Atmospheric
Measurement
Techniques

A technique for rapid source apportionment applied to ambient organic aerosol measurements from a thermal desorption aerosol gas chromatograph (TAG)

Yaping Zhang[1], Brent J. Williams[1], Allen H. Goldstein[2], Kenneth S. Docherty[3,4,a], and Jose L. Jimenez[3,4]

[1]Department of Energy, Environmental, and Chemical Engineering, Washington University in St. Louis,
St. Louis, Missouri, USA
[2]Department of Environmental Science, Policy, & Management, University of California, Berkeley, California, USA
[3]Cooperative Institute for Research in the Environmental Sciences (CIRES), University of Colorado,
Boulder, California, USA
[4]Dept. of Chemistry & Biochemistry, University of Colorado, Boulder, Colorado, USA
[a]currently at: Alion Science and Technology, US EPA Office of Research and Development,
Research Triangle Park, North Carolina, USA

Correspondence to: Brent J. Williams (brentw@wustl.edu)

Abstract. We present a rapid method for apportioning the sources of atmospheric organic aerosol composition measured by gas chromatography–mass spectrometry methods. Here, we specifically apply this new analysis method to data acquired on a thermal desorption aerosol gas chromatograph (TAG) system. Gas chromatograms are divided by retention time into evenly spaced bins, within which the mass spectra are summed. A previous chromatogram binning method was introduced for the purpose of chromatogram structure deconvolution (e.g., major compound classes) (Zhang et al., 2014). Here we extend the method development for the specific purpose of determining aerosol samples' sources. Chromatogram bins are arranged into an input data matrix for positive matrix factorization (PMF), where the sample number is the row dimension and the mass-spectra-resolved eluting time intervals (bins) are the column dimension. Then two-dimensional PMF can effectively do three-dimensional factorization on the three-dimensional TAG mass spectra data. The retention time shift of the chromatogram is corrected by applying the median values of the different peaks' shifts. Bin width affects chemical resolution but does not affect PMF retrieval of the sources' time variations for low-factor solutions. A bin width smaller than the maximum retention shift among all samples requires retention time shift correction. A

six-factor PMF comparison among aerosol mass spectrometry (AMS), TAG binning, and conventional TAG compound integration methods shows that the TAG binning method performs similarly to the integration method. However, the new binning method incorporates the entirety of the data set and requires significantly less pre-processing of the data than conventional single compound identification and integration. In addition, while a fraction of the most oxygenated aerosol does not elute through an underivatized TAG analysis, the TAG binning method does have the ability to achieve molecular level resolution on other bulk aerosol components commonly observed by the AMS.

1 Introduction

Atmospheric aerosols can impact human health (Dominici et al., 2006; Gauderman et al., 2015), atmospheric visibility (Sun et al., 2006; Junjun et al., 2014), the water cycle, and climate change (IPCC, 2013). Anthropogenic and natural sources emit primary pollutants, which undergo atmospheric chemical and physical transformation to produce secondary pollutants. Organic aerosols (OAs) account for 20–70 % of total fine aerosols (PM$_1$) (Jimenez et al., 2009; Mur-

phy et al., 2006). Their chemical composition can comprise thousands of organic compounds, whose sources and transformations are not fully understood due to their complexity and dynamic chemical properties and gas–particle partitioning (Hallquist et al., 2009; Goldstein and Galbally, 2007). While inorganic ions, elemental carbon (EC)/organic carbon (OC), OC functional groups, and trace metals from offline filters analyses of atmospheric aerosol have been used for source apportionment (Chueinta et al., 2000; Ito et al., 2004; Lee et al., 1999; Ramadan et al., 2000; Ahlm et al., 2013), recently high-time-resolution Aerodyne aerosol mass spectrometer (AMS) mass spectra and the Aerosol Chemical Speciation Monitor (ACSM) have been extensively used to determine the major components of ambient OA (Zhang et al., 2011; Ng et al., 2011). Online and offline measurements of molecular level marker molecules have also been used to apportion the major chemical components and source attributions of these organic aerosols (Schauer et al., 1996; Jaeckels et al., 2007; Zhang et al., 2014; Williams et al., 2010, 2014).

The Aerodyne AMS is a widely used instrument for aerosol analysis due to its capability to quantitatively characterize the size-resolved bulk composition of PM_1 (Canagaratna et al., 2007). AMS reports the bulk (also size-resolved) composition of PM_1 in the form of ensemble mass spectra, which are generated from the linear superposition of the mass spectra of individual compounds. Positive matrix factorization (PMF), a multivariate factor analysis method (Paatero, 1997; Ulbrich et al., 2009), is applied to the ensemble mass spectra and deconvolves the spectra into several factors with approximately constant mass spectra and consistent temporal behavior. Each of these factors can represent hundreds to thousands of organic compounds from a source or atmospheric process. The use of this technique has been growing rapidly in the last 10 years due to its broad applicability (Zhang et al., 2011). However, AMS inherently has limited chemical resolution because it reports ensemble mass spectra with high fragmentation, and some important aspects of the sources and processes affecting OA are difficult to resolve using only AMS data. To obtain higher chemical resolution, another online technique, called thermal desorption gas chromatograph (TAG), was combined with mass spectrometry (GC–MS) to separate and measure individual compounds (Williams et al., 2006). The organic matter presenting similar mass spectra cannot be resolved by AMS. However, this type of material can be resolved by gas chromatography separation. For example, all alkanes show similar mass spectral patterns with dominant mass spectral peaks at m/z 43, 57, 71, 85, etc. AMS only can separate alkanes from other chemical classes with different functional groups, such as organic acids. AMS cannot separate different alkanes which are all grouped into one component called hydrocarbon-like organic aerosol (HOA) (Zhang et al., 2011), although various alkanes may come from different sources. However, TAG can resolve all alkanes through gas chromatography separation and preserve individual compound information to deter-

mine potential temporal variability differences (Williams et al., 2006, 2010).

TAG is a fully automated, field-deployable instrument that can provide molecular level separation of organic aerosols with 1 h time resolution to help identify specific aerosol source signatures and atmospheric transformation processes, e.g. through PMF analysis of a suite of individual integrated compounds (Williams et al., 2006). In 2008, Goldstein et al. (2008) developed a two-dimensional gas chromatograph combined with an in situ TAG collection system, which can speciate more organic compounds in atmospheric aerosols than TAG alone (Goldstein et al., 2008). In 2013, the Semi-Volatile (SV-)TAG was developed to extend TAG's capability to include quantitative characterization of semi-volatile organic compounds (Zhao et al., 2013), and in 2014 Isaacman et al. (2014) introduced a technique for online chemical derivatization on the SV-TAG system to improve quantification of oxygenated molecules (Isaacman et al., 2014). Recently, a combined TAG-AMS instrument which can simultaneously measure the bulk and speciated composition of organic aerosols has been presented (Williams et al., 2014). The advantage of providing speciated timelines for organic chemicals is significant; however the time required for chromatographic peak identification, integration, and confirmation of integration quality for hundreds of compounds limits the wider application of using a chromatographic approach for aerosol analysis. Additionally, a significant fraction of GC–MS data are typically present as an unresolved complex mixture (UCM) when analyzing ambient aerosol samples (Williams et al., 2006, 2010, 2014, 2016), and peak integration methods typically ignore the material that is not resolved. Therefore, rapid techniques for comprehensively analyzing the complete chromatographically separated mass spectral data, including UCM signal, may broaden the application of the various TAG measurement methods.

One such method was recently introduced, in which each chromatogram is evenly divided into bins and PMF is performed on the covariance of signal vs.retention time to deconvolute the chromatogram into homologous compound series, individual compounds, and multiple UCM components (Zhang et al., 2014). In order to pursue source apportionment, we initially attempted to input these deconvolution chromatogram factors to a second PMF analysis. However, such an approach does not achieve source apportionment results. For example, the factor of homologous compound series contains the full range of compounds (such as C_{12}–C_{40} alkanes). The second PMF cannot effectively separate different sources which may contain just a specific portion of homologous compound series (for example, one source may contain C_{12}–C_{30} alkanes, whereas another source contains C_{31}–C_{40} alkanes).

Here, we present an alternative method specifically designed for source apportionment of ambient organics measured by TAG, where PMF is performed on the covariance of species from different sources to deconvolve the study

period into major contributing sources or aerosol transformation processes. We investigate the data matrix for binning mass spectra, the retention time shift correction, and bin resolution, and we compare the results of the method against those from factor analysis of conventional TAG resolved compound integration method and AMS.

2 Methods

2.1 TAG instrument

Williams et al. describe the TAG instrument in detail (Williams et al., 2006) as applied during the Study of Organic Aerosol at Riverside (SOAR) 2005. Briefly, particles with diameters less than 1.5 μm are humidified and impacted onto a collection and thermal desorption (CTD) cell at 30 °C. The CTD cell is then heated to 310 °C, and the particles are thermally desorbed into a helium carrier gas that transports them into a GC oven at 45 °C, where they re-condense onto the head of the GC column. After sample injection, the GC oven slowly heats to 310 °C, and the compounds eluting through a 30 m low-polarity column are then detected by a quadrupole mass spectrometer. The TAG is fully automated; achieves hourly time resolution; and can cycle between ambient samples (particles + adsorbing semi-volatile gases), filtered samples (adsorbing semi-volatile gases), denuded samples (particles only), cell blanks (no collection), and syringe-injected liquid calibration standards.

2.2 PMF

For PMF analysis, the input data matrix, $\mathbf{X}$, with dimensions of n rows and m columns, is factorized into two matrices – the time series matrix $\mathbf{G}$ ($n \times p$) and the chemical profile matrix $\mathbf{F}$ ($p \times m$), where p is the number of factors – and a third matrix, the residual matrix, $\mathbf{E}$ ($n \times m$):

$$\mathbf{X} = \mathbf{GF} + \mathbf{E}. \tag{1}$$

The PMF model is fitted by weighted least squares, with the weights based on the estimated uncertainties of the individual input matrix data points (σ_{ij} being the estimated uncertainty for data point x_{ij}). Thus mathematical formulation of PMF is to minimize the following function, Q:

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{e_{ij}^2}{\sigma_{ij}^2}, \tag{2}$$

subject to $g_{ik} \geq 0$ and $f_{kj} \geq 0$, $\tag{3}$

where e_{ij}, g_{ik}, and f_{kj} are elements of matrices $\mathbf{E}$, $\mathbf{G}$, and $\mathbf{F}$, and σ_{ij} is the estimated uncertainties of x_{ij}, which is an element of matrix $\mathbf{X}$ (Paatero, 1997).

