural and marine site in Northern Europe, *Atmos. Chem. Phys. Discuss.*, doi:10.5194/acp-2016-413, ~~2016b~~2017a.
- Bozzetti, C., El Haddad, I., Salameh, D., Daellenbach, K. R., Fermo, P., Gonzalez, R., Minguillón, M. C., Iinuma, Y.,
20 Poulain, L., Müller, E., Slowik, J. G., Jaffrezo, J.-L., Baltensperger, U., Marchand, N., and Prévôt, A. S. H.: Organic aerosol source apportionment by offline-AMS over a full year in Marseille, *Atmos. Chem. Phys.*, ~~in prep.~~17, 8247–8268, <https://doi.org/10.5194/acp-17-8247-2017>.
- 25 Bruns, E. A., El Haddad, I., Slowik, J. G., Kilic, D., Klein, F., Baltensperger, U., and Prévôt, A. S. H.: Identification of significant precursor gases of secondary organic aerosols from residential wood combustion, *Sci. Rep.*, 6, 27781, doi:10.1038/srep27881, 2016.
- Canagaratna, M. R., Jayne, J. T., Jimenez, J. L., Allan, J. D., Alfarra, M. R., Zhang, Q., Onasch, T. B., Drewnick, F., Coe, H., Middlebrook, A., Delia, A., Williams, L. R., Trimborn, A. M., Northway, M. J., DeCarlo, P. F., Kolb, C. E., Davidovits, P. and Worsnop, D. R.: Chemical and microphysical characterization of ambient aerosols with the Aerodyne aerosol mass spectrometer, *Mass Spectrom. Rev.* 26:185–222, doi:10.1002/mas.20115, 2007.
- 30

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

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.

5 Canonaco, F., Slowik, J. G., Baltensperger, U., and Prévôt, A. S. H.: Seasonal differences in oxygenated organic aerosol composition: implications for emissions sources and factor analysis. *Atmos. Chem. Phys.* 15, 6993-7002, doi:10.5194/acp-15-6993-2015, 2015.

10 Cavalli, F., Viana, M., Yttri, K. E., Genberg, J., and Putaud, J. P.: Toward a standardised thermal-optical protocol for measuring atmospheric organic and elemental carbon: the EUSAAR protocol, *Atmos. Meas. Tech.*, 3, 79-89, doi:10.5194/amt-3-79-2010, 2010.

15 [Chen, Q., Miyazaki, Y., Kawamura, K., Matsumoto, K., Coburn, S. C., Volkamer, R., Iwamoto, Y., Kagami, S., Deng, Y., Ogawa, S., Ramasamy, S., Kato, S., Ida, A., Kajii, Y., and Mochida, M.: Characterization of chromophoric water-soluble organic matter in urban, forest, and marine aerosols by HR-ToF-AMS analysis and excitation emission matrix spectroscopy, *Environ. Sci. Technol.*, 50, 10,351–10,360, doi: 10.1021/acs.est.6b01643, 2016.](#)

20 Crippa, M., Canonaco, F., Lanz, V. A., Äijälä, M., Allan, J. D., Carbone, S., Capes, G., Ceburnis, D., Dall'Osto, M., Day, D. A., DeCarlo, P. F., Ehn, M., Eriksson, A., Freney, E., Hildebrandt Ruiz, L., Hillamo, R., Jimenez, J. L., Junninen, H., Kiendler-Scharr, A., Kortelainen, A.-M., Kulmala, M., Laaksonen, A., Mensah, A. A., Mohr, C., Nemitz, E., O'Dowd, C., Ovadnevaite, J., Pandis, S. N., Petäjä, T., Poulain, L., Saarikoski, S., Sellegri, K., Swietlicki, E., Tiitta, P., Worsnop, D. R., Baltensperger, U., and Prévôt, A. S. H.: Organic aerosol components derived from 25 AMS data sets across Europe using a consistent ME-2 based source apportionment approach, *Atmos. Chem. Phys.*, 14, 6159-6176, doi:10.5194/acp-14-6159-2014, 2014.

25 Crippa, M., Canonaco, F., Slowik, J. G., El Haddad, I., DeCarlo, P. F., Mohr, C., Heringa, M. F., Chirico, R., Marchand, N., Temime-Roussel, B., Abidi, E., Poulain, L., Wiedensohler, A., Baltensperger, U., and Prévôt, A. S. H.: Primary and secondary organic aerosol origin by combined gas-particle phase source apportionment, *Atmos. Chem. Phys.*, 13, 8411-8426, doi:10.5194/acp-13-8411-2013, 2013a.

30 Crippa, M., El Haddad, I., Slowik, J. G., DeCarlo, P.F., Mohr, C., Heringa, M. F., Chirico, R., Marchand, N., L., Sciare, J., Baltensperger, U., and Prévôt, A. S. H.: Identification of marine and continental aerosol sources in Paris using high resolution aerosol mass spectrometry, *J. Geophys. Res.*, 118, 1950-1963, doi:10.1002/jgrd.50151, 2013b.

Crippa, M., DeCarlo, P. F., Slowik, J. G., Mohr, C., Heringa, M. F., Chirico, R., Poulain, L., Freutel, F., Sciare, J., Cozic, J., Di Marco, C. F., Elsasser, M., Nicolas, J. B., Marchand, N., Abidi, E., Wiedensohler, A., Drewnick, F., Schneider, J., Borrmann, S., Nemitz, E., Zimmermann, R., Jaffrezo, J.-L., Prévôt, A. S. H., and Baltensperger, U.

Formatted

5 Crippa, M., DeCarlo, P. F., Slowik, J. G., Mohr, C., Heringa, M. F., Chirico, R., Poulain, L., Freutel, F., Sciare, J., Cozic, J., Di Marco, C. F., Elsasser, M., Nicolas, J. B., Marchand, N., Abidi, E., Wiedensohler, A., Drewnick, F., Schneider, J., Borrmann, S., Nemitz, E., Zimmermann, R., Jaffrezo, J.-L., Prévôt, A. S. H., and Baltensperger, U.: Wintertime aerosol chemical composition and source apportionment of the organic fraction in the metropolitan area of Paris, Atmos. Chem. Phys., 13, 961-981, doi:10.5194/acp-13-961-2013, 2013c.

10 Daellenbach, K. R., Bozzetti, C., ~~Krepeleova~~Krepeleová, A., Canonaco, F., ~~Huang, R.-J.~~Wolf, R., Zotter, P., Fermo, P., Crippa, M., Slowik, J. G., Sosedova, Y., Zhang, Y., Huang, R.-J., Poulain, L., Szidat, S., Baltensperger, U., El Haddad, I., and Prévôt, A. S. H., and El Haddad, I. .:
15 Characterization and source apportionment of organic aerosol using offline aerosol mass spectrometry, Atmos. Meas. Tech., 9, 23-39, doi:10.5194/amt-9-23-2016, 2016.

Formatted: pb_toc_link, Font: Times New Roman, English (U.K.)

Formatted: pb_toc_link, Font: Times New Roman, English (U.K.)

Formatted: pb_toc_link, Font: Times New Roman, English (U.K.)

Formatted: pb_toc_link, Font: Times New Roman, English (U.K.)

Formatted: pb_toc_link, Font: Times New Roman, English (U.K.)

Formatted: pb_toc_link, Font: Times New Roman, English (U.K.)

20 DeCarlo, P. F., Dunlea, E. J., Kimmel, J. R., Aiken, A. C., Sueper, D., Crouse, J., Wennberg, P. O., Emmons, L., Shinozuka, Y., Clarke, A., Zhou, J., Tomlinson, J. Collins, D. R., Knapp, D., Weinheimer, A. J., Montzka, D. D., Campos, T., and Jimenez, J. L.: Fast airborne aerosol size and chemistry measurements above Mexico city and Central Mexico during the MILAGRO campaign, Atmos. Chem. Phys., 8, 4027-4048, doi:10.5194/acp-8-4027-2008, 2008.

25 El Haddad, I., Marchand, N., D'Anna, B., Jaffrezo, J.-L., and Wortham, H.: Functional group composition of organic aerosol from combustion emissions and secondary processes at two contrasted urban environments, Atmos. Environ., 75, 308– 320, doi:10.5194/acp-13-7875-2013, 2013.

Elser, M., Bozzetti, C., El-Haddad, I., Maasikmets, M., Teinemaa, E., Richter, R., Wolf, R., Slowik, J. G., Baltensperger, U., and Prévôt, A. S. H.: Urban increments of gaseous and aerosol pollutants and their sources using mobile aerosol mass spectrometry measurements, Atmos. Chem. Phys., 16, 7117-7134, doi:10.5194/acp-16-7117-2016 , 2016a.

30 Elser, M., Huang, R.-J., Wolf, R., Slowik, J. G., Wang, Q., Canonaco, F., Li, G., Bozzetti, C., Daellenbach, K. R., Huang, Y., Zhang, R., Li, Z., Cao, J., Baltensperger, U., El-Haddad, I., and Prévôt, A. S. H.: New insights into PM2.5 chemical composition and sources in two major cities in China during extreme haze events using aerosol mass spectrometry, Atmos. Chem. Phys., 16, 3207-3225, doi:10.5194/acp-16-3207-2016, 2016b.

Favez, O., El Haddad, I., Piot, C., Boréave, A., Abidi, E., Marchand, N., Jaffrezo, J.-L., Besombes, J.-L., Personnaz, M.-B., Sciare, J., Wortham, H., George, C., and D'Anna, B.: Inter-comparison of source apportionment models for the estimation of wood burning aerosols during wintertime in an Alpine city (Grenoble, France), *Atmos. Chem. Phys.*, 10, 5295–5314, doi:10.5194/acp-10-5295-2010, 2010.

5

Fröhlich, R., Cubison, M. J., Slowik, J. G., Bukowiecki, N., Prévôt, A. S. H., Baltensperger, U., Schneider, J., Kimmel, J. R., Gonin, M., Rohner, U., Worsnop, D. R., and Jayne, J. T.: The ToF-ACSM: a portable aerosol chemical speciation monitor with TOFMS detection, *Atmos. Meas. Tech.*, 6, 3225-3241, doi:10.5194/amt-6-3225-2013, 2013.

10 Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), *Atmos. Chem. Phys.*, 6, 3181-3210, doi:10.5194/acp-6-3181-2006, 2006.

15 Herich, H., Gianini, M. F. D., Piot, C., Močnik, G., Jaffrezo, J.-L., Besombes, J.-L., Prévôt, A. S. H., Hueglin, C.: Overview of the impact of wood burning emissions on carbonaceous aerosols and PM in large parts of the Alpine region, *Atmos. Environ.*, 89, 64-75, doi:10.1016/j.atmosenv.2014.02.008, 2014

20 Holzke, C., Hoffmann, T., Jaeger, L., Koppmann, R., and Zimmer, W.: Diurnal and seasonal variation of monoterpene and sesquiterpene emissions from Scots pine (*Pinus sylvestris* L.), *Atmos. Environ.*, 40, 3174-3185, doi:org/10.1016/j.atmosenv.2006.01.039, 2006.

25 Huang, R.-J., Zhang, Y., Bozzetti, C., Ho, K.-F., Cao, J., Han, Y., Dällenbach, K. R., Slowik, J. G., Platt, S. M., Canonaco, F., Zotter, P., Wolf, R., Pieber, S. M., Bruns, E. A., Crippa, M., Ciarelli, G., Piazzalunga, A., Schwikowski, M., Abbaszade, G., Schnelle-Kreis, J., Zimmermann, R., An, Z., Szidat, S., Baltensperger, U., Haddad, I. E., and Prévôt, A. S. H.: High secondary aerosol contribution to particulate pollution during haze events in China, *Nature*, 514, 218-222, doi:10.1038/nature13774, 2014.

