Chemical Characteristics of Organic Aerosols in Shanghai: A Study by Ultrahigh-Performance Liquid Chromatography Coupled With Orbitrap Mass Spectrometry

X. Wang, N. Hayeck, M. Brueggemann, L. Yao, H. Chen, C. Zhang, C. Emmelin, Jie Chen, C. George, L. Wang

Chemical Characteristics of Organic Aerosols in Shanghai: A Study by Ultra-High-Performance Liquid Chromatography Coupled with Orbitrap Mass Spectrometry

Xinke Wang[1], Nathalie Hayeck[2], Martin Brüggemann[2], Lei Yao[1], Hangfei Chen[1], Ci Zhang[1], Corinne Emmelin[2], Jianmin Chen[1,3], Christian George[2 *], Lin Wang[1,3 *]

[1] *Shanghai Key Laboratory of Atmospheric Particle Pollution and Prevention (LAP[3]), Department of Environmental Science & Engineering, Fudan University, Shanghai 200433, China*

[2] *Univ Lyon, Université Claude Bernard Lyon 1 CNRS, IRCELYON, F-69626, Villeurbanne, France*

[3] *Institute of Atmospheric Sciences, Fudan University, Shanghai 200433, China*

* *Corresponding Authors: C.G., email, Christian.George@ircelyon.univ-lyon1.fr; L.W., email, lin_wang@fudan.edu.cn.*

Key Points

(1) Characterize organic aerosols in Shanghai using UHPLC-Orbitrap MS and determine their potential sources.

(2) Variations in organic aerosol composition are observed among different months and between daytime and nighttime.

(3) Epoxide-intermediated routes and ammonia-carbonyl chemistry likely dominate the formation of OSs and nitrogen-containing species, respectively.

Abstract. $PM_{2.5}$ filter samples were collected in July and October 2014 and January and April 2015 in urban Shanghai, and analyzed using ultra-high-performance liquid chromatography (UHPLC) coupled with heated electrospray ionization (HESI) and Orbitrap mass spectrometry (MS). The measured chromatogram-mass spectra were processed by a non-target screening approach to identify significant signals. In total, 810-1510 chemical formulas of organic compounds in the negative polarity (ESI-) and 860-1790 in the positive polarity (ESI+), respectively, were determined. The chemical characteristics of organic aerosols (OAs) in Shanghai varied among different months and between daytime and nighttime. In the January samples, organics were generally richer in terms of both number and abundance, whereas those in the July samples were far lower. Organic compounds containing only carbon, hydrogen, and oxygen in addition to sulfur-containing organics detected by ESI- were found to be more abundant in the daytime samples, suggesting a photochemical source, whereas organic compounds containing both nitrogen and sulfur were more abundant in the nighttime samples, due to nocturnal nitrate radical chemistry. A significant number of mono- and polycyclic aromatic compounds, and nitrogen- and sulfur-containing heterocyclic compounds were detected in all samples, indicating that biomass burning and fossil fuel combustion made important contributions to the OAs in urban Shanghai. Additionally, precursor-product pair analysis indicates that the epoxide pathway was an important formation route for organosulfates observed in Shanghai. Moreover, a similar analysis suggests that 35-57% of nitrogen-containing compounds detected in ESI+ could be formed through reactions between ammonia and carbonyls. Our study presents a comprehensive view of OAs in urban Shanghai, which helps to understand their characteristics and sources.

1. Introduction

OAs often represent a main fraction (20%-90%) of submicron atmospheric particulate mass [*Jimenez et al.*, 2009; *Kroll and Seinfeld*, 2008] and play an important role in determining the climatic and health effects of atmospheric aerosols [*Pöschl*, 2005]. Atmospheric OAs can be divided according to their sources into two categories: primary organic aerosols (POAs) are directly emitted into the atmosphere from natural or anthropogenic sources, whereas secondary organic aerosols (SOAs) are formed in the atmosphere from precursor gases [*Seinfeld and Pankow*, 2003]. The complexity of OAs dramatically increases through multiple chemical reactions of complex primary organic compounds in the atmosphere [*Goldstein and Galbally*, 2007] and/or atmospheric ageing of OAs [*Jimenez et al.*, 2009; *Rudich et al.*, 2007]. Hence, elucidating the chemical composition of the organic fraction in the atmospheric particles at the molecular level is a vital but challenging task.

Traditional methods of OA analysis such as gas chromatography (or liquid chromatography) interfaced with mass spectrometry (GC/MS or LC/MS) can only identify a limited number of highly complex organic compounds in fine aerosol samples, because authentic standards are required to provide retention times in chromatograms and fragmentation patterns in mass spectra. In recent years, methods based on ultra-high-resolution mass spectrometry (UHRMS, e.g., Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) and Orbitrap mass spectrometry) coupled with soft ionization techniques (e.g., electrospray ionization, ESI) have been introduced to deliver the exact molecular formulas of the organic components in aerosol particles. The high mass resolution and high mass accuracy of UHRMS allows in many cases the determination of the elemental composition of unknown organic components without the need of authentic standards [*Nizkorodov et al.*, 2011].

Previously, ESI-UHRMS has been used to identify secondary organic compounds generated in laboratory experiments. Reinhardt et al. [2007] identified hundreds of organic compounds generated from α-pinene ozonolysis in a smog chamber using FTICR-MS. Later similar techniques were widely applied to investigate the chemical composition of SOAs generated from oxidation of isoprene, monoterpenes, diesel fuel,

etc., under different conditions [*Bateman et al.*, 2009; *Kourtchev et al.*, 2015; *Nguyen et al.*, 2010; *Nguyen et al.*, 2011; *Romonosky et al.*, 2017; *Walser et al.*, 2008].

Aerosol samples collected in field campaigns have been analyzed using ESI-HRMS with the objectives to obtain a comprehensive elemental composition of ambient OAs and their plausible sources [*Altieri et al.*, 2012; *Kourtchev et al.*, 2013; *Lin et al.*, 2012a; *Lin et al.*, 2012b; *O'Brien et al.*, 2013; *Rincón et al.*, 2012; *Roach et al.*, 2010; *Tao et al.*, 2014]. For example, Rincón et al. [2012] identified thousands of organic compounds in the ambient aerosols from Cambridge, UK and showed that summer samples generally contained more organic components than winter samples. The variation in aerosol composition at the molecular level can be utilized to elucidate the potential sources of ambient aerosols. O'Brien et al. [2014] observed clear diurnal variations of OAs in Bakersfield, CA, USA and presented evidence of local and long range transport of SOAs from both biogenic and anthropogenic sources. Later this technique has been extended to characterize the elemental composition of organic components in aerosols emitted from a specific source [*Kourtchev et al.*, 2016; *Laskin et al.*, 2009; *Tong et al.*, 2016]. For example, Tong et al. [2016] studied the composition of fine particles at a road tunnel (Queensway) in Birmingham, UK, and concluded that many oxidized particulate mono-aromatic, polycyclic aromatic hydrocarbons, and nitrooxy-aromatics originated from vehicle emissions.

Recently, ESI-UHRMS has been combined with LC in a number of studies [*Lin et al.*, 2016; *Vogel et al.*, 2016a; *Vogel et al.*, 2016b; *Wang et al.*, 2016], which allows separation of isomers and, hence, offers an additional dimension of information in comparison with previous studies using solely direct injection ESI-UHRMS. Another advantage of LC separation is to reduce ion suppression during ionization and to achieve semi-quantification of organic compounds. For example, Wang et al. [2016] identified about 200 formulas of particulate organosulfates (OSs) and dozens of formulas of nitrooxy-OSs, with various numbers of isomers for each determined formula using UHPLC coupled to ESI-Orbitrap MS. It was shown that the abundances of nitrooxy-OSs in the nighttime samples were larger than those in the daytime samples

in Shanghai, highlighting the formation of nitrooxy-OSs from the NO_3 nighttime chemistry.

Shanghai is a megacity with extensive emissions of anthropogenic pollutants including particulate matters, volatile organic compounds (VOCs), sulfur dioxide, and nitrogen oxides [*Huang et al.*, 2011; *Wang et al.*, 2013]. In previous studies, ESI-UHRMS has been used to determine hundreds of OSs in atmospheric aerosol samples in Shanghai and elucidate their seasonal and diurnal variations [*Tao et al.*, 2014; *Wang et al.*, 2016]. Nonetheless, other classes of organic components such as highly oxidized molecules (HOMs) that likely possess extremely low vapor pressure leading to a high mass fraction in the condensed phase [*Molteni et al.*, 2016; *Rissanen et al.*, 2014] remain elusive. In this study, we determined chemical formulas of compounds containing only carbon, hydrogen, and oxygen (CHO), sulfur-containing organics (CHOS), nitrogen-containing organics (CHON), compounds containing both nitrogen and sulfur (CHONS), and organics without oxygen (CHNS* including CHN (containing only carbon, hydrogen, and nitrogen), CHS (containing only carbon, hydrogen, and sulfur) and CHNS (containing carbon, hydrogen, nitrogen and sulfur)) in aerosol particles collected in Shanghai, using UHPLC coupled to Orbitrap MS. All species detected by ESI- were labeled with "-" (e.g., CHO-, CHOS-, CHON-, CHONS- and CHNS*-) and those detected by ESI+ were labeled with "+" (e.g., CHO+, CHOS+, CHON+, CHONS+ and CHNS*+). In addition, the temporal variations of these species were characterized. Precursor-product pair analysis was utilized to elucidate the formation routes of certain classes of organic compounds. Furthermore, potential sources for OAs in Shanghai will be discussed. Therefore, compared with the previous studies [*Tao et al.*, 2014; *Wang et al.*, 2016], this study presents a more comprehensive view of OAs in urban Shanghai.

2. Material and methods

2.1 Sample Collection

Twelve-hour ambient aerosol samples were collected onto 90 mm prebaked quartz-

fiber filters (Whatman Company, UK) during 26 to 30 July 2014 (sample ID: July daytime samples, JUD; July nighttime samples, JUN), 19 to 23 October 2014 (sample ID: October daytime samples, OCD; October nighttime samples, OCN), 9 to 11 and 15 to 16 January 2015 (sample ID: January daytime samples, JAD; January nighttime samples, JAN), and 17 to 21 April 2015 (sample ID: April daytime samples, APD; April nighttime samples, APN), using a middle-flow impact aerosol sampler (Qingdao Hengyuan Tech Co., Ltd., HY-100) operating at 100 L min^{-1}. Daytime samples were collected between 8 am-8 pm local time, whereas nighttime samples were collected between 8 pm-8 am the next day. The sampling site was located on the rooftop of a teaching building at Fudan University (31°18′N, 121°30′E), about 20 m above ground with surrounding residential and commercial properties and a major highway to the south of the site [*Xiao et al.*, 2015; *Ma et al.*, 2014]. Blank samples were taken following the same procedure except that no air was drawn through the filter substrate. After sample collection, filters were stored at -20 °C in a freezer before further analysis. Table S1 provides a comparison of air quality and meteorological conditions during the sampling days in Shanghai.

2.2 Sample analysis

A quarter of each quartz filter was extracted twice with 6 mL of acetonitrile (Optima® LC/MS, Fischer Scientific, USA) and agitated for 20 min on an orbital shaker set at 1000 rpm. The combined extracts were filtered through a 0.2 μm PTFE membrane (13 mm, Pall Corporation, USA) using a glass syringe, and then blown to almost dryness under a gentle stream of N_2. The extracts were then reconstituted in 1 mL of a 1:1 v/v mixture of water (Optima® LC/MS, Fischer Scientific, USA) and acetonitrile (Optima® LC/MS, Fischer Scientific, USA). For the analysis, 200 μL of the reconstituted extract was diluted by adding 100 μL of water. Five μL of the diluted solution was analyzed by UHPLC (Dionex 3000, Thermo Scientific, USA) coupled to a Q-Exactive Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo scientific, USA). Three replicate analyses were performed for each sample.

In addition, pentafluorobenzylhydroxylamine derivatization was used to identify organic compounds with carbonyl functional groups [*Borrás and Tortajada-Genaro*, 2012]. A volume of 200 µL of the reconstituted extract were mixed with 800 µL of o-(2, 3, 4, 5, 6-Pentafluorobenzyl) hydroxylamine hydrochloride (PFBHA, Sigma Aldrich, ≥ 99.0%) solutions (1mg/mL). The mixture was left in darkness at room temperature for 24 hours, then analyzed using the same procedure as previously described.

Analytes were separated using a Waters Acquity HSS T3 column (1.8 µm, 100×2.1 mm) with mobile phases consisting of (A) 0.1% formic acid in water (Optima® LC/MS, Fischer Scientific, USA) and (B) 0.1% formic acid in acetonitrile (Optima® LC/MS, Fischer Scientific, USA). Gradient elution was performed by the A/B mixture at a total flow rate of 300 µL/min: 1% B for 2 min, a linear gradient to 100% B in the next 11 min, 100% B for another 2 min, back to 1% B in 0.1 min, and then 1% B for 6.9 min. The Q-Exactive Hybrid Quadrupole-Orbitrap mass spectrometer was equipped with a heated electrospray ionization source, using a spray voltage of -2.6 and 3.2 kV for ESI- and ESI+, respectively. The mass resolution was 140,000 at *m/z* 200 and the scanning range was set to *m/z* 50–750. The Q-Exactive mass spectrometer was externally mass calibrated daily using a 2 mM sodium acetate solution that provides a series of negative and positive adduct ions in the range of *m/z* 50–750.

2.3 Data processing

The obtained chromatogram-mass spectra were analyzed by a non-target approach software (MZmine 2.21), which provides the core functionality for MS data processing: raw data import, peak detection, shoulder peaks filtering, chromatogram building, chromatogram deconvolution, deisotoping, search of adducts and peak complexes, alignment, gap filling, identification, and duplicate peak filtering [*Hu et al.*, 2016; *Katajamaa et al.*, 2006; *Pluskal et al.*, 2010]. The detailed processing steps and settings can be found in the Supplementary Information. The output of MZmine 2.21 includes the *m/z* ratios, formulas, retention times, and abundances of detected organic compounds. Molecular formulas were expressed as $C_cH_hO_oN_nS_s$, where c, h, o, n, and

s correspond to the numbers of carbon, hydrogen, oxygen, nitrogen, and sulfur atoms in the molecular formula, respectively. Spectra of blank samples were processed using the same settings.

