

Interactive comment on “Source apportionment of size and time resolved trace elements and organic aerosols from an urban courtyard site in Switzerland” by A. Richard et al.

A. Richard et al.

agnes.richard@psi.ch

Received and published: 2 July 2011

Dear Reviewer,

We thank you very much for your helpful and constructive comments and incorporated them to increase the quality of our manuscript. We answer to the comments point by point in the following:

General comments:

C5848

The handling of replaced RDI values is highly questionable. In using interpolated values in place of missing values, the values adjacent to the missing one are, in effect, being used twice each by the PMF algorithm and can therefore skew the solution. While substitution in this manner is permissible (due to the algorithm not being able to handle missing data points), the uncertainty associated with the inserted points should be set to a very high number (effectively infinity), such that they do not have any influence on the final solution.

In our current solution the uncertainty was set to 10% of the maximum concentration of that species. This is assuming already a large error, since the "missing values" are those, where the fit in the raw data spectra did not yield a reasonable value, which means that the concentration of that species at that certain date is very low. Nevertheless we have repeated all calculations assuming an uncertainty of a very high number ($100000 \mu\text{g m}^{-3}$). We have checked and listed the number of replaced values in Table 1, to give an overview. As already obvious from the Table, most values are replaced in the intermediate size range and only here we observed a tiny difference in the final PMF solution when increasing the uncertainty to $100000 \mu\text{g m}^{-3}$, see Figures 1-4. The solutions for the largest and the smallest size range didn't change at all, that's why we believe that the chosen uncertainty calculation for replacements of missing values is adequate.

The reason why we came up with the approach of replacing missing values with the adjacent values and increasing the corresponding uncertainty (it is actually a big advantage of PMF that missing values can be retained in the data set, rather than excluding a whole species from the data set!) is that we followed the experience of several authors (see following list) who applied the same approach:

- Kim and Hopke 2007, Source Identification of Atlanta Aerosol by Positive Matrix Factorization: "Missing values were replaced by the geometric mean of the measured values, and their accompanying uncertainties were set at 4 times this

C5849

geometric mean value."

- Huang, S.; Rahn, K.A.; Arimoto, R. Testing and Optimizing Two Factor-Analysis Techniques on Aerosol at Narragansett, Rhode Island; Atmos. Environ. 1999, 33, 2169-2185.
- Reff et al. 2007: "PMF2 requires that values be present in all entries of a data matrix for analysis. Missing species measurements in individual samples must, therefore, be dealt with in some way. Three approaches have been used in previous work. The first approach is to eliminate samples (rows of the data matrix) for which any measurement is missing. This is the approach generally used when either a key species or several species are missing measurements. A second approach is to eliminate the species (columns of the data matrix) from the PMF analysis completely. This is typically used when a large percentage of species' observations are missing. The third approach is to impute a value and associate a large uncertainty with this value so it has little influence in the PMF modeling. A common procedure for this third approach is to impute either the arithmetic or geometric mean species value for missing values and use a multiple, such as 3 or 4, of the mean concentration for the uncertainty value."
- Kim, E.; Larson, T.V.; Hopke, P.K.; Slaughter, C.; Sheppard, L.E.; Claiborn, C. Source Identification of PM_{2.5} in an Arid Northwest U.S. City by Positive Matrix Factorization; Atmos. Res. 2003, 66, 291-305.
- Polissar, A.V.; Hopke, P.K.; Malm, W.C.; Sisler, J.F. The Ratio of Aerosol Optical Absorption Coefficients to Sulfur Concentrations, as an Indicator of Smoke from Forest Fires when Sampling in Polar Regions, Atmos. Environ. 1996, 30, 1147-1157.
- Song, X.H.; Faber, N.M.; Hopke, P.K.; Seuss, D.T.; Prather, K.A.; Schauer, J.J.; Cass, G.R. Source Apportionment of Gasoline and Diesel by Multivariate Calibration Based on Single Particle Mass Spectral Data; Anal. Chim. Acta 2001, 446, 327-343.

bration Based on Single Particle Mass Spectral Data; Anal. Chim. Acta 2001, 446, 327-343.