In this paper, the PMF2 algorithm, a PMF model solver, was used for solving Eq. (1). A custom software tool (PMF Evaluation Tool, PET, version 2.06; Ulbrich et al., 2009) in

Igor Pro (version 6.3, WaveMetrics, Inc.) was used to evaluate PMF outputs and related statistics.

### 2.3 Error estimation for the PMF model

The PMF model is fitted by weighted least squares, and the weights are based on the estimated uncertainties of the input matrix data points. Here we discuss the uncertainty for each ion peak of TAG data. Generally, the uncertainty of an ion signal can be estimated as the square root of the number of ions counted, based on Poisson statistics (Allan et al., 2003), which is referred to as the ion counting (IC) error method. Alternatively, the uncertainty can also be expressed as

$$\text{Unc} = \tag{4}$$
$$\begin{cases} 2 \times \text{MDL}, & \text{if signal} \leq \text{MDL} \\ \sqrt{(\text{signal} \times \text{error fraction})^2 + (\text{MDL})^2} & \text{if signal} > \text{MDL}, \end{cases}$$

where error fraction is reported 10 % for TAG ambient air samples (Williams et al., 2006, 2010); MDL is method detection limit. A detailed description about how to retrieve TAG MDL from ambient measurement is included in the Supplement. PMF on TAG bins' mass spectra was not sensitive to the choice of either of the two error methods above (Zhang et al., 2014). Choosing a method depends on the availability of input data for different error methods. For example, TAG data used in this paper were measured by an Agilent quadrupole mass spectrometer (5973 QMS), which reports not ion counts but rather only an adjusted relative abundance. Thus the MDL error method is used here.

2.4 Data analysis

The original chromatogram binning method for the purpose of chromatogram structure deconvolution utilizes the TAG data set from SOAR 2005 (Zhang et al., 2014). Further development here of the chromatogram binning method for the purpose of source apportionment uses the same data set. Detailed information regarding the SOAR field site and auxiliary measurements can be found in Williams et al. (2010) and Docherty et al. (2008, 2011). The TAG SOAR sampling sequence has been described in detail previously (Zhang et al., 2014). In summary, the TAG system sampled ambient (gas + particle) and filtered ambient (gas-only) data, from which the subtraction yielded particle-only data (with GC column bleed signal also being subtracted by this method). The PMF analysis on binned particle-only data is called TAG–Bin PMF. TAG–Bin PMF indicates a binning method for source apportionment unless stated otherwise. In the original TAG source apportionment paper using individual resolved compounds, 123 major compounds were identified and integrated using the Agilent ChemStation software. A detailed description of the compounds' integration and the PMF on the integrated compounds are given in Williams et al. (2010). Here, the PMF results from analyzing the data
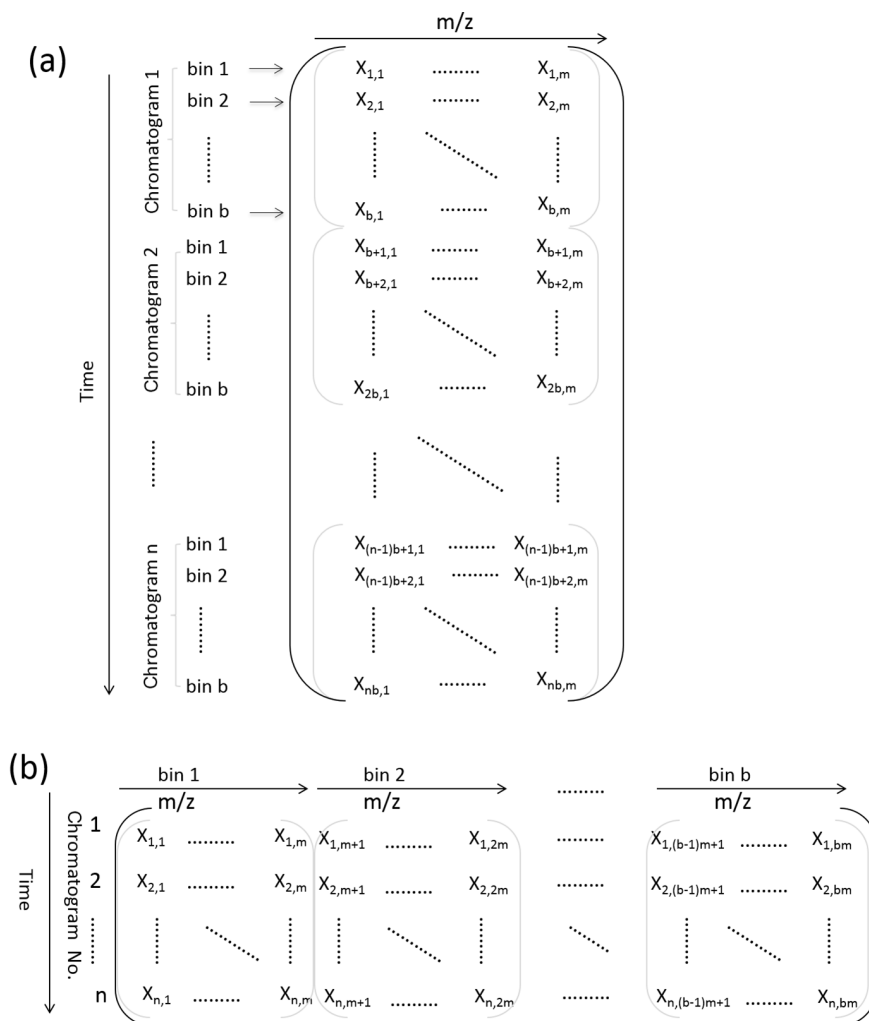
Figure 1. PMF data matrices of **(a)** the binning method for deconvolution of chromatograms and **(b)** source apportionment. The parameter n is the chromatogram number for each hour, b is the bin number, and m is the index of mass spectrum m/z.

set of the 123 resolved compounds is called TAG–Integrated PMF. In the discussion of bin resolution, two-factor solutions are chosen for AMS and TAG–Bin (Integrated) PMF results – hydrocarbon-like OA (HOA) and oxygenated OA (OOA) for comparison. In the comparison of source factors, six-factor solutions are chosen for TAG–Bin and TAG–Integrated, and they are both compared to AMS six-component solutions, as reported by Docherty et al. (2011). Table 1 shows the six AMS components. The comparison between AMS and TAG is based on the assumption that the composition measured by AMS and TAG at PM_1 and $PM_{1.5}$, respectively, overlap significantly.

Table 1. AMS six components of PMF analysis by Docherty et al. (2011).

Factor	Full name	% OA contribution
cLV–OOA	Composite low-volatility oxygenated organic aerosol	31.3
LOA–AC	Amine containing local OA	4.4
LOA-2	Local OA 2	2.6
SV–OOA	Semi-volatile oxygenated OA	14.4
MV–OOA	Medium-volatility oxygenated OA	30.2
HOA	Hydrocarbon-like OA	13.8
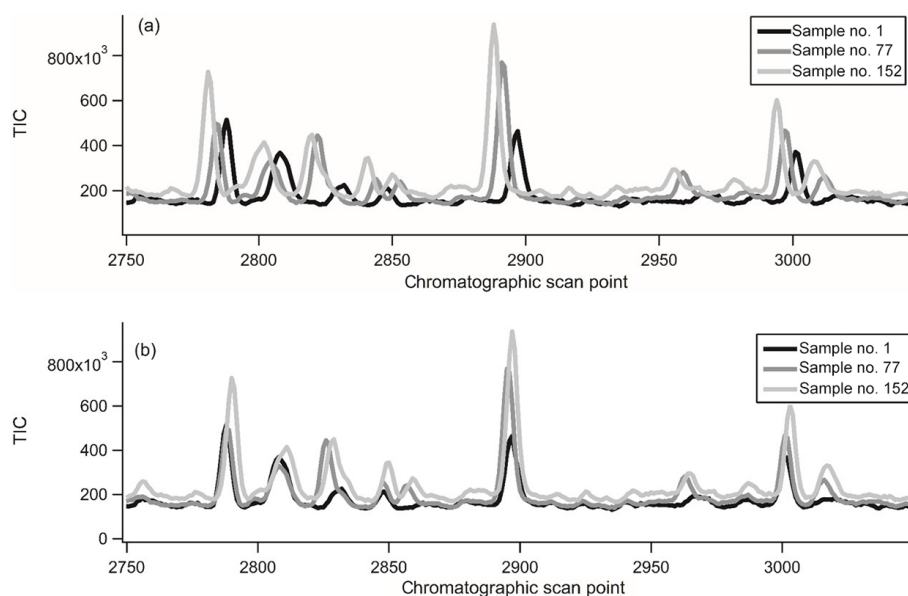
Figure 2. Excerpts of chromatogram for the beginning, middle, and end of the study focus period **(a)** before retention time shift, and **(b)** after retention time shift. The sample collection times for no. 1, 77, and 152 are 29 July 2005 at 00:00, 3 August 2005 at 08:00, and 8 August 2005 at 07:00, respectively.

3 Results and discussion

### 3.1 PMF binning data matrix for source apportionment

For source apportionment, PMF works on the covariance of samples collected at different times. Therefore, the row dimension of binning data for the PMF matrix is only the sample number, and the column dimension is the bins' mass spectra in retention time order. Figure 1 shows the PMF binning data matrix for source apportionment (this study) compared to the previous PMF binning matrix for chromatogram structure (Zhang et al., 2014). In the binning method for source apportionment, the row dimension is the sample number from 1 to n; in the column dimension the first bin's mass spectra range from 1 to m, and the second bin's mass spectra range from $m+1$ to $2 \times m$, where m is the mass spectra m/z index.