For the chromatogram-mass spectra analysis, the abundance of a compound refers to the average area of its chromatographic peak from the three repetitions and the number of isomers for a compound is based on the number of chromatographic peaks observed for a given *m/z* value. The obtained number of isomers may vary significantly when the separation method is further optimized [*Wang et al.*, 2016]. Also, the defective shapes of some peaks may lead to uncertainties in the number of isomers. The compound list for the blank samples was compared to those for ambient samples, and only compounds with a sample-to-blank abundance ratio ≥ 10 were retained. Subsequently the abundance of the retained compounds in the ambient samples was blank-corrected. Then, the arbitrary abundances of all isomers for a given formula were added up. The intensity variability due to instrumental performance was taken into account through daily analysis of standards.

Ring and double bond equivalence (RDBE) provides information on the number of rings and double bonds in the molecules. RDBE is usually calculated as:

$$\mathrm{RDBE} = \frac{2c+2+n-h}{2} \qquad \text{(Eq. 1)}$$

However, it should be noted that RDBE for compounds with heteroatoms (e.g., O, S) may not accurately indicate the level of unsaturation. For example, the two S=O double bonds in sulfate groups are not taken into account when calculating the unsaturation degree of OSs based on Eq. (1) [*Wang et al.*, 2016].

The aromaticity equivalent (Xc) has been suggested to improve identification and characterization of mono- and polycyclic aromatic compounds [*Tong et al.*, 2016; *Yassine et al.*, 2014]. Xc for compounds only containing carbon, hydrogen, nitrogen, oxygen, sulfur, and phosphorus can be calculated as follows:

$$\mathrm{Xc} = \frac{3(\mathrm{RDBE}-(p\times o+q\times s))-2}{\mathrm{RDBE}-(p\times o+q\times s)} \qquad \text{(Eq. 2)}$$

where p and q correspond to the fraction of oxygen and sulfur atoms, respectively, involved in π-bond structures of a compound that varies depending on the category of

a compound. For example, carboxylic acids, esters, and nitro compounds are characterized by p=q=0.5. For compounds containing carbonyl, nitroso, cyano, hydroxyl, or ether functionalities, p and q could be either 1 or 0. In this study, p=q=0.5 was used for compounds detected in ESI-, because ESI- is more sensitive to compounds containing carboxylic groups [*Kourtchev et al.*, 2016; *Tong et al.*, 2016]. However, p=q=1 was selected in ESI+, because compounds with a large diversity of functional groups can possibly be detected. The selection of these values for p and q is attempting to avoid an overestimation of the amount of mono- and polycyclic aromatics, however, it could have led to underestimations [*Yassine et al.*, 2014]. For molecular formulas with an odd number of oxygen or sulfur, the sum (p×o+q×s) in Eq. (2) was rounded to the next lower integer [*Yassine et al.*, 2014]. Xc ≥ 2.50 and Xc ≥ 2.71, respectively, are proposed as unambiguous minimum criteria for the presence of mono- and polycyclic aromatics.

Carbon oxidation state (OS_C) was introduced as a preferred metric for the degree of oxidation of organic species in the atmosphere, compared to the simple but more error-prone use of O/C ratios. For organic compounds composed of carbon, hydrogen and reduced oxygen, OS_C can be calculated through the following equation [*Kroll et al.*, 2011]:

$$\mathrm{OSc} \approx 2\mathrm{O/C} - \mathrm{H/C} \qquad \text{(Eq. 3)}$$

where O/C and H/C are the elemental ratio of oxygen-to-carbon and hydrogen-to-carbon, respectively.

The Kendrick mass defect (KMD) is very useful to differentiate a group of similar compounds among a large set of molecular formulas obtained by UHRMS [*Hughey et al.*, 2001; *Kendrick*, 1963]. In this study, CH_2 (14.00000) was chosen as a base unit, and Eq. 4 and Eq. 5 were used to calculate the Kendrick mass (KM_{CH2}) and the Kendrick mass defect (KMD_{CH2}), respectively:

$$\mathrm{KM_{CH_2}} = \mathrm{Observed\ \ Mass} \times \left(\frac{14.00000}{14.01565}\right) \qquad \text{(Eq. 4)}$$

$$\mathrm{KMD_{CH_2}} = \mathrm{Nominal\ \ Mass} - \mathrm{KM_{CH_2}} \qquad \text{(Eq. 5)}$$

where “Observed Mass” is the mass measured by the mass spectrometer, and “Nominal

Mass" is the rounded integer mass of a compound [*Wang et al.*, 2016].

3. Results and discussion

3.1 General Characteristics

A main emphasis of this study is to determine the broad and complex composition of organic compounds in ambient aerosol samples from Shanghai, and compare the characteristics among different months and between daytime and nighttime. Hence, the list of molecular formulas from each sample was compared with that from the blank sample as discussed above, and then inter-compared with those from samples collected in the same month. Only compounds observed in all daytime or all nighttime samples for a particular month were considered as representative species and included for further discussion. As discussed in the previous session, LC separation helps reduce ion suppression effects and achieve semi-quantification of organic compounds. In the following, we attempt to draw a statistical and comprehensive picture of the organic formulas in Shanghai samples, and the number of isomers is not regarded as a key output at this point.

The assigned formulas can be subdivided into five groups: compounds containing only carbon, hydrogen, and oxygen (CHO), sulfur-containing organics (CHOS), nitrogen-containing organics (CHON), compounds containing both nitrogen and sulfur (CHONS), and organics without oxygen (CHNS* including CHN, CHS and CHNS). The number of molecular formulas tentatively identified in each month including the number for each subgroup is listed in Tables S2-S5. Overall, 810-1510 and 860-1790 molecular formulas of organic compounds with various numbers of isomers for each formula were determined in ESI- and ESI+, respectively. The greatest number of formulas was detected in the January samples in both ESI- and ESI+, whereas the number of formulas in the July samples was far lower compared to those in the other three months, showing an opposite trend to previous measurements in Cambridge, UK [*Rincón et al.*, 2012]. This contrasting observation can be explained by the backward trajectories from the sampling site depicted in Figure S1. These trajectories show that

during winter the air masses originated from northern China with extensive emissions of anthropogenic pollutants from fossil fuel combustion and biomass burning [*Huang et al.*, 2014; *Zha et al.*, 2014] resulting in the highest concentrations of $PM_{2.5}$ and correspondingly a high number of organic compounds, whereas most of air masses came from the clean ocean with the lowest concentrations of $PM_{2.5}$ and particulate organics during the July sampling period (Figure S1 and Table S1). In addition, the boundary layer is normally higher in summer and lower in winter, leading to a higher abundance of organics in January in Shanghai. Nonetheless, it should be noted that the number of formulas did not show a prominent diurnal difference (Tables S2-S5).

Mass spectra of urban atmospheric OAs were reconstructed in both ESI- (Figure 1) and ESI+ (Figure 2), with the sizes of pie charts being proportional to the sum of peak areas of all species for a sub-group in the five samples of a particular month. It should be noted that liquid chromatography in front of the mass spectrometer could greatly resolve ion suppression effect, but uncertainties were still existing in comparing the abundances of different categories due to the different response factor. In this study, all species were tentatively assumed to have same signal response to briefly compare the total peak areas of all organics in different samples. In both ESI- and ESI+, the January samples showed the highest abundance of organic compounds, whereas the abundance of organic compounds in the July samples was far lower than those in the other three months, similarly to the trend in the number of organic compounds identified (Tables S2-S5). In Figure 1, CHO- accounted for a significant percentage in terms of abundance in each month and its fraction was the highest in the July samples. Thus, daytime samples were generally richer in CHO- than nighttime samples, indicating a photochemical formation route of CHO- compounds. The abundance of CHOS- showed the same trend with that of CHO-, also suggesting a photochemical formation route of these multifunctional compounds [*Nozière et al.*, 2010; *Rincón et al.*, 2012; *Schindelka et al.*, 2013]. In contrast, the abundance of CHONS- was far higher in nighttime samples than those in daytime samples, probably due to nighttime NO_3 chemistry [*Hatch et al.*, 2011; *Wang et al.*, 2016]. The abundance of CHON- showed no obvious diurnal

variations, suggesting diverse sources.

For each of the organic compounds with the strongest arbitrary abundance in Figure 1, the formula, RDBE, Xc, and potential source/precursor are listed in Table 1. Most of these compounds were probably aromatics ($Xc \geq 2.50$, such as $C_6H_5NO_3$ (nitrophenol)) with a high degree of unsaturation. Their potential sources include biomass burning and vehicular emissions, important sources for organic compounds in urban atmosphere [*Iinuma et al.*, 2010; *Lin et al.*, 2012a; *Mohr et al.*, 2013; *Simoneit et al.*, 2003; *Vogel et al.*, 2016a]. On the other hand, K' ($C_8H_{12}O_5$) and M' ($C_{10}H_{17}NO_7S$) were oxidation products of monoterpenes [*Gómez-González et al.*, 2012; *Riva et al.*, 2015; *Surratt et al.*, 2008], suggesting that biogenic sources could have made a contribution to the loading of OAs.

A major difference between the mass spectra measured in ESI- and ESI+ is the higher percentage of CHN and CHNS compounds in ESI+, in terms of both number and abundance (Tables S2-S5, Figure 1 and Figure 2). This is mainly due to the fact that amines are easily detected in ESI+ [*Rincón et al.*, 2012]. In Figure 2, the abundance of CHO+ and CHNS*+ accounted for 76-90% of the total, far greater than the others. The total peak area of CHON+ was less than 20% and their number accounted for more than 31% (Tables S2-S5). Both the number and abundance of CHOS+ and CHONS+ represented a few percent.

For each of the ESI+ compounds with the strongest arbitrary abundance, the formula, RDBE, Xc, and potential source/precursor are also listed in Table 1. Most of these compounds contained reduced nitrogen based on the O/N ratios, indicating amines or amides probably from direct emissions of industry and tobacco smoke [*Ge et al.*, 2011; *Schmeltz and Hoffmann*, 1977]. In addition, a number of aliphatic alcohols and carbonyls with long carbon chains showed high abundances, probably from the vehicle emissions [*Jakober et al.*, 2006; *Williams et al.*, 2012].

The numbers of formulas present in both daytime and nighttime samples accounted for about 69-85% of total compounds detected in ESI- and 66-79% of total compounds detected in ESI+ in the four months samples, indicating a less prominent difference of

the organic species between daytime and nighttime samples. Therefore, in the following discussion, formulas and their abundances of daytime and nighttime samples were combined for simplicity, just as if a 24 h sample were collected instead of a 12 h daytime sample and a 12h nighttime sample. On the other hand, only 10-13% of compounds were detected in both polarities, suggesting a general significant difference between compounds detected by ESI- and ESI+. This is due to the differences between the negative and positive ionization mechanisms of the electrospray (e.g., due to the gas phase acidity of the analytes). ESI- is sensitive to deprotonated organic acids, whereas positively charged ions such as protonated organic bases are analyzed by ESI+ [*Lin et al.*, 2012a; *Cech and Enke*, 2001].

3.2 CHO compounds

In total, 333-502 CHO- and 365-603 CHO+ compounds were determined. In Figure 1 and Figure 2, these CHO- and CHO+ compounds, shown in purple, covered wide ranges of molecular weights from 60 to 400 Da and 70 to 500 Da, and the mass spectra peaks with the highest abundances were centered around 120-250 Da and 140-280 Da, respectively.

Figure 3 presents the OS_C vs carbon number plots for the CHO- compounds for each month, with the sizes of symbols being proportional to the fourth root of the abundance of a compound and the color-coding reflecting the Xc values calculated from Eq. (2). According to the criteria for the presence of mono- and polycyclic aromatics, grey circles represent aliphatic compounds ($Xc < 2.50$), whereas orange, green, blue, and red circles correspond to aromatics with a benzene core structure ($2.50 \leq Xc < 2.71$), with a naphthalene ($2.71 \leq Xc < 2.80$) core structure, with an anthracene ($2.80 \leq Xc < 2.83$) core structure, and with a pyrene core structure ($2.83 \leq Xc < 2.92$) [*Yassine et al.*, 2014], respectively.

The detected CHO- compounds exhibited up to 24 carbon atoms and covered a wide mass range from 50 to 400 Da (Figure 1), showing OS_C from -1.8 to +1.5 in the four months (Figure 3). In agreement, Tong et al. [2016] obtained the similar trend of

376 OS_C vs carbon number in a previous study on the molecular composition of OAs at road
377 tunnel sites. Many earlier studies suggest that a molecule with an OS_C between -1 and
378 +1 and with less than 14 carbon atoms can be attributed to semi-volatile and/or low-
379 volatility oxidized organic aerosol components (SV-OOA and LV-OOA), which are
380 produced by multistep oxidation reactions. In contrast, molecules with an OS_C between
381 -0.5 and -1.5 and with more than 6 carbon atoms should be regarded as primary biomass
382 burning organic aerosol (BBOA), which is directly emitted into the atmosphere
383 [*Kourtchev et al.*, 2015; *Kourtchev et al.*, 2016; *Kroll et al.*, 2011]. In this study, most
384 of the CHO- compounds were in the region of SV-OOA and LV-OOA, accounting for
385 56%, 67%, 71%, and 57% of the total CHO- compounds in January, April, July, and
386 October samples, respectively. It seems clear that the majority of mono-aromatics gave
387 signals in region (A-) (Figure 3) where the OS_C lies between -0.5 and 1.0 and the carbon
388 number is between 7 and 15. Consistent with the lower OS_C and longer carbon numbers
389 of polycyclic aromatics, the majority of these signals were found in region (B-) where
390 the OS_C is between -0.75 and 0.5 and the carbon numbers are between 12 and 24. In
391 addition, an obvious variation in the number and abundance of mono- and polycyclic
392 aromatics was observed among different months. For the January samples, the number
393 of mono- and polycyclic aromatics accounted for 24% and 27% of total CHO-
394 compounds, respectively, greater than those in the other three months. The number of
395 mono- and polycyclic aromatics in the July samples was the lowest, being only 22%
396 and 7% of total CHO- compounds, respectively.

