

Supplement of Atmos. Chem. Phys., 15, 11885–11907, 2015  
<http://www.atmos-chem-phys.net/15/11885/2015/>  
doi:10.5194/acp-15-11885-2015-supplement  
© Author(s) 2015. CC Attribution 3.0 License.



*Supplement of*

## **Black carbon surface oxidation and organic composition of beech-wood soot aerosols**

**J. C. Corbin et al.**

*Correspondence to:* J. C. Corbin ([joel.corbin@psi.ch](mailto:joel.corbin@psi.ch)) and A. A. Mensah ([amewu.mensah@env.ethz.ch](mailto:amewu.mensah@env.ethz.ch))

The copyright of individual parts of the supplement might differ from the CC-BY 3.0 licence.

This supplement contains a detailed description of the SP-AMS data analysis procedures used, and additional plots for the PMF models presented in the paper including residual plots and additional time-series details.

## Extended AMS Data Analysis Details

### PIKA fitting procedure

#### Adjustments to the Peak Width determination

The width of ion peaks in the AMS typically has a linear dependence on  $m/z$ , as discussed and shown by DeCarlo et al. (2006). To allow for the analysis of overlapping mass-spectral peaks, a linear peak-width function  $w(m/z)$  is determined by least-squares regression of measured peak widths against  $m/z$  in the PIKA software. In the standard software, peak widths are “measured” by performing unconstrained Gaussian fits to user-defined reference peaks and averaging the results. The reference peaks are required to be (i) well-resolved from their neighbours and (ii) present at sufficient intensity to allow a successful fit. Normally, peaks which do not meet these two criteria are excluded from analysis by inspection of the raw mass spectrum for case (i) and omitting the entire mass spectrum from analysis for case (ii) (Sueper et al., 2011).

In the present data set, the above two criteria were met by very few mass spectra. While several peaks met criterion (i), they were not all present simultaneously, so that applying criterion (ii) resulted in a very small number of reference data. To address this issue, we changed the PIKA software so that the 25% trimmed mean (interquartile mean) rather than the mean was used to average peak-width measurements. This allowed the entire data set to be included in the peak-width analysis. The results were not sensitive to the exact value used for trimming. Exclusion of outliers at this stage of the analysis is justified because these outliers represent the imprecision of the fitting routine [normally due to violation of criterion (ii) above] and not variability in the instrument itself.

The overall peak-width function was thus determined as follows. Clearly isolated peaks were used to generate a list of candidate ions, and a provisional peak-width function  $w(m/z)$  was found. This original candidate-ion list contained only about five peaks, and no peaks in the range  $20 < m/z < 50$ . The list was then extended stepwise by adding peaks which were not perfectly isolated, but whose neighbours appeared to be well-resolved. Of these peaks, most followed the trend of the clearly-resolved peaks. Those that did not were found to have suffered from interference in special situations, for example,  $\text{Zn}^+$  ions were not always well-resolved from  $\text{SO}_2^+$  ions. The resulting averaged peak-widths were weighted by their trimmed variance (the variance of the trimmed values, Wilks, 2011) to determine  $w(m/z)$  via lin-

ear regression against  $m/z$ . The relative uncertainty in the fitted slope and intercept was 2.8% and 0.4% respectively for the SP-AMS. For the AMS, these values were 2.3% and 0.34%. Since the exact value of this result depended on the user-chosen candidate ions and on the value used for trimming, the overall uncertainty  $\sigma_w$  was taken as 2.5%.

Three other considerations were kept in mind while determining the peak-width function. First, ions generated largely from gas-phase species (e.g.  $\text{Ar}^+$ ,  $\text{CO}_2^+$ ) were found to have a different trend than particulate ions and were avoided. (These ions were also avoided during  $m/z$  calibration.) This difference is likely due to the particular mass-spectrometer tuning of our instrument (DeCarlo et al., 2006). Second, ions at very low  $m/z$  ( $\leq 13$ ) such as  $\text{C}^+$  were typically represented by just one or two detector bins because of their short time-of-flight and the finite temporal resolution of the detector. The “width” of such a signal is not measurable. Third, peaks which were largely associated with ash –  $\text{K}^+$ ,  $\text{Cl}^+$ ,  $\text{Na}^+$  and others – also showed different shapes and were excluded.