The authors provide insufficient information regarding their choices of the number of factors. They should explain why the higher-order solutions were rejected. In particular, there is evidence to suggest that the HOA factor contains a contribution from cooking. Lanz et al. (2007) found that this could be resolved if higher-order solutions were used. Also, Allan et al. (2010; <http://www.atmos-chem-phys.net/10/647/2010/acp-10-647-2010.html>) were able to see this with 4-factor solutions in London and Manchester.

We will include the following discussion about the choice of the final number of factors in the supplementary material.

Determination of the number of factors of the AMS-PMF solution:

To identify the most appropriate number of factors, PMF was run by varying the number of factors between 2 and 7. The most reasonable number of factors was determined following the interpretability of the identified factors (according to the methods applied by Lanz et al. (2007) for measurements at the same sampling site) and the criteria of the "Uniqueness of the derived factors" as discussed by Allan et al. (2010).

In the following the different results for varying number of factors are presented, all calculations are performed with $F_{\text{peak}} = 0$.

2-factor solution:

In the 2-factor solution the first factor presents the typical features shown by OOA: its mass spectrum is clearly dominated by 44 (and therefore also 28, which is calculated from the m/z 44 peak), while time series of this factor correlate with secondary

AMS-inorganics ($R^2 = 0.7$ for NH_4^+ , 0.64 for SO_4^{2-} and 0.59 for NO_3^-). Concerning the second factor, its mass spectrum is dominated by m/z 's 27, 29, 41, 43, 55, 57, and presents also relevant peaks at m/z 's 60 and 73. These are fragments mainly related to both HOA and BBOA. Due to the presence of a mixed factor the 2-factor solution cannot be considered the optimal one.

3-factor solution:

The 3-factor solution yielded oxygenated (OOA), hydrocarbon-like (HOA) and biomass burning (BBOA) organic aerosol. All the factor profiles and time series were mostly unique without evident correlations of different factors. The identified factors are in agreement with the same three sources detected by Lanz et al. (2008) in an earlier winter measurement campaign at the same site.

4-factor solution:

Increasing the number of factors to 4 did not yield a reasonable solution since the OOA factor was split into a profile dominated by 44-28 and a second one dominated by 28-29-43-44 (Fig. 7). Furthermore the time series are not unique anymore, since the new factor dominated by 44-28 and the BBOA factor have the same time trend in the beginning of the measurement campaign as can be seen from Figure 8.

5-factor solution:

Increasing the number of factors even one step further, did not yield any reasonable improvements either, since now there are three factors showing a similar time trend in the beginning of the measurement campaign. Furthermore, this new factor shows again mainly peaks for 28 and 44 and therefore does not contain any new information and is also not related to a food cooking factor.

C5852

6- and higher factor solutions:

These solutions lead to increased splitting of the three identified sources (OOA, HOA, BBOA), but not to a food cooking factor.

In conclusion, no food cooking factor could be identified with higher-order solutions. Lanz et al. (2008) did not find a cooking factor for Zürich Kaserne in winter either, even with the hybrid model approach. In contrast, a food cooking factor was found by Lanz et al. (2007) in summer and is rather attributed to leisure activities such as barbecues and charbroiling than domestic cooking. It therefore has to be considered as a seasonal source only. In the surrounding of Zürich Kaserne there are hardly any open places where food is sold, such as hot-dog-stands. In addition, the majority of restaurant kitchens are equipped with filters and these facts presumably prevent the identification of a cooking factor at this site in winter. Allan et al. (2010), supplementary material, noted that "... if a solution set with a given number of factors was deemed unreliable, those with greater than this number also failed." This was also found in this study.