30 Iinuma, Y., Engling, G., Puxbaum, H., and Herrmann, H.: A highly resolved anion-exchange chromatographic method for determination of saccharidic tracers for biomass combustion and primary bio-particles in atmospheric aerosol, *Atmos. Environ.*, 43, 1367–1371, 2009.

Jaffrezo, J. L., Calas, T., and Bouchet, M.: Carboxylic acids measurements with ionic chromatography, *Atmos. Environ.*, 32, 2705–2708, doi:10.1016/S1352-2310(98)00026-0, 1998.

Formatted: Font color: Auto

Formatted: Font color: Auto

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prévôt, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., Dunlea, E. J., Huffman, J. A., Onasch, T. B.,
5 Alfara, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimo, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb, C. E., Baltensperger, U., and Worsnop, D. R.: Evolution of organic aerosols in the atmosphere, *Science*, 326, 1525–1529, doi:10.1126/science.1180353, 2009.

10

Kirchstetter, T., Harley, R., Kreisberg, N., Stolzenburg, M., and Hering, S.: On-road measurement of fine particle and nitrogen oxide emissions from light- and heavy-duty motor vehicles, *Atmos. Environ.*, 33(18), 2955–2968, 1999.

Kunit,M. and Puxbaum, H.: Enzymatic determination of the cellulose content of atmospheric aerosols, *Atmos. Environ.*, 30,
15 1233–1236, doi:10.1016/1352-2310(95)00429-7, 1996.

Lanz, V. A., Alfara, M. R., Baltensperger, U., Buchmann, B., Hueglin, C., and Prévôt, A. S. H.: Source apportionment of submicron organic aerosols at an urban site by factor analytical modelling of aerosol mass spectra, *Atmos. Chem. Phys.*, 7,
1503–1522, doi:10.5194/acp-7-1503-2007, 2007.

20

Lanz, V. A., Alfara, M. R., Baltensperger, U., Buchmann, B., Hueglin, C., Szidat, S., Wehrli, M. N., Wacker, L., Weimer,
S., Caseiro, A., Puxbaum, H., and Prévôt, A. S. H.: Source attribution of submicron organic aerosols during wintertime
inversions by advanced factor analysis of aerosol mass spectra, *Environ. Sci. Technol.*, 42, 214–220,
doi:10.1021/es0707207, 2008.

25

Leaitch, W. R. Macdonald, A. M., Brickell, P. C., Liggio, J., Sjostedt, S. J., Vlasenko, A., Bottenheim, J. W., Huang, L., Li,
S.-M., Liu, P. S. K., Toom-Sauntry, D., Hayden, K. A., Sharma, S., Shantz, N. C., Wiebe H. A., Zhang, W., Abbatt, J. P. D.,
Slowik, J. G., Chang, Rachel, Y.-W., Russell, L. M., Schwartz, R. E., Takahama, S., Jayne, J. T., and Ng, N. L.: Temperature
response of the submicron organic aerosol from temperate forests, *Atmos. Environ.*, 45, 6696-6704,
30 doi:10.1016/j.atmosenv.2011.08.047, 2011.

Mihara, T. and Mochida, M.: Characterization of solvent-extractable organics in urban aerosols based on mass spectrum
analysis and hygroscopic growth measurement, *Envir. Sci. Tech.*, 45, 9168–9174, doi:10.1021/es201271w, 2011.

Formatted: Justified

- Minguillón, M. C., Perron, N., Querol, X., Szidat, S., Fahrmi, S. M., Alastuey, A., Jimenez, J. L., Mohr, C., Ortega, A. M., Day, D. A., Lanz, V. A., Wacker, L., Reche, C., Cusack, M., Amato, F., Kiss, G., Hoffer, A., Decesari, S., Moretti, F., Hillamo, R., Teinilä, K., Seco, R., Peñuelas, J., Metzger, A., Schallhart, S., Müller, M., Hansel, A., Burkhardt, J. F., Baltensperger, U., and Prévôt, A. S. H.: Fossil versus contemporary sources of fine elemental and organic carbonaceous particulate matter during the DAURE campaign in Northeast Spain, *Atmos. Chem. Phys.*, 11, 12067–12084, doi:10.5194/acp-11-12067-2011, 2011.
- Mohr, C., DeCarlo, P. F., Heringa, M. F., Chirico, R., Slowik, J. G., Richter, R., Reche, C., Alastuey, A., Querol, X., Seco, R., Penuelas, J., Jimenez, J. L., Crippa, M., Zimmermann, R., Baltensperger, U., and Prevot, A. S. H.: Identification and quantification of organic aerosol from cooking and other sources in Barcelona using aerosol mass spectrometer data, *Atmos. Chem. Phys.*, 12, 1649-1665, doi:10.5194/acp-12-1649-2012, 2012.
- Mohr, C., Richter, R., DeCarlo, P. F., Prévôt, A. S. H., and Baltensperger, U.: Spatial variation of chemical composition and sources of submicron aerosol in Zurich during wintertime using mobile aerosol mass spectrometer data, *Atmos. Chem. Phys.*, 11, 7465-7482, doi:10.5194/acp-11-7465-2011, 2011.
- 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, *Environ. Sci. Technol.*, 45, 910-916, doi:10.1021/es102951k, 2010.
- Ng, N. L., Herndon, S. C., Trimborn, A., Canagaratna, M. R. Croteau, P. L., Onasch, T. B. Sueper, D., Worsnop, D. R., Zhang, Q., Sun, Y. L., and Jayne, J. T.: An Aerosol Chemical Speciation Monitor (ACSM) for routine monitoring of the composition and mass concentrations of ambient aerosol, *Aerosol Sci. Tech.*, 45, 770–784, doi:10.1080/02786826.2011.560211, 2011.
- Oderbolz, D. C., Aksoyoglu, S., Keller, J., Barmpadimos, I., Steinbrecher, R., Skjøth, C. A., Plaß-Dülmer, C., and Prévôt, A. S. H.: A comprehensive emission inventory of biogenic volatile organic compounds in Europe: improved seasonality and land-cover, *Atmos. Chem. Phys.*, 13, 1689-1712, doi:10.5194/acp-13-1689-2013, 2013.
- Paatero, P., Eberly, S., Brown, S. G., and Norris, G. A.: Methods for estimating uncertainty in factor analytic solutions, *Atmos. Meas. Tech.*, 7, 781-797, doi:10.5194/amt-7-781-2014, doi:10.5194/amt-7-781-2014, 2014.
- Paatero, P.: The multilinear engine - A table-driven, least squares program for solving multilinear problems, including the n-way parallel factor analysis model, *J. Comput. Graph. Stat.*, 8, 854-888, doi:10.2307/1390831, 1999.

- Paatero, P.: Least squares formulation of robust non-negative factor analysis, *Chemometr. Intell. Lab.*, 37, 23–35, doi: 10.1016/S0169-7439(96)00044-5, 1997.
- 5 Piazzalunga, A., Fermo, P., Bernardoni, V., Vecchi, R., Valli, G., and De Gregorio, M. A.: A simplified method for levoglucosan quantification in wintertime atmospheric particulate matter by high performance anion-exchange chromatography coupled with pulsed amperometric detection, *Intern. J. Environ. Anal. Chem.*, 90, 12, 934-947, doi:10.1080/03067310903023619, 2010.
- 10 Piazzalunga, A., Bernardoni, V., Fermo, P., and Vecchi, R.: Optimisation of analytical procedures for the quantification of ionic and carbonaceous fractions in the atmospheric aerosol and applications to ambient samples, *Anal., Bioanal. Chem.*, 56, 30-40, doi:10.1007/s00216-012-6433-5, 2013
- Pieber, S. M., El Haddad, I., Slowik, J. G., Canagaratna, M. R., Jayne, J. T., Platt, S. M., Bozzetti, C., Daellenbach, K. R.,
15 Fröhlich, R., Vlachou, A., Klein, F., Dommen, J., Miljevic, B., Jimenez, J. L., Worsnop, D. R., Baltensperger, U., and Prévôt, A. S. H.: Inorganic salt interference on CO₂⁺ in aerodyne AMS and ACSM organic aerosol composition studies, *Environ. Sci. Technol.*, 50, 10494-10503, doi: 10.1021/acs.est.6b01035, 2016.
- Sanchez-Ochoa, A., Kasper-Giebl, Puxbaum, H., Gelencser, A., Legrand, M., Pio, C.: Concentration of atmospheric
20 cellulose: A proxy for plant debris across a west-east transect over Europe, *J. Geophys. Res.*, 112, D23S08, doi: 10.1029/2006JD008180, 2007.
- Schurgers, G., Arneth, A., Holzinger, R., and Goldstein, A. H.: Process-based modelling of biogenic monoterpene emissions combining production and release from storage, *Atmos. Chem. Phys.*, 9, 3409-3423, doi:10.5194/acp-9-3409-2009, 2009.
25
- Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R., and Jimenez, J. L.: Interpretation of organic components from Positive Matrix Factorization of aerosol mass spectrometric data, *Atmos. Chem. Phys.*, 9, 2891–2918, doi:10.5194/acp-9-2891-2009, 2009.
- 30 Vlachou, A., Daellenbach K. R., Bozzetti C., Szidat S., Salazar G., Agrios K., Jaffrezo J. L., Baltensperger U., El Haddad I. and Prévôt A. S. H., Radiocarbon based source apportionment of carbonaceous aerosols in Magadino, Switzerland. in prep.
- Zar, J.H., *Biostatistical Analysis*, 4th ed., 929 pp., Prentice Hall, Englewood Cliffs, New Jersey, 1999.

- Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Ulbrich, I. M., Ng, N. L., and Worsnop, Y. S.: Understanding atmospheric organic aerosols via factor analysis of aerosol mass spectrometry: a review, *Anal. Bioanal. Chem.*, 401, 3045–3067, doi:10.1007/s00216-011-5355-y, 2011.
- 5 Zorn, S. R., Drewnick, F., Schott, M., Hoffmann, T., and Borrmann, S.: Characterization of the South Atlantic marine boundary layer aerosol using an aerodyne aerosol mass spectrometer, *Atmos. Chem. Phys.*, 8, 4711–4728, doi:10.5194/acp-8-4711-2008, 2008.
- Zotter, P., Ciobanu, V. G., Zhang, Y. L., El-Haddad, I., Macchia, M., Daellenbach, K. R., Salazar, G. A., Huang, R.-J.,
10 Wacker, L., Hueglin, C., Piazzalunga, A., Fermo, P., Schwikowski, M., Baltensperger, U., Szidat, S., and Prévôt, A. S. H.: Radiocarbon analysis of elemental and organic carbon in Switzerland during winter-smog episodes from 2008 to 2012 – Part 1: Source apportionment and spatial variability, *Atmos. Chem. Phys.*, 14, 13551–13570, doi:10.5194/acp-14-13551-2014, 2014.
- 15 Yttri, K. E., Simpson, D., Nøjgaard, J. K., Kristensen, K., Genberg, J., Stenström, K., Swietlicki, E., Hillamo, R., Aurela, M., Bauer, H., Offenberg, J. H., Jaoui, M., Dye, C., Eckhardt, S., Burkhardt, J. F., Stohl, A., and Glasius, M.: Source apportionment of the summer time carbonaceous aerosol at Nordic rural background sites, *Atmos. Chem. Phys.*, 11, 13339–13357, doi:10.5194/acp-11-13339-2011, 2011.
- 20 Zotter, P., Zhang, Y.L., El-Haddad, I., Ciobanu, V.G., Daellenbach, K. R., Salazar, G. A., Wacker, L., Herich, H., Hueglin, C., Baltensperger, U., Szidat, S., and Prévôt A. S. H.: Radiocarbon Analysis of Carbonaceous Aerosols in Switzerland from 2008 to 2012 – Part II: Seasonal and yearly variability, *Atmos. Chem. Phys.*, in prep.