#### Adjustments to the pseudogaussian peak-shape function

PIKA uses a pseudogaussian “peak shape” function  $f_0$  to represent the shape of mass spectral peaks (DeCarlo et al., 2006; Sueper et al., 2011). In this study, the  $f_0$  was relatively similar to a true Gaussian function  $G_0$ . The  $f_0$  was defined according to standard procedures, by averaging the isolated peaks described in Section 5 and excluding peaks which showed interference from neighbouring ions. Also following standard procedures, the  $f_0$  was forced to be monotonic and to have values greater-than-or-equal-to the equivalent Gaussian function for values  $> 3$  standard deviations away from the mode of the equivalent  $G_0$ . Finally, an additional adjustment was performed where the  $f_0$  was smoothed by a binomial filter for values  $> 1.3$  standard deviations away from the  $G_0$  mode. Because the slope of  $f_0$  was high at 1.3 standard deviations, the smoothed ( $> 1.3$  deviations) and unsmoothed data ( $< 1.3$  deviations) could be directly combined without introducing discontinuities to the final  $f_0$ .

#### PMF Modelling details

As mentioned in Section 2.5, the PMF solutions may be considered “true” solutions in the sense that there was no rotational ambiguity (Paatero and Tapper, 1994; Paatero et al., 2002) because of the large number of zeroes measured (Paatero et al., 2002). These zeroes reflected, for example, the facts that BC is produced during flaming combustion and that volatile organics were mainly emitted at the start of a burn.

In general, PMF provided good fits to each measured ion in terms of the residual distribution (Zhang et al., 2011), with two notable exceptions. First, residuals of the highest-signal ions (as seen in Figs. 3 and 4 of the main paper) showed sig-

nificant outliers from high-loading periods. This may have reflected the fact that SP-AMS background measurements occurred on a timescale comparable to the spikes associated with many of these high-loading signals. Second, for  $\text{CO}^+$  the opposite behaviour was observed: its uncertainty-weighted residuals were much smaller than expected. This suggests that the variability in fitted  $\text{CO}^+$  signals was smaller than expected, possibly due to an influence from its poorly-resolved neighbour,  $\text{N}_2^+$ , or due to the assignment of large uncertainties to the smaller  $\text{CO}^+$  signals (Section 2.2).

The POM–Start and POM–Flame mass spectra were quite similar. Some mixing or splitting of these factors (Ulbrich et al., 2009), likely related to the fact that wood-burning emissions are not a linear sum of distinct PM components, was therefore suspected. We recombined the POM–Start, POM–Flame, and Fresh–BC factors with arbitrary time series into an artificial data matrix, using time series of linear and sine functions, and found that PMF successfully resolved freshBC but not always POM–Start and POM–Flame from this matrix. We therefore suspect that the mass spectra of POM–Start and POM–Flame would change slightly if the present experiments were repeated.

### Choice of factors

#### Fresh emissions

For the AMS, a one-factor PMF solution was able to fit the fresh-combustion data with a  $Q/Q_{\text{exp}}$  of 2. The ratio  $Q/Q_{\text{exp}}$  is unity for a perfect PMF model (Paatero and Tapper, 1994; Ulbrich et al., 2009). (A “perfect” PMF model is one where the data comprises a linear combination of fixed factors and all uncertainties have been correctly specified.) This already-low  $Q/Q_{\text{exp}}$  indicated a relatively homogeneous OM composition.

With two factors, PMF provided a good fit to the AMS data with no apparent rotational ambiguity and  $Q/Q_{\text{exp}} = 0.99$ . Tentatively increasing the number of factors to three allowed PMF to better explain the pyrolysis spikes at the start of each batch combustion (Fig. 1) and lowered  $Q/Q_{\text{exp}}$  to 0.78. As the model thus appeared to overexplain the data and explaining these spikes was not the goal of this analysis, the two-factor solution was used.

For the SP-AMS, the residuals of PMF analyses with less than six factors were dominated by the first two measurements of the day. These residuals were explained only with a seven-factor solution, which allowed the initial signals to be identified as originating from the tinder used to start the fire. The Tinder factor, discussed in detail below, was not present at any other time and its representation was not the goal of this analysis. So the first two measurements were downweighted threefold (Paatero and Hopke, 2003) and the PMF analysis was repeated. After excluding these points, a two-factor solution yielded  $Q/Q_{\text{exp}}$  of 2.4 and considerable residual structure (in both dimensions,  $m/z$  and time)

so a three-factor solution was sought. The three-factor solution reduced  $Q/Q_{\text{exp}}$  by 30%. A tentative four-factor solution was explored and rejected for the same reason as for the AMS.

#### Aged emissions

For the aged-emissions experiment, the choice of PMF factors was similar to the fresh emissions. However, the failed-start burn (c.f. Fig. 2) introduced additional complications into the analysis because the stove was opened to add tinder and manually ignite the fire.