The AMS PMF output seems to feature a very significant spike late on Dec 10 in the HOA series. Given that this is unlikely to be representative of the general behaviour of HOA, it should have been filtered. Either way, the scale of its trace on Fig 8 should be expanded, because the reader isn't able to easily see its temporal behavior currently.

This significant spike (three values out of 5000 values in total, which do not significantly influence the mean value) can be related to a diesel vehicle directly in the courtyard of Zürich Kaserne passing directly in front of the measurement container (tracked by a webcam that is installed at the measurement site). The calculations were always performed in the robust mode (see manuscript page 9, Sect. 2.3) where this outlier is

C5853

automatically excluded from the analysis.

The scale of Figure 8 was changed according to the suggestions.

Specific comments:

P3732, L11: Given the number of artefacts associated with the Aethalometer (<http://www.atmos-meas-tech.net/3/457/2010/amt-3-457-2010.html>) the authors should specify if any corrections were made.

Yes, the Aethalometer data were corrected with the correction method suggested by Weingartner et al. (2003). We added a sentence to the manuscript: "Due to shadowing effects that can occur on the quartz fiber filters with increasing loads of light absorbing particles, a correction as proposed by Weingartner et al. (2003) with an f parameter of 1.2 was applied to the Aethalometer data for the calculation of BC concentrations."

P2732, L25: At certain sites, the local primary emissions become more distinct during the winter rather than the summer. The authors should cite any previous studies where they have determined this to not be the case here.

Lanz et al. (2007, 2008) also found that sources might not be separated as clearly in winter as in summer since time series are less dependent from meteorological conditions as compared to winter times with inversions. ("Unlike in summer, primary pollutants from different sources, therefore, can show similar temporal variabilities at the background site of Zurich-Kaserne in winter, limiting the applicability of bilinear mixing models."). Moreover, Lanz et al. (2008) only reached a separation of primary components in winter through the use of a hybrid approach with ME-2, where an

C5854

priori HOA profile was fixed. We therefore added the respective citations in the new version of the manuscript.

P3739, L7: The authors state that many different combinations of data sets were tested, but do not discuss the results of these tests.

We inserted a sentence to the revised manuscript: "The result of these tests was a less clear separation of sources, or elements being mixed into factor profiles that do not match the emission source."

P3751, L27: Agreement between PMF2 and ME2 does not confirm the robustness of the solution; it merely shows that that the two algorithms are producing equivalent results. To properly test numerical robustness, the authors should try varying the initialisation seed or perform bootstrapping analysis.

We have changed the sentence to: "Calculations were accomplished with both the PMF2 and the ME-2 algorithm, which produced equivalent results and showed good agreement. Calculations were performed with varying seed values and no significant difference in the solutions was observed."

Technical corrections:

P3737, L22: Rather than the Igor Pro version number, the version number of the analysis code should be given. Also, the web address given is obsolete and does not host TOF-AMS analysis code; the reader should be directed to <http://cires.colorado.edu/jimenezgroup/ToFAMSResources/ToFSoftware/index.html>

C5855

instead.

We agree and replaced the IGOR pro version number to the number of the analysis code as well as the URL in the text, so that it reads now: ... and was then transferred into organic-equivalent concentrations (org eq. μgm^{-3}) using the ToF-AMS Analysis toolkit version 1.47 based on the work of Allan et al. 2003 (<http://cires.colorado.edu/Jimenez-group/ToFAMSResources/ToFSoftware/index.html>).

P3738, L12: Presumably, the PMF2-equivalent script was used for ME2. In which case, this should be stated.

We changed the sentence to: "In this study two programs to solve Eq. (2) were tested with the same RDI input data: PMF2 (version 4.2, Paatero, 2004) and the Multilinear Engine (ME-2, Paatero, 1999, PMF2-equivalent script)."

P3745, L10: Technically, ME2 was the algorithm used by Lanz et al., not the model.