3.2 Retention time shift correction

By the nature of the data format used for source apportionment, where the retention time is in the column dimension of the PMF input data matrix, the PMF solution produces a signal chromatogram for each factor, with a fixed signal vs. retention time for the whole study period. However, this is not strictly true for the actual chromatogram samples as recorded, because their retention time will shift from sample to sample due to different sample mass loadings, chemical composition, aerosol water content, and column condition. Figure 2a shows example compound retention time shifts for the beginning, middle, and end of the study focus period. An approximately constant profile, which is an assumption of the PMF model, is required for a successful factor analysis. Therefore, the retention time shift along the sample number dimension should be corrected before PMF analysis. In this study we approximately correct retention time shift by calculating the median values of the retention time shifts along the retention time dimension and correct the retention times accordingly. Figure 2b shows the peak retention time after shift correction using the median value: the peaks display a closer overlap after correction. While this correction greatly improves the retention time shift, it is not an exact correction. As discussed in further detail in Sect. 3.3, multiple scan points will be summed (binned), and retention time shift corrections become less necessary when using larger bin widths. Therefore, with the combination of median value correction and large bin width, the retention time shift issue will be addressed.
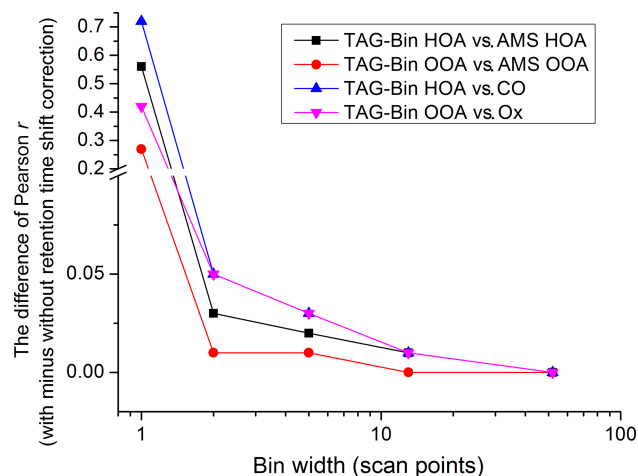
Figure 3 shows the evolution of the median retention time shift during the study time period. The 123 major compounds, used in a prior source apportionment study (Williams et al., 2010) and representing a wide range of the nonpolar and polar compounds, were used to calculate these median values of retention time shifts. A positive median value means the chromatographic peaks shift to the right of the first sample (the elution runs slower), whereas a negative value means peaks shift to the left of the first sample (an earlier, or faster, elution). Almost all the samples in Fig. 3 shift to the left of the first sample. In general, the median values show a linear relationship with the sample number during the study time period. The full range of median shifts among
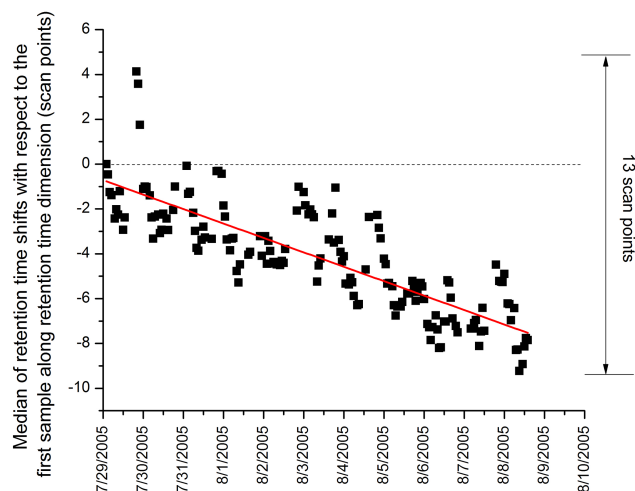
Figure 3. Median values of retention time shifts with respect to the first sample. The positive median value means the chromatographic peaks shift to the right of the first sample (the elution runs slower), whereas the negative value means peaks shift to the left (faster elution). The range of median shifts among all samples during the study period is 13 scan points.

Figure 4. The difference of Pearson r values obtained with retention time shift correction subtracted from values obtained with no retention time shift correction, of TAG–Bin HOA (OOA) time series compared to AMS HOA (OOA), and CO (O_x) time series.

all samples is 13 scan points (corresponding to 36.4 s), and this shift is likely due to a slow change in the condition of the column as the study progressed. However, daily retention time variability is also observed. Figure S2 in the Supplement shows the median variability with respect to the linear fitting line in Fig. 3 and the total ion signal of the TAG samples during the study time period. The median variability is highly anti-correlated ($r = -0.81$) with the total signal of the TAG samples – a metric for aerosol mass loading on the TAG system. The elution runs slower when TAG has less mass loading, whereas the elution runs faster when TAG has a higher mass loading. Mass overloading of the GC system corresponds to the saturation of column stationary phase and can change the peak shape and retention time (Zenkevich and Pavlovskii, 2015). This can be explained by lowered interaction between each molecule in the samples and the stationary column phase when more molecules are present (in a larger sample), allowing material to pass through the column slightly faster.

PMF results with and without retention time shift correction are compared in detail for different bin widths in Sect. 3.3. For future users, the internal standards, the external standards, or the major compounds in samples, all of which work well in automated integration software due to high signal-to-background ratios, can be used to estimate retention time shifts. If desired, additional retention time shift precision can be achieved by including both the long-term median shift from column condition and the daily shift due to sample size. However, it will be shown later that high-precision retention time shifts would only be required if op-

erating this PMF method with very high bin resolution and are not necessary in most analyses of interest.

In addition to retention time shifts, detector response and GC column conditions can drift over the course of a study. Individual bin response factors have not been developed for this data set due to limited calibration standards applied during the SOAR study. Online internal standard injections are now possible (Isaacman et al., 2011), and a complex mixture of various polarity and volatility molecules would need to be analyzed as surrogate species to represent bin response, allowing for a scale to interpolate and track bin response drift corrections. Given the shorter study focus period analyzed here and the relatively high correlations observed below between several TAG–Bin components and AMS components, the bin response correction does not appear critical here, but it should be included in future applications of this method.

3.3 Bin resolution

The effects of different bin widths, with and without retention time (r.t.) shift correction, are compared here. Figure 4 shows the Pearson correlation coefficient difference Δr (= with r.t. shift correction − without r.t. shift correction) of time series for four pairs – TAG–Bin HOA vs. AMS HOA, TAG–Bin OOA vs. AMS OOA, TAG–Bin HOA vs. CO, and TAG–Bin OOA vs. O_x. Carbon monoxide (CO) has been shown to correlate highly with primary organic aerosol concentrations during the SOAR study, and odd oxygen ($O_x = O_3 + NO_2$) has been shown to correlate with secondary organic aerosol concentrations (Docherty et al., 2011). The Δr values are all 0 at bin width 52 (meaning there are 52 scan points summed within each single bin) for the four pairs, which means that retention time correction and no correction produce the same correlations. For the correlations of the four pairs, the Δr
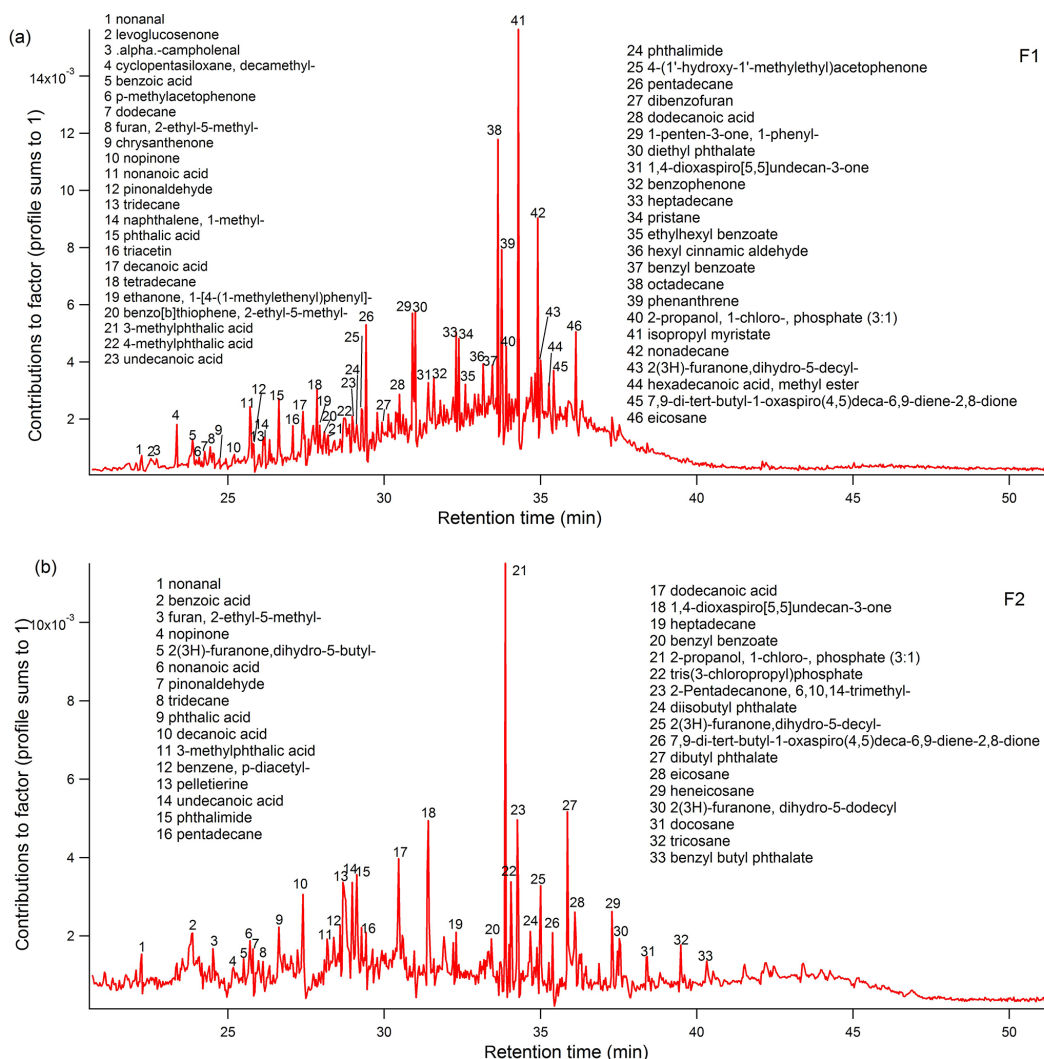
Figure 5.

shows little increase (0.0075 on average) as the bin width decreases from 52 to 13, whereas the Δr begins to increase (0.03 on average) from the bin width 13 to 2 and dramatically increases (0.46 on average) from the bin width 2 to 1. The reason Δr begins to increase at the bin width 13 is that the retention time shift among all samples is 13, which is shown in Fig. 3. TAG–Bin with the bin width larger than the total retention time shift is not sensitive to the retention time shift correction, whereas TAG–Bin with bin widths smaller than the total shift is sensitive to the correction, and bin width 1 (where each scan point is retained) is extremely sensitive to the retention time shift correction. In this case, without prior retention time shift correction, the user would certainly not want to use every MS scan point in a PMF analysis and would need to exceed bin widths of 13 scan points to minimize the impact of retention time shifts on PMF results.

