# Response to Reviewer 1

**Overall comment:** Yu et al reports observations of organic aerosol, both primary and secondary, collected on filters for different seasons/time periods of 2017 in Beijing, China. They report water soluble organic carbon (WSOC), its hydrophobic and hydrophilic portions, water soluble ions, total $PM_{2.5}$, total organic carbon (OC), and total elemental carbon (EC). Further, they report tracers associated to different sources (levoglucosan, cholesterol, phthalic acid, 4-methyl-5-nitrocatechol, 2-methylerythritol, 3-hydroxyglutaric acid, and *cis*-pinonic acid). They use the tracers to differentiate sources of OC and WSOC during the seasons via "CO-scaled" concentrations, day and night ratios, correlation coefficients with various meteorological and chemical properties of aerosol ("acidity" and liquid water content), and positive matrix factorization. They find that aqueous chemistry explains a large portion of the secondary organic carbon during most seasons except summer, where photochemistry explains an important biogenic portion. They also find differences in the sources between the seasons (biomass burning vs dust vs vegetation). Overall, the paper is important and of interest to Atmsopheric Chemistry and Physics community as there is general overall uncertainty in the sources of organic aerosol in urban environments, especially during all seasons and high pollution events. The paper will be of value once the authors address the comments below.

**Response:** We deeply appreciate the reviewer’s time and efforts devoted to improving this manuscript. We have provided our responses point by point below each comment, and have carefully revised the paper according to the reviewer’s valuable suggestions.

1. Section 2.1 Sampling: Further information is needed here for the readers to have a better understanding of how the aerosol was collected–Was there a drier in-line prior to be collected on the filters? Was there a denuder to scrub gases prior to the filter to minimize gas-particle partitioning? Was there an impactor or cyclone for size selection? Further, of importance, was there any analysis of potential reactions that occurred on the filters prior to sampling?

**Response:** Thanks for the reviewer's comment. Our sampler included a $PM_{2.5}$ impactor, but no in-line drier or denuder was used in this study.

Sampling of organic carbon is accompanied by both positive and negative artifacts. The positive artifact is due to the adsorption of gaseous organics to the sampling filter, and the negative artifact is caused by the evaporation of collected particulate organic carbon. To eliminate the positive artifact, a denuder can be placed upstream the sample filter to remove the gaseous organics by diffusion to the adsorbent surface (Cheng et al., 2009). The use of a denuder in the sampling system has been reported in previous studies (Eatough et al., 1993, 1999; Mader et al., 2001; Matsumoto et al., 2003; Viana et al., 2006; Cheng et al., 2009, 2010, 2012; Kristensen et al., 2016). The use of a denuder may induce a larger negative artifact, however, as the removal of gaseous organics can enhance the evaporation of particulate OC. Thus a backup filter should also be included in the sampling system (Cheng et al., 2009). Besides, the flow rate passing through the denuder was very low in most studies (Matsumoto et al., 2003; Viana et al., 2006; Cheng et al., 2009, 2010, 2012; Kristensen et al., 2016). This might be due to the significantly decreased removal efficiency of the denuder as the air flow rate increased (Cui et al., 1998; Ding et al., 2002). To collect enough samples for the accurate measurement of trace organic species, the flow rate of 1.05 $m^3$ $min^{-1}$ was chosen in this study. The air flow rate of about 1.05 $m^3$ $min^{-1}$ has been frequently used in the field sampling of organic aerosols (Kawamura et al., 2013; Verma et al., 2012, 2015; Li et al., 2018; Ma et al., 2018; Huang et al., 2020). At this flow rate, a denuder with a high removal efficiency is hardly commercially available.

Nevertheless, we were aware of the potential sampling artifacts and attempted to estimate the sampling artifacts of OC based on the literature results. Firstly, the adsorption behavior of OC might vary with meteorological conditions. Besides, the OC fractions with different volatility show different adsorption behavior. Cheng et al. (2015) compared the concentrations of different OC fractions (OC1, OC2, OC3, OC4) on bare quartz filters with those on denuded quartz filters in the four seasons of Beijing, and the results are summarized in Table S1. The contributions of different OC fractions measured in this study are also listed in Table S1.

**Table S1** The ratio of the OC concentrations on the bare quartz filters to those on the denuded quartz filters in Cheng et al. (2015), as well as the contribution of different OC fractions measured in this study.

| | The ratio of OC on bare quartz filters to denuded quartz filters (Cheng et al., 2015) | | | | The contribution of different OC fractions measured in this study | | | |
|---|---|---|---|---|---|---|---|---|
| | OC1 | OC2 | OC3 | OC4 | OC1 | OC2 | OC3 | OC4 |
| Winter | 1.27 | 1.03 | 1.02 | 1.05 | 10.8 % | 19.6 % | 24.7 % | 44.8 % |
| Spring | 2.05 | 1.05 | 1.00 | 1.01 | 3.9 % | 27.2 % | 43.1 % | 25.7 % |
| Summer | 2.45 | 1.60 | 1.17 | 1.08 | 4.4 % | 37.6 % | 36.0 % | 22.0 % |
| Autumn | 2.08 | 1.05 | 0.99 | 1.01 | 7.9 % | 26.5 % | 40.2 % | 25.3 % |

McDow (1986) systematically investigated the effect of sampling procedure on the OC measurement. The adsorption of organic vapors on bare quartz filters ($C_{postive\ artifact}$) was a function of the sampling duration (t) multiplied by the face velocity (v) as follows:

$$C_{postive\,artifact} = \sum_i \rho_i \frac{1-e^{-\varepsilon_i vt}}{\varepsilon_i vt} \qquad (1)$$

where the face velocity (v, cm $s^{-1}$) is the ratio of the flow rate ($cm^3$ $s^{-1}$) to the sampling area of the filter ($cm^2$), $\rho_i$ is the concentration of adsorptive vapor *i* (g $cm^{-3}$), and $\varepsilon_i$ is a constant which can be defined as:

$$\varepsilon_i = \frac{1}{l}[\beta f(t)t_o \exp\left(\frac{Q_A}{RT}\right) + 1] \qquad (2)$$

where *l* is the effective filter thickness. The average thickness of the quartz filter used in this study was 463 µm. The other parameters are all constants.

Hence, it can be calculated that $\varepsilon_i > 1/l > 20$ $cm^{-1}$, and $1-e^{-\varepsilon vt} \approx 1$. Therefore, the positive artifact ($C_{positive\ artifact}$) is inversely proportional to the product of the sampling duration and the face velocity (v×t). The face velocity of Cheng et al. (2015) was 9.8 cm $s^{-1}$, while that in our study was 47.3 cm $s^{-1}$. The sampling duration of Cheng et al. (2015) was 24 h, while that in our study was 12 h. That is to say, the positive artifact of Cheng et al. (2015) was about 2.4 times higher than that in our study.

Based on the literature results and taking into account all the factors (seasons, OC fractions, sampling procedure), the contribution of positive artifact to the measured OC was estimated to be 2.3 %, 1.4 %, 9.9 %, and 2.2 % in winter, spring, summer and autumn respectively in this study, which is roughly acceptable.

To further estimate the impact of gas-particle partitioning and potential reactions occurring on the filters, we overlapped two quartz filters and took samples at a flow rate of 1.05 $m^3$ $min^{-1}$ for a duration of 12 h. The organic tracers selected in this study were measured in both filters. The organic tracers on the backup filters typically originate from three sources: (1) adsorption of the organic vapors in the atmosphere; (2) adsorption of the semi-volatile species evaporated from the front filter; (3) secondary formation from the adsorbed organic vapors on the backup filter. Except for *cis*-pinonic acid, the tracer concentrations on the backup filter were all less than 5 % of those on the front filters, while the concentration of *cis*-pinonic acid on the backup filter was 21.6 % of that on the front filter. This result suggested that the sampling procedure in this study might bring some uncertainties for the measurement of *cis*-pinonic acid, and the sampling artifact was not significant for the other organic tracers.

## References

Cheng, Y., He, K., Duan, F., Zheng, M., Ma, Y., and Tan, J.: Positive sampling artifact of carbonaceous aerosols and its influence on the thermal-optical split of OC/EC, Atmos. Chem. Phys., 9, 7243-7256, 2009.

Cheng, Y., He, K., Duan, F., Zheng, M., Ma, Y., Tan, J., and Du, Z.: Improved measurement of carbonaceous aerosol: evaluation of the sampling artifacts and inter-comparison of the thermal-optical analysis methods, Atmos. Chem. Phys., 10, 8533-8548, doi:10.5194/acp-10-8533-2010, 2010.

Cheng, Y., Duan, F., He, K., Du, Z., Zheng, M., and Ma, Y.: Sampling artifacts of organic and inorganic aerosol: Implications for the speciation measurement of particulate matter, Atmos. Environ., 55, 229-233, doi:10.1016/j.atmosenv.2012.03.032, 2012.

Cheng, Y., and He, K.: Uncertainties in observational data on organic aerosol: An annual perspective of sampling artifacts in Beijing, China, Environ. Pollut., 206, 113-121, doi:10.1016/j.envpol.2015.06.012, 2015.

Cui, W., Eatough, D. J., and Eatough, N. L.: Fine particulate organic material in the Los Angeles Basin-I: Assessment of the high-volume Brigham Young University Organic Sampling System, BIG BOSS, J. Air Waste Manage Assoc., 48, 1024-1037, doi:10.1080/10473289.1998.10463760, 1998.

Ding, Y., Pang, Y., and Eatough, D. J.: High-volume diffusion denuder sampler for the routine monitoring of fine particulate matter: I. design and optimization of the PC-BOSS, Aerosol Sci. Tech., 36, 369–382, 2002.

Eatough, D. J., Wadsworth, A., Eatough, D. A., Crawford, J. W., Hansen, L. D., and Lewis, E. A.: A multiple-system, multichannel diffusion denuder sampler for the determination of fine particulate organic material in the atmosphere, Atmos. Environ., 27A, 1213–1219, 1993.

Eatough, D. J., Obeidi, F., Pang, Y., Ding, Y., Eatough, N. L., and Wilson, W. E.: Integrated and real-time diffusion denuder sampler for $PM_{2.5}$, Atmos. Environ., 33, 2835–2844, 1999.

Kawamura, K., Tachibana, E., Okuzawa, E., Aggarwal, S. G., Kanaya, Y., and Wang, Z.: High abundances of water-soluble dicarboxylic acids, ketocarboxylic acids and α-dicarbonyls in the mountaintop aerosols over the North China Plain during wheat burning season, Atmos. Chem. Phys., 13, 8285-8302, doi:10.5194/acp-13-8285-2013, 2013.

Kristensen, K., Bilde, M., Aalto, P. P., Petaja, T., and Glasius, M.: Denuder/filter sampling of organic acids and organosulfates at urban and boreal forest sites: Gas/particle distribution and possible sampling artifacts, Atmos. Environ., 130, 36-53, doi:10.1016/j.atmosenv.2015.10.046, 2016.

Mader, B. T., Flagan, R. C., and Seinfeld, J. H.: Sampling atmospheric carbonaceous aerosols using a particle trap impactor/denuder sampler, Environ. Sci. Technol., 35, 4857-4867, 2001.

Matsumoto, K., Hayano, T., and Uematsu, M: Positive artifact in the measurement of particulate carbonaceous substances using an ambient carbon particulate monitor, Atmos. Environ., 37, 4713-4717, doi:10.1016/j.atmosenv.2003.07.006, 2003.

McDow, S. R.: The effect of sampling procedures on organic aerosol measurement, Ph.D. dissertation, Oregon Graduate Center, 1987.

Viana, M., Chi, X., Maenhaut, W., Cafmeyer, J., Querol, X., Alastuey, A., Mikuska, P., and Vecera, Z.: Influence of sampling artifacts on measured PM, OC, and EC levels in carbonaceous aerosols in an urban area, Aerosol Sci. Tech., 40, 107-117, 2006.

2. Line 121: It is unclear what the standards curves were of (the tracers reported throughout paper or other standards), and what is meant by "standard curves with five to seven concentration gradients were re-established." What was re-established?

**Response:** Thanks for the reviewer's comment. The standard curves were made by the silylation derivatives of the organic tracers, and the detailed information of these standards is listed in Table S2. "Re-established" means that we measured the derivative products of the standard solutions each time before measuring the ambient samples. We have avoided the unclear expression and revised as follows:

“The authentic standards (Table S2) were dissolved in anhydrous pyridine, and diluted to five to seven different concentrations. Then 100 µL of the standard solutions were reacted with 200 µL silylating reagent (BSTFA: TMCS = 99:1) at 75°C for a duration of 70 min. After cooling down to the ambient temperature, these solutions containing derivative products were diluted to 1 mL with n-hexane, and measured by GC/MS/MS right before the analysis of ambient samples. The $R^2$ of the derivative products were above 0.99, indicating good linearities of these standard curves. ”

3. Line 122: What authentic standards? What company/purity?

**Response:** The information of the authentic standards is provided in Table S2 below.

**Table S2** The detailed information of the authentic standards used in this study.

| **Authentic standard** | **Molecular formula** | **CAS number** | **Company** | **Purity** |
|---|---|---|---|---|
| Levoglucosan | $C_6H_{10}O_5$ | 498-07-7 | Sigma-Aldrich | 99 % |
| Cholesterol | $C_{27}H_{46}O$ | 57-88-5 | Sigma-Aldrich | 93 % |
| Phthalic acid | $C_8H_6O_4$ | 88-99-3 | Sigma-Aldrich | 99 % |
| 4-Methyl-5-nitrocatechol | $C_7H_7NO_4$ | 68906-21-8 | Toronto Research Chemicals | 98 % |
| 2-Methylerythritol | $C_5H_{12}O_4$ | 58698-37-6 | Sigma-Aldrich | 90 % |
| 3-Hydroxyglutaric acid | $C_5H_8O_5$ | 638-18-6 | Sigma-Aldrich | 95 % |
| cis-Pinonic acid | $C_{10}H_{16}O_3$ | 61826-55-9 | Sigma-Aldrich | 98 % |

4. Section 2.3: Please describe or cite the PMF software used. PMF 5.0 is not enough to understand how positive matrix factorization was actually conducted.

**Response:** Thanks for this kind suggestion. We have added a brief introduction to the PMF software in the manuscript. Besides, we have added the detailed information on the uncertainty calculation of the input data, the selection criteria for the optimal solution, the diagnostic plots and error estimation in the supplementary material.

5. Line 140: I highly recommend the use of "aerosol acidity," as defined in this line, due to the discussion from Pye et al. 2020 (https://acp.copernicus.org/articles/20/4809/2020/). The ratio here does not define acidity, and is analytically challenging to say if it is defining the amount of hydronium

ions in the aerosol phase, as the hydronium ions may be a very low detection limits that cannot be quantified due to propagation of uncertainty.

**Response:** We deeply appreciate the reviewer's valuable suggestion. According to Pye et al. (2020), we changed the ratio of $R_{A/C}$ to the approximate value of aerosol pH ($pH_F$) to denote the aerosol acidity. The $pH_F$ value was estimated using the ISORROPIA-II model. And the molality of $H^+$ ($m_{H^+}$), which was calculated by $m_{H^+} = 10^{-pH_F}$, was used instead for the correlation analysis in the revised manuscript.

6. Section 3: Though an important and valuable aspect of this whole manuscript is that the filters were collected during different seasons, I highly recommend the authors soften the language throughout that the results "reflect" a specific season or are similar or different to other studies. Since it's only for one year and approximately 2 weeks for each season. The limited data makes it hard to say how typical the results are and this should be discussed/emphasized throughout (instead of general statements that in fall this is what is observed/happens).

**Response:** Thanks for the reviewer's valuable suggestion. We have changed the terms of "winter/spring/summer/autumn" to "January/April/July/October" or "during the sampling periods in winter/spring/summer/autumn" throughout the revised manuscript, in order to be more specific.

7. Another area I suggest the authors be careful in their discussion of r values, as majority of the values they report lead to Rˆ2 values less than 0.5 (thus explaining less than 50% of the variability observed).

**Response:** We agree with the reviewer that most of the r values led to $R^2$ less than 0.5, thus could only explain less than 50 % of the variability observed, and were insufficient to reach clear conclusions. Therefore, in the revised manuscript, we only explained the r results and avoided making definite conclusions based on the r values.

8. Line 195: Since the authors are comparing OC from emissions inventory to Fig. 2, I would recommend converting the emissions to OC-to-CO ratios. Also, I would

recommend adding these ratios, if possible, to Fig. 2, for direct comparisons with observations.

**Response:** We are deeply grateful for the reviewer's valuable suggestion. As suggested by the reviewer, we have added the emissions of OC from the Fire Inventory to Figure 2 in the revised manuscript. As a tracer for biomass burning, levoglucosan in the atmosphere may derive from both residential biofuel burning and open biomass burning. To determine whether open biomass burning was the dominant type of biomass burning in Beijing, the seasonal variation of the CO-scaled concentration of levoglucosan was compared with that of the OC emission amount from the Fire Inventory in this study. The CO-scaled concentration of levoglucosan showed a totally different seasonal variation trend from that of the OC emission from open biomass burning, therefore, we speculated that open biomass burning was not the dominant category of biomass burning in Beijing. After careful consideration, we thought that it was not necessary to convert the emissions to OC-to-CO ratios. Instead, the absolute emission amounts of OC from the Fire Inventory were added.