For the AMS PMF analysis, a one-factor solution gave  $Q/Q_{\text{exp}} = 2.1$ , reflecting the overall chemical homogeneity of the OM. This homogeneity was expected given the extensive oxidation observed within the MSC and discussed in Corbin et al. (2015a). Adding a second factor improved the model fit, but the residual mass spectrum still contained significant structure. A three-factor solution explained and removed this structure by adding a factor to describe the tinder. Adding further factors served to better explain the failed-start burn, which emitted spherical OM particles with little BC (Corbin et al., 2015a) and is therefore beyond the scope of this analysis. The three-factor solution is reported.

For the SP-AMS, it was expected that four factors would be required (three AMS factors plus BC). However, a four-factor solution did not identify the tinder factor, and a five-factor solution was therefore explored. The five-factor solution succeeded in isolating a tinder factor. It also contained two similar OM factors (unc centred correlation  $> 0.87$  in time and  $> 0.90$  in  $m/z$  space) reflecting the starting phase (c.f. POM–Start above). Since the four-factor solution had already provided a satisfactory description of the starting phase and the goal of this analysis is to explore the composition of BC, the two starting-phase factors are reported as one by summing their mass loadings and presenting the (comparable) mass spectrum of the more-abundant factor.

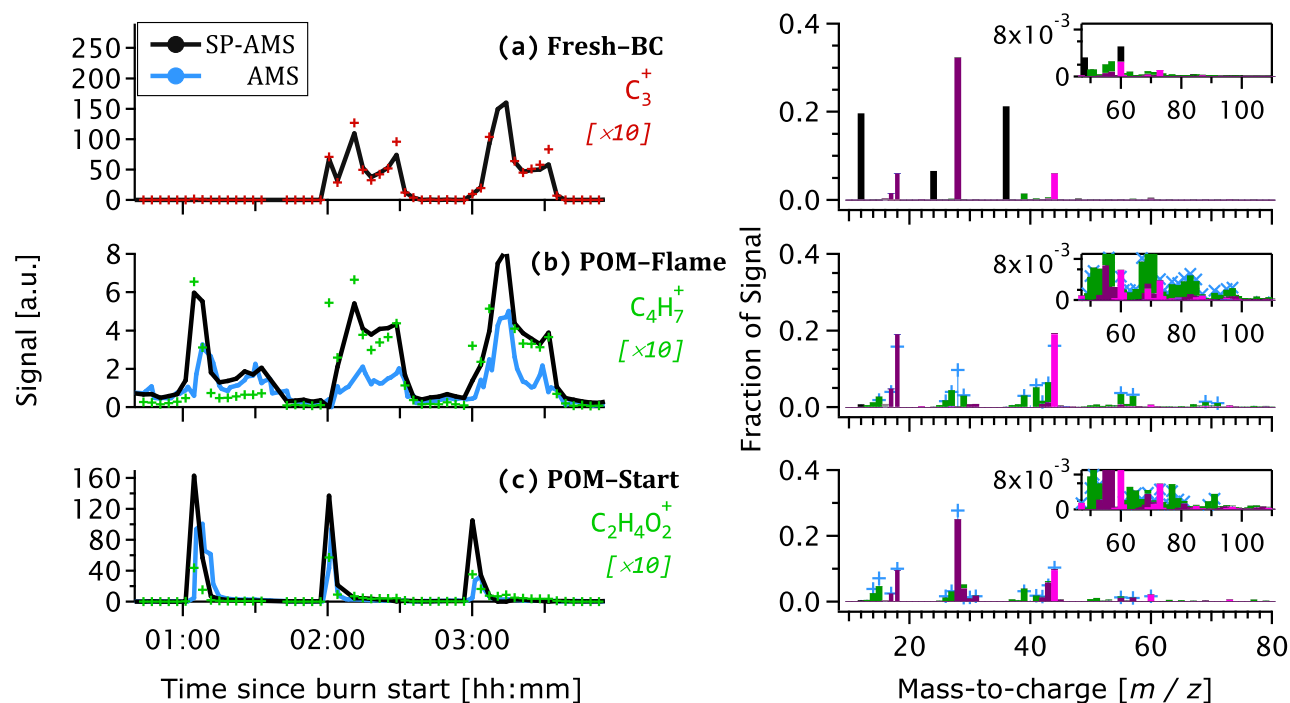
#### Residual plots

For the residual plots, the scaled residuals are presented as a function of  $m/z$  and of time. The residuals as a function of time show that the majority of the unexplained variance occurs at the start of each burn. Adding more factors typically changed only this feature (affecting the POM/OOM–Start factors) but not the other features of the model (BC and POM/OOM–Flame factors). That is, the majority of the model uncertainty was affected by POM/OOM–Start, which was ultimately the simplest factor to interpret. Therefore, statistics comparing model fits (such as  $Q / Q_{\text{exp}}$ ) would mainly convey the variability in POM/OOM–Start and mislead the reader. Additional information is available from the authors.