Figure S3 shows the correlation coefficient r of the four pairs with retention time shift correction (all analyses be-

low are performed using r.t. shift correction). r increases only slightly (0.04 on average) as the bin width decreases from 52 to 1. The bin width being smaller, with correspondingly higher chemical resolution, does not increase r very much for this simple two-factor PMF solution (that requires limited chemical resolution). The bin's mass spectrum with different bin widths is an ensemble mass spectrum derived from the linear superposition of the different mass spectra of individual compounds. PMF, a multi-linear model, can deconvolve the ensemble mass spectra, such as AMS mass spectra and TAG bins' mass spectra, into groups of mass spectra, which provide chemical information on the sources and atmospheric aging processes. Thus, ideally, different bin widths, which affect chemical resolution, will not affect PMF performed to retrieve the factors' time series. The slight increase (0.04) of r for bin widths from 52 to 1 may be because of better PMF error estimation, and better PMF fits for small peaks will be obtained when a small bin width is used. This

explanation is supported by the fact that r increases more for TAG–Bin OOA (0.06 on average) than for TAG–Bin HOA (0.015 on average) as bin width decreases from 52 to 1. The compounds in the TAG–Bin OOA group have overall lower signal peaks than the compounds in the HOA group do and are better fit by small bin widths.

Theoretically, at least five scan points can define a peak; practically, more than 10 scan points are found in the compound peaks. Thus, for future users, a bin width of more than five scan points is recommended because the smaller bin width requires significant computational power and takes exponentially more time for PMF fitting. The retention time shift correction is strongly required when the total retention time shift among all samples is larger than the bin width.

3.4 Profiles of a six-factor solution to the TAG–Bin method

The six-factor PMF analysis with bin width equal to five scan points and retention time shift correction is applied to the TAG–Bin data matrix. According to the source apportionment of this TAG data set by Williams et al. (2010) using the integration method, the potential sources for TAG collected samples could be local vehicle, secondary organic aerosol, food cooking, biomass burning, etc. Here we do not relate these factors to sources again. But the major compounds (and UCM) in each factor's chromatogram are described here for later comparisons with TAG–Integrated and AMS results. As shown in Fig. 5, the first factor (F_1) contains the highest contributions from isopropyl myristate, alkanes (octadecane, nonadecane, pentadecane, eicosane, heptadecane, pristane), phenanthrene, diethyl phthalate, 1-phenyl-1-penten-3-one, 2-propanol,1-chloro-,phosphate (3 : 1), etc. The resulting chromatogram of factor two (F_2) is dominated by the compounds 2-propanol,1-chloro-,phosphate (3 :1˙), dibutyl phthalate, 1,4-dioxaspiro[5,5]undecan-3-one, 2-pentadecanone,6,10,14-trimethyl, carboxylic acids (dodecanoic acid, undecanoic acid, decanoic acid), tris(3-chloropropyl)phosphate, dihydro-5-decyl-2(3H)-furanone, pelletierine, phthalimide, etc. Factor three (F_3) contains major compounds: 2-propanol,1-chloro-,phosphate (3 : 1), alkanes (nonadecane, heneicosane, eicosane, heptadecane), phthalic acid, phthalimide, phenanthrene, triacetin, dibutyl phthalate, decanoic acid, p-diacetylbenzene, etc. Factor four (F_4) shows major resolved peaks mostly composed of oxygenated molecules: 2-pentadecanone,6,10,14-trimethyl, phthalates (diisobutyl phthalate, dibutyl phthalate), carboxylic acids (nonanoic acid, decanoic acid, dodecanoic acid, hexadecanoic acid, methyl ester), phthalic acids (phthalic acid, methyl phthalic acid), triacetin, furanones (dihydro-5-heptyl-2(3H)-furanone, dihydro-5-octyl-2(3H)-furanone, dihydro-5-undecyl-2(3H)-furanone), etc. Besides the resolved peaks, F_4 has a portion of UCM in the retention time range of 35–43 min, with the mass spectrum featuring m/z of 43, 55, 69, 81, 95, 109, etc. Factor five's (F_5) chro-

matogram has major compounds: 2-pentadecanone,6,10,14-trimethyl, alkanes (nonadecane, octadecane, eicosane, heneicosane), phenanthrene, benzyl benzoate, 2-propanol,1-chloro-,phosphate (3 : 1), 1,4-dioxaspiro[5,5]undecan-3-one, phthalates (diisobutyl phthalate, dibutyl phthalate, diethyl phthalate), dihydro-5-decyl-2(3H)-furanone, nonanoic acid, pinonaldehyde, pelletierine, nonanal, benzoic acid, etc. Factor six (F_6) shows major resolved peaks mostly composed of alkanes (C17–C29), isopropyl myristate, phenanthrene, etc. In addition, F_6 also contains a big portion of UCM in the retention time range of 41–51 min, with the mass spectrum dominated by m/z 43, 57, 69, 83, 97, 111, etc.

3.5 Binning method for source apportionment compared to previously developed method for chromatogram deconvolution

The chromatogram binning method has two applications for TAG data: chromatogram deconvolution described in detail in Zhang et al. (2014) and source apportionment presented here. The PMF factor chromatograms and time series of both the six- and 20-factor solutions are compared here. Figure S4 displays the six-factor chromatograms and mass spectral profiles for the chromatogram deconvolution method. The 20-factor chromatograms and mass spectral profiles for the chromatogram deconvolution method are presented in Zhang et al. (2014). In the six-factor solution for chromatogram deconvolution, two of the six factors are mainly resolved compounds (one is the alkane compound class; the other is mostly phthalic acid compound classes), and the others are predominantly composed of UCM. In the 20-factor solution for chromatogram deconvolution, more compound classes are separated as single factors – alkanes, carboxylic acids, furanones, phthalates, cylcyclohexanes, etc. – as well as several individual compound factors. In the binning method for source apportionment, the six-factor solution was previously described in Sect. 3.4. For the 20-factor solution, the compounds in each factor are marked in Fig. S5. Compared to the previous binning method for chromatogram deconvolution, this method for source apportionment tends to load many of the compounds into multiple factors since many of the compounds can be due to multiple sources. PMF factors resulting from the source apportionment method contain a greater diversity of compound types that correlate over sample time and represent a mixed chemical profile for specific source types or aerosol processes. The binning method for chromatogram deconvolution found major chromatogram components and individual factors were dominated by major compound classes with similar mass spectral features (e.g., alkanes series, acid series, phthalate series).