397 Not surprisingly, CHO+ compounds came with lower OS_C (from -1.75 to 0.5)
398 compared with CHO- compounds (Figure 4). In addition, most of the CHO+
399 compounds were in the region of BBOA, making up to 56%, 55%, 63%, and 60% of
400 the total CHO+ compounds in January, April, July, and October samples, respectively.
401 These high percentages suggest that primary organic compounds directly emitted from
402 biomass burning were preferably detected in ESI+. A large number of mono- and
403 polycyclic aromatics were located in the regions (A+) and (B+) in Figure 4, showing
404 the same trend in number as those detected in ESI-, i.e., the highest for the January

samples (20% for the mono-aromatics and 18% for the polycyclic aromatics) and the lowest for the July samples (18% for the mono-aromatics and 7% for the polycyclic aromatics). However, while regions (A+) and (B+) had similar ranges of carbon numbers compared to (A-) and (B-), lower ranges of OS_C were observed (-1.0-0.25 for both region (A+) and region (B+)), indicating that the mono- and polycyclic aromatics detected in ESI+ were more reduced than those detected in ESI-.

Triangles in Figure 4 denote the presence of carbonyl functional group(s) in the compound based on pentafluorobenzylhydroxylamine derivatization. A total of 132-183 CHO+ compounds contained at least one carbonyl functional group, accounting for 30-40% of the total in Shanghai. However, it should be noted that carbonyl groups in a compound with a neutral mass greater than 554 Da could not be differentiated using pentafluorobenzylhydroxylamine derivatization, since the scanning range of MS was limited to *m/z* 50–750.

Van Krevelen (VK) diagrams can be used to describe the overall compositional characteristics of complex organic mixtures. The most oxidized species lie at the lower right, whereas the most reduced/saturated species lie at the upper left of the VK diagram where the H/C ratio of each compound in the mass spectrum is plotted vs its O/C ratio [*Nizkorodov et al.*, 2011; *Nozière et al.*, 2015], as shown in Figure 5 and Figure S2. In Figure 5, Figure S2 and Table S2-S5, O/C values for CHO- compounds were obviously greater than those of CHO+ compounds, likely due to the more favorable ionization of carboxylic acids in negative polarity.

HOMs that could play an important role in the new particle formation [*Bianchi et al.*, 2016] and contribute to the loading of OAs [*Molteni et al.*, 2016; *Rissanen et al.*, 2014] are clearly visualized in the VK diagrams. Defined as formulas with $O/C \geq 0.6$ and/or $OS_C \geq 0$, HOMs can be subdivided into three groups: those containing a great fraction of oxidized functional groups (acids, carbonyls) in region (1) ($O/C \geq 0.6$ and $OS_C \geq 0$), those containing a great fraction of reduced functional groups (for instance, alcohols, esters, peroxides) in region (2) ($O/C \geq 0.6$ and $OS_C < 0$), and those with a moderate level of oxygenation in region (3) ($OS_C \geq 0$ and $H/C \leq 1.2$) [*Tu et al.*, 2016].

The number and percentage of HOMs in region (1), region (2) and region (3) for each month are tabulated in Table S6. HOMs accounted for 33-40% of the CHO- compounds in ESI-, and 8-12% in ESI+. More than 50% of HOMs were located in region (3) for both polarities. Recently, Molteni et al. [2016] identified more than 100 HOMs derived from OH radical-initiated reactions of anthropogenic volatile organic compounds (AVOCs, such as benzene, toluene, xylenes, and trimethylbenzenes). Between 31 and 44 of them were present in our samples (in ESI-), accounting for 18-29% of total numbers and 11-31% of total abundances of HOMs-. The highest abundance in the July samples and the lowest in the January samples suggest that photooxidation of AVOCs is probably an important source for HOMs in Shanghai.

3.3 CHON compounds

CHON compounds represented a significant percentage of organic compounds in terms of both number and abundance (Figure 1, Figure 2 and Tables S2-S5). Overall, 168-373 CHON- compounds were observed, compared to 308-698 CHON+ compounds. In general, the CHON compounds covered wide mass ranges with averages around 225 Da in ESI- and 240 Da in ESI+, respectively (Figure 1 and Figure 2).

As can be seen from Figure S3, O/C values for the majority of CHON- compounds were in the range of 0-1.2. Since the ranges of O/C values were too broad to compare between different regions, thus, average O/C values were determined values around 0.5 in each month. Slightly lower values () were previously reported for measurements in the Pearl River Delta region, but higher values were found in Cambridge, UK and Bakersfield, CA, USA [*Kourtchev et al.*, 2016; *Lin et al.*, 2012a; *O'Brien et al.*, 2014; *Rincón et al.*, 2012]. In addition, only around 50% of CHON- compounds were characterized by O/N $\geq$ 3, which is far lower than that in Bakersfield, CA, USA, and possibly indicating more reduced nitrogen-containing compounds in Shanghai. In all four months, aromatics accounted for more than 65% of the CHON- compounds, of which 23-27% accounted for polycyclic aromatics. The majority of aromatics were in

the triangle region between 0.3 < H/C < 2 and 0 < O/C < 0.9. Nitrooxy-aromatic compounds such as nitrophenols ($C_6H_5NO_3$), methylnitrophenol ($C_7H_7NO_3$), nitrocatechol ($C_6H_5NO_4$) and methyl-nitrocatechols ($C_7H_7NO_4$) were always abundant in Shanghai (Figure 1 and Table 1), and were in the past often attributed to OAs from biomass burning [*Iinuma et al.*, 2010; *Kitanovski et al.*, 2012a; *Kitanovski et al.*, 2012b; *Kourtchev et al.*, 2016; *Mohr et al.*, 2013]. Furthermore, these compounds were proposed to be potential contributors to light absorption in organic aerosol particles [*Laskin et al.*, 2015]. It should be noted that Xc values for $C_6H_5NO_4$ and $C_7H_7NO_4$ are lower than 2.50, suggesting that the amount of aromatics could have been underestimated. In addition, the polycyclic aromatics were characterized by lower H/C and O/C ratios, such as nitronapthol ($C_{10}H_7NO_3$) and methylnitronapthol ($C_{11}H_9NO_3$) in the triangle region of 0.3 < H/C < 1.2 and 0 < O/C < 0.6.

In Figure S4, the majority of the CHON+ compounds are located in the range of 0.3 < H/C < 3 and 0 < O/C < 0.6, consistent with previous studies [*O'Brien et al.*, 2013; *Rincón et al.*, 2012], indicating that they were more reduced and saturated than CHON- compounds. 28-110 CHON formulas with more than one oxygen atom were detected in both ESI- and ESI+ in the four month samples. Thus, these compounds might represent amino acids, containing both acidic (-COOH) and basic ($-NH_2$) functionalities [*Lin et al.*, 2012a], which were previously identified as an important class of dissolved organic nitrogen in aerosol particles [*Zhang et al.*, 2002].

CHON+ with O/N ≤ 2 showed a decreasing trend of 80%, 78%, 69%, and 61% of the total CHON+ compounds in the January, April, October, and July samples, respectively, which might be caused by stronger photochemical degradation in July and October (Table S1). Additionally, the average values of O/N were around 1.6, 1.8, 2.0, and 2.5, respectively, showing the same trend as that of CHON-. Apart from that, the most abundant CHON+ compounds in the July and October samples possessed higher O/N ratios and larger carbon numbers than those in the January and April samples, indicating different sources for these compounds. 22%, 23%, 15%, and 17% of CHON+ compounds were mono-aromatics, and 19%, 9%, 12%, and 13% were polycyclic

aromatics for January, April, July, and October, respectively. Benzamide (C_7H_7NO), 4-Hydroxy-benzene acetonitrile (C_8H_7NO), and 3-(4-Hydroxyphenyl) propionitrile (C_9H_7NO), from biomass burning [*Ma and Hays*, 2008] were high in abundances during all four sampling periods.

3.4 CHOS and CHONS compounds

Depending on the collection month, there were 307-413 CHOS- compounds and 169-300 CHONS- compounds, whose molecular weights centered around 120-440 Da and 160-450 Da, respectively (Figure 1). The CHONS- compounds showed higher masses than CHOS- because of the presence of one or two additional nitrate groups [*Tao et al.*, 2014; *Wang et al.*, 2016]. The majority of the CHOS- and CHONS- compounds possessed higher H/C and O/C values ($0.3 < H/C < 3$ and $0 < O/C < 2$) than the CHO- compounds (Figure 5 and Figure S5). Only about 18% of sulfur-containing compounds detected in ESI-, i.e., CHOS- and CHONS-, probably belonged to mono- ($2.50 \leq Xc < 2.71$) or polycyclic ($2.71 \leq Xc$) aromatics with an average RDBE of 10, and the majority of them were in the triangle region of $0.3 < H/C < 1.6$ and $0 < O/C < 0.6$, indicating their low degrees of saturation and oxidation.

About 81-84% of CHOS- and 53-65% of CHONS- compounds possessed enough oxygen atoms to allow assignment of $-OSO_3H$ and/or $-ONO_2$ groups in their formulas, which can be regarded as OSs or nitrooxy-OSs [*Lin et al.*, 2012b; *O'Brien et al.*, 2014; *Wang et al.*, 2016]. Comparing our results for $CHOS_1$- compounds ($O/S \geq 4$) to an earlier study on SOA from biodiesel and diesel fuel [*Blair et al.*, 2017], a sum of 213-254 identical formulas were found, accounting for 70-80% of the total number and 69-82% of the total abundance of the $CHOS_1$- compounds observed here. Nonetheless, previous studies have shown that many of these species can also be formed from biogenic precursors, e.g., $C_5H_{10}O_5S$ (derived from isoprene), $C_7H_{12}O_6S$ (derived from α-pinene) and $C_9H_{16}O_6S$ (derived from limonene or β-caryophyllene) [*Chan et al.*, 2011; *Riva et al.*, 2016; *Surratt et al.*, 2008]. The possibly mixed anthropogenic and biogenic origin of OSs leads to a great complexity and, thus, uncertainty in judging their sources.

On the other hand, OSs of $C_5H_{12}O_7S$ (derived only from biogenic isoprene to the best of our knowledge) and nitrooxy-OSs of $C_{10}H_{17}NO_7S$ (derived only from biogenic α-pinene, β-pinene, α-terpinene, or terpinolene to the best of our knowledge) [*Surratt et al.*, 2008] showed high abundances in each of the four month samples, clearly indicating an influence of biogenic emissions on organic aerosol chemistry in Shanghai.

Lin et al. [2012b] and O'Brien et al. [2014] concluded the importance of the epoxide formation pathway for OSs and nitrooxy-OSs by examining the presence of precursor-product pairs of the CHOS- (or CHONS-) and the corresponding CHO- (CHON-) compounds. The epoxide pathway necessitates that both OS and a corresponding polyol are present, because nucleophilic attack to an epoxide from both sulfates and water molecules can occur in the aerosol phase. In this study, a similar analysis indicates that 51-57% of OSs existed with corresponding CHO- polyols (Table 2), similar to what was observed in Bakersfield [*O'Brien et al.*, 2014] but slightly lower than those in the Pearl River Delta Region and Taiwan [*Lin et al.*, 2012b], suggesting that the epoxide formation pathway is an important formation route for the OSs observed in urban Shanghai. Recently a direct formation of OSs from HOMs and/or peroxyradicals was suggested [*Brüggemann et al.*, 2017]. Because a number of HOMs were detected in all Shanghai samples, as discussed earlier, the direct formation mechanism may have also played a significant role in OSs formation in addition to the epoxide formation pathway. In contrast to the CHOS-, only a small fraction of nitrooxy-OSs (3-11%) existed with the corresponding CHON- polyols, similar to those in the Pearl River Delta Region (3-15%) and Taiwan (0-6%) [*Lin et al.*, 2012b] but far lower than those in Bakersfield (27-42%) [*O'Brien et al.*, 2014], implying that the epoxide pathway is insignificant for the formation of nitrooxy-OSs in Shanghai. In addition, it was shown that organonitrates usually hydrolyze more rapidly than OSs [*Hu et al.*, 2011]. Hence, hydrolysis of nitrooxy-OSs could have also contributed to the abundance of OSs. In our study, on average 78-83% of the nitrooxy-OSs and their corresponding OSs were present (Table 2), highlighting the hydrolysis of the nitrooxy groups.

In ESI+, sulfur-containing compounds including CHOS+ and CHONS+ accounted

for the lowest percentage, in terms of both number (Tables S2-S5) and abundance (Figure 2). Only 41-140 CHOS+ and 80-133 CHONS+ were detected in Shanghai samples. The majority of CHOS+ and CHONS+ compounds were characterized with low degrees of oxidation and saturation, and with an average RDBE of 8 and 5, respectively. In addition, more than 85% of the sulfur-containing compounds possessed $(S+N)/O \geq 0.5$, indicating that reduced nitrogen or sulfur functional groups were present in the chemical structure. 36-56% of CHOS+ and CHONS+ compounds had $Xc \geq 2.50$, indicating that they were aromatics containing reduced sulfur such as $C_{13}H_8OS$ (Xc = 2.75), $C_{14}H_{10}OS$ (Xc = 2.75), and $C_{16}H_{12}OS$ (Xc = 2.78) which were from fossil fuel combustion [*Mead et al.*, 2015].

3.5 CHNS, CHN and CHS compounds

Only 20-30 signals for CHNS*- compounds were detected in Shanghai, representing less than 0.4% of the total abundance of organic compounds (Figure 1). On the other hand, CHNS*+ accounted for around 20% of the total number (Tables S2-S5) and 22-44% of the total abundance (Figure 2), respectively.