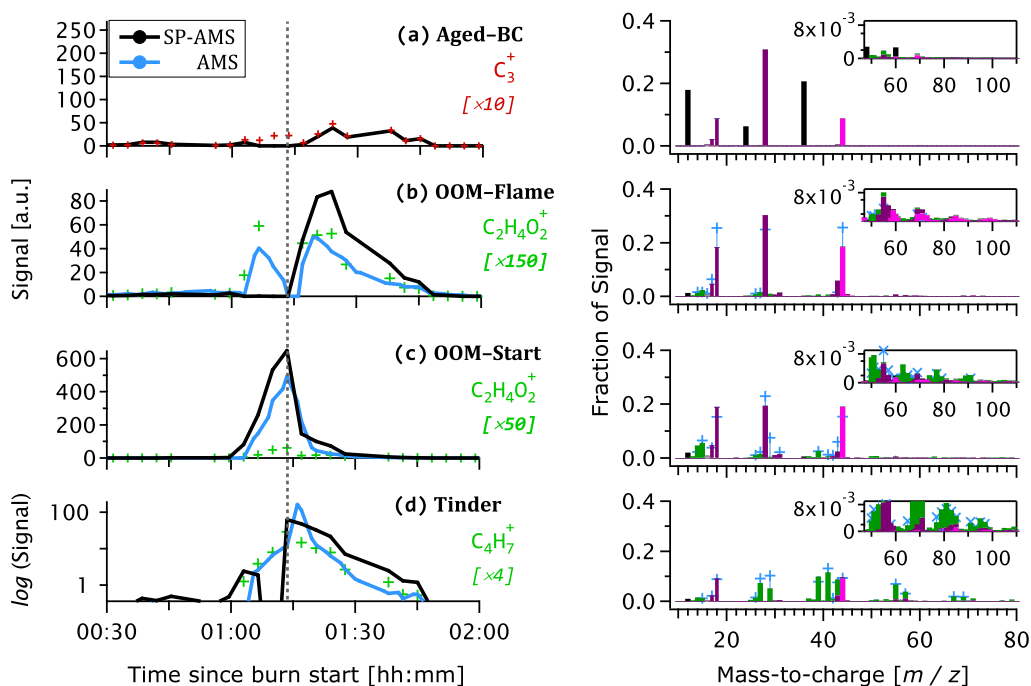
Note that the lack of rotational freedom of the model discussed in the main paper meant that exploring the dataset for

local minima via FPEAK runs (Ulbrich et al., 2009; Paatero, 2000; Paatero et al., 2002; Zhang et al., 2011) always converged to similar solutions, therefore no plots are included here. For the aged SP-AMS data, SEED runs gave slightly different results for smaller numbers of factors, but not for the presented five-factor solution. For the other three data sets, SEED runs using the presented number of factors gave similar results. As stated in the main paper, the large number of zeroes in this data set and the non-negativity constraint of PMF led to negligible ambiguity in the PMF solutions (Paatero and Tapper, 1994; Paatero et al., 2002). Rather, it was common that PMF did not converge for certain FPEAK or SEED values, seemingly at random.