We agree and we changed it: "However, a hybrid receptor model solved by ME-2 including HOA as a priori source was used there, whereas no initial assumptions about sources were made here."

Fig. 3: The y axes should extend to the full range of the data.

We agree and we changed the axis of the figure.

C5856

Figs (general): Much of the text on the axis labels are so small as to be unreadable.

We agree and we changed the labels of the figures.

Figs. 10, 11 and 10S: The variations in font size of the Pearson's r values only serve to make the figure harder to read (especially on figure 10S). I would make the font size uniform.

We decided to keep the changing font size, so that the higher correlations are more easily visible.

References:

J. D. Allan, P. I. Williams, W. T. Morgan, C. L. Martin, M. J. Flynn, J. Lee, E. Nemitz, G. J. Phillips, M. W. Gallagher, and H. Coe; Contributions from transport, solid fuel burning and cooking to primary organic aerosols in two UK cities, *Atmos. Chem. Phys.*, 10, 647-668, 2010, SI.

Lanz, V. A., Alfarra, M. R., Baltensperger, U., Buchmann, B., Hueglin, C., and Prevot, 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.

Lanz, V. A., Alfarra, M. R., Baltensperger, U., Buchmann, B., Hueglin, C., Szidat, S., Wehrli, M. N., Wacker, L., Weimer, S., Caseiro, A., Puxbaum, H., and Prevot, 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,

C5857

2008.

Weingartner, E., Saathoff, H., Schnaiter, M., Streit, N., Bitnar, B., and Baltensperger, U.: Absorption of light by soot particles: determination of the absorption coefficient by means of aethalometers, *J. Aerosol Sci.*, 34, 10, 1445-1463, 2003.

Interactive comment on *Atmos. Chem. Phys. Discuss.*, 11, 3727, 2011.

C5858

Table 1. Number of replaced values. (The dash ("-") means that this element is not included in the calculations for the corresponding size range.

Element	PM _{10-2.5}	PM _{2.5-1}	PM _{1-0.1}
Al	22	-	-
Si	24	22	56
P	-	108	44
S	2	0	0
Cl	2	122	-
K	2	2	0
Ca	0	2	10
Ti	2	0	10
V	-	-	46
Cr	-	2	-
Mn	4	4	30
Fe	0	0	0
Co	32	78	-
Ni	-	52	66
Cu	0	0	0
Zn	0	0	0
Sr	32	10	60
Zr	0	2	0
Mo	8	2	0
Sn	0	0	0
Sb	0	0	0
Ba	8	2	24
Pb	-	20	2
Total	138	428	348

C5859

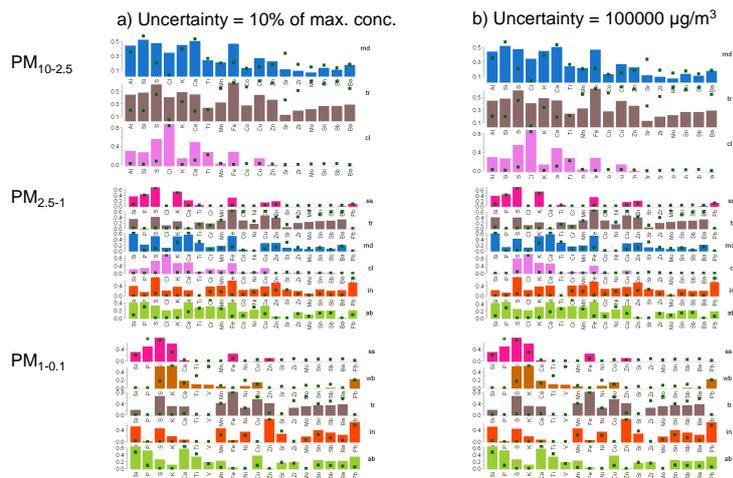


Figure 1: Comparison of PMF calculations with a) uncertainty of missing values replaced by 10% of maximum concentration of that species and b) uncertainty replaced by a very high number (100000 $\mu\text{g m}^{-3}$) for all three size ranges. Abbreviations for factors mean: md = mineral dust, tr = traffic, cl = de-icing salt, ss = secondary sulfate, in = industrial, ab = anthropogenic background and wb = wood burning.