Similarly, 115-304 CHN+ compounds were detected and many of them were likely amines, which are readily protonated [*Rincón et al.*, 2012]. Figure 6 shows the KMD diagrams for the CHN+ compounds, giving the most intense signals in the molecular weight range from 120 to 220 Da. In the KMD diagrams, all molecules differing only in the number of $–CH_2$ groups would exhibit the same KMD_{CH2} values and appear on the same horizontal line, suggesting that these molecules have same degree of unsaturation. In Figure 6, aliphatic compounds ($Xc < 2.50$, grey circles), such as triethylamine ($C_6H_{15}N$) and dodecanenitrile ($C_{12}H_{23}N$), appeared in the KMD_{CH2} range of -0.01-0.05; aromatics compounds with a benzene core structure ($2.50 \leq Xc < 2.71$, orange circles), such as nicotine ($C_{10}H_{14}N_2$) and dimethyl toluidine ($C_9H_{13}N$), were observed in the KMD_{CH2} range of 0.05-0.10; and polycyclic aromatics with a naphthalene core structure ($2.71 \leq Xc < 2.80$, green circles), such as 1-naphthalenamine ($C_{10}H_9N$), including nitrogen-heterocyclic compounds such as quinoline (C_9H_7N), were

found in the KMD_{CH2} range of 0.09-0.14. Apart from that, a few polycyclic aromatic compounds with an anthracene core structure ($2.80 \leq Xc < 2.83$, in the KMD_{CH2} range of 0.13-0.16, blue circles) or with a pyrene core structure ($2.83 \leq Xc < 2.92$, in the KMD_{CH2} range of 0.15-0.25, red circles) were detected. These results suggest that the higher values of KMD_{CH2} correspond to the more complex aromatic compounds. Overall, the number of aliphatics accounted for 20-28% of the total CHN+ compounds, in contrast to 30-37% for mono-aromatics and 36-49% for the polycyclic aromatics. The January samples contained the most aromatic CHN+ in terms of number and abundance, which were probably from biomass burning (Figure 1) [*Ge et al.*, 2011; *Schmeltz and Hoffmann*, 1977; *Simoneit et al.*, 2003].

Further VK diagrams for the CHN+ compounds are shown in Figure S6 where the N/C ratio is plotted versus the H/C ratio. These plots clearly separate the different classes of CHN+ according to the number of nitrogen atoms and the degree of unsaturation. Aromatic compounds, especially the polycyclic aromatic compounds, can be found at the lower left corner of the VK diagram, due to their low degree of saturation.

In areas where anthropogenic NH_3 emissions mix with anthropogenic or biogenic SOAs, reduced nitrogen can be added to organic molecules via imine formation reactions, replacing carbonyl oxygen atoms with NH groups (R1) [*Nizkorodov et al.*, 2011; *O'Brien et al.*, 2013]. This pathway was verified previously in laboratory studies through heterogeneous uptake and reaction of gaseous ammonia with wet SOA particles [*Laskin et al.*, 2010]. The resulting imines can undergo additional cyclization reactions (R2) in the condensed aqueous phase SOAs through reactions with a second carbonyl group in the same compound. In addition, the resulting imines can also react with another carbonyl-containing compound forming substituted imines (R3). Then, the substituted imines from the R2 and R3 reactions can undergo additional disproportionation reactions (R4) [*Laskin et al.*, 2010; *Nizkorodov et al.*, 2011]. The schematic about R1-R4 reactions has been given in Figure S7 in the Supplementary Information. The potential relevance of these reactions in forming nitrogen-containing compounds including CHN+, CHON+, CHONS+, and CHNS+, was investigated

according to the mass difference between the reactants and products in reactions R1-R4 and the pre-requisite that only 178-254 compounds were found to contain at least one carbonyl functional group using pentafluorobenzylhydroxylamine derivatization. The number of precursor-product pairs for reactions R1-R4 is tabulated in Table 3, in which "Total unique formulas" shows the total number of unique nitrogen-containing compounds that can be explained by one of the reactions R1-R4, due to large overlaps. Approximately 35-57% of the nitrogen-containing compounds, accounting for 34-70% of the total abundance, could be explained by this type of chemistry. The results showed clear seasonal variations, the highest for the January samples and the lowest for the July samples in terms of both number and abundance, indicating that these reactions played more important roles in the formation of compounds containing reduced nitrogen groups in the January and April samples. It is likely that ammonia was involved in the chemical transformation of particulate organics in Shanghai, the same as in Bakersfield [*O'Brien et al.*, 2013].

About 48-67 CHS+ formulas were detected, containing likely thiols, thioethers (R-S-R), thiophenols or disulfides (R-S-S-R) [*Rincón et al.*, 2012]. Almost all tentatively identified CHS+ were reduced S-containing aromatics with $Xc \geq 2.50$ and the majority of them (82%-91%) belonged to polycyclic aromatics ($Xc \geq 2.71$). For example, $C_{12}H_8S$, which was the most abundant CHS+ compound in the January, April, and October samples and the second most abundant CHS+ compound in July samples, was likely dibenzothiophene from the vehicle emissions [*Wingfors et al.*, 2001].

In addition, 29-47 CHNS+ formulas have been determined, possibly indicating the presence of sulfur-containing amines. 48%, 26%, 15%, and 23% of CHNS*+ compounds were aromatics ($Xc \geq 2.50$) in the January, April, July, and October samples, respectively, also showing the same trend with other aromatics.

3.6 Pollution case in October, 2014

One pollution case was observed in October 2014, during which a significant contrast in air quality was observed between clean and polluted days (Table S1 and

Figure 7a). Hence, this case is selected to elucidate the evolution of OAs. Figure 7a shows the average concentrations of O_3, SO_2, NO_2 and $PM_{2.5}$ in the same time interval as that for the October filter samples. On polluted days, SO_2, NO_2 and $PM_{2.5}$ were in high mass concentrations. Additionally, the particulate organics were also richer in terms of both number (Figure 7b) and abundance (Figure 7c). The abundance of selected tentatively identified tracers was further examined in Figure 7d: $C_6H_{10}O_5$ likely correspond to anhydrosugars with several isomers including levoglucosan, mannosan, galactosan and 1,6-anhydro-β-D-glucofuranose that are regarded as markers for biomass burning [*Kourtchev et al.*, 2016; *Pashynska et al.*, 2002]; $C_6H_5NO_4$ is probably isomers of nitrocatechols from mixed anthropogenic sources, e.g. biomass and vehicular emissions; $C_7H_7NO_4$ is likely methyl-nitrocatechols usually selected as an important marker for biomass burning OAs, which is formed from *m*-cresol emitted during biomass burning as well as diesel exhaust [*Iinuma et al.*, 2010; *Lin et al.*, 2016]; 3-methyl-1, 2, 3-butanetricarboxylic acid (3-MBTCA) can potentially explain the molecular formula $C_8H_{12}O_6$, an OH radical-initiated oxidation product of α- and β-pinene regarded as a tracer for biogenic SOA [*Kourtchev et al.*, 2013; *Szmigielski et al.*, 2007]; and $C_5H_{12}O_7S$ could be isoprene epoxydiol organosulfate ester (IEPOX-OS) formed through reactions between SOx (SO_2 + SO_4^{2-}) and isoprene oxidation products [*Budisulistiorini et al.*, 2015; *Pye et al.*, 2013; *Surratt et al.*, 2008].The assignment of the exact identification for these molecular formulas was not based on MS/MS analysis or standard analysis, however, the suggested compounds (except for 3-MBTCA) have been previously identified in high concentrations in the Shanghai aerosol samples [*Li et al.*, 2016; *Ma et al.*, 2014], adding the credential to our assignment. Compared with those in the daytime of October 19th, the abundances of anhydrosugars, nitrocatechols, methyl-nitrocatechols, 3-MBTCA, and IEPOX-OS increased by a factor of 19, 21, 12, 25, and 68 respectively on the polluted day (October 21st) with the highest $PM_{2.5}$ concentrations, indicating that both anthropogenic and biogenic emissions made contributions to the OAs in this pollution case.

4. Conclusions

The molecular composition of the organic fraction in atmospheric particles collected in July and October 2014 and January and April 2015 at an urban site in Shanghai was investigated using UHPLC coupled to HESI-Orbitrap MS. Compared with direct injection measurements by UHRMS in previous studies [*Lin et al.*, 2012b; *Lin et al.*, 2012a; *O'Brien et al.*, 2014; *O'Brien et al.*, 2013], the UHPLC separation used in our study reduced potential ion suppression effects and, thus, allowed us to retrieve semi-quantitative information on organic compounds in the collected particles. In total, 810-1510 (ESI-) and 860-1790 (ESI+) chemical formulas of organic compounds were assigned.

The chemical characteristics of OAs in urban Shanghai showed clear differences among different months and between daytime and nighttime. January samples contained the most organics in terms of both number and abundance. In contrast, the number and abundance of organic compounds in the July samples were by far the lowest. In addition, a significant number of mono-aromatics and polycyclic aromatics were detected in each of the four months' samples. Again, the highest amount was observed for the January samples (42% of the total number in ESI- and 49% of the total number in ESI+) and the lowest for the July samples (32% of the total number in ESI- and 35% of the total number in ESI+), suggesting that biomass burning and fossil fuel combustion made important contributions to the OAs in urban Shanghai. In general, the number of CHO- and CHOS- compounds was higher for the daytime samples, whereas CHONS- compounds were predominant in the nighttime samples. This contrasting temporal behavior might indicate different formation pathways of these compound classes, e.g., daytime photochemistry versus nitrate radical nighttime chemistry.

We hypothesize that a large range of CHO species in Shanghai were associated with BBOA directly emitted into the atmosphere. Subsequent processing by multistep oxidation reactions might have then resulted in the production of SV-OOA and LV-OOA compounds. In our study, a significant number of HOMs were determined. These compounds may have been derived from oxidation reactions between OH radicals and

aromatics such as benzene, toluene, xylenes, and trimethylbenzenes. The presence of these HOMs in Shanghai samples suggests that photooxidation of AVOCs might be an important source for highly oxygenated compounds in urban environments.

The CHOS and CHONS compounds were mainly detected in the ESI-, thus, most of them were attributed to OSs and nitrooxy-OSs. Since a significant number of OSs could derive from both anthropogenic and biogenic precursors, it is premature to judge their exact sources at this point. Nonetheless, a precursor-product pair analysis pointed towards the conclusion that the epoxide intermediate pathway dominated for the formation of OSs. A similar precursor-product pair analyses suggested that 35-57% of the nitrogen-containing compounds (accounting for 34-70% of the total abundances) detected in ESI+ might have been formed through ammonia-carbonyl chemistry, followed by cyclization and disproportionation reactions. Compared to a previous study in Bakersfield, CA, USA, far more reduced nitrogen-containing compounds were detected [*O'Brien et al.*, 2014; *O'Brien et al.*, 2013], highlighting one of the major characteristics of urban aerosols in Shanghai.

In summary, anthropogenic emissions were shown to have a larger impact on organic aerosol composition in urban Shanghai compared with biogenic primary and secondary sources. Nonetheless, regional transport could have been at least partly responsible for the observed seasonal variations in chemical characteristics of OAs, because many of the most abundant organics observed in our study (Table 1) are likely formed from biomass burning processes, which are normally fairly negligible in the Shanghai area [*Li et al.*, 2016].

Acknowledgement

This study was financially supported by the National Natural Science Foundation of China (No. 21222703, 21561130150, 41575113, & 91644213), the Ministry of Science & Technology of China (2012YQ220113-4), and the Cyrus Tang Foundation (No. CTF-FD2014001). L.Wang thanks the Royal Society-Newton Advanced Fellowship

(NA140106). C. George thanks the support by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement 290852-AIRSEA, and the Marie Curie International Research Staff Exchange project AMIS (Grant 295132). The data generated from this study are available from the authors upon request (*lin_wang@fudan.edu.cn*).

References:

Altieri, K. E., M. G. Hastings, A. J. Peters, and D. M. Sigman (2012), Molecular characterization of water soluble organic nitrogen in marine rainwater by ultra-high resolution electrospray ionization mass spectrometry, *Atmos. Chem. Phys.*, *12*(7), 3557-3571.

Bateman, A. P., S. A. Nizkorodov, J. Laskin, and A. Laskin (2009), Time-resolved molecular characterization of limonene/ozone aerosol using high-resolution electrospray ionization mass spectrometry, *Phys. Chem. Chem. Phys.*, *11*(36), 7931-7942.

Bianchi, F., et al. (2016), New particle formation in the free troposphere: A question of chemistry and timing, *Science*, *352*(6289), 1109-1112.

Blair, S. L., et al. (2017), Molecular Characterization of Organosulfur Compounds in Biodiesel and Diesel Fuel Secondary Organic Aerosol, *Environ. Sci. Technol.*, *51*(1), 119-127.

Borrás, E., and L. A. Tortajada-Genaro (2012), Determination of oxygenated compounds in secondary organic aerosol from isoprene and toluene smog chamber experiments, *Int. J. Environ. Anal. Chem.*, *92*(1), 110-124.

Brüggemann, M., et al. (2017), Real-time detection of highly oxidized organosulfates and BSOA marker compounds during the F-BEACh 2014 field study, *Atmos. Chem. Phys.*, *17*(2), 1453-1469.

Budisulistiorini, S. H., et al. (2015), Examining the effects of anthropogenic emissions on isoprene-derived secondary organic aerosol formation during the 2013 Southern Oxidant and Aerosol Study (SOAS) at the Look Rock, Tennessee ground site, *Atmos. Chem. Phys.*, *15*(15), 8871-8888.

Cech, N. B., and C. G. Enke (2001), Practical implications of some recent studies in electrospray ionization fundamentals, *Mass Spectrom. Rev.*, *20*(6), 362-387.

Chan, M. N., et al. (2011), Influence of aerosol acidity on the chemical composition of secondary organic aerosol from β-caryophyllene, *Atmos. Chem. Phys.*, *11*(4), 1735-1751.

Cheng, Y., S.-M. Li, and A. Leithead (2006), Chemical Characteristics and Origins of Nitrogen-Containing Organic Compounds in PM2.5Aerosols in the Lower Fraser Valley, *Environ. Sci. Technol.*, *40*(19), 5846-5852.

Gómez-González, Y., W. Wang, R. Vermeylen, X. Chi, J. Neirynck, I. A. Janssens, W. Maenhaut, and M. Claeys (2012), Chemical characterisation of atmospheric aerosols during a 2007 summer field campaign at Brasschaat, Belgium: sources and source processes of biogenic secondary organic aerosol, *Atmos. Chem. Phys.*, *12*(1), 125-138.

Ge, X., A. S. Wexler, and S. L. Clegg (2011), Atmospheric amines – Part I. A review, *Atmos. Environ.*, *45*(3), 524-546.

