Atmos. Chem. Phys. Discuss., 9, 6739–6785, 2009
www.atmos-chem-phys-discuss.net/9/6739/2009/

Atmospheric Chemistry and Physics Discussions

This discussion paper is/has been under review for the journal *Atmospheric Chemistry and Physics (ACP)*. Please refer to the corresponding final paper in *ACP* if available.

Simultaneous factor analysis of organic particle and gas mass spectra: AMS and PTR-MS measurements at an urban site

J. G. Slowik[1,3], A. Vlasenko[1,3], M. McGuire[2,3], G. J. Evans[2,3], and J. P. D. Abbatt[1,3]

[1]Department of Chemistry, University of Toronto, Toronto, Canada
[2]Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, Canada
[3]Southern Ontario Centre for Atmospheric Aerosol Research, University of Toronto, Toronto, Canada

Received: 19 February 2009 – Accepted: 2 March 2009 – Published: 11 March 2009

Correspondence to: J. P. D. Abbatt (jabbatt@chem.utoronto.ca)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

During the winter component of the SPORT (Seasonal Particle Observations in the Region of Toronto) field campaign, particulate non-refractory chemical composition and concentration of selected volatile organic compounds (VOCs) were measured by an Aerodyne time-of-flight aerosol mass spectrometer (AMS) and a proton transfer reaction-mass spectrometer (PTR-MS), respectively. Sampling was performed in downtown Toronto ~15 m from a major road. The mass spectra from the AMS and PTR-MS were combined into a unified dataset, which was analyzed using positive matrix factorization (PMF). The two instruments were given equal weight in the PMF analysis by application of a scaling factor to the uncertainties of each instrument. A residual based metric, $\Delta\overline{e_{sc}}$, was used to evaluate the relative weight. The PMF analysis yielded a 5-factor solution that included factors characteristic of regional transport, local traffic emissions, charbroiling, and oxidative processing. The unified dataset provides information on particle and VOC sources and atmospheric processing that cannot be obtained from the datasets of the individual instruments, such as apportionment of oxygenated VOCs to direct emission sources vs. secondary reaction products, improved correlation of oxygenated aerosol factors with photochemical age, and increased detail regarding the composition of oxygenated organic aerosol factors. This analysis represents the first application of PMF to a unified AMS/PTR-MS dataset.

1 Introduction

Air pollutants have important effects on ecosystems, human health, atmospheric visibility, and climate change. Organic pollutants exist in both the gas and particle phases and vary in terms of their composition and source. Both particulate organic species and volatile organic compounds (VOCs) may enter the atmosphere either as a result of primary emissions such as fossil fuel combustion or through secondary processes such as gas-phase or heterogeneous chemical reactions. A quantitative understand-

ACPD

9, 6739–6785, 2009

ing of VOC and particulate organic sources and atmospheric processing is necessary to reduce uncertainties in global climate models and for the development of pollution mitigation strategies to improve air quality (Kanakidou et al., 2005).

One approach to estimating the effects of source contributions and atmospheric processing to particle and VOC composition and concentration is through the use of receptor modeling techniques such as positive matrix factorization (PMF) (Paatero, 1997; Paatero and Tapper, 1994) and UNMIX (Lewis et al., 2003). Multivariate statistical techniques are used to deconvolve a time series of simultaneous measurements into a set of factors and their time-dependent concentrations. These factors may then be related to emissions sources, chemical composition, and/or atmospheric processing, depending on their specific characteristics. Because receptor models require no a priori knowledge of meteorological conditions or emissions inventories, they are ideal for use in locations where emissions inventories are poorly characterized or highly complicated (e.g. urban areas), or where atmospheric processing plays a major role.

Factor analysis techniques have previously been applied to a range of VOC measurements (Buzcu and Frazier, 2006; Holzinger et al., 2007; Lanz et al., 2008b), yielding factors related to atmospheric processing and sources such as traffic and biogenic emissions. Although PMF has previously been applied to particle measurements (Owega et al., 2004), a detailed treatment of the organic component has only recently been attempted. Lanz et al. (2007) applied PMF to organic aerosol mass spectra obtained from an aerosol mass spectrometer (AMS), obtaining six distinct factors relating to aerosol composition, volatility, and specific sources such as charbroiling and wood burning emissions. Zhang et al. (2005) developed a technique for deconvolving AMS mass spectra into oxygenated organic aerosol (OOA) and hydrocarbon-like organic aerosol (HOA) using *m/z* 44 (CO_2^+) and *m/z* 57 ($C_4H_9^+$, $C_3H_5O^+$) as OOA and HOA tracers. Other studies have typically included selected AMS mass spectral fragments in receptor modeling (typically restricted to inorganic species, *m/z* 44, and *m/z* 57) (Buset et al., 2006; Quinn et al., 2006), classified organics based on their thermal properties (Zhao and Hopke, 2006), or treated the organics as a single species for analysis.

ACPD
9, 6739–6785, 2009

Recent studies indicate that the traditional binary treatments of atmospheric organics as either gases or particles may be inadequate (Robinson et al., 2007). A proposed alternative is the treatment of organic species through the use of a volatility basis set (Donahue et al., 2006), in which the partitioning behavior of organics is considered over a range of volatilities. Such issues highlight the need for analytical approaches capable of simultaneous, cohesive analysis of gas and particle data. One such approach is presented here, through the application of the PMF receptor modeling technique to coupled gas and particle data.

In this experiment, simultaneous measurements of the mass spectra of particulate organics and VOCs were obtained using an Aerodyne aerosol mass spectrometer (AMS) and a proton transfer reaction mass spectrometer (PTR-MS). The measurements from these two instruments were combined into a single dataset and analyzed using PMF. This analysis yielded factors related to emission sources and chemical composition, specifically the degree of oxygenation. These factors were compared to the results obtained from PMF analysis conducted separately on the individual AMS and PTR-MS datasets. This is the first application of PMF analysis to a combined AMS/PTR-MS dataset.

2 Materials and methods

2.1 Sampling and instrumentation

During the winter component of the SPORT (Seasonal Particle Observations in the Region of Toronto) field campaign (22 January 2007 to 5 February 2007), a time-of-flight aerosol mass spectrometer (C-ToF-AMS) (Aerodyne Research, Inc., Billerica, MA, USA) and a proton transfer reaction-mass spectrometer (PTR-MS) (Ionicon Analytik, Innsbruck, Austria) were deployed in downtown Toronto (Wallberg Building, University of Toronto). The sampling inlet consisted of a 4 in diameter circular duct located ~5 m above ground and ~15 m from a major road. Ambient air was sampled contin-

uously at a rate of 300 L/min through a 10.2 cm diameter duct. The AMS sampling line was ~7 m long and constructed from stainless steel and conductive tubing. The PTR-MS utilized a Teflon sampling line with a length of ~2.5 m.

The literature provides detailed descriptions of the AMS (Canagaratna et al., 2007; Drewnick et al., 2005) and PTR-MS (Hansel et al., 1995; Lindinger et al., 1998). The AMS provides size-resolved, non-refractory composition of particles, while the PTR-MS provides the concentrations of VOCs with a proton affinity greater than that of water. The AMS and PTR-MS recorded data on 1 min and 30 s time intervals, respectively, which was re-averaged into 15 min time intervals because of signal-to-noise considerations. For analysis of the both the individual and unified datasets, time periods containing mass spectra from only one instrument were excluded, yielding a total of 1148 analyzed mass spectra.

AMS data analysis was performed using the ToF-AMS Analysis Toolkit v.1.44 (D. Sueper, University of Colorado-Boulder, Boulder, CO, USA) for the Igor Pro software package (Wavemetrics, Inc., Portland, OR, USA). The organic components of *m/z*'s≤300 were included in the PMF analysis. Mass fragments containing no organic signal were excluded, resulting in 270 analyzed *m*/z's. At *m/z*'s that contain signal from both inorganic and organic ions, the organic contribution was determined through a fragmentation pattern-based analysis routine (Allan et al., 2004). The procedure for calculating AMS uncertainties is described in detail in the literature (Allan et al., 2003) and summarized briefly as follows. The distribution of ion signals recorded for a given ensemble are represented as a Poisson distribution and convolved with a detector-dependent Gaussian distribution representing the variation in signal obtained for a single ion. During operation, the particle beam is alternately blocked (yielding a background measurement) and unblocked. Uncertainties are calculated independently for each mode and summed in quadrature, yielding the expression $\Delta I_d = \alpha \frac{\sqrt{I_o + I_b}}{\sqrt{t_s}}$. Here I_o and I_b are the ion signals in the unblocked and blocked (background) positions, t_s is the sampling time, and α is a factor accounting for the width of the Gaussian ion signal

distribution.

Due to signal-to-noise constraints imposed by the 30 s sampling intervals, the PTR-MS was not used to scan the entire mass spectrum and instead was set to measure specific masses. Ions at *m/z* 31 (formaldehyde), 43 (alkyl fragments, propylene, acetic acid, acetone, peroxyacetyl nitrate (PAN)), 45 (acetaldehyde), 59 (acetone, propanal, glyoxal), 61 (acetic acid), 73 (methyl ethyl ketone (MEK), methylglyoxal, butanal), 79 (benzene), 93 (toluene), 107 (xylenes, ethyl benzene, benzaldehyde), and 121 (trimethyl benzene, ethyl toluene, propyl benzene) were included in the PMF analysis (de Gouw and Warneke, 2007), while *m/z* 33 (methanol), 37 (water dimer), 42 (acetonitrile), 69 (acetone), and 129 (naphthalene) were measured but excluded from the PMF due to poor signal-to-noise (*m/z* 42, 69, 129), signal exclusively due to the reagent ion (*m/z* 37), or problems with the measurement dynamic range due to persistent local sources (*m/z* 33 (methanol), from windshield washer fluid). Uncertainties for the PTR-MS were calculated from background levels and Poisson ion counting statistics as described in the literature and summarized below (de Gouw et al., 2003). Typical uncertainty values were in the range of 2 to 18% of signal, depending on the *m/z*. Background levels were obtained by sampling through a charcoal denuder. The overall uncertainty is given by $\Delta(I-I_c)=\sqrt{\frac{I}{\tau}+\frac{I_c}{\tau_c}}$, where I is the ion signal, τ is the dwell time, and the "c" subscript denotes background measurements.

2.2 Positive Matrix Factorization (PMF)

The AMS and PTR-MS mass spectral time series and uncertainties obtained as described above were analyzed using the PMF2 software package (P. Paatero, U. of Helsinki, Finland), together with a modified version of the CU AMS PMF Tool (D. Sueper, I. Ulbrich, University of Colorado-Boulder, Boulder, CO, USA). Two methods of analysis were employed. In the first method, PMF was separately applied to the AMS and PTR-MS data. In the second method, the data from the two instruments was combined into a single dataset, and PMF was applied to this unified dataset.

The PMF model is described in detail in the literature (Paatero, 1997; Paatero and Tapper, 1994). Here we provide a brief summary and discuss the special considerations required to apply PMF to the unified dataset. PMF operates on the input data matrix **X** and the corresponding uncertainty matrix **S**. In the present study, **X** is the time series of mass spectra collected by the AMS and/or PTR-MS. The matrix **S** therefore contains the uncertainty in the measurement of the signal of each *m/z* at every point in time. The PMF model is described by the matrix equation:

$$X = GF + E \quad (1)$$

Here the columns of the **G** matrix contain the factor time series and the rows of the **F** matrix contain the factor mass spectra. The number of factors in a solution is determined by the user, through criteria discussed later. The **E** matrix contains the residuals and is defined by Eq. (1). The PMF model solves Eq. (1) by using a weighted least-squares algorithm to minimize the sum of squares, Q, defined as:

$$Q = \sum_i \sum_j \left(e_{ij}/s_{ij}\right)^2 \quad (2)$$

Here e_{ij} are the elements of the residual matrix **E**, and s_{ij} are the elements of the uncertainty matrix **S**. In an ideal solution, the PMF model reproduces the input data to the limit imposed by the measurement uncertainty; that is, $|e_{ij}|=s_{ij}$. The theoretical value of Q, denoted Q_{expected}, is therefore equal to the number of elements in the input matrix **X**. In practice, Q is expected to be somewhat larger than Q_{expected} for ambient data because the data cannot be perfectly represented by a finite number of factors.