The six-factor time series of the binning method for source apportionment (Table 2) and chromatogram deconvolution (Table S1 in the Supplement) are compared to the six-factor time series of AMS PMF factors, which are considered as the source components. It is noted that due to the AMS in-
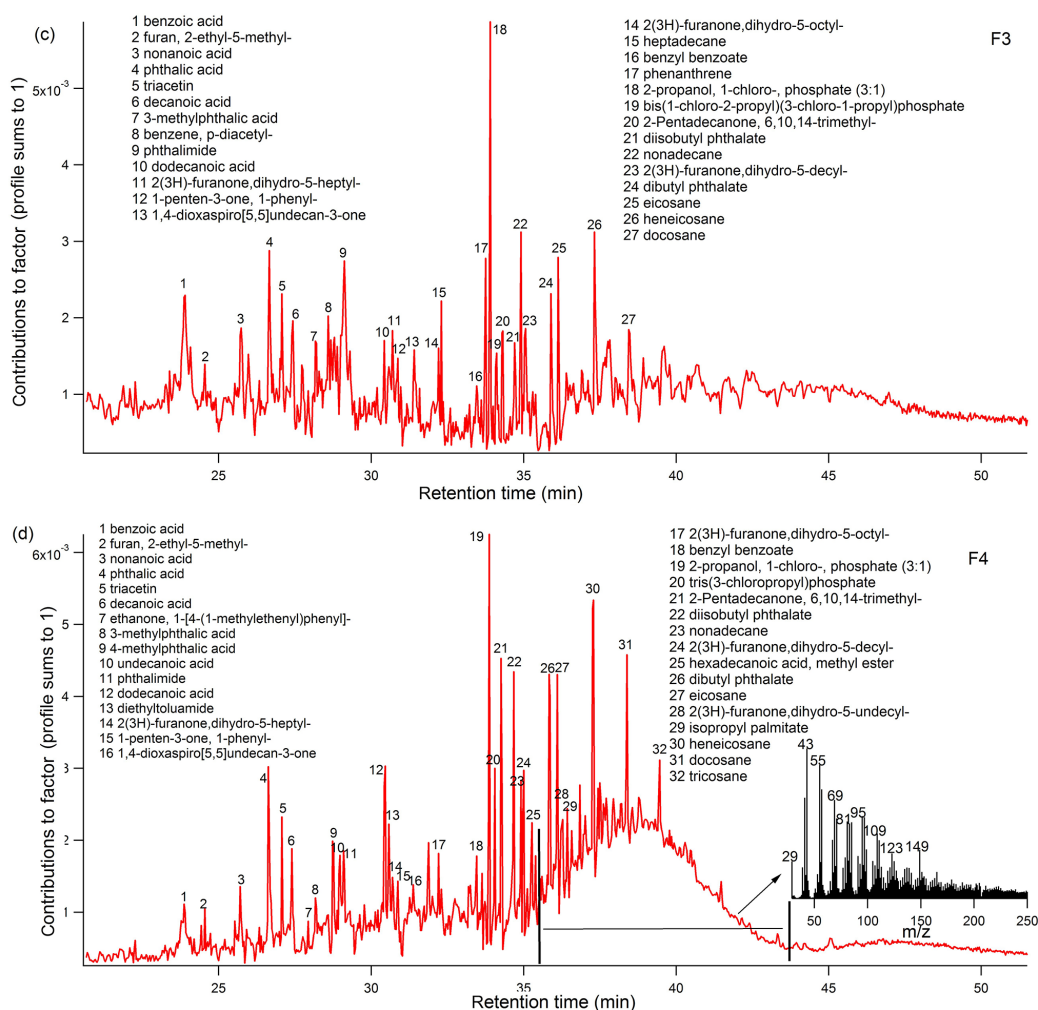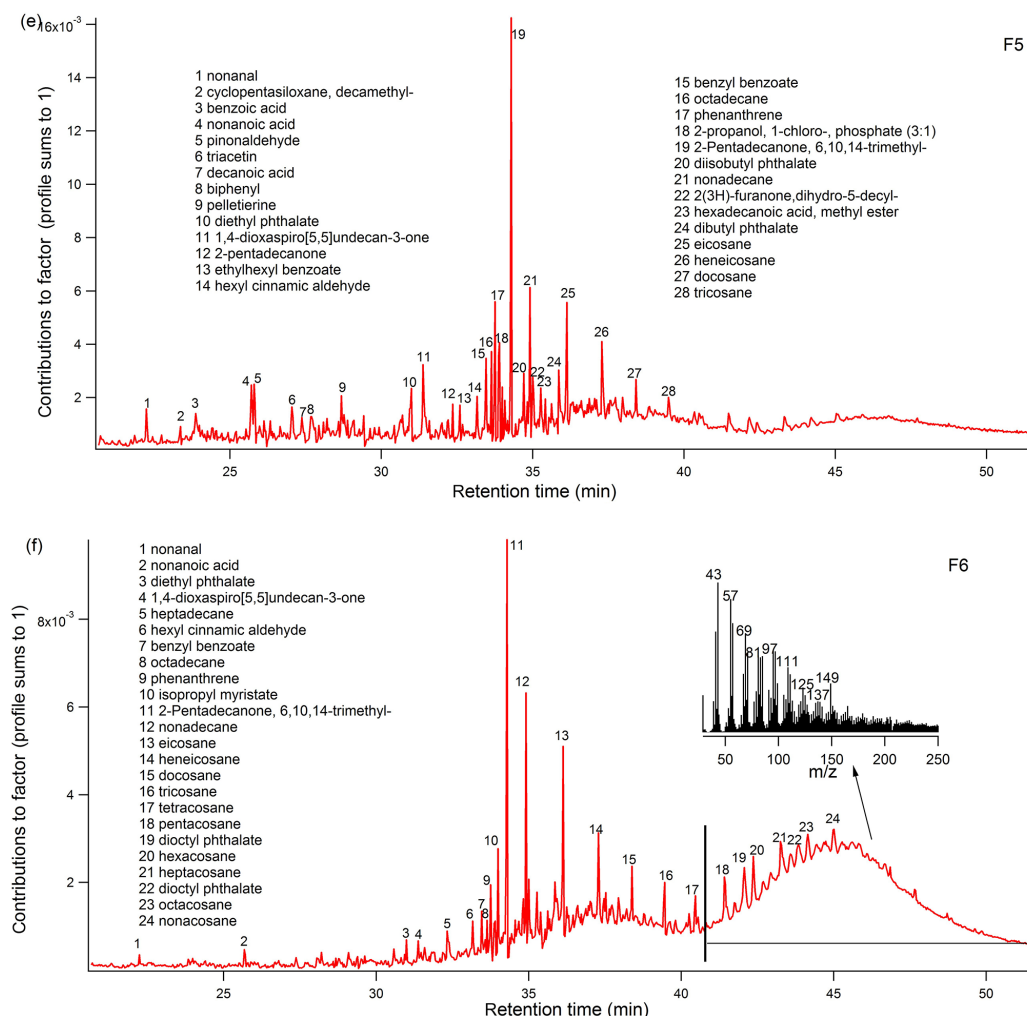
Figure 5.

Table 2. The correlation coefficient r of factors' time series of the PMF six-factor solution between AMS and TAG–Bin. TAG–Bin indicates the TAG binning method for source apportionment. A bin width of five scan points with retention time shift correction is used here. The italic font represents the maximum value in each column.

Pearson r		AMS six components					
		cLV–OOA	LOA–AC	LOA2	SV–OOA	MV–OOA	HOA
	F_1	.08	*.38*	−.32	.26	−.03	.31
TAG–	F_2	.45	.01	.11	-.09	.39	.12
Bin	F_3	−.15	.26	.11	.05	−.44	−.32
six factors	F_4	*.55*	−.26	−.26	−.41	*.87*	−.12
	F_5	−.51	−.20	*.20*	*.63*	−.39	−.05
	F_6	−.09	−.01	−.18	−.15	-.01	*.80*
Maximum		*.55*	*.38*	*.20*	*.63*	*.87*	*.80*

strument's quantitative ability we make comparisons to major components of OA as determined by AMS PMF. However, the AMS measurement offers limited chemical separation compared to the TAG system and is not capable of de-

termining many likely sources that contribute to atmospheric OA. While we utilize AMS PMF components here as an independent third method for comparison, it is likely possible that TAG measurements are capable of resolving additional

Figure 5. The chromatogram profiles of the TAG–Bin six-factor solution. A bin width of five scan points and retention time shift correction are used here. The mass spectra in panels **(d)** (factor four) and **(f)** (factor six) are the summed mass spectra of what are largely UCM components in the indicated retention time range.

sources or transformation processes compared to AMS. The maximum correlation coefficients (r) with AMS factors in each table (Tables 2 and S1) are summarized in Fig. S6. Most factors in the binning method for source apportionment display a better correlation with AMS factors than the factors in the chromatogram deconvolution method, indicating the binning method for source apportionment is superior for the purpose of source apportionment. Tables S2 and S3 show the correlation coefficients r of the AMS six components with the 20-factor solution of TAG–Bin for chromatogram deconvolution and source apportionment, respectively. While the binning method for chromatogram deconvolution displayed some high correlations with a few of the AMS PMF components, this method used a 20-factor PMF solution to more completely separate chemical components, and resulting individual compound classes can have a high correlation with AMS components similar to how individual marker compounds can have a high correlation with AMS components. Here (Table S3) it is observed that individual factors from the source apportionment method are more distinct and tend to correlate highest with a single AMS component, as opposed to correlating with multiple components as was observed in the chromatogram deconvolution method (Table S2). Additionally, there are some factors that correlate even with the minor AMS components (e.g., LOA–AC, LOA2, SV–OOA) when using the source apportionment method. To do an independent separation of major sources using only TAG data, this new binning method for source apportionment must be applied.
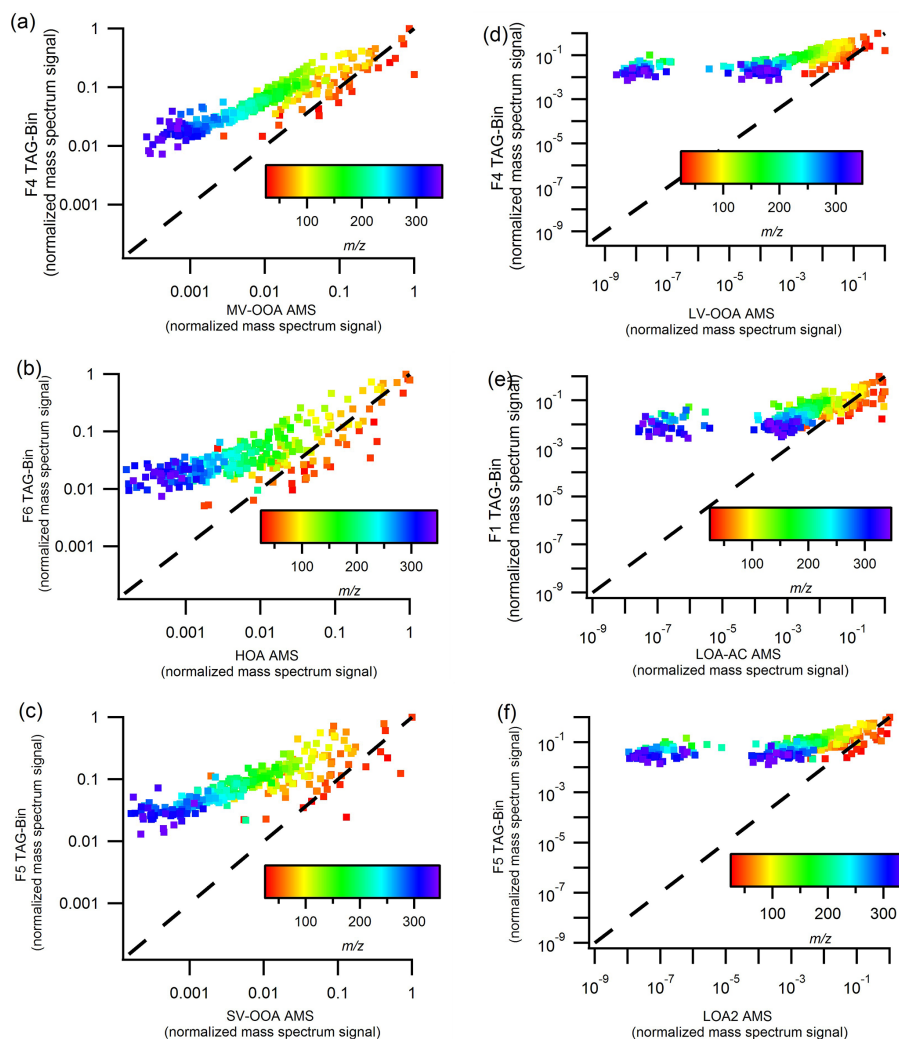
Figure 6. The mass spectra comparison of the six pairs of TAG–Bin (assigned by factor number) vs. AMS – F_4 vs. MV–OOA, F_6 vs. HOA, F_5 vs. SV–OOA, F_4 vs. LV–OOA, F_1 vs. LOA–AC, and F_5 vs. LOA2. Each pair shows the maximum Pearson r in Table 2. The mass spectra (m/z 29–343 is shown in the color scale) is the normalized signal in log scale.