Goldstein, A. H., and I. E. Galbally (2007), Known and unexplored organic constituents in the earth's atmosphere, *Environ. Sci. Technol.*, *41*(5), 1514-1521.

763 Hatch, L. E., J. M. Creamean, A. P. Ault, J. D. Surratt, M. N. Chan, J. H. Seinfeld, E. S. Edgerton, Y. Su, and
764 K. A. Prather (2011), Measurements of isoprene-derived organosulfates in ambient aerosols by
765 aerosol time-of-flight mass spectrometry - part 1: single particle atmospheric observations in Atlanta,
766 *Environ. Sci. Technol.*, *45*(12), 5105-5111.
767 Hu, K. S., A. I. Darer, and M. J. Elrod (2011), Thermodynamics and kinetics of the hydrolysis of
768 atmospherically relevant organonitrates and organosulfates, *Atmos. Chem. Phys.*, *11*(16), 8307-8320.
769 Hu, M., M. Krauss, W. Brack, and T. Schulze (2016), Optimization of LC-Orbitrap-HRMS acquisition and
770 MZmine 2 data processing for nontarget screening of environmental samples using design of
771 experiments, *Anal. Bioanal. Chem.*, *408*(28), 7905-7915.
772 Huang, C., C. H. Chen, L. Li, Z. Cheng, H. L. Wang, H. Y. Huang, D. G. Streets, Y. J. Wang, G. F. Zhang, and
773 Y. R. Chen (2011), Emission inventory of anthropogenic air pollutants and VOC species in the Yangtze
774 River Delta region, China, *Atmos. Chem. Phys.*, *11*(9), 4105-4120.
775 Huang, R. J., et al. (2014), High secondary aerosol contribution to particulate pollution during haze
776 events in China, *Nature*, *514*(7521), 218-222.
777 Hughey, C. A., C. L. Hendrickson, R. P. Rodgers, A. G. Marshall, and K. N. Qian (2001), Kendrick mass
778 defect spectrum: A compact visual analysis for ultrahigh-resolution broadband mass spectra, *Anal.*
779 *Chem.*, *73*(19), 4676-4681.
780 Iinuma, Y., O. Boge, R. Grafe, and H. Herrmann (2010), Methyl-nitrocatechols: atmospheric tracer
781 compounds for biomass burning secondary organic aerosols, *Environ. Sci. Technol.*, *44*(22), 8453-
782 8459.
783 Jakober, C. A., M. J. Charles, M. J. Kleeman, and P. G. Green (2006), LC-MS analysis of carbonyl
784 compounds and their occurrence in diesel emissions, *Anal. Chem., 78*(14), 5086-5093.Jimenez, J. L.,
785 et al. (2009), Evolution of organic aerosols in the atmosphere, *Science*, *326*(5959), 1525-1529.
786 Katajamaa, M., J. Miettinen, and M. Oresic (2006), MZmine: toolbox for processing and visualization of
787 mass spectrometry based molecular profile data, *Bioinf.*, *22*(5), 634-636.
788 Kendrick, E. (1963), A mass scale based on CH2=14.0000 for high resolution mass apectrometry of
789 organic compounds, *Anal. Chem.*, *35*(13), 2146-2154.
790 Kitanovski, Z., I. Grgic, R. Vermeylen, M. Claeys, and W. Maenhaut (2012a), Liquid chromatography
791 tandem mass spectrometry method for characterization of monoaromatic nitro-compounds in
792 atmospheric particulate matter, *J. Chromatogr. A*, *1268*, 35-43.
793 Kitanovski, Z., I. Grgic, F. Yasmeen, M. Claeys, and A. Cusak (2012b), Development of a liquid
794 chromatographic method based on ultraviolet-visible and electrospray ionization mass spectrometric
795 detection for the identification of nitrocatechols and related tracers in biomass burning atmospheric
796 organic aerosol, *Rapid Commun. Mass Spectrom.*, *26*(7), 793-804.
797 Kourtchev, I., J. F. Doussin, C. Giorio, B. Mahon, E. M. Wilson, N. Maurin, E. Pangui, D. S. Venables, J. C.
798 Wenger, and M. Kalberer (2015), Molecular composition of fresh and aged secondary organic aerosol
799 from a mixture of biogenic volatile compounds: a high-resolution mass spectrometry study, *Atmos.*
800 *Chem. Phys.*, *15*(10), 5683-5695.
801 Kourtchev, I., S. Fuller, J. Aalto, T. M. Ruuskanen, M. W. McLeod, W. Maenhaut, R. Jones, M. Kulmala,
802 and M. Kalberer (2013), Molecular composition of boreal forest aerosol from Hyytiala, Finland, using
803 ultrahigh resolution mass spectrometry, *Environ. Sci. Technol.*, *47*(9), 4069-4079.
804 Kourtchev, I., et al. (2016), Molecular composition of organic aerosols in central Amazonia: an ultra-
805 high-resolution mass spectrometry study, *Atmos. Chem. Phys.*, *16*(18), 11899-11913.
806 Kroll, J. H., et al. (2011), Carbon oxidation state as a metric for describing the chemistry of atmospheric

organic aerosol, *Nature Chem.*, *3*(2), 133-139.
Kroll, J. H., and J. H. Seinfeld (2008), Chemistry of secondary organic aerosol: Formation and evolution of low-volatility organics in the atmosphere, *Atmos. Environ.*, *42*(16), 3593-3624.
Laskin, A., J. Laskin, and S. A. Nizkorodov (2015), Chemistry of atmospheric brown carbon, *Chem. Rev.*, *115*(10), 4335-4382.
Laskin, A., J. S. Smith, and J. Laskin (2009), Molecular characterization of nitrogen-containing organic compounds in biomass burning aerosols using high-resolution mass spectrometry, *Environ. Sci. Technol.*, *43*(10), 3764-3771.
Laskin, J., A. Laskin, P. J. Roach, G. W. Slysz, G. A. Anderson, S. A. Nizkorodov, D. L. Bones, and L. Q. Nguyen (2010), High-resolution desorption electrospray ionization mass spectrometry for chemical characterization of organic aerosols, *Anal. Chem.*, *82*(5), 2048-2058.
Li, X., L. Jiang, L. P. Hoa, Y. Lyu, T. Xu, X. Yang, Y. Iinuma, J. Chen, and H. Herrmann (2016), Size distribution of particle-phase sugar and nitrophenol tracers during severe urban haze episodes in Shanghai, *Atmos. Environ.*, *145*, 115-127.
Lin, P., P. K. Aiona, Y. Li, M. Shiraiwa, J. Laskin, S. A. Nizkorodov, and A. Laskin (2016), Molecular characterization of brown carbon in biomass burning aerosol particles, *Environ. Sci. Technol.*, *50*(21), 11815-11824.
Lin, P., A. G. Rincon, M. Kalberer, and J. Z. Yu (2012a), Elemental composition of HULIS in the Pearl River Delta Region, China: results inferred from positive and negative electrospray high resolution mass spectrometric data, *Environ. Sci. Technol.*, *46*(14), 7454-7462.
Lin, P., J. Z. Yu, G. Engling, and M. Kalberer (2012b), Organosulfates in humic-like substance fraction isolated from aerosols at seven locations in East Asia: a study by ultra-high-resolution mass spectrometry, *Environ. Sci. Technol.*, *46*(24), 13118-13127.
Ma, Y., and M. D. Hays (2008), Thermal extraction-two-dimensional gas chromatography-mass spectrometry with heart-cutting for nitrogen heterocyclics in biomass burning aerosols, *J. Chromatogr. A*, *1200*(2), 228-234.
Ma, Y., X. K. Xu, W. H. Song, F. H. Geng, and L. Wang (2014), Seasonal and diurnal variations of particulate organosulfates in urban Shanghai, China, *Atmos. Environ.*, *85*, 152-160.
Mead, R. N., J. D. Felix, G. B. Avery, R. J. Kieber, J. D. Willey, and D. C. Podgorski (2015), Characterization of CHOS compounds in rainwater from continental and coastal storms by ultrahigh resolution mass spectrometry, *Atmos. Environ.*, *105*, 162-168.
Mohr, C., et al. (2013), Contribution of nitrated phenols to wood burning brown carbon light absorption in Detling, United Kingdom during winter time, *Environ. Sci. Technol.*, *47*(12), 6316-6324.
Molteni, U., F. Bianchi, F. Klein, I. El Haddad, C. Frege, M. J. Rossi, J. Dommen, and U. Baltensperger (2016), Formation of highly oxygenated organic molecules from aromatic compounds, *Atmos. Chem. Phys. Discuss.*, 1-39, doi:10.5194/acp-2016-1126.
Nguyen, T. B., A. P. Bateman, D. L. Bones, S. A. Nizkorodov, J. Laskin, and A. Laskin (2010), High-resolution mass spectrometry analysis of secondary organic aerosol generated by ozonolysis of isoprene, *Atmos. Environ.*, *44*(8), 1032-1042.
Nguyen, T. B., J. Laskin, A. Laskin, and S. A. Nizkorodov (2011), Nitrogen-containing organic compounds and oligomers in secondary organic aerosol formed by photooxidation of isoprene, *Environ. Sci. Technol.*, *45*(16), 6908-6918.
Nizkorodov, S. A., J. Laskin, and A. Laskin (2011), Molecular chemistry of organic aerosols through the application of high resolution mass spectrometry, *Phys. Chem. Chem. Phys.*, *13*(9), 3612-3629.

851 Nozière, B., S. Ekström, T. Alsberg, and S. Holmström (2010), Radical-initiated formation of
852 organosulfates and surfactants in atmospheric aerosols, *Geophys. Res. Lett.*, *37*(5),
853 doi:10.1029/2009gl041683.
854 Nozière, B., et al. (2015), The molecular identification of organic compounds in the atmosphere: state
855 of the art and challenges, *Chem. Rev.*, *115*(10), 3919-3983.
856 O'Brien, R. E., A. Laskin, J. Laskin, C. L. Rubitschun, J. D. Surratt, and A. H. Goldstein (2014), Molecular
857 characterization of S- and N-containing organic constituents in ambient aerosols by negative ion
858 mode high-resolution Nanospray Desorption Electrospray Ionization Mass Spectrometry: CalNex
859 2010 field study, *J. Geophys. Res. Atmos.*, *119*(22), 12706-12720, doi:10.1002/2014jd021955.
860 O'Brien, R. E., A. Laskin, J. Laskin, S. Liu, R. Weber, L. M. Russell, and A. H. Goldstein (2013), Molecular
861 characterization of organic aerosol using nanospray desorption/electrospray ionization mass
862 spectrometry: CalNex 2010 field study, *Atmos. Environ.*, *68*, 265-272.
863 Pöschl, U. (2005), Atmospheric aerosols: Composition, transformation, climate and health effects,
864 *Angew. Chem. Int. Ed.*, *44*(46), 7520-7540.
865 Pashynska, V., R. Vermeylen, G. Vas, W. Maenhaut, and M. Claeys (2002), Development of a gas
866 chromatographic/ion trap mass spectrometric method for the determination of levoglucosan and
867 saccharidic compounds in atmospheric aerosols. Application to urban aerosols, *J. Mass Spectrom.*,
868 *37*(12), 1249-1257.
869 Pluskal, T., S. Castillo, A. Villar-Briones, and M. Oresic (2010), MZmine 2: modular framework for
870 processing, visualizing, and analyzing mass spectrometry-based molecular profile data, *BMC Bioinf.*,
871 *11*, doi:10.1186/1471-2105-11-395.
872 Pye, H. O., et al. (2013), Epoxide pathways improve model predictions of isoprene markers and reveal
873 key role of acidity in aerosol formation, *Environ. Sci. Technol.*, *47*(19), 11056-11064.
874 Reinhardt, A., C. Emmenegger, B. Gerrits, C. Panse, J. Dommen, U. Baltensperger, R. Zenobi, and M.
875 Kalberer (2007), Ultrahigh mass resolution and accurate mass measurements as a tool to characterize
876 oligomers in secondary organic aerosols, *Anal. Chem.*, *79*(11), 4074-4082.
877 Rincón, A. G., A. I. Calvo, M. Dietzel, and M. Kalberer (2012), Seasonal differences of urban organic
878 aerosol composition - an ultra-high resolution mass spectrometry study, *Environ. Chem.*, *9*(3), 298-
879 319.
880 Rissanen, M. P., et al. (2014), The formation of highly oxidized multifunctional products in the ozonolysis
881 of cyclohexene, *J. Am. Chem. Soc.*, *136*(44), 15596-15606.
882 Riva, M., S. H. Budisulistiorini, Z. Zhang, A. Gold, and J. D. Surratt (2016), Chemical characterization of
883 secondary organic aerosol constituents from isoprene ozonolysis in the presence of acidic aerosol,
884 *Atmos. Environ.*, *130*, 5-13.
885 Riva, M., E. S. Robinson, E. Perraudin, N. M. Donahue, and E. Villenave (2015), Photochemical aging of
886 secondary organic aerosols generated from the photooxidation of polycyclic aromatic hydrocarbons
887 in the gas-phase, *Environ. Sci. Technol.*, *49*(9), 5407-5416.
888 Roach, P. J., J. Laskin, and A. Laskin (2010), Molecular characterization of organic aerosols using
889 nanospray-desorption/electrospray ionization-mass spectrometry, *Anal. Chem.*, *82*(19), 7979-7986.
890 Romonosky, D. E., Y. Li, M. Shiraiwa, A. Laskin, J. Laskin, and S. A. Nizkorodov (2017), Aqueous
891 photochemistry of secondary organic aerosol of alpha-pinene and alpha-humulene oxidized with
892 ozone, hydroxyl radical, and nitrate radical, *J. Phys. Chem. A*, *121*(6), 1298-1309.
893 Rudich, Y., N. M. Donahue, and T. F. Mentel (2007), Aging of organic aerosol: Bridging the gap between
894 laboratory and field studies, *Annu. Rev. Phys. Chem.*, *58*, 321-352.

Schindelka, J., Y. Iinuma, D. Hoffmann, and H. Herrmann (2013), Sulfate radical-initiated formation of isoprene-derived organosulfates in atmospheric aerosols, *Faraday Discuss.*, *165*, 237-259.

Schmeltz, I., and D. Hoffmann (1977), Nitrogen-containing compounds in tobacco and tobacco smoke, *Chem. Rev.*, *77*(3), 295-311.