In considering the solutions to the unified dataset, an important consideration is the relative weight given to each instrument. Because the AMS data contains 270 *m/z*'s, whose time trends are not totally independent, while the PTR-MS data contains only 10, the PMF model weights the AMS data more heavily than the PTR-MS. The relative weight of the two instruments may be determined by examining the scaled residuals, e_{ij}/s_{ij}, for each instrument. When the AMS and PTR-MS data are equally

ACPD
9, 6739–6785, 2009

weighted in the solution, the value of the scaled residuals is independent of the measuring instrument. To evaluate this requirement, we define the quantity $\Delta\overline{e_{sc}}$ as:

$$\Delta\overline{e_{sc}} = \overline{\left(\frac{|e_{ij}|}{s_{ij}}\right)_{\mathrm{AMS}}} - \overline{\left(\frac{|e_{ij}|}{s_{ij}}\right)_{\mathrm{PTR}}} \quad (3)$$

If $\Delta\overline{e_{sc}}$=0, the AMS and PTR-MS data is balanced in the PMF solution. Values of $\Delta\overline{e_{sc}}$<0 indicate that the PTR-MS is underweighted (because the residuals for the PTR-MS are larger than for the AMS), while $\Delta\overline{e_{sc}}$>0 indicates that the AMS is underweighted.

The constraint applied to the fit of each *m/z* is determined by the uncertainties. Here we have adopted the approach of weighting the PTR-MS by application of the factor C_{PTR} to the uncertainties. That is, if the uncertainties for unweighted PTR-MS data as a function of *m/z* and time (*t*) are given by $\Delta I(m/z, t)$ (where *I* is the signal intensity), for weighted data the PTR-MS uncertainties are given by $\Delta I(m/z, t)\times C_{\mathrm{PTR}}$. Note that the AMS uncertainties are always given by $\Delta I(m/z, t)$, regardless of C_{PTR}. As C_{PTR} is increased, the PTR-MS is given more weight in the PMF solution, and $\Delta\overline{e_{sc}}$ increases.

Note that C_{PTR} is not used in calculating the $\Delta\overline{e_{sc}}$ value in Eq. (3). Inclusion of C_{PTR} in this calculation could affect the value of $\Delta\overline{e_{sc}}$ without producing changes in the **F** and **G** matrices. For example, for very low or very high values of C_{PTR}, only one instrument is being significantly considered in the solution. In this scenario, a small change in C_{PTR} would not affect the solution because the relative weight of data from the considered instrument would remain the same. However, the change in C_{PTR} would still affect the $\Delta\overline{e_{sc}}$ value. To prevent such an artifact, the s_{ij} values from the unweighted solution are used in Eq. (3).

For the analysis of the unified AMS/PTR-MS dataset, the true mode of the PMF2 engine was utilized instead of the robust mode. The robust mode downweights data where the value of e_{ij}/s_{ij} exceeds predefined limits. Because the C_{PTR} weighting method artificially increases the residuals of the PTR-MS, the robust mode introduces bias by systematically preventing PTR-MS data from pulling with their full weight. As

ACPD
9, 6739–6785, 2009

discussed later, this prevents balanced solutions (i.e. $\Delta\overline{e_{sc}}=0$) from being obtained. Analysis of the individual AMS and PTR-MS datasets was conducted using both the true and robust modes.

Matrix rotations were explored by varying the PMF2 fPeak parameter around zero. For values of fPeak that yielded convergent solutions with reasonable Q-values, it was found that the effect of matrix rotation on the solutions was significantly less than the effect of variations in C_{PTR}. For simplicity, we present here only solutions obtained using fPeak=0.

3 Results and discussion

We first present results obtained from PMF analysis of the individual AMS and PTR-MS datasets. We then discuss the PMF analysis of the unified AMS/PTR-MS dataset in terms of (1) selection and evaluation of solutions, (2) physical interpretation of the extracted factors, and (3) comparison of the information yielded by the individual and unified analyses.

3.1 AMS dataset

The application of PMF to a dataset of AMS organic mass spectra has been previously described in detail (Lanz et al., 2007; Ulbrich et al., 2008), and a similar approach was used in the present study. A crucial consideration is the number of factors used in the PMF model. This number is somewhat subjective because the PMF model can be run with an arbitrary number of factors and no foolproof method for determining the "correct" number of factors exists. As will be described, a 6-factor solution was chosen based on the effects of the number of factors on the residual time series and Q, correlations between the factor time series and external tracers, and physical interpretation of the factor mass spectra.

To facilitate comparison with the unified dataset, the AMS dataset was analyzed

utilizing both the "true" and "robust" modes of the PMF2 engine. (As explained in the Methods section, the unified dataset could only be analyzed meaningfully in the "true" mode.) For the 6-factor solution to the AMS dataset, the following factors were obtained for both modes: (1 and 2) oxygenated organic aerosol, split across 2 factors ("OOA-Ia" and "OOA-Ib"); (3) charbroiling; (4) hydrocarbon-like organic aerosol (HOA); (5) wood burning; and (6) an oxygenated factor of unknown origin associated exclusively with winds from the North/Northeast. The time series of the OOA-Ia and OOA-Ib factors are highly correlated, suggesting that this factor splitting is an artifact of the PMF solution (Ulbrich et al., 2008) and not two unique factors. However, increasing the number of factors beyond the point where the splitting is observed continues to yield meaningful factors. Therefore, the OOA-Ia and OOA-Ib factors are here combined into a single "OOA-I" factor. The "true" and "robust" modes yielded nearly identical results, with the only significant differences being the nature of the OOA-Ia/OOA-Ib split (with the recombined OOA-I again nearly identical) and the order in which new factors were resolved as the number of factors in the solution was increased. The results presented below were obtained using the "true" mode.

Figure 1 shows a plot of the scaled residual time series (that is, $\sum_j \left(e_{ij}/s_{ij}\right)^2$ where j denotes the *m/z*'s) for the 6-factor solution. The inset to Fig. 1 shows Q/Q_{expected} as a function of the number of factors, where Q_{expected} is estimated as the number of elements in the input matrix. Structure is evident in the residual time series, primarily in the form of short-duration, high-intensity periods. Although the AMS data was averaged to 15 min time intervals for the PMF analysis, examination of the organic time series at 1 min sampling intervals shows that these features are caused by intense concentration spikes of <1 min duration. These spikes are due to emissions from nearby point sources, particularly a roadside hot dog stand and passing vehicles. Their prevalence in the residuals suggests fluctuations in the emission signatures of these sources that are not fully represented by a single factor. During these periods, the scaled residuals are dominated by hydrocarbon peaks.

ACPD
9, 6739–6785, 2009

Figure 2 shows the effect of the number of selected factors on the residual time series. Here we plot the change in the scaled residual time series, $\Delta\left(\sum_j \left(e_{ij}/s_{ij}\right)^2\right)$, between two solutions with different numbers of factors. That is, in Fig. 2a, we plot the residuals for the 4-factor solution minus those from the 5-factor solution. Thus the structure evident in Fig. 2a has been removed from the residuals by increasing the number of factors from 4 to 5. The key feature of this figure is that the structure present in Fig. 2a and b includes both the spikes described above and features covering longer time periods. However, when the number of factors is increased above 6 (Fig. 2c and d), the only significant changes are in the magnitude of the spikes. These spikes are also present in extracted factors (especially the charbroiling factor), suggesting that the 7 and 8-factor solutions are representing variations in source profiles as a linear combination of factors. At present, such splitting cannot be interpreted meaningfully, and the number of factors is capped at six.

The mass spectra and time series of the 6-factor AMS solution are shown in Fig. 3a and b, respectively. Figure 3b also contains the time series for correlated tracer species. The mass spectra are normalized such that the sum of each spectrum across all *m/z*'s is equal to 1. The time series are reported in terms of mass concentration (μg/m^3). We again note although 6 factors were obtained in this solution, only 5 are displayed because factor $F1_{AMS}$ ("OOA-I") is a combination of two factors (OOA-Ia and OOA-Ib) believed to be a single factor split by the PMF algorithm. All AMS reference spectra described below and in Sect. 3.3.2 were obtained from the AMS Spectral Database (Ulbrich et al., 2007).

Factor $F1_{AMS}$ is similar to oxygenated organic aerosol (OOA) factors obtained from AMS data in previous field studies (Zhang et al., 2007) using either the two-component deconvolution technique (Zhang et al., 2005) or PMF analysis (R^2=0.96 vs. Zurich winter OOA, Lanz et al., 2008a, and R^2=0.90 vs. both Zurich summer OOA-I, Lanz et al., 2007, and Pittsburgh OOA, Zhang et al., 2005). It is characterized by strong signal at *m/z* 18 (H_2O^+) and 44 (CO_2^+), and low signal at *m/z*'s resulting from hydrocarbon

ACPD
9, 6739–6785, 2009

fragmentation. The time series of this factor correlates with particulate sulfate (Fig. 3b, R^2=0.71) and back trajectories passing over industrial regions to the west/southwest of Toronto (trajectories were calculated using the NOAA HYSPLIT model). $F1_{AMS}$ accounts for 41.9% of the total organic mass.

Factor $F2_{AMS}$ is similar to hydrocarbon-like organic aerosol (HOA) factors previously observed in the field (Zhang et al., 2007) R^2=0.97 vs. Zurich winter HOA, Lanz et al. (2008a), R^2=0.95 vs. Pittsburgh HOA, Zhang et al. (2005), and R^2=0.91 vs. Zurich summer HOA-I, Lanz et al. (2007). Unlike $F1_{AMS}$, $F2_{AMS}$ has negligible signal at *m/z* 44. The organic functionality can be investigated by analyzing the delta (Δ) patterns (McLafferty, 1980), where $\Delta = m/z - 14n + 1$ (n is an integer). $F2_{AMS}$ is dominated by the Δ=0 (i.e., *m/z* 27, 41, 55, 69, ...) and Δ=2 (*m/z* 29, 43, 57, 71, ...) series, which are characteristic of alkenes and alkanes, respectively. The time series of this factor correlates with NO_x (Fig. 3b) and is elevated overnight and during the morning rush hour. Factor $F2_{AMS}$ accounts for 16.6% of the total organic mass.

Factor $F3_{AMS}$ is attributed to charbroiling emissions. The mass spectrum of $F3_{AMS}$ shows strong signals at many of the same *m/z*'s as $F2_{AMS}$. The mass spectrum is correlated with reference spectra for charbroiling emissions (R^2=0.90) and HOA (R^2=0.90 vs. Zurich winter HOA, Lanz et al., 2008a, and Pittsburgh HOA, Zhang et al., 2005). A difference between $F2_{AMS}$ and $F3_{AMS}$ is in the relative intensities of the Δ=0 and Δ=2 series. Whereas in $F2_{AMS}$ (HOA) the Δ=0 and Δ=2 series each comprised 26% of the factor mass spectrum, in $F3_{AMS}$ they represent 37% and 16% of the spectrum, respectively. Additionally, the diurnal profiles of the two factors are distinct, with $F3_{AMS}$ exhibiting strong signal around noon. Most of the $F3_{AMS}$ signal is concentrated in intense spikes of <1 min duration, which occur exclusively during the operation of a roadside hot dog stand ~25 m from the sampling inlet. Day-to-day variation in $F3_{AMS}$ is determined by the number of detected particles, not particle size, suggesting that the variation is driven by street-level mixing dynamics. $F3_{AMS}$ comprises 26.1% of the total organic mass.