25

30

|

Formatted: Justified

Table 1: Study sites with geographical location and classification

Site (station code)	Classification	General location	altitude <u>Altitude</u>
Basel, St. Johann (bas)	Urban/background	North of Alps/Swiss plateau	308 m-
Bern, Bollwerk (ber)	Urban/traffic	North of Alps/Swiss plateau	506 m-
Frauenfeld, <u>Bahnhofstr.</u> (fra)	Suburban/background	North of Alps/Swiss plateau	403 m-
Payerne (pay)	Rural/background	North of Alps/Swiss plateau	539 m-
St. Gallen- er , <u>Rorschacherstr.</u> (gal)	Urban/traffic	North of Alps/Swiss plateau	457 m-
Zurich, Kaserne (zue)	Urban/background	North of Alps/Swiss plateau	457 m-
Vaduz, Austrasse (vad)	Urban/traffic	North of Alps/alpine valley	706 m-
Magadino, Cadenazzo (mag)	Rural/background	South of Alps	254 m-
San Vittore, <u>Zentrum</u> (vi)	Rural/traffic	South of Alps/alpine valley	330 m-

- Formatted: Justified
- Formatted: Justified
- Formatted: Justified
- Formatted: Justified
- Formatted: Justified
- Formatted: Justified
- Formatted: Justified
- Formatted: Justified
- Formatted: Justified
- Formatted: Justified
- Formatted: Justified
- Formatted: Justified
- Formatted: Justified
- Formatted: Justified
- Formatted: Justified
- Formatted: Justified

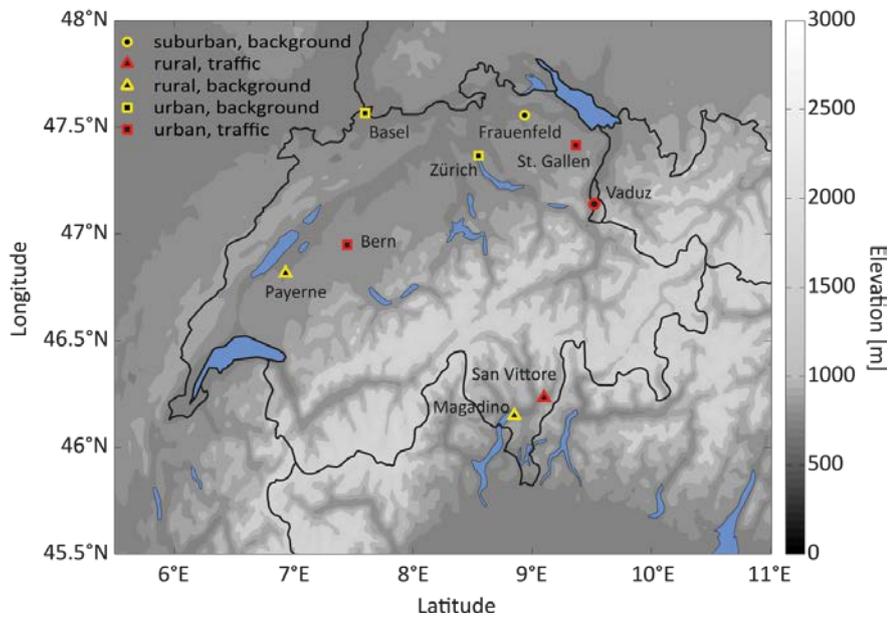


Figure 1: Map of study area with locations of sites indicating their characteristics. The topography is displayed as meters above sea level.

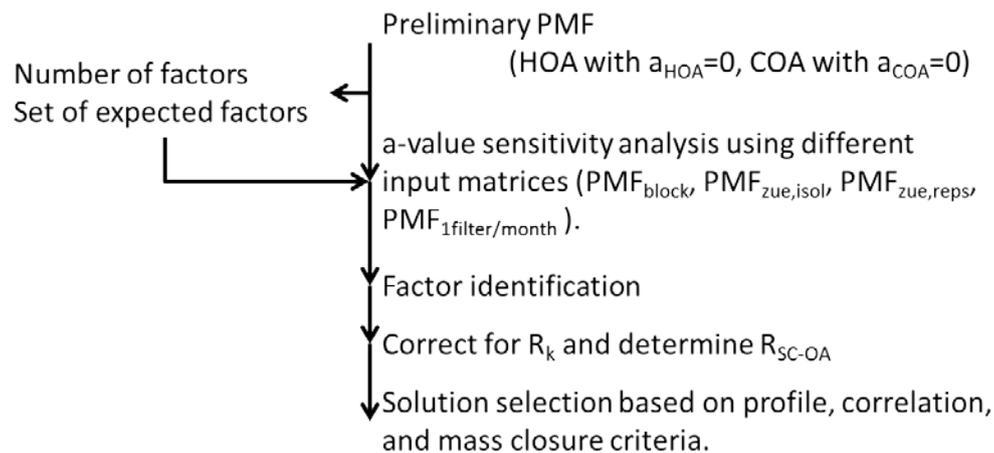
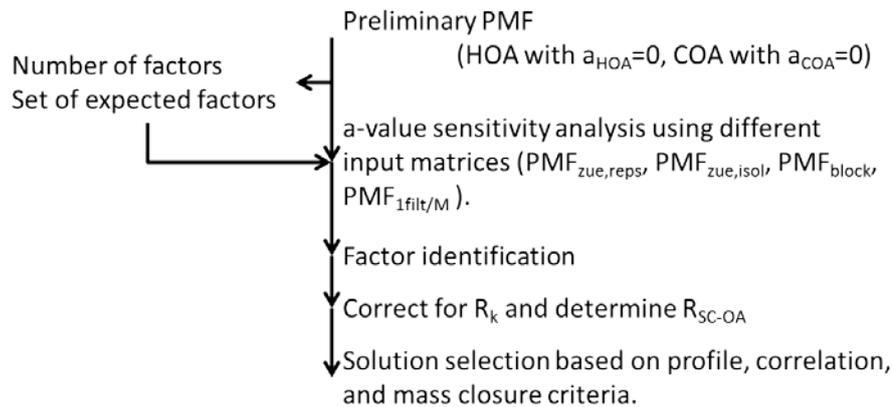
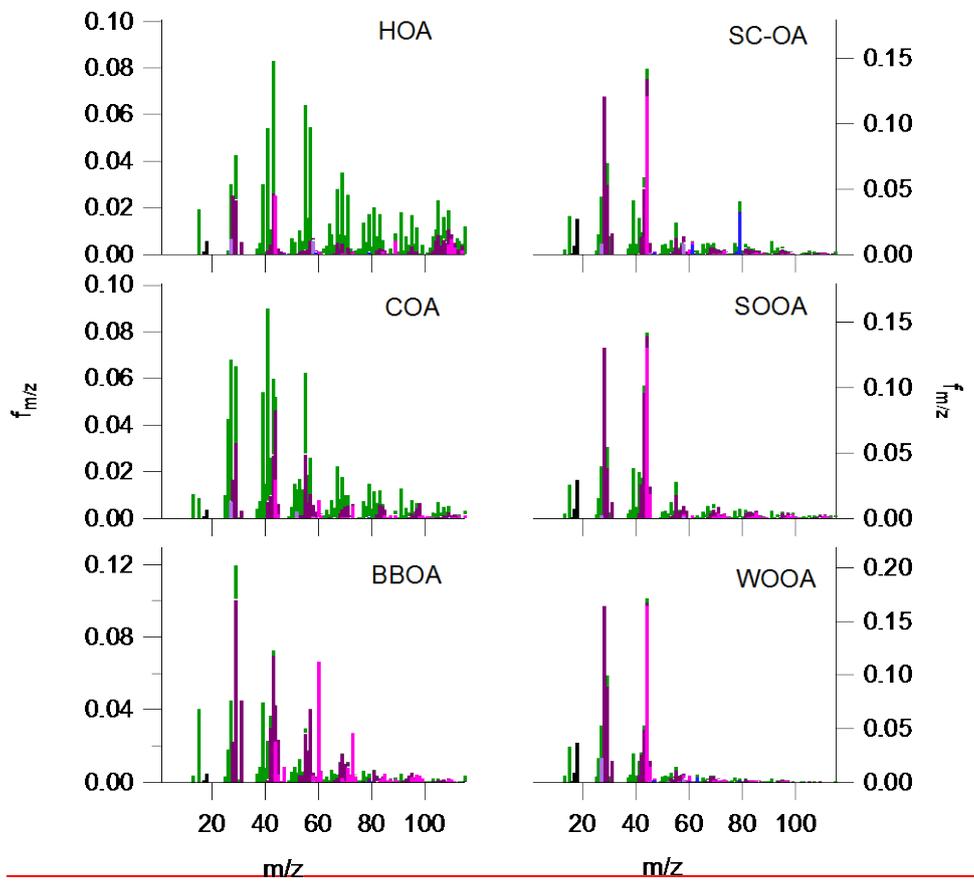


Figure 2: Step-by-step outline of adopted source apportionment approach (factor recoveries R_k). a_{HOA} and a_{COA} represent the a-value applied for HOA and COA, respectively.