Seinfeld, J. H., and J. F. Pankow (2003), Organic atmospheric particulate material, *Annu. Rev. Phys. Chem.*, *54*, 121-140.

Simoneit, B. R. T., A. I. Rushdi, M. R. bin Abas, and B. M. Didyk (2003), Alkyl Amides and Nitriles as Novel Tracers for Biomass Burning, *Environ. Sci. Technol.*, *37*(1), 16-21.

Surratt, J. D., et al. (2008), Organosulfate formation in biogenic secondary organic aerosol, *J. Phys. Chem. A*, *112*(36), 8345-8378.

Szmigielski, R., et al. (2007), 3-methyl-1,2,3-butanetricarboxylic acid: An atmospheric tracer for terpene secondary organic aerosol, *Geophys. Res. Lett.*, *34*(24), doi:10.1029/2007gl031338.

Tao, S., X. Lu, N. Levac, A. P. Bateman, T. B. Nguyen, D. L. Bones, S. A. Nizkorodov, J. Laskin, A. Laskin, and X. Yang (2014), Molecular characterization of organosulfates in organic aerosols from Shanghai and Los Angeles urban areas by nanospray-desorption electrospray ionization high-resolution mass spectrometry, *Environ. Sci. Technol.*, *48*(18), 10993-11001.

Tong, H., I. Kourtchev, P. Pant, I. J. Keyte, I. P. O'Connor, J. C. Wenger, F. D. Pope, R. M. Harrison, and M. Kalberer (2016), Molecular composition of organic aerosols at urban background and road tunnel sites using ultra-high resolution mass spectrometry, *Faraday Discuss.*, *189*, 51-68.

Tu, P., W. A. Hall, and M. V. Johnston (2016), Characterization of highly oxidized molecules in fresh and aged biogenic secondary organic aerosol, *Anal. Chem.*, *88*(8), 4495-4501.

Vogel, A. L., J. Schneider, C. Muller-Tautges, T. Klimach, and T. Hoffmann (2016a), Aerosol chemistry resolved by mass spectrometry: Insights into particle growth after ambient new particle formation, *Environ. Sci. Technol.*, *50*(20), 10814-10822.

Vogel, A. L., et al. (2016b), Aerosol chemistry resolved by mass spectrometry: Linking field measurements of cloud condensation nuclei activity to organic aerosol composition, *Environ. Sci. Technol.*, *50*(20), 10823-10832.

Walser, M. L., Y. Desyaterik, J. Laskin, A. Laskin, and S. A. Nizkorodov (2008), High-resolution mass spectrometric analysis of secondary organic aerosol produced by ozonation of limonene, *Phys. Chem. Chem. Phys.*, *10*(7), 1009-1022.

Wang, L., H. H. Du, J. M. Chen, M. Zhang, X. Y. Huang, H. B. Tan, L. D. Kong, and F. H. Geng (2013), Consecutive transport of anthropogenic air masses and dust storm plume: Two case events at Shanghai, China, *Atmos. Res.*, *127*, 22-33.

Wang, X. K., S. Rossignol, Y. Ma, L. Yao, M. Y. Wang, J. M. Chen, C. George, and L. Wang (2016), Molecular characterization of atmospheric particulate organosulfates in three megacities at the middle and lower reaches of the Yangtze River, *Atmos. Chem. Phys.*, *16*(4), 2285-2298.

Williams, P. I., J. D. Allan, P. Lobo, H. Coe, S. Christie, C. Wilson, D. Hagen, P. Whitefield, D. Raper, and L. Rye (2012), Impact of alternative fuels on emissions characteristics of a gas turbine engine - part 2: volatile and semivolatile particulate matter emissions, *Environ. Sci. Technol.*, *46*(19), 10812-10819.

Wingfors, H., A. Sjodin, P. Haglund, and E. Brorstrom-Lunden (2001), Characterisation and determination of profiles of polycyclic aromatic hydrocarbons in a traffic tunnel in Gothenburg, Sweden, *Atmos. Environ.*, *35*(36), 6361-6369.

Xiao, S., et al. (2015), Strong atmospheric new particle formation in winter in urban Shanghai, China, *Atmos. Chem. Phys.*, *15*(4), 1769-1781.

939 Xu, W., and R. Zhang (2012), Theoretical investigation of interaction of dicarboxylic acids with common
940 aerosol nucleation precursors, *J. Phys. Chem. A*, *116*(18), 4539-4550.
941 Yao, L., et al. (2016), Detection of atmospheric gaseous amines and amides by a high-resolution time-
942 of-flight chemical ionization mass spectrometer with protonated ethanol reagent ions, *Atmos. Chem.*
943 *Phys.*, *16*(22), 14527-14543.
944 Yassine, M. M., M. Harir, E. Dabek-Zlotorzynska, and P. Schmitt-Kopplin (2014), Structural
945 characterization of organic aerosol using Fourier transform ion cyclotron resonance mass
946 spectrometry: Aromaticity equivalent approach, *Rapid Commun. Mass Spectrom.*, *28*(22), 2445-2454.
947 Zha, S. P., T. T. Cheng, J. Tao, R. J. Zhang, J. M. Chen, Y. W. Zhang, C. P. Leng, D. Q. Zhang, and J. F. Du
948 (2014), Characteristics and relevant remote sources of black carbon aerosol in Shanghai, *Atmos. Res.*,
949 *135*, 159-171.
950 Zhang, Q., C. Anastasio, and M. Jimemez-Cruz (2002), Water-soluble organic nitrogen in atmospheric
951 fine particles (PM2.5) from northern California, *J. Geophys. Res. Atmos*, *107*(D11),
952 doi:10.1029/2001jd000870.

Table 1. Potential Identities and Sources for the Most Intense Species From ESI- (Figure 1) and ESI+ (Figure 2) Mass Spectra.

ID	Neutral Mass	Formula	RDBE	Xc	Potential Identity	Potential Source/ Precursor	Reference
A’	118.0264	$C_4H_6O_4$	2	0.00	Succinic acid	-	[*Xu and Zhang*, 2012]
B’	122.0366	$C_7H_6O_2$	5	2.50	Benzoic acid	Naphthalene	[*Riva et al.*, 2015]
C’	139.0268	$C_6H_5NO_3$	5	2.50	Nitrophenol	Biomass burning	[*Mohr et al.*, 2013]
D’	153.0094	$C_3H_7NO_4S$	1	0.00	-	-	-
E’	153.0424	$C_7H_7NO_3$	5	2.50	Methylnitrophenol	Biomass burning	[*Mohr et al.*, 2013]
F’	155.0217	$C_6H_5NO_4$	5	2.33	Nitrocatechols	Biomass burning and vehicle emissions	[*Kourtchev et al.*, 2016] and references therein
G’	166.0264	$C_8H_6O_4$	6	2.50	Phthalic acid	Naphthalene	[*Riva et al.*, 2015]
H’	167.0580	$C_8H_9NO_3$	5	2.50	Benzene compound	-	[*Lin et al.*, 2012a]
I’	169.0473	$C_7H_7NO_4$	5	2.33	Methyl-nitrocatechols	Biomass burning and diesel exhaust	[*Iinuma et al.*, 2010]
J’	183.0166	$C_7H_5NO_5$	6	2.50	3-/ 5-Nitrosalicylic acid	Biomass burning	[*Kitanovski et al.*, 2012a]
K’	188.0681	$C_8H_{12}O_5$	3	1.00	Unknown terpenoic acid	α-pinene; β-pinene	[*Gómez-González et al.*, 2012]
L’	210.0921	$C_8H_{18}O_4S$	0	0.00	Aliphatic organosulfate	Biodiesel and diesel fuel	[*Blair et al.*, 2017]
M’	295.0721	$C_{10}H_{17}NO_7S$	3	0.00	Pinanediol mononitrate	α-pinene; β-pinene; α-terpinene;	[*Gómez-González et al.*,

						Terpinolene	2012]
a'	89.0838	$C_4H_{11}NO$	0	0.00	Dimethylethanolamine; Isobutanolamine	CO_2 capture; Industry	[*Ge et al.*, 2011] and references therein
b'	101.1201	$C_6H_{15}N$	0	0.00	Triethylamine; Hexylamine; Dipropylamine; Diisopropylamine; n-Propylisopropylamine	Cattle; swine; solvent; sewage; ambient air; Tobacco smoke	[*Ge et al.*, 2011] and references therein
c'	113.0838	$C_6H_{11}NO$*	2	1.00	ε-Caprolactam	Industry	[*Cheng et al.*, 2006]
d'	129.0577	C_9H_7N	7	2.71	Quinoline	Tobacco smoke	[*Schmeltz and Hoffmann*, 1977]
e'	143.0733	$C_{10}H_9N$	7	2.71	1-Naphthalenamine; 2-Naphthylamine	Aluminum smelter; tobacco smoke	[*Ge et al.*, 2011] and references therein
f'	148.0159	$C_8H_4O_3$	7	2.50	Phthalic anhydride	Naphthalene	[*Riva et al.*, 2015]
g'	162.1154	$C_{10}H_{14}N_2$	5	2.60	Nicotine	Tobacco smoke	[*Schmeltz and Hoffmann*, 1977]
h'	162.1251	$C_8H_{18}O_3$	0	0.00	Aliphatic alcohol	-	[*Yao et al.*, 2016]

i'	176.1044	$C_8H_{16}O_4$*	1	0.00	Aliphatic carbonyl	-	[*Lin et al.*, 2012a]
j'	181.1826	$C_{12}H_{23}N$	2	2.00	Dodecanenitrile	Biomass burning	[*Simoneit et al.*, 2003]
k'	198.1614	$C_{12}H_{22}O_2$*	2	0.00	Aliphatic carbonyl	-	[*Lin et al.*, 2012a]
l'	204.0783	$C_{12}H_{12}O_3$*	7	2.50	Benzene compound	-	[*Lin et al.*, 2012a]
m'	213.2451	$C_{14}H_{31}N$	0	0.00	C_{14} Amine	-	[*Roach et al.*, 2010]
n'	248.1981	$C_{13}H_{28}O_4$	0	0.00	Aliphatic alcohol	-	-
o'	278.1512	$C_{16}H_{22}O_4$	6	2.00	Aliphatic acid	-	[*Lin et al.*, 2012a]

* A molecule contains at least one carbonyl functional group.

Table 2. Number of (Nitrooxy-) OSs From Different Formation Pathways.

Time	Number (Percentage) of OSs From Plausible Reactant−Product Pairs		
	CHO ($+SO_4^{2-}$) → CHOS	CHON ($+SO_4^{2-}$) → CHONS	CHONS ($+H_2O$, $-HNO_3$) → CHOS
April, 2015	153 (53%)	12 (10%)	99 (79%)
July, 2014	139 (54%)	3 (3%)	76 (80%)
October, 2014	176 (57%)	10 (7%)	108 (78%)
January, 2015	176 (51%)	17 (11%)	132 (83%)

Table 3. Number of Nitrogen-Containing Compounds Detected in ESI+ From Plausible Reactant−Product Pairs.

Time	Number (Percentage) of Nitrogen-Containing Compounds From (Both Instances)Plausible Reactant−Product Pairs				
	R1	R2	R3	R4	Total Unique Formulas
April, 2015	131 (22%)	116 (19%)	152 (25%)	209 (35%)	331 (55%)*
July, 2014	85 (15%)	84 (15%)	87 (15%)	118 (21%)	202 (35%)*
October, 2014	151 (16%)	123 (13%)	228 (24%)	276 (29%)	424 (45%)*
January, 2015	194 (17%)	142 (13%)	438 (39%)	466 (41%)	651 (57%)*

* The percentage is calculated based on the sum of all N-containing compounds (i.e. CHON, CHONS, CHNS and CHN). CHNS and CHN are both included together with CHS in the value of CHNS* in table S1.

Figure Captions

Figure 1. Mass spectra of detected CHO-, CHOS-, CHON-, CHONS-, and CHNS*- reconstructed from extracted ion chromatograms (UHPLC-Orbitrap MS analysis, ESI-). Note that the average abundance of $C_6H_5NO_3$ in the January nighttime samples (the highest one from all samples) was set arbitrarily to 100%. The sizes of pie charts are proportional to the added-up abundances of all species for a sub-group in the five samples of a particular month.

Figure 2. Mass spectra of detected CHO+, CHOS+, CHON+, CHONS+, and CHNS*+ reconstructed from extracted ion chromatograms (UHPLC-Orbitrap MS analysis, ESI+). Note that the average abundance of $C_{10}H_{14}N_2$ in the January nighttime samples (the highest one from all samples) was set arbitrarily to 100%. The sizes of pie charts are proportional to the added-up abundances of all species for a sub-group in the five samples of a particular month.

Figure 3. Carbon oxidation state (OS_C) plots for CHO- compounds. The color-coding indicates the Xc values calculated from Eq. (2). The size of symbols is proportional to the fourth root of the abundance of a compound.

Figure 4. Carbon oxidation state (OS_C) plots for CHO+. The color-coding indicates the Xc values calculated from Eq. (2) and triangles denote compounds containing at least one carbonyl functional group. The size of symbols is proportional to the fourth root of the abundance of a compound.

Figure 5. Van Krevelen (VK) diagrams of CHO- compounds. The color-coding indicates the Xc values calculated from Eq. (2). The size of symbols is proportional to the fourth root of the abundance of a compound.

Figure 6. CH_2-Kendrick diagrams of CHN+. The color-coding indicates the Xc values

calculated from Eq. (2). The size of symbols is proportional to the fourth root of the abundance of a compound.

Figure 7. (a) The average concentrations of O_3, SO_2, NO_2 and $PM_{2.5}$ in the same time interval as that for the October filter samples. (b) Number of ESI- and ESI+ species in different categories. (c) Abundance of ESI- and ESI+ species in different categories. (d) Abundance of selected tentatively identified ESI- tracers in the October samples. The sign of asterisk (*) denotes compounds detected in ESI+.