Factors $F4_{AMS}$ and $F5_{AMS}$ are more difficult to validate due to their lower concen-

ACPD
9, 6739–6785, 2009

trations and the absence of satisfactory tracer species. Factor $F4_{AMS}$ is tentatively attributed to biomass burning emissions. Some features of the $F4_{AMS}$ time series are correlated with the AMS potassium measurement (see Fig. 3b). However, AMS potassium measurements are not quantitative because of multiple ionization processes and high instrument background. Further, during the study period, potassium signals could also be obtained from road salt. The $F4_{AMS}$ mass spectrum correlates only moderately well with previously extracted wood-burning factors (Lanz et al., 2007, 2008a) (R^2~0.5). However, burning signatures vary significantly with fuel type and burn conditions (Weimer et al., 2008). $F4_{AMS}$ is the only factor with a significant contribution from *m/z* 60, which is frequently used as a tracer for levoglucosan and an indicator of biomass burning (1.8% of the factor spectrum vs. 0.7% for $F1_{AMS}$, for which *m/z* 60 has the next largest contribution). For the Lanz et al. (2007, 2008a) wood-burning factors, *m/z* 60 is between 1.4% (winter) and 3.2% (summer) of the spectrum. Factor $F4_{AMS}$ accounts for 7.6% of the organic mass.

A unique feature of the $F5_{AMS}$ mass spectrum is the prominent signal at *m/z* 56 (16% of total). The presence of *m/z* 44 indicates oxygenation, suggesting that *m/z* 56 may be influenced by $C_3H_4O^+$ fragment, obtained from alkylcycloalkanones. However, contributions from $C_4H_8^+$ (cycloalkanes and branched alkenes), or $C_3H_6N^+$ (cyclic amines) cannot be ruled out. The $F5_{AMS}$ time series does not correlate with any available tracer species. However, it does depend strongly on wind direction and is observed only during north/northeast winds. This specificity suggests the factor is due to a specific point source of primary emissions. The factor comprises 7.7% of the organic mass.

3.2 PTR-MS dataset

A 5-factor solution was selected for the PTR-MS dataset, using similar criteria as described above for the AMS dataset. The dataset was analyzed using both the "true" and "robust" modes, and results were found to be very similar, with the only significant difference being the the order in which new factors were resolved as the number of factors in the solution was increased. The results presented below were obtained using

the "true" mode.

Figure 4 is a plot of the scaled residual time series, $\sum_j \left(e_{ij}/s_{ij}\right)^2$, for the 5-factor solution. The inset to Fig. 4 shows Q/Q_{expected} as a function of the number of factors. As was the case for the AMS dataset, significant structure remains in the residuals. In Fig. 5, the effect of the number of resolved factors on the residual time series is shown by plotting the quantity $\Delta\left(\sum_j \left(e_{ij}/s_{ij}\right)^2\right)$ as was previously done for the AMS dataset in Fig. 2. The structure in the time series evident in Fig. 5a, b, and c indicates improvement in the solution by increasing the number of resolved factors. A further increase to 6 factors has little additional effect (Fig. 5d), with the improvements existing mostly as small modifications to periods where much larger gains were made at lower numbers of factors (compare for example the structure on 2 January 2007 in Fig. 5a and d). Similar time series are obtained as the number of factors increases above 6, suggesting that these higher-order solutions represent variations in source profiles, rather than providing factors than can be meaningfully interpreted.

The mass spectra and time series for the 5-factor solution to the PTR-MS dataset are presented in Fig. 6a and b, respectively. The mass spectra are normalized such that the sum of each spectrum is equal to 1, while the time series are reported in units of ppb. An alternate picture of the factor mass spectra is provided in Fig. 7. Here we report the fraction of the signal at a given *m/z* apportioned to each factor. The differences between the value of each *m/z* stack and 1 is the signal at that *m/z* remaining in the residuals.

Factor F1_{PTR} is identified with traffic emissions. As shown in Fig. 7, factor F1_{PTR} contains 63–71% of the signal from aromatic compounds at or above *m/z* 93 and 42% of the benzene signal (*m/z* 79). The mass spectrum (Fig. 6a) shows a toluene/benzene ratio of 2.52, suggesting relatively fresh emissions. The toluene/benzene ratio can be used as a photochemical clock, because these two aromatics are typically emitted by the similar sources, but toluene has a shorter lifetime (Roberts et al., 1984). In the

ACPD
9, 6739–6785, 2009

present study, source emissions were estimated to have a toluene/benzene ratio of ~4.0, and the ratio approaches zero with increasing photochemical age. The source emission ratio is consistent with previous measurements of fresh traffic emissions (de Gouw et al., 2005; Kristensson et al., 2004). The $F1_{PTR}$ factor peaks during the morning and evening rush hours (3–5 times nighttime values) and is slightly elevated during the rest of the day. As shown in Fig. 6b, the factor correlates strongly with NO_x (R^2=0.64).

The mass spectrum of $F2_{PTR}$ is dominated by signal at *m/z* 61 (acetic acid), and includes 59% of the acetic acid signal. The factor time series (Fig. 6b) correlates well with AMS OOA ($F1_{AMS}$) for most of the study, suggesting a contribution from transported, well-processed air. However, this correlation breaks down during the period of 29–31 January, where strong spikes in the $F2_{PTR}$ time series are not reflected in the OOA data. Such short-lived and intense features in the $F2_{PTR}$ time series likely indicate a local source. Further, the factor mass spectrum (Fig. 6a) shows a toluene/benzene ratio of 2.78, which is consistent with fresh emissions, although the aromatics constitute only a small fraction of the factor mass spectrum and the factor contains only ~10% of the aromatic signal. Acetic acid is a product of ambient photochemical reactions, but has also been observed in emissions from spark-ignition engines (Zervas et al., 2001). The information provided above suggests that $F2_{PTR}$ is influenced by both local emissions and transported air, and that these contributions cannot be decoupled through PMF using only the PTR-MS dataset. As will be discussed later, the effects of these sources can be largely decoupled in the unified AMS/PTR-MS dataset.

Similar issues arise for factor $F3_{PTR}$. As was the case for $F2_{PTR}$, the $F3_{PTR}$ time series correlates well with OOA for much of the study, but does not correlate during spikes in the $F3_{PTR}$ intensity. The toluene/benzene ratio is 1.86, which although lower than either $F1_{PTR}$ or $F2_{PTR}$, is still high enough to suggest some contributions from local emissions sources. The factor mass spectrum is dominated by acetone and includes 56% of the total acetone signal. $F3_{PTR}$ also contains strong signals at *m/z* 45 (acetaldehyde) and *m/z* 73 (methyl ethyl ketone, methylglyoxal, butanal). Figure 7 shows

ACPD
9, 6739–6785, 2009

that while acetaldehyde is primarily found in other factors, most of the non-residual *m/z* 73 is assigned to $F3_{PTR}$. Several of the time series spikes during the 31 January–2 February period are attributed to known local painting activity. Both acetone and methyl ethyl ketone have primary emissions sources (including paint solvents), but are also generated as photochemical reaction products. Similar to $F2_{PTR}$, $F3_{PTR}$ is affected by both local emissions and transported air, and the two effects can be decoupled only in the unified AMS/PTR-MS dataset.

$F4_{PTR}$ contains 63% of the acetaldehyde signal (*m/z* 45), a VOC oxidation product with a lifetime of less than a day. The signals at *m/z* 31, 59, and 61 are attributed to formaldehyde, acetone (with potential minor contributions from propanal and glyoxal), and acetic acid, respectively, which are all produced from VOC oxidation. In contrast to the factors described above, $F4_{PTR}$ has a low toluene/benzene ratio (0.42), indicating that this factor is more photochemically aged than, for example, $F1_{PTR}$. The relatively short lifetimes of formaldehyde and acetaldehyde suggest that $F4_{PTR}$ is the result of local oxidation. As discussed in Sect. 3.3.3 in connection with the unified dataset, the amount of formaldehyde apportioned to secondary oxidation is consistent with previous urban studies during winter. A difficulty in interpreting factor $F4_{PTR}$ is that no correlated tracer species were found. However, this is not surprising given that the available tracer species were expected to correlate with either direct emissions or transported, aged air. $F4_{PTR}$ is anticorrelated with temperature (R^2=0.50), as shown in Fig. 6b.

Like $F4_{PTR}$, factor $F5_{PTR}$ consists primarily of long- and short-lived oxygenated compounds, although acetaldehyde is proportionally lower. Given the oxygenated nature of the other constituents in $F5_{PTR}$, the signal at *m/z* 43 is probably from the CH_3CO^+ ion, which results from a variety of oxygenated compounds, including peroxyacetyl nitrate (PAN), acetone, and acetic acid (de Gouw and Warneke, 2007). As was the case for $F4_{PTR}$, $F5_{PTR}$ does not correlate with any available tracers. However, it exhibits no temperature dependence. The absence of distinct events in the time series and the low toluene/benzene ratio of 0.65 suggest that the factor is not being dominated by a local point source.

ACPD
9, 6739–6785, 2009

3.3 Unified AMS/PTR-MS dataset

3.3.1 Selection and evaluation of solution

As stated in the Methods section, the AMS and PTR-MS datasets were weighted equally in solutions to the unified dataset by applying the weighting factor C_{PTR} to the instrument uncertainties. As a result, the solution space to be explored for the unified dataset has two dimensions: (1) the number of output factors and (2) the value of C_{PTR}. The Methods section (Eq. 3) also introduced the metric $\Delta\overline{e_{sc}}$, which is used to evaluate the relative weight of the AMS and PTR-MS datasets. We note again that the datasets are equally weighted when $\Delta\overline{e_{sc}}$=0.

Figure 8 shows the value of $\Delta\overline{e_{sc}}$ as a function of the number of output factors (x-axis) and C_{PTR} (colored traces). The following values of C_{PTR} were tested: 1, 2, 5, 10, 20, 50, 100, 200, 500, and 1000. For each value, the 1- to 10-factor solutions were analyzed. Points within this matrix that are absent from Fig. 8 correspond to non-convergent solutions.

Several features are evident in Fig. 8. First, we note that for most of the C_{PTR} traces, $\Delta\overline{e_{sc}}$ becomes closer to zero as the number of resolved factors increases. This is the expected behavior, because the degrees of freedom in the solution is increasing. Second, it is evident that the set of solutions for C_{PTR}=1 is never balanced, with $\Delta\overline{e_{sc}}$ values ranging from −6.2 (2 factors) to −1.5 (10 factors). The figure shows that C_{PTR} values between 10 and 20 lead to balanced solutions. The solution at C_{PTR}=10 is presented below. Solutions for C_{PTR}=5 and C_{PTR}=20 are presented in the Supplemental Information, see: http://www.atmos-chem-phys-discuss.net/9/6739/2009/acpd-9-6739-2009-supplement.pdf.

As was done for the individual AMS and PTR-MS datasets, the number of factors in the solution to the unified dataset is selected through analysis of the residual time series, factor mass spectra, and factor profiles. As will be discussed, the 5-factor solution was selected as optimal and is presented below. In Fig. 9, we show the scaled residual

ACPD
9, 6739–6785, 2009

time series, $\sum_j \left(e_{ij}/s_{ij}\right)^2$, for the 5-factor solution. In the top panel, the residuals are plotted for the entire unified solution, while in the lower two panels they are segregated by instrument. The ratio Q/Q_{expected} as a function of the number of factors is shown in the inset. Note that the difference in scale between the PTR-MS residuals in this figure vs. those in Fig. 5 (i.e. the PTR-MS dataset) is due to the inclusion of C_{PTR} in the residual calculation for the unified dataset. The PTR-MS scaled residuals in the unified dataset are larger than those in the individual PTR-MS dataset by a factor of 600; of this a factor of 100 is due to the inclusion of C_{PTR}=10 in the calculation, while the remaining factor of 6 is due to changes in e_{ij} between the two solutions. The AMS residuals are larger in the unified dataset by a factor of 17 relative to the individual dataset.