Fig. 1.

C5860

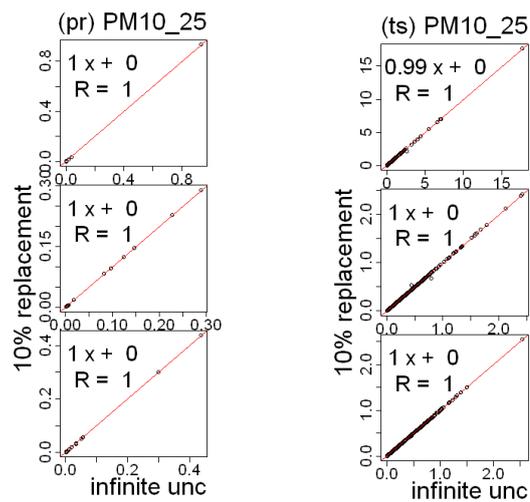


Figure 2: Correlations of profiles (panel (pr)) and time series (panel (ts)) of the two PMF solutions with "infinite" uncertainty and the previous 10% of maximum concentration replacement for stage $\text{PM}_{10-2.5}$.

Fig. 2.

C5861

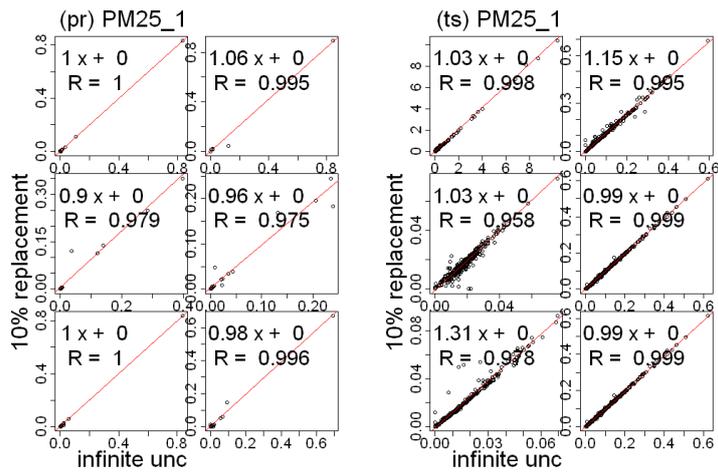


Figure 3: Correlations of profiles (panel (pr)) and time series (panel (ts)) of the two PMF solutions with "infinite" uncertainty and the previous 10% of maximum concentration replacement for stage PM_{2.5-1}.

Fig. 3.

C5862

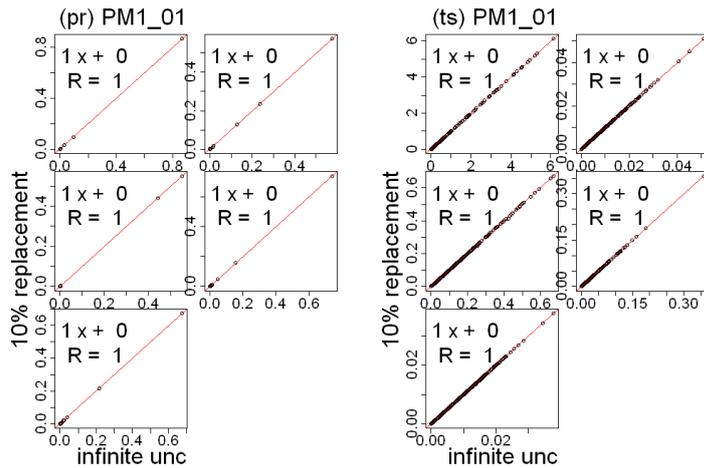


Figure 4: Correlations of profiles (panel (pr)) and time series (panel (ts)) of the two PMF solutions with "infinite" uncertainty and the previous 10% of maximum concentration replacement for stage PM_{1-0.1}.