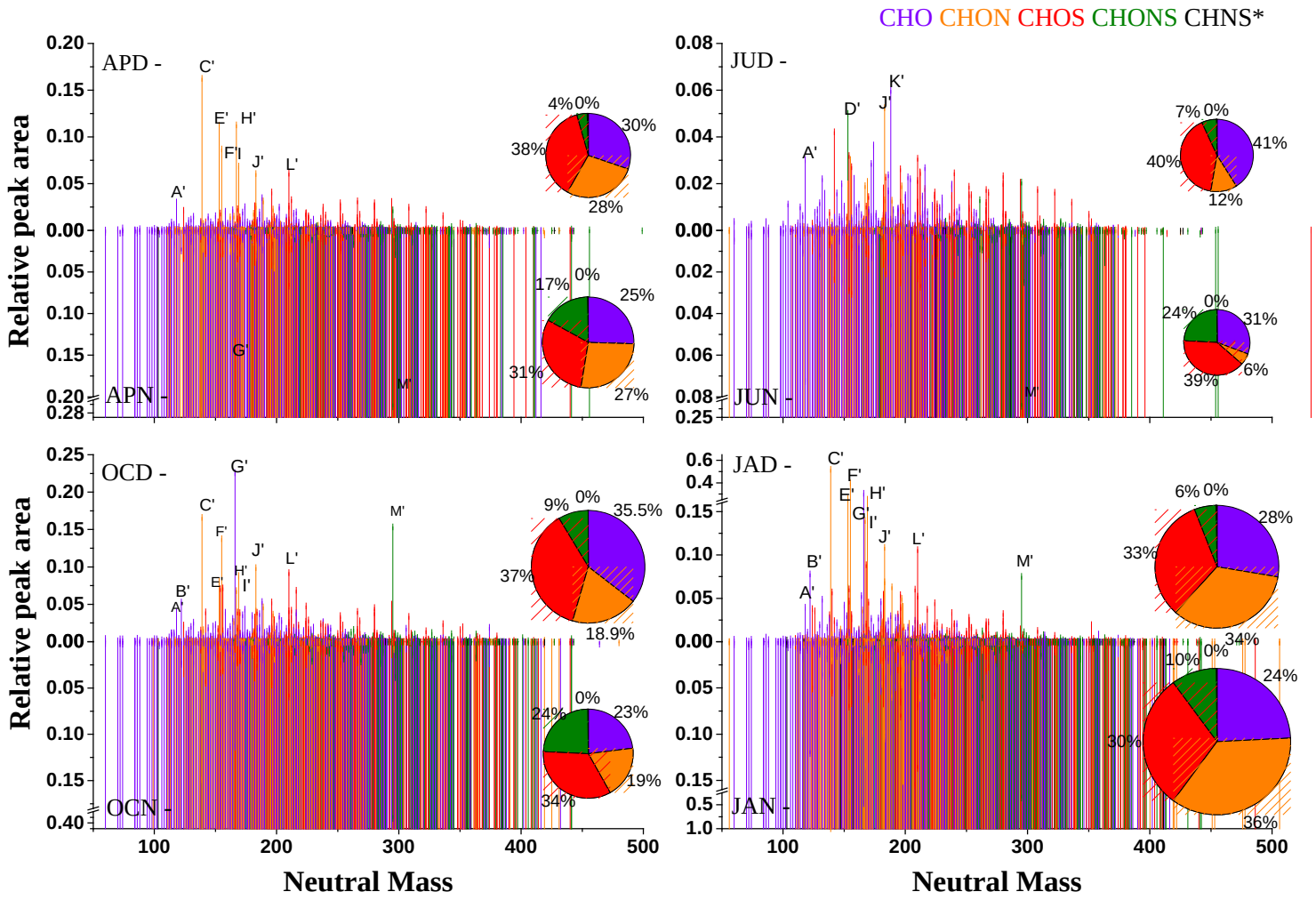
CHO CHON CHOS CHONS CHNS*
APD -
APN -
JUD -
JUN -
OCD -
OCN -
JAD -
JAN -
Relative peak area
Neutral Mass
4% 0% 30% 38% 28%
17% 0% 25% 31% 27%
7% 0% 41% 40% 12%
24% 0% 31% 39% 6%
9% 0% 35.5% 37% 18.9%
24% 0% 23% 34% 19%
6% 0% 28% 33% 34%
10% 0% 24% 30% 36%

Figure 1

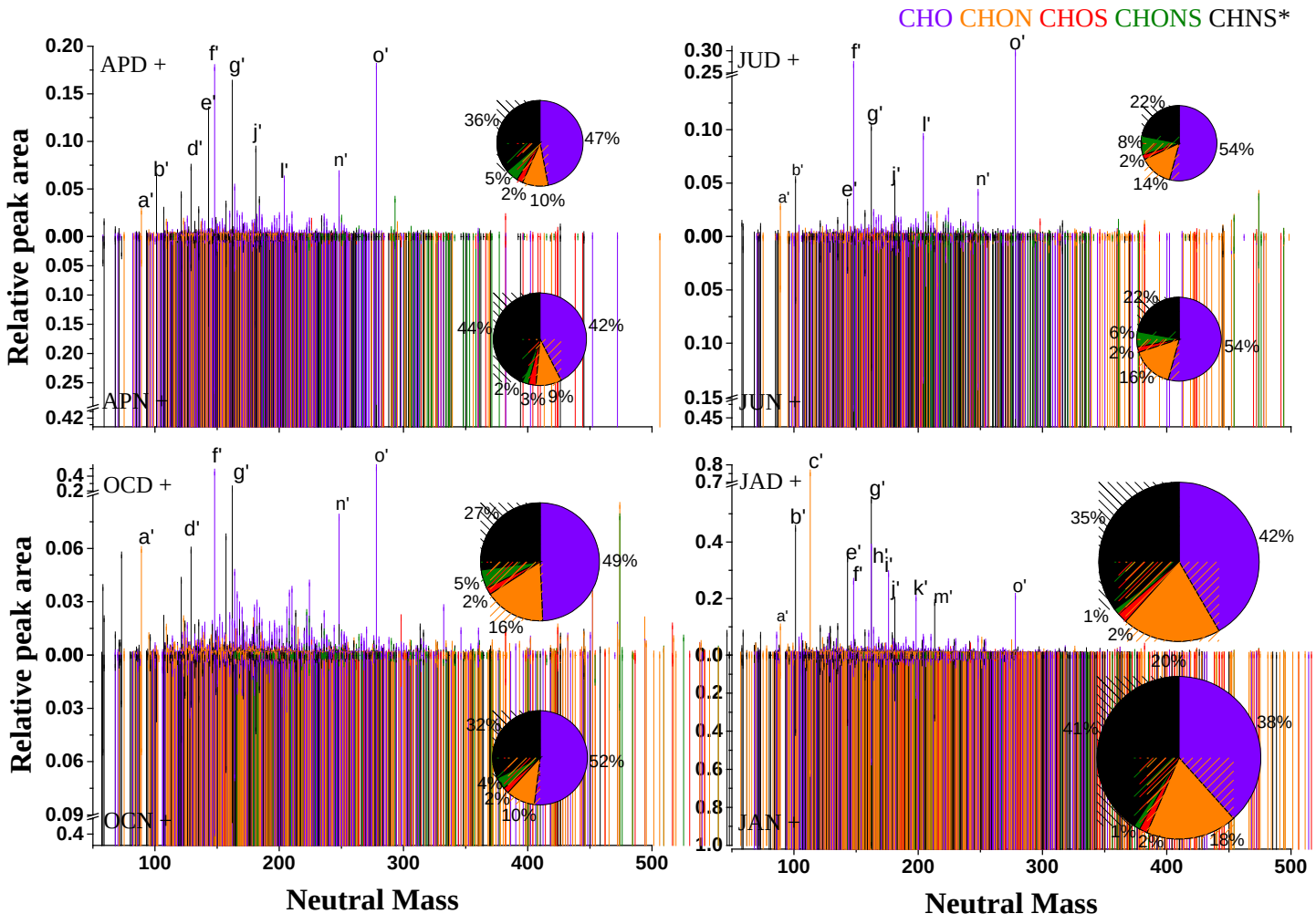
CHO CHON CHOS CHONS CHNS*
APD +
APN +
JUD +
JUN +
OCD +
OCN +
JAD +
JAN +
Relative peak area
Neutral Mass

Figure 2

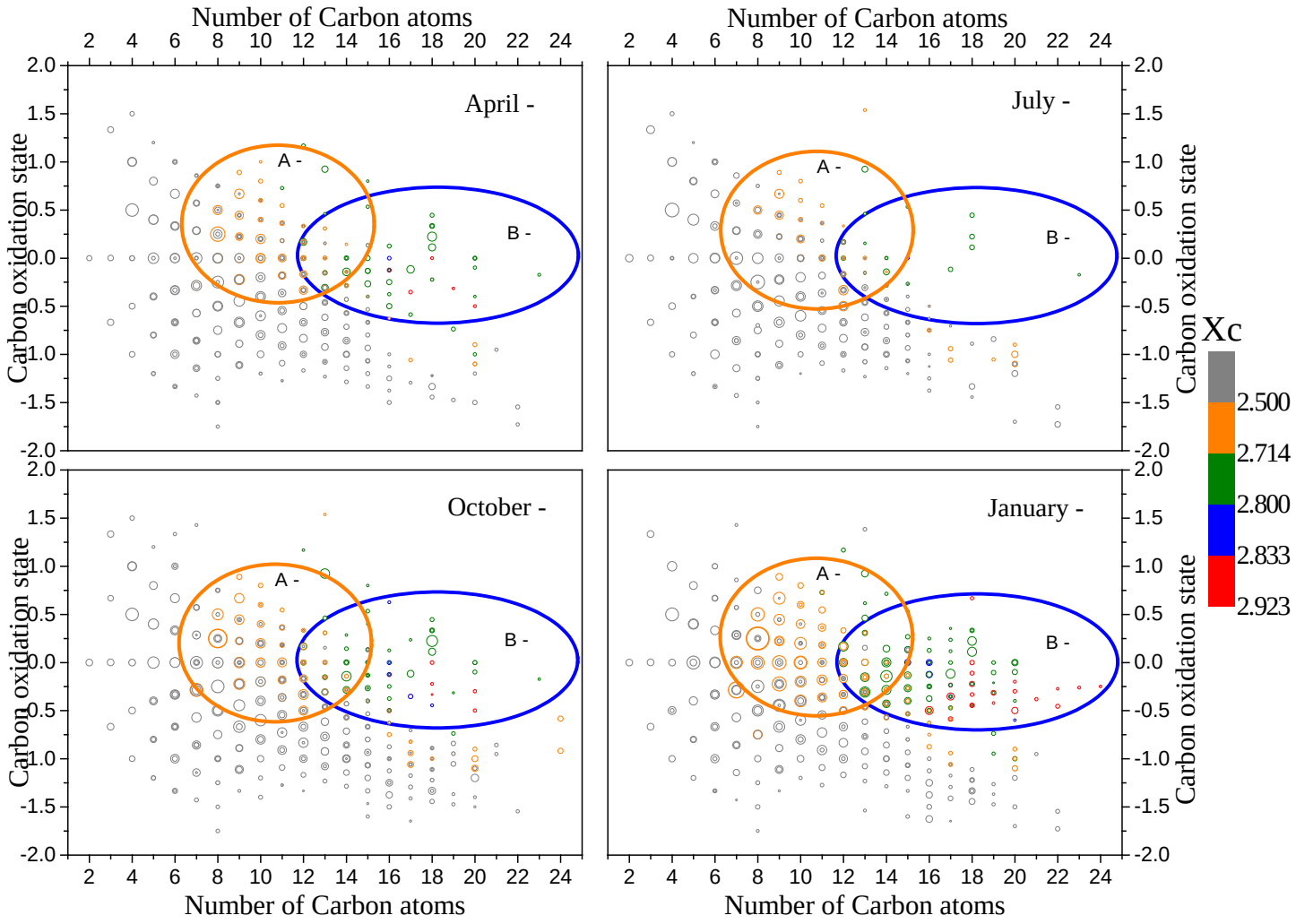
Number of Carbon atoms
Carbon oxidation state
April -
July -
October -
January -
A -
B -
Xc
2.500
2.714
2.800
2.833
2.923

Figure 3

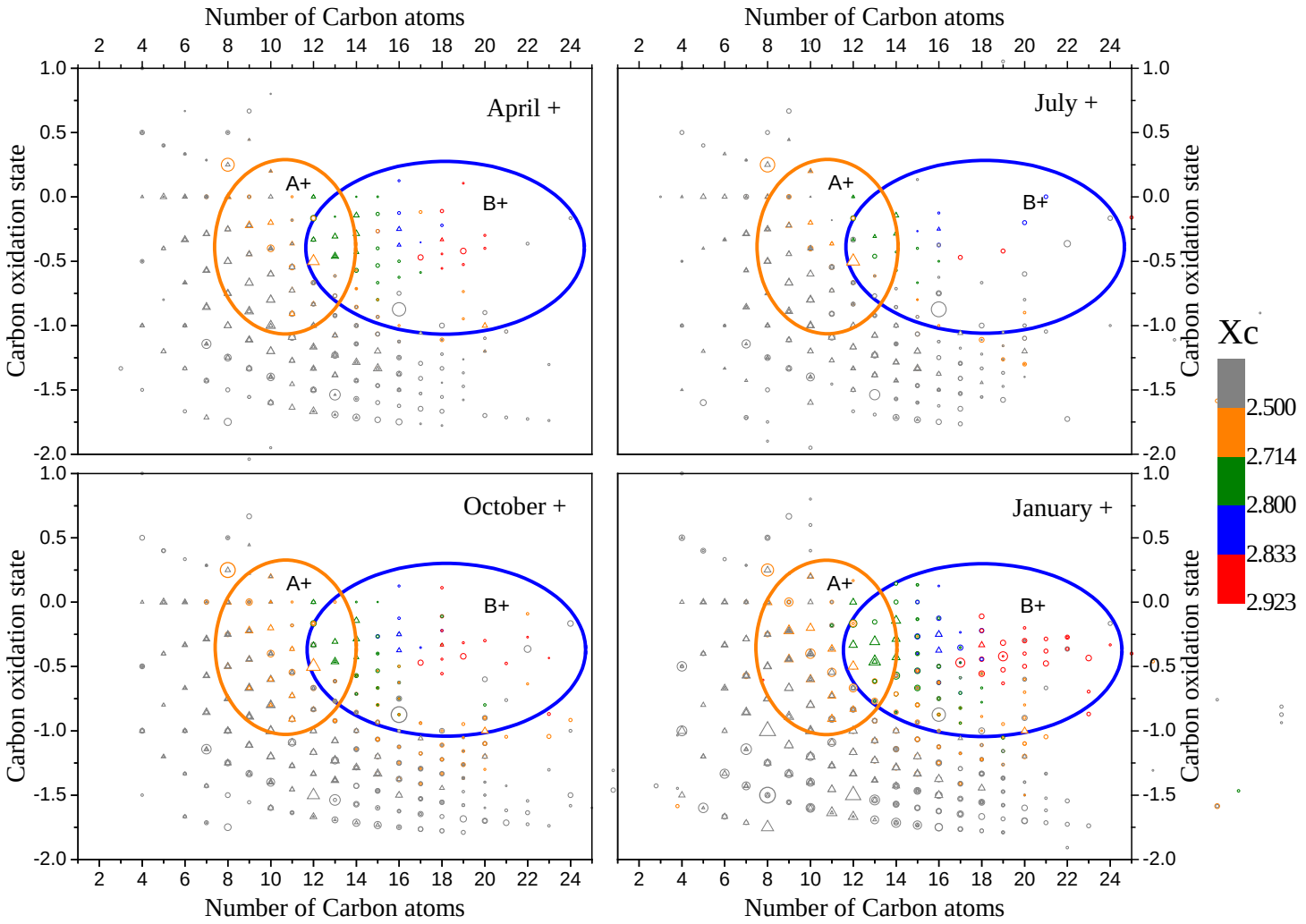
Number of Carbon atoms
Carbon oxidation state
April +
July +
October +
January +
A+
B+
Xc
2.500
2.714
2.800
2.833
2.923