The PTR-MS scaled residual time series in Fig. 9 is qualitatively similar to that obtained from the individual PTR-MS dataset (Fig. 4). For both the individual and unified datasets, the AMS residuals include strong concentration spikes, which were identified with charbroiling emissions in the individual AMS dataset. A difference is the inclusion in the AMS unified residuals of longer-duration events on 28 January, which correspond to the unidentified northerly point source, F5_{AMS}.

In Fig. 10, we plot the change in the scaled residual time series, $\Delta\left(\sum_j \left(e_{ij}/s_{ij}\right)^2\right)$, as a function of the number of resolved factors. Here the AMS and PTR-MS residuals are plotted separately in the left and right-hand columns, respectively. Structure is evident in the AMS time series for increases up to 5 factors. Selecting a larger number of factors mostly causes changes in the magnitude of the concentration spikes. For the PTR-MS, there is no clear-cut point at which the changes in the residual structure decrease. However, for solutions containing at least 5 factors, $\Delta\left(\sum_j \left(e_{ij}/s_{ij}\right)^2\right)$ is consistently dominated by structure in the period of 29 to 31 January. Further, in-

ACPD
9, 6739–6785, 2009

creasing the number of factors beyond 5 leads to AMS factors with very similar mass spectra. As discussed previously, such splitting behavior has been associated with excessive numbers of factors (Ulbrich et al., 2008). The 6-factor solution also contains a split PTR-MS factor. The number of factors is therefore selected at 5, although the possibility of meaningful factors in higher-order solutions cannot be completely ruled out.

3.3.2 Physical interpretation of factors

The factors obtained from the 5-factor solution to the unified dataset at C_{PTR}=10 are discussed below in terms of the results shown in Figs. 11 and 12. Figure 11 contains the factor profiles and time series, plotted as follows. In Fig. 11a, the AMS spectra are shown as sticks (left and bottom axes) and the PTR-MS spectra as bars (top and right axes). Signal intensity is normalized such that each factor mass spectrum of the individual instrument sums to 1. In Fig. 11b, the time series are plotted (black traces, left axes) in terms of mass concentration (μg/m^3) for the AMS and mixing ratio (ppb) for the PTR-MS. Selected tracer species (NO_x, particulate NO_3^-) are also shown (gray traces, right axis). In Fig. 12, we plot the fraction of signal apportioned to the AMS and to each PTR-MS *m/z*. The factors are discussed individually below. As previously stated, all AMS reference spectra described below were obtained from the AMS Spectral Database (Ulbrich et al., 2007).

Factor $F1_{UN}$ is attributed mostly to charbroiling emissions, although it likely also contains traffic emissions, as discussed in the following section. The factor time series and AMS mass spectrum of factor $F1_{UN}$ are both very similar to that of the charbroiling factor $F3_{AMS}$ (time series R^2=0.97; mass spectrum R^2=0.92). The time trend of $F1_{UN}$ correlates with that of PTR-MS *m/z* 69 (R^2=0.53), which was excluded from the PMF analysis due to low signal-to-noise. Several compounds contribute to *m/z* 69, including furan, which is produced during meat cooking (Lee, 1999) and other combustion-related processes (Andreae and Merlet, 2001; Beychok, 1987). Previous discussions

ACPD
9, 6739–6785, 2009

of the $F3_{AMS}$ factor hold for $F1_{UN}$, notably that the particle mass spectrum is characteristic of aliphatic hydrocarbons and that the time series is dominated by short-duration concentration spikes clustered in the early afternoon. $F1_{UN}$ accounts for 47.2% of the AMS mass, compared to only 26.1% for $F3_{AMS}$. As shown in Fig. 12, $F1_{UN}$ contains 18% to 29% of the PTR-MS aromatics, depending on the *m/z*. As discussed for $F3_{AMS}$, day-to-day variation in $F1_{UN}$ is likely governed by local mixing dynamics.

Factor $F2_{UN}$ is associated with traffic emissions. The factor time series is correlated with NO_x (Fig. 11b), which was also the case for the HOA factor $F2_{AMS}$ and traffic factor $F1_{PTR}$. The AMS mass spectrum is also similar to that of $F2_{AMS}$ (R^2=0.83), indicating a hydrocarbon-dominated particle composition. The PTR-MS mass spectrum is somewhat different between $F2_{UN}$ and $F1_{PTR}$. The mass spectrum has a larger fraction of acetone and acetaldehyde relative to the aromatics and other oxygenated species. The toluene/benzene ratio is 3.42, suggesting fresh emissions. Further, the spectrum contains 13% of the benzene, and 20% to 25% of the *m/z*'s corresponding to higher aromatics (Fig. 12).

The AMS mass spectrum of factor $F3_{UN}$ closely resembles that of $F1_{AMS}$ (OOA) (R^2=0.995) and previously reported OOA-I spectra (Lanz et al., 2007; Ulbrich et al., 2008). Similar to $F1_{AMS}$, the oxygenated species at *m/z* 18 and 44 are prominent, and the time trend correlates with AMS nitrate (Fig. 11b) and sulfate (not shown, but correlated with $F1_{AMS}$ in Fig. 3b). The mass spectrum and correlations suggest that the factor is generated by secondary oxidation reactions. The PTR-MS mass spectrum is unique to the unified dataset and dominated by signals attributable to oxygenated species, with acetone (*m/z* 59) as the strongest signal. These species are consistent with secondary oxidation, though they also have direct emissions sources. The factor is attributed to regional transport of secondary organic aerosol. The apportionment of oxygenated VOCs such as acetone to this type of factor is an important feature of the unified dataset solution and is discussed further in the following section.

Factor $F4_{UN}$ is consistent with local oxidation. The PTR-MS mass spectrum is similar to that of $F4_{PTR}$, and is dominated by acetaldehyde. As shown in Fig. 12, the other

ACPD
9, 6739–6785, 2009

contributing VOCs are oxygenated (most notably formaldehyde) and signal from the aromatics is negligible. Compared to $F3_{UN}$, $F4_{UN}$ has a stronger contribution from oxygenated VOCs with shorter lifetimes. The dominant species in $F4_{UN}$, acetaldehyde and formaldehyde, have lifetimes of less than a day, while acetic acid and acetone, which are more prominent in $F3_{UN}$, have lifetimes on the order of a few weeks. The AMS mass spectrum somewhat resembles that of the OOA-like factors ($F3_{UN}$ and $F1_{AMS}$), but has proportionally more signal at *m/z* 43, 55 and in the C_4-C_6 range, indicating less oxygenation. The spectrum is similar to the OOA-II factors observed in Zurich (Lanz et al., 2007) and Pittsburgh (Ulbrich et al., 2008), which have been attributed to oxygenated organics that are more volatile and/or result from more recent oxidation than the OOA-I in $F3_{UN}$ and $F1_{AMS}$. The correlations in the present study with shorter-lived oxygenated VOCs and anticorrelation with temperature suggest that both the oxidation timescale and volatility contribute here to the factor time series. The particle spectrum of $F4_{UN}$ is unique to the unified dataset.

As shown in Fig. 12, factor $F5_{UN}$ contains almost 50% of the aromatic signal, indicating fresh, primary emissions. This is supported by the high toluene/benzene ratio (2.19). The oxygenated VOCs included in this factor have both primary and secondary sources. The AMS mass spectrum (Fig. 11a) is unique in containing a significant contribution from *m/z* 56, which can result from the fragmentation of cyclic ketones. However, this factor includes only ~5% of the AMS signal, which is sufficiently small to render the AMS component unreliable (Ulbrich et al., 2008). The factor time series does not correlate with any available tracer species. However, the dominant features in the time series are correlated with features in $F2_{PTR}$ and $F5_{PTR}$, which were believed to result from a mixture of local sources and transported reaction products. Because of the large fraction of aromatics assigned to $F5_{UN}$ (Fig. 12), it is likely that this factor contains the primary contribution to $F2_{PTR}$ and $F5_{PTR}$, with the secondary contributions assigned to $F3_{UN}$ (long-lived oxidation products/OOA-I) and $F4_{UN}$ (short-lived oxidation products/OOA-II), respectively.

3.3.3 Comparison of unified and individual datasets

For the first time, PMF has been applied to a unified AMS/PTR-MS dataset. This analysis provides information on particle and VOC sources and atmospheric processing that cannot be obtained from the datasets of the individual instruments.

The unified dataset enables apportionment of the oxygenated VOCs as primary emissions or secondary reaction products, which was not possible with the individual PTR-MS dataset. Metrics that may be used to identify factors resulting from primary emissions are: (1) aromatic VOCs, which are exclusively primary emissions, (2) hydrocarbon-like particulate organics, (3) correlation between the factor time series and a primary tracer, and (4) diurnal patterns, where the primary source can be identified and has a distinct diurnal emission pattern. In the unified solution, factors $F1_{UN}$ and $F2_{UN}$ satisfy all four metrics, while $F5_{UN}$ satisfies (1), (2) and (3) cannot be evaluated due to low signal in the AMS and lack of any correlated tracer, and (4) cannot be assessed due to the uncertain source identity. Further, the oxygenated character of the AMS and PTR-MS factor profiles for $F3_{UN}$ and $F4_{UN}$ suggest these factors are secondary reaction products, as does the time series correlation of $F3_{UN}$ with nitrate and sulfate.

Use of these classifications provides on approach to estimating the fraction of each VOC resulting from primary emissions ($F1_{UN}$, $F2_{UN}$, $F5_{UN}$) vs. secondary reactions ($F3_{UN}$, $F4_{UN}$). For example, Fig. 12 indicates that 55% of formaldehyde (*m/z* 31) is directly emitted, compared to only 35% of acetaldehyde (*m/z* 45). The primary contribution of formaldehyde in this study (55%) is consistent with values reported for other urban locations. Best agreement is found with a winter study in Rome (Possanzini et al., 2002) (65–70%) and an April 2003 study in Mexico City (Garcia et al., 2006) (42 to 63%). Photochemical production makes a larger contribution during the summer, as evidenced by lower primary contributions during campaigns in Rome (Possanzini et al., 2002) (10–20%), Houston (Friedfeld et al., 2002) (22%), and a rural site near Vancouver (Li et al., 1997) (33%). Approximately 60% of both acetone and acetic acid

ACPD
9, 6739–6785, 2009

are attributed to primary emissions, which is slightly higher than the 52% reported for acetone at a site outside Vancouver during summer (Li et al., 1997). All of the *m/z*'s corresponding to non-benzene aromatics have >95% attributed to primary sources. Benzene is more complex, with only ~70% attributed to direct emission factors. This may be because of interferences at *m/z* 79, or because the benzene lifetime is sufficiently long to persist into the time scale of the secondary factors. This consideration also holds for the longer-lived oxygenated species (i.e. acetone and acetic acid), meaning that the values presented above should be considered as lower limits for the direct emission contribution.

The primary/secondary VOC analysis presented above is only possible through the unified dataset. In the individual PTR-MS dataset, several factors were identified as having contributions from both primary and secondary sources. Inclusion of the AMS data, where the classification of hydrocarbon vs. OOA is more closely related to primary vs. secondary sources, directs the PMF deconvolution of the unified dataset along these lines. Similarly, the AMS primary/secondary classification is enhanced by the inclusion of PTR-MS data. This is shown in Fig. 13, where the aerosol mass fraction attributed to OOA factors is plotted as a function of the toluene-to-benzene ratio for the AMS and unified datasets. The figure shows a significantly tighter correlation between the OOA mass fraction and the toluene/benzene ratio for the unified dataset (R^2=0.53 vs. R^2=0.06 for the AMS dataset). Correlation between these two quantities is expected because the toluene/benzene ratio is inversely related to photochemical age (Roberts et al., 1984) and the oxygen content of organic aerosol is known to increase with photochemical age.