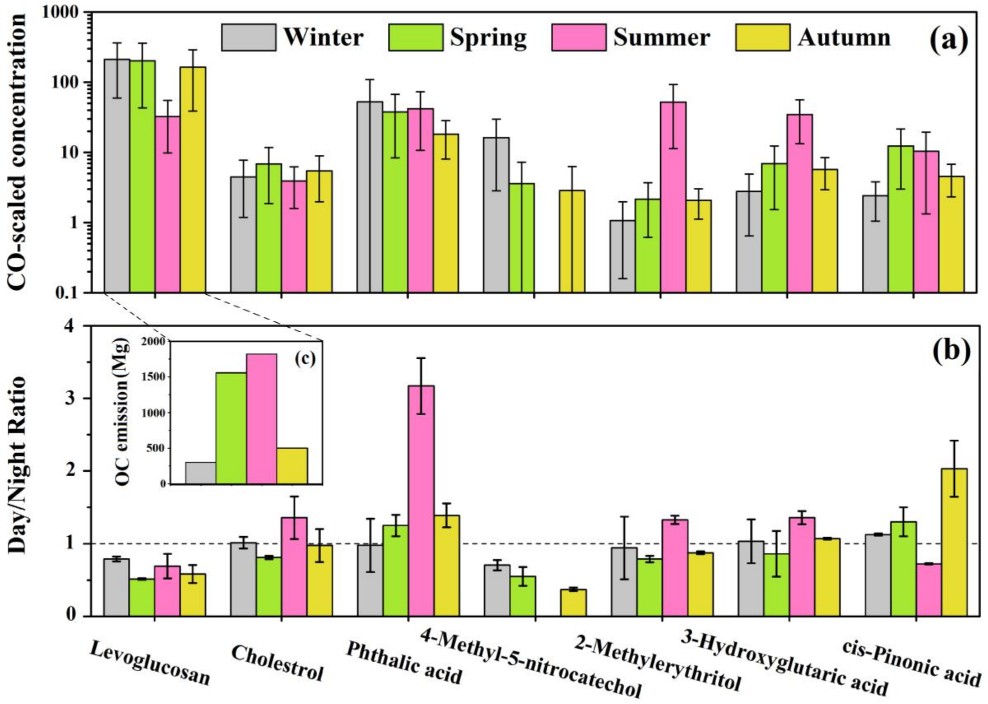


**Figure 2.** (a) The CO-scaled concentration of the identified organic tracers; (b) The day to night ratios of the measured concentrations of the organic tracers; (c) The OC emission amounts from open biomass burning provided by the Fire Inventory (FINN)

in Beijing during the sampling periods in four seasons of 2017.

9. Line 199: Is it possible to get emission inventory values of residential biofuel combustion and coal combustion to compare with the OC from open biomass burning?

**Response:** The total emission amounts of OC from the residential sources in January, April, July and October of 2016 can be obtained from the multi-resolution emission inventory for China (MEIC, http://www.MEICmodel.org). However, the residential OC from MEIC is a total amount of biofuel combustion, coal combustion, other fossil fuel combustion, etc. Unfortunately, the detailed OC emission amount from each residential source is not available. Therefore, the respective OC emission inventory values of residential biofuel combustion and coal combustion were not provided in the manuscript.

10. Line 207: It is unclear how aromatics form SOA to impact WSOC during winter, as the photochemistry is greatly reduced. Could the authors provide more discussion concerning this?

**Response:** We deeply appreciate the reviewer's valuable suggestion. The following discussion has been added in the revised manuscript.

Although the concentrations of oxidants were usually lower in winter due to the weaker solar radiation, a previous observation found that the ·OH concentration in Beijing was significantly higher than that in New York, Birmingham and Tokyo, and was nearly 1 order of magnitude higher than that predicted by global models in northern China in winter (Tan et al., 2018). Zhang et al. (2019) indicated that HONO, which was mainly from the heterogeneous reactions of $NO_2$ and traffic emissions, was the major precursor of ·OH in winter. According to the WRF-Chem model simulation, HONO resulted in a significant enhancement (5-25 $\mu g\ m^{-3}$) of SOA formation (most of which were from the aromatic precursors) during a haze episode in winter in the Beijing-Tianjin-Hebei region (Zhang et al., 2019). Besides, some recent studies suggested that the brown carbon-derived singlet molecular oxygen ($^1O_2^*$) in aerosol liquid water could react rapidly with the electron-rich organics such as PAHs, thus facilitate the aromatic SOA formation (Kaur et al., 2019; Manfrin et al., 2019). This process might be more significant in winter, when the concentration of HULIS was much higher than that in other seasons. Therefore, both the enhanced levels of oxidants including ·OH and $^1O_2^*$ and the higher concentrations of aromatic precursors in winter contributed to the enhanced aromatic SOA formation during the study period in winter.

**References**

Tan, Z., Rohrer, F., Lu, K., Ma, X., Bohn, B., Broch, S., Dong, H., Fuchs, H., Gkatzelis, G. I., Hofzumahaus, A., Holland, F., Li, X., Liu, Y., Liu, Y., Novelli, A., Shao, M., Wang, H., Wu, Y., Zeng, L., Hu, M., Kiendler-Scharr, A., Wahner, A., and Zhang, Y.: Wintertime photochemistry in Beijing: observations of $RO_x$ radical concentrations in the North China Plain during the BEST-ONE campaign, Atmos. Chem. Phys., 18, 12391-12411, doi:10.5194/acp-18-12391-2018, 2018.

Zhang, J., Chen, J., Xue, C., Chen, H., Zhang, Q., Liu, X., Mu, Y., Guo, Y., Wang, D., Chen, Y., Li, J., Qu, Y., and An, J.: Impacts of six potential HONO sources on $HO_x$ budgets and SOA formation during a wintertime heavy haze period in the North China Plain, Sci. Total Environ., 681, 110-123, doi:10.1016/j.scitotenv.2019.05.100, 2019.

11. Line 249: It is surprising that the authors are saying that gas-phase photooxidation was not the dominant formation mechanism of secondary organic carbon. I can see maybe WSOC, but seeing all secondary organic carbon is a big statement. Especially, since the authors go on in line 254-55 to say photochemistry plays a role.

**Response:** We deeply appreciate the reviewer's valuable comment. We agree with the reviewer that it was not proper to conclude that "gas-phase photooxidation was not the dominant formation mechanism of SOC" merely based on the result that "WSOC/OC did not show any significant positive correlation with $O_3$ concentrations". The corresponding statement has been deleted in the revised manuscript. Besides, "SOC" has been corrected to "SOC in WSOC" to be more specific.

12. Line 339: Source 3 did not show the highest contribution in winter....highest contribution of what?

**Response:** We feel sorry for the unclear expression. This sentence has been revised as: "The primary emission strength of coal combustion was the strongest in winter among four seasons, since the domestic heating activities required extra amounts of coal combustion in this season. However, the contribution of Factor 3 to WSOC during the study period in winter was not the highest among four seasons, implying that there could be other sources beyond coal combustion included in Factor 3.

13. Line 416-419: I would recommend caution here, as other hypothesis have been stated for reasons in differences between chamber SOA and ambient SOA, including losses of vapors to the walls and autoxidation (which has been shut down in chamber experiments due to too high $NO_x$ levels and/or too high aerosol loadings).

**Response:** We deeply appreciate the reviewer's valuable suggestion. This statement has been revised as: "Previous observation suggested that a large fraction of ambient SOA was more oxidized than those generated in the dry smog chambers, where SOA could only be produced through gas-phase oxidation (Aiken et al., 2008). There have been some hypothesis for the difference between the chamber SOA and ambient SOA, such as the losses of vapors to the walls, and the autoxidation due to the uncertainties in chamber radical environment (McVay et al., 2016; Thornto et al., 2020). Besides, the results of this study also indicated that the aqueous-phase processing, which can produce more hydrophilic SOA, may be one of the reasons for the discrepancy in the oxidation degrees of ambient SOA and chamber SOA (Ervens et al., 2011)."

**References**

McVay, R.C., Zhang, X., Aumont, B., Valorso, R., Camredon, M., La, Y. S., Wennberg, P. O., and Seinfeld, J. H.: SOA formation from the photooxidation of α-pinene: systematic exploration of the simulation of chamber data, Atmos. Chem. Phys., 16, 2785-2802, doi:10.5194/acp-16-2785-2016, 2016.

Thornton, J. A., Shilling, J. E., Shrivastava, M., D'Ambro, E. L., Zawadowicz, M. A., and Liu, J.: A near-explicit mechanistic evaluation of isoprene photochemical secondary organic aerosol formation and evolution: simulations of multiple chamber experiments with and without added $NO_x$, ACS Earth Space Chem., 4, 1161-1181, doi: 10.1021/acsearthspacechem.0c00118, 2020.

14. Table 1: I would recommend somehow highlighting which values show statistical

differences between day and night and between seasons. Also, I would recommend including average CO mixing ratios.

**Response:** We deeply appreciate the reviewer’s suggestion. The average values which showed statistical differences among seasons and between day and night have been highlighted as follows, and the average CO concentrations have also been included in Table 1 in the revised manuscript.

**Table 1** The average concentrations and standard deviations of the identified carbonaceous species in PM$_{2.5}$ during the sampling periods in four seasons.

| **Compounds (μg m$^{-3}$)** | **Winter** | | | **Spring** | | | **Summer** | | | **Autumn** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Daytime** | **Nighttime** | **Mean** | **Daytime** | **Nighttime** | **Mean** | **Daytime** | **Nighttime** | **Mean** | **Daytime** | **Nighttime** | **Mean** |
| CO | 1.7 ± 1.7 | 2.2 ± 2.8 | 1.9 ± 2.3$^{a}$ | 0.6 ± 0.4$^{*}$ | 0.8 ± 0.4$^{*}$ | 0.7 ± 0.4$^{b}$ | 1.1 ± 0.4 | 0.9 ± 0.2 | 1.0 ± 0.3$^{b}$ | 1.1 ± 0.4$^{*}$ | 1.4 ± 0.4$^{*}$ | 1.2 ± 0.4$^{b}$ |
| PM$_{2.5}$ | 120 ± 107 | 147 ± 154 | 133 ± 131$^{a}$ | 60.6 ± 36.2 | 64.5 ± 34.8 | 62.5 ± 34.9$^{b}$ | 59.8 ± 28.6 | 51.9 ± 20.6 | 55.8 ± 24.8$^{b}$ | 75.2 ± 58.1 | 81.1 ± 50.8 | 78.2 ± 53.7$^{b}$ |
| OC | 20.1 ± 19.2 | 21.0 ± 24.8 | 20.6 ± 21.9$^{a}$ | 7.9 ± 2.6 | 9.5 ± 3.4 | 8.7 ± 3.1$^{b}$ | 8.7 ± 3.4$^{*}$ | 6.8 ± 4.3$^{*}$ | 7.8 ± 3.9$^{b}$ | 9.4 ± 3.8 | 10.1 ± 3.7 | 9.7 ± 3.7$^{b}$ |
| EC | 3.9 ± 3.1 | 4.7 ± 5.8 | 4.3 ± 4.6$^{a}$ | 1.9 ± 1.1$^{*}$ | 2.7 ±1.4$^{*}$ | 2.3 ± 1.3$^{b, c}$ | 1.4 ± 1.0 | 1.3 ± 1.0 | 1.3 ± 1.0$^{c}$ | 2.4 ± 1.4$^{*}$ | 3.4 ± 1.7$^{*}$ | 2.9 ± 1.6$^{b}$ |
| OC/EC | 4.6 ± 1.1 | 4.3 ± 1.3 | 4.5 ± 1.2$^{a, b}$ | 5.2 ± 2.1 | 4.4 ± 2.5 | 4.8 ± 2.3$^{a, b}$ | 6.7 ± 3.9 | 5.3 ± 4.2 | 6.1 ± 4.1$^{a}$ | 4.4 ± 1.6$^{*}$ | 3.3 ± 0.9$^{*}$ | 3.8 ± 1.4$^{b}$ |
| WSOC | 11.4 ± 11.3 | 12.0 ± 16.4 | 11.7 ± 13.9$^{a}$ | 4.1 ± 2.0 | 4.7 ± 2.6 | 4.4 ± 2.3$^{b}$ | 5.3 ± 2.1$^{*}$ | 4.0 ± 2.7$^{*}$ | 4.7 ± 2.5$^{b}$ | 4.7 ± 3.0 | 4.9 ± 2.8 | 4.8 ± 2.8$^{b}$ |
| WSOC/OC | 0.53 ± 0.08 | 0.51 ± 0.08 | 0.52 ± 0.08$^{b}$ | 0.50 ± 0.10 | 0.47 ± 0.14 | 0.49 ± 0.12$^{b}$ | 0.62 ± 0.11 | 0.59 ± 0.10 | 0.60 ± 0.11$^{a}$ | 0.47 ± 0.12 | 0.46 ± 0.12 | 0.46 ± 0.12$^{b}$ |
| MH-WSOC | 7.9 ± 7.6 | 8.0 ± 10.3 | 8.0 ± 8.9$^{a}$ | 2.8 ± 1.3 | 2.9 ± 1.6 | 2.9 ± 1.5$^{b}$ | 4.1 ± 1.2 | 3.4 ± 1.6 | 3.8 ± 1.5$^{b}$ | 2.9 ± 1.6 | 2.9 ± 1.3 | 2.9 ± 1.5$^{b}$ |
| SH-WSOC | 3.2 ± 3.8 | 4.0 ± 6.1 | 3.6 ± 5.0$^{a}$ | 1.3 ± 0.9 | 1.8 ± 1.1 | 1.6 ± 1.0$^{b}$ | 1.2 ± 1.0$^{*}$ | 0.7 ± 1.1$^{*}$ | 1.0 ± 1.1$^{b}$ | 1.8 ± 1.4 | 2.0 ± 1.5 | 1.9 ± 1.4$^{b}$ |
| **Organic tracers (ng m$^{-3}$)** | | | | | | | | | | | | |
| Levoglucosan | 307 ± 300 | 388 ± 394 | 349 ± 348$^{a}$ | 100 ± 87.8$^{*}$ | 194 ± 175$^{*}$ | 147 ± 144$^{b}$ | 23.6 ± 11.0 | 34.2 ± 24.2 | 28.9 ± 19.3$^{c}$ | 136 ± 102$^{*}$ | 234 ± 125$^{*}$ | 185 ± 123$^{b}$ |
| Cholesterol | 5.0 ± 3.0 | 4.9 ± 3.3 | 4.9 ± 3.1$^{a, b}$ | 3.9 ± 1.9 | 4.8 ± 2.5 | 4.3 ± 2.3$^{b}$ | 4.1 ± 2.4 | 3.0 ± 1.1 | 3.6 ± 1.9$^{b}$ | 6.1 ± 4.4 | 6.3 ± 3.1 | 6.2 ± 3.8$^{a}$ |
| Phthalic acid | 88.7 ± 84.8 | 90.8 ± 121 | 89.8 ± 103$^{a}$ | 27.3 ± 20.8 | 21.9 ± 14.0 | 24.6 ± 17.7$^{b}$ | 55.9 ± 22.0$^{*}$ | 17.6 ± 9.1$^{*}$ | 36.8 ± 25.5$^{b}$ | 27.6 ± 21.8 | 19.9 ± 13.3 | 23.8 ± 18.2$^{b}$ |
| 4-Methyl-5-nitrocatechol | 24.7 ± 26.4 | 35.2 ± 41.0 | 30.1 ± 34.5$^{a}$ | 1.8 ± 1.9 | 3.3 ± 2.7 | 2.6 ± 2.4$^{b}$ | 0.1 ± 0.3 | 0.0 ± 0.0 | 0.1 ± 0.2$^{b}$ | 1.6 ± 1.2$^{*}$ | 4.4 ± 3.6$^{*}$ | 3.0 ± 3.0$^{b}$ |
| 2-Methylerythritol | 2.1 ± 2.3 | 2.2 ± 3.5 | 2.2 ± 2.9$^{b}$ | 1.2 ± 0.6 | 1.5 ± 0.8 | 1.4 ± 0.7$^{b}$ | 55.4 ± 48.5 | 41.6 ± 34.6 | 48.5 ± 42.0$^{a}$ | 2.3 ± 1.1 | 2.6 ± 1.2 | 2.5 ± 1.1$^{b}$ |
| 3-Hydroxyglutaric acid | 4.4 ± 3.9 | 4.2 ± 5.0 | 4.3 ± 4.5$^{b}$ | 4.2 ± 2.8 | 4.9 ± 5.1 | 4.6 ± 4.0$^{b}$ | 37.1 ± 22.7$^{*}$ | 27.3 ± 18.5$^{*}$ | 32.2 ± 20.9$^{a}$ | 7.5 ± 4.6 | 7.0 ± 4.4 | 7.2 ± 4.5$^{b}$ |
| *cis*-Pinonic acid | 3.3 ± 2.4 | 3.0 ± 2.1 | 3.2 ± 2.2$^{c}$ | 9.0 ± 6.0$^{*}$ | 6.9 ± 3.6$^{*}$ | 7.9 ± 5.0$^{a}$ | 7.3 ± 4.2 | 10.1 ± 6.0 | 8.7 ± 5.3$^{a}$ | 7.3 ± 3.0$^{*}$ | 3.6 ± 0.8$^{*}$ | 5.5 ± 2.9$^{b}$ |

[a, b, c] We performed one-way analysis of variance (ANOVA) to evaluate whether these mean values showed statistically significant differences ($p<0.05$) between two seasons. If two seasonal average values have one or more same superscripts, it means that they did not show significant differences ($p>0.05$). In contrast, if two average values do not have any same superscript, it means that they showed significant differences ($p<0.05$). For example, the PM$_{2.5}$ concentration was significantly higher in winter than in other seasons, but it did not show significant difference between spring and summer, spring and autumn, or summer and autumn. Besides, the EC concentration did not show significant difference between spring and summer, as well as between spring and autumn, but it showed significant difference between summer and autumn.

[*] We also performed paired t test to evaluate whether daytime and nighttime values showed statistically significant differences ($p<0.05$). The values with $^{*}$ as their superscripts showed statistically significant differences between day and night.