Figure 4

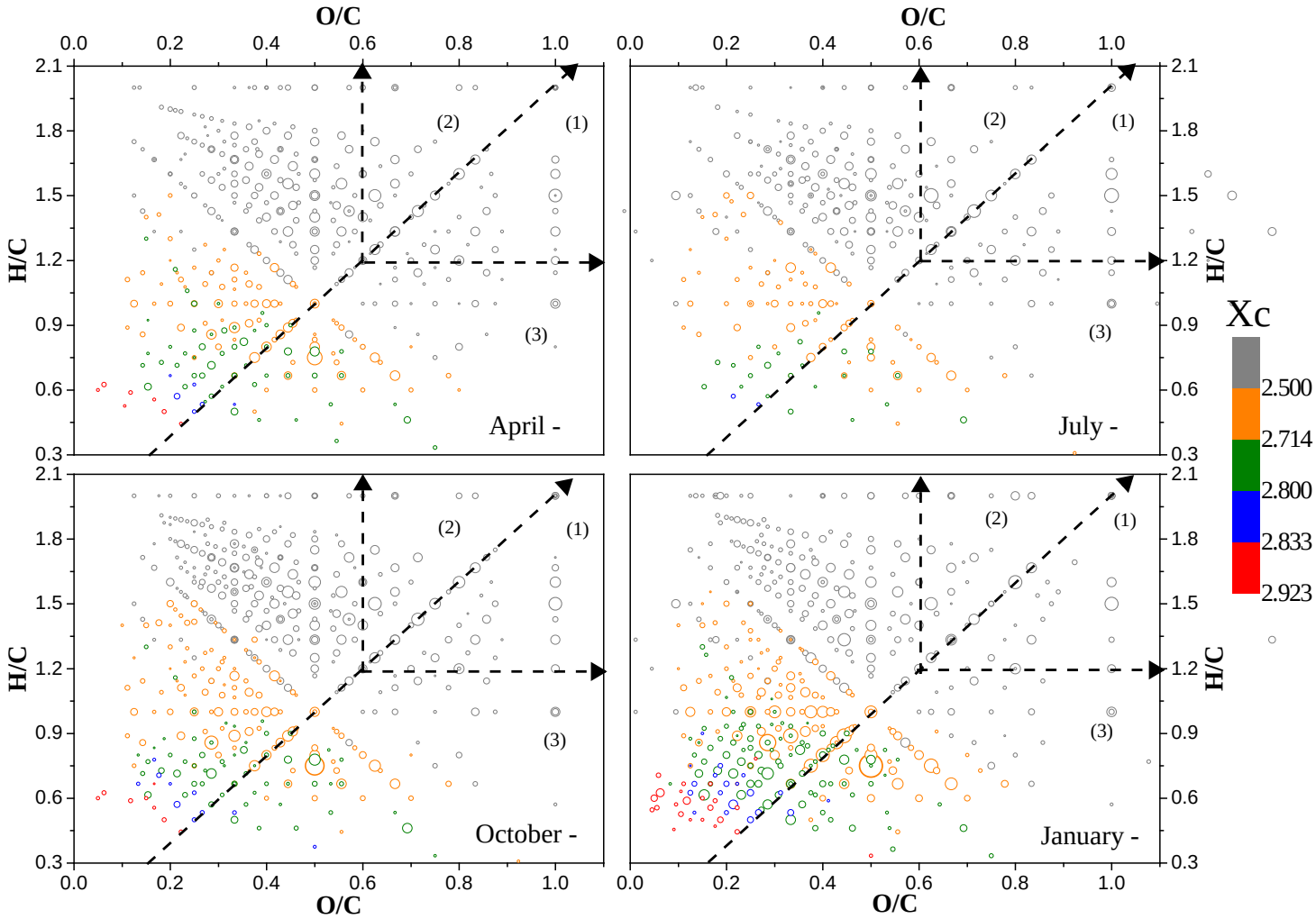
O/C
H/C
April -
July -
October -
January -
(1)
(2)
(3)
Xc
2.500
2.714
2.800
2.833
2.923

Figure 5

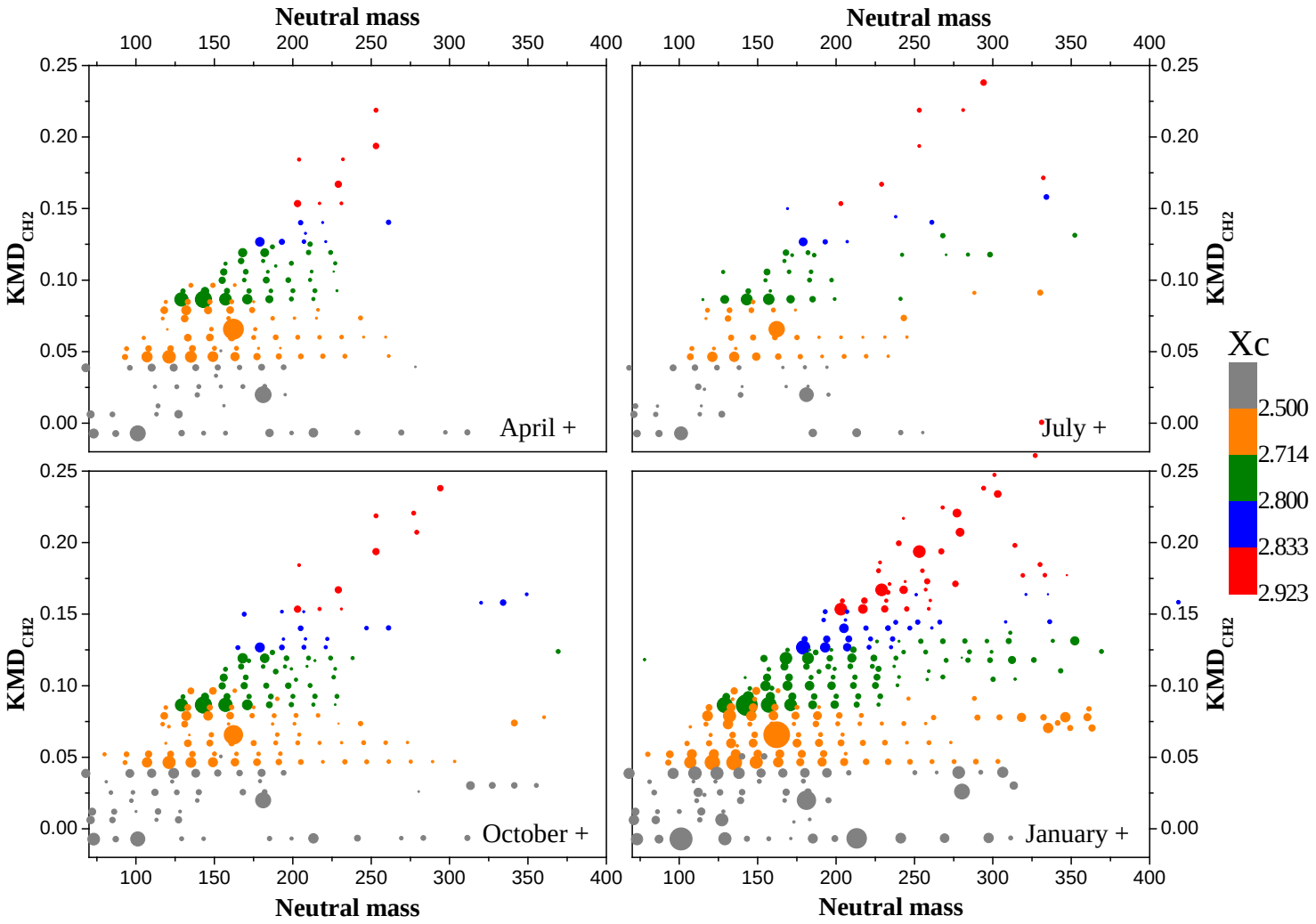
Neutral mass
KMD CH2
April +
July +
October +
January +
Xc
2.500
2.714
2.800
2.833
2.923

Figure 6

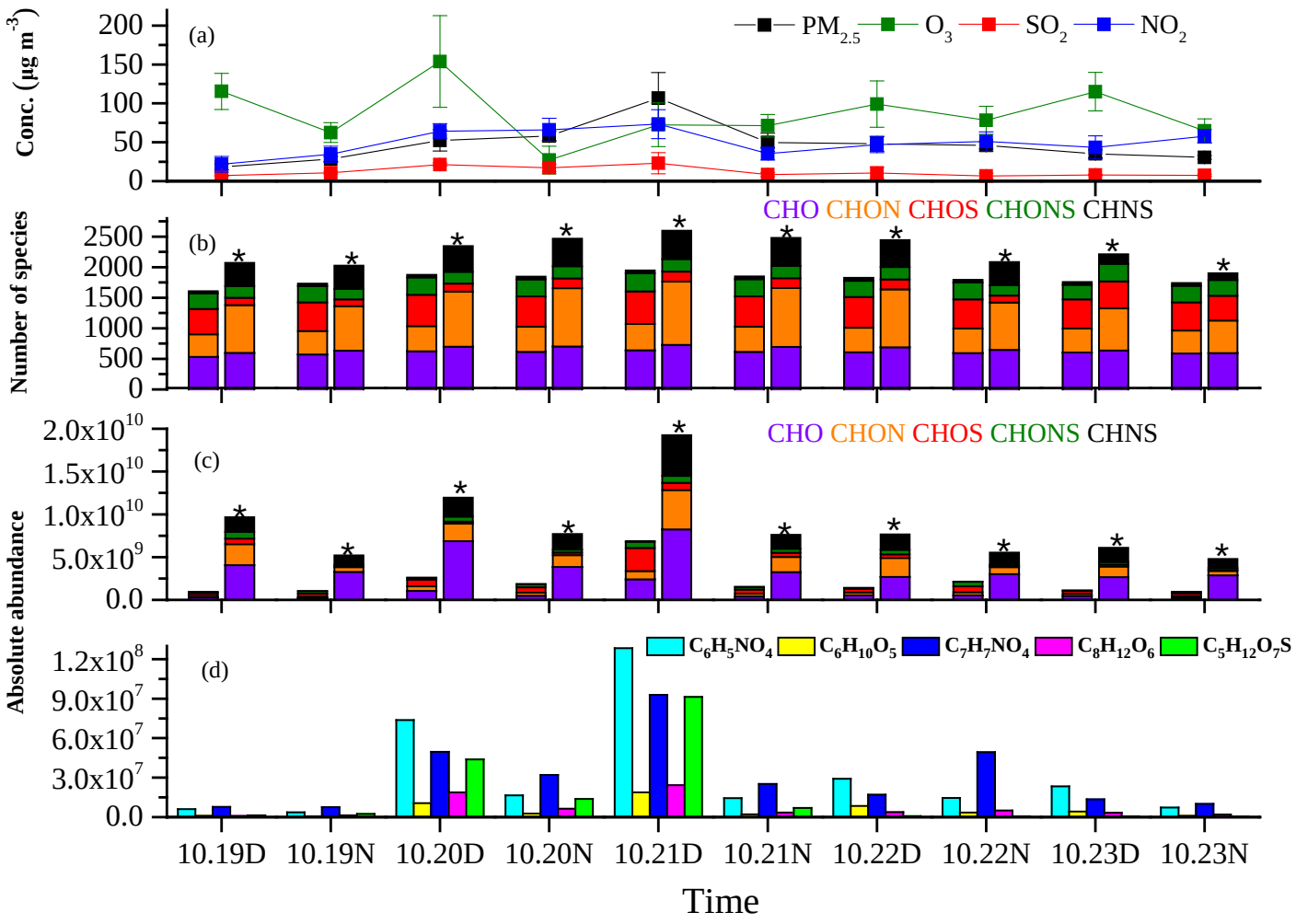
(a)
Conc. (μg m-3)
200
150
100
50
0
PM2.5
O3
SO2
NO2
(b)
Number of species
2500
2000
1500
1000
500
0
CHO CHON CHOS CHONS CHNS
(c)
Absolute abundance
2.0x10^10
1.5x10^10
1.0x10^10
5.0x10^9
0.0
CHO CHON CHOS CHONS CHNS
(d)
1.2x10^8
9.0x10^7
6.0x10^7
3.0x10^7
0.0
C6H5NO4
C6H10O5
C7H7NO4
C8H12O6
C5H12O7S
10.19D
10.19N
10.20D
10.20N
10.21D
10.21N
10.22D
10.22N
10.23D
10.23N
Time

Figure 7