Through inclusion of the PTR-MS data in the unified dataset, the AMS OOA is resolved into OOA-I and OOA-II components. As noted in the previous section, both OOA-I and OOA-II have been extracted in several AMS datasets (Lanz et al., 2007; Ulbrich et al., 2008). However, the OOA-II could not be distinguished in the present dataset using only AMS data. The key factor in resolving OOA-II in the unified dataset is the distinction between short-lived photochemical reaction products (formaldehyde,

ACPD
9, 6739–6785, 2009

acetaldehyde) vs. long-lived products (acetone, acetic acid). The OOA-I/OOA-II distinction is important because it provides information about the relative age of the particulate organics: time periods with a higher ratio of OOA-I to OOA-II are generally more processed. Lanz et al. (2008a) encountered a similar situation with respect to a wood-burning factor that was not resolvable from PMF of AMS-only data, but was extracted by forcing the mass spectrum of one factor towards the desired profile. The present study indicates that inclusion of tracer species in the PMF analysis is useful in extracting hard-to-resolve factors in such cases.

Another important feature of the unified solution is the re-mixing or disappearance of several factors resolved in the individual datasets. For the PTR-MS factors, the prominent features in the time series are generally redistributed among factors in ways that enhance interpretation of the dataset, as described above in terms of primary vs. secondary species. However, different behavior is observed with respect to the distribution of mass between $F1_{AMS}$ (charbroiling) vs. $F2_{AMS}$ (HOA) into $F1_{UN}$ and $F2_{UN}$, and the disappearance from the unified dataset solutions of $F4_{AMS}$ (biomass burning) and $F5_{AMS}$ (northeast point source).

The charbroiling and HOA factors have very similar mass spectra and cannot be reliably distinguished under some circumstances (Ulbrich et al., 2008). For the AMS dataset, the proximity of the sampling inlet to a major charbroiling source contributes to a distinctive time series for $F1_{AMS}$, aiding in the separation of the $F1_{AMS}$ and $F2_{AMS}$ factors. While the factors remain separated in the unified dataset, the mass fraction of each changes significantly. The charbroiling and HOA factors respectively contain 26.1% and 16.6% of the AMS signal in the individual dataset, vs. 47.3% and 5.8% in the unified dataset. This suggests that the charbroiling vs. HOA apportionment is not fully quantitative in either dataset.

Factors $F4_{AMS}$ (biomass burning) and $F5_{AMS}$ (northeast point source) do not resemble any of the factors resolved in the unified dataset. These factors contained relatively small mass fractions in the AMS (7.6% and 7.7%, respectively). This locates them very close to the ~5% limit for AMS factor extraction observed in synthetic datasets

ACPD
9, 6739–6785, 2009

(Ulbrich et al., 2008). While the analysis presented for the AMS dataset suggests that these factors are in fact real, they make only minor contributions to the AMS data and do not exert sufficient influence on Q to be resolved. The major features of the $F5_{AMS}$ time series (Fig. 3b) are contained in the AMS residuals for the unified dataset (Fig. 9). However, variations in the other factors exert a stronger influence on Q.

4 Conclusions

We present the first application of positive matrix factorization (PMF) to a unified AMS/PTR-MS dataset. The datasets are equally weighted through the use of a residual based metric, $\Delta\overline{e_{sc}}$. This method can be directly applied to any dataset containing two mass spectrometers, and is readily generalized to account for three or more instruments. The unified dataset provides a complementary approach to analysis of the individual instrument datasets. In this study, the previously identified oxygenated aerosol factors OOA-I and OOA-II could be distinguished only within the unified dataset. Further, the unified dataset greatly enhanced interpretation of oxygenated VOC sources, apportioning them into primary sources vs. secondary reaction products. Minor factors in the individual dataset of one instrument lacking corresponding tracers in the other may not be resolvable in the unified dataset.

Acknowledgements. Partial operational funding was provided by the Canadian Foundation for Climate and Atmospheric Sciences (CFCAS) and the Natural Sciences and Engineering Council of Canada (NSERC). This project has also received funding support from the Ontario Ministry of the Environment. Such support does not indicate endorsement by the Ministry of the contents of this material. Funding for SOCAAR was provided by the Canadian Foundation for Innovation, the Ontario Innovation Trust, and the Ontario Research Fund.

References

Allan, J. D., Coe, H., Bower, K. N., Alfarra, M. R., Delia, A. E., Jimenez, J. L., Middlebrook, A. M., Drewnick, F., Onasch, T. B., Canagaratna, M. R., Jayne, J. T., and Worsnop, D. R.: Technical Note: Extraction of Chemically Resolved Mass Spectra from Aerodyne Aerosol Mass Spectrometer Data, J. Aerosol Sci., 35, 909–922, 2004.

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.

Andreae, M. O. and Merlet, P.: Emission of Trace Gases and Aerosols from Biomass Burning, Global Biogeochem. Cy., 15, 955–966, 2001.

Beychok, M. R.: A Data Base for Dioxin and Furan Emissions from Municipal Refuse Incinerators, Atmos. Environ., 21, 29–36, 1987.

Buset, K. C., Evans, G. J., Leaitch, W. R., Brook, J. R., and Toom-Sauntry, D.: Use of Advanced Receptor Modelling for Analysis of an Intensive 5-Week Aerosol Sampling Campaign, Atmos. Environ., 40, 482–499, 2006.

Buzcu, B. and Frazier, M. P.: Source Identification and Apportionment of Volatile Organic Compounds in Houston, TX, Atmos. Environ., 40, 2385–2400, 2006.

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. M., 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, 2007.

de Gouw, J. A., Goldan, P. D., Warneke, C., Kuster, W. C., Roberts, J. M., Marchewka, M., Bertman, S. B., Pszenny, A. A. P., and Keene, W. C.: Validation of Proton Transfer Reaction-Mass Spectrometry (PTR-MS) Measurements of Gas-Phase Organic Compounds in the Atmosphere During the New England Air Quality Study (NEAQS) in 2002, J. Geophys. Res., 108, 4682, doi:10.1029/2003JD003863, 2003.

de Gouw, J. A., Middlebrook, A. M., Warneke, C., Goldan, P. D., Kuster, W. C., Roberts, J. M., Fehsenfeld, F. C., Worsnop, D. R., Canagaratna, M. R., Pszenny, A. A. P., Keene, W. C., Marchewka, M., Bertman, S. B., and Bates, T. S.: Budget of Organic Carbon in a Polluted Atmosphere: Results from the New England Air Quality Study in 2002, J. Geophys. Res.,

110, D16305, doi:10.1029/2004JD005623, 2005.
de Gouw, J. A. and Warneke, C.: Measurements of Volatile Organic Compounds in the Earth's Atmosphere using Proton-Transfer-Reaction Mass Spectrometery, Mass Spectrom. Rev., 26, 223–257, 2007.
Donahue, N. M., Robinson, A. L., Stanier, C. O., and Pandis, S. N.: Coupled Partitioning, Dilution, and Chemical Aging of Semivolatile Organics, Environ. Sci. Technol., 40, 2635–2643, 2006.
Drewnick, F., Hings, S. S., DeCarlo, P. F., Jayne, J. T., Gonin, M., Fuhrer, K., Weimer, S., Jimenez, J. L., Demerjian, K. L., Borrmann, S., and Worsnop, D. R.: A New Time-of-Flight Aerosol Mass Spectrometer (ToF-AMS) – Instrument Description and First Field Deployment, Aerosol Sci. Technol., 39, 637–658, 2005.
Friedfeld, S., Fraser, M., Ensor, K., Tribble, S., Rehle, D., Leleux, D., and Tittle, F.: Statistical Analysis of Primary and Secondary Atmospheric Formaldehyde, Atmos. Environ., 36, 4767–4775, 2002.
Garcia, A. R., Volkamer, R., Molina, L. T., Molina, M. J., Samuelson, J., Mellqvist, J., Galle, B., Herndon, S. C., and Kolb, C. E.: Separation of emitted and photochemical formaldehyde in Mexico City using a statistical analysis and a new pair of gas-phase tracers, Atmos. Chem. Phys., 6, 4545–4557, 2006, http://www.atmos-chem-phys.net/6/4545/2006/.
Hansel, A., Jordan, A., Holzinger, R., Prazeller, P., Vogel, W., and Lindinger, W.: Proton Transfer Reaction Mass Spectrometry: On-line Trace Gas Analysis at ppb Level, Int. J. Mass Spectrom., 149/150, 609–619, 1995.
Holzinger, R., Millet, D. B., Williams, B., Lee, A., Kreisberg, N., Hering, S. V., Jimenez, J. L., Allan, J. D., Worsnop, D. R., and Goldstein, A. H.: Emission, Oxidation, and Secondary Organic Aerosol Formation of Volatile Organic Compounds as Observed at Chebogue Point, Nova Scotia, J. Geophys. Res., 112, D10S24, doi:10.1029/2006JD007599, 2007.
Kanakidou, M., Seinfeld, J. H., Pandis, S. N., Barnes, I., Dentener, F. J., Facchini, M. C., Van Dingenen, R., Ervens, B., Nenes, A., Nielsen, C. J., Swietlicki, E., Putaud, J. P., Balkanski, Y., Fuzzi, S., Horth, J., Moortgat, G. K., Winterhalter, R., Myhre, C. E. L., Tsigaridis, K., Vignati, E., Stephanou, E. G., and Wilson, J.: Organic aerosol and global climate modelling: a review, Atmos. Chem. Phys., 5, 1053–1123, 2005, http://www.atmos-chem-phys.net/5/1053/2005/.
Kristensson, K., Johansson, C., Westerholm, R., Swietlicki, E., Gidhagen, L., Wideqvist, U.,

and Vesely, V.: Real-world Traffic Emission Factors of Gases and Particles Measured in a Road Tunnel in Stockholm, Sweden, Atmos. Environ., 38, 657–673, 2004.
Lanz, V. A., Alfarra, 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, 2007, http://www.atmos-chem-phys.net/7/1503/2007/.
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, 2008a.
Lanz, V. A., Hueglin, C., Buchmann, B., Hill, M., Locher, R., Staehelin, J., and Reimann, S.: Receptor modeling of C_2C_7 hydrocarbon sources at an urban background site in Zurich, Switzerland: changes between 1993–1994 and 2005–2006, Atmos. Chem. Phys., 8, 2313–2332, 2008b, http://www.atmos-chem-phys.net/8/2313/2008/.
Lee, S. H.: Emissions from Street Vendor Cooking Devices (Charcoal Grilling), US EPA, Washington DC, USA, 1999.
Lewis, C. W., Norris, G. A., Conner, T. L., and Henry, R. C.: Source Apportionment of Phoenix $PM_{2.5}$ Aerosol with the UNMIX Receptor Model, J. Air Waste Manage., 53, 325–338, 2003.
Li, S.-M., Anlauf, K. G., Wiebe, H. A., Bottenheim, J. W., Shepson, P. B., and Biesenthal, T. A.: Emission Ratios and Photochemical Production Efficiencies of Nitrogen Oxides, Ketones, and Aldehydes in the Lower Fraser Valley during the Summer Pacific 1993 Oxidant Study, Atmos. Environ., 31, 2037–2048, 1997.
Lindinger, W., Hansel, A., and Jordan, A.: Proton-Transfer Reaction Mass Spectrometry (PTR-MS): On-line Monitoring of Volatile Organic Compounds at pptv Levels, Chem. Soc. Rev., 27, 347–354, 1998.
McLafferty, F. W.: Interpretation of Mass Spectra, University Science Books, Mill Valley, CA, USA, 1980.
Owega, S., Khan, B. U. Z., D'Souza, R., Evans, G. J., Fila, M., and Jervis, R. E.: Receptor Modeling of Toronto PM2.5 Characterized by Aerosol Laser Ablation Mass Spectrometry, Environ. Sci. Technol., 38, 5712–5720, 2004.
Paatero, P.: Least Squares Formulation of Robust Non-Negative Factor Analysis, Chemometr. Intell. Lab., 37, 23–35, 1997.