■ CH, ■ CHO_{z=1}, ■ CHO_{z>1}, ■ CHN, ■ CS, ■ HxO



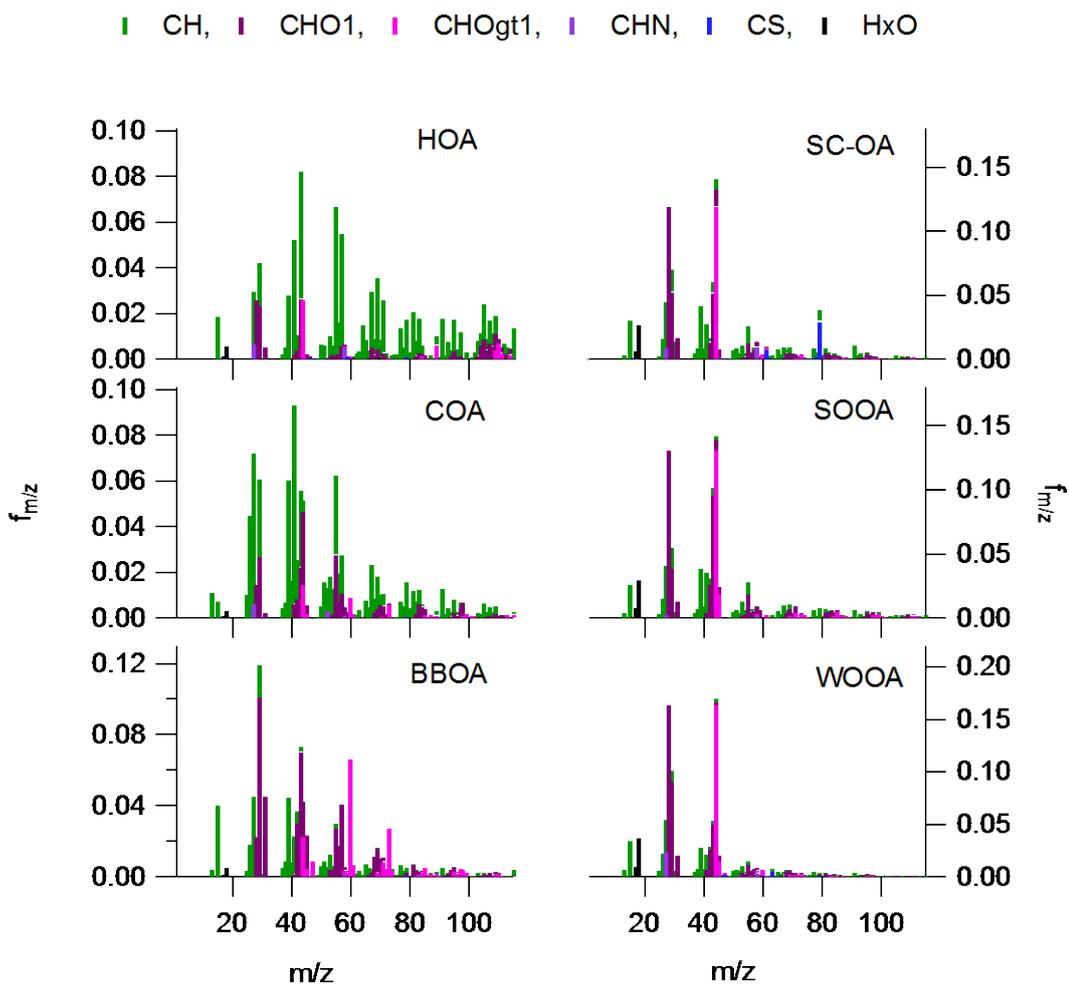
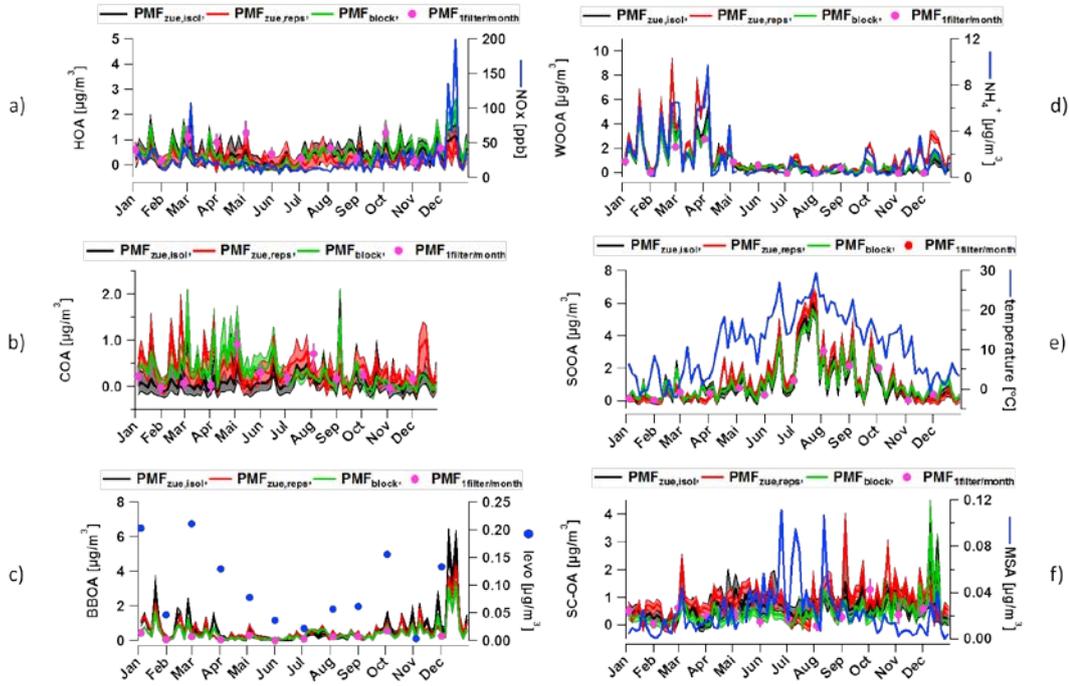


Figure 3: PMF factor profiles of HOA, COA, BBOA, SOOA, WOOA, SC-OA, color-coded with ion family of $\text{PMF}_{\text{block}}$ (average). $f_{m/z}$ is the relative intensity at a specific mass-to-charge ratio (m/z).



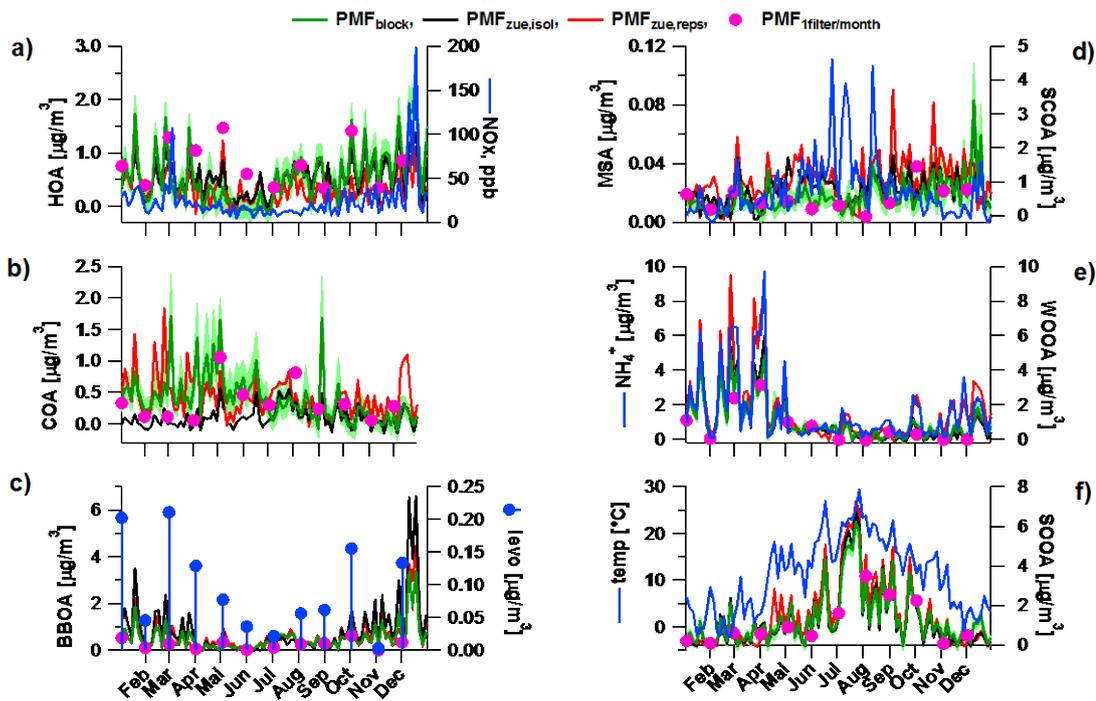
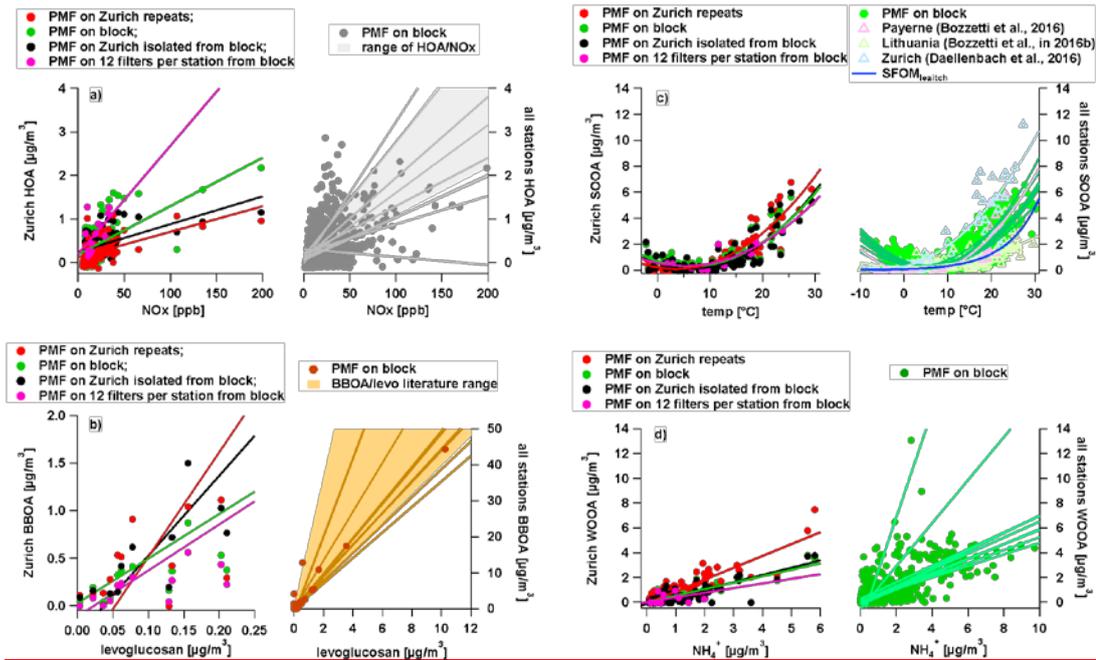
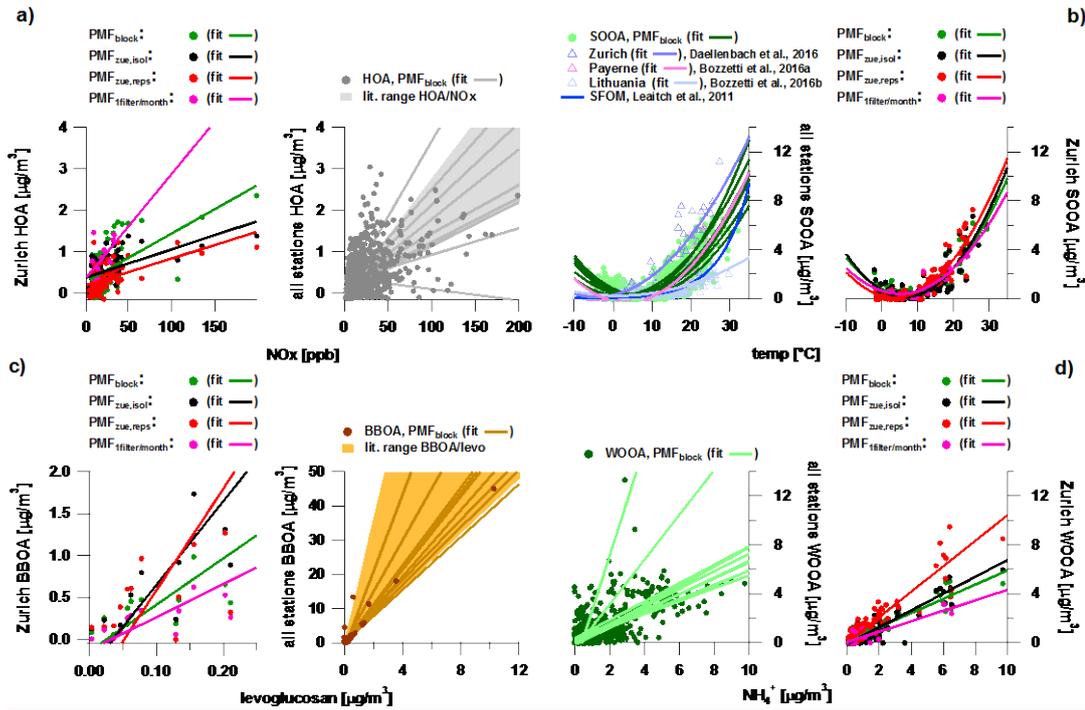


Figure 4: HOA, COA, BBOA, ~~SC-OA~~WOOA, SOOA, and ~~WOOA~~SC-OA and respective marker concentrations as a function of time for Zurich in 2013. Depicted are the median factor time series results for the different PMF datasets (median) including the uncertainties for PMF_{block} (first and third quartile) (green: PMF_{block}, black: PMF_{zue,isol}, red: PMF_{zue,repr}, pink bullets: ~~PMF_{filter/month}~~PMF_{filter/month}).

5

Formatted: Justified





5 | Figure 5: Scatter-plots for the different extreme sensitivity tests for Zurich and for all sites for $\text{PMF}_{\text{block}}$ median concentrations):
a) HOA vs NO_x , b) BBOA vs levoglucosan, c) SOOA vs temperature, d) WOOA vs NH_4^+ .

Formatted: Justified

Table 2: Comparison of factor time series to reference data for different PMF input datasets runs (by Pearson and Spearman correlation coefficient, R_p and R_s). Displayed are the results for PMF_{block} unless stated otherwise.