3.6 Source factor comparisons between AMS and the TAG binning method for source apportionment

Table 2 shows the Pearson correlation coefficient r of six-factor PMF time series between AMS and the TAG–Bin source apportionment method. Three pairs of TAG–Bin (assigned by factor number) and AMS factors display good correlations: $r = 0.87$ for F_4 vs. MV–OOA, $r = 0.80$ for F_6 vs. HOA, and $r = 0.63$ for F_5 vs. SV–OOA. Figures 6a–c and S7a–c show the mass spectra comparison of those three pairs in two different ways. For all of the three pairs, TAG–Bin vs. AMS for $m/z < 100$ follows the line $y = x$, whereas for $m/z > 100$ it is above the line $y = x$. The patterns of mass spectra for $m/z < 100$ are similar between TAG–Bin and AMS. For $m/z > 100$, TAG–Bin is much higher than AMS (also suggested in Fig. S7). The TAG system has been reported to have higher contributions from larger fragments

when compared to AMS mass spectra, likely due to lower temperature of molecules during evaporation and fragmentation (Williams et al., 2014). Here, F_4 (paired with MV–OOA) is mostly composed of oxygenated compounds – carboxylic acids, phthalic acids, triacetin, furanones, etc. Besides the resolved compounds, F_4 also contains a portion of UCM with a similar mass spectrum to AMS MV–OOV (see Figs. 5d and 7d). F_6 (paired with HOA) contains a suite of alkanes (C_{17}–C_{29}) as well as a portion of UCM with a similar mass spectrum to AMS HOA (see Figs. 5f and 7e). Finally, F_5 (paired with SV–OOA) contains a large number of semi-volatile compounds: nonanoic acid, pinonaldehyde, pelletierine, nonanal, benzoic acid, etc.
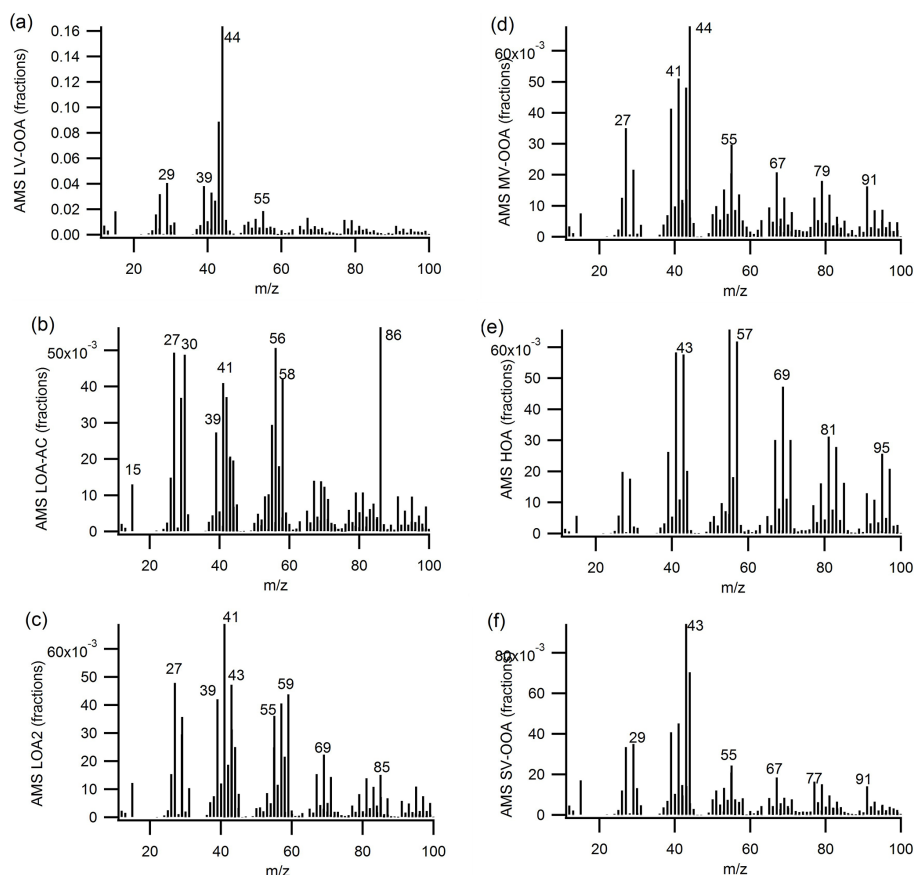
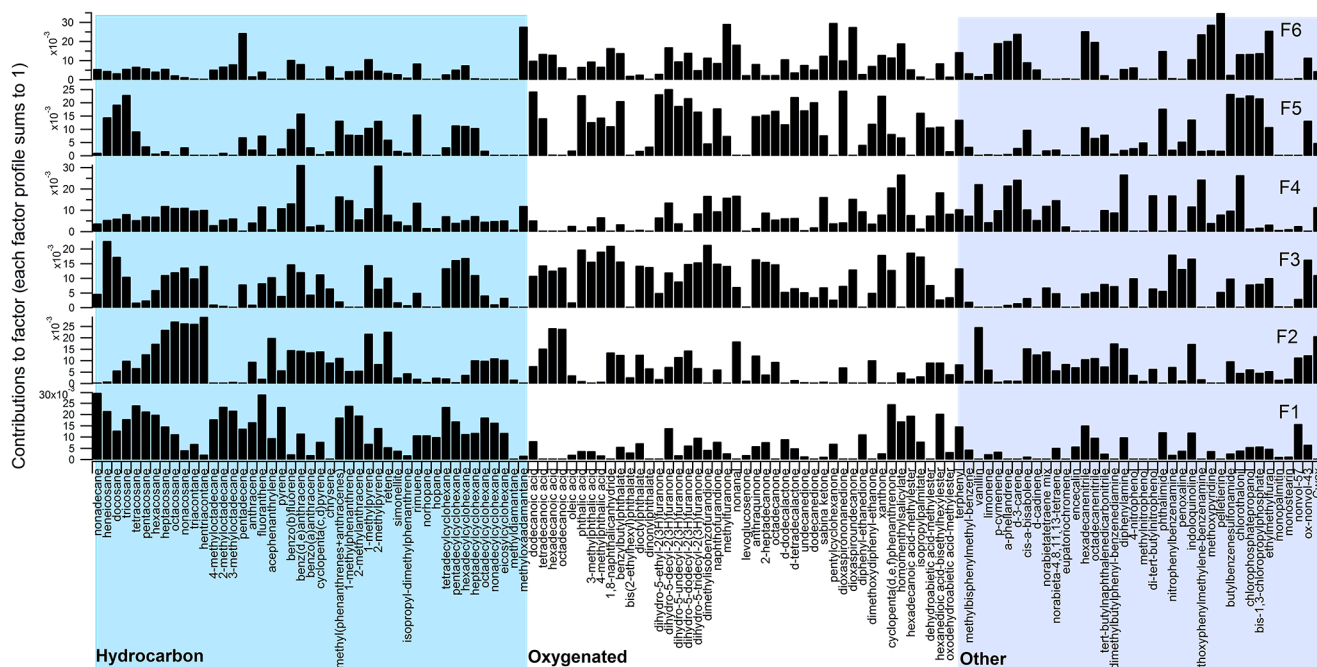**Figure 7.** The mass spectra of AMS six components.

Figure 8. Profiles of the TAG–Integrated six-factor solution.

Table 3. The correlation coefficient r of factors' time series of the PMF six-factor solution between AMS and TAG–Integrated. The TAG–Integrated indicates the TAG integration method. The italic font represents the maximum value in each column.

Pearson r		AMS six components					
		cLV–OOA	LOA–AC	LOA2	SV–OOA	MV–OOA	HOA
	F_1	.07	.11	$-.17$	$-.32$	$-.01$	*.71*
TAG–	F_2	$-.27$	$-.03$	.08	.19	$-.21$	.53
Integrated	F_3	.54	$-.22$	$-.26$	$-.31$	*.78*	.06
six factors	F_4	$-.59$	$-.27$	.05	*.63*	$-.36$	$-.02$
	F_5	.46	$-.13$	$-.03$	$-.39$	.59	$-.58$
	F_6	$-.03$	*.45*	*.18*	.49	$-.44$	.21
Maximum		*.54*	*.45*	*.18*	*.63*	*.78*	*.71*
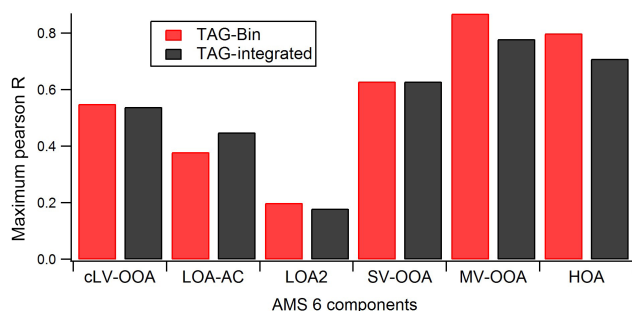
Figure 9. Profiles of the TAG–Integrated six-factor solution.

3.7 PMF profiles from the six-factor solution of the TAG–Integrated method

The six-factor PMF is applied to the more traditional TAG–Integrated method. The factor number is assigned to each factor. Figure 8 shows the chemical profiles of the TAG–Integrated six-factor solution. F_1 is mostly composed of the hydrocarbons – alkanes, polycyclic aromatic hydrocarbons (PAHs), cyclohexanes, etc. F_2 is dominated by larger alkanes. F_3 is featured in the oxygenated compounds - carboxylic acids, phthalic acids, furanones, etc. F_4 has the major compounds – terpenes, xanthone, cyclopenta(d,e,f)phenanthrenone, N-(1,3-dimethylbutyl)-N'-phenyl-1,4-benzenediamine, etc. F_5 is dominated by the oxygenated compounds – phthalic acids, furanones, ketones, sulfur-chlorine-phosphorus-containing compounds, etc. F_6 has high loadings of the nitrogen-containing compounds, furanones, and ketones.