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, 1994.

Possanzini, M., Palo, V. D., and Cecinato, A.: Sources and Photodecomposition of Formaldehyde and Acetaldehyde in Rome Ambient Air, Atmos. Environ., 36, 3195–3201, 2002.

Quinn, P. K., Bates, T. S., Coffman, D., Onasch, T. B., Worsnop, D. R., Baynard, T., de Gouw, J. A., Goldan, P. D., Kuster, W. C., Williams, E., Roberts, J. M., Lerner, B., Stohl, A., Pettersson, A., and Lovejoy, E. R.: Impacts of Sources and Aging on Submicrometer Aerosol Properties in the Marine Boundary Layer Across the Gulf of Maine, J. Geophys. Res., 111, D23S36, doi:10.1029/2006JD007582, 2006.

Roberts, J. M., Fehsenfeld, F. C., Liu, S. C., Bollinger, M. J., Hahn, C., Albritton, D. L., and Sievers, R. E.: Measurements of Aromatic Hydrocarbon Ratios and NO_x Concentrations in the Rural Troposphere: Estimates of Air Mass Photochemical Age and NO_x Removal Rate, Atmos. Environ., 18, 2421–2432, 1984.

Robinson, A. L., Donahue, N. M., Shrivastava, M. K., Weitkamp, E. A., Sage, A. M., Grieshop, A. P., Lane, T. E., Pierce, J. R., and Pandis, S. N.: Rethinking Organic Aerosol: Semivolatile Emissions and Photochemical Aging, Science, 315, 1259–1262, 2007.

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. Discuss., 8, 6729–6791, 2008, http://www.atmos-chem-phys-discuss.net/8/6729/2008/.

Ulbrich, I. M., Lechner, M., and Jimenez, J. L.: AMS Spectral Database, available online at: http://cires.colorado.edu/jimenez-group/AMSsd/, last access: 2009, 2007.

Weimer, S., Alfarra, M. R., Schreiber, D., Mohr, M., Prevot, A. S. H., and Baltensperger, U.: Organic Aerosol Mass Spectral Signatures from Wood-Burning Emissions: Influence of Burning Conditions and Wood Type, J. Geophys. Res., 113, D10304, doi:10.1029/2007JD009309, 2008.

Zervas, E., Montagne, X., and Lahaye, J.: C-C Organic Acid Emissions from an SI Engine: Influence of Fuel and Air/Fuel Equivalence Ratio, Environ. Sci. Technol., 35, 2746–2751, 2001.

Zhang, Q., Alfarra, M. R., Worsnop, D. R., Allan, J. D., Coe, H., Canagaratna, M. R., and Jimenez, J. L.: Deconvolution and Quantification of Hydrocarbon-like and Oxygenated Organic Aerosols Based on Aerosol Mass Spectrometry, Environ. Sci. Technol., 39, 4938–4952, 2005.

Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Allan, J. D., Coe, H., Ulbrich, I., Alfarra, M. R., Takami, A., Middlebrook, A. M., Sun, Y. L., Dzepina, K., Dunlea, E., Docherty, K., DeCarlo, P. F., Salcedo, D., Onasch, T. B., Jayne, J. T., Miyoshi, T., Shimono, A., Hatakeyama, S., Takegawa, N., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K. L., Williams, P., Bower, K., Bahreini, R., Cottrell, L., Griffin, R. J., Rautiainen, J., Sun, J. Y., Zhang, Y. M., and Worsnop, D. R.: Ubiquity and Dominance of Oxygenated Species in Organic Aerosols in Anthropogenically-influenced Northern Hemisphere Midlatitudes, Geophys. Res. Lett., 34, L13801, doi:10.1029/2007GL029979, 2007.

Zhao, W. X. and Hopke, P. K.: Source Identification for Fine Aerosols in Mammoth Cave National Park, Atmos. Res., 80, 309–322, 2006.

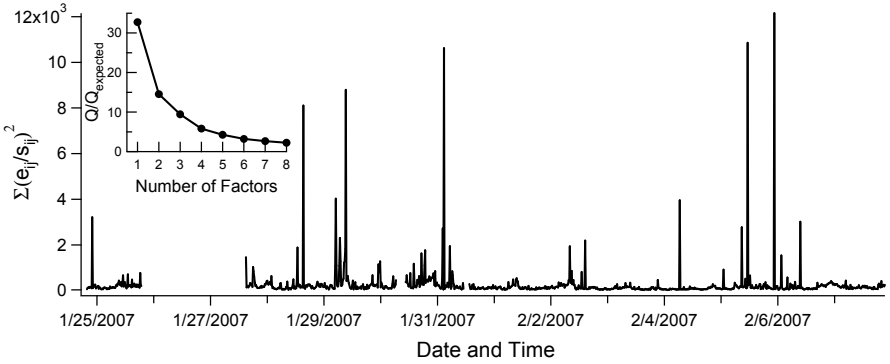

Fig. 1. Residual time series for the 6-factor solution to the AMS dataset. The inset shows the ratio $Q/Q_{expected}$ as a function of the number of factors in the solution.

ACPD
9, 6739–6785, 2009

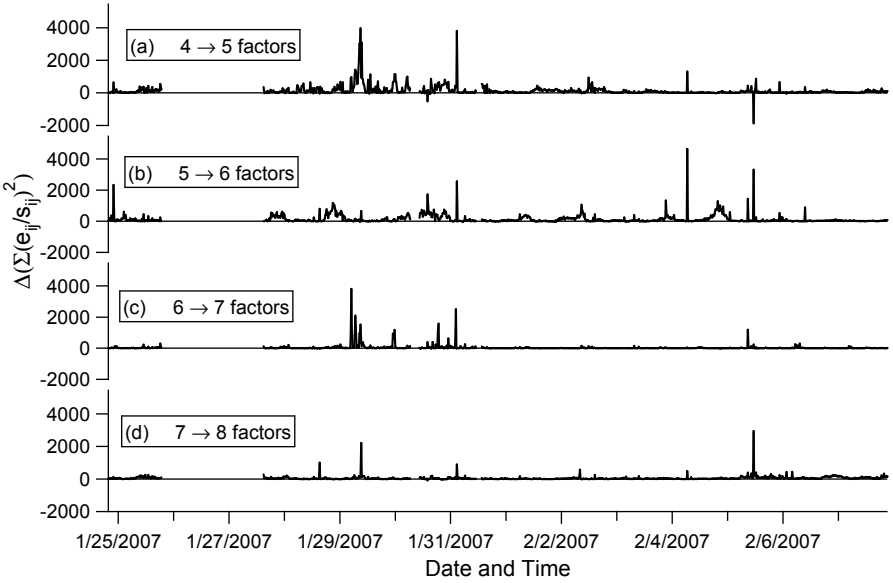

Fig. 2. Effect of the number of factors contained in a solution on the residual time series for the AMS dataset.

ACPD
9, 6739–6785, 2009

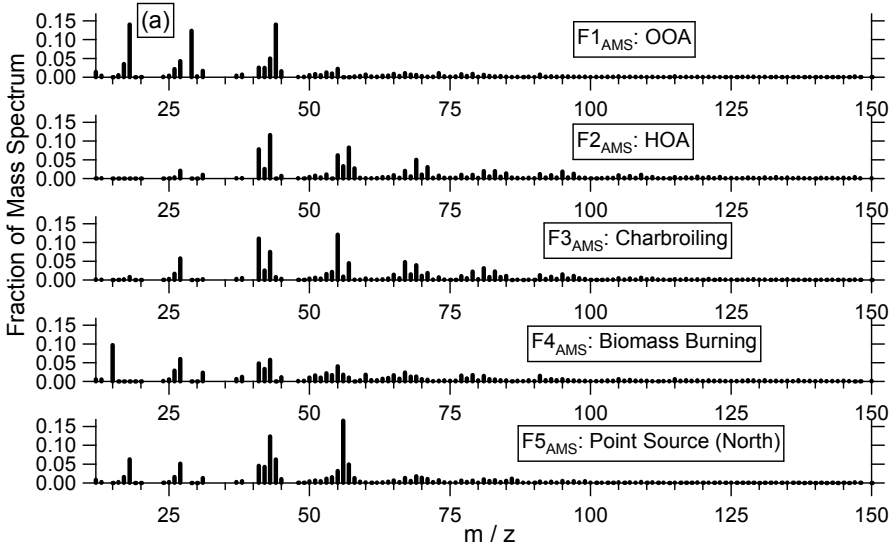

Fig. 3a. Mass spectra **(a)** and time series **(b)** for the PMF solution to the AMS dataset. (b) includes the time series both for the PMF factors (black traces, left axis) and selected tracer species (gray traces, right axis).

ACPD
9, 6739–6785, 2009

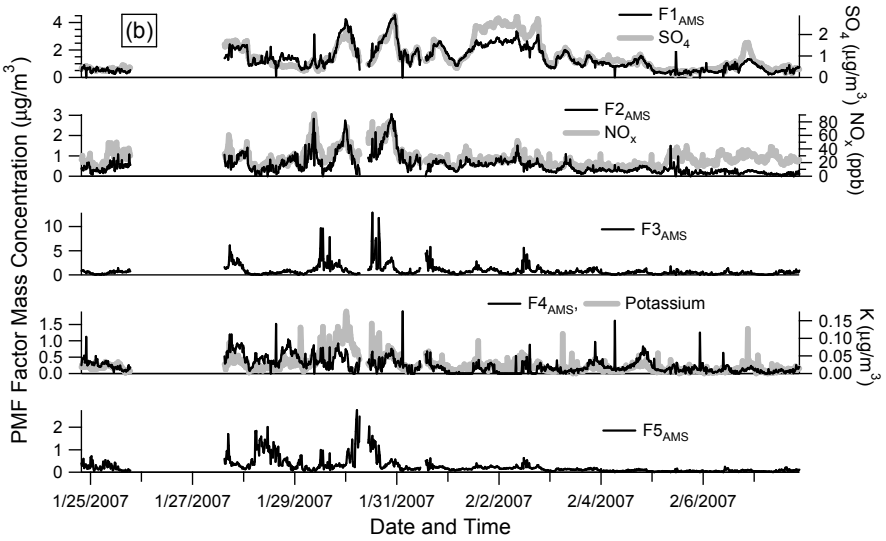

Fig. 3b. Continued.

ACPD
9, 6739–6785, 2009

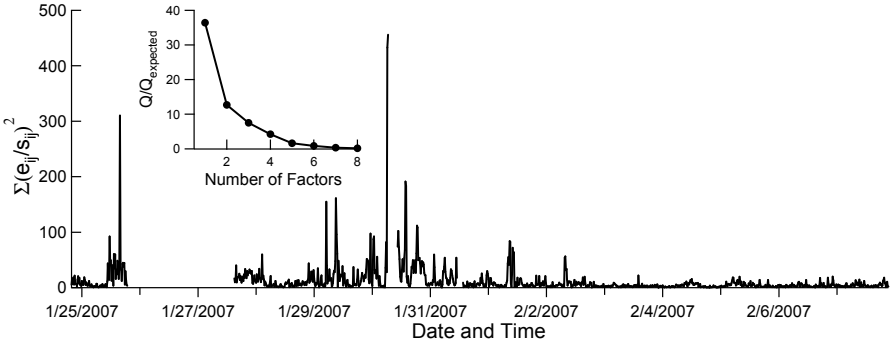

Fig. 4. Residual time series for the 5-factor solution to the PTR-MS dataset. The inset shows the ratio $Q/Q_{expected}$ as a function of the number of factors in the solution.