R^2 (number of points)	HOA vs NOx, R_p^2	BBOA vs levo, R_p^2	WOOA vs NH_4^+ , R_p^2	SOOA vs T, R_s	SC-OA vs NOx, R_p^2
Basel	0. 39 <u>31</u> (91)	0.91 (11)	0. 65 <u>66</u> (91)	0. 65 <u>70</u> (91)	0. 25 <u>17</u> (91)
Bern	0. 23 [†] <u>22</u> [†] (90)	0.48 (12)	0.53 (90)	0. 58 <u>63</u> (90)	0. 20 <u>17</u> (90)
Frauenfeld	0. 41 <u>40</u> (89)	0. 74 <u>73</u> (12)	0. 76 <u>77</u> (90)	0. 60 <u>63</u> (90)	0. 37 <u>28</u> (89)
St. Gallen	0. 22 <u>23</u> (91)	0. 40 <u>39</u> (12)	0.78 (91)	0. 68 <u>72</u> (91)	0. 59 <u>50</u> (91)
Magadino	0. 21 <u>18</u> (91)	0. 54 <u>55</u> (12)	0. 52 <u>54</u> (91)	0. 66 <u>72</u> (91)	0. 70 <u>63</u> (91)
Payerne	0. 49 <u>48</u> (91)	0. 67 <u>65</u> (12)	0. 46 <u>44</u> (91)	0. 63 <u>68</u> (91)	0. 27 <u>17</u> (91)
Vaduz	0. 39 <u>38</u> (91)	0.90 (12)	0.77 (91)	0. 65 <u>68</u> (91)	0. 54 <u>46</u> (91)
San Vittore	0. 01 <u>02</u> (90)	0.99 (12)	0.36 (90)	0. 72 <u>76</u> (68)	0. 00 <u>01</u> (90)
Zurich					
PMF _{block}	0. 38 <u>35</u> (91)	0. 42 <u>43</u> (12)	0.79 (90)	0. 61 <u>65</u> (91)	0. 50 <u>40</u> (91)
PMF _{zue,isoisol}	0. 28 <u>29</u> (91)	0.59 (12)	0. 83 <u>82</u> (90)	0. 63 <u>66</u> (91)	0. 34 <u>27</u> (91)
PMF _{zue,reprtps} (only 12 points)	0. 33 <u>32</u> (12)	0.23 (12)	0.84 (12)	0.85 (12)	0. 02 <u>01</u> (12)
PMF _{filt/m} PMF _{lfj} lter/month	0.30 (91)	0.44 (12)	0. 78 <u>77</u> (90)	0.59 (91)	0. 63 <u>53</u> (91)

[†] 1 outlier removed.

Formatted: Justified

Formatted: Justified

Formatted: Justified

Formatted: Justified

Formatted: Justified

Formatted: Justified

Formatted: Justified

Formatted: Justified

Formatted: Justified

Formatted: Justified

Formatted: Justified

Formatted: Justified

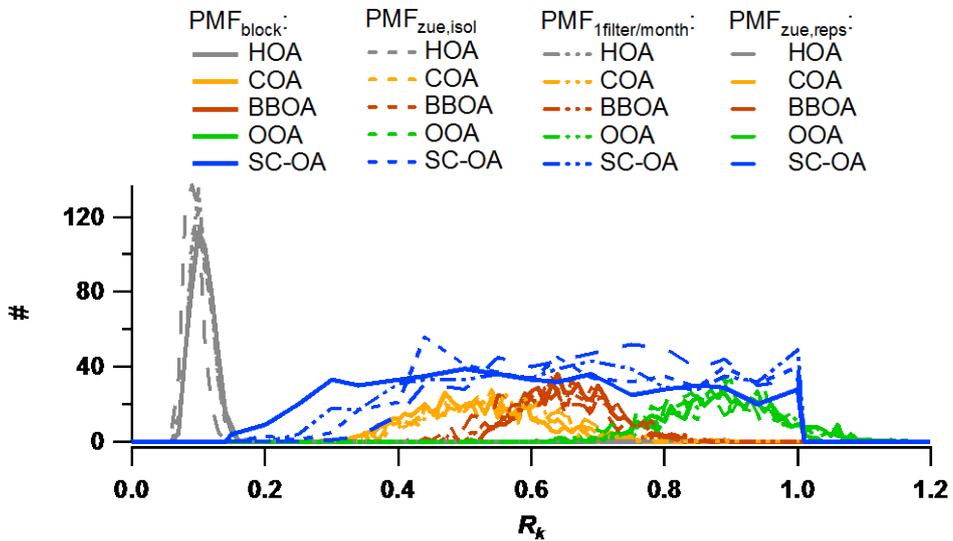
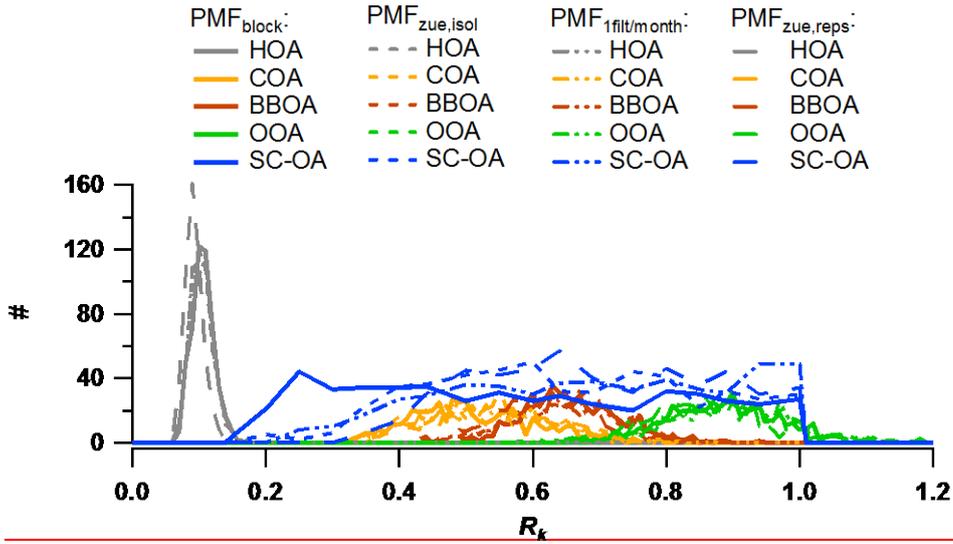
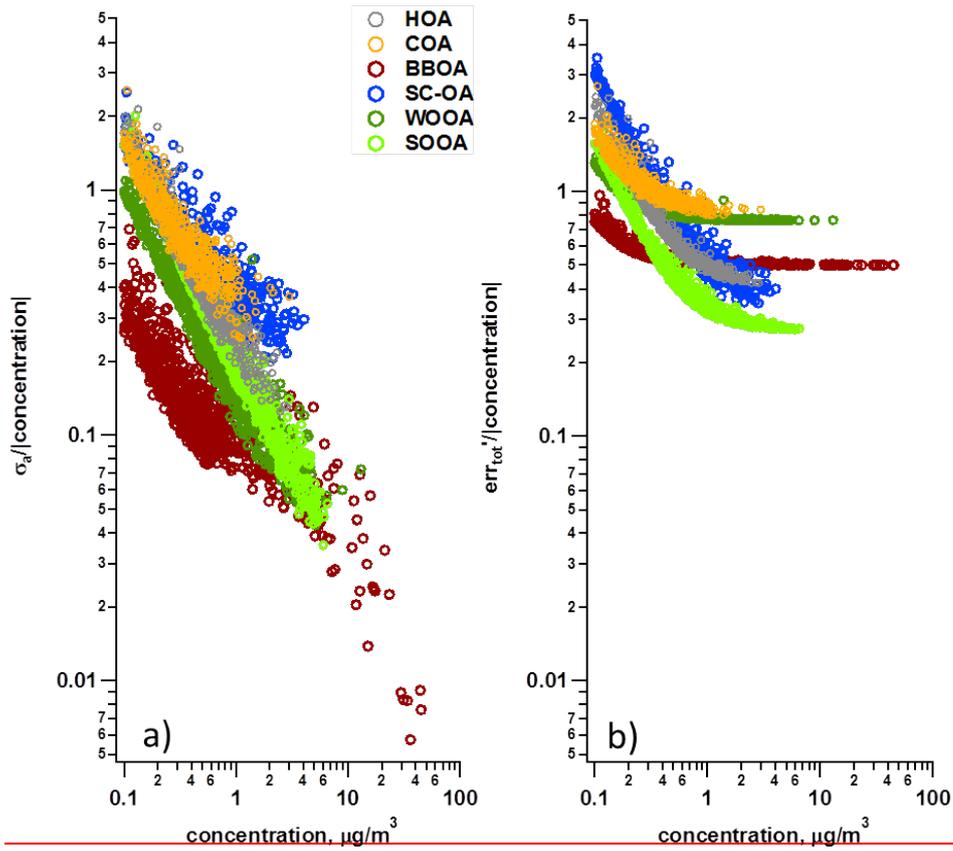


Figure 6: Distributions of R_k for HOA, COA, BBOA, OOA (WOOA plus SOOA) and SC-OA (500 pairs). A priori information for HOA, COA, BBOA, and OOA on R_k are used from Daellenbach et al., 2016, with propagated errors and biases, while R_{SC-OA} is determined in this study. Distributions of all factors have a resolution of $dR_k=0.01$ except for $dR_{SC-OA}=0.05$.