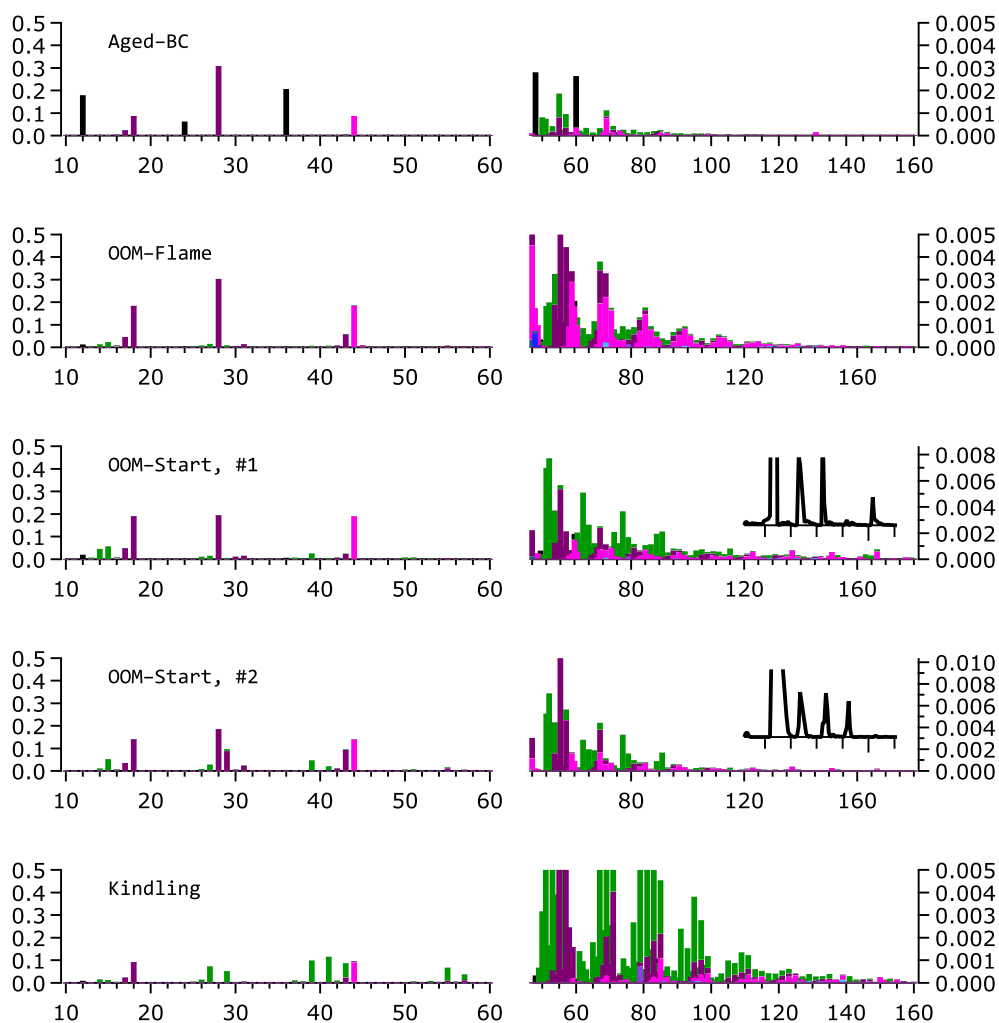
The final figure shows the mass spectrum of the two SP-AMS OOM–Start factors before their combination. The combined (summed) time series is presented in the main paper and the mass spectrum of OOM–Start1 is presented. As stated above, these factors are considered a “splitting” (Ulbrich et al., 2009) of the 3-factor solution, but a 5-factor solution was required to extract the Tinder factor.



**Figure S1.** The same fresh-emissions time series shown in the main paper, showing the POM-Start data which extended off the axis in that figure. The time axes have been correspondingly magnified.



**Figure S2.** The same aged-emissions time series shown in the main paper, showing the OOM-Start data which extended off the axis in that figure. The time axes have been correspondingly magnified. The vertical grey line illustrates the opening of the stove door to add tinder and manually ignite the fire.



**Figure S3.** Mass spectra for the 5-factor SP-AMS aged PMF solution. The two insets show the time series of the OOM-Start factors prior to combination, on the same axis as the time series in the main paper.