ACPD
9, 6739–6785, 2009

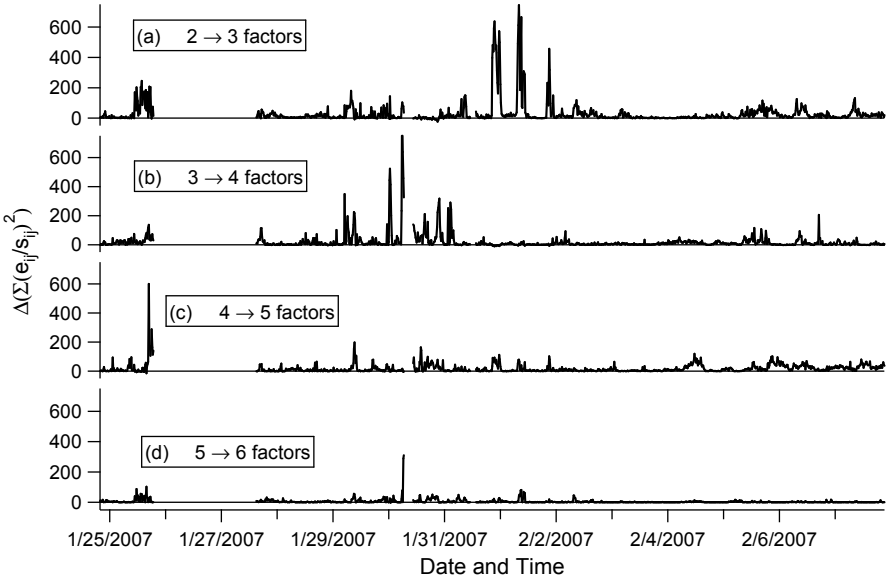

Fig. 5. Effect of the number of factors contained in a solution on the residual time series for the PTR-MS dataset.

ACPD
9, 6739–6785, 2009

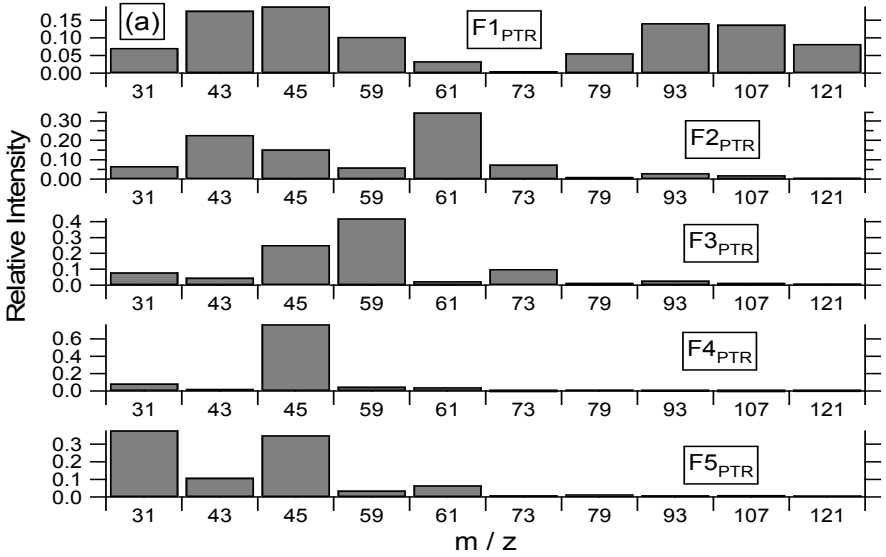

Fig. 6a. Mass spectra **(a)** and time series **(b)** for the PMF solution to the AMS dataset. (b) includes the time series both for the PMF factors (black traces, left axis) and selected tracer species (gray traces, right axis). Note that the temperature axis (vs. $F4_{PTR}$) is reversed.

ACPD
9, 6739–6785, 2009

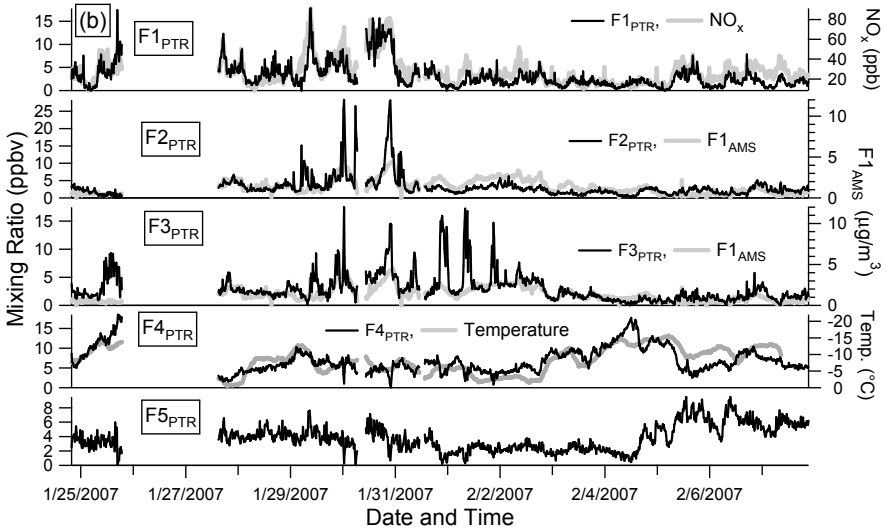

Fig. 6b. Continued.

ACPD
9, 6739–6785, 2009

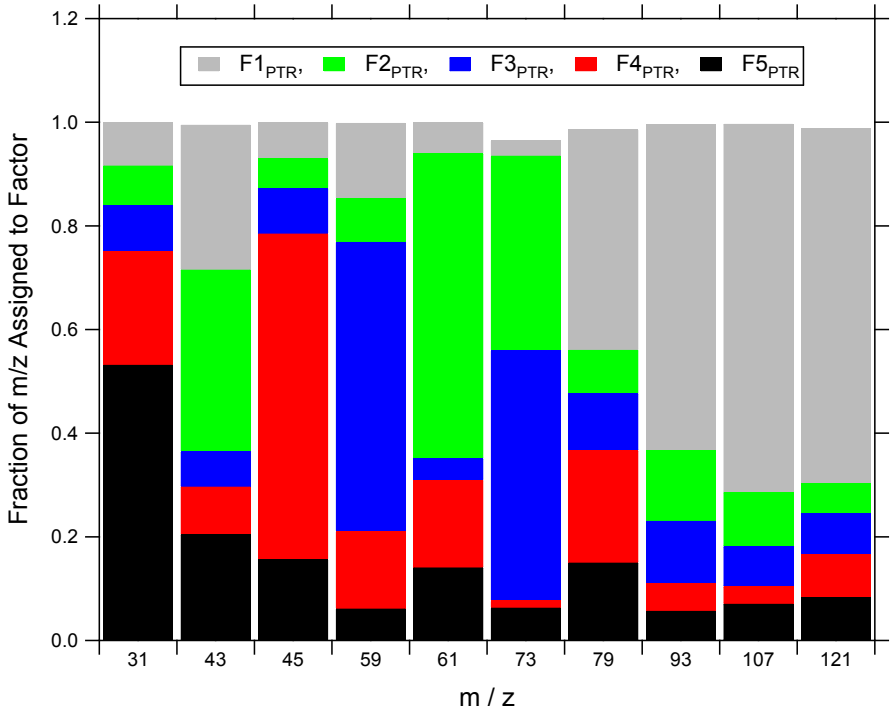

Fig. 7. Mass spectra for the 5-factor solution to the PTR-MS dataset represented as the fraction of signal at each *m/z* assigned to the specified factor. Differences between the stacked total and 1 are due to signal contained in the residuals.

ACPD
9, 6739–6785, 2009

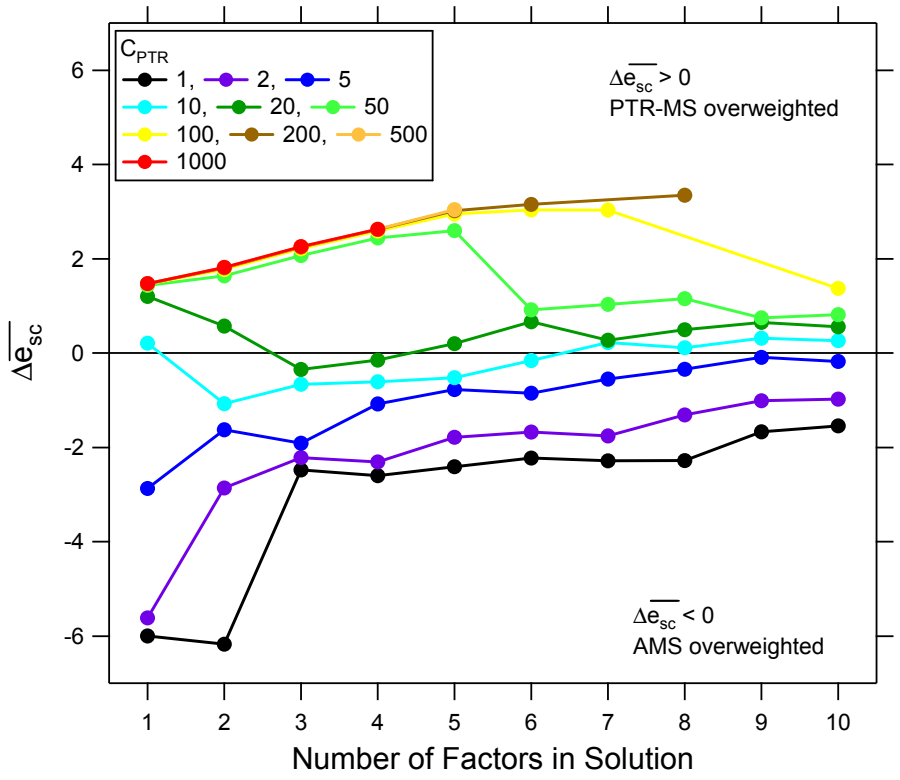

Fig. 8. Change in the mean scaled residual between the AMS and PTR-MS ($\Delta\overline{e_{sc}}$) as a function of the number of factors in the solution (x-axis) and the PTR-MS weighting factor (C_{PTR}). Instruments carry their natural weight at C_{PTR}=1; the instruments are balanced at $\Delta\overline{e_{sc}}$=0.

ACPD
9, 6739–6785, 2009

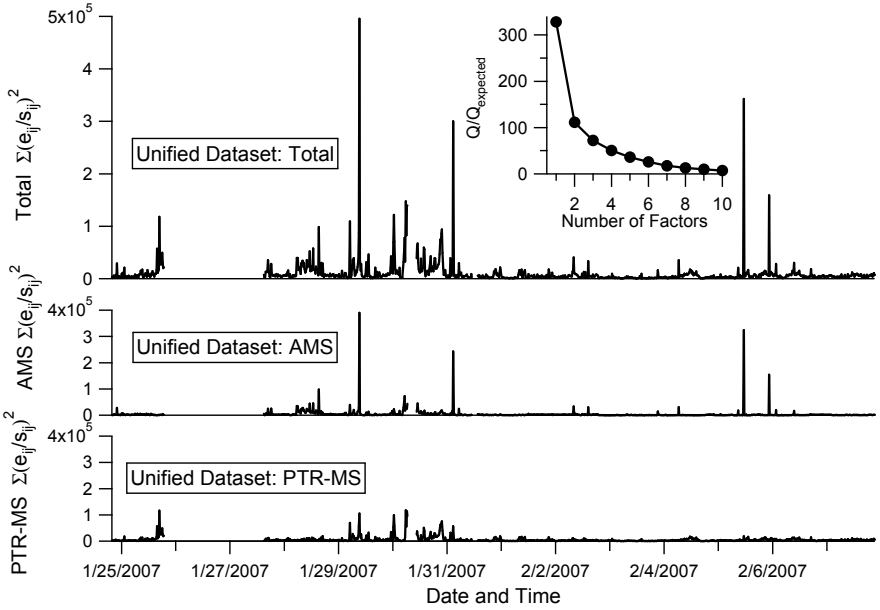

Fig. 9. Residual time series for the 5-factor solution to the unified dataset. The inset shows the ratio $Q/Q_{expected}$ as a function of the number of factors in the solution. Separate residual traces are shown for total, AMS, and PTR-MS residuals.