15. Fig. 1: I would recommend including a line that shows the average and standard deviation for the WSOC/OC values. Currently, eyeballing the values in Fig. 1, they look fairly similar in all seasons.

**Response:** We deeply appreciate this nice suggestion. As suggested by the reviewer, we have added four lines which represented the average WSOC/OC values in each season, and four shaded areas that showed the standard deviations of WSOC/OC values.

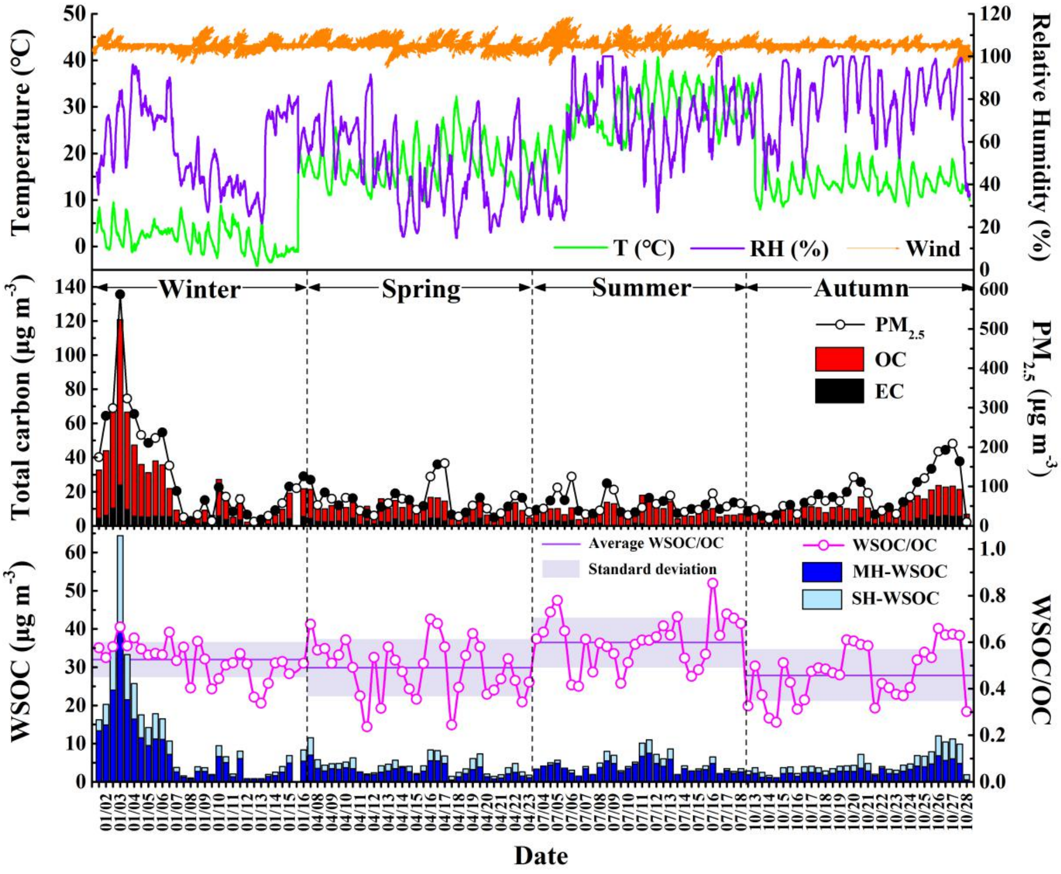


**Figure 1**. Temporal variations of meteorological parameters, the mass concentrations of $PM_{2.5}$, OC, EC, WSOC and WSOC/OC ratio in Beijing during the sampling periods in four seasons of 2017.

16. Minor: Line 30: replace “takes up” with “composes”;

Line 117: replace “entirely dryness” with either "entirely dry" or "entire dryness";

Line 121: replace “T The” with “The”;

Line 248: replace “did not appear any” with “did not have any”;

Line 280: replace “association” with “correlation”;

Line 294: replace “appeared” with “showed”;

Line 340: replace “except” with “beyond”;

Line 377: remove “Nevertheless”;

Line 399: remove “of” after Both;

Line 408: believe “C” is missing after “SO”;

Line 427: replace “was in consistent” with “was consistent”.

**Response:** We deeply appreciate the reviewer’s careful and detailed comments. We have corrected these errors in the revised manuscript. Thank you very much!

# Response to Reviewer 2

**Overall comment:** This work reported 4-season filter-based WSOC measurements including tracer measurements and group separation of the aqueous extracts into so-called hydrophobic and hydrophilic fractions by the SPE method. The sources of WSOC were speculated based on some correlations with $O_3$, RH, ALWC etc. The authors also conducted the PMF analysis to evaluate the source contributions. The problem is the quality of data analysis and discussion. Many of the discussions were not logically presented. Loss terms (e.g., photolysis, chemical reactions, gas-particle partitioning) were generally ignored. Conclusions about the relative contributions of photochemical vs aqueous pathways were made mainly on the basis of simple correlations with $O_3$ or ALWC etc., which can be largely uncertain especially for the winter-haze episodes when all components of $PM_{2.5}$ including primary species were correlated with ALWC or RH. There is also a lack of sufficient information to validate the PMF analysis in this study. The presented PMF results seem quite arbitrary.

**Response:** We deeply appreciate the reviewer's rigorous consideration and valuable comments, which enabled us to essentially improve the analysis and interpretation of the data. We have fully considered and carefully addressed all the comments raised by the reviewer, and have thoroughly revised the manuscript accordingly. The potential effect of photolysis, chemical reactions and gas-particle partitioning have been discussed in the revised manuscript. The relative contributions of photochemical vs aqueous pathways have been more cautiously discussed in the revised manuscript, and the correlation analysis between the ratio of SOA tracers to OC with RH or ALWC has been added to further support some of our speculations. Moreover, for the PMF analysis, we have provided detailed information on the uncertainty calculation of the input data, the selection criteria for the optimal PMF solution, the diagnostic plots and the error estimation of the PMF results in the revised manuscript. Below the comments are our responses point by point, and the revisions have been indicated in the revised manuscript.

**Comment 1**: Page 1, Line 14; Page 2, Line 53-57; Page 4, Line 106-112: Different SPE columns and extraction procedures (e.g., pH) result in various fractions of the WSOC (Sullivan et al., 2006). The authors used SPE (Oasis HLB) to separate the "hydrophilic and hydrophobic" fractions of WSOC. However, as described by Kiss et al. (2002), the one-step SPE on Oasis HLB column is to separate the WSOC into moderately hydrophilic (retained on the column) and strongly hydrophilic (passed through the column) fractions. I think it is wrong to simply assign the retained fraction herein as "hydrophobic" or "mainly HULIS" and the passed-through fraction as typical "hydrophilic (short-chain dicarboxylic acids and saccharides)".

**Response:** We deeply appreciate the reviewer's valuable comment. We used the same SPE column and separating procedure as described by Kiss et al. (2002) to separate different portions of WSOC in this study. As suggested by the reviewer, we have changed the term "hydrophobic WSOC" to "moderately hydrophilic WSOC", and "hydrophilic WSOC" to "strongly hydrophilic WSOC" throughout the revised manuscript.

To correctly identify different portions of WSOC, we have also reviewed the literature for support. As concluded by Kiss et al. (2002), the moderately hydrophilic WSOC (retained on the Oasis HLB column) is composed of humic-like substances (HULIS). This method is commonly used for the determination of atmospheric HULIS (Lin et al., 2010; Lin and Yu, 2011; Fan et al., 2016; Ma et al., 2018; Sengupta et al., 2018). Lin et al. (2010) suggested that several anhydrosugars (levoglucosan, xylose, sucrose) and short-chain organic acids (oxalic acid, succinic acid, malic acid) were present in the passed-through fraction. Therefore, we thought that it should be proper to assign the retained WSOC fraction as "mainly HULIS", and the passed-through portion as "strongly hydrophilic WSOC (short-chain dicarboxylic acids and saccharides)".

In Line 54-57, the references we cited (Verma et al., 2012, 2015; Yu et al., 2018) all used C-18 silica gel SPE columns to separate different WSOC fractions. Varga et al. (2001) compared the performance of C-18 silica gel columns and Oasis HLB columns, and suggested that Oasis HLB columns retained the same fraction of organic matter as silica-based columns. Hence, the statement that "previous studies have revealed that the moderately and strongly hydrophilic fractions of WSOC show significantly different intrinsic oxidative potential, thus would pose different effects on human health" should be proper here.

**References**

Kiss, G., Varga, B., Galambos, I., and Ganszky, I.: Characterization of water-soluble organic matter isolated from atmospheric fine aerosol, J. Geophys. Res. Atmos., 107(D21), 8339, doi:10.1029/2001JD000603, 2002.

Lin, P., Huang, X., He, L., and Yu, J.: Abundance and size distribution of HULIS in ambient aerosols at a rural site in South China, J. Aerosol. Sci., 41, 74-87, doi: 10.1016/j.jaerosci.2009.09.001, 2010.

Lin, P., and Yu, J.: Generation of reactive oxygen species mediated by humic-like substances in atmospheric aerosols, Environ. Sci. Technol., 45, 10362-10368, doi: 10.1021/es2028229, 2011.

Sengupta, D., Samburova, V., Bhattarai, C., Kirillova, E., Mazzoleni, L., Iaukea-Lum, M., Watts, A., Moosmuller, H., and Khlystov, A.: Light absorption by polar and non-polar aerosol compounds from laboratory biomass combustion, Atmos. Chem. Phys., 18, 10849-10867, doi:10.5194/acp-18-10849-2018, 2018.

Varga, B., Kiss, G., Ganszky, I., Gelencser, A., and Krivacsy, Z.: Isolation of water soluble organic matter from atmospheric aerosol, Talanta, 55, 561-572, 2011.

**Comment 2:** Introduction: Previous understanding of the characteristics of WSOC and its separated fractions as well as their primary and secondary sources were poorly summarized in the current Introduction section.

**Response:** We deeply appreciate the reviewer's valuable comment. The composition and sources of WSOC and its separated fractions have been summarized and added in the Introduction section of the revised manuscript. The corresponding revision is as follows.

Based on the solid phase extraction (SPE) by the Oasis HLB column, WSOC can be divided into the moderately hydrophilic fraction and strongly hydrophilic fraction (Varga et al., 2001; Kiss et al., 2002). The moderately hydrophilic fraction of WSOC mainly consists of humic-like substances (HULIS), which are an unresolved mixture of polycyclic ring structures with substituted hydroxyl, carboxyl, carbonyl, methoxy, and ester groups (Kiss et al., 2002; Lin et al., 2010; Fan et al., 2012). In addition, some smaller molecular with well-defined structures, such as phthalic acid and suberic acid, may also constitute a minor portion of the moderately hydrophilic fraction (Lin et al., 2010). The strongly hydrophilic fraction mainly includes low-molecular-weight organic acids (such as oxalic acid, succinic acid, malic acid) and anhydrosugars (such as levoglucosan, xylose, sucrose) (Lin et al., 2010). Previous studies have revealed that the moderately and strongly hydrophilic fractions of WSOC show significantly different intrinsic oxidative potential, thus would pose different effects on human health (Verma et al., 2012, 2015; Yu et al., 2018). However, source contributions of the moderately and strongly hydrophilic WSOC were scarcely investigated and compared in previous research.

Previous studies have suggested that coal combustion (Zhang et al., 2018; Li et al., 2019c), traffic emissions (Kawamura and Kaplan, 1987; Li et al., 2019c), residual oil combustion (Kuang et al., 2015), cooking (Qiu et al., 2020), soil dust and sea salts (Huang et al., 2006) can all contribute to WSOC. Nevertheless, it is most commonly recognized that WSOC mainly derives from the direct emissions of biomass burning and secondary organic aerosols (SOA) (Ding et al., 2008; Feng et al., 2013; Du et al., 2014; Zhang et al., 2018). For example, Du et al. (2014) suggested that SOA, biomass burning and other primary combustion sources contributed about 54 %, 40 % and 6 % respectively to WSOC in Beijing during 2010-2011. Zhang et al. (2018) also indicated that the sum of biomass burning and SOA contributed more than 80 % to WSOC in Beijing, Shanghai, Guangzhou and Xi’an in the wintertime of 2013. In recent years, the adjustment of energy and industrial structures as well as the effective control of the open burning activities in the surrounding areas of Beijing have posed significant impact on the source emissions. The average $PM_{2.5}$ concentration in Beijing has been greatly reduced from 89.5 $\mu g \cdot m^{-3}$ in 2013 to 58.0 $\mu g \cdot m^{-3}$ in 2017 since the implement of the Action Plan of Air Pollution Prevention and Control in 2013 (Cheng et al., 2019). Meanwhile, it has been reported that the oxidant concentrations were enhanced accompanying the decrease of $PM_{2.5}$ level, which might promote the SOA formation (Feng et al., 2019). Consequently, the sources and composition of WSOC in Beijing may show significant changes due to the control policies and enhanced atmospheric oxidizing capacity in the surrounding areas in recent years. Therefore, it is necessary to compare the source contributions of WSOC with those in the previous studies.

## References

Ding, X., Zheng, M., Yu, L., Zhang, X., Weber, R. J., Yan, B., Russell, A. G., Edgerton, E. S., and Wang, X.: Spatial and seasonal trends in biogenic secondary organic aerosol tracers and water-soluble organic carbon in the southeastern United States, Environ. Sci. Technol., 42, 5171-5176, doi:10.1021/es7032636, 2008.

Fan, X., Song, J., and Peng, P.: Comparison of isolation and quantification methods to measure humic-like substances (HULIS) in atmospheric particles, Atmos. Environ., 60, 366-374, doi:10.1016/j.atmosenv.2012.06.063, 2012.

Feng, J., Li, M., Zhang, P., Gong, S., Zhong, M., Wu, M., Zheng, M., Chen, C., Wang, H., and Lou, S.: Investigation of the sources and seasonal variations of secondary organic aerosols in $PM_{2.5}$ in Shanghai with organic tracer, Atmos. Environ., 79, 614-622, doi:10.1016/j.atmosenv.2013.07.022, 2013.

Huang, X., Yu, J., He, L., and Yuan, Z.: Water-soluble organic carbon and oxalate in aerosols at a coastal urban site in China: Size distribution characteristics, sources, and formation mechanisms, J. Geophys. Res. Atmos., 111, D22212, doi: 10.1029/2006JD007408, 2006.

Kawamura, K. and Kaplan, I. R.: Motor exhaust emissions as a primary source for dicarboxylic acids in Los-Angeles ambient air, Environ. Sci. Technol., 21, 105-110, 1987.

Kuang, B., Lin, P., Huang, X., and Yu, J.: Sources of humic-like substances in the Pearl River Delta, China: positive matrix factorization analysis of $PM_{2.5}$ major components and source markers, Atmos. Chem. Phys., 15, 1995-2008, doi: 10.5194/acp-15-1995-2015, 2015.

**Comment 3:** Page 4, Line 97 and Page 5, Line 124-125: How were the field blanks collected before and after sampling? What exactly were corrected?

**Response:** Thanks for the reviewer's careful concern. We have clarified the corresponding statement as follows.

The field blanks were collected before and after each sampling period, and a total of 8 field blank samples were obtained. The blank filters were put on the filter holder of the $PM_{2.5}$ sampler without pumping for 1 min, then stored and analyzed together with the ambient samples. Since two field blanks were obtained during each sampling period, the average concentrations of the targeted compounds on these two field blanks were used to correct the measured concentrations of the ambient samples in the corresponding season. Therefore, the reported concentrations of the targeted species were the measured concentrations on each ambient sample minus the average concentration on the two field blank samples.

**Comment 4:** Page 5, Line 129-131: Detailed information about the PMF analysis should be provided. The authors said that "the uncertainties were calculated referring to the measured RSD data of chemical analysis and previous studies". It is unclear to me whether this is a right approach. What do "the measured RSD data of chemical analysis and previous studies" mean specifically? Also, the authors said "The PMF model was run repeatedly to obtain a clear and reasonable source profile". How? The reasons of the selection of the numbers of PMF factors as well as the PMF uncertainty estimates and diagnostics are necessary.

**Response:** We deeply appreciate the reviewer's thoughtful comment. We have provided detailed information on the uncertainty calculation of input data, the selection criteria for the optimal solution and diagnostic plots. Besides, we have conducted the error estimates for the selected PMF solution, and found some uncertainties involved in this solution. To improve the reliability of the selected solution, we have added some constraints on the source profiles based on the priori information of the sources. Details of the PMF analysis are as follow.

## 1. Uncertainties of the input data

According to the User Guide of PMF5.0 (Norris et al., 2014), the uncertainties of the target species can be calculated as follow:

$$Unc = 5/6 \times MDL \ (c \leq MDL) \qquad (1)$$

$$Unc = \sqrt{(P \times c)^2 + (0.5 \times MDL)^2} \ (c > MDL) \qquad (2)$$

where $Unc$ is the data uncertainty, c is the concentration of target species, MDL is the method detection limit, and P is the error fraction. Since the User Guide did not give the calculation method for the error fraction (P), we estimated the P values referring to the measured relative standard deviations (RSD) of the target species. The RSD values were calculated by measuring six identical portions of an ambient sample. P was set as 10 % when RSD < 10 %, and 15 % or 20 % when RSD > 10 %.

## 2. Selection of base solutions

The chemical components input into the PMF model were selected based on our understanding of the possible WSOC sources (Norris et al., 2014). Interpretability was usually considered to be the most important factor for selecting the optimum solution (Shrivastava et al., 2007; Huang et al., 2014). Interpretable solutions are those which group tracers from different sources into distinct factors, while those solutions grouping tracers from multiple sources into one factor, distributing tracers for one source across multiple factors, or including factors with no distinct grouping of species are judged less interpretable (Shrivastava et al., 2007; Sowlat et al., 2016). In some literature, the optimal solution was defined as that with the maximum number of factors which had distinctive groupings of species, and explained at least 90 % of the total variable (Shrivastava et al., 2007).