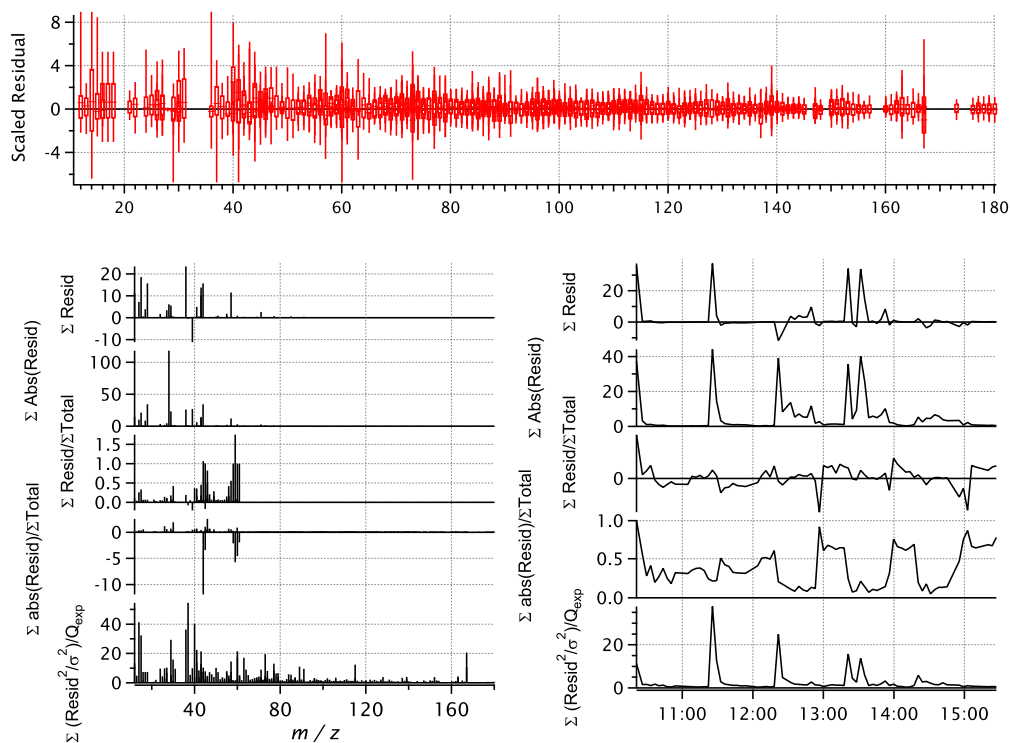


Figure S4. Residuals of the SP-AMS PMF model for fresh emissions.

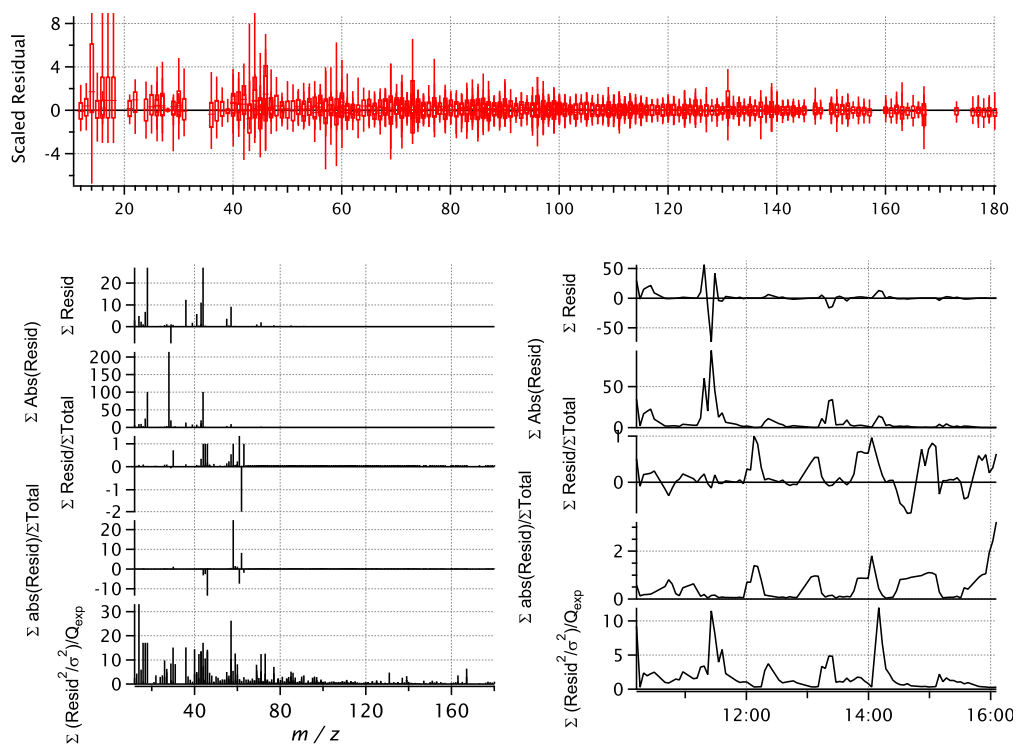


Figure S5. Residuals of the SP-AMS PMF model for aged emissions.