|

← Formatted: Justified



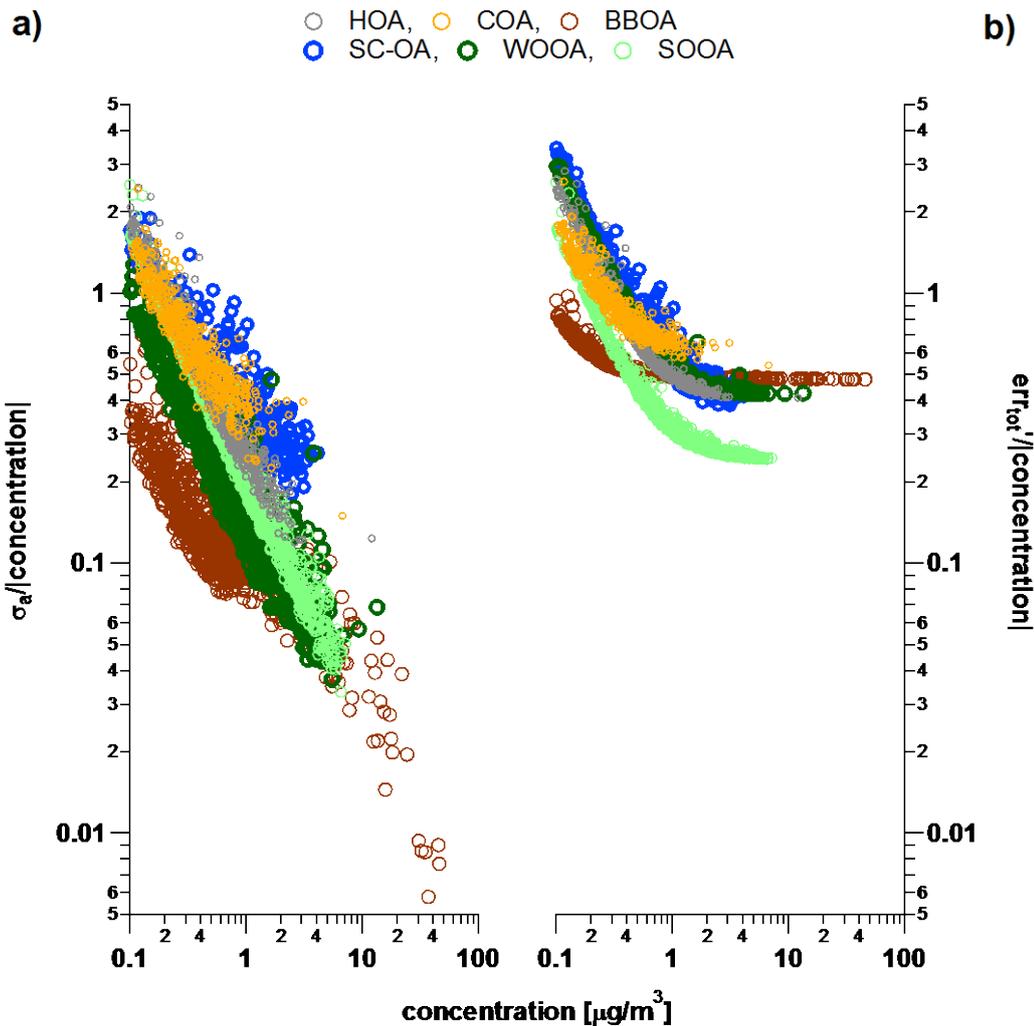
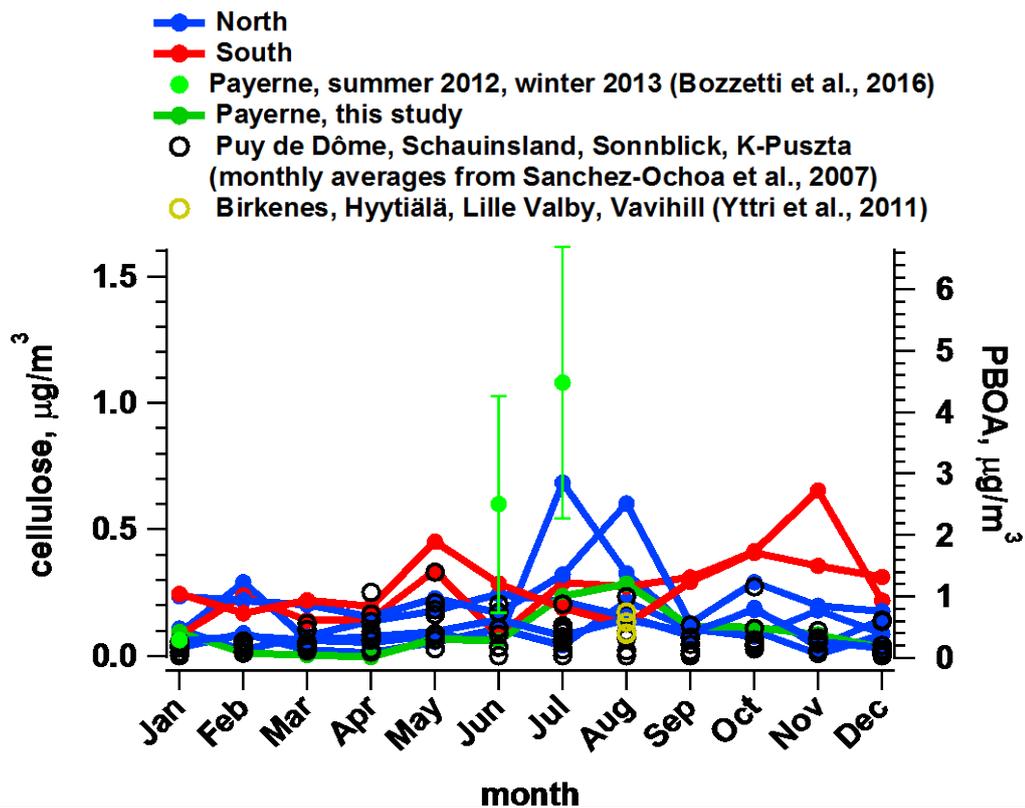


Figure 7: Relative σ_a (a) and err_{tot}' (b) for factor concentrations $> 0.1 \mu\text{g}/\text{m}^3$ as a function of factor concentration. err_{tot}' includes the uncertainties from a -value/seed variability and R_k , and the different PMF datasets.



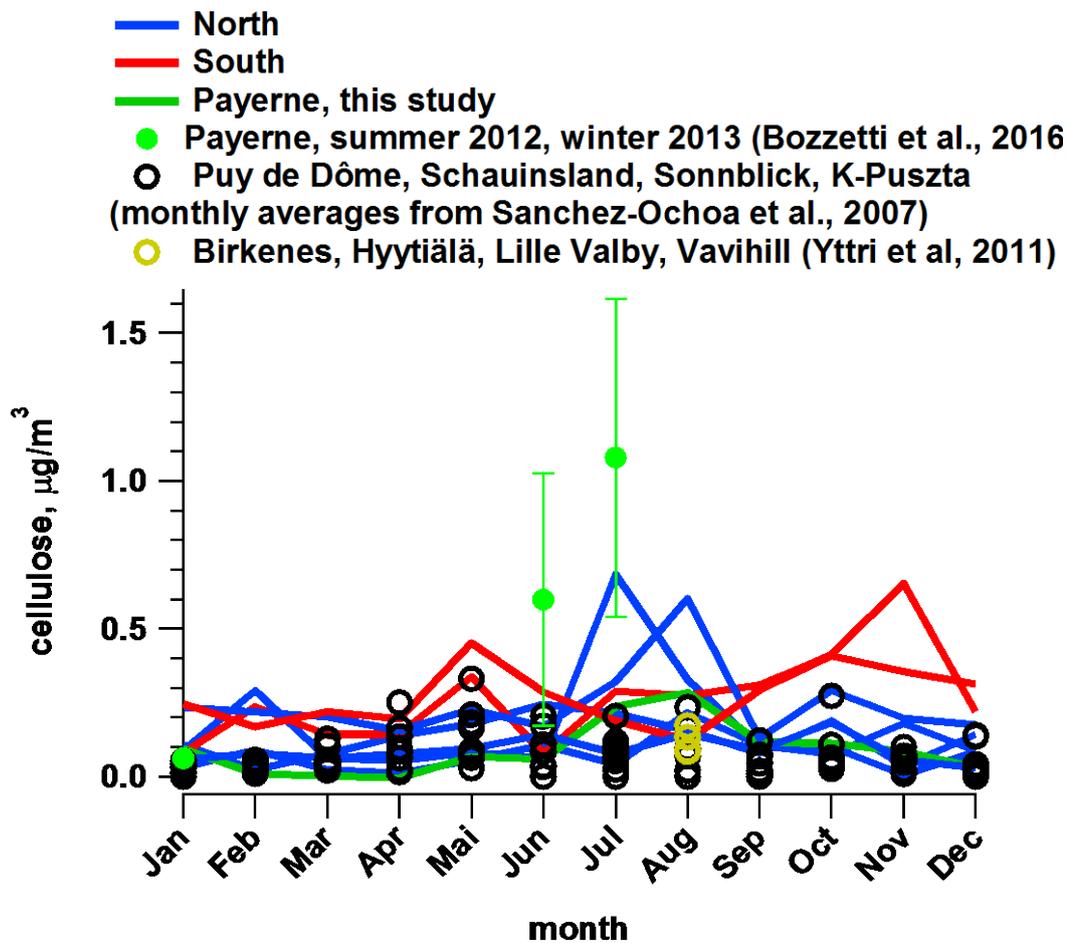
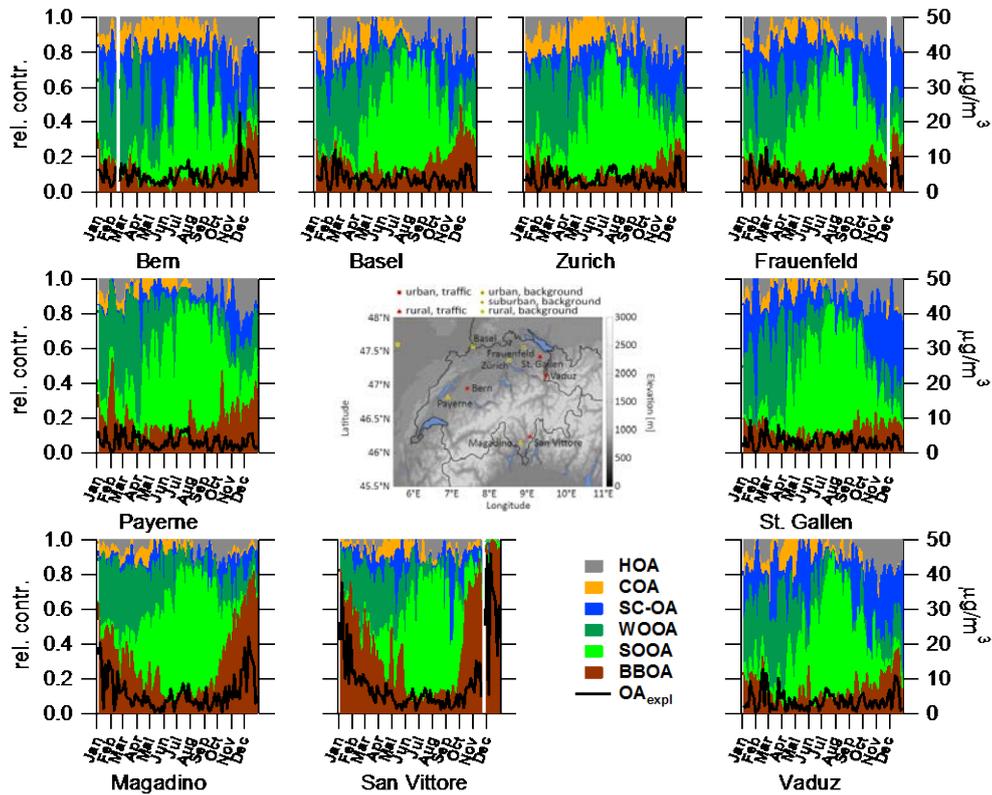


Figure 8: Cellulose concentrations ~~and PBOA concentrations (estimated based on cellulose concentrations)~~ as a function of the season and site. For comparison literature data from other years is added European sites: Payerne (Bozzetti et al., 2016, error bars representing the standard deviation of the measurements in June and July), Puy de Dôme, Schauinsland, Sonnblick, K-Puszta (Sanchez-Ochoa et al., 2007), Birkenes, Hyytiälä, Lille Valby, and Vavihill (Yttri et al., 2011).

5

Formatted: Justified



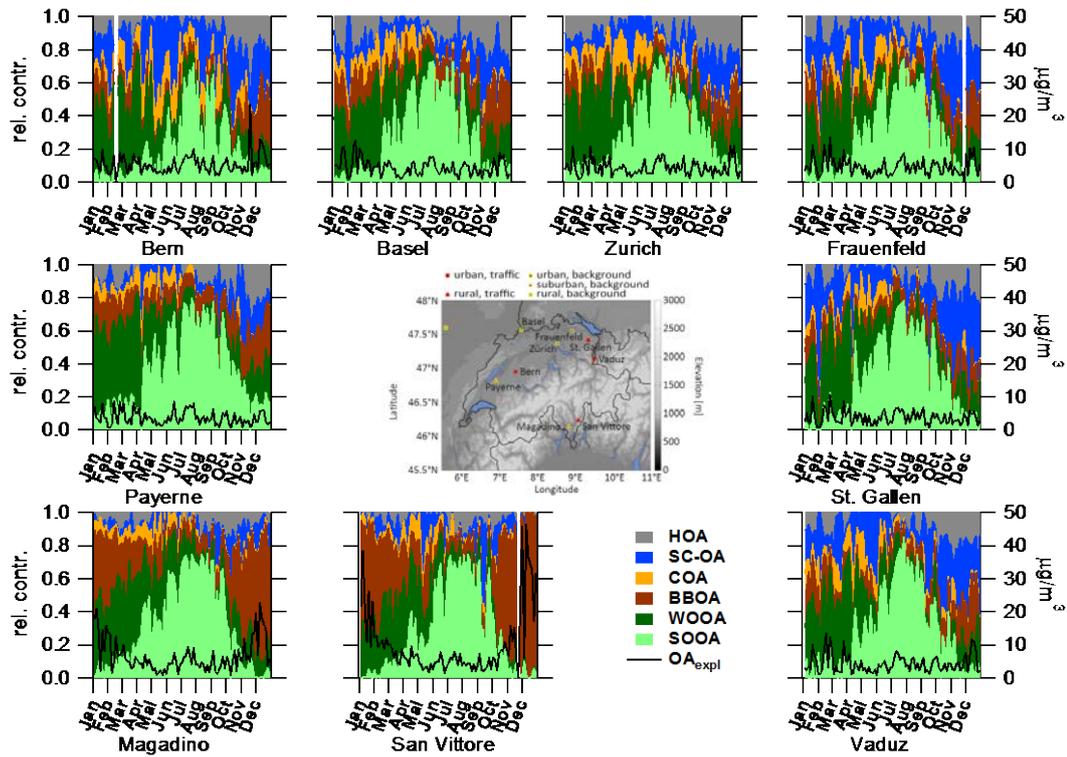


Figure 9: Map of Switzerland with yearly cycles. Negative concentrations were set to 0 prior to normalization for display. **The OA mass explained by the source apportionment analysis is termed OA_{expl} .**

Formatted: Justified

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)

|

Formatted: Justified

- Number of factors:

Based on the input data for PMF_{block} , we evaluate the influence of the number of factors, p , on the $Q_i/Q_{i,exp}$ -contribution- $Q_{i,j}$. 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.