3.8 Source factor comparisons between TAG binning and integration methods

Just as Table 2 had listed the Pearson correlation coefficient r of six-factor PMF time series between AMS and the TAG–Bin source apportionment method, Table 3 shows the r of six-factor time series between AMS and TAG–Integrated (conventional compound analysis). Factor numbers are as-

signed by each PMF analysis, and factor numbers will not be reported in the same order for different PMF methods. Also, it is not expected that PMF results from different instruments (AMS vs. TAG), and different input data matrices from the same instrument (TAG–Bin vs. TAG–Integrated) will divide covarying factors identically. The maximum r (in the column dimension of tables) with AMS factors in each table is displayed in Fig. 9 for the purpose of comparing TAG–Bin source apportionment and TAG–Integrated results. For both comparisons in Tables 2 and 3, four (LOA–AC, SV–OOA, MV–OOA, and HOA) of the six maximum r in the column dimension are also the maximum r in the row dimension (AMS factor's maximum r with each TAG factor). Similar maximum correlation r for TAG–Bin and TAG–Integrated indicates that TAG–Bin source apportionment shows similar performance to TAG–Integrated. The maximum r pairs with AMS MV–OOA are TAG–Bin F_4 and TAG–Integrated F_3, which share many of the same compounds – carboxylic acids, phthalic acids, and furanones. The maximum r pairs with AMS HOA are TAG–Bin F_6 and TAG–Integrated F_1. They also present many of the same compounds – alkanes and PAHs. In addition, TAG–Bin F_6 has better r with AMS HOA than TAG–Integrated F_1, and TAG–Bin F_4 has better r with MV–OOA than TAG–Integrated F_3. The improved correlation ($\Delta r = 0.09$ on average) is because TAG–Bin F_6 and F_4 have a portion of UCM, with the mass spectra similar to the AMS HOA and MV–OOA, respectively. In addition, TAG–Bin and TAG–Integrated factors have good correlations ($r > 0.6$) with AMS MV–OOA, SV–OOA, and HOA factors, suggesting that the TAG system as operated during the SOAR study was good at measuring components which are related to MV–OOA, SV–OOA, and HOA. Furthermore, TAG–Bin and TAG–Integrated factors have lower correlations ($r < 0.5$) with AMS LOA–AC and LOA2 factors. LOA–AC and LOA2 factors account only for 7 % of total AMS mass, and it is expected that PMF results from different instruments, such as TAG and AMS, would produce lower correlations for minor factors such as these. The maximum r pairs with AMS LOA–AC are TAG–Bin F_1 and TAG–Integrated F_6. The TAG–Integrated F_6 has better r than

Pearson r		TAG–Bin six factors					
		F_1	F_2	F_3	F_4	F_5	F_6
	F_1	.2	−.05	−.21	−.02	−.18	*.88*
TAG–	F_2	0	.23	−.51	−.17	.35	.57
Integrated	F_3	.01	.36	−.51	.76	−.4	.2
six factors	F_4	−.08	−.43	*.07*	−.5	*.84*	.05
	F_5	−.02	.06	.06	*.68*	−.46	−.44
	F_6	*.33*	*.48*	.01	−.44	.17	−.12
Maximum		*.33*	*.48*	*.07*	*.68*	*.84*	*.88*

TAG–Bin F_1. Many nitrogen-containing compounds, which are highlighted in the TAG–Integrated method using the normalized abundance as the PMF input by each compound's maximum raw signal, are loaded into TAG–Integrated F_6. However, those compounds that have low absolute signals in the raw chromatogram are buried in the chromatogram profiles of the TAG–Bin method, which uses the raw signal as the PMF input. For the AMS LV–OOA factor, TAG–Bin and TAG–Integrated factors display mid-range correlations ($0.5 < r < 0.6$), as many compounds in the LV–OOA category likely either undergo thermal decomposition or do not transfer through the 30 m TAG separation column. To address low detection of this analytically challenging OA fraction, subsequent TAG field deployments have applied a range of methods to increase detection through online derivatization in some cases (Isaacman et al., 2014), or shorter GC column lengths to enhance recovery of oxygenated material (Martinez et al., 2016); in other cases the thermal decomposition products from heating of low-volatility and highly functionalized OA have been detected and analyzed (Williams et al., 2016).

Table 4 shows the direct comparison between TAG–Bin and TAG–Integrated, without any reference to AMS PMF results. Three pairs of TAG–Bin and TAG–Integrated have good correlations ($r > 0.75$). These three pairs are associated with MV–OOA, SV–OOA, and HOA (AMS factors) as suggested in Tables 2 and 3. The TAG system during the SOAR field study was good at measuring species in MV–OOA, SV–OOA, and HOA categories, which have high relative abundance and mid- to low polarity. There is only one pair with a low correlation ($r = 0.07$). Two potential reasons for this are presented here. Firstly, TAG–Bin used the whole chromatogram signal as PMF input, whereas TAG–Integrated only used the 123 resolved compound signals as input. Different signal (mass) input may affect how PMF resolves factors. Secondly, as mentioned above, small peaks present in

chromatograms are amplified in the TAG–Integrated method, whereas the signal and variability present in these small peaks will be buried in the large, comprehensive signal that is utilized by the TAG–Bin method. This could also produce different factor solutions between TAG–Bin and TAG–Integrated.

Although the binning and integration methods have similar performance, the binning method requires limited data pre-processing and incorporates the entirety of the data set, allowing for a comprehensive and rapid method for utilizing chromatographically separated mass spectral data in factor analyses for the purpose of organic aerosol source identification. By incorporating all of the GC–MS signal, the binning method does not risk missing an important compound or series of compounds as could easily occur for the traditional single-compound method since the input compounds are chosen by the operator. In terms of chemical resolution, the binning method can get down to the molecular level with appropriate retention time shifting. In this case, the operator can then go back and identify important compounds within each of the factors, after PMF analysis has decided they are defining species for a resulting component.

When this binning method for source apportionment is applied to future ambient data sets, the user will need to determine the appropriate number of PMF factors to choose for a solution. Each data set will be different, and ultimately the operator will need to use his or her own discretion by utilizing all information available. In general, too few factors will combine sources or transformation processes that share either chemical profile similarities or temporal similarities, and too many factors will begin to cause "factor splitting", where what should have been a single component is divided into multiple components based on very minor differences. The original TAG compound integration PMF solution published by Williams et al. (2010) found that nine factors best described the analyzed data set. Here we presented a six-factor solution and a 20-factor solution to TAG PMF analysis using the binning method for source apportionment. This is an appropriate range to explore for urban, suburban, to rural locations where you would expect at least six major OA source contributors or atmospheric transformation processes that would alter chemical profiles. Urban locations may contain 20 or more contributing sources; however with that many factors it is likely that PMF would begin to cause factor splitting of major sources before separating some of the minor contributing sources. Previous AMS PMF analyses have used higher-factor solutions to separate minor contributing sources and then manually recombined major factors that had been split by the high-factor solution (e.g., Docherty et al., 2011). This is an option for TAG PMF analyses as well, and, given the enhanced chemical resolution of the TAG, additional contributing sources are expected to be identified.

4 Conclusions and implications

In the chromatogram binning method for source apportionment, the whole chromatogram was divided into evenly spaced bins, within which mass spectra were summed to form a bin's mass spectrum. PMF was applied to separate the sources according to their covariance. The row dimension of the PMF binning data matrix is the sample number, and the column dimension is mass spectra eluting time bins. The retention time shift with respect to the first sample was investigated in both the retention time and the sample number dimensions. The median value of the retention time shifts in each sample is used to correct the major shifts of the chromatographic peaks. The effects of different bin widths, with and without retention time shift correction, were compared. When the bin width was smaller than the retention time shift among all samples, the retention time shift correction was required. Bin width, which affects chemical resolution, does not affect the PMF retrieval of the factors' time series for low-factor simple solutions. In multiple-source comparisons, the binning method had similar performance to the conventional compound integration method, but the binning method incorporates the entirety of the data set, can be fully automated, and requires limited data pre-processing prior to PMF analyses.

Future applications of this method should continue to apply retention time shifts when necessary and should incorporate new relationships using regularly injected calibration standards to develop bin-specific response factors, especially when longer study periods susceptible to larger drifts are to be analyzed. In the future, it will be of great interest to investigate if these TAG PMF components can provide additional factor/source resolution as compared to the bulk components currently derived by AMS PMF. Two binning methods, for chromatogram structure (Zhang et al., 2014) or study time structure (the source apportionment method presented here), have now been shown to operate well for the TAG GC–MS data, and the approach should be of interest for any measurement technique (mass spectrometry or spectroscopy) with an additional separation dimension(s) (volatility, hygroscopicity, electrical mobility, etc.).

5 Data availability

The data used in this study is available from authors upon request.

The Supplement related to this article is available online at doi:10.5194/amt-9-5637-2016-supplement.

References

Allan, J. D., Jimenez, J. L., Williams, P. I., Alfarra, M. R., Bower, K. N., Jayne, J. T., Coe, H., and Worsnop, D. R.: Quantitative sampling using an Aerodyne aerosol mass spectrometer – 1. Techniques of data interpretation and error analysis, J. Geophys. Res.-Atmos., 108, 4090, doi:10.1029/2002jd002358, 2003.

Ahlm, L., Shakya, K. M., Russell, L. M., Schroder, J. C., Wong, J. P. S., S. J., Hayden, K. L., Liggio, J., Wentzell, J. J. B., Wiebe, H. A., Mihele, C., Leaitch, W. R., and Macdonald, A. M.: Temperature-dependent accumulation mode particle and cloud nuclei concentrations from biogenic sources during WACS 2010, Atmos. Chem. Phys., 13, 3393–3407, doi:10.5194/acp-13-3393-2013, 2013.

Canagaratna, M. R., Jayne, J. T., Jimenez, J. L., Allan, J. D., Alfarra, M. R., Zhang, Q., Onasch, T. B., Drewnick, F., Coe, H., Middlebrook, A., Delia, A., Williams, L. R., Trimborn, A. M., Northway, M. J., DeCarlo, P. F., Kolb, C. E., Davidovits, P., and Worsnop, D. R.: Chemical and microphysical characterization of ambient aerosols with the aerodyne aerosol mass spectrometer, Mass Spectrom. Rev., 26, 185–222, doi:10.1002/mas.20115, 2007.