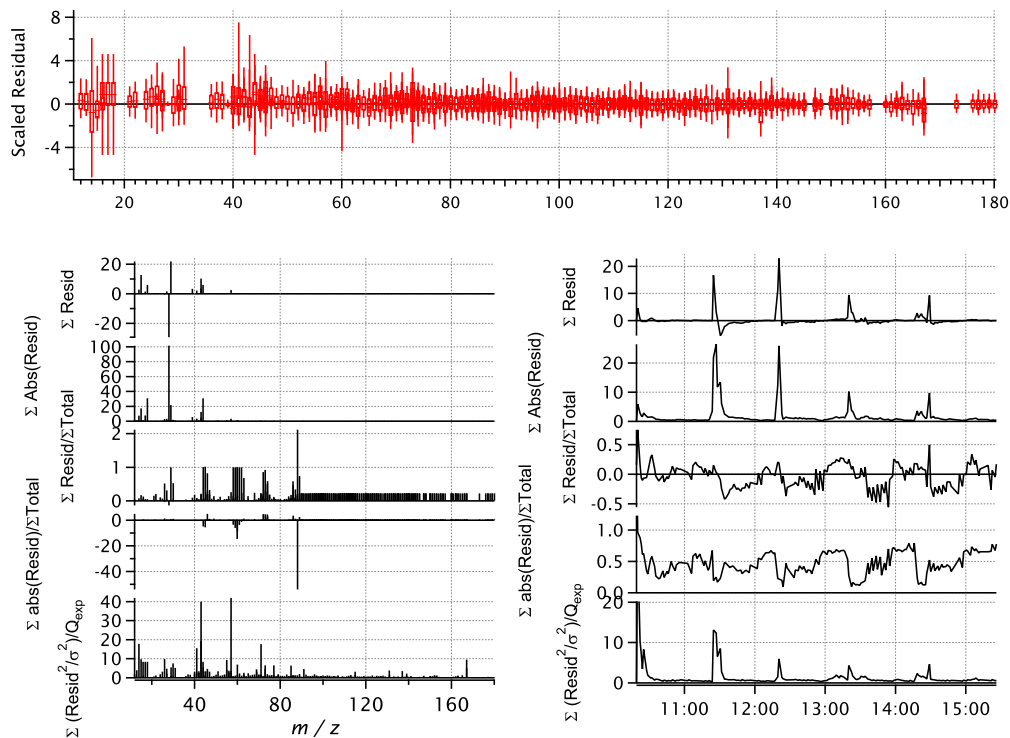


Figure S6. Residuals of the AMS PMF model for fresh emissions.

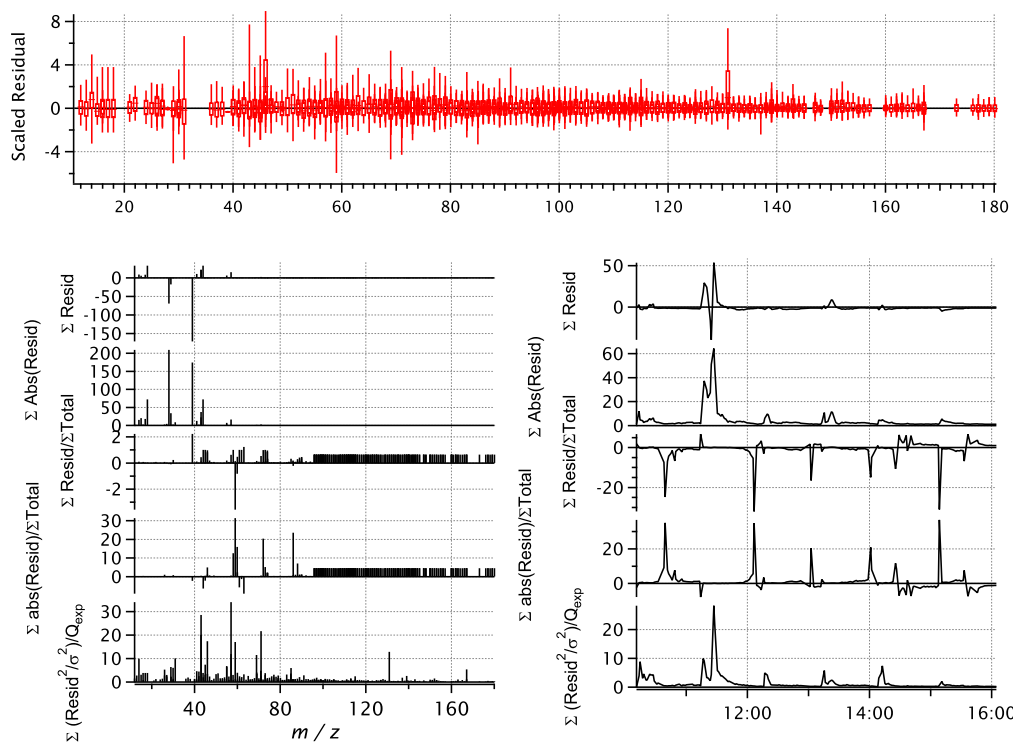
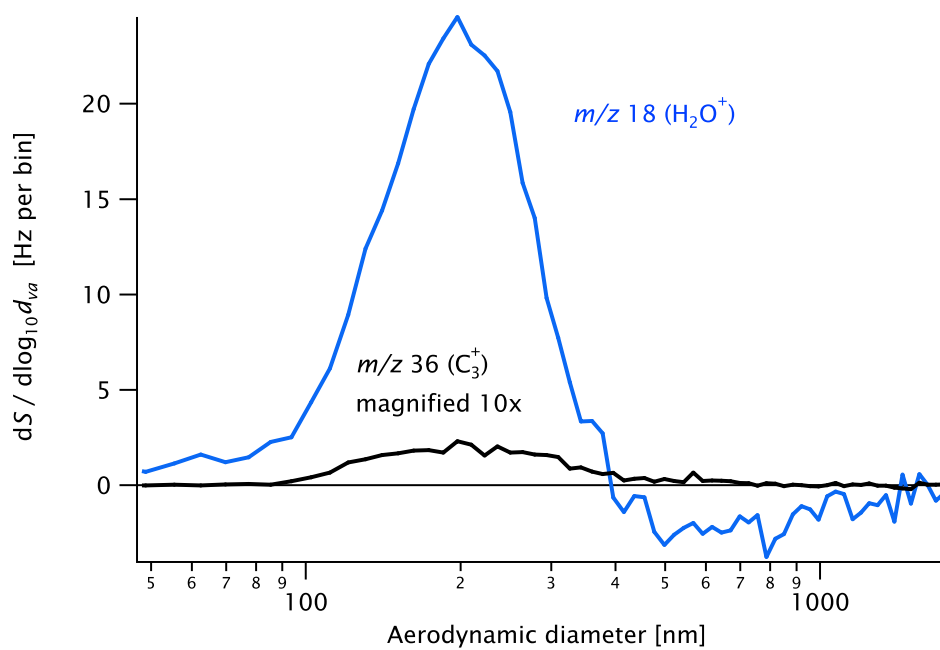