Fig. 4.

C5863

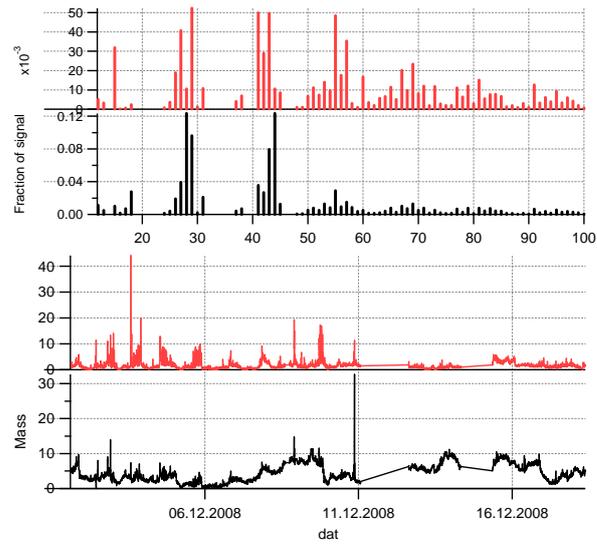


Figure 5: 2-factor solution, red = BBOA and black = HOA.

Fig. 5.

C5864

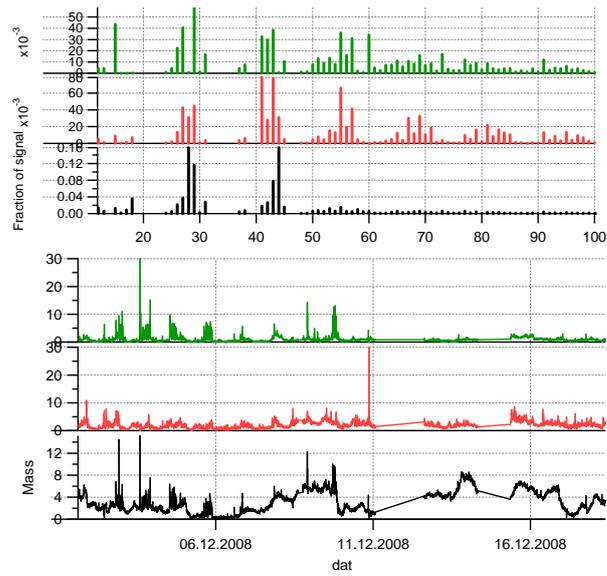


Figure 6: 3-factor solution: green = BBOA, red = HOA, black = OOA.

Fig. 6.

C5865

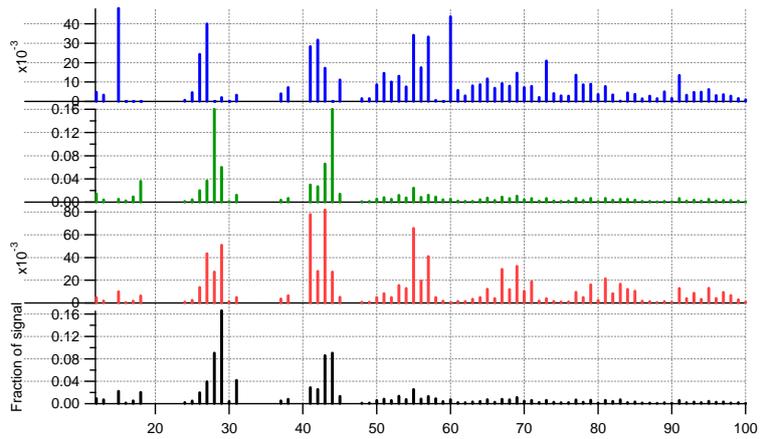


Figure 7: Factor profiles for the 4-factor solution of AMS compounds: BBOA (blue), first OOA (green), HOA (red) and the second OOA (black) factor.