Formatted: English (U.K.)

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

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

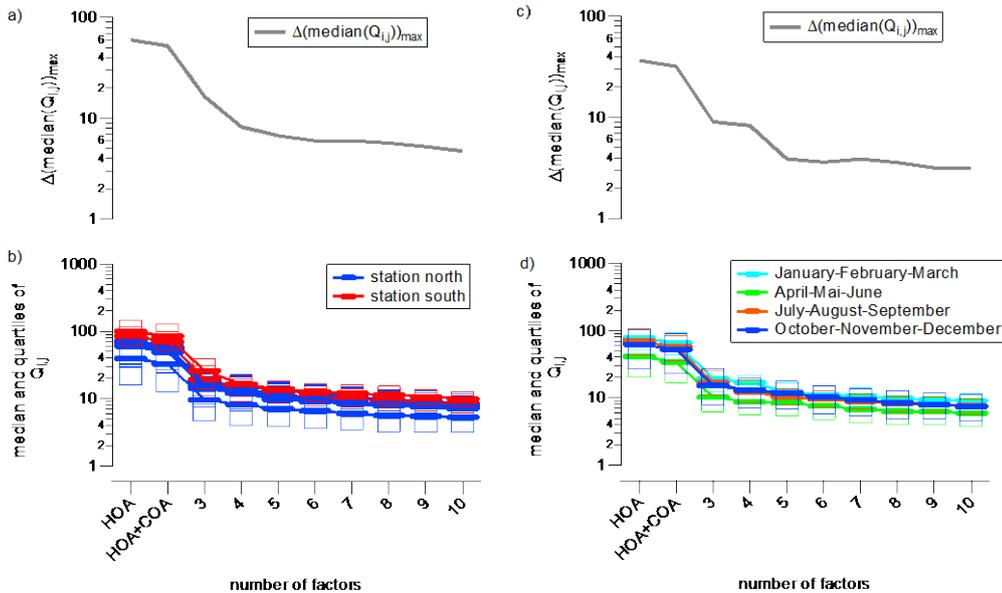


Figure S1: $Q_{i,j}$ as a function of the number of factors for a reference experiment with all data used in PMF (9 sites, full year 2013, HOA and COA constrained with $\alpha=0.0$ (b and d)). $\Delta(\text{median}(Q_{i,j}))_{\text{max}}$ is evaluated for the different periods during the year 2013 (January-February-March, 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-S1 shows the $Q_i/Q_{i,exp}$ -contribution- $Q_{i,j}$ 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 $Q_i/Q_{i,exp}$ -contributions-median $Q_{i,j}$. $\Delta(Q_i/\text{median}(Q_{i,exp}))_{\text{max}}$ shows the difference in the median

Formatted: Font: Not Italic

$Q_i/Q_{i,exp,j}$ between groups of points like sites or season. The smaller the $\Delta(\text{median}(Q_i/Q_{i,exp,j}))_{\text{max}}$, 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).

Formatted: Font: Italic

Formatted: Font: Not Italic

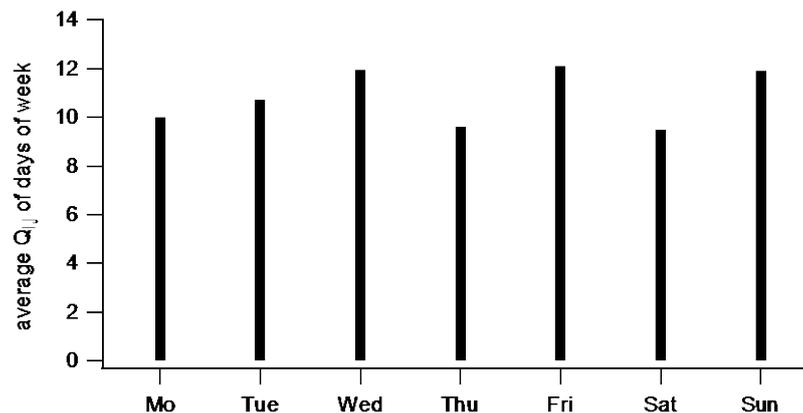
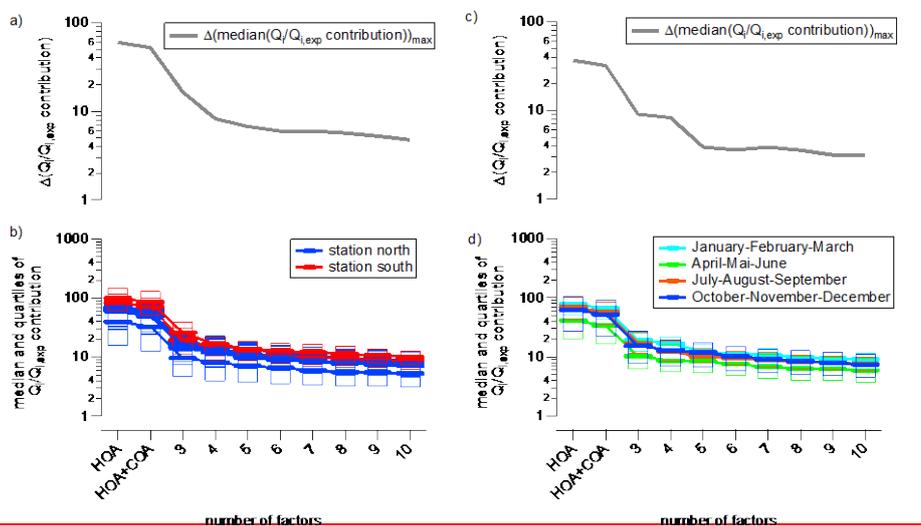


Figure S2: $Q_i/Q_{i,exp,j}$ as a function of the day of the week.

5 However, for PMF_{block} also with 6 factors, the average $O_{i,j}$ 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_{blocks} , the average $O_{i,j}$ for Zurich is slightly reduced for the same number of factors ~~for a reference experiment with all data used~~ when only including 1 site in PMF (9 sites, full $PMF_{zue,isol}$, $PMF_{zue,reprs}$, average $O_{i,j}$ 6). In this study, we analyse yearly cycles and, thereby, assume constant factor profiles throughout the year 2013, which can contribute to $Q > 1$.

Formatted: Font: 10 pt, Not Bold

Formatted: Font: 10 pt, Not Bold

Formatted: Font: 10 pt, Not Bold

10 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 $O_{i,j}$ 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 $O_{i,j}$ than hydrocarbon ions (CH, only C and H) and oxygenated ions ($CHO_{z=1}$ with 1 oxygen and $CHO_{z>1}$ with more than 1 oxygen). Since the time series of $CH_3SO_2^+$ is event-driven, the high $O_{i,j}$ of this ion hints to the fact that PMF is unable to accurately resolve all of these events.

15 The average $O_{i,j}$ 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).

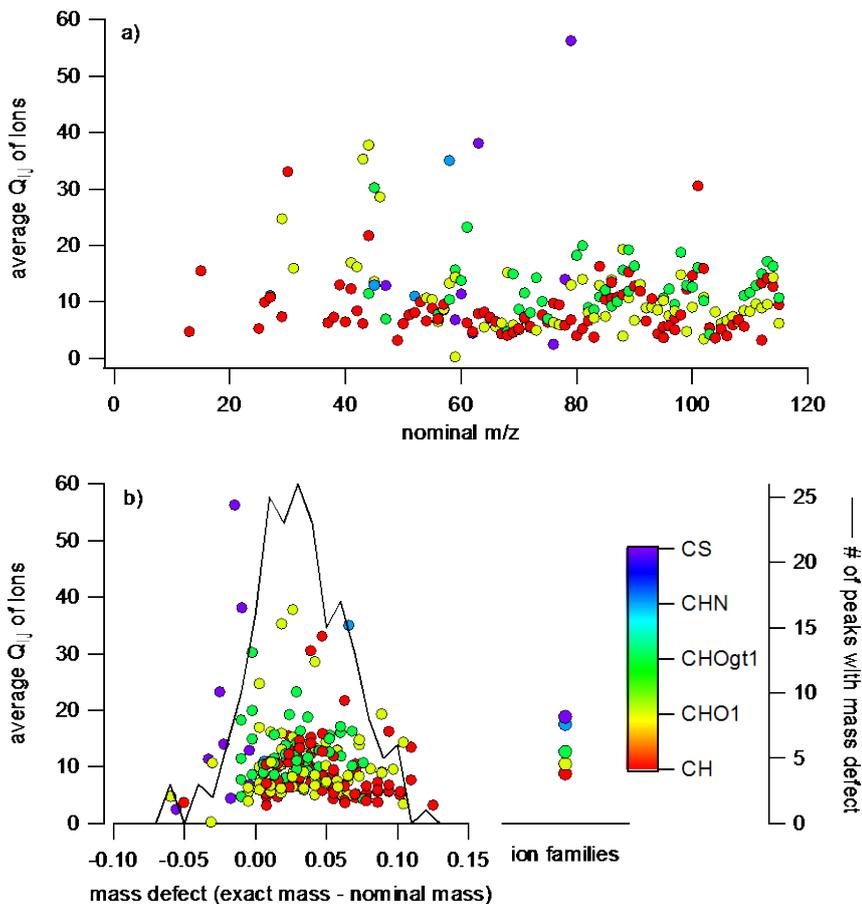


Figure S3: a) Average O_{ij} of ions in PMF_{block} as a function of their mass-to-charge ratio (m/z). The ions are color-coded with their composition (CH: ions consisting only of C and H; CHO1: ions consisting of C, H, and 1 O; CHOgt1: ions consisting of C, H, and more than 1 O; CHN: ions consisting of C, H, N, (and O); CS: ions consisting of C, H, S, (and O)). b) Average O_{ij} 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 O_{ij} of the ion families is displayed separately.

Cumulative density functions for the a-values of HOA and COA constrained with $a=0$ are presented for the accepted solutions in Fig. S4. We found that 80% of the accepted solutions have an a-value < 0.5 and AO/O_{exp} is evaluated for > 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

Formatted: Font: 10 pt, Not Bold, English (U.S.)

Formatted: Font: 10 pt, Not Bold, English (U.S.)

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 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_p/Q_{exp} and the lowest. factors do not show more than 1 distinct mode, indicating that we do not have several populations of solutions.

Formatted: Font: 10 pt, Not Bold, English (U.S.)