Figure S7. Residuals of the AMS PMF model for aged emissions.





**Figure S8.** Aerodynamic size distribution (calibrated range 125–724.5 nm) of signals at  $m/z$  18 and 36, identified as  $\text{H}_2\text{O}^+$  and  $\text{C}_3^+$  respectively via high-resolution analysis. Raw, unsmoothed data acquired over 20 seconds are shown.

## References

- Corbin, J. C., Keller, A., Sierau, B., Lohmann, U., and Mensah, A. A.: Wood-stove and pellet-burner organic aerosol emissions aged in a continuous-flow photoreactor, in press at *Aerosol Sci. Technol.*, 2015a.
- Corbin, J. C., Sierau, B., Gysel, M., Laborde, M., Keller, A., Kim, J., Petzold, A., Onasch, T. B., Lohmann, U., and Mensah, A. A.: Mass spectrometry of refractory black carbon particles from six sources: carbon-cluster and oxygenated ions, *Atmos. Chem. Phys.*, 14, 2591–2603, doi:10.5194/acp-14-2591-2014, 2015b.
- DeCarlo, P. F., Kimmel, J. R., Trimborn, A., Northway, M. J., Jayne, J. T., Aiken, A. C., Gonin, M., Fuhrer, K., Horvath, T., and Docherty, K. S.: Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer, *Anal. Chem.*, 78, 8281–8289, doi:10.1021/ac061249n, 2006.
- Paatero, P.: PMF User's Guide, Helsinki: University of Helsinki, 2000.
- Paatero, P. and Hopke, P. K.: Discarding or downweighting high-noise variables in factor analytic models, *Anal. Chim. Acta*, 490, 277–289, doi:10.1016/S0003-2670(02)01643-4, 2003.
- Paatero, P. and Tapper, U.: Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, *Environmetrics*, 5, 111–126, doi:10.1002/env.3170050203, 1994.
- Paatero, P., Hopke, P. K., Song, X.-H., and Ramadan, Z.: Understanding and controlling rotations in factor analytic models, *Chemom. Intell. Lab. Syst.*, 60, 253–264, doi:10.1016/S0169-7439(01)00200-3, 2002.
- Sueper, D., Jimenez, J. L., Aiken, A., and DeCarlo, P.: PIKA ToF-AMS High Resolution Analysis Software, [cires.colorado.edu/jimenez-group/ToFAMSResources/ToFSoftware/PikaInfo](http://cires.colorado.edu/jimenez-group/ToFAMSResources/ToFSoftware/PikaInfo), last accessed October 2014, 2011.
- Ulbrich, I., Canagaratna, M., Zhang, Q., Worsnop, D., and Jimenez, J.: 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.
- Wilks, D.: *Statistical Methods in the Atmospheric Sciences*, International Geophysics, Elsevier Science, Oxford, UK, 2011.
- Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Ulbrich, I. M., Ng, N. L., Worsnop, D. R., and Sun, Y.: 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.