Chueinta, W., Hopke, P. K., and Paatero, P.: Investigation of sources of atmospheric aerosol at urban and suburban residential areas in Thailand by positive matrix factorization, Atmos. Environ., 34, 3319–3329, 2000.

Docherty, K. S., Stone, E. A., Ulbrich, I. M., DeCarlo, P. F., Snyder, D. C., Schauer, J. J., Peltier, R. E., Weber, R. J., Murphy, S. M., Seinfeld, J. H., Grover, B. D., Eatough, D. J., and Jimenez, J. L.: Apportionment of Primary and Secondary Organic Aerosols in Southern California during the 2005 Study of Organic Aerosols in Riverside (SOAR-1), Environ. Sci. Technol., 42, 7655–7662, doi:10.1021/es8008166, 2008.

Docherty, K. S., Aiken, A. C., Huffman, J. A., Ulbrich, I. M., DeCarlo, P. F., Sueper, D., Worsnop, D. R., Snyder, D. C., Peltier, R. E., Weber, R. J., Grover, B. D., Eatough, D. J., Williams, B. J., Goldstein, A. H., Ziemann, P. J., and Jimenez, J. L.: The 2005 Study of Organic Aerosols at Riverside (SOAR-1): instrumental intercomparisons and fine particle composition, Atmos. Chem. Phys., 11, 12387–12420, doi:10.5194/acp-11-12387-2011, 2011.

Dominici, F., Peng, R. D., Bell, M. L., Pham, L., McDermott, A., Zeger, S. L., and Samet, J. M.: Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases, JAMA-J. Am. Med. Assoc., 295, 1127–1134, doi:10.1001/jama.295.10.1127, 2006.

Gauderman, W. J., Urman, R., Avol, E., Berhane, K., Mc-Connell, R., Rappaport, E., Chang, R., Lurmann, F., and Gilliland, F.: Association of Improved Air Quality with Lung Development in Children, New Engl. J. Med., 372, 905–913, doi:10.1056/NEJMoa1414123, 2015.

Goldstein, A. H. and Galbally, I. E.: Known and unexplored organic constituents in the earth's atmosphere, Environ. Sci. Technol., 41, 1514–1521, doi:10.1021/es072476p, 2007.

Goldstein, A. H., Worton, D. R., Williams, B. J., Hering, S. V., Kreisberg, N. M., Panic, O., and Gorecki, T.: Thermal desorption comprehensive two-dimensional gas chromatography for in-situ measurements of organic aerosols, J. Chromatogr. A, 1186, 340–347, doi:10.1016/j.chroma.2007.09.094, 2008.

Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N. M., George, C., Goldstein, A. H., Hamilton, J. F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M. E., Jimenez, J. L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, Th. F., Monod, A., Prévôt, A. S. H., Seinfeld, J. H., Surratt, J. D., Szmigielski, R., and Wildt, J.: The formation, properties and impact of secondary organic aerosol: current and emerging issues, Atmos. Chem. Phys., 9, 5155–5236, doi:10.5194/acp-9-5155-2009, 2009.

IPCC: Climate Change 2013: The Physical Science Basis, 2013.

Isaacman, G., Kreisberg, N. M., Worton, D. R., Hering, S. V., and Goldstein, A. H.: A versatile and reproducible automatic injection system for liquid standard introduction: application to in-situ calibration, Atmos. Meas. Tech., 4, 1937–1942, doi:10.5194/amt-4-1937-2011, 2011.

Isaacman, G., Kreisberg, N. M., Yee, L. D., Worton, D. R., Chan, A. W. H., Moss, J. A., Hering, S. V., and Goldstein, A. H.: Online derivatization for hourly measurements of gas- and particle-phase semi-volatile oxygenated organic compounds by thermal desorption aerosol gas chromatography (SV-TAG), Atmos. Meas. Tech., 7, 4417–4429, doi:10.5194/amt-7-4417-2014, 2014.

Ito, K., Xue, N., and Thurston, G.: Spatial variation of PM$_{2.5}$ chemical species and source-apportioned mass concentrations in New York City, Atmos. Environ., 38, 5269–5282, 2004.

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

Junjun, D., Zhenyu, X., Bingliang, Z., and Ke, D.: Comparative study on long-term visibility trend and its affecting factors on both sides of the Taiwan Strait, Atmos. Res., 143, 266–278, doi:10.1016/j.atmosres.2014.02.018, 2014.

Jaeckels, J. M., Bae, M. S., and Schauer, J. J.: Positive matrix factorization (PMF) analysis of molecular marker measurements to quantify the sources of organic aerosols, Environ. Sci. Technol., 41, 5763–5769, 2007.

Lee, E., Chan, C. K., and Paatero, P.: Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong, Atmos. Environ., 33, 3201–3212, 1999.

Martinez, R. E., Williams, B. J., Zhang, Y., Hagan, D., Walker, M., Kreisberg, N. M., Hering, S. V., Hohaus, T., Jayne, J. T., and Worsnop, D. R.: Development of a volatility and polarity separator (VAPS) for volatility- and polarity-resolved organic aerosol measurement, Aerosol Sci. Technol., 50, 255–271, 2016.

Murphy, D. M., Cziczo, D. J., Froyd, K. D., Hudson, P. K., Matthew, B. M., Middlebrook, A. M., Peltier, R. E., Sullivan, A., Thomson, D. S., and Weber, R. J.: Single-particle mass spectrometry of tropospheric aerosol particles, J. Geophys. Res.-Atmos., 111, D23s32, doi:10.1029/2006jd007340, 2006.

Ng, N. L., Herndon, S. C., Trimborn, A., Canagaratna, M. R., Croteau, P. L., Onasch, T. B., Sueper, D., Worsnop, D. R., Zhang, Q., Sun, Y. L., and Jayne, J. T.: An Aerosol Chemical Speciation Monitor (ACSM) for Routine Monitoring of the Composition and Mass Concentrations of Ambient Aerosol, Aerosol Sci. Technol., 45, 780–794, doi:10.1080/02786826.2011.560211, 2011.

Paatero, P.: Least squares formulation of robust non-negative factor analysis, Chemometr. Intell. Lab., 37, 23–35, doi:10.1016/s0169-7439(96)00044-5, 1997.

Ramadan, Z., Song, X. H., and Hopke, P. K.: Identification of sources of Phoenix aerosol by positive matrix factorization, J. Air Waste Manage. Assoc., 50, 1308–1320, 2000.

Schauer, J. J., Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Simoneit, B. R. T.: Source apportionment of airborne particulate matter using organic compounds as tracers, Atmos. Environ., 30, 3837–3855, 1996.

Sun, Y. L., Zhuang, G. S., Tang, A. H., Wang, Y., and An, Z. S.: Chemical characteristics of PM$_{2.5}$ and PM$_{10}$ in haze-fog episodes in Beijing, Environ. Sci. Technol., 40, 3148–3155, doi:10.1021/es051533g, 2006.

Ulbrich, I. M., Canagaratna, M. R., Zhang, Q., Worsnop, D. R., and Jimenez, J. L.: Interpretation of organic components from Positive Matrix Factorization of aerosol mass spectrometric data, Atmos. Chem. Phys., 9, 2891–2918, doi:10.5194/acp-9-2891-2009, 2009.

Williams, B. J., Goldstein, A. H., Kreisberg, N. M., and Hering, S. V.: An in-situ instrument for speciated organic composition of atmospheric aerosols: Thermal Desorption Aerosol GC / MS-FID (TAG), Aerosol Sci. Technol., 40, 627–638, doi:10.1080/02786820600754631, 2006.

Williams, B. J., Goldstein, A. H., Kreisberg, N. M., Hering, S. V., Worsnop, D. R., Ulbrich, I. M., Docherty, K. S., and Jimenez, J. L.: Major components of atmospheric organic aerosol in southern California as determined by hourly measurements of source marker compounds, Atmos. Chem. Phys., 10, 11577–11603, doi:10.5194/acp-10-11577-2010, 2010.

Williams, B. J., Jayne, J. T., Lambe, A. T., Hohaus, T., Kimmel, J. R., Sueper, D., Brooks, W., Williams, L. R., Trimborn, A. M., Martinez, R. E., Hayes, P. L., Jimenez, J.

L., Kreisberg, N. M., Hering, S. V., Worton, D. R., Goldstein, A. H., and Worsnop, D. R.: The First Combined Thermal Desorption Aerosol Gas Chromatograph–Aerosol Mass Spectrometer (TAG-AMS), Aerosol Sci. Technol., 358-370, doi:10.1080/02786826.2013.875114, 2014.

Williams, B. J., Zhang, Y., Zuo, X., Martinez, R. E., Walker, M. J., Kreisberg, N. M., Goldstein, A. H., Docherty, K. S., and Jimenez, J. L.: Organic and inorganic decomposition products from the thermal desorption of atmospheric particles, Atmos. Meas. Tech., 9, 1569–1586, doi:10.5194/amt-9-1569-2016, 2016.

Zenkevich, I. G. and Pavlovskii, A. A.: Overloading control of gas chromatographic systems, J. Sep. Sci., 38, 2848–2856, doi:10.1002/jssc.201401471, 2015.

Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Ulbrich, I. M., Ng, N. L., Worsnop, D. R., and Sun, Y. L.: 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.

Zhang, Y., Williams, B. J., Goldstein, A. H., Docherty, K., Ulbrich, I. M., and Jimenez, J. L.: A Technique for Rapid Gas Chromatography Analysis Applied to Ambient Organic Aerosol Measurements from the Thermal Desorption Aerosol Gas Chromatograph (TAG), Aerosol Sci. Technol. 48, 1166–1182, doi:10.1080/02786826.2014.967832, 2014.

Zhao, Y. L., Kreisberg, N. M., Worton, D. R., Teng, A. P., Hering, S. V., and Goldstein, A. H.: Development of an In Situ Thermal Desorption Gas Chromatography Instrument for Quantifying Atmospheric Semi-Volatile Organic Compounds, Aerosol Sci. Technol., 47, 258–266, doi:10.1080/02786826.2012.747673, 2013.