ACPD
9, 6739–6785, 2009

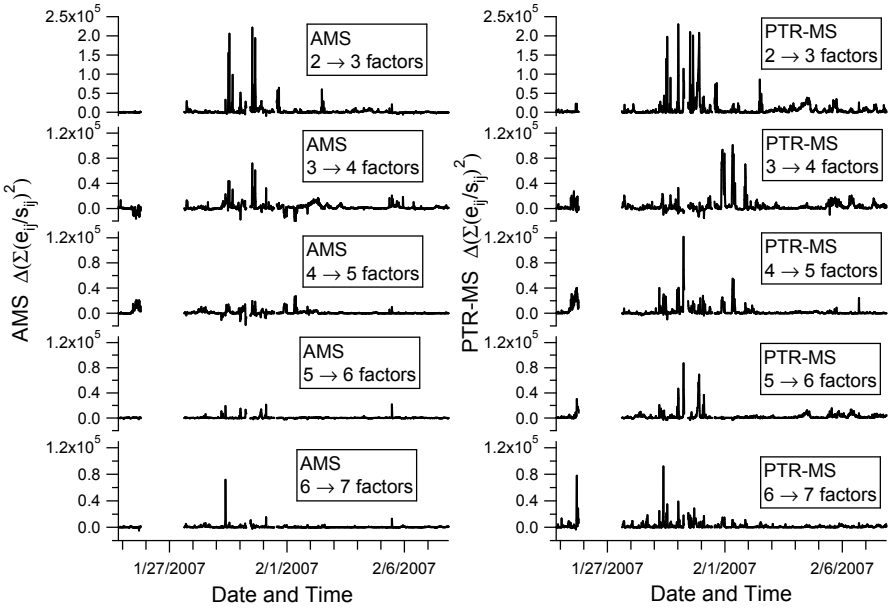

Fig. 10. Effect of the number of factors contained in a solution on the residual time series for the unified dataset. AMS and PTR-MS residuals are separately in the left and right-hand columns, respectively. Note the 2→3 factor plots are on a different vertical scale.

ACPD
9, 6739–6785, 2009

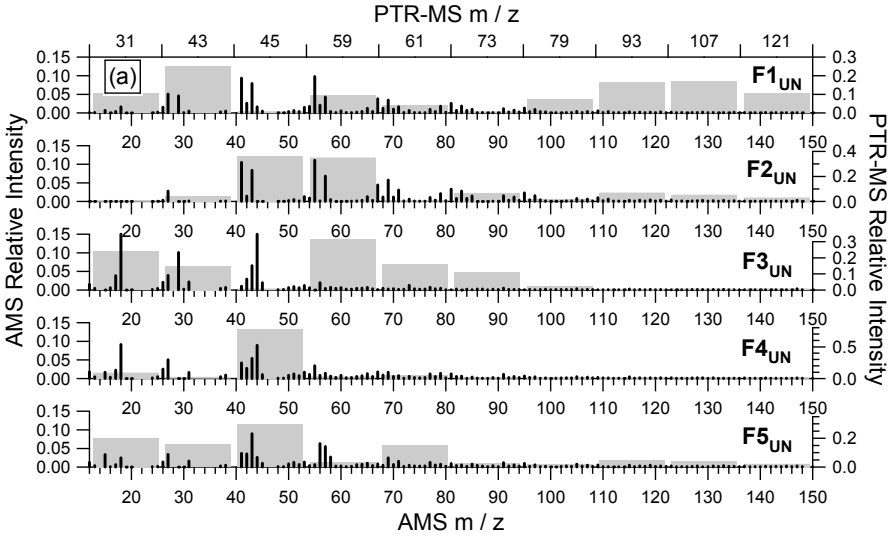

Fig. 11a. Mass spectra **(a)** and time series **(b)** for the PMF solution to the unified dataset. (b) includes the time series both for the PMF factors (black traces, left axis) and selected tracer species (gray traces, right axis). Note the temperature axis (vs. $F4_{UN}$) is reversed.

ACPD
9, 6739–6785, 2009

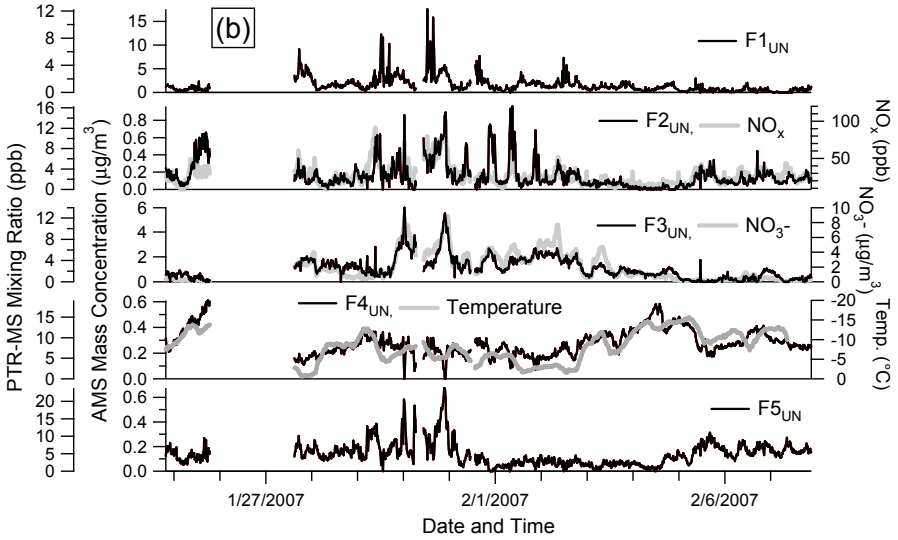

Fig. 11b. Continued.

ACPD
9, 6739–6785, 2009

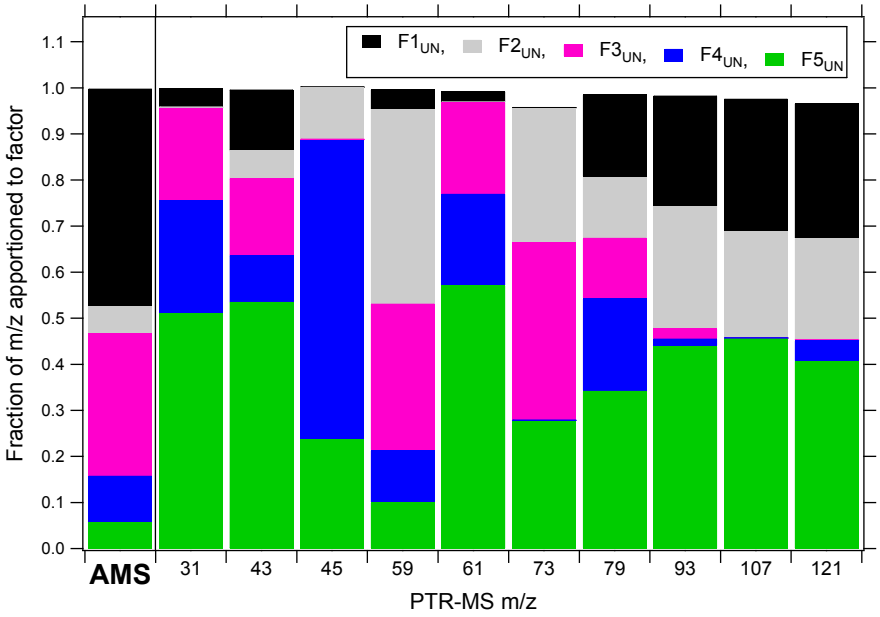

Fig. 12. Mass spectra for the 5-factor solution to the unified dataset represented as the fraction of signal in the AMS and at each PTR-MS *m/z* assigned to the specified factor. Differences between the stacked total and 1 are due to signal contained in the residuals.

ACPD
9, 6739–6785, 2009

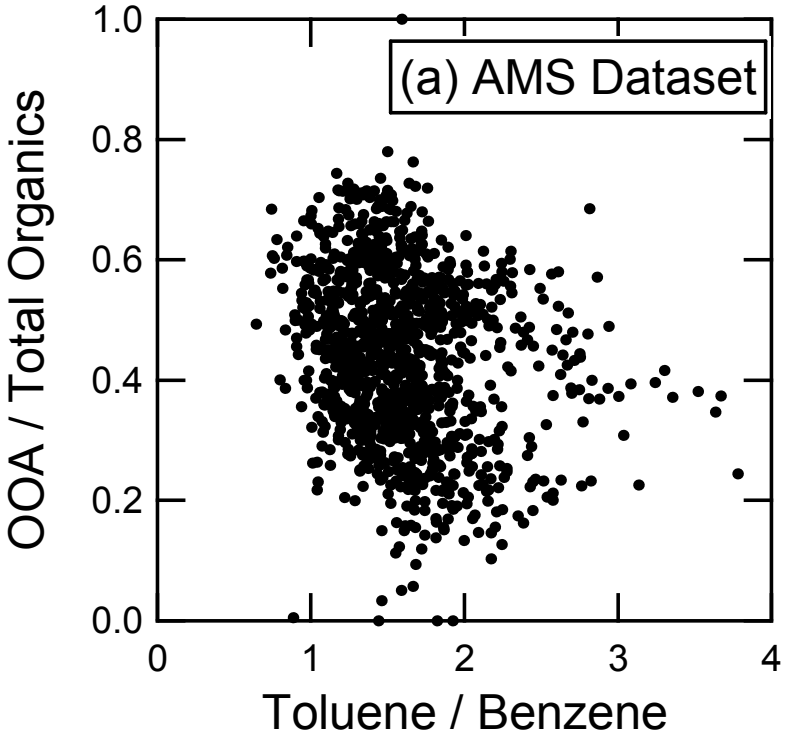

Fig. 13a. Mass fraction of oxygenated aerosol as a function of the toluene/benzene ratio. The toluene/benzene ratio is inversely related to photochemical age.

ACPD
9, 6739–6785, 2009

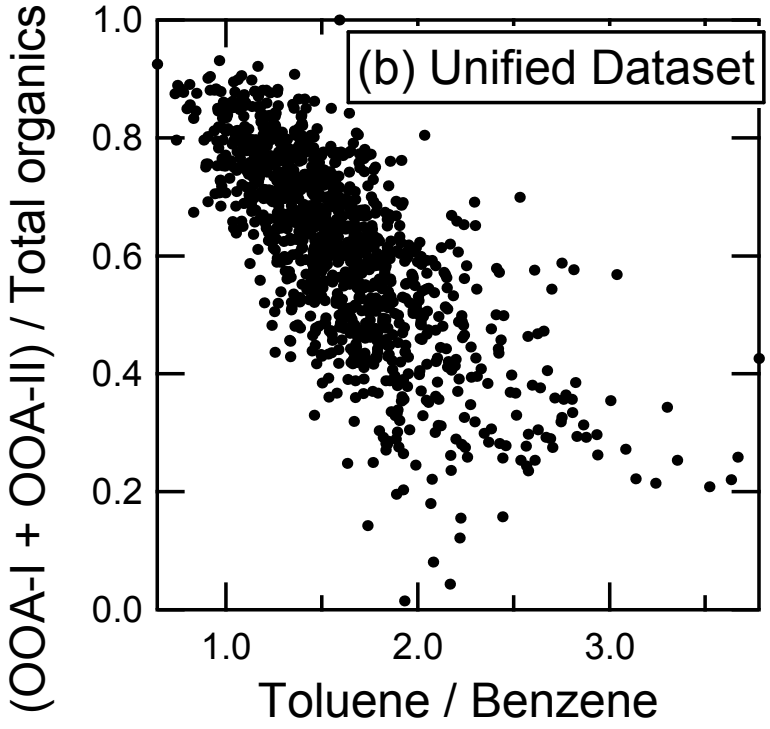

Fig. 13b. Continued.

ACPD
9, 6739–6785, 2009