5

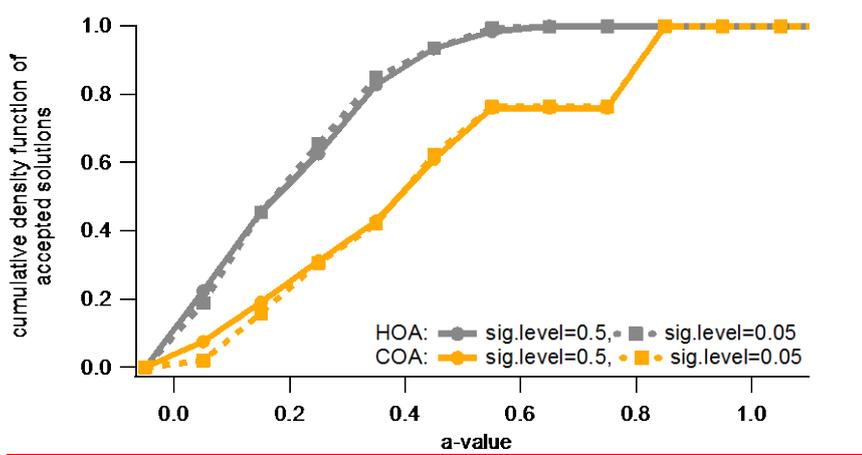


Figure S4: Cumulative density functions of a-values for HOA and COA for the accepted solutions.

The yearly average factor concentrations of all selected $\text{PMF}_{\text{block}}$ solutions after R_f correction are shown 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.

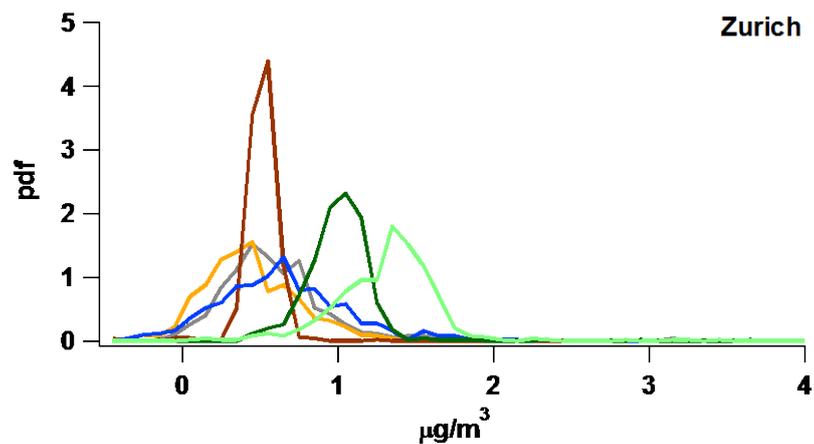


Figure S5: Histograms of yearly average factor concentrations of all selected $\text{PMF}_{\text{block}}$ solutions (after R_f correction).

5

Formatted: Font: 10 pt, Not Bold

Formatted: Normal, Space After: 10 pt

Formatted: Left, Space After: 0 pt, Line spacing: single

- Quality assessment of solutions:

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

Table S1.1S1: set of **acceptance** criteria used ~~to discard bad solutions~~. **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		
	$f(\text{CO}_2^+)$	$f(\text{C}_2\text{H}_4\text{O}_2^+)$
HOA	<0.4	<0.004
COA	<0.4	<0.01
Criteria on time series		
HOA	$r(\text{HOA}, \text{NOx}) > 0$ & $r(\text{HOA}, \text{NOx}) > r(\text{COA}, \text{NOx})$	
BBOA	$r(\text{BBOA}, \text{levo}) > 0$	
SC-OA	$r(\text{SC-OA}, \text{CH}_3\text{SO}_2^+) > 0$	
Mass closure criteria		
OC _{res} total	$Q_{25}(\text{OC}_{\text{res}} - \text{OC}) < 0$ & $Q_{75}(\text{OC}_{\text{res}} - \text{OC}) > 0$	
Magadino winter, Magadino summer	$Q_{25}(\text{OC}_{\text{res}} - \text{OC}) < 0$ & $Q_{75}(\text{OC}_{\text{res}} - \text{OC}) > 0$	
Zurich winter, Zurich summer	$Q_{25}(\text{OC}_{\text{res}} - \text{OC}) < 0$ & $Q_{75}(\text{OC}_{\text{res}} - \text{OC}) > 0$	
Magadino, Zurich	$Q_{25}(\text{OC}_{\text{res}} - \text{OC}) < 0$ & $Q_{75}(\text{OC}_{\text{res}} - \text{OC}) > 0$	
HOC<median, HOC>median	$Q_{25}(\text{OC}_{\text{res}} - \text{OC}) < 0$ & $Q_{75}(\text{OC}_{\text{res}} - \text{OC}) > 0$	
COC<median, COC>median	$Q_{25}(\text{OC}_{\text{res}} - \text{OC}) < 0$ & $Q_{75}(\text{OC}_{\text{res}} - \text{OC}) > 0$	
BBOC<median, BBOC>median	$Q_{25}(\text{OC}_{\text{res}} - \text{OC}) < 0$ & $Q_{75}(\text{OC}_{\text{res}} - \text{OC}) > 0$	
SC-OC<median, SC-OC>median	$Q_{25}(\text{OC}_{\text{res}} - \text{OC}) < 0$ & $Q_{75}(\text{OC}_{\text{res}} - \text{OC}) > 0$	
WOOC<median, WOOC>median	$Q_{25}(\text{OC}_{\text{res}} - \text{OC}) < 0$ & $Q_{75}(\text{OC}_{\text{res}} - \text{OC}) > 0$	
SOOC<median, SOOC>median	$Q_{25}(\text{OC}_{\text{res}} - \text{OC}) < 0$ & $Q_{75}(\text{OC}_{\text{res}} - \text{OC}) > 0$	
for PMF with 12 filters per site summer Magadino and Zurich	$Q_{25}(\text{OC}_{\text{res},i} - \text{OC}_i) < 0$ & $Q_{75}(\text{OC}_{\text{res},i} - \text{OC}_i) > 0$	
for PMF with 12 filters per site winter Magadino and Zurich	$Q_{25}(\text{OC}_{\text{res},i} - \text{OC}_i) < 0$ & $Q_{75}(\text{OC}_{\text{res},i} - \text{OC}_i) > 0$	

Formatted: Justified

Formatted: Justified

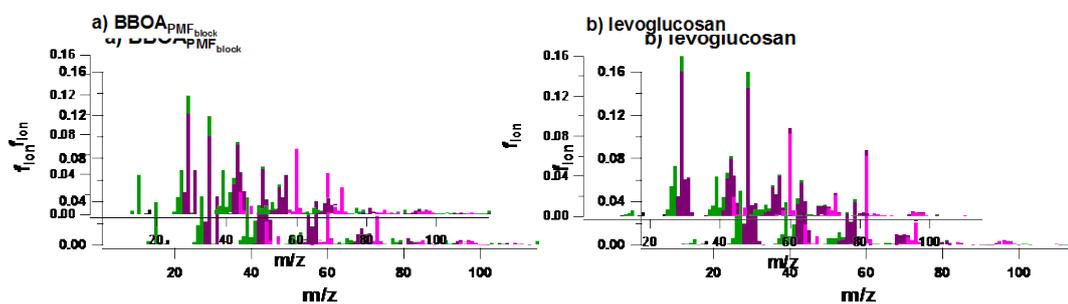
Formatted: Justified

|

Formatted: Justified

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

Figure S1.2S6 demonstrated the high similarity between the retrieved BBOA signature and the mass spectrum of nebulized levoglucosan.



5

10

Figure S1.2S6: 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.

Formatted: Font: 9 pt, Bold

Formatted: No underline

• 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 $R_{s,SOOA,O_3}=0.33$, Fig. S7.b) and Ox ($R_{s,SOOA,Ox}=0.38$, Fig. S7.c).

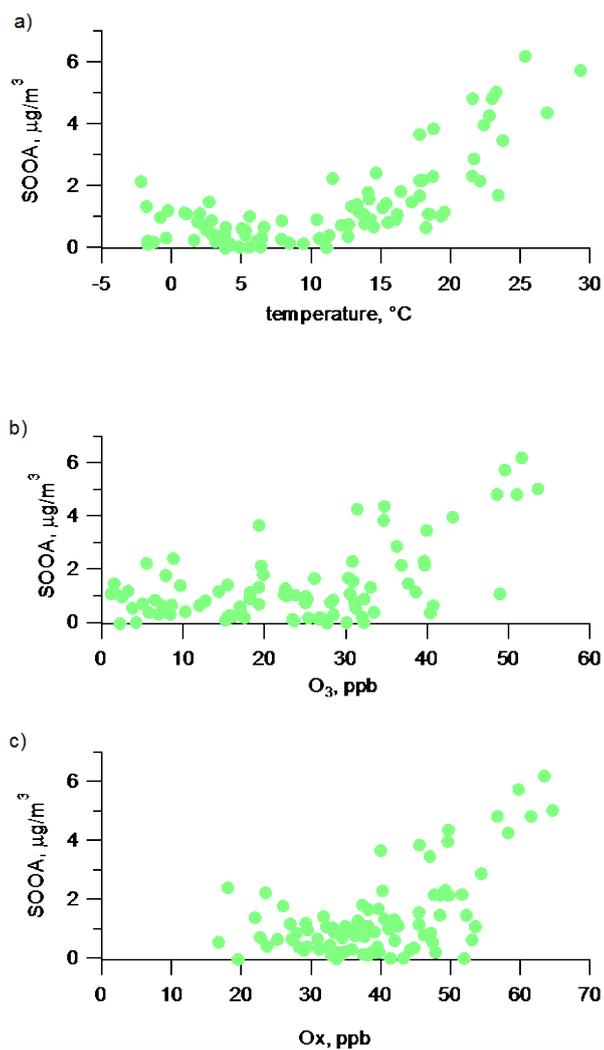


Figure S7: SOOA concentrations compared to temperature, ozone, and Ox (O_3+NO_2) for Zurich.

References:

5 ~~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., 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/aes.est.5b05960, 2016.~~

• Sensitivity to significance level of statistical tests in PMFblock:

For $\text{PMF}_{\text{block}}$, 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_n) 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.

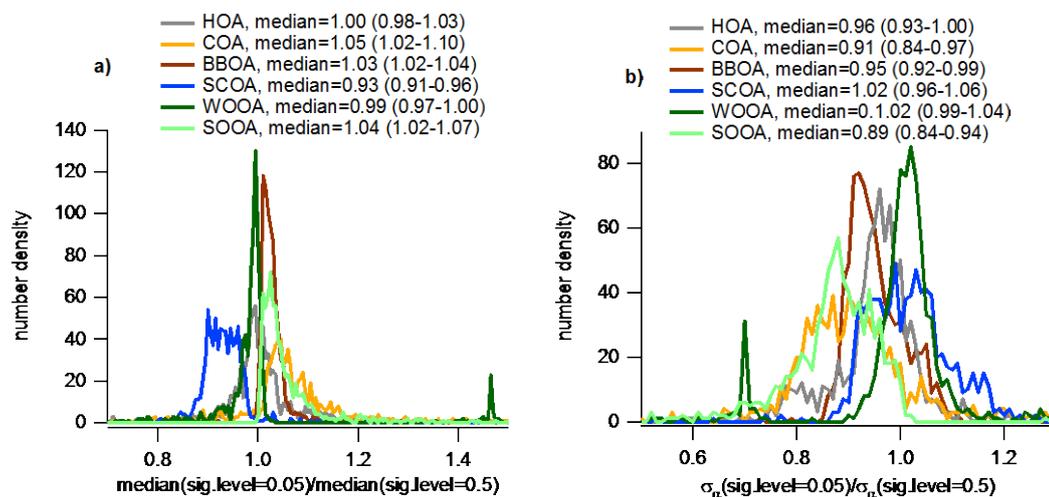


Figure S8: number density functions of source apportionment results obtained using a significance level of 0.05 normalized to results obtained using a significance level of 0.5: a) Comparison of factor concentrations b) Comparison of uncertainty estimate (σ_n).

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.