In this study, PMF was run repeatedly by changing the number of factors and the start seed numbers. The base solution was selected based on: (1) the interpretability of the derived factor profiles and the temporal variations of source contributions; (2) the reconstruction of the total variable and $R^2$ of input organic tracers ($R^2$>0.90); (3) the scaled residuals of the input species. As presented in Figure S1, the 7-factor solution separated cholesterol (the tracer for cooking) into multiple sources. It was difficult to explain why cholesterol appeared in the factor profiles of biomass burning, dust and fresh biogenic SOC. Besides, this solution led to poor fits for cholesterol ($R^2 = 0.28$) and *cis*-pinonic acid ($R^2 = 0.32$), which were the key tracers selected in this study. Therefore, the 7-factor solution was not selected. As shown in Figure S2, the 8-factor solution also distributed cholesterol into multiple factors. This solution also resulted in a poor fit ($R^2 = 0.28$) for cholesterol. Therefore, the 8-factor solution was not chosen in this study. As shown in Figures S5-7, the solutions with 4 to 6 factors all showed poor interpretability for the derived factor profiles and poor fits for the key organic tracers. The 10-factor solution involved a factor without any tracer of high loading to indicate a specific source, thus could not be explained. By comparing the results with different factor numbers, the solution with 9 factors (Figure S3) was thought to be the most interpretable one.

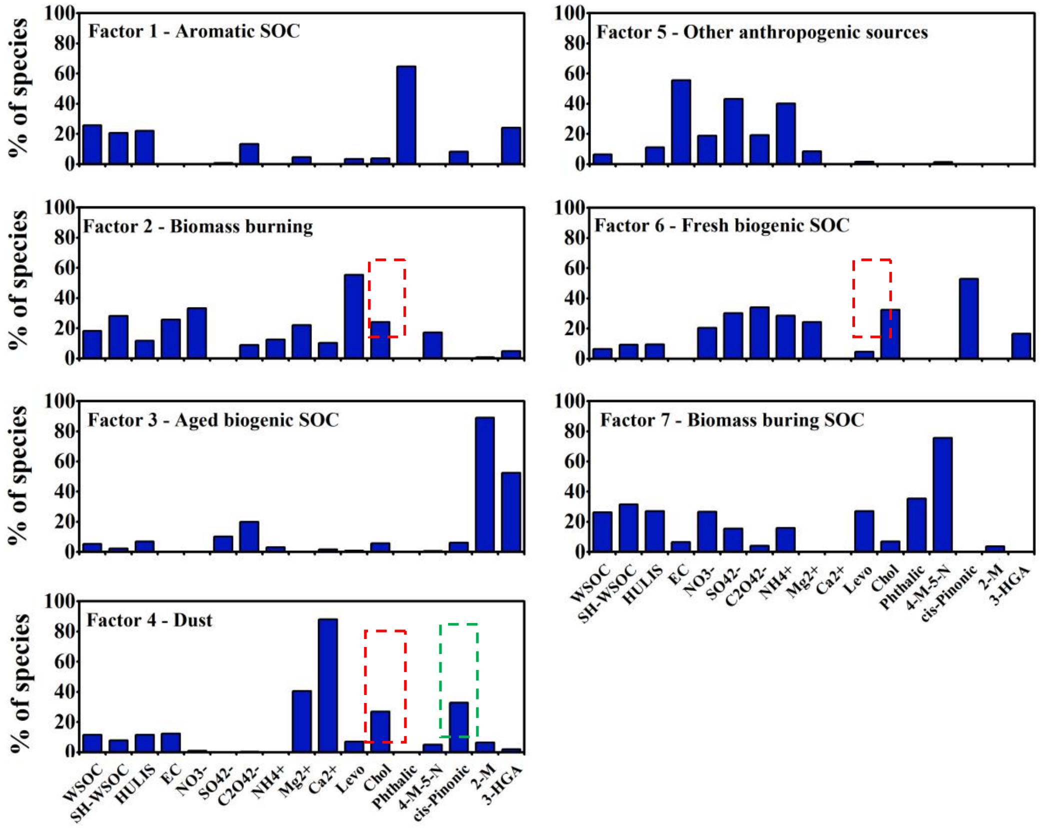


**Figure S1**. A 7-factor solution resolved by the PMF model.

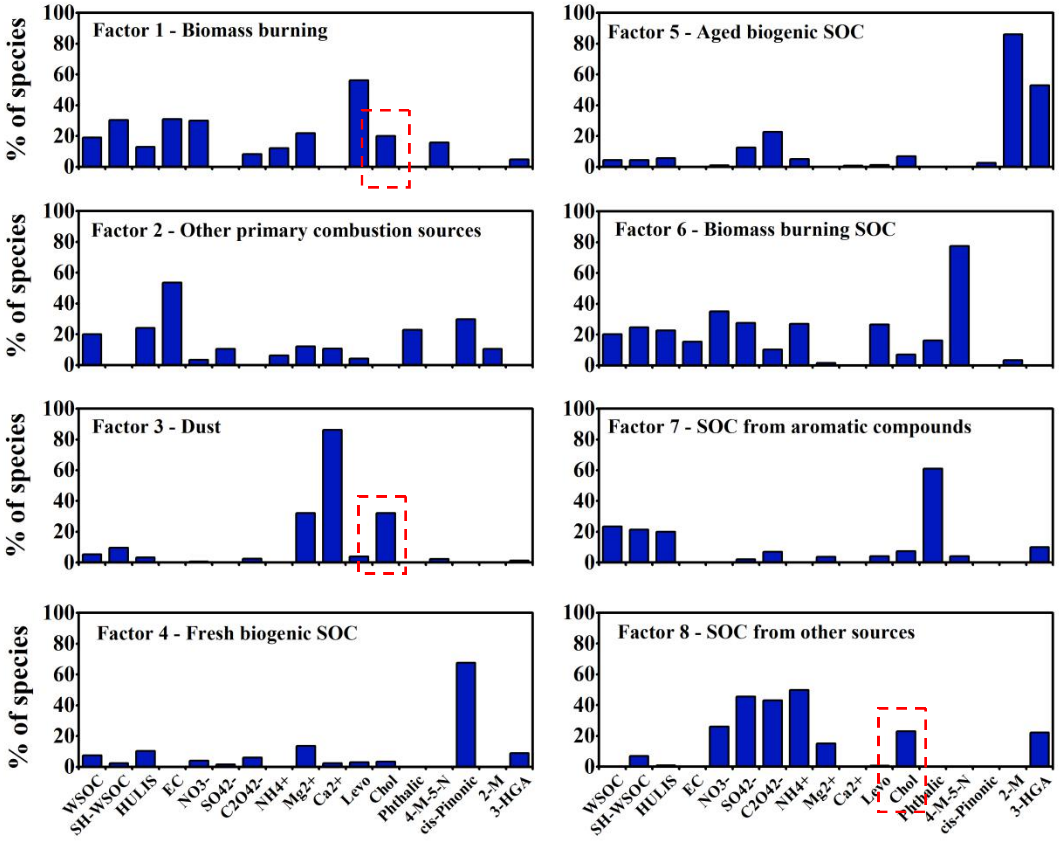


**Figure S2**. An 8-factor solution resolved by the PMF model.

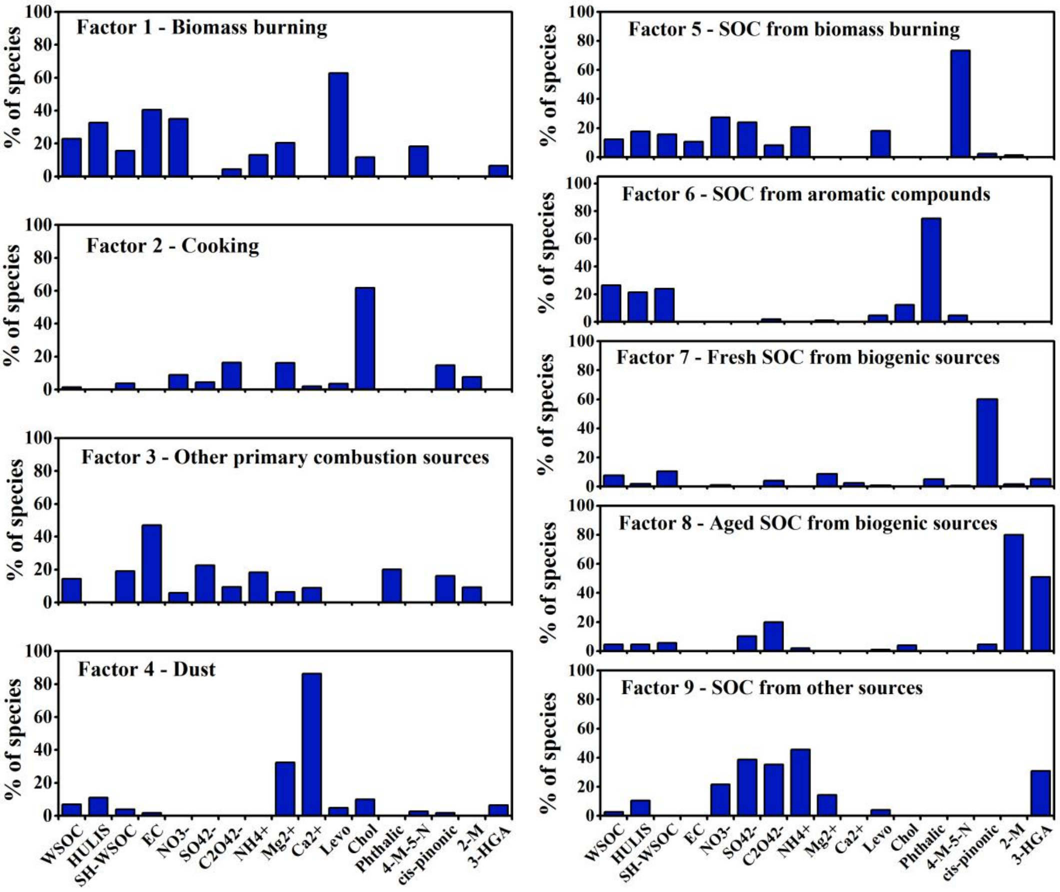


**Figure S3**. A 9-factor solution resolved by the PMF model.

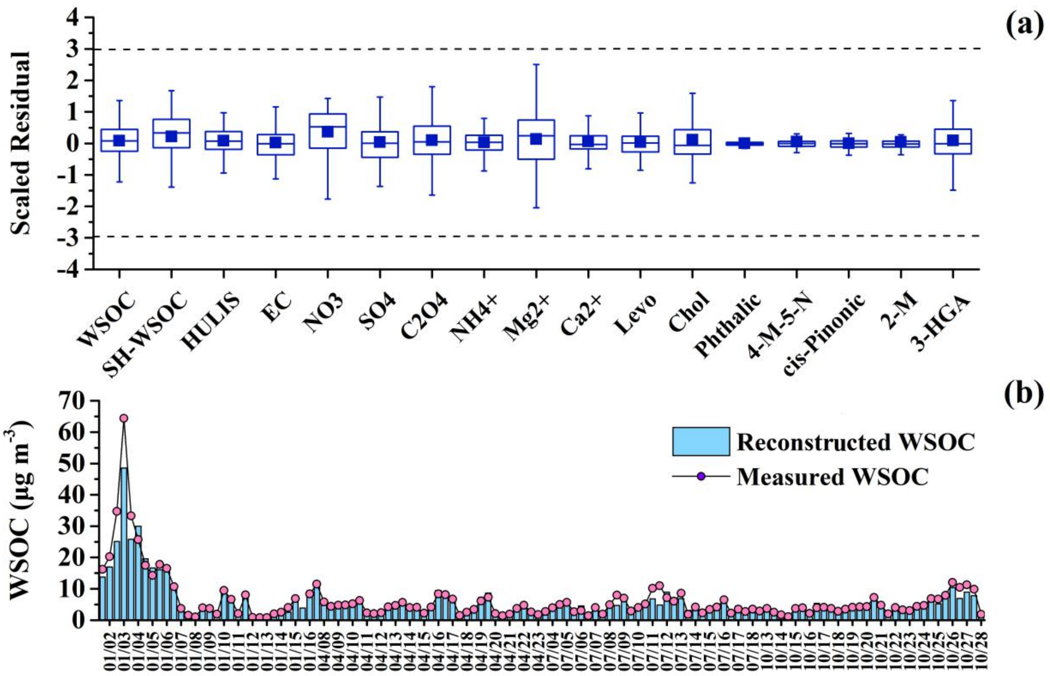


**Figure S4.** (a) The box plots showing the distributions of the scaled residuals for each species; (b) The time series of the measured WSOC and the reconstructed WSOC based on the 9-factor solution.

## 3. Diagnostics for the base model run

The selected 9-factor solution was converged, and Q(Robust) was similar to Q(True). As shown in Figure S4, most of the input species showed normally distributed residuals between -2.0 and +2.0, indicating that these species were well modeled. The $R^2$ of WSOC and HULIS were 0.94 and 0.93, respectively, and the $R^2$ for all the organic tracers were higher than 0.96, again suggesting that these species were well modeled.

## 4. Error estimation

The selected base solution was subjected to displacement (DISP) and bootstrap (BS) tests for error estimation. For the DISP test, the percent change in Q (%dQ) was less than 0.1 %, indicating that this solution was the global minimum (Paatero et al., 2014). No factor swapped for any value of $dQ^{max}$, indicating little rotational ambiguity in this solution (Paatero et al., 2014). For the BS test, the factor of “cooking” was mapped 79 % of the runs, the factor of “other primary combustion sources” was mapped 69 % of the runs, while other factors were mapped more than 91 % of the runs. The BS results indicated some uncertainties for the factors of cooking and other primary combustion sources, while the other factors were relatively stable.

Brown et al. (2015) indicated that the unstable solution might be due to too many factors involved. To investigate the effect of factor number on the stability of solutions, the BS results for solutions with different factor numbers were compared and shown in Table S3. As shown in Table S3, reducing the number of factors did not significantly increase the successful rates of BS mapping, but decreased the interpretability of the derived factor profiles. As recommended by the previous studies (Norris et al., 2014; Paatero et al., 2014), some constraints can be defined based on the priori information of the sources to reduce the variability of the solution.

**Table S3** The successful rates of BS mapping for the solutions with different numbers of factors. The values no more than 85 % were shown in bold.

| BS mapping | 3-factor solution | 4-factor solution | 5-factor solution | 6-factor solution | 7-factor solution | 8-factor solution | 9-factor solution |
|---|---|---|---|---|---|---|---|
| Factor 1 | 93 % | 93 % | 95 % | 95 % | **82 %** | 99 % | 99 % |
| Factor 2 | **83 %** | **85 %** | 99 % | 98 % | 88 % | 95 % | **79 %** |
| Factor 3 | 100 % | 95 % | **78 %** | 99 % | **74 %** | 100 % | 100 % |
| Factor 4 | | 87 % | **74 %** | **75 %** | **79 %** | 98 % | **69 %** |
| Factor 5 | | | 94 % | 91 % | 92 % | 92 % | 96 % |
| Factor 6 | | | | **81 %** | 99 % | **67 %** | 96 % |
| Factor 7 | | | | | 98 % | 100 % | 100 % |
| Factor 8 | | | | | | 98 % | 91 % |
| Factor 9 | | | | | | | 99 % |

## 5. Constrained model run

Bozzetti et al. (2017) exploited the markers' source specificity to set constraints for the profiles, so as to solve the problem of large mixtures of PMF factors associated with contributions of markers from different sources. They treated the contribution of unrelated source-specific markers as zero for each source, while non-source-specific variables were freely apportioned by the PMF algorithm. In addition, they set constraints for primary markers and combustion-related markers that can be seen as negligible in the secondary factors.

In the constrained model run, we set the constraints similar to those of Bozzetti et al. (2017), with a slight difference that we set the constraints by "soft pulling" so as to obtain a stable solution with a minimal change in the Q-value (dQ). The constraints were set as: (1) Levoglucosan was pulled up maximally with a limit of 0.25 % dQ for the factor of "primary biomass burning"; (2) Cholesterol was pulled up maximally with a limit of 0.50 % dQ for the factor of "cooking"; (3) Sulfate, *cis*-pinonic acid and 2-methylerythritol were pulled down maximally with limits of 0.25 % dQ for the factor of "other primary combustion sources"; (4) Phthalic acid was pulled up maximally with a limit of 0.25 % dQ for the factor profile of aromatic SOA. The dQ(Robust) for all the constraints were 0.93 % in the final constrained model run, which was acceptable (below 1 %) as recommended by the PMF user guide (Norris et al., 2014). As shown in Table S4, all the factors were mapped more than 94 % of the runs, suggesting that this solution was stable. Thus the constrained 9-factor solution was chosen as the final solution.

**Table S4** The successful rates of BS mapping for the constrained 9-factor solution.

| BS Mapping: | Base Factor 1 | Base Factor 2 | Base Factor 3 | Base Factor 4 | Base Factor 5 | Base Factor 6 | Base Factor 7 | Base Factor 8 | Base Factor 9 | Unmapped |
|---|---|---|---|---|---|---|---|---|---|---|
| Boot Factor 1 | 97 % | 0 | 0 | 0 | 3 % | 0 | 0 | 0 | 0 | 0 |
| Boot Factor 2 | 1 % | 95 % | 0 | 0 | 0 | 0 | 0 | 1 % | 3 % | 0 |
| Boot Factor 3 | 0 | 0 | 100 % | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Boot Factor 4 | 4 % | 0 | 0 | 94 % | 0 | 1 % | 0 | 1 % | 0 | 0 |
| Boot Factor 5 | 3 % | 0 | 0 | 0 | 97 % | 0 | 0 | 0 | 0 | 0 |
| Boot Factor 6 | 0 | 0 | 0 | 1 % | 0 | 99 % | 0 | 0 | 0 | 0 |
| Boot Factor 7 | 0 | 0 | 0 | 0 | 0 | 0 | 100 % | 0 | 0 | 0 |
| Boot Factor 8 | 0 | 0 | 0 | 0 | 0 | 1 % | 0 | 97 % | 2 % | 0 |
| Boot Factor 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 % | 0 |

## 6. Factor identification

The source profiles of the final solution are shown in Figure 4. Factor 1 showed high levels of levoglucosan and EC, thus was interpreted as the direct emissions from biomass burning. Factor 2 exhibited a high level of cholesterol, thus was regarded as cooking. Factor 3 exhibited a large fraction of EC that could not be explained by the direct emissions from biomass burning, suggesting that it was primary emissions from other combustion sources, such as coal combustion, traffic emissions and waste burning. Factor 4 was featured by high loadings of $Mg^{2+}$ and $Ca^{2+}$, thus was considered as dust. No significant EC but high fractions of 4-methyl-5-nitrocatechol and phthalic acid were observed in Factor 5 and Factor 6, respectively, which were recognized as SOC from biomass burning (biomass burning SOC) and aromatic precursors (aromatic SOC), respectively. Factor 7 exhibited a high level of *cis*-pinonic acid, thus was explained as the freshly generated biogenic SOC. Factor 8 was characterized by high fractions of 2-methylerythritol and 3-hydroxyglutaric acid, which are the end products from isoprene and monoterpenes respectively, thus was identified as the aged biogenic SOC. Note that 3-hydroxyglutaric acid and *cis*-pinonic acid were not grouped into one factor though they are both SOA tracers of monoterpenes, owing to their different oxidation degree as discussed above. Factor 9 covered the secondary components (such as $SO_4^{2-}$, $NO_3^-$, $NH_4^+$ and $C_2O_4^{2-}$) that can not be well explained by the identified sources above, thus was considered to be SOC from other sources. More detailed discussion is provided below the next comment (Comment 14).

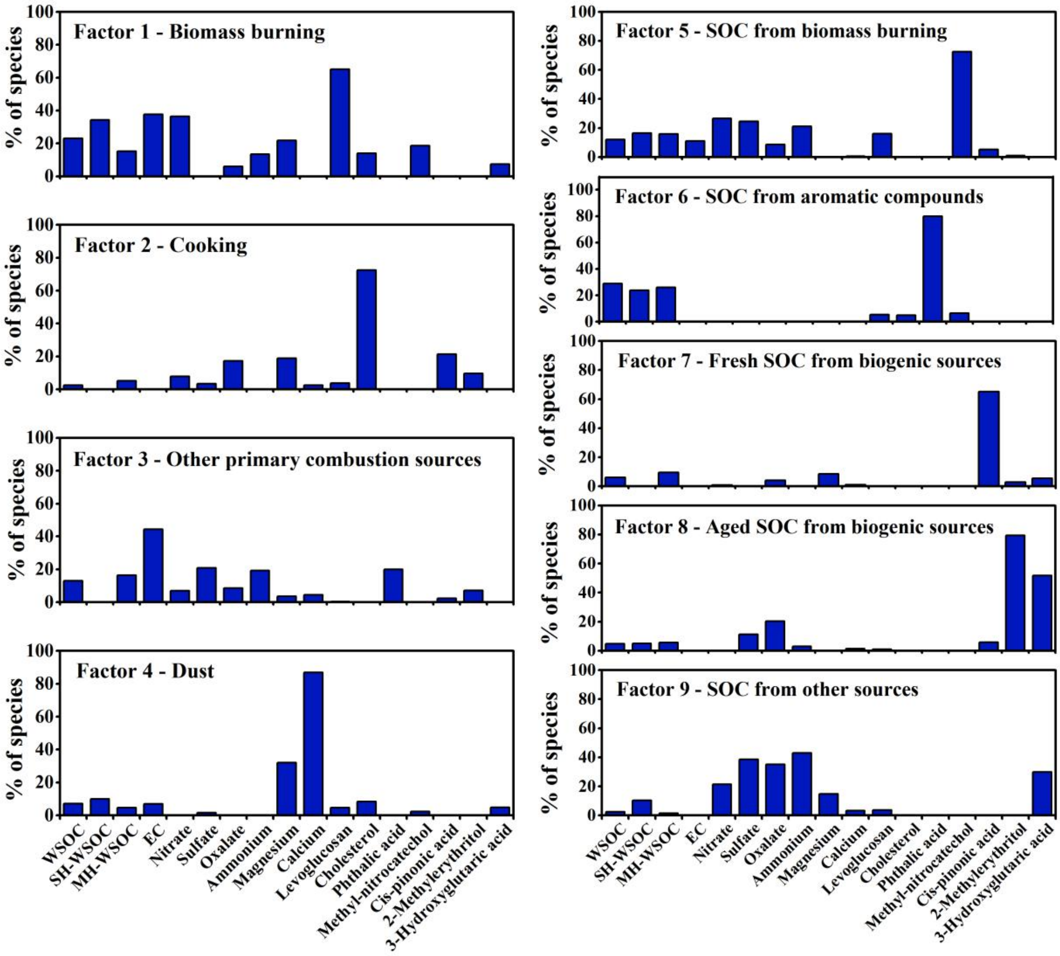


**Figure 4**. The constrained 9-factor solution resolved by the PMF model.

## References

Bozzetti, C., Sosedova, Y., Xiao, M., Daellenbach, K. R., Ulevicius, V., Dudoitis, V., Mordas, G., Bycenkiene, S., Plauskaite, K., Vlachou, A., Golly, B., Chazeau, B., Besombes, J. L., Baltensperger, U., Jaffrezo, J. L., Slowik, J. G., Haddad, I. E., and Prevot, A. S. H.: Argon offline-AMS source apportionment of organic aerosol over yearly cycles for an urban, rural, and marine site in northern Europe, Atmos. Chem. Phys., 17, 117-141, doi:10.5194/acp-17-117-2017, 2017.

Brown, S. G., Eberly, S., Paatero, P., and Norris, G. A.: Methods for estimating uncertainty in PMF solutions: Examples with ambient air and water quality data and guidance on reporting PMF results, Sci. Total Environ., 518-519, 626-635, doi:10.1016/j.scitotenv.2015.01.022, 2015.

Norris, G., Duvall, R., Brown, S., and Bai, S.: EPA positive matrix factorization (PMF) 5.0 fundamentals and user guide, US Environmental Protection Agency, Office of Research and Development, Washington, DC, 2014.

Paatero, P., Eberly, S., Brown, S. G., and Norris, G. A.: Methods for estimating uncertainty in factor analytic solutions, Atmos. Meas. Tech., 7, 781-797, doi: 10.5194/amt-7-781-2014, 2014.

Shrivastava, M. K., Subramanian, R., Rogge, W. F., and Robinsen, A. L.: Sources of organic aerosol: Positive matrix factorization of molecular marker data and comparison of results from different source apportionment models, Atmos. Environ., 41, 9353-9369, doi:10.1016/j.atmosenv.2007.09.016, 2007.

Sowlat, M. H., Hasheminassab, S., and Sioutas, C.: Source apportionment of ambient particle number concentrations in central Los Angeles using positive matrix factorization (PMF), Atmos. Chem. Phys., 16, 4849-4866, doi: 10.5194/acp-16-4849-2016, 2016.

**Comment 14:** Page 11, Line 314-327: The interpretation of the PMF factors is over simple here. It looks like the authors intentionally choose a PMF solution that has separate factors for individual tracers. However, how do the tracers correlate each other? (1) Previous studies found that methyl-nitrocatechol correlates with AMS BBOA and levoglucosan (Linuma et al. 2010; Mohr et al. 2013). If the two temporally correlates, the split of the biomass burning factor into two (primary and secondary) may be highly uncertain given the small sample size of this study. (2) Factor 7, 8, and 9 are all associated with biogenic SOA tracers. It was said in Page 10, Line 297-298 that 3-hydroxyglutaric acid correlated strongly with 2-methylerythritol. Then how and why to separate Factor 8 and 9. Is 3-hydroxyglutaric acid a unique tracer for monoterpene SOA? For day and night samples which did not maintain much oxidation process information (meaning first-generation vs multi-generation), I am surprised that there were two monoterpene SOA factors (one is marked by cis-pinonic acid and the other is marked by 3-hydroxyglutaric acid). (3) For Factor 3, the profile has significant contributions of secondary species, is it really primary?

**Response:** We deeply appreciate the reviewer's valuable comment. We understood the reviewer's concerns and would like to respond to the above concerns point by point.

## 1. The separation of primary biomass burning and secondary biomass burning SOA.

Indeed, levoglucosan correlated strongly with 4-methyl-5-nitrocatechol ($r=0.87$, $p<0.01$) in this study. However, even if we reduced the factor number from nine to five, levoglucosan and 4-methyl-5-nitrocatechol could not be merged into one source (Figure S1, 2, 5, 6). As presented in Figure S2, for the 8-factor solution, levoglucosan had a high loading in Factor 1, and 4-methyl-5-nitrocatechol showed a high loading in Factor 6. This solution distributed cholesterol into multiple sources and led to a poor fit for cholesterol ($R^2=0.28$), and the factor profile was less interpretable compared to the 9-factor solution. As shown in Figure S1, for the 7-factor solution, levoglucosan exhibited a high loading in Factor 2, and 4-methyl-5-nitrocatechol showed a high fraction in Factor 7. Cholesterol was again separated into distinct sources, leading to poor fits for both cholesterol ($R^2 = 0.28$) and *cis*-pinonic acid ($R^2 = 0.32$). As presented in Figure S5, for the 6-factor solution, levoglucosan had a high loading in Factor 4, and 4-methyl-5-nitrocatechol had a high fraction in Factor 6. Cholesterol, *cis*-pinonic acid and phthalic acid were divided into different sources, resulting in poor fits for cholesterol ($R^2 = 0.26$) and *cis*-pinonic acid ($R^2 = 0.25$). For the 5-factor solution (Figure S6), levoglucosan showed a high loading in Factor 3, and 4-methyl-5-nitrocatechol had a high fraction in Factor 5. Compared to the 6-factor solution, biomass burning SOC and aromatic SOC were merged into one factor, which was interpreted as the anthropogenic SOC. When the factor number decreased to four, levoglucosan and 4-methyl-5-nitrocatechol were merged into the same factor (Figure S8). However, this solution was less interpretable, and resulted in poorer fits for most of the species (cholesterol: $R^2=0.17$; *cis*-pinonic acid: $R^2=0.25$; $Ca^{2+}$: $R^2=0.67$; $Mg^{2+}$: $R^2=0.73$; $NO_3^-$: $R^2=0.75$; etc). Furthermore, the slope of the fitting equation for observed and predicted 4-methyl-5-nitrocatechol was only 0.31. And the high value of 4-methyl-5-nitrocatechol in winter was not reproduced by the 4-factor solution. Hence, the 4-factor solution was also excluded in this study.

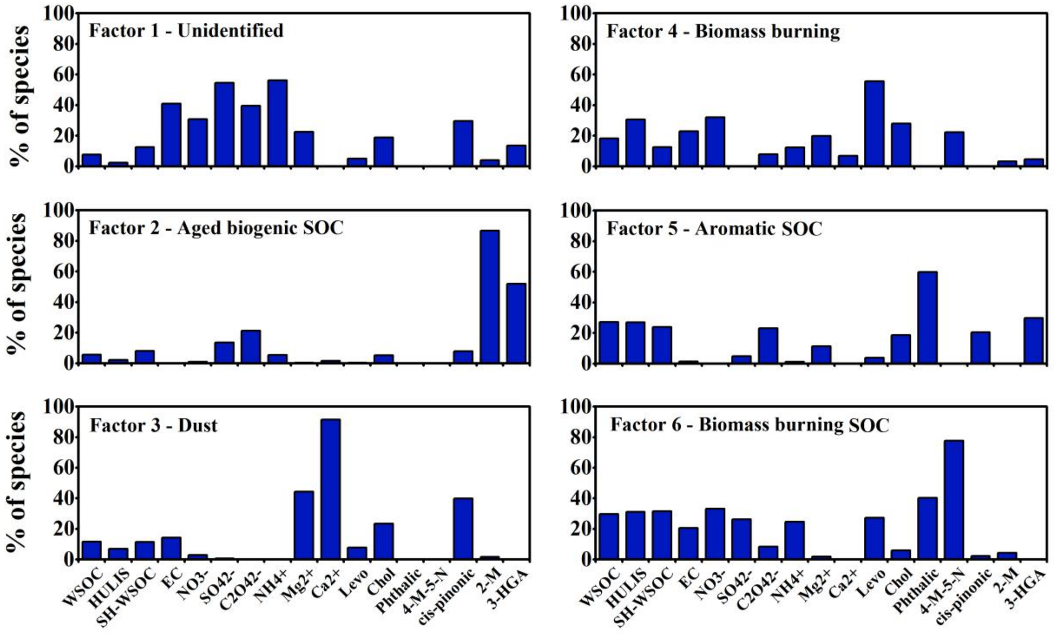


**Figure S5**. A 6-factor solution resolved by the PMF model.

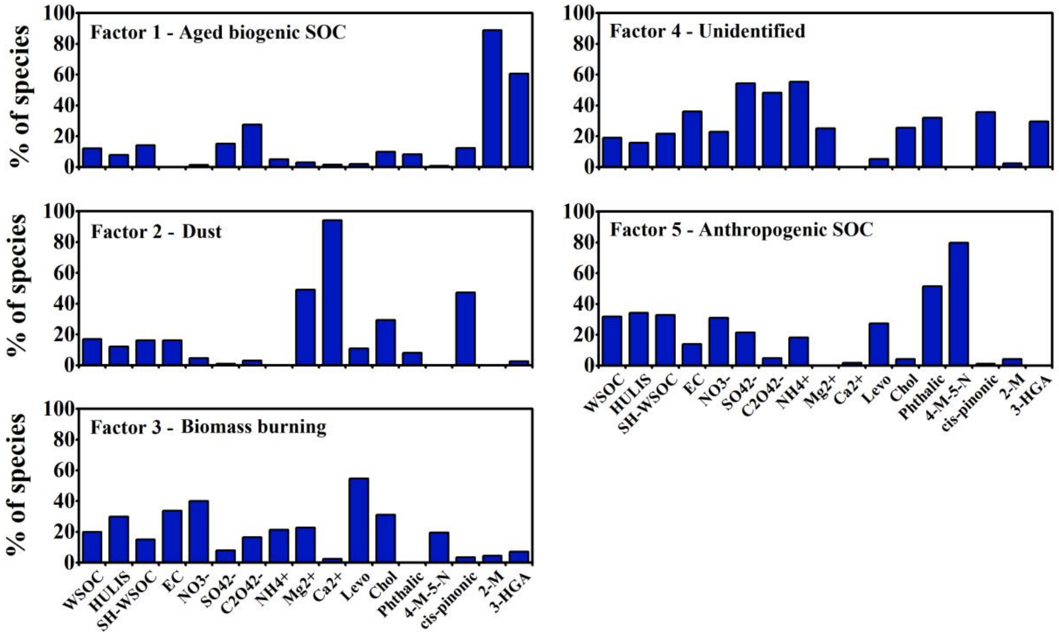


**Figure S6**. A 5-factor solution resolved by the PMF model.

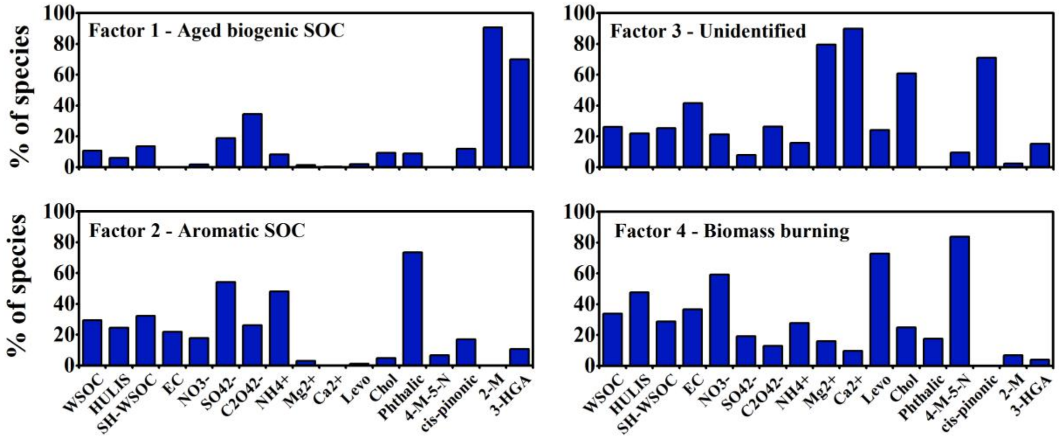


**Figure S7**. A 4-factor solution resolved by the PMF model.

It was indeed interesting that levoglucosan and 4-methyl-5-nitrocatechol were not distributed in the same factor, though they showed strong correlation with each other. We attempted to explain this phenomenon as follows. The ratio of 4-methyl-5-nitrocatechol to levoglucosan had significantly higher values ($p<0.01$) in winter than in other seasons (winter: 0.071 ± 0.029; other seasons: 0.010 ± 0.009), implying different types of biomass burning sources (primary and secondary). If they were merged into the same source, the ratio of 4-methyl-5-nitrocatechol to levoglucosan was supposed to be stable throughout the year, which was not the truth. As shown in Figure S8, the linear regression between levoglucosan and 4-methyl-5-nitrocatechol on the whole year scale (the blue line) was largely affected by the high values in winter. However, according to the uncertainty estimation method for the input species (Equation 2), the species with lower concentrations usually have lower uncertainties, thus may have a larger impact on the Q value. Taking the 4-factor solution (Figure S7) as an example, when the two tracers were merged into one factor, to minimize the Q value, the algorithm in the PMF model tended to assign a low value for the ratio of 4-methyl-5-nitrocatechol to levoglucosan in the factor profile of biomass burning (i.e. 0.024 in Factor 4). As presented in Figure S8, this ratio (orange line) was closer to the regression slope in other seasons (0.017), but much lower than that in winter (0.096). Consequently, the high concentration of 4-methyl-5-nitrocatechol in winter could not be reproduced at all by such PMF solution. In conclusion, the solution which merged these two tracers into the same factor might bring about large uncertainties, and fail to reproduce the peak values of 4-methyl-5-nitrocatechol over the study period in winter.

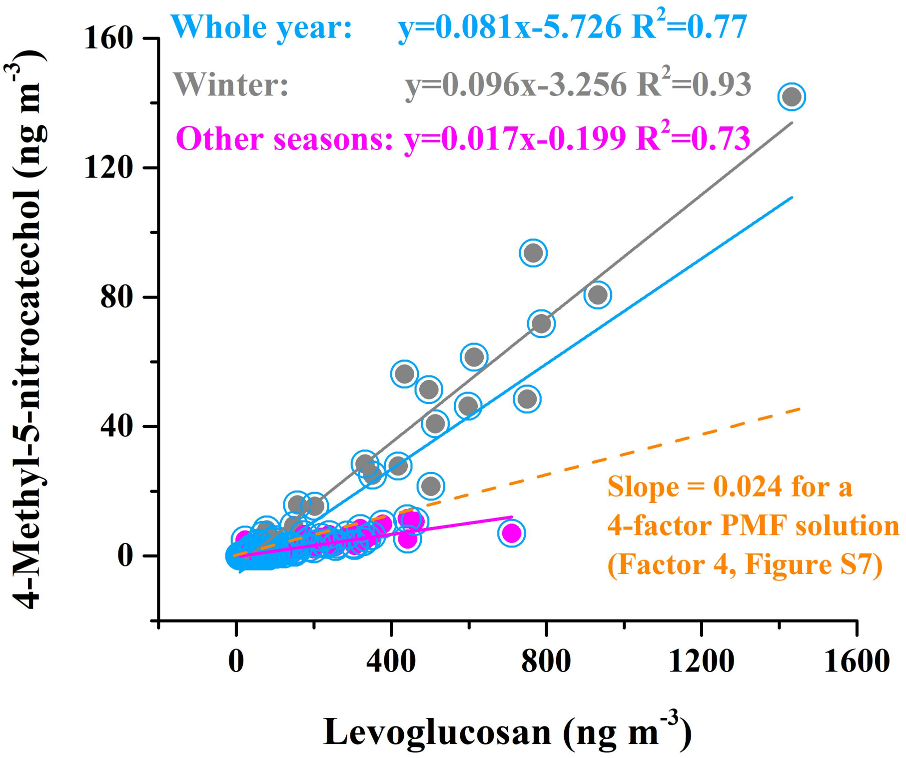


**Figure S8** The relationship between levoglucosan and 4-methyl-5-nitrocatechol.

Fresh biomass burning emissions show a high fraction of anhydrosugar, such as levoglucosan. The relative intensity of anhydrosugar decreased due to the degradation or oxidation reactions (Gilardon et al., 2016; Sengupta et al., 2020). Chamber studies indicated that substantial amounts of nitrogen-containing organic compounds, such as nitrophenols and methyl-nitrocatechols, were generated during aging (Bertrand et al., 2018; Hartikainen et al., 2018). 4-Methyl-5-nitrocatechol was recommended to be the secondary tracer for aged biomass burning SOA (Bertrand et al., 2018). In the previous studies (Gilardon et al., 2016; Zhou et al., 2017; Srivastava et al., 2018), the biomass burning source was also separated into primary and secondary fractions with the PMF model. In Gilardon et al. (2016), the factor profile of primary biomass burning was featured by a high loading of anhydrosugar (signal at m/z=60) but a low level of the aged OA (signal at m/z=44), while the factor profile of biomass burning

SOA was featured by a high fraction of the aged OA (signal at m/z=44) but a low fraction of anhydrosugar (signal at m/z=60) (Gilardon et al., 2016). Srivastava et al. (2018) also used levoglucosan and 4-methyl-5-nitrocatechol as tracers to differentiate the primary biomass burning and biomass burning SOA using the PMF model (92 samples, which was less than that in our study). In this study, as shown in Figure 4, Factor 1 showed high fractions of levoglucosan and EC, but a low fraction of 4-methyl-5-nitrocatechol, thus was considered as the direct emission from biomass burning. The concentration ratio of levoglucosan to WSOC in this factor was 0.085 μg $\mu g^{-1}$, which was similar to that measured in the primary combustion of crop straws (0.097 μg $\mu g^{-1}$), wood (0.081 μg $\mu g^{-1}$) and leaves (0.095 μg $\mu g^{-1}$) in North China (Yan et al., 2018). Factor 5 had a high level of 4-methyl-5-nitrocatechol, but low levels of EC and levoglucosan, thus was identified as biomass burning SOA.

## 2. The interpretation of Factor 7, Factor 8, and Factor 9.

As presented in Figures S1-3 and Figures S5-7, even if we reduced the factor number from nine to four, 2-methylerythritol, 3-hydroxyglutaric acid and *cis*-pinonic acid could not be merged into the same factor. Large fractions of 3-hydroxyglutaric acid and 2-methylerythritol were usually grouped into one factor, since they strongly correlated with each other ($r=0.94$, $p<0.01$). *Cis*-pinonic acid could not be distributed in this factor since it correlated less strongly with 2-methylerythritol ($r=0.51$, $p<0.01$) and 3-hydroxyglutaric acid ($r=0.58$, $p<0.01$). As stated in Section 3.2, *cis*-pinonic acid is a lower-generation oxidative product from monoterpenes, while 2-methylerythritol and 3-hydroxyglutaric acid are more aged products from isoprene and monoterpenes, respectively (Kourtchev et al., 2009). Hence, Factor 7 with a high level of *cis*-pinonic acid was interpreted as the fresh biogenic SOC, and Factor 8 with high loadings of 2-methylerythritol and 3-hydroxyglutaric acid was interpreted as aged biogenic SOC. As shown in Figure 5, the seasonal variation of their source contributions also supported this interpretation.

Since the major fraction of 3-hydroxyglutaric acid was distributed in Factor 8, it was not proper to interpret Factor 9 as monoterpene SOC. In fact, as shown in Figure

S1-3 and Figure S5-7, a minor fraction of 3-hydroxyglutaric acid was always distributed in factors other than the biogenic SOC. It was more interpretable when this minor fraction of 3-hydroxyglutaric acid was distributed in the same factor together with $SO_4^{2-}$, $NO_3^-$, $NH_4^+$ and $C_2O_4^{2-}$. In this case, Factor 9 of the selected 9-factor solution could be interpreted as a mixed secondary source and explain the secondary species that were not well fitted by other identified secondary sources. Similar factor profile has also been resolved in the literature, and was usually interpreted as the "inorganic-rich SOA" (Huang et al., 2014).

## 3. The interpretation of Factor 3 (Other primary combustion sources).

As shown in Figure 4, for the constrained 9-factor solution, Factor 3 exhibited a significant fraction of EC from sources other than the direct emissions of biomass burning, implying that it could be associated with the primary emissions from other combustion sources, such as coal combustion, traffic emissions, and waste burning, etc. Indeed, a minor fraction of $SO_4^{2-}$ (20.8 %), $NH_4^+$ (19.3 %) and phthalic acid (20.0 %) were also distributed in Factor 3. Previous studies have found that $SO_4^{2-}$, $NO_3^-$, $NH_4^+$ and phthalic acid can also be directly emitted from coal combustion (Zhang et al., 2008; Dai et al., 2019) and traffic emissions (Al-Naiema and Stone, 2017; Hao et al., 2019). Dai et al. (2019) suggested that primary $SO_4^{2-}$ accounted for 38.9 % and 16.9 % to the total $SO_4^{2-}$ in $PM_{2.5}$ during the heating and non-heating seasons respectively. Accordingly, such distribution of $SO_4^{2-}$ (20.8 %) in Factor 3 was acceptable. Similar loadings of $SO_4^{2-}$ and $NH_4^+$ were also found in the factor profile of coal combustion in the previous source apportionment study (Huang et al., 2014). To sum up, it was reasonable that a minor fraction of $SO_4^{2-}$, $NH_4^+$ and phthalic acid presented in the factor profile of Factor 3, i.e., other primary combustion sources.

## References

Bertrand, A., Stefenelli, G., Jen, C. N., Pieber, S. M., Bruns, E. A., Ni, H., Temimeroussel, B., Slowik, J. G., Goldstein, A. H., Haddad, I. E., Baltensperger, U., Prevot, A. S. H., Wortham, H., and Marchandet, N.: Evolution of the chemical fingerprint of biomass burning organic aerosol during aging, Atmos. Chem. Phys., 18, 7607-7624, doi:10.5194/acp-18-7607-2018, 2018.

Dai, Q., Bi, X., Song, W., Li, T., Liu, B., Ding, J., Xu, J., Song, C., Yang, N., Schulze,

B. C., Zhang, Y., Feng, Y., and Hopke, P. K.: Residential coal combustion as a source of primary sulfate in Xi'an, China, Atmos. Environ., 196, 66-76, doi: 10.1016/j.atmosenv.2018.10.002, 2019.

Hao, Y., Gao, C., Deng, S., Yuan, M., Song, W., Lu, Z., and Qiu, Z.: Chemical characterisation of $PM_{2.5}$ emitted from motor vehicles powered by diesel, gasoline, natural gas and methanol fuel, Sci. Total Environ., 674, 128-139, doi: 10.1016/j.scitotenv.2019.03.410, 2019.

Sengupta, D., Sengupta, V., Bhattarai, C., Watts, A. C., Moosmuller, H., and Khlystov, A. Y.: Polar semivolatile organic compounds in biomass-burning emissions and their chemical transformations during aging in an oxidation flow reactor, Atmos. Chem. Phys., 20, 8227-8250, doi:10.5194/acp-20-8227-2020, 2020.

Srivastava, D., Favez, O., Bonnaire, N., Lucarelli, F., Lucarelli, M., Perraudin, E., Gros, V., Villenave, E., and Albinet, A.: Speciation of organic fractions does matter for aerosol source apportionment. Part 2: Intensive short-term campaign in the Paris area (France), Atmos. Environ., 634, 267-278, doi: 10.1016/j.scitotenv.2018.03.296, 2018.

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, doi:10.1021/es7022576, 2008.

Zhou, S., Collier. S., Jaffe, D. A., Briggs, N. L., Hee, J., Sedlacek III, A. J., Kleinman, L., Onasch, T. B., and Zhang, Q.: Regional influence of wildfires on aerosol chemistry in the western US and insights into atmospheric aging of biomass burning organic aerosol, Atmos. Chem. Phys., 17, 2477-2493, doi: 10.5194/acp-17-2477-2017, 2017.

**Comment 5:** Page 5, Line 140-141: The equation calculates ion balance not “aerosol acidity”.

**Response:** We deeply appreciate the reviewer’s valuable suggestion. According to Pye et al. (2020), we changed the ratio of $R_{A/C}$ to the approximate value of aerosol pH ($pH_F$) to denote the aerosol acidity. The $pH_F$ value was estimated using the ISORROPIA-II model. And the molality of $H^+$ ($m_{H^+}$), which was calculated by $m_{H^+} = 10^{-pH_F}$, was used instead for the correlation analysis in the revised manuscript.

**References**

Pye, H. O. T., Nenes, A., Alexander, B., Ault, A. P., Barth, M. C., Clegg, S. L., Collett, J. L., Fahey, K. M., Hennigan, C. J., Herrmann, H., Kanakidou, M., Kelly, J. T., Ku, I. T., McNeill, V. F., Riemer, N., Schaefer, T., Shi, G., Tilgner, A., Walker, J.

T., Wang, T., Weber, R., Xing, J., Zaveri, R. A., and Zuend, A.: The acidity of atmospheric particles and clouds, Atmos. Chem. Phys., 20, 4809-4888, doi:10.5194/acp-20-4809-2020, 2020.

**Comment 6:** Page 6, Line 166: Why would the reduction of open biomass burning lead to decreased WSOC/OC ratios? Please clarify and cite references to support the reduction of open biomass burning.

**Response:** Thanks for the reviewer's valuable comment. We deleted this sentence after careful consideration. Some previous studies suggested that the ratios of WSOC/OC in biomass burning aerosols were higher than those from other primary combustion sources (Ruellan and Cachier, 2001; MayolBracero et al., 2002; Yang et al., 2020). However, the formation of SOC can also significantly influence the ratio of WSOC/OC (Jaffrezo et al., 2005; Deshmukh et al., 2016). The influencing factors of the WSOC/OC ratio are actually complex, thus we have deleted this uncertain speculation in the revised manuscript.

**References**

Deshmukh, D. K., Kawamura, K., Lazaar, M., Kunwar, B., and Boreddy, S. K. R.: Dicarboxylic acids, oxoacids, benzoic acid, α-dicarbonyls, WSOC, OC, and ions in spring aerosols from Okinawa Island in the western North Pacific Rim: size distributions and formation processes, Atmos. Chem. Phys., 16, 5263-5282, doi: 10.5194/acp-16-5263-2016, 2016.

Jaffrezo, J. L., Aymoz, G., Delaval, C., and Cozic, J.: Seasonal variations of the water soluble organic carbon mass fraction of aerosol in two valleys of the French Alps, Atmos. Chem. Phys., 5, 2809-2821, 2005.

MayolBracero, O. L., Guyon, P., Graham, B., Roberts, G., Andreae, M. O., Decesari, S., Facchini, M. C., Fuzzi, S., and Artaxo, P.: Water-soluble organic compounds in biomass burning aerosols over Amazonia 2. Apportionment of the chemical composition and importance of the polyacidic fraction, J. Geophys. Res. Atmos., 107, D20, 8091, doi:10.1029/2001JD000522, 2002.

Ruellan, S., and Cachier, H.: Characterisation of fresh particulate vehicular exhausts near a Paris high flow road, Atmos. Environ., 35, 453-468, 2011.

Yang, Y., Qin, J., Qi, T., Zhou, X., Chen, R., Tan, J., Xiao, K., Ji, D., He, K., and Chen, X.: Fluorescence characteristics of particulate water-soluble organic compounds emitted from coal-fired boilers, Atmos. Environ., 223, 117297, doi: 10.1016/j.atmosenv.2020.117297, 2020.

**Comment 7:** Page 6, Line 173-176: The correlations (r = 0.44-0.58) are not strong. I think it is difficult to conclude that OA became more hygroscopic as pollution aggravated. Indeed, if primary sources make a large contribution, e.g., in winter when coal combustion was enhanced, OA might not be more hygroscopic although its concentration became greater.

**Response:** We deeply appreciate the reviewer's valuable suggestion. We have deleted this statement in the revised manuscript.

**Comment 8:** Page 6, Line 180: Perhaps remove "ideal". Levoglucosan is not chemically inert. It is also not a unique tracer for biomass burning. As the authors mentioned in Page 7, Line 199, biofuel and coal burning are also sources of levoglucosan.

**Response:** We agree with the reviewer's valuable suggestion and have removed "ideal" in the revised manuscript.

**Comment 9:** Methylnitrocatechol is not necessarily secondary. Wang X et al. (2017) showed primary emissions of methylnitrocatechol from biomass burning. Coal burning is also a source. The statement that "4-methyl-5-nitrocatechol is a good indicator for biomass burning SOA" is perhaps inappropriate.

**Response:** Thanks for the reviewer's valuable comment. We have removed the word "good" and revised the statement to: "4-methyl-5-nitrocatechol can be used as a tracer for biomass burning SOA (Iinuma et al., 2010; Bertrand et al., 2018; Srivastava et al., 2018)".

Iinuma et al. (2010) proposed that 4-methyl-5-nitrocatechol could be used as a tracer for biomass burning SOA. Bertrand et al. (2018) investigated the evolution of chemical fingerprint of biomass burning OA during aging, and sorted the tracers into three types: primary, non-conventional primary and secondary tracers. They also concluded that 4-methyl-5-nitrocatechol was the secondary tracer for biomass burning, and could be used as an aged biomass burning tracer in source apportionment studies. Srivastava et al. (2018) also used 4-methyl-5-nitrocatechol in the PMF model as the tracer for biomass burning SOA.

Wang et al. (2017) reported that the ratio of primary 4-methyl-5-nitrocatechol to OC varied greatly for different biomass types and burning conditions. The ratio for corncob under the flaming condition (0.87 ng $\mu g^{-1}$) was much higher than that for the other biomass types or other burning conditions (0.11 ng $\mu g^{-1}$ on average). Therefore, we used the ratio of primary 4-methyl-5-nitrocatechol to OC for corncob burning (0.87 ng $\mu g^{-1}$) to roughly estimate the maximum contribution of primary biomass burning to the concentration of 4-methyl-5-nitrocatechol in the atmosphere. Recent source apportionment studies in Beijing suggested that biomass burning OA contributed at most 14 % of total OA (Sun et al., 2018; Tang et al., 2018; Duan et al., 2020). It was thus estimated that at most 16.2 % of atmospheric 4-methyl-5-nitrocatechol was from primary biomass burning. Therefore, it should be proper to use 4-methyl-5-nitrocatechol to indicate biomass burning SOA.

**References**

Wang, X., Gu, R., Wang, L., Xu, W., Zhang, Y., Chen, B., Li, W., Xue, L., Chen, J., and Wang, W.: Emissions of fine particulate nitrated phenols from the burning of five common types of biomass, Environ. Pollut., 230, 405-412, doi: 10.1016/j.envpol.2017.06.072, 2017.

**Comment 10:** Page 7, Line 200-203; Line 210-218; Page 8, Line 237-243: Errors can be propagated to the ratios in Figure 2b so that the day/night CO-scaled ratios can be discussed statistically (not just seasonal mean values). The authors said the CO-scaled concentration of cholesterol was close for the whole sampling period. However, panel a is in log scale. I think it is hard to conclude that 4 vs 7 (75% difference) is "close".

**Response:** We are very thankful for the reviewer's valuable suggestion. The error bars have been added to Figure 2b, and the results have been discussed statistically in the revised manuscript. We agree that the CO-scaled concentration of cholesterol was not that close for the whole sampling period, and we have made the corresponding revision as: "The CO-scaled concentration of cholesterol during the study periods in spring and autumn were significantly higher ($p<0.05$) than that in winter and summer."

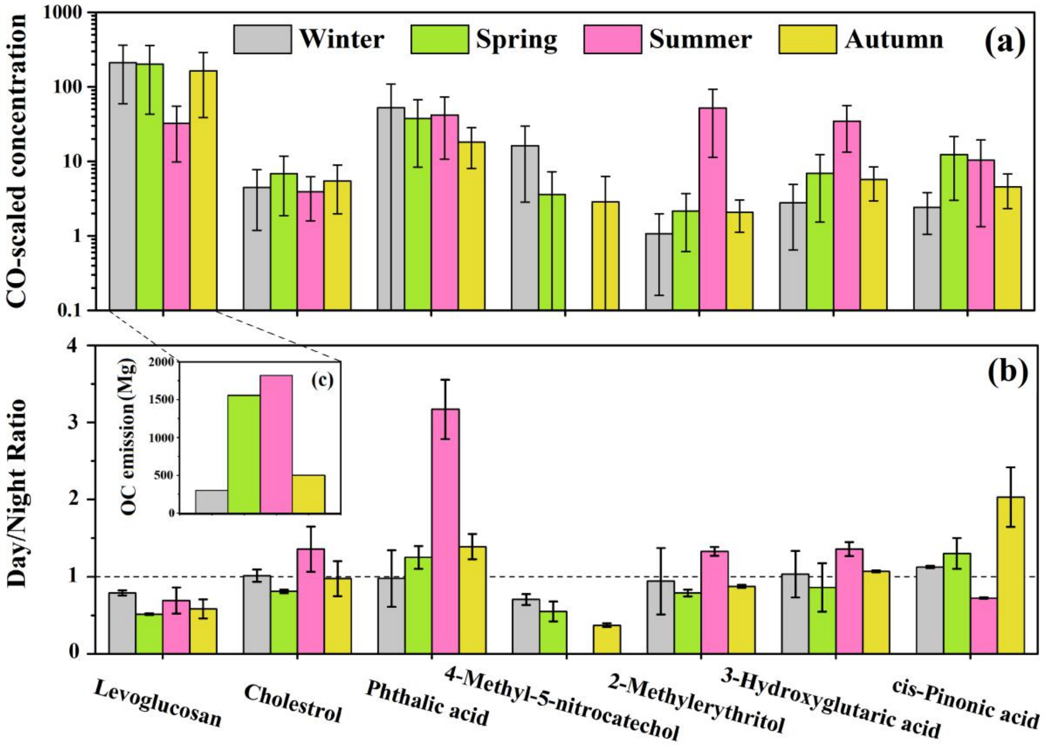


**Figure 2.** (a) The CO-scaled concentration of the identified organic tracers; (b) The day to night ratio of the measured concentrations of the organic tracers; (c) The OC emission amount from open biomass burning provided by the Fire Inventory (FINN) in Beijing during the sampling periods in four seasons of 2017.

**Comment 11:** Besides the statistical issue, other problems exist for the conclusions made on the basis of day/night ratios. First, biomass burning is not the only source of CO. The <1 day/night CO-scaled ratios of levoglucosan can be simply caused by enhanced CO emissions at night from other sources when the biomass burning contributions were constant. Similarly for cooking. Second, from emissions to concentrations, many factors are involved. Biomass burning is not a local source in Beijing. Similar strength of emissions may lead to different concentrations in Beijing because of the atmospheric aging and dilution during the regional transport process. Also, scaling secondary tracers by CO has complicated meaning, especially for biogenic-related tracers. I can't understand the logic behind Line 236-243. Not to mention that the phthalic acid concentrations may be affected by the OA concentrations due to gas-particle partitioning and the photolysis of

4-methyl-5-nitrocatechol might be significant to affect its daytime concentrations.

**Response:** We deeply appreciate the reviewer's valuable comment. We agree with the reviewer that there were many factors involved from emissions to atmospheric concentrations, thus it is hard to draw certain conclusions. Therefore, we focused on the result demonstration and explanation, instead of reaching certain conclusions in the revised manuscript.

As we know, the observed concentrations of organic tracers in different seasons were not only controlled by their emission strengths or secondary formation rates, but also affected by the atmospheric dilution conditions or boundary layer height. As a tracer for traffic exhaust, carbon monoxide (CO) shows relatively stable emission rates in different seasons, and is quite inertial to chemical reactions. Thus we used the CO-scaled concentration of the organic tracers (the concentration of organic tracers divided by CO concentration) to correct for different diffusion conditions in four seasons, so as to better discuss their differences in emission strengths or secondary production rates (Kleinman et al., 2008; Aiken et al., 2009).

As mentioned by the reviewer, the emission rates of CO might be different between day and night, thus it was not meaningful to compare the day/night ratios of the CO-scaled concentrations of the organic tracers. Therefore, in the revised manuscript, we demonstrated the day/night ratios of the observed concentrations of the organic tracers (not scaled by CO) instead. The planetary boundary layer (PBL) heights (derived from the WRF model) and wind speeds were shown in Table S6. The diffusion conditions were typically worse at night. Therefore, if the day/night ratio was much higher than 1, it could be attributed to stronger source emission or secondary formation during the daytime. However, if the day/night ratio was lower than 1, it cannot be concluded that the source emission strength or secondary formation rate was stronger at night. Besides, the impact of gas-particle partitioning on phthalic acid and the photolysis of 4-methyl-5-nitrocatechol were also discussed in the revised manuscript.

**Table S6** The mean values and standard deviations (SD) of planetary boundary layer

(PBL) heights and wind speeds in Beijing during the study periods in four seasons.

| | | Winter | | Spring | | Summer | | Autumn | |
|---|---|---|---|---|---|---|---|---|---|
| | | **Day** | **Night** | **Day** | **Night** | **Day** | **Night** | **Day** | **Night** |
| **PBL (m)** | **Mean** | 309.7 | 283.2 | 1149.4 | 303.9 | 871.6 | 210.5 | 423.2 | 90.9 |
| | **SD** | 236.7 | 259.1 | 841.1 | 363.2 | 405.6 | 164.3 | 278.9 | 90.4 |
| **WS (m $s^{-1}$)** | **Mean** | 1.14 | 0.83 | 1.75 | 0.83 | 1.03 | 0.63 | 0.72 | 0.39 |
| | **SD** | 0.79 | 0.61 | 0.71 | 0.75 | 0.68 | 0.44 | 0.62 | 0.25 |

**References:**

Aiken, A. C., Salcedo, D., Cubison, M. J., Huffman, J. A., DeCarlo, P. F., Ulbrich, I. M., Docherty, K. S., Sueper, D., Kimmel, J. R., Worsnop, D. R., Trimborn, A., Northway, M., Stone, E. A., Schauer, J. J., Volkamer, R. M., Fortner, E., Foy, B., Wang, J., Laskin, A., Shutthanandan, V., Zheng, J., Zhang, R., Gaffney, J., Marley, N. A., Paredes-Miranda, G., Arnott, W. P., Molina, L. T., Sosa, G., and Jimenez, J. L.: Mexico City aerosol analysis during MILAGRO using high resolution aerosol mass spectrometry at the urban supersite (T0) - Part 1: Fine particle composition and organic source apportionment, Atmos. Chem. Phys. 9, 6633-6653, 2009.

Kleinman, L. I., Springston, S. R., Daum, P. H., Lee, Y. N., Nunnermacker, L. J., Senum, G. I., Wang, J., Weinstein-Lloyd, J., Alexander, M. L., Hubbe, J., Ortega, J., Canagaratna, M. R., and Jayne, J.: The time evolution of aerosol composition over the Mexico City plateau, Atmos. Chem. Phys. 8, 1559-1575, 2008.

**Comment 12:** Page 7, Line 205-207: The cited reference is only for PAH. How about other aromatic precursors (e.g., single-ring aromatics)? Do those typical SOA precursors show higher concentrations in winter? Please note that oxidation conditions are also important when discussing about the secondary formation potential of SOA. In winter, the oxidant concentrations (e.g., OH) might be lower.

**Response:** We deeply appreciate the reviewer's valuable comments. Based on the VOC measurements in Beijing, some studies reported that single-ring aromatics (such as benzene, toluene, xylenes, ethylbenzene, etc) also exhibited higher concentrations in winter (Li et al., 2020a, b), while others reported a different result (Liu et al., 2020). Thus the related expression has been corrected to: "In winter, the enhanced emissions due to residential heating and adverse atmospheric diffusion conditions led to higher concentrations of PAHs (Feng et al., 2018; Sun et al., 2018), thus might facilitate the PAH-derived SOA formation." In fact, a significant increase of toluene-derived SOA (indicated by DHOPA) in winter has also been observed in northern cities of China due to residential heating (Ding et al., 2017).

We agree with the reviewer that the concentrations of oxidants might be lower in winter due to weak solar radiation. Discussion on the oxidation conditions in winter has also been added in the revised manuscript. A previous observation found that the ·OH concentration in Beijing was significantly higher than that in New York, Birmingham and Tokyo, and was nearly 1 order of magnitude higher than that predicted by global models in northern China in winter (Tan et al., 2018). Zhang et al. (2019) indicated that HONO, which was mainly from the heterogeneous reactions of $NO_2$ and traffic emissions, was the major precursor of ·OH in winter. According to the WRF-Chem model simulation, HONO resulted in a significant enhancement (5-25 μg $m^{-3}$) of SOA formation (most of which were from the aromatic precursors) during a haze episode in winter in the Beijing-Tianjin-Hebei region (Zhang et al., 2019). Besides, some recent studies suggested that the brown carbon-derived singlet molecular oxygen ($^1O_2^*$) in aerosol liquid water could react rapidly with the electron-rich organics such as PAHs, thus facilitate the aromatic SOA formation (Kaur et al., 2019; Manfrin et al., 2019). This process might be more significant in winter, when the concentration of HULIS was much higher than that in other seasons.

**Response:**

Ding, X., Zhang, Y., He, Q., Yu, Q., Wang, J., Shen, R., Song, W., Wang, Y., and Wang, X.: Significant increase of aromatics-derived secondary organic aerosol during fall to winter in China, Environ. Sci. Technol., 51, 7432-7441, doi: 10.1021/acs.est.6b06408, 2017.

Li, Q., Su, G., Li, C., Liu, P., Zhao, X., Zhang, C., Sun, X., Mu, Y., Wu, M., Wang, Q., and Sun, B.: An investigation into the role of VOCs in SOA and ozone production in Beijing, China, Sci. Total Environ., 720, 137536, doi: 10.1016/j.scitotenv.2020.137536, 2020a.

Li, C., Li, Q., Tong, D., Wang, Q., Wu, M., Sun, B., Su, G., and Tan, L.: Environmental impact and health risk assessment of volatile organic compound emissions during different seasons in Beijing, J. Environ. Sci., 93, 1-12, doi: 10.1016/j.jes.2019.11.006, 2020b.

Liu, Y., Song, M., Liu, X., Zhang, Y., Hui, L., Kong, L., Zhang, Y., Zhang, C., Qu, Y., An, J., Ma, D., Tan, Q., and Feng, M.: Characterization and sources of volatile organic compounds (VOCs) and their related changes during ozone pollution days in 2016 in Beijing, China, Environ. Pollut., 257, 113599, doi:

10.1016/j.envpol.2019.113599, 2020.

Tan, Z., Rohrer, F., Lu, K., Ma, X., Bohn, B., Broch, S., Dong, H., Fuchs, H., Gkatzelis, G. I., Hofzumahaus, A., Holland, F., Li, X., Liu, Y., Liu, Y., Novelli, A., Shao, M., Wang, H., Wu, Y., Zeng, L., Hu, M., Kiendler-Scharr, A., Wahner, A., and Zhang, Y.: Wintertime photochemistry in Beijing: observations of $RO_x$ radical concentrations in the North China Plain during the BEST-ONE campaign, Atmos. Chem. Phys., 18, 12391-12411, doi:10.5194/acp-18-12391-2018, 2018.

Zhang, J., Chen, J., Xue, C., Chen, H., Zhang, Q., Liu, X., Mu, Y., Guo, Y., Wang, D., Chen, Y., Li, J., Qu, Y., and An, J.: Impacts of six potential HONO sources on HOx budgets and SOA formation during a wintertime heavy haze period in the North China Plain, Sci. Total Environ., 681, 110-123, doi: 10.1016/j.scitotenv.2019.05.100, 2019.

**Comment 13:** Section 3.2.2 in Page 8-10: Conclusions in this section are generally arbitrary. Although correlations sometimes help diagnostics, connections between $O_3$, ALWC, RH, T and the tracer species as well as WSOC/OC are not simple/obvious. For example, in Line 248-251, it was said that "WSOC/OC did not correlate with $O_3$, suggesting that gas-phase photooxidation was not the dominant formation mechanism of SOC". Why? Do the authors assume that WSOC are SOC and gas-phase photooxidation is equivalent to $O_3$? What about terms other than chemical production in the mass balance (e.g., photolysis, primary contributions, and so on? Besides, the correlation isn't strong (r=0.5) when the authors sometimes said "significantly correlated". Such kind of correlations might be used as non-conflict evidence for explaining the formation pathways but definitely insufficient to make any conclusions. A common argument is that in winter Beijing all components of $PM_{2.5}$ often correlate with RH and ALWC, even for primary OA. It is not surprise that 4-methyl-5-nitrocatechol correlate with RH and ALWC. The study done by Wang L et al. (2018) suggest coal and traffic contributions to 4-methyl-5-nitrocatechol were the dominant sources in northern China.

**Response:** We deeply appreciate the reviewer's valuable comments. We agree with the reviewer that it was inappropriate to conclude that "gas-phase photooxidation was not the dominant formation mechanism of SOC" merely based on the result that "WSOC/OC did not show any significant positive correlation with $O_3$ concentration". Therefore, we have deleted the related discussion in the revised manuscript. We also agree with the reviewer that correlation analysis was insufficient to make certain conclusions. Therefore, in the revised manuscript, we focused on the result explanation instead of making certain conclusions in most cases. Besides, the potential impact of photolysis and gas-particle partitioning was also discussed in the revised manuscript.

Indeed, in the wintertime of Beijing, $PM_{2.5}$ often correlates well with RH and ALWC, which is also the case even for primary OA (Sun et al., 2013). In some previous studies, the concentration of SOA was normalized by HOA (Sun et al., 2013), CO (Kleinman et al., 2008; Aiken et al., 2009), EC (Zheng et al., 2015) or OA (Xu et al., 2017), to eliminate the diffusion effect and better evaluate their secondary production as far as possible. In this study, not only 4-methyl-5-nitrocatechol correlated strongly with RH ($r=0.69$, $p<0.01$) and ALWC ($r=0.88$, $p<0.01$), but also the ratios of 4-methyl-5-nitrocatechol/OC (RH: $r=0.73$, $p<0.01$; ALWC: $r=0.82$, $p<0.01$), 4-methyl-5-nitrocatechol/EC (RH: $r=0.57$, $p<0.01$; ALWC: $r=0.73$, $p<0.01$), 4-methyl-5-nitrocatechol/CO (RH: $r=0.64$, $p<0.01$; ALWC: $r=0.81$, $p<0.01$), all correlated significantly with RH and ALWC during the sampling period in winter. Therefore, we speculated that the aqueous-phase oxidation probably facilitated the formation of 4-methyl-5-nitrocatechol during the study period in winter, and these correlations were added in the revised manuscript.

The work done by Wang et al. (2018) did not directly measure nitrophenols of the source samples, but reached their conclusion based on the PMF result (that is, the correlations between 4-methyl-5-nitrocatechol and $SO_2$ or NO). Besides, levoglucosan was not measured in their study, thus we could not compare its correlations with $SO_2$ or NO and levoglucosan. The result of Wang et al. (2018) was different from the other studies which commonly suggested that 4-methyl-5-nitrocatechol mainly derived from secondary formation (Iinuma et al., 2010; Bertrand et al., 2018; Wang et al., 2019). In this study, correlation between 4-methyl-5-nitrocatechol and levoglucosan ($r=0.87$) was higher than those with $SO_2$ ($r=0.59$) or CO ($r=0.44$). Therefore, coal combustion and traffic emissions should not be the dominant sources of 4-methyl-5-nitrocatechol in Beijing during the sampling period in 2017 in this study.

## References

Sun, Y., Wang, Z., Fu, P., 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, doi:10.1016/j.atmosenv.2013.06.019, 2013.

Wang, L., Wang, X., Gu, R., Wang, H., Yao, L., Wen, L., Zhu, F., Wang, W., Xue, L., Yang, L., Lu, K., Chen, J., Wang, T., Zhang, Y., and Wang, W.: Observations of fine particulate nitrated phenols in four sites in northern China: concentrations, source apportionment, and secondary formation, Atmos. Chem. Phys, 18, 4349-4359, doi:10.5194/acp-18-4349-2018, 2018.

Zheng, G., Duan, F., Su, H., Ma, Y., Cheng, Y., Zheng, B., Zhang, Q., Huang, T., Kimoto, T., Chang, D., Poschl, U., Cheng, Y., and He, K.: Exploring the severe winter haze in Beijing: the impact of synoptic weather, regional transport and heterogeneous reactions, Atmos. Chem. Phys, 15, 2969-2983, doi:10.5194/acp-15-2969-2015, 2015.

**Comment 15:** Technical remarks:

Page 7, Line 211: “secondary formation of aromatic SOA” - SOA is secondary.

Page 8, Line 239: “the diurnal patterns were close to 1” – What does this mean?

Page 9, Line 252: LWC has already defined.

**Response:** We are deeply thankful for the reviewer’s comments. (1) The expression “secondary formation of aromatic SOA” has been corrected to “formation of aromatic SOA”. (2) The expression “the diurnal patterns were close to 1” has been corrected to “the day to night ratios were close to 1”. (3) The “LWC” in the bracket has been deleted.

# Characteristics, primary sources and secondary formation of water soluble organic aerosols in downtown Beijing

Qing Yu[1,2], Jing Chen[1,2], Weihua Qin[1,2], Siming Cheng[1,2], Yuepeng Zhang[1,2], Yuewei Sun[1,2], Ke Xin[1,2], Mushtaq Ahmad[1,2]

[1]State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Beijing Normal University, Beijing 100875, China.
[2]Center of Atmospheric Environmental Studies, Beijing Normal University, Beijing 100875, China.

*Correspondence to*: Jing Chen (jingchen@bnu.edu.cn)

**Abstract.** Water soluble organic compounds (WSOC) account for a large proportion of aerosols and play a critical role in various atmospheric chemical processes. In order to investigate the primary sources and secondary production of WSOC in downtown Beijing, the day and night $PM_{2.5}$ samples in January (winter), April (spring), July (summer) and October (autumn) of 2017 were collected and analyzed for WSOC and organic tracers in this study. WSOC was dominated by its moderately hydrophilic fraction, and showed the highest concentration in January and comparable levels in April, July and October of 2017. Some typical organic tracers were chosen to evaluate the emission strength and secondary formation of WSOC. Seasonal variation of the organic tracers suggested significantly enhanced formation of anthropogenic secondary organic aerosols (SOA) during the sampling period in winter, and obviously elevated biogenic SOA formation during the sampling period in summer. These organic tracers were applied into positive matrix factorization (PMF) model to calculate the source contributions of WSOC as well as its moderately and strongly hydrophilic portions. The secondary sources contributed more than 50 % to WSOC, with higher contributions during the sampling periods in summer (75.1 %) and winter (67.4 %), and the largest contributor was aromatic SOC. Besides, source apportionment results under different pollution levels suggested that controlling biomass burning and aromatic precursors would be effective to reduce WSOC during the haze episodes in cold seasons. The impact factors for the formation of different SOA tracers, total secondary organic carbon (SOC) as well as moderately and strongly hydrophilic SOC were also investigated. The acid-catalyzed heterogeneous or aqueous-phase oxidation appeared to dominate in the SOC formation during the sampling period in winter, while the photochemical oxidation played a more critical role during the sampling period in summer. Moreover, photooxidation played a more critical role in the formation of moderately hydrophilic SOC, while the heterogeneous or aqueous-phase reactions posed more vital effects on the formation of strongly hydrophilic SOC.

## 1 Introduction

Organic compounds account for a considerable fraction (20-60 %) of atmospheric aerosols (Huang et al., 2014; Zhang et al., 2020), and water-soluble organic carbon (WSOC) generally composes 30-70 % of organic carbon (OC) (Zhang et al., 2018; Yang et al., 2019; Chen et al., 2020). WSOC in aerosols is active in light adsorption (Yan et al., 2015; Geng et al., 2020), thus may make significant impact on radiative forcing and global climate change (Andreae and Gelencser, 2006). Meanwhile, the photoexcitation of water-soluble brown carbon (BrC) can generate oxidants in aerosols and cloud/fog droplets (Manfrin et al., 2019; Kaur et al., 2019), which can promote atmospheric chemical reactions and aging processes of organic aerosols. Overall, WSOC is widely involved in the cloud processes and heterogeneous reactions due to its surface activity and water solubility (Ervens et al., 2011; George et al., 2015), thus plays a significant role in severe haze episodes (Cheng et al., 2015; Wu et al., 2019; Ma et al., 2020). Besides, WSOC is closely linked to the oxidative potential of aerosols, posing adverse health outcomes (Verma et al., 2012; Chen et al., 2019; Wang et al., 2020). Therefore, it is of great significance to study the characteristics, primary sources and secondary production of WSOC in atmospheric particulate matter.

Based on the solid phase extraction (SPE) by the Oasis HLB column, WSOC can be divided into the moderately hydrophilic fraction and strongly hydrophilic fraction (Varga et al., 2001; Kiss et al., 2002). The moderately hydrophilic fraction of WSOC mainly consists of humic-like substances (HULIS), which are an unresolved mixture of polycyclic ring structures with substituted hydroxyl, carboxyl, carbonyl, methoxy, and ester groups (Kiss et al., 2002; Lin et al., 2010; Fan et al., 2012). In addition, some smaller molecules with well-defined structures, such as phthalic acid and suberic acid, may also constitute a minor portion of the moderately hydrophilic fraction of WSOC (Lin et al., 2010). The strongly hydrophilic fraction mainly consists of low-molecular-weight organic acids (such as oxalic acid, succinic acid, malic acid) and anhydrosugars (such as levoglucosan, xylose, sucrose) (Lin et al., 2010). Previous studies have revealed that the moderately and strongly hydrophilic fractions of WSOC show significantly different intrinsic oxidative potentials, thus would pose different effects on human health (Verma et al., 2012, 2015; Yu et al., 2018). However, source contributions of the moderately and strongly hydrophilic WSOC were scarcely investigated and compared in previous research.

So far, the most widely used source apportionment approaches of organic aerosols include the chemical mass balance (CMB) model coupled with tracer-yield method (Guo et al., 2012; Islam et al., 2020), and the aerosol mass spectrometry combining positive matrix factorization (AMS-PMF) method (Hu et al., 2017; Sun et al., 2018; Shen et al., 2019). CMB model is used to quantify the primary sources of organic aerosols, while the tracer-yield method is applied to calculate the contributions of secondary sources. However, CMB model requires local source profiles, and the tracer-yield experiments conducted under simple chamber conditions usually ignore cloud and aqueous-phase processes, leading to large uncertainties when applying these results to the real atmosphere (Kleindienst et al., 2007; Feng et al., 2013). The AMS-PMF method is based on the mass spectra of organic aerosols and massive data analysis, which can avoid such disadvantages. However, online AMS can not differentiate WSOC from water-insoluble OC, not to mention the moderately hydrophilic and strongly hydrophilic WSOC. Besides, AMS-PMF method typically classifies secondary organic aerosols (SOA) into two categories, the less oxidized oxygenated organic aerosols (LO-OOA) and the more oxidized oxygenated organic aerosols (MO-OOA), which is unable to distinguish SOA from different precursors or sources in most cases. To raise effective control measures targeting on the specific SOA precursors, some recent studies introduced the SOA tracers into PMF model to investigate the secondary sources of organic aerosols (Kang et al., 2018a, b; Geng et al., 2020).

Previous studies have suggested that coal combustion (Zhang et al., 2018; Li et al., 2019a, c), traffic emissions (Kawamura and Kaplan, 1987; Li et al., 2019c), residual oil combustion (Kuang et al., 2015), cooking (Qiu et al., 2020), soil dust and sea salts (Huang et al., 2006) can all contribute to WSOC. However, it is most commonly recognized that WSOC mainly derives from biomass burning and SOA (Ding et al., 2008; Feng et al., 2013; Du et al., 2014; Zhang et al., 2018). For example, Du et al. (2014) suggested that SOA, biomass burning and other primary combustion sources contributed about 54 %, 40 % and 6 % respectively to WSOC in Beijing during 2010-2011. Zhang et al. (2018) also indicated that the sum of biomass burning and SOA contributed more than 80 % to WSOC in Beijing, Shanghai, Guangzhou and Xi'an in the wintertime of 2013. In recent years, the adjustment of energy and industrial structures as well as the effective control of the open burning activities in the surrounding areas of Beijing have posed significant impact on the source emissions. The average $PM_{2.5}$ concentration in Beijing has been greatly reduced from 89.5 $\mu g \cdot m^{-3}$ in 2013 to 58.0 $\mu g \cdot m^{-3}$ in 2017 since the implement of the Action Plan of Air Pollution Prevention and Control in 2013 (Cheng et al., 2019). Meanwhile, it has been reported that the oxidant concentrations were enhanced accompanying the decrease of $PM_{2.5}$ level, which might promote the SOA formation (Feng et al., 2019). Consequently, the sources and composition of WSOC in Beijing may show significant changes due to the control policies and enhanced atmospheric oxidizing capacity in the surrounding areas in recent years. Therefore, it is necessary to compare the source contributions of WSOC with those in the previous studies.

A large fraction of WSOC is formed in the atmosphere, and WSOC greatly overlaps with SOA (Zhang et al., 2018). Because most of the organic aerosols remain unidentified at a molecular level, it is difficult to thoroughly understand the formation mechanisms of secondary WSOC or SOA. An effective approach to explore the formation mechanisms of SOA is to classify SOA into several categories based on their specific properties. For example, the formation mechanisms of secondary organic carbon (SOC) in the moderately and strongly hydrophilic fractions of WSOC may differ from each other due to their different water-solubility in the cloud droplets or the aqueous phase of aerosols. Furthermore, the formation mechanisms may also be different for SOA from different precursors, which originate from various sources and show disparate chemical structures and properties (Sun et al., 2016; Cheng et al., 2018). While gas-phase photooxidation of volatile organic compounds (VOCs) is an important formation pathway of SOA, direct observations have proved that the aqueous-phase reactions of VOCs from biomass burning contribute remarkably to SOA formation (Gilardoni et al., 2016). In this regard, SOA tracers can provide implications for, though may not fully represent, the formation mechanism of SOA from corresponding precursors. Besides, in recent years, haze episodes still occur frequently in Beijing in cold seasons. The humid meteorological conditions as well as the high concentrations of VOC precursors and HONO during haze episodes may pose unique effects on the formation of some secondary components in WSOC (Li et al., 2019b; Yang et al., 2019; Zhang et al., 2019). Previous research found that haze events were usually accompanied with elevated WSOC/OC ratio (Cheng et al., 2015; Yang et al., 2019) and enhanced SOA production (Huang et al., 2014; Li et al., 2019b). Comparing source contributions of WSOC under different pollution levels would help to better understand the complex properties of WSOC, and put forth effective control measures to reduce WSOC during haze events.

In this work, the 12-hour day and night $PM_{2.5}$ samples in Beijing in January, April, July and October of 2017 were collected and analyzed for WSOC and organic tracers. The characteristics of WSOC and the selected organic tracers were investigated, and the contributions of primary and secondary sources to WSOC and its moderately and strongly hydrophilic fractions were quantified. The key influencing factors for the formation of different SOA tracers, the moderately hydrophilic SOC and strongly hydrophilic SOC were explored and compared, so as to gain insights into the possible formation mechanism of different types of SOA.

## 2. Experimental

### 2.1 Sampling

Field sampling was performed on the roof of a twenty-meter-high building in the campus of Beijing Normal University in downtown Beijing, which is considered to be a representative urban site. Fine particulate matter ($PM_{2.5}$) were sampled in four seasons of 2017 using a high-volume sampler (TH-1000C, Wuhan Tianhong Instruments Co. Ltd, China) equipped with an $PM_{2.5}$ impactor at a flow rate of 1.05 $m^3$ $min^{-1}$. The potential sampling error was discussed and estimated in the supplementary materials (S1). The sampling periods included January 2nd-16th, April 7th-23rd, July 3rd-18th and October 12th -28th in winter, spring, summer and autumn of 2017, respectively. Sampling was conducted during 8:00-19:30 in the daytime and during 20:00-7:30 at night. A total of 124 effective $PM_{2.5}$ samples were obtained in this study. The field blank samples were collected before and after each sampling period, and a total of 8 field blank samples were obtained. The blank filters were put on the filter holder of $PM_{2.5}$ sampler without pumping for 1 min, then stored and analyzed together with the ambient samples. All the samples were gathered on the quartz filters (PALLFLEX) which were pre-baked at above 500 °C for at least 4 hours before sampling. Before and after sample collections, the quartz filters were weighed by an analytical balance for three times after stabilizing under conditions of fixed temperature (20 ± 1 °C) and humidity (40 ± 2 %) for 24 h. After that, the sampled filters were stored under dark conditions below -20 ℃ until being analyzed.

### 2.2 Chemical analysis

To measure the values of OC and EC, a part (0.296 cm$^2$) of each filter was detected using a DRI 2001A carbon analyzer with thermal/optical reflectance (TOR) protocol. The analysis of WSOC and water-soluble ions followed the same procedure as in our previous research (Chen et al., 2014), and the details can be found in the supplementary materials (S2-1).

WSOC was further divided into its moderately and strongly hydrophilic fractions by the solid phase extraction (SPE). The moderately hydrophilic fraction of WSOC (MH-WSOC) was directly measured by the following procedure. Briefly, a punch of the sampled filter was shredded into tiny pieces and extracted using 20 mL ultrapure water for a duration of 30 min, and then filtrated via a 0.45 μm PTFE filter. The extract was acidified to pH=2.0 with HCl (1mol L$^{-1}$), then passed through a SPE cartridge (Oasis HLB, 30 μm, Waters). After that, the SPE column was rinsed with 3 mL water, then eluted with 1.5 mL MeOH (containing 3 % $NH_3$) for three times. The eluent was blown to dryness and redissolved in 20 mL ultrapure water, and measured by a total organic carbon (TOC) analyzer (Shimadzu TOC-L CPN). The strongly hydrophilic fraction of WSOC (SH-WSOC) was calculated as the total WSOC minus the moderately hydrophilic WSOC.

Seven organic tracers, including levoglucosan, cholesterol, 4-methyl-5-nitrocatechol, phthalic acid (Ph), 2-methylerythritol, 3-hydroxyglutaric acid, and *cis*-pinonic acid (details shown in Section 3.2), were analyzed for each sample in this study. A punch of each filter was ultrasonically extracted twice in 10 mL MeOH for 20 min (below 20 °C). The combined extracts were filtrated, concentrated and stored in dark place at -20 °C until further derivation. Afterwards, the concentrates were blown to entire dryness with gentle ultrapure nitrogen ($N_2$), then redissolved in 100 μL pyridine and reacted with 200 μL silylating reagent (BSTFA with 1 % TMCS) at 75°C for 70 min. After cooling down to the room temperature, the derivative products were diluted with n-hexane and immediately analyzed using the GC/MS/MS equipped with a JA-5MS column. The detailed parameter setting of GC/MS/MS can be found in the supplementary materials (S2-2). The authentic standards (Table S2) were dissolved in anhydrous pyridine, and diluted to five to seven different concentrations. Then 100 μL of the standard solutions were reacted with 200 μL silylating reagent (BSTFA: TMCS = 99:1) at 75°C for a duration of 70 min. After cooling down to the ambient temperature, these solutions containing derivative products were diluted to 1 mL with n-hexane, and measured by GC/MS/MS right before the analysis of ambient samples. The $R^2$ of the derivative products were all above 0.99, indicating good linearities of these standard curves.

To ensure quality of the measurement, the recovery rates were determined together with ambient samples, by measuring the authentic standards (Table S2) spiked onto the pre-combustioned quartz filters. The recovery rates were all in the range of 70 %-110 %, and the relative standard deviations (RSD) were all below 15 %. Besides, since two field blank samples were obtained during each sampling period, the average concentrations of the targeted compounds on these two field blanks were used for the correction in the corresponding season. The concentrations of the targeted compounds on the field blanks were all close to zero. The reported concentrations of the targeted compounds were: the measured concentrations on each ambient sample minus the average concentration on the two field blank samples.

### 2.3 PMF source apportionment

The positive matrix factorization (PMF) model is an effective multivariate factor analysis tool which can decompose a matrix of speciated sample data (X) into two matrices: source contributions (G) and source profiles (F) (Norris et al., 2014). In this study, PMF 5.0 was applied for the source apportionment of WSOC in $PM_{2.5}$ during the sampling periods in four seasons of 2017 with a total of 124 samples. Seventeen species were input into PMF model, including WSOC, MH-WSOC, SH-WSOC, elemental carbon, sulfate, nitrate, oxalate, ammonium, magnesium, calcium and seven organic tracers. The PMF model was run repeatedly by changing the number of factors, and the derived solutions were compared. The base solution was selected based on: (1) the interpretability of the factor profiles and temporal variations of source contributions; (2) the reconstruction of the total variable and the $R^2$ of the input organic tracers ($R^2$>0.90); (3) the scaled residuals of the input species. Afterwards, the selected base solution was subjected to the displacement (DISP) and bootstrap (BS) tests for error estimation. To reduce the variability of the base solution, some constraints were defined based on the priori information on the sources (Norris et al., 2014; Paatero et al., 2014; Bozzetti et al., 2017). The detailed information about the uncertainty calculation for the input data, the selection criteria for the optimal solution, the diagