New insights into $PM_{2.5}$ chemical composition and sources in two major cities in China during extreme haze events using aerosol mass spectrometry

M. Elser[1], R.-J. Huang[1,2], R. Wolf[1], J.G. Slowik[1], Q.-Y. Wang[2], F. Canonaco[1], G. H. Li[2], C. Bozzetti, K.R. Daellenbach, Y. Huang[2], R.-J. Zhang[3], Z.-Q. Li[4], J.J. Cao[2], U. Baltensperger[1], I. El-Haddad[1] and A.S.H. Prévôt[1,2]

[1]{Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, 5232, Villigen PSI, Switzerland}

[2]{State Key Laboratory of Loess and Quaternary Geology and Key Laboratory of Aerosol Chemistry and Physics, Institute of Earth Environment, Chinese Academy of Sciences, 710075, Xi'an, China}

[3]{Key Laboratory of Regional Climate-Environment Research for Temperate East Asia, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China}

[4]{State Environmental Protection Key Laboratory of Satellite Remote Sensing, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China}

Correspondence to: R.-J. Huang (rujin.huang@ieecas.cn), I. El-Haddad (imad.el-haddad@psi.ch) and A. S. H. Prévôt (andre.prevot@psi.ch)

Abstract

During winter 2013-2014 aerosol mass spectrometer (AMS) measurements were conducted for the first time with a novel $PM_{2.5}$ (particulate matter with aerodynamic diameter ≤ 2.5 µm) lens in two major cities of China: Xi'an and Beijing. We denote the periods with visibility below 2 km as extreme haze and refer to the rest as reference periods. During the measurements in Xi'an an extreme haze covered the city for about a week and the total non-refractory (NR)-$PM_{2.5}$ mass fraction reached peak concentrations of over 1000 µg m^{-3}. During the measurements in Beijing two extreme haze events occurred, but the temporal extent and the total concentrations reached during these events were lower than in Xi'an. Average $PM_{2.5}$ concentrations of 537 ± 146 µg m^{-3} and 243 ± 47 µg m^{-3} (including NR species and equivalent black carbon (eBC)) were recorded during the extreme haze events in Xi'an and Beijing, respectively. During the reference periods the measured average concentrations were 140 ± 99 μg m^{-3} in Xi'an and 75 ± 61 μg m^{-3} in Beijing. The relative composition of the NR-PM$_{2.5}$ evolved substantially during the extreme haze periods, with increased contributions of the inorganic components (mostly sulfate and nitrate). Our results suggest that the high relative humidity present during the extreme haze events had a strong effect on the increase of sulfate mass (via aqueous phase oxidation of sulfur dioxide). Another relevant characteristic of the extreme haze is the size of the measured particles. During the extreme haze events, the AMS showed much larger particles, with a volume weighted mode at about 800 to 1000 nm, in contrast to about 400 nm during reference periods. These large particle sizes made the use of the PM$_{2.5}$ inlet crucial, especially during the severe haze events, where 39 ± 5 % of the mass would have been lost in the conventional PM$_1$ (particulate matter with aerodynamic diameter $\leq$ 1 μm) inlet. A novel positive matrix factorization procedure was developed to apportion the sources of organic aerosols (OA) based on their mass spectra using the multilinear engine (ME-2) controlled via the source finder (SoFi). The procedure allows an effective exploration of the solution space, a more objective selection of the best solution and an estimation of the rotational uncertainties. Our results clearly show an increase of the oxygenated organic aerosol (OOA) mass during extreme haze events. The contribution of OOA to the total OA increased from the reference to the extreme haze periods from 16.2 ± 1.1 % to 31.3 ± 1.5 % in Xi'an and from 15.7 ± 0.7 % to 25.0 ± 1.2 % in Beijing. By contrast, during the reference periods the total OA mass was dominated by domestic emissions of primary aerosols from biomass burning in Xi'an (42.2 ± 1.5 % of OA) and coal combustion in Beijing (55.2 ± 1.6 % of OA). These two sources are also mostly responsible for extremely high polycyclic aromatic hydrocarbon (PAH) concentrations measured with the AMS (campaign average of 2.1 ± 2.0 μg m^{-3} and frequent peak concentrations above 10 μg m^{-3}). To the best of our knowledge, this is the first dataset where the simultaneous extraction of these two primary sources could be achieved in China by conducting on-line AMS measurements at two areas with contrasted emission patterns.

1 Introduction

China, the fastest developing country in the history of the world, has been facing severe pollution problems in the last decades in response to rapid industrialization and economic growth. These problems include soil, water and air contamination. In terms of air pollution, large parts of the country are frequently affected by heavy smog events, causing widespread environmental and health issues. A recent study that investigated the link between premature mortality and several emission sources in urban and suburban environments, estimated 1.3 million premature deaths in China in 2010 due to outdoor air pollution (Lelieveld et al., 2015).

Atmospheric $PM_{2.5}$ (particulate matter with aerodynamic diameter ≤ 2.5 μm) affects climate, visibility and human health. The $PM_{2.5}$ fraction is widely used as an air quality metric, as long-term exposure to this fraction has been linked to increased lung cancer rates (Hu and Jiang, 2014), acute bronchitis and asthma (J. J. Gao et al., 2015). Moreover, some of the known $PM_{2.5}$ combustion sources in China have been shown to dominate emissions of carcinogenic species, including polycyclic aromatic hydrocarbons (PAH) (Xu et al., 2006; Zhang and Tao, 2009; Huang et al., 2014; Wei et al., 2015).

In China, severe pollution events often occur during wintertime, when stagnant meteorological conditions confine the gas- and particle-phase pollutants at the ground level. The particles can either be directly emitted as primary aerosols (e.g. particles emitted from combustion sources) or formed in the atmosphere by condensation of oxidation products of sulfur dioxide, nitrogen oxides and volatile organic compounds (secondary aerosol).

The first step for developing air pollution control strategies requires the identification of the major sources and processes producing airborne particles. Most previous aerosol studies in the areas of Xi'an and Beijing, two major Chinese cities, are based on filter measurements (Cao et al., 2012; Wang et al., 2013; Huang et al., 2014; Ho et al., 2015; M. Gao et al., 2015; Wang et al., 2015a; Xu et al., 2015; Yang et al., 2015). Carbonaceous materials, water-soluble ions (e.g. sulfate (SO_4^{2-}), nitrate (NO_3^-) and ammonium (NH_4^+)) and mineral dust have been found to be major constituents of fine particles in both cities during wintertime. During haze days, elevated concentrations of secondary ion species contribute considerably to the decrease in visibility (J. J. Gao et al., 2015; Zhang et al., 2015a). High relative humidity resulting in enhanced water uptake by the hygroscopic aerosol particles and formation of secondary aerosol have been suggested as an important factor during haze events in China (Sun et al., 2013; Wang et al., 2015b). Using aerosol mass spectrometry measurements of filters collected in four different cities in China during a haze event including Beijing and Xi’an, Huang et al. (2014) showed that the haze can be driven to a large extent by secondary aerosol formation. Furthermore, dust-related particles and biomass burning were identified as major contributors to the primary aerosol in Xi'an, while coal combustion particles dominated the primary aerosol in Beijing. However, filter measurements have a limited time resolution, and are often insufficient to characterize the rapid evolution of atmospheric aerosols or distinguish emission sources. Furthermore, filter measurements may suffer from sampling and measurement artifacts (e.g., loss of semi-volatile species due to evaporative processes during the sampling or absorption of gases into the filter material). Real-time measurements of aerosol particle composition with high time resolution have been conducted in wintertime in Beijing using various online mass spectrometers (e.g. Sun et al., 2013b; 2014; Zhang et al., 2014). These studies identified several primary sources, including traffic, coal combustion and cooking emissions. Sun et al. (2013b) found coal combustion particles to dominate the organic aerosol (OA) in Beijing in wintertime (on average 33 % of the OA) and enhanced contribution of this factor during polluted periods. Lower contributions of coal combustion aerosol were found in measurements performed in January 2013 (Zhang et al., 2014; Sun et al., 2014), with coal combustion explaining 15 and 19 % of the total OA, respectively. Among all three studies, the average contribution of traffic to the OA varied between 11 and 18 %, while cooking emissions explained between 12 and 20 % of the OA. However, all these studies failed to resolve a factor related to biomass burning, which is known to be a major particle source in winter. In addition, each study reported two to three oxygenated OA (OOA) components resulting from secondary processes. Secondary organic aerosol (SOA) was found to dominate the OA mass concentrations in January 2013 (54 % of OA in Zhang et al. (2014) and 55 % in Sun et al. (2014)), with increased relative contribution during more polluted days. Similar real-time measurements in other Chinese cities, including Xi’an, are scarce, preventing an accurate assessment of the spatial variation of the aerosol composition and sources in China during haze events.

Despite the widespread use of $PM_{2.5}$ as an air quality standard, previous online aerosol mass spectrometry measurements have only been able to measure the submicron fraction. In this work we present the first online high-resolution time-of-flight aerosol mass spectrometer (HR-ToF-AMS) measurements of the non-refractory (NR)-$PM_{2.5}$ fraction in two large Chinese cities during extreme and moderate pollution periods. The application of novel source apportionment techniques using the multi-linear engine tool (ME-2) provides an improved identification and quantification of OA sources compared to conventional positive matrix factorization (PMF) analysis. PAHs were quantified from AMS data and the contributions from their corresponding sources were determined.

2 Methodologies

2.1 Measurement campaign

Measurements were performed from 13 December 2013 to 6 January 2014 in Xi’an (34.23ºN, 108.88ºE, 10 m above ground level) and from 9 to 26 January 2014 in Beijing (40.00ºN, 116.38ºE, 20 m above ground level).

Xi’an, with over 8 million inhabitants in 2013, is the largest city in western China. Besides the local anthropogenic emissions, the region is often affected by the transport of dust particles from the Gobi desert and by the accumulation of pollutants when stagnant air masses are confined in the Guanzhoung basin. The sampling site was located within the High-Tech area south-west from the urban core, surrounded by many office buildings, some factories and construction sites. Nearby streets were sporadically watered during high pollution periods to minimize road dust resuspension.

Beijing, the capital of China, with over 20 million inhabitants in 2013 is one of the largest and most populated cities in the world. The city is located in a plain that opens to the east and to the south and is surrounded by mountains from the southwest to the northeast. The sampling site in Beijing was located in a residential area north of the urban core, near the Olympic Park. A large canteen was located about 20 m from the sampling site. Moreover, two main roads were situated at about 800 m in south and west directions. In Beijing, equivalent black carbon (eBC) concentration was measured in a separate location with similar characteristics, situated 2.8 km south from the sampling site described above.

2.2 Instrumentation

2.2.1 Aerosol Mass Spectrometer

An HR-ToF-AMS (Aerodyne Research Inc.) was deployed for online measurements of size segregated mass spectra of NR-$PM_{2.5}$. A detailed description of the instrument can be found elsewhere (Jayne et al., 2000; DeCarlo et al., 2006; Canagaratna et al., 2007). Briefly, a particle beam sampled through an aerodynamic lens is either alternately blocked and unblocked, yielding the bulk particle mass spectra (MS mode) or modulated by a spinning chopper wheel (~125 Hz), yielding size-resolved spectra (PToF mode) (Drewnick et al., 2005). In both operational modes, NR particles are flash vaporized by impaction on a heated tungsten surface (heated to about 600 ºC) at ~10^{-7} Torr. The resulting gases are ionized by electron ionization (EI, ~70 eV) and the mass-to-charge ratios (*m/z*) of the resulting fragments are determined by the ToF mass spectrometer.

The AMS was alternated between lower and higher mass resolution modes (V and W modes, respectively) each minute. For each of these modes, the AMS measured 25 seconds in MS mode, 25 seconds in PToF mode and the remaining 10 seconds were used to change configuration between V and W modes.

While commonly only the PM_1 fraction (particulate matter with aerodynamic diameter ≤ 1 µm) is accessible by the AMS, for this work we have equipped the AMS with a recently developed aerodynamic lens that extends the measured particle size to $PM_{2.5}$. This lens transmits efficiently particles between 80 nm and up to at least 3 µm and was well characterized by Williams et al. (2013). A ~ 2 m length copper tube (12 mm outer diameter) was used to connect a total suspended particles (TSP) inlet to the AMS. As the flow into the AMS is relatively low (~0.8 l min^{-1}), an auxiliary flow of ~ 4 l min^{-1} was maintained in this line in order to decrease particles losses. After the split with the auxiliary flow and before entering the AMS, the sampling air was drawn through a nafion drier in order to reduce uncertainties in the bounce-related collection efficiency (CE_b) and RH-dependent differences in aerodynamic sizing.

AMS data were analyzed in Igor Pro 6.3 (WaveMetrics) using the SQUIRREL (version 1.52L) and PIKA (1.11L) analysis software. Standard relative ionization efficiencies (RIE) were assumed for organics, nitrate and chloride (RIE = 1.4, 1.1 and 1.3, respectively) and experimentally determined for sulfate and ammonium (RIE = 1.48 and 3.37, respectively). The CE_b was corrected for by using a composition dependent collection efficiency (CDCE) algorithm by Middlebrook et al. (2012). Detailed studies of particle bounce in the AMS were not conducted for the larger particles sampled here; additionally, in principle the $PM_{2.5}$ lens could enhance particle bounce by operating at a higher pressure and thus yielding increased particle velocity for a given size. However, the comparison between total measured $PM_{2.5}$ mass (NR species and eBC, 24 hours average) with gravimetric measurements on filters (collected also in the $PM_{2.5}$ range at room temperature) suggests that these losses are not significant in the current study (see Fig. S1). Specifically, during the period for which filter measurements are available, the total measured $PM_{2.5}$ from AMS and Aethalometer is similar or greater than the gravimetric measurements. The comparison between these two techniques is biased by the higher cut-off of the $PM_{2.5}$ lens of the AMS combined with the presence of large particles during the extreme haze events, the increased loss of ammonium nitrate on the filters during extreme haze events, and the deposition of dust, other refractory compounds and water on the filters. However, the comparison evidences that in our case there is no important loss of mass in the AMS due to enhanced particle bounce.

2.2.2 Aethalometer

Two Aethalometers (Magee Scientific) were deployed for the determination of the aerosol attenuation at seven different wavelengths with a time resolution of 1 minute. From the change in the light attenuation, optical properties of the aerosol and eBC concentrations were retrieved. An Aethalometer model AE-33 was deployed in Xi'an, whereas a model AE-31 was deployed in Beijing. The newly developed model AE-33 uses a double spot technique for real-time loading compensation (Drinovec et al., 2015), while the data of the AE-31 was manually corrected for this effect after acquisition using the procedure of Weingartner et al. (2003). A $PM_{2.5}$ cyclone was located in front of the main inlet of the Aethalometers. The particles were transmitted from the cyclone to the Aethalometer through ~ 3 m of copper tube (12 mm outer diameter) at a flowrate of ~ 4 l min^{-1}.

2.3 Source apportionment techniques

2.3.1 OA source apportionment

Source apportionment was performed on the organic AMS data using PMF as implemented by the multilinear engine (ME-2; Paatero, 1997) and controlled via the interface SoFi coded in Igor Wavemetrics (Source Finder; Canonaco et al., 2013).

PMF is a bilinear unmixing receptor model which enables describing the variability of a multivariate database as the linear combination of static factor profiles and their corresponding time series. This is achieved by solving Eq. (1), where $\mathbf{X}$ is the measurement matrix (consisting of i rows and j columns), $\mathbf{G}$ contains the factor time series, $\mathbf{F}$ the factor profiles and $\mathbf{E}$ the model residuals. The model uses a least squares approach to iteratively minimize the object function Q (Eq. 2), defined as the sum of the squared residuals (e_{ij}) weighted by their respective uncertainties (σ_{ij}).

$$\mathbf{X} = \mathbf{G} \times \mathbf{F} + \mathbf{E} \tag{1}$$

$$Q = \sum_{\mathrm{i}=1}^{\mathrm{m}} \sum_{\mathrm{j}=1}^{\mathrm{n}} \left(\frac{e_{ij}}{\sigma_{ij}}\right)^2 \quad (2)$$

In our case, the model input consists of a data and error matrix of OA mass spectra, where the rows represent the time series and the columns contain the ions fitted in high resolution (HR) for the V mode data. Considering only the mass from the HR fits (up to *m/z* 115), 10 ± 8 % of the OA mass was excluded. The initial error values were calculated by the HR-AMS data analysis software previously described (PIKA) and a minimum error corresponding to the measurement of a single ion was enforced (Ulbrich et al., 2009). Further, as suggested by Paatero and Hopke (2003), variables with signal-to-noise ratio (SNR) lower than 0.2 were removed and variables with SNR between 0.2 and 2 were down-weighted by increasing their error by a factor of 3. Finally, all variables directly related to *m/z* 44 in the organic fragmentation table (i.e. *m/z*'s 16, 17, 18 and 28) (Allan et al., 2004) were excluded for the PMF analysis to appropriately weight the variability of *m/z* 44 in the algorithm and were reinserted post-analysis. After the aforementioned corrections were applied, the final input matrix contained 270 ions and 50909 points in time (with steps of 60 seconds).

PMF was solved using the multi-linear engine (ME-2, Paatero 1999), which in contrast to unconstrained PMF analyses enables complete exploration of the rotational ambiguity (i.e. different combinations of the matrices **G** and **F** can give solutions with the same mathematical quality) of the solution space. For computational efficiency, in this study this was achieved by directing the solution towards environmentally meaningful rotations using the *a* value approach. This method constrains one or more output factor profiles to fall within a predetermined range governed by the combination of an input profile and a range-defining scalar (*a* value). For example, in the case in which a factor profile (f_{j}, where *j* indicates the *m/z* of the ions fitted in high resolution) is constrained with a certain *a* value (*a*), the following condition needs to be fulfilled:

$$f_{\mathrm{j,sol}} = f_{\mathrm{j}} \pm a \times f_{\mathrm{j}} \quad (3)$$

The number of factors in PMF is determined by the user and the solutions of the model are not mathematically unique, due to rotational ambiguity. Therefore, it is very important to use criteria such as chemical fingerprint of the factor profiles, diurnal cycles and correlations between the time series of factors and external measurements to support factor identification and interpretation (Ulbrich et al., 2009, Canonaco et al. 2013).

2.3.2 eBC source apportionment

As mentioned in section 2.2.2, light attenuation was measured at seven wavelengths ranging from the ultraviolet to near-infrared, namely 370, 470, 520, 590, 660, 880 and 950 nm. The eBC from wood burning (eBC_{wb}) can be distinguished from eBC emitted by traffic(eBC_{tr}) by exploiting the enhanced absorption of eBC_{wb} in the ultraviolet range (Sandradewi et al., 2008). This method for the separation of wood burning and traffic contributions to eBC was developed and successfully applied in environments where no other major combustion sources were present. However, in China coal is widely used in some industrial sectors and for heating purposes, too. The optical properties of eBC from coal combustion remain very uncertain and the effect of coal emissions on the different wavelengths of the Aethalometer is still unknown. Yang et al. (2009) estimated an Angstrom exponent for coal burning aerosol of 1.46 from ambient measurements. This value falls between the more established Angstrom exponents for traffic (α_{tr} = 0.9 to 1.1) and wood burning (α_{wb} = 1.6 to 1.8) (Zotter et al., in prep). Thus, the presence of coal emissions makes the use of the method for eBC source apportionment rather uncertain.

The OA source apportionment results (see Sect 2.4.2) show rather high mass loadings of coal combustion OA (CCOA) in Beijing (23.4 ± 0.6 $\mu g\ m^{-3}$ on average and 53.8 ± 1.3 % relative contribution to OA mass), but a very low mass fraction from this source in Xi'an (10.5 ± 0.4 $\mu g\ m^{-3}$ on average and only 9.2 ± 0.3 % relative contribution to OA mass). Using these results, we estimated the ratio eBC/CCOA to be 0.037 ± 0.006 in Beijing. This was accomplished by fitting eBC as a linear combination of the three identified combustion sources: traffic (hydrocarbon-like OA, HOA), biomass burning (biomass burning OA, BBOA) and coal combustion (CCOA). Although the eBC measurements in Beijing were conducted at 2.8 km south from our sampling site, the reconstruction of the eBC concentration based on OA primary fractions from ME-2 shows a very good agreement with the measured eBC (see Fig. S2) and the obtained eBC/CCOA ratio is in good agreement with previous values reported in literature (Zhang et al., 2008). Using the ratio eBC/CCOA obtained for Beijing, we estimate that coal combustion contributed on average only 2.2 ± 1.4 % to the measured eBC in Xi'an. Moreover, also for Xi’an the reconstruction of eBC by means of the combustion OA sources is very successful (as shown in Fig. S2). Therefore, we conclude that the method described in Sandradewi et al. (2008) to separate eBC_{wb} and eBC_{tr} can be reasonably applied to our data from Xi'an, but not in the case of Beijing.

For the eBC source apportionment in Xi'an, Angstrom exponents of 0.9 and 1.7 were considered for traffic and wood burning, respectively, following the suggestions in Zotter et al. (in prep) presenting a re-evaluation of the method developed in Sandradewi et al. (2008). The eBC_{wb} to BBOA ratio was found to be 0.14, which is in good agreement with previous reported values (Gilardoni et al., 2011; Zotter et al., 2014). The ratio eBC_{tr} to HOA was 0.79, which is lower than the ratios reported in previous European studies (El Haddad et al., 2013 and references therein) but is in good agreement with results derived from measurements in China (Huang et al., 2012; Zhou et al., 2014). This difference in the eBC_{tr} to HOA ratio at the two locations is most probably related to the higher percentage of gasoline vehicles in China compared to Europe. Specifically, according to DeWitt et al. (2015), an eBCtr to HOA of around 0.8 corresponds to a diesel fuel share of about 40 to 50 %. This estimation is in good agreement with results from Gentner et al. (in prep), where a diesel fuel share of around 40 % was estimated for 2013 in China, dominated by heavy and medium duty vehicles.

2.4 AMS-PAH quantification

PAH concentrations were quantitatively determined from the high resolution AMS data. All details about the method used can be found in Bruns et al. (2015) and references therein. The following PAH molecular ions $[M]^+$ were fitted in the HR spectra: $[C_{10}H_8]^+$ (128), $[C_{12}H_8]^+$ (152), $[C_{14}H_8]^+$ (176), $[C_{14}H_{10}]^+$ (178), $[C_{16}H_{10}]^+$ (202), $[C_{18}H_{10}]^+$ (226), $[C_{18}H_{12}]^+$ (228), $[C_{20}H_{12}]^+$ (252), $[C_{22}H_{12}]^+$ (276), $[C_{22}H_{14}]^+$ (278), $[C_{24}H_{12}]^+$ (300) and $[C_{24}H_{14}]^+$ (302), with the nominal mass in parentheses. In addition to the aforementioned molecular ions, also other associated fragments were considered, including $[M\text{-}H]^+$, $[M\text{-}2H]^+$, $[M]^{2+}$ and $[M\text{-}H]^{2+}$ and the ^{13}C-isotopes of singly charged ions. To reduce uncertainty in the quantification of the associated ions, the ratios between molecular ions and their corresponding fragments were determined during periods with high PAH concentrations and then applied to the entire data set. Due to possible interference with ions from non-PAH compounds, the fragments that presented low correlation with their corresponding molecular ions (i.e. R2 below 0.6 for C13H6 to C13H7 and C16H6 to C16H7 ratios) were not taken into account in the analysis. In cases of overlap between a molecular ion and associated fragments (e.g. [M-2H]+ from [C14H10]+ overlaps with the molecular ion [C14H8]+) the fragments were not included. As for the non-PAH organics, the RIE for PAHs was assumedto be 1.4 and the dependency of the collection efficiency (CE_b) on the chemical composition of the aerosol was estimated using a composition dependent collection efficiency (CDCE) algorithm (Middlebrook et al., 2012).

Previous works found that PAH quantified from AMS measurements were systematically higher than PAH determined from filter measurements (Bruns et al., 2015). These differences were mostly attributed to filter artifacts (predominantly negative artifacts, i.e. volatilization of PAH on the filter surface). However, also the AMS-PAH analysis is subject to uncertainties. The PAH RIE considered in this work (1.4) is at the lower end of the values found in literature (e.g. Dzepina et al. (2007) measured RIEs between 1.35 and 2.1 for four PAH standards), and the reported PAH values would decrease by 33.3 % if a higher RIE (2.1) would be considered. Moreover, our results might be affected by interferences with fragments from other PAH classes at the quantified m/z's. To assess the contribution of these interferences we have analyzed the electron impact ionization mass spectrum (Linstrom et al., 2016) of major PAHs derivatives including alkylated, oxygenated and nitro –PAHs (43 compounds). The fragmentation patterns of alkylated and oxygenated –PAHs suggests that their response at the m/z's of interest is highly unlikely. On the other hand, nitro-PAHs exhibit a small response at some fragments of interest, but yield mainly fragments at odd m/z. In addition to the molecular ions mentioned above, we have considered as PAHs the ions [C11H7]+ (139), [C13H7]+ (163), [C13H9]+ (165), [C15H9]+ (189) and [C16H7]+ (199) and their related associated fragments. These fragments at odd masses have been observed to derive from the fragmentation of PAHs using laser desorption (Bente et al., 2009); they contribute to 32% of the total PAH mass quantified here. Due to all the limitations mentioned above, PAHs' concentrations presented here have to be considered as absolute highest estimates and we will refer to this chemical family as AMS-PAHs.

3 Interpretation and optimization of OA source apportionment

A key consideration for PMF analysis is the number of factors selected by the user. As currently no methodical and completely objective approach exists for choosing the right number of factors, this selection is generally evaluated through comparisons of the time series of the factor and external tracers as well as the analyses of factor mass spectra and diurnal patterns. In this work we present a detailed source apportionment that has been optimized to minimize the user subjectivity on the solution and better estimate the uncertainties of the final solution.

In a first step, we examined a range of solutions with two to eight factors from unconstrained runs (see Fig. S3). The solution that best represented the dataset is the five-factor solution, which yields factors interpreted as hydrocarbon-like OA (HOA), biomass burning OA (BBOA), coal combustion OA (CCOA), cooking OA (COA) and oxygenated OA (OOA). The HOA profile is distinguished by the presence of alkyl fragment signatures (Ng et al., 2011), with prominent contributions of non-oxygenated species at *m/z* 43 ($C_3H_7^+$), *m/z* 55 ($C_4H_7^+$) and *m/z* 57 ($C_4H_9^+$). BBOA is characterized by the presence of signals at *m/z* 60 ($C_2H_4O_2^+$) and *m/z* 73 ($C_3H_5O_2^+$), which are known fragments from anhydrous sugars present in biomass smoke (Alfarra et al., 2007). The CCOA profile, which has been previously identified in other locations (Dall'Osto et al., 2013), is characterized by the presence of unsaturated hydrocarbons, with higher explained variability of these unsaturated fragments at higher *m/z*.The COA profile is very similar to the HOA spectra but has higher contributions of the oxygenated ions at *m/z* 55 ($C_3H_3O^+$) and *m/z* 57 ($C_3H_5O^+$). Finally, the OOA profile is characterized by a very high *m/z* 44 (CO_2^+). COA is not resolved in solutions with a lower number of factors. Meanwhile, when a six-factor solution is considered, OOA splits into two factors with very similar profiles and whose time series reflect the change in the instrument tuning (Fig. S3). Further increasing the number of factors doesn't improve the interpretation of the data, as the new factor time series and spectral profiles are highly correlated with those extracted from lower order solutions and cannot be explicitly associated to distinct sources or processes.

Although the unconstrained five-factor solution appears to be a reasonable representation of the data, the mass spectral profiles indicate mixing between the sources. This is specifically the case between HOA and BBOA, where the HOA profile contains a higher than expected contribution of $C_2H_4O_2^+$ (*m/z* 60), and between COA and OOA, with a rather high contribution of CO_2^+ (*m/z* 44) in the COA profile. Precisely, in the unconstrained solution the *m/z* 60 in HOA is 0.009 ± 0.001 % (standard deviation from average over 10 seed runs), compared to 0.002 ± 0.002 % obtained from the average of multiple ambient datasets (Ng et al., 2011). Likewise, the *m/z* 44 in the unconstrained COA profile is 0.069 ± 0.001 %, compared to 0.013 ± 0.004 % obtained as an average of previously reported COA spectra (He et al., 2010; Crippa et al., 2013; Wolf, 2014).

To decrease the influence of BBOA on the apportionment of HOA, we constrained HOA using the profile from Crippa et al. (2013), which is characterized by a minor contribution of *m/z* 60. Note that while other approaches were explored throughout the entire analysis, including the use of other HOA profiles or increase of the factor number, the BBOA-HOA separation couldn't be significantly improved. As the vehicular fleet in China and Europe are significantly different, - e.g. higher diesel contribution in Europe, the use of a European profile to apportion traffic emissions in China could introduce significant errors. However, the comparison between HOA spectra from Europe (fleet dominated by diesel) and from the United states (fleet dominated by gasoline), shows that the variability among two European spectra (Mohr et al., 2012 and Crippa et al., 2013) is comparable to the variability among HOA spectra from the United States and Europe (Docherty et al., 2011 and Crippa et al., 2013). This was evaluated by means of cosine similarity analysis, which resulted in $\Theta_{\mathrm{HOA(Europe\text{-}Europe)}} = 0.93$ and $\Theta_{\mathrm{HOA(Europe\text{-}US)}} =0.92$. Thus, we show that HOA emissions from different types of cars have similar profiles. Although constraining the HOA improves the HOA-BBOA separation, it compromises the apportionment of cooking emissions, with a higher background mass and unexpectedly high concentration overnight in the diurnal trend. To avoid the mixing of COA with other sources, the COA profile of Crippa et al. (2013) was constrained. While some differences are expected between the Chinese and European cooking activities, cosines similarity analysis indicate very good correlations ($\Theta_{\mathrm{COA(Europe\text{-}China)}} = 0.97$ on average) between the COA profile from Paris and four spectra from different types of Chinese cooking (CC1 to CC4 in He et al., 2010). Moreover, the use of *a* values allows for a certain re-adjustment of the input profiles (for both HOA and COA), minimizing the effect of using a non-local input profiles. In the following we discuss the sensitivity of the results to the *a* values used to constrain the HOA and COA factor profiles.

Considering *a* values between 0 and 1 with a step of 0.1 for both HOA and COA yields 121 possible combinations of *a* values. A set of three criteria was established to assess the solutions that represent environmentally better the OA fractions.

1) Minimization of m/z 60 in HOA. A threshold for the maximal fractional contribution of *m/z* 60 in HOA was set to 0.006 based on profiles derived from multiple ambient datasets (mean ± 2σ from Ng et al., 2011). The fractional contribution of *m/z* 60 to the normalized HOA profiles varied between 0.0016 and 0.0092 % over the full *a* value space. This criterion eliminated all solutions with an *a* value for HOA of 1, as shown in Fig. S4.

2) Optimization of COA diurnals. Unambiguous chemical markers for cooking emissions are not yet clearly established, hindering their use for the optimization of the apportionment of this source. A valuable characteristic for the identification of COA is the analysis of its diurnal trends: near the emissions source (e.g. in an urban area) COA typically has a distinctive diurnal with maxima at lunch and dinner times. In order to categorize the solutions, a novel approach using cluster analysis was utilized. The normalized COA diurnals of all studied *a* value combinations were grouped using k-means cluster analysis. This technique aims at grouping the observations into k clusters, by minimizing the first term (T1) from the cost function (CF) shown in Eq. (6). This term represents the sum of the Euclidian distances between each observation (x_i) and its respective cluster center (μ_{zi}). The results from the cluster analysis are shown in Fig. 1, for two-, three-, and four-cluster solutions. For each solution, the first panel shows all diurnals pertaining to the different clusters, the second plot shows the diurnal pattern of the cluster center and the third plot shows the clusters' attribution in the *a* value space. An issue encountered in cluster analysis is the determination of the number of clusters (k) that best describes the data. Increasing k decreases T1, while adding complexity to the solution. A common approach to select the optimal number of clusters is to explicitly penalize the higher order solutions for complexity by using the Bayesian information criteria (BIC). This penalty is introduced with the second term (T2) in Eq. (6), given by the product of the number of clusters (k) and the logarithm of the dimensionality of the cluster (D=24 hours in our case):

$$\mathrm{CF} = \mathrm{T1} + \mathrm{T2} = \|x_i - \mu_{zi}\|^2 + k \times \log(D) \qquad (6)$$

Figure 2A shows a minimum in the cost function at three clusters, which are therefore selected to represent the different types of COA diurnals. From the three-cluster solution in Fig. 1, the diurnals of the purple cluster exhibit a higher background concentration over the full day, which are difficult to reconcile with the expected COA emission trends. The red and blue clusters have both lower background values; however the blue cluster has some peaks over the night hours that aren't expected from COA emissions. Moreover, the solutions in the red cluster are more similar to previous reported COA spectra (He et al., 2010; Crippa et al., 2013; Wolf, 2014), as they have a lower contribution of *m/z* 44 compared to the solutions in the other two clusters (see Fig. S5). Specifically, the average relative contribution of *m/z* 44 in the COA spectra from literature previously mentioned is 0.013 ± 0.004 %, which is in good agreement with the relative contribution of 0.013 ± 0.002 % found for the red cluster. As the spectrum for the blue and purple clusters have higher contributions of *m/z* 44 (0.026 ± 0.008 % and 0.025 ± 0.019 %, respectively), only the solutions belonging to the red cluster are considered as good solutions. A disadvantage of the k-means algorithm is that the solution space might have several local minima and therefore the result could strongly depend on the initialization. Hence, 100 random initializations of the algorithm were performed and only the *a* value combinations that fell into the red cluster more than 95 % of the time were retained as good solutions. Combining these results with the criterion previously applied on the HOA profile, we obtained the range of accepted *a* values combinations shown in Fig. 2B.

3) Factor-tracer correlation. The following correlations between the identified primary sources and the external tracers were considered:

$$PAH_{calculated}\,(\mathrm{t}) = \mathrm{a} \times \mathrm{BBOA(t)} + \mathrm{b} \times \mathrm{CCOA(t)} + \mathrm{c} \times \mathrm{HOA(t)} \tag{7}$$

$$\mathrm{eBC}_{tr\ calculated}\,(\mathrm{t}) = \left(\frac{\mathrm{eBC}_{tr}}{HOA}\right)_{median} \times \mathrm{HOA(t)} \tag{8}$$

$$\mathrm{eBC}_{wb\ calculated}\,(\mathrm{t}) = \left(\frac{\mathrm{eBC}_{wb}}{BBOA}\right)_{median} \times \mathrm{BBOA(t)} \tag{9}$$

In all cases, low concentration points (below the 5th percentile, P05) were discarded. Note that the separation between $\mathrm{eBC_{tr}}$ and $\mathrm{eBC_{wb}}$ was only possible with the data collected in Xi'an (see Sect 2.3.2). Moreover, as the $\mathrm{eBC_{wb}}$ doesn't follow the BBOA time series during the haze event (see discussion in Sect 4.2), only data from the reference period was considered for this analysis. The linear relation between AMS-PAH and BBOA, CCOA and HOA will be discussed in detail in the source apportionment result section (Sect 4.3).

For each of these parameters, the normalized difference, *S*, between the measured and calculated marker concentrations was retrieved for all accepted *a* value combinations using Eq. (10). The standard deviations of *S*, which are considered as an estimate of the variability between the factor and its corresponding marker, were combined in quadrature as shown in Eq. (11):

$$S = \frac{F_{measured} - F_{calculated}}{F_{measured}}, \quad \text{with } F\text{=AMS-PAH, eBC}_{\text{tr}} \text{ and eBC}_{\text{wb}} \tag{10}$$

$$\sigma_{ALL} = \sqrt{(\sigma_{AMS-PAH})^2 + (\sigma_{eBC_{tr}})^2 + (\sigma_{eBC_{wb}})^2} \tag{11}$$

where σ_{ALL} is the object function that needs to be minimized for the optimization of the selected solutions and is represented with a color scale in Fig. 2C. The standard deviations of the individual parameters ($\sigma_{AMS\text{-}PAH}$, σ_{eBCtr} and σ_{eBCwb}) within the accepted *a* value space are shown in the supplementary information (Fig. S6). As seen from Fig. 2C, the solution obtained using an *a* value of 0.9 for the HOA profile and 0.6 for the COA profile, has the minimum σ_{ALL} (σ_{min} =0.94). In order to establish the stability of the solution at a certain *a* value combination with respect to the measurement uncertainty, we examined the variability of σ_{min} for the best solution, by reinitializing 50 times the ME-2 algorithm with different input matrices. For each repetition, the elements of the OA input matrix were varied within twice their uncertainties (OA(i,j) ± $2OA_{error}(i,j)$). All of the 50 solutions satisfied the two criteria previously described (i.e. minimization of *m/z* 60 in HOA and optimization of COA diurnal) and σ_{ALL} presented 7.5 % variability among the 50 iterations. Considering all solutions inside the 95 % confidence interval (i.e. twice its variability, $\sigma_{ALL} < \sigma_{min}$ + 15.0 %) to represent the data equally well compared to the best solution, all *a* value combinations within the red region in Fig. 2C were retained. All results presented hereon are averaged over all this possible *a* value combinations, and their standard deviation is considered as our best estimation of ME-2 errors. Note that these errors are very likely lower estimates of the model uncertainties, as the solution space could not be fully explored. The error bars in Fig. 3 represent the variability of each *m/z* fraction (standard deviation) across all good solutions in the *a* value space. As this retained range of solution is a direct consequence of our input error estimate, we assessed the sensitivity of our results to the input errors by running the algorithm by varying the OA input matrix within smaller limits (OA ± $1OA_{error}$). This lead to similar results as the method described above, with the only difference that two additional *a* value combinations (marked with the dashed line in the left corner of Fig. 2C) would not have been considered as good solutions in this case.

Compared to the unconstrained PMF (average over 10 seeds), the optimized solution (average over all good *a* value combinations) has more genuine factor profiles (Fig. 3), with decreased contributions of *m/z* 60 in the HOA spectra (from 0.009 ± 0.001 % to 0.003 ± 0.001 %) and of *m/z* 44 in the COA spectra (from 0.069 ± 0.001 % to 0.013 ± 0.002 %). In terms of the relative contributions of the different sources to the total OA (Fig. S7), the optimized solution yielded significantly lower COA (7.0 ± 1.1 % vs. 19.9 ± 0.1% in the unconstrained PMF) and HOA (15.1 ± 1.6 % vs. 25.1 ± 0.1% in the unconstrained PMF). Moreover, σ_{ALL}, the object function that we seek to minimize, decreases considerably from 3.3 ± 0.1 in the unconstrained solution to 1.0 ± 0.1 in the optimized solution. In terms of the model mathematical performance, there is only a moderate increase in the residuals in the optimized solution compared to the unconstrained run. Specifically, *Q* normalized by its expected value (*Q/Qexp*) (Paatero and Hopke, 2009) increases from 7.5 ± 0.1 in the unconstrained solution to 8.5 ± 0.4 in the optimized solution. The correlations between the OA factors from the optimized solution and its corresponding tracers are presented in Fig. S8.The correlation parameters (R^2 and slope) are reported in Table S1 for the unconstrained and optimized solutions. Compared to the unconstrained solution, the correlations between COA and its marker ($C_6H_{10}O$) are higher in the optimized solution, while the correlations between OOA and NH_4 are slightly lower in this case, especially during the haze events.

4 Results and discussion

4.1 Bulk PM2.5 chemical composition

Figure 4A shows the temporal variations of the non-refractory (NR) chemical components measured by the AMS (OA, SO_4, NO_3, NH_4 and Cl) and eBC concentrations during the measurement periods in Xi'an and Beijing. The periods highlighted with a blue background relate to extreme haze events, which are defined by a visibility (reported in the top panel) below 2 km (Zhang et al., 2015b). We recognize that the reduction of the visibility is partially due to the increase of the aerosol water content as a result of the increase in the RH. However, during the extreme haze periods a significant increase in the total aerosol burden is observed, with total $PM_{2.5}$ mass reaching peak concentrations above 1000 µg m^{-3} in Xi'an. Regarding the chemical composition, Fig. 4A shows an increase in the inorganic aerosol species during the extreme haze periods, while organic aerosols dominate the particle mass in the reference periods (i.e. visibility above 2 km). In the medium panel of Fig. 4A, the ratio PM_1 to $PM_{2.5}$ mass is reported. This ratio was obtained from the integration of the collected PToF data. The size distributions of each species were integrated over the full measured size range (up to 6000 nm) to determine the total mass measured with the $PM_{2.5}$ lens and until 800 nm as an estimation of the mass that the commonly used PM_1 lens would have detected (the 50 % cut-off diameter of the PM_1 inlet is at about 800 nm vacuum aerodynamic diameter, Liu et al., 2007). As mentioned in section 2.2.1, the actual upper cut-off of the $PM_{2.5}$ inlet has been determined to be above 2.5 µm (Williams et al., 2013). This comparison between PM fractions might suffer from the slow evaporation of some particles in PToF mode, which would lead to a higher apparent d_{va} and a calculated higher than true mass loss in the PM_1 lens. On the other hand, it is possible that super-micron particles are more prone to particle bounce (Liu et al., 2013). Despite the uncertainties related to this calculation, the importance of measuring the $PM_{2.5}$ fraction in China is clear, especially during extreme haze events, where 39 ± 5 % of the mass would have been lost in the PM_1 lens.

As different emission sources can be present in the two measurement locations and some characteristics of the aerosols are expected to be distinct during the extreme haze periods, results are presented below for four different time frames: (1) Xi'an extreme haze (17 December to 26 December 2013), (2) Xi'an reference (13 December 2013 to 6 January 2014, excluding extreme haze), (3) Beijing extreme haze (15 January to 17 January 2014, with a small gap of some hours) and (4) Beijing reference (9 January to 26 January 2014, excluding extreme haze). Table S2 contains a summary of the mean concentrations of all measured compounds and OA sources during the four time periods.

The median diurnal trends of the AMS species and eBC are shown in the top panel of Fig. 5 (see the 25th and 75th percentiles in Fig. S9). The extreme haze events in Beijing occurred twice over night and therefore the diurnals are incomplete and hard to interpret. The diurnal trends are rather flat during the extreme haze in Xi’an, and exhibit more variation (with maximum concentrations at night) for the reference periods in Xi'an and Beijing. This variation is strongly influenced by the evolution of the planetary boundary layer height (which governs the vertical dilution of pollutants) and by the diurnal cycle of the emissions. During the reference periods, the increased solar radiation induces the development of the mixing layer during daytime, and therefore the dilution of the pollutants. At night, the pollutant concentrations increase as a result of additional emissions in an increasingly shallower planetary boundary layer. During extreme haze periods, less solar radiation reaches the Earth’s surface (see Fig. S10) and therefore dilution is reduced and particle concentrations remain elevated throughout the day.

Another important characteristic of the extreme haze events is the size of the measured particles. On average larger particles were detected during the extreme haze periods (size distribution modes at about 800 nm in Xi’an and between 800 and 1000 nm in Beijing) than during the reference periods (distribution modes at about 400 nm for both cities, Fig. 5).

As mentioned above, the mass of all aerosol components increased considerably during the extreme haze periods. The measured mean concentrations (and standard deviations) were 537 ± 146 $\mu g\ m^{-3}$ and 243 ± 47 $\mu g\ m^{-3}$ during the extreme haze periods in Xi’an and Beijing, and 140 ± 99 $\mu g\ m^{-3}$ and 75 ± 61 $\mu g\ m^{-3}$ for the reference periods in Xi’an and Beijing, respectively. As shown by the relative contributions in the bottom panel of Fig. 5, the increase in mass during the extreme haze events is particularly high for the inorganic species (NO_3, SO_4, NH_4 and Cl) and therefore, the resulting ratio of inorganic (inorg) to organic (org) species is much higher during the extreme haze periods. Specifically, inorg/org ratios of 1.3 and 1.4 were obtained for the extreme haze conditions in Xi'an and Beijing, while the ratio dropped to 0.6 and 0.8 for the reference periods in Xi'an and Beijing, respectively. The mass concentration and relative contribution of eBC is higher in Xi'an than in Beijing, probably due to a higher contribution of older cars in Xi'an.

4.2 OA sources

In this section the final results of the OA source apportionment are presented. All results are averages of all *a* value combinations that were accepted in the methodology described in section 3. The absolute concentrations and relative contributions of the OA sources over time are shown in Fig. 6A together with the time series of external tracers. The absolute concentrations of the sources have rather small variability across all good solutions (see Fig. S11). The mean OA concentrations and relative contributions of the identified OA factors are shown in Fig. 6B for the four periods of interest. Lastly, the daily patterns of the absolute concentrations of the identified OA factors and the external tracers are reported in Fig. 7. Similar to the inorganic species and total OA, the diurnals of the OA factors are partially driven by the PBL dynamics, with increased dilution during daytime and accumulation of the particulate mass overnight. Nevertheless, some factor-dependent differences are evident. The 25th and 75th percentiles of these diurnals and the standard deviation among all considered *a* value combinations are reported in Fig. S12. Potential Source Contribution Function (PSCF) analysis was performed to explore the geographical origin of the air masses during the measurements and to identify source regions and other transport-related pollution events (see Fig. S13).

OOA. A pronounced increase in the OOA mass concentration is observed during the extreme haze periods (blue background in Figure 6A), reflecting the importance of secondary organic aerosol formation under these conditions. This increase in mass is also reflected in an enhanced relative contribution of OOA to total OA during the extreme haze periods (from 16.2 ± 1.1 % to 31.3 ± 1.5 % in Xi'an and from 15.7 ± 0.7 % to 25.0 ± 1.2 % in Beijing). In terms of the diurnal trends, for all examined periods OOA concentrations remain rather constant during the day with only a slight increase in the late morning and afternoon. These increases are most probably related to regional production of OOA due to enhanced photochemical activity. These results are in agreement with the PSCF results, where shorter backward air mass trajectories during the extreme haze period in Xi'an indicate that regional emissions (within around 1000 kilometers) might play a dominant role during the extreme haze. OOA concentrations are higher with northeast winds during the haze period in Xi'an (which might indicate a source region as there isn't a characteristic diurnal variation for the wind direction) while for the reference periods in Xi'an and Beijing the OOA shows rather homogeneous spatial distributions.

COA. The COA average relative contribution to total OA is generally low for the extreme haze periods (3.6 ± 0.5 % in Xi'an, 5.8 ± 1.0 % in Beijing) and around 10% (9.3 ± 1.6 % in Xi'an, 11.5 ± 1.9 % in Beijing) for the reference periods. For all four periods, COA shows a very distinct diurnal trend with very strong peaks at lunch (between noon and 13:00 local time LT) and dinner (19:00 to 20:00 LT) times. A small increase in the COA concentrations is also observed in the morning (6:00 to 7:00 LT), coinciding with breakfast time. The fragment ion $C_6H_{10}O^+$ at *m/z* 98 has been suggested among others as a marker ion for the COA factor (Sun et al., 2011; Crippa et al., 2013). Nevertheless, the correlation between these two components is very poor, mostly during the extreme haze period in Xi'an (R^2=0.21, see Table S1). This low correlation is mostly due to increased concentrations of $C_6H_{10}O^+$ over the night hours, when COA decreases to background concentrations. Those night peaks coincide with the enhancement of BBOA, which in our case explains on average about 40 % of the mass of the $C_6H_{10}O^+$ fragment. For all periods considered in the PSCF analyses, COA appears to be a local source.

CCOA. Coal emissions are high in Beijing, dominating OA burden with contributions greater than 45 % of the OA mass (46.8 ± 1.2 % and 55.2 ± 1.6 % for extreme haze and reference periods, respectively). In comparison, in Xi'an CCOA is of lower importance (5.7 ± 0.1 % and 14.0 ± 0.6 % for extreme haze and reference periods, respectively). The CCOA mass slightly increases during the haze periods (more clearly seen in the case of Xi'an), probably due to the accumulation of primary emissions during the stagnant conditions. CCOA concentrations decrease substantially during day time, due to dilution of the emissions in a deeper PBL. CCOA concentrations peak in the morning (at around 9:00 LT) and at night (starting to rise at 18:00 LT), probably due to domestic heating activities. Moreover, the CCOA is characterized by a strong peak in concentrations at around 3:00 to 4:00 LT, especially during the extreme haze period in Beijing. This peak, which is also present in the corresponding BBOA diurnal, might result from the late night burning emissions in a shallower boundary layer or from the advection of evening emissions from the surrounding areas. The PSCF results indicate that the high concentrations of CCOA (and BBOA) measured at the sampling site in Beijing might be related to air masses coming from southwest of the sampling site (from the Hebei region).

BBOA. Unlike CCOA, BBOA is much more important in Xi'an, comprising about 40 % of the OA mass in the two considered periods (43.4 ± 1.1 % and 42.2 ± 1.5 % for extreme haze and reference periods, respectively). In Beijing instead, BBOA represents less than 15 % of the total OA (13.8 ± 0.8 % and 8.9 ± 0.3 % for extreme haze and reference periods, respectively). Accordingly, while combustion emissions from domestic heating and cooking predominate the organic aerosol mass at both locations, our results highlight the clear difference in the type of fuel used for burning, with a higher fraction of coal burned in Beijing vs. a higher fraction of biomass burned in Xi'an. BBOA primary emissions appear to accumulate under the stagnant conditions during severe haze events. In particular, in the last days of extreme haze in Xi'an, very high concentrations of BBOA are perceived without a significant increase in eBC_{wb}. Nonetheless, the temporal evolution of BBOA correlates with the ion $C_2H_4O_2^+$ at *m/z* 60 (overall R^2=0.97), confirming the assignment to BBOA. This specific episode might be related to special conditions with lower amounts of eBC_{wb} (e.g. from smoldering conditions) or absorbing wood burning organic carbon (e.g. from smoldering conditions or aged emissions – which would result in an overestimation of eBC_{tr}). The characteristics in the diurnal trends of BBOA are similar to those found in CCOA. The dilution of the particles in a deeper PBL during day time results in a decrease in the BBOA concentration at around 16:00 LT, while peaks related to residential heating appear in the morning (between 9:00 to 10:00 LT) and at night (starting to rise at 18:00 LT). As already mentioned there is a strong peak at around 3:00 to 4: 00 LT in the BBOA signal, which is probably related to the late night biomass burning emissions in a shallower PBL. In Xi'an, the PSCF results show that high concentrations of BBOA (and also HOA and CCOA) are observed when the air parcels are transported to the sampling site from northwest, indicating a possible major local pollution area northwest of the sampling site. In Beijing BBOA seems to be transported together with CCOA from the southwestern province of Hebei.

HOA. Despite the larger vehicle fleet in Beijing, higher concentrations of HOA are noticeable in Xi'an, possibly owing to a higher contribution of older cars. Accordingly, HOA is the third contributing source in Xi'an, explaining about 15 % of the OA mass (16.0 ± 1.6 % and 18.3 ± 1.9 % for extreme haze and reference periods, respectively). By contrast, in Beijing, HOA is the least important source together with COA, explaining only around 8 % of the OA mass (8.6 ± 1.3 % and 8.7 ± 1.2 % for extreme haze and reference periods, respectively). An increase in HOA levels can be noticed during the haze periods, related to the accumulation of primary emissions under stagnant conditions. The HOA diurnals show peaks during morning and evening rush-hours (7:00 to 8:00 and 8:00 LT, respectively), as is typically the case for traffic-related pollutants. Additional peaks are observed in the HOA during the night hours (between 23:00 and 6:00 LT). These peaks might be related to truck activity, which is strongly enhanced during the night hours as in both cities truck activity is banned during the day. While during the extreme haze event in Xi’an the PSCF results indicate an HOA source region northwest from the sampling site, homogeneous distributions of the HOA factor are found for the reference periods in both Xi’an and Beijing, indicating a homogeneous distribution of this source.

4.3 Effect of relative humidity on aerosol composition

As previously mentioned, periods identified as extreme haze in this study are characterized by high RH (see Fig. 4). We examine in Fig. 8 the impact of RH on aerosol concentration and composition following the approach proposed by Sun et al. (2013). As we have identified different emission patters in the two cities and the RH was only few times above 60% in Beijing, the analysis is only performed for the case of Xi’an. In Fig. 8a, we display the mass concentrations of the NR aerosol species (top) and of the identified OA sources (bottom) as a function of RH, with RH bins of 10 % increments. The absolute mass concentrations of all aerosol species increase during extreme haze events, hence with RH. OA and sulfate show the largest increases compared to nitrate, ammonium and chloride. Among the OA sources, BBOA is strongly enhanced at higher RH, followed by OOA and HOA. In contrast, haze conditions seem to have a marginal impact on CCOA and COA concentrations. To exclude accumulation and dilution effects, we normalize in Fig. 8b aerosol species and OA fractions by the sum of the primary OA fractions (i.e. HOA, BBOA, CCOA and COA), as a surrogate for secondary aerosol precursors. For a better illustration, all the ratios were further normalized by their value at the first RH bin (10-20 %). At RH below 50 %, none of the normalized aerosol species show a clear trend with increasing RH, while at higher RH only sulfate shows a sizeable increase. A change in the emissions of the different primary sources is also observed, with an increase in the BBOA and HOA contributions during extreme haze events, as described in the previous section. Specifically, during the extreme haze period the contributions of BBOA and HOA to the total primary OA increase from 47 to 64 % and from 17 to 21 %, respectively, while the contributions of CCOA and COA decrease from 14 to 6 % and from 22 to 8 %, respectively. Therefore, the normalization of the OA fractions produces an apparent increase in HOA and BBOA and a decrease in CCOA and COA with higher RH. These effects can't therefore be unequivocally attributed to the change in RH. More importantly, although the OOA mass concentration increases from about 10 to 60 µg m^{-3} when RH varies from 50 to 90 %, when normalized to its potential precursors, OOA does not show significant variability with RH. This suggests that unlike sulfate, whose production is highly enhanced in the aqueous phase at high RH, OOA production rates seem to be independent of RH.

The strong increase of the normalized sulfate at high RH suggests that aqueous phase oxidation of SO_2 could be an important process during extreme haze events. To investigate the oxidation degree of sulfur at different RH, the sulfur oxidation ratio (F_{SO4}, Sun et al., 2006) was calculated according to Eq. (12) (where n is the molar concentration) and is reported in Fig. 9 as a function of RH (note that this plot contains the full campaign data).

$$F_{SO_4} = \frac{n[SO_4]}{n[SO_4]+n[SO_2]} \tag{12}$$

As seen in Fig. 9, F_{SO4} has a clear exponential trend with RH. At RH below 50 % F_{SO4} is rather constant and low (about 0.045 on average), while for higher RH the oxidation ratio rapidly increases reaching 0.62 on average for the last RH bin (90-100 %). This extremely high oxidation degree of sulfur under high RH is an indication that aqueous phase production of sulfate might play a very important role during extreme haze events in China, in good agreement with the results reported by Sun et al. (2013) for wintertime in Beijing.

4.4 AMS-PAH sources

To identify all sources emitting PAHs, PMF was performed using the OA matrix as an input, with an additional column containing the total AMS-PAH mass concentration calculated from the AMS. AMS-PAH errors were calculated assuming a Poisson distribution and the goodness of the combination of the two datasets (OA and AMS-PAH) was evaluated examining the model residuals. The AMS-PAH weighted residuals, which are reported in Fig. S14A, are distributed around zero. However a small increase in their weighted residuals (Fig. S14B) is observed over night. The average AMS-PAH attribution was 28.9 ± 0.4 % to BBOA, 57.0 ± 0.7 % to CCOA and 14.1 ± 0.4 % to HOA (errors denote the standard deviation from 10 seed runs). The same combined input matrix was afterwards tested in the ME-2 approach, with the HOA profile constrained with an *a* value of 0.9, the COA profile constrained with an *a* value of 0.6, and the AMS-PAHs unconstrained in all factors. Also in this case the AMS-PAHs were attributed to these three combustion sources with similar results (28.6 ± 0.4 % to BBOA, 62.0 ± 0.1 % to CCOA and 9.4 ± 0.3 % to HOA, with errors being the standard deviation among 10 seed runs).

Hence the measured AMS-PAHs in our dataset can be fully attributed to biomass burning, coal burning and traffic emissions. Using a linear regression of BBOA, CCOA and HOA to fit the measured AMS-PAHs (see Eq. (7) in Sect 3) very similar attributions of the mass are found (27.6 ± 0.7 % attributed to BBOA, 66.4 ± 0.4 % to CCOA and 6.0 ± 0.5 % to HOA). The result of this fit (averaged over all good *a* value combinations) is shown in Fig. 10A together with the total mass of the measured AMS-PAHs. As it can be seen from this time series, the linear regression can reconstruct the measured AMS-PAHs very precisely (R^2=0.94) and peaks of over 10 µg m^{-3} of AMS-PAHs can be attributed to the combined biomass burning, coal combustion and traffic emissions. The high AMS-PAH concentrations lead to high AMS-PAH to OA ratios (1.9 ± 0.7 % in Xi’an and 4.4 ± 2.2 % in Beijing) compared to previous reported values for Europe. This can be related to the different methodologies used to measure PAHs (e.g. volatilization of semi-volatile PAHs and oxidation of unstable PAHs on filters). Moreover, in our case combustion sources, especially coal burning, explain a very large fraction of OA, which would enhance the AMS-PAH to OA ratio compared to Europe. In this regard, Chen et al., (2005) reported mean PAH to OC ratios of 28 % (i.e. PAH/OA of 17.5 % assuming OA/OC of 1.6) for bituminous coal and 0.8 % (i.e. PAH/OA of 0.5 %) for anthracite. Considering that a mixture of these two types of coal is used in the cities considered in this work and that the relative contribution of coal to the total OA is higher in Beijing than in Xi’an, the obtained AMS-PAH to OA ratios seem reasonable.

Fig. 10B presents the average AMS-PAHs concentrations and relative contributions of the three combustion sources to the measured AMS-PAHs for the different measurement periods. During the extreme haze event in Xi’an, 63.8 ± 1.1 % of AMS-PAHs are attributed to biomass burning, 25.3 ± 0.4 % to coal combustion, and the rest (10.9 ± 0.9 %) to traffic emissions. For the reference period the contribution of coal increases to about 55.9 ± 0.9 %, the biomass burning influence decreases to around 36.4 ± 1.4 % and the traffic remains a minor contributor, explaining about 7.7 ± 0.8 % of the AMS-PAHs mass. In Beijing, coal emissions completely dominate over biomass burning and traffic emissions, and on average 88.1 ± 1.1 % and 94.1 ± 0.3 % of the measured AMS-PAHs can be attributed to coal during the extreme haze and reference periods, respectively. These results are in agreement with Huang et al. (2014), showing that coal burning emission is an asymmetric source of PAHs, compared to other combustion emissions.

5 Discussion and conclusions

This work presents a thorough analysis of extreme haze events (visibility below 2km) which occurred in Xi’an and Beijing during winter 2013-2014. Online aerosol mass spectrometer analyses provided a detailed characterization of the chemical composition and size distribution of the aerosol components during the different measurement periods.

The extreme haze events were produced by a combination of primary emissions of particulate matter, generation of secondary aerosol, and stagnant meteorological conditions which confined the pollutants in the basin. Under such conditions, the mass concentrations of all aerosol components strongly increased, with resulting average $PM_{2.5}$ mass concentrations of 537 ± 146 $\mu g\ m^{-3}$ in Xi’an and 243 ± 47 $\mu g\ m^{-3}$ in Beijing (in contrast to 140 ± 99 $\mu g\ m^{-3}$ and 75 ± 61 $\mu g\ m^{-3}$ average $NR\text{-}PM_{2.5}$ mass measured during the reference periods in Xi’an and Beijing, respectively). Among all aerosol components, sulfate and nitrate show the strongest enhancements during the extreme haze periods. Moreover, source apportionment of the organic aerosol (OA) fraction shows that also the formation of oxygenated organic aerosols (OOA) is strongly enhanced during the haze events. The high relative humidity characteristic of the periods with extreme haze was shown to favor the heterogeneous oxidation of SO_2 on deliquesced aerosols and can therefore drive the drastic increase in sulfate concentrations. In contrast, aqueous phase processing appears not to significantly affect the formation of OOA and the other inorganic species.

Another distinct feature of the aerosols during extreme haze events is their larger size compared to particles during lower pollution periods (the distribution mode of all NR-aerosol compounds shifts from around 400 nm during the reference periods to about 800 to 1000 nm during extreme haze events in both cities). The growth of the particles is associated with high secondary aerosol fractions and condensation of semi-volatile compounds on preexisting particles. Given the large mean aerosol diameters found during the extreme haze periods, the use of a $PM_{2.5}$ inlet for the AMS was a crucial point for the meaningfulness of our results, as 39 ± 5 % of the mass would have been neglected if a standard PM_1 inlet had been deployed.

The use of a novel source apportionment technique (ME-2), together with a systematic analysis focused on minimizing the effect of user subjectivity on the solution, allowed for the separation of the several primary sources of OA in the two cities under study. Compared to previous studies at a single site where the simultaneous extraction of coal and biomass burning factors is exceedingly challenging, this separation was possible here by including measurements at two contrasted sites with different exposure patterns. Our source apportionment results suggest that biomass burning (from domestic heating and agricultural activities) is a major source of OA in Xi'an during wintertime (representing 42 to 43 % of the OA mass), while coal emissions (from domestic heating, cooking and industrial processes) dominate the OA mass in Beijing (47 to 55 %). Coal combustion and biomass burning were also the major cause for very high concentrations of AMS-PAHs (on average 2.1 $\mu g\ m^{-3}$), known to be highly carcinogenic. Moreover, PSCF analyses indicate that these coal-related particles are at least partially transported from the industrial province of Hebei to the highly populated capital. Oxygenated organic aerosol (related to secondary processes) is also found to be an important contributor to the measured OA mass, mostly during the extreme haze events. The relative contribution of OOA increases from 16 to 31 % and from 16 to 25 % of OA mass during the extreme haze events in Xi'an and Beijing, respectively. Traffic emissions have a slightly larger impact in Xi'an (representing 15 to 20 % of the OA) than in Beijing (about 9 % of the OA), while cooking is a rather minor source in both cities (explaining 5 to 10 % of the OA mass). Considering these results, major efforts should be put into regulating more thoroughly the biomass and coal burning activities widely spread in urban areas in China and regulate the gaseous precursor emissions of organic and inorganic aerosols.

Acknowledgements

The research leading to these results received funding from the "Strategic Priority Research Program" of the Chinese Academy of Sciences (Grant No. XDB05000000) and National Natural Science Foundation of China (Grant No. 41403110). JGS acknowledges the support of the Swiss National Science Foundation (Starting Grant No. BSSGI0 155846). IE-H acknowledges the support of the Swiss National Science Foundation (IZERZ0 142146). The authors gratefully acknowledge the NOAA Air Resources Laboratory (ARL) for the provision of the HYSPLIT transport and dispersion model and READY website (http://www.ready.noaa.gov) used in this publication.

References

Alfarra, M. R., Prévôt, A. S. H., Szidat, S., Sandradewi, J., Weimer, S., Lanz, V. A., Schreiber, D., Mohr, M., and Baltensperger, U.: Identification of the mass spectral signature of organic aerosols from wood burning emissions, Environ. Sci. Technol., 41, 5770–5777, 2007.

Allan, J. D., Delia, A. E., Coe, H., Bower, K. N., Alfarra, M. R., Jimenez, J. L., Middlebrook, A. M., Drewnick, F., Onasch, T. B., Canagaratna, M. R., Jayne, J. T., and Worsnopf, D. R.: A generalized method for the extraction of chemically resolved mass spectra from Aerodyne aerosol mass spectrometer data, J. Aerosol Sci., 35, 909–922, 2004.

Baek, S. O., Field, R. A., Goldstone, M. E., Kirk, P. W., Lester, J. N., and Perry, R.: A review of atmospheric polycyclic aromatic hydrocarbons: sources, fate and behavior, Water Air Soil Poll., 60, 279–300, 1991.

Bente, M., Sklorz, M., Streibel, T., and Zimmermann, R.: Thermal desorption-multiphoton ionization time-of-flight mass spectrometry of individual aerosol particles: a simplified approach for online single-particle analysis of polycyclic aromatic hydrocarbons and their derivatives, Anal. Chem., 81, 2525–2536, 2009.

Bressi, M., Sciare, J., Ghersi, V., Mihalopoulos, N., Petit, J.-E., Nicolas, J. B., Moukhtar, S., Rosso, A., Féron, A., Bonnaire1, N., Poulakis, E., and Theodosi, C.: Sources and geographical origins of fine aerosols in Paris (France), Atmos. Chem. Phys., 14, 8813–8839, 2014.

Bruns, E. A., Krapf, M., Orasche, J., Huang, Y., Zimmermann, R., Drinovec, L., Močnik, G., El-Haddad, I., Slowik, J. G., Dommen, J., Baltensperger, U., and Prévôt, A. S. H: Characterization of primary and secondary wood combustion products generated under different burner loads, Atmos. Chem. Phys., 15, 2825–2841, 2015.

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, 2007.

Canonaco, F., Crippa, M., Slowik, J. G., Baltensperger, U., and Prévôt, A. S. H.: SoFi, an IGOR-based interface for the efficient use of the generalized multilinear engine (ME-2) for the source apportionment: ME-2 application to aerosol mass spectrometer data, Atmos. Meas. Tech., 6, 3649–3661, 2013.

Cao, J. J., Wang, Q. Y., Chow, J. C., Watson, J. G., Tie, X. X., Shen, Z. X., Wang, P., and An, Z.: Impacts of aerosol compositions on visibility impairment in Xi'an, China, Atmos. Environ., 59, 559–566, 2012.

Chen, Y., Sheng, G., Bi, X., Feng, Y., Mai, B., and Fu, J.: Emission factors for carbonaceous particles and polycyclic aromatic hydrocarbons from residential coal combustion in China, Environ. Sci. Technol., 39, 1861-1867, 2005.

Chen, S. C. and Liao, C. M.: Health risk assessment on human exposed to environmental polycyclic aromatic hydrocarbons pollution sources, Sci. Total Environ, 366, 112–123, 2006.

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

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

Dall’Osto, M., Ovadnevaite, J., Ceburnis, D., Martin, D., Healy, R. M., O’Connor, I. P., Kourtchev, I., Sodeau, J. R., Wenger, J. C., and O’Dowd, C.: Characterization of urban aerosol in Cork city (Ireland) using aerosol mass spectrometry, Atmos. Chem. Phys., 13, 4997–5015, doi:10.5194/acp-13-4997-2013, 2013.

DeCarlo, P. F., Kimmel, J. R., Trimborn, A., Northway, M. J., Jayne, J. T., Aiken, A. C., Gonin, M., Fuhrer, K., Horvath, T., Docherty, K. S., Worsnop, D. R., and Jimenez, J. L.:

Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer, Anal. Chem., 78, 8281-8289, 2006.

DeWitt, H. L., Hellebust, S., Temime-Roussel, B., Ravier, S., Polo, L., Jacob, V., Buisson, C., Charron, A., André, M., Pasquier, A., Besombes, J. L., Jaffrezo, J. L., Wortham, H., and Marchand, N.: Near-highway aerosol and gas-phase measurements in a high-diesel environment, Atmos. Chem. Phys., 15, 4373-4387, 2015.

Draxler, R. and Hess, G. D.: An overview of the HYSPLIT 4 modeling system for trajectories, dispersion and deposition, Aust. Meteor. Mag., 47, 295–308, 1998.

Drewnick, F., Hings, S. S., DeCarlo, P., 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.

Drinovec, L., Močnik, G., Zotter, P., Prévôt, A. S. H., Ruckstuhl, C., Coz, E., Rupakheti, M., Sciare, J., Müller, T., Wiedensohler, A., and Hansen, A. D. A.: The “dual-spot” Aethalometer: an improved measurement of aerosol black carbon with real-time loading compensation, Atmos. Meas. Tech., 8, 43–55, 2015.

Dzepina, K., Arey, J., Marr, L. C., Worsnop, D. R., Salcedo, D., Zhang, Q., Onasch, T. B., Molina, L. T., Molina, M. J., and Jimenez, J. L.: Detection of particle-phase polycyclic aromatic hydrocarbons in Mexico City using an aerosol mass spectrometer, Int. J. Mass Spectrom., 263, 152–170, 2007.

El Haddad, I., D'Anna, B., Temime-Roussel, B., Nicolas, M., Boreave, A., Favez, O., Voisin, D., Sciare, J., George, C., Jaffrezo, J.-L., Wortham, H., and Marchand, N.: Towards a better understanding of the origins, chemical composition and aging of oxygenated organic aerosols: case study of a Mediterranean industrialized environment, Marseille, Atmos. Chem. Phys., 13, 7875-7894, doi:10.5194/acp-13-7875-2013, 2013.

Eriksson, A. C., Nordin, E. Z., Nyström, R., Pettersson, E., Swietlicki, E., Bergvall, C., Westerholm, R., Boman, C., and Pagels, J. H.: Particulate PAH emissions from residential biomass combustion: time-resolved analysis with aerosol mass spectrometry, Environ. Sci. Technol., 2014, 48, 7143–7150, 2014.

Gao, J. J., Tian, H. Z., Cheng, K., Lu, L., Zheng, M., Wang, S. X., Hao, J. M., Wang, K., Hua, S. B., Zhu, C. Y., and Wang, Y: The variation of chemical characteristics of PM2.5 and

PM10 and formation causes during two haze pollution events in urban Beijing, China, Atmos. Environ., 107, 1–8, 2015.

Gao, M., Guttikunda, S. K., Carmichael, G. R., Wang, Y., Liu, Z., Stanier, C. O., Saide, P. E., and Yu, M.: Health impacts and economic losses assessment of the 2013 severe haze event in Beijing area, Sci. Total Environ., 511, 553–561, 2015.

Gentner, D. R., Jathar, S. H., Gordon, T. D., Bahreini, R., Day, D. A., El Haddad, I., Hayes, P. L., Platt, S. M., Pieber, S. M., de Gouw, J., Goldstein, A. H., Harley, R. A., Jimenez, J. L., Prévôt, A. S. H., and Robinson, A. L.: Contribution of gasoline and diesel motor vehicles to secondary organic aerosol formation in urban areas worldwide: a review, In prep.

Gilardoni, S., Vignati, E., Cavalli, F., Putaud, J. P., Larsen, B. R., Karl, M., Stenstrom, K., Genberg, J., Henne, S., and Dentener, F.: Better constraints on sources of carbonaceous aerosols using a combined ^{14}C – macro tracer analysis in a European rural background site, Atmos. Chem. Phys., 11, 5685–5700, 2011.

He, L.-Y., Lin, Y., Huang, X.-F., Guo, S., Xue, L., Su, Q., Hu, M., Luan, S.-J., and Zhang, Y.-H.: Characterization of high-resolution aerosol mass spectra of primary organic aerosol emissions from Chinese cooking and biomass burning, Atmos. Chem. Phys., 10, 11535–11543, 2010.

Ho, K. F., Huang, R.-J., Kawamura, K., Tachibana, E., Lee, S. C., Ho, S. S. H., Zhu, T., and Tian, L.: Dicarboxylic acids, ketocarboxylic acids, α-dicarbonyls, fatty acids and benzoic acid in PM2.5 aerosol collected during CAREBeijing-2007: an effect of traffic restriction on air quality, Atmos. Chem. Phys., 15, 3111–3123, 2015.

Hu, D. and Jiang, J.: PM2.5 pollution and risk for lung cancer: a rising issue in China, J. Environ. Prot., 5, 731–738, 2014.

Huang, X. F., He, L. Y., Hu, M., Canagaratna, M. R., Sun, Y., Zhang, Q., Zhu, T., Xue, L., Zeng, L. W., Liu, X. G., Zhang, Y. H., Jayne, J. T., Ng, N. L., and Worsnop, D. R.: Highly time-resolved chemical characterization of atmospheric submicron particles during 2008 Beijing Olympic Games using an Aerodyne High-Resolution Aerosol Mass Spectrometer, Atmos. Chem. Phys., 10, 8933–8945, 2010.

Huang, X. F., He, L. Y., Xue, L., Sun, T. L., Zeng, L. W., Gong, Z. H., Hu, M., and Zhu, T.: Highly time-resolved chemical characterization of atmospheric fine particles during 2010 Shanghai World Expo, Atmos. Chem. Phys., 12, 4897–4907, 2012.

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

HYSPLIT4: Hybrid single-particle Lagrangian integrated trajectory model, NOAA Air Resources Laboratory: http://www.arl.noaa.gov/ready/hysplit4.html, last access: 25 May 2015.

Jayne, J. T., Leard, D. C., Zhang, X., Davidovits, P., Smith, K. A., Kolb, C. E., and Worsnop, D. R.: Development of an aerosol mass spectrometer for size and composition. Analysis of submicron particles, Aerosol Sci. Technol., 33, 49–70, 2000.

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 Prévôt, A. S. H.: Source attribution of submicron organic aerosols during wintertime inversions by advanced factor analysis of aerosol mass spectra, Environ. Sci. Technol., 42, 214–220, 2008.

Lanz, V. A., Prévôt, A. S. H., Alfarra, M. R., Weimer, S., Mohr, C., DeCarlo, P. F., Gianini, M. F. D., Hueglin, C., Schneider, J., Favez, O., D'Anna, B., George, C., and Baltensperger, U.: Characterization of aerosol chemical composition with aerosol mass spectrometry in Central Europe: an overview, Atmos. Chem. Phys., 10, 10453–10471, 2010.

Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D., and Pozzer, A.: The contribution of outdoor air pollution sources to premature mortality on a global scale, Nature, 525, 367–371, 2015.

Liu, P. S. K., Deng, R., Smith, K. A., Williams, L. R., Jayne, J. T., Canagaratna, M. R., Moore, K., Onasch, T. B., Worsnop, D. R., and Deshle, T.: Transmission efficiency of an aerodynamic focusing lens system: comparison of model calculations and laboratory measurements for the Aerodyne aerosol mass spectrometer, Aerosol Sci. Technol., 41, 721–733, 2007.

Liu, Q., Sun, Y., Hu, B., Liu, Z., Akio, S., and Wang, Y: In situ measurement of PM1 organic aerosol in Beijing winter using a high-resolution aerosol mass spectrometer, Chin. Sci. Bull., 57, 819–826, 2012.

Liu, S., Russell, L. M., Sueper, D. T., and Onasch, T. B.: Organic particle types by single-particle measurements using a time-of-flight aerosol mass spectrometer coupled with a light scattering module, Atmos. Meas. Tech., 6, 187–197, 2013.

Middlebrook, A. M., Bahreini, R., Jimenez, J. L., and Canagaratna, M. R.: Evaluation of composition-dependent collection efficiencies for the Aerodyne aerosol mass spectrometer using field data, Aerosol Sci. Technol., 46, 258–271, 2012.

Murray, C. J., Ezzati, M., Flaxman, A. D., Lim, S., Lozano, R., Michaud, C., Naghavi, M., Salomon, J. A., Shibuya, K., Vos, T., Wikler, D., and Lopez, A. D.: GBD 2010: design, definitions, and metrics, Lancet, 380, 2063–2066, 2013.

Ng, N. L., Canagaratna, M. R., Jimenez, J. L., Zhang, Q., Ulbrich, I. M., and Worsnop, D. R.: Real-time methods for estimating organic component mass concentrations from aerosol mass spectrometer data, Environ. Sci. Technol., 45, 910–916, 2011.

Linstrom, P.J., and Mallard, W.G. (Eds.): NIST Chemistry WebBook, NIST standard reference database number 69, National Institute of Standards and Technology, Gaithersburg MD, 20899, http://webbook.nist.gov, last access 2 February 2016.

Paatero, P.: Least squares formulation of robust non-negative factor analysis, Chemom. Intell. Lab. Syst., 37, 23–35, 1997.

Paatero, P. and Hopke, P. K.: Discarding or downweighting highnoise variables in factor analytic models, Anal. Chim. Acta, 490, 277–289, 2003.

Paatero, P. and Hopke, P. K.: Rotational tools for factor analytic models, J. Chemometrics, 23, 91–100, 2009.

Polissar, A. V., Hopke, P. K., Paatero, P., Kaufmann, Y. J., Hall, D. K., Bodhaine, B. A., Dutton, E. G., and Harris, J. M.: The aerosol at Barrow, Alaska: long-term trends and source locations, Atmos. Environ., 33, 2441–2458, 1999.

Rondia, D., Cooke, M., and Haroz, R. K. (Eds.): Mobile source emissions including policyclic organic species, NATO ASI Series, 112, D. Reidel Publishing Company, Liège, Belgium, 1983.

Sandradewi, J., Prévôt, A. S. H., Szidat, S., Perron, N., Alfarra, M. R., Lanz, V. A., Weingartner, E., and Baltensperger, U.: Using aerosol light absorption measurements for the quantitative determination of wood burning and traffic emission contributions to particulate matter, Environ. Sci. Technol., 42, 3316–3323, 2008.

Schoeny, R., and Poirier. K.: Provisional guidance for quantitative risk assessment of polycyclic aromatic hydrocarbons. U.S. Environmental Protection Agency, http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=49732, last access: 11 May 2015, 1993.

Sun, Y. L, Zhuang, G., Tang, A., Wang, Y., and An, Z.: Chemical characteristics of PM2.5 and PM10 in haze-fog episodes in Beijing, Environ. Sci. Technol., 40, 3148–3155, 2006.

Sun, Y. L., Zhang, Q., Schwab, J. J., Demerjian, K. L., Chen, W. N., Bae, M.-S., Hung, H.-M, Hogrefe, O., Frank, B., Rattigan, O.V., and Lin, Y.-C.: Characterization of the sources and processes of organic and inorganic aerosols in New York City with a high-resolution time-of-flight aerosol mass spectrometer, Atmos. Chem. Phys., 11, 1581–1602, 2011.

Sun, Y. L., Wang, Z. F., Fu, P. Q., Jiang, Q., Yang, T., Li, J., and Ge, X.: The impact of relative humidity on aerosol composition and evolution processes during wintertime in Beijing, China, Atmos. Environ., 77, 927–934, 2013.

Sun, Y. L., Wang, Z. F., Fu, P. Q., Yang, T., Jiang, Q., Dong, H. B., Li, J., and Jia, J. J.: Aerosol composition, sources and processes during wintertime in Beijing, China, Atmos. Chem. Phys., 13, 4577–4592, 2013b.

Sun, Y., Jiang, Q., Wang, Z., Fu, P., Li, J., Yang, T., and Yin, Y.: Investigation of the sources and evolution processes of severe haze pollution in Beijing in January 2013, J. Geophys. Res. Atmos., 119, 4380–4398, 2014.

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, 2009.

Wang, H., Dwyer-Lindgren, L., Lofgren K. T., Knoll Rajaratnam, J., Marcus, J. R., Levin-Rector, A., Levitz, C. E., Lopez, A. D., and Murray C. J. L.: Age-specific and sex-specific mortality in 187 countries, 1970–2010: a systematic analysis for the Global Burden of Disease Study, Lancet, 380, 2071–2094, 2010.

Wang, P., Cao, J. J., Shen, Z. X., Han, Y. M., Lee, S. C., Huang, Y., Zhu, C. S., Wang, Q. Y., Xu, H. M., and Huang, R.-J.: Spatial and seasonal variations of PM2.5 mass and species during 2010 in Xi'an, China, Sci. Total Environ., 508, 477–487, 2015.

Wang, Q., Chen, X., He, G. L., Lin, S. B., Liu, Z., and Xu, D. Q.: Study on characteristics of elements in PM2.5 during haze-fog weather in winter in urban Beijing, Spectrosc. Spect. Anal., 33, 1441–1445, 2013.

Wang, Y. H., Liu, Z. R., Zhang, J. K., Hu, B., Ji, D. S., Yu, Y. C., and Wang, Y. S.: Aerosol physicochemical properties and implications for visibility during an intense haze episode during winter in Beijing, Atmos. Chem. Phys., 15, 3205–3215, 2015.

Wei, C., Han, Y. M., Bandowe, B. A. M., Cao, J. J., Huang, R.-J., Ni, H. Y., Tian, J., and Wilcke, W.: Occurrence, gas/particle partitioning and carcinogenic risk of polycyclic aromatic hydrocarbons and their oxygen and nitrogen containing derivatives in Xi’an, central China, Sci. Total Environ., 505, 814–822, 2015.

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

Williams, L. R., Gonzalez, L. A., Peck, J., Trimborn, D., McInnis, J., Farrar, M. R., Moore, K. D., Jayne, J. T., Robinson, W. A., Lewis, D. K., Onasch, T. B., Canagaratna, M. R., Trimborn, A., Timko, M. T., Magoon, G., Deng, R., Tang, D., de la Rosa Blanco, E., Prévôt, A. S. H., Smith, K. A., and Worsnop, D. R.: Characterization of an aerodynamic lens for transmitting particles greater than 1 micrometer in diameter into the Aerodyne aerosol mass spectrometer, Atmos. Meas. Tech,, 6, 3271–3280, 2013.

Wolf, R.: Extended Source Apportionment of Organic Aerosols, Unpublished PhD thesis, Swiss Federal Institute of Technology in Zurich, 2014.

Xu, S. S., Liu, W. X., and Tao, S.: Emission of polycyclic aromatic hydrocarbons in China, Environ. Sci. Technol., 40, 702–708, 2006.

Xu, Z. J., Wen, T. X., Li, X. R., Wang, J. G., and Wang, Y. S.: Characteristics of carbonaceous aerosols in Beijing based on two-year observation, Atmos. Pollut. Res., 6, 202–208, 2015.

Yang, M., Howell, S. G., Zhuang, J., and Huebert, B. J.: Attribution of aerosol light absorption to black carbon, brown carbon, and dust in China – interpretations of atmospheric measurements during EAST-AIRE, Atmos. Chem. Phys., 9, 2035–2050, 2009.

Yang, Y., Liu, X., Qu., Y., Wang, J., An, J., Zhang, Y., Zhang, F.: Formation mechanism of continuous extreme haze episodes in the megacity Beijing, China, in January 2013, Atmos. Res., 155, 192–203, 2015.

Zhang, Y., Schauer, J. J., Zhang Y., Zeng, L., Wei, Y., Liu, Y., and Shao, M.: Characteristics of particulate carbon emissions from real-world Chinese coal combustion, Environ. Sci. Technol., 42, 5068–5073, 2008.

Zhang, Y. and Tao, S.: Global atmospheric emission inventory of polycyclic aromatic hydrocarbons (PAHs) for 2004, Atmos. Environ., 43, 812–819, 2009.

Zhang, J. K., Sun, Y., Liu, Z. R., Ji, D. S., Hu, B., Liu, Q., and Wang, Y. S.: Characterization of submicron aerosols during a month of serious pollution in Beijing, 2013, Atmos. Chem. Phys., 14, 2887–2903, 2014.

Zhang, Q., Shen, Z. X., Cao, J. J., Zhang, R. J., Zhang, L. M., Huang, R.-J., Zheng, C. J., Wang, L. Q., Liu, S. X., Xu. H. M.: Variations in PM2.5, TSP, BC, and trace gases (NO_2, SO_2, and O_3) between haze and non-haze episodes in winter over Xi’an, China, Atmos. Environ., 112, 64–71, 2015a.

Zhang, Q., Quan, J., Tie, X., Li, X., Liu, Q., Gao, Y., and Zhao, D.: Effects of meteorology and secondary particle formation on visibility during heavy haze events in Beijing, China, Sci. Total Environ., 502, 578–584, 2015b.

Zhou, R., Wang, S., Shi, C., Wang, W., Zhao, H., Liu, R., Chen, L., and Zhou, B.: Study on the traffic air pollution inside and outside a road tunnel in Shanghai, China, PLoS One, 9, e112195, 2014.

Zotter, P., Ciobanu, V. G., Zhang, Y. L., El-Haddad, I., Macchia, M., Daellenbach, K. R., Salazar, G. A., Huang, R. J., Wacker, L., Hueglin, C., Piazzalunga, A., Fermo, P., Schwikowski, M., Baltensperger, U., Szidat, S., and Prévôt, A. S. H.: Radiocarbon analysis of elemental and organic carbon in Switzerland during winter-smog episodes from 2008 to 2012 – Part 1: Source apportionment and spatial variability, Atmos. Chem. Phys., 14, 13551–13570, 2014.

Zotter, P., Herich, H., Gysel, M., El-Haddad, I., Zhang, Y.L., Močnik, G., Hüglin, C., Baltensperger, U., Szidat, S., and Prévôt, A. S. H.: Evaluation of the absorption Ångström exponents for traffic and wood burning in the Aethalometer based source apportionment using radiocarbon measurements of ambient aerosol, In preparation.

Figures

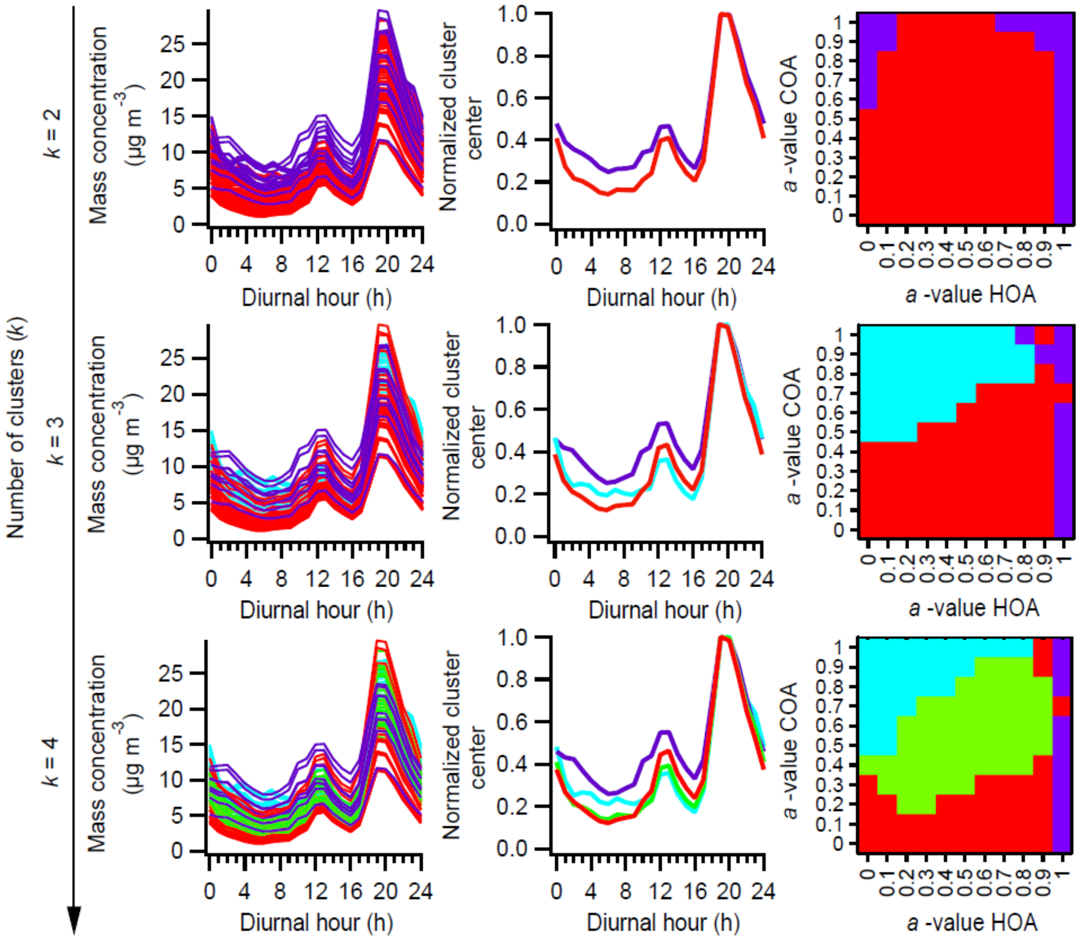

Figure 1. Cluster analysis on the normalized COA diurnal trends from the evaluated *a* value space. Results from two-, three- and four-cluster solutions are presented (from top to bottom), using different colors to represent the clusters. For each solution the left plot represents all COA diurnals colored with the color of the corresponding cluster, the center plot represents the normalized diurnals of the cluster centers and right plot represents the cluster assignment for all considered *a* value combinations.

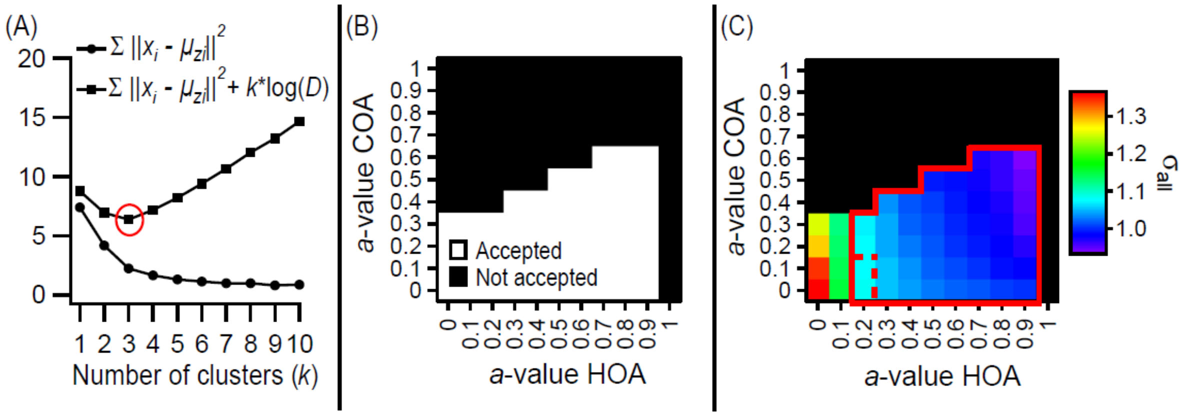

Figure 2. (A) The first term T1 (circles) and the total cost function CF (T1+T2) (squares) defined in Eq. (6) for increasing number of clusters (*k*); (B) Final accepted *a* value range after applying the criteria 1 and 2; (C) Color coded quadratic sum of standard deviations between external tracers and correlated primary OA sources. The region surrounded with red line includes all final accepted *a* value combinations obtained from the study of the variability of the solution among 50 solutions with pseudo-randomly modified input matrix within twice its errors ($OA \pm 2OA_{error}$). Dashed line in left lower corner shows the two *a* value combinations that would not have been considered as good solutions if smaller changes are allowed to the input matrix ($OA(i,j) \pm 1OA_{error}(i,j)$) and only 10 solutions are considered for the variability study.

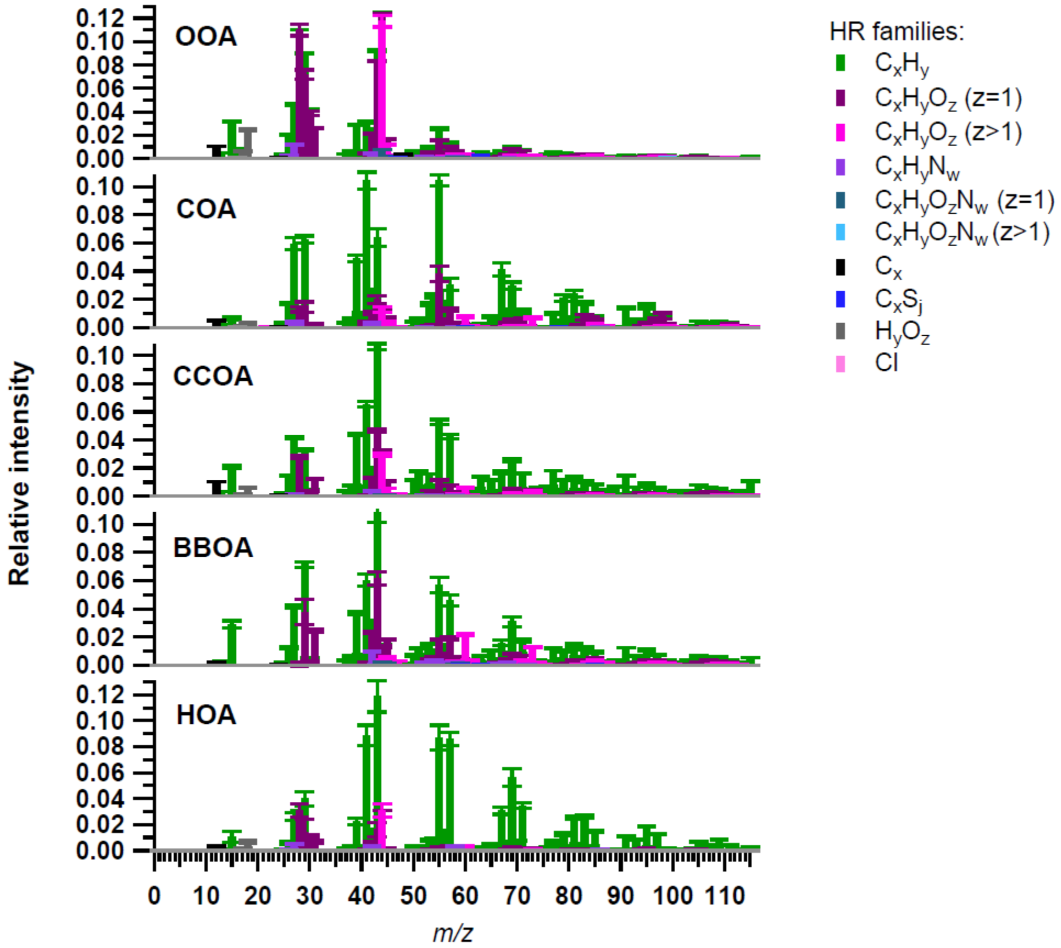

Figure 3. Mass spectra of the five identified OA factors color coded with the chemical families; Spectra are averaged over all good *a* value combinations and error bars represent the standard deviation of each *m/z* over all the accepted solutions.

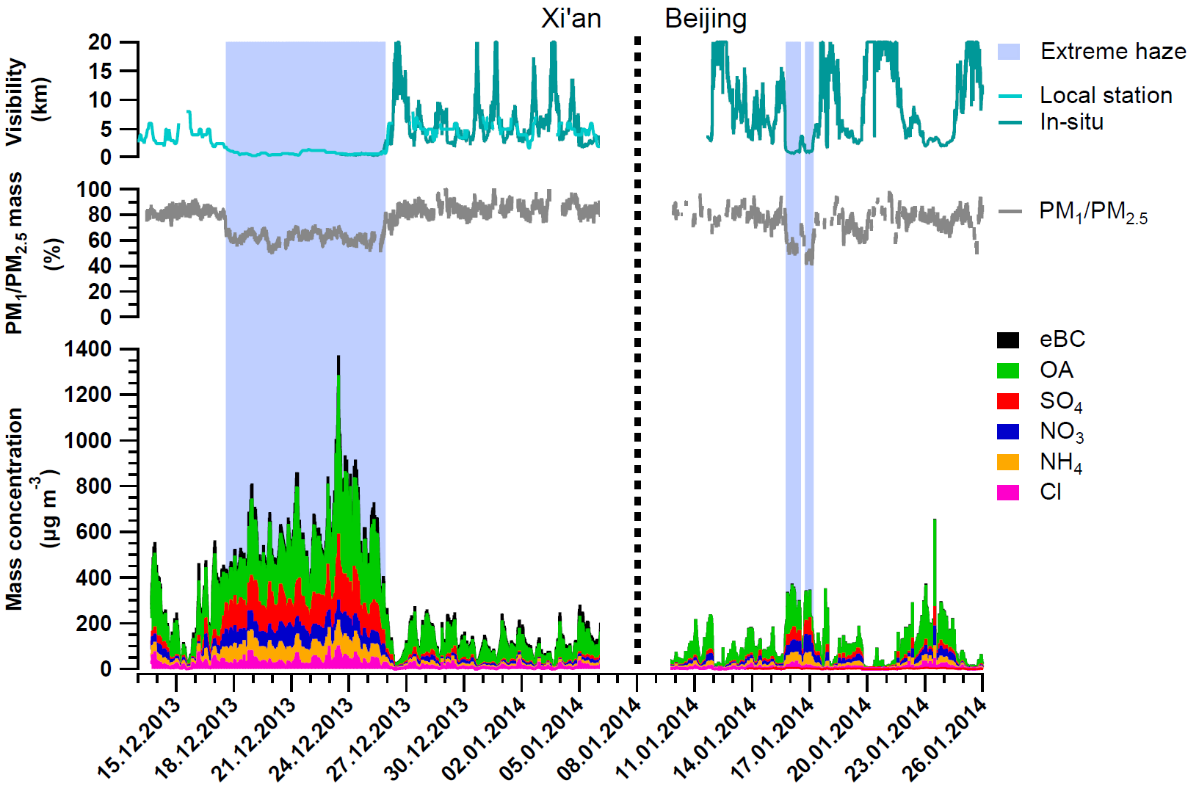

Figure 4. Bottom panel: Time series of the AMS species and eBC mass concentrations for the full measurement period; Medium panel: Ratio between the NR mass that would have been measured using a PM_1 lens and the total mass measured with the $PM_{2.5}$ lens; Top panel: Measured visibility.

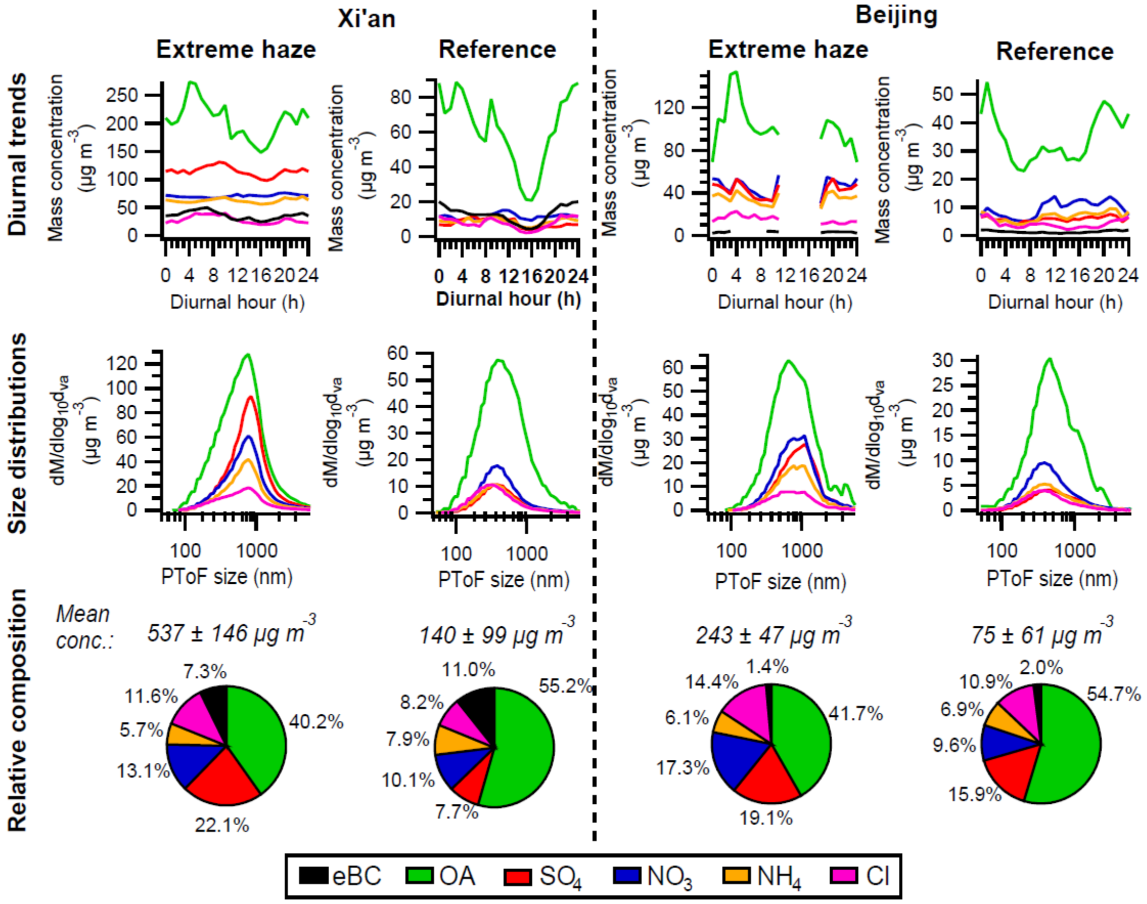

Figure 5. Diurnal trends, size distributions, mean concentrations (NR-PM2.5 plus eBC mass) and relative contributions of the AMS species and eBC for the four periods. Note: Size distributions only available for AMS species.

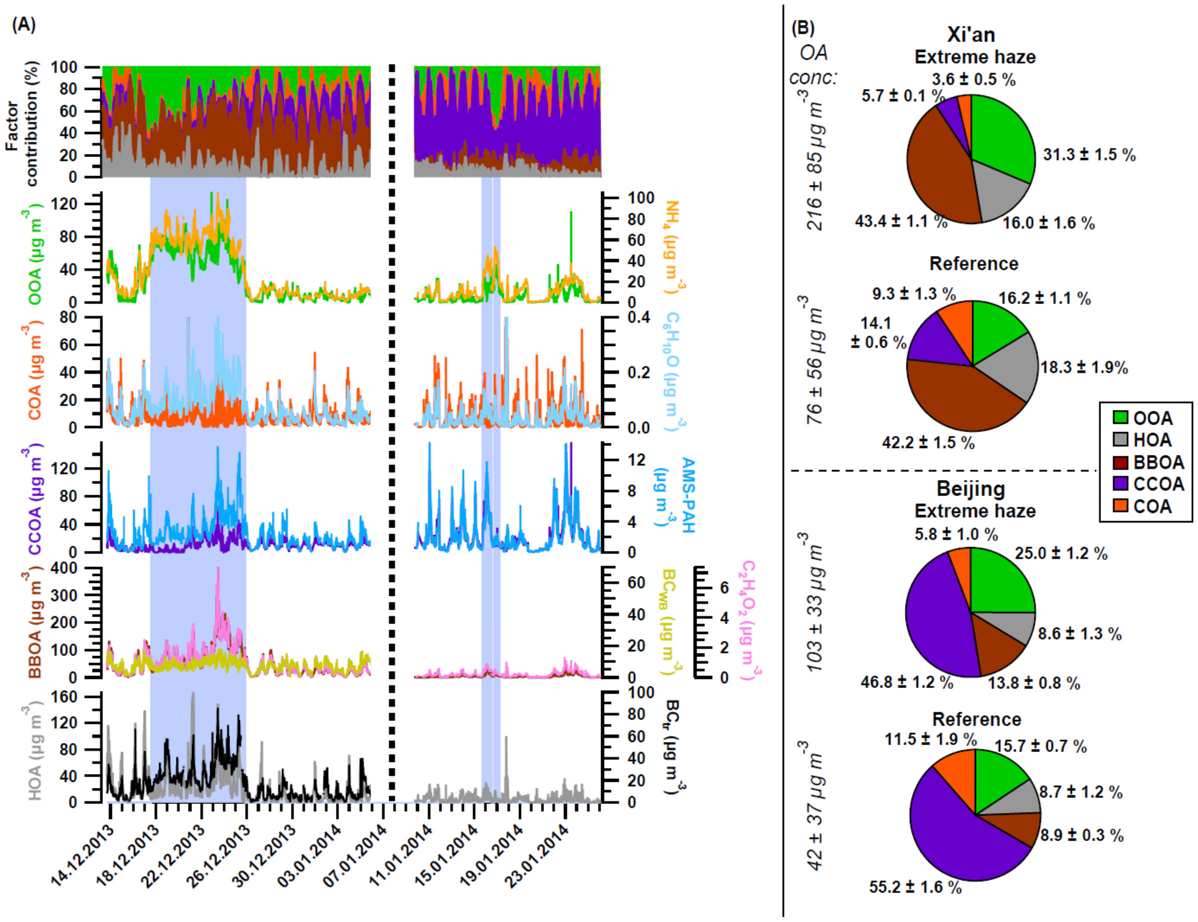

Figure 6. A) Time series of the OA sources and the external tracers and relative contribution of the different sources over time; B) Mean relative contributions of the OA sources for the four periods of interest. Errors represent the standard deviation among all good *a* value combinations.

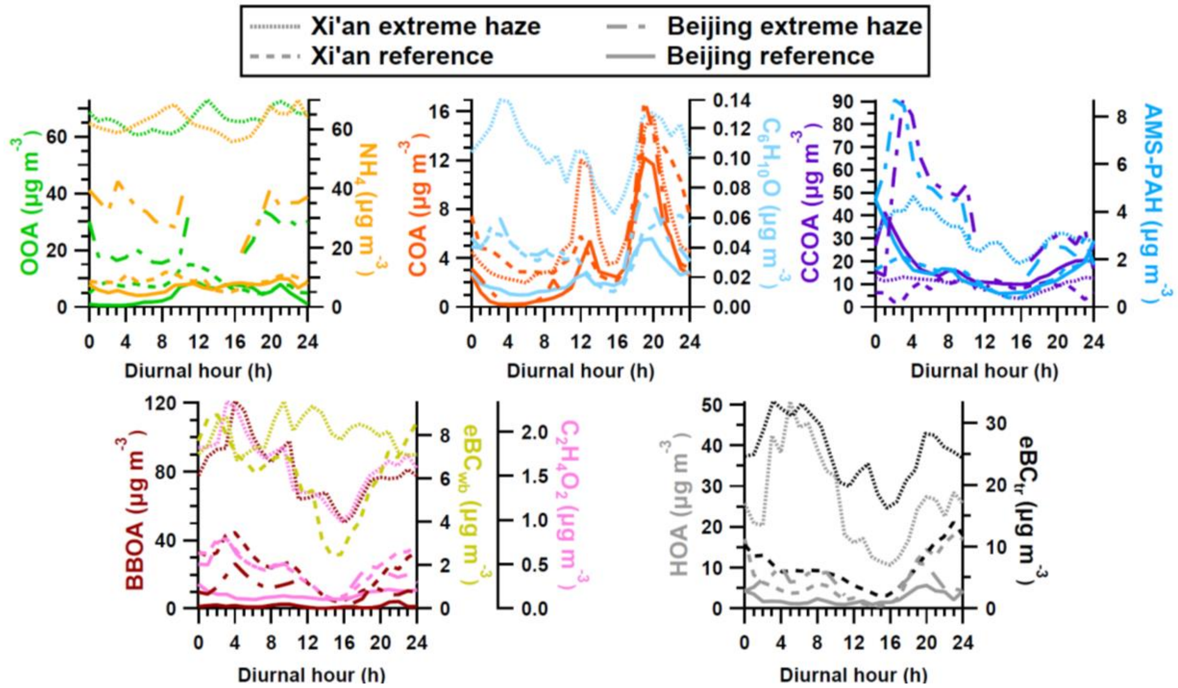

Figure 7. Median daily pattern (absolute concentrations) of the OA sources and external tracers for the four periods of interest.

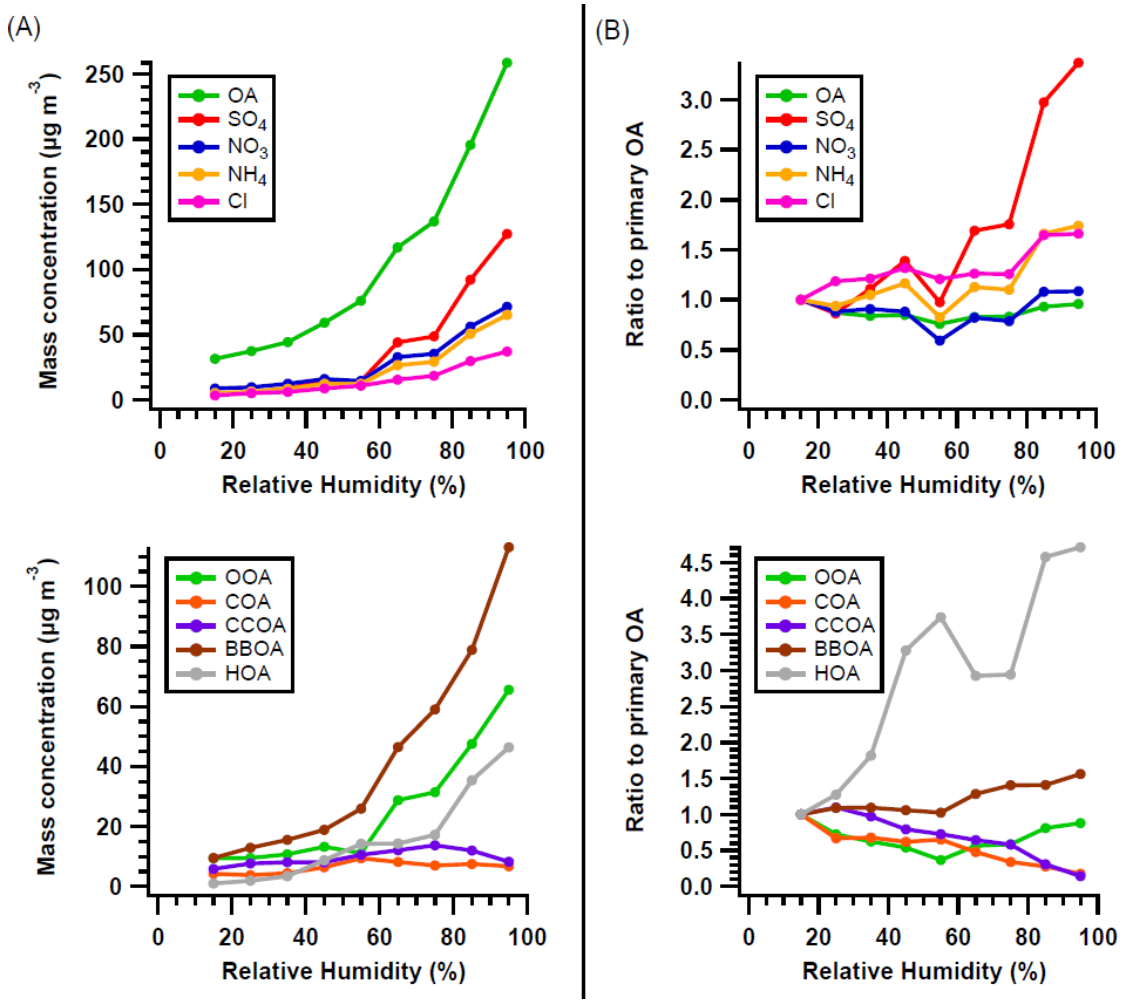

Figure 8. (A) Average mass concentration of the organic and inorganic species (top) and of the different OA sources (bottom) as a function of RH. (B) Average mass concentration of the OA and inorganic species (top) and of the OA sources (bottom) normalized to the sum of all primary sources (HOA, BBOA, CCOA and COA) and represented as a function of RH. For better representation of the trends, all components were further normalized to their corresponding value in the first RH bin. Note: RH bins of 10 % in all cases.

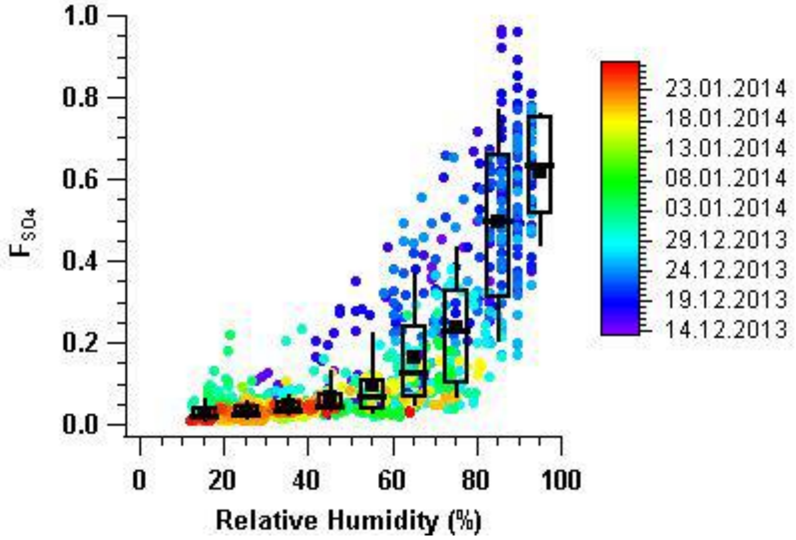

Figure 9. Sulfur oxidation ratio (F_{SO4}) as a function of RH, color coded by date. Mean (square), median (middle horizontal line), 25th and 75th percentiles (P25-P75, box) and 10th and 90th percentiles (P10-P90, whiskers) are reported for each RH bin (10 % step).

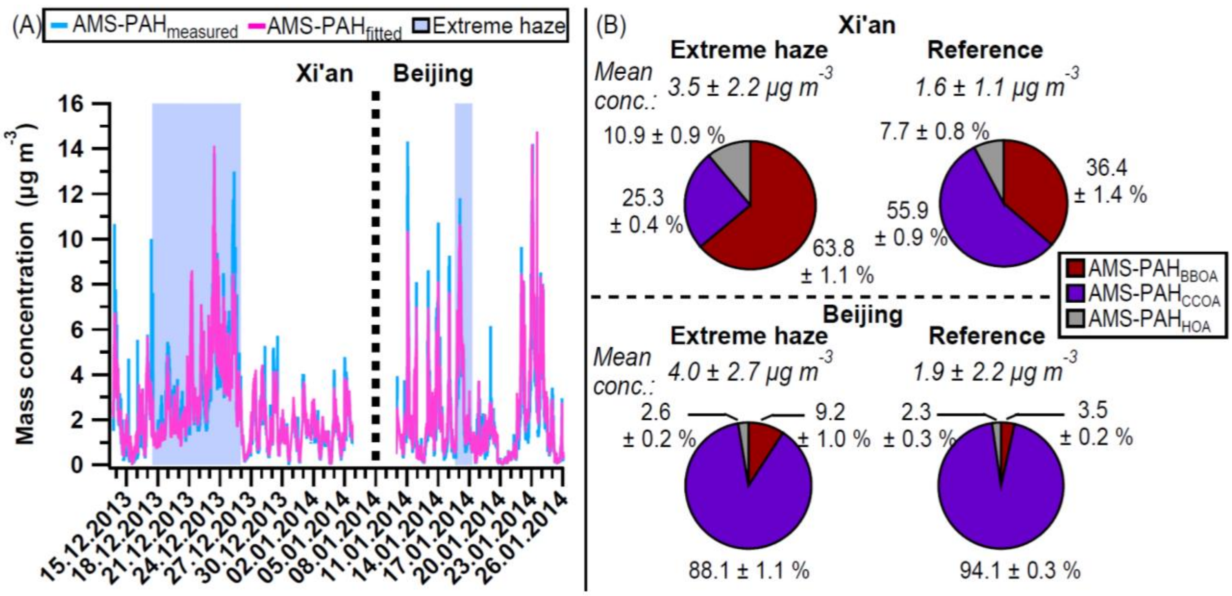

Figure 10. (A) Time series of the measured (blue) and fitted AMS-PAH (pink); (B) Average AMS-PAH concentrations and relative contributions of the different sources to the measured AMS-PAHs for the four periods of interest. Errors represent the standard deviation among all good *a* value combinations.