Fig. 7.

C5866

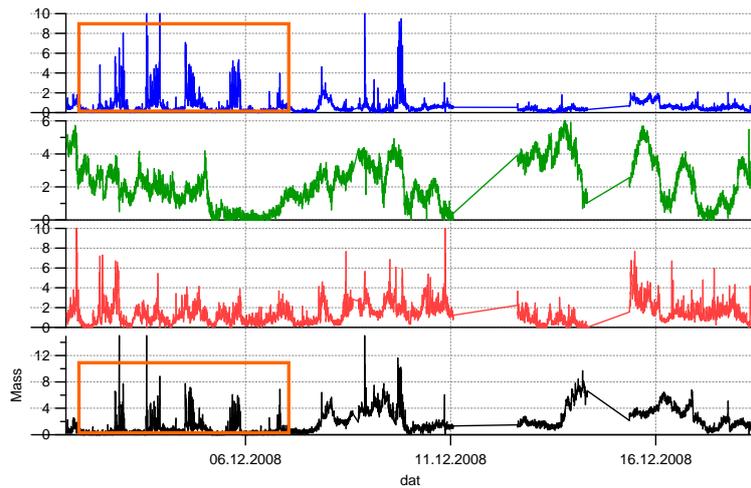


Figure 8: Time series of 4-factor solution of AMS compounds: BBOA (blue), first OOA (green), HOA (red) and the second OOA (black) factor.

Fig. 8.

C5867

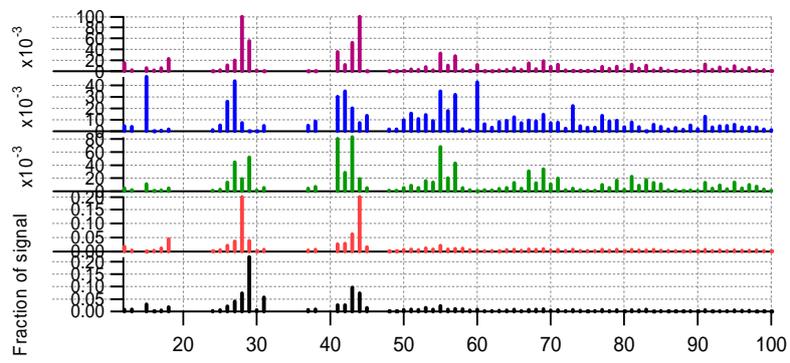


Figure 9: Profiles of 5-factor solution.

Fig. 9.

C5868

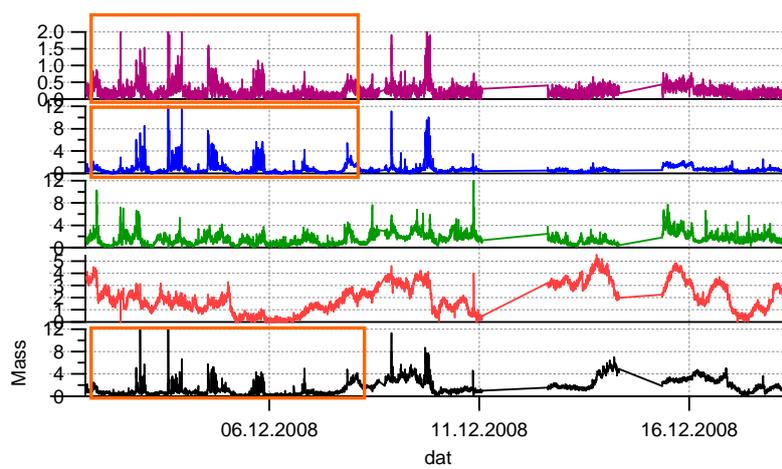


Figure 10: Time series of 5-factor solution.

Fig. 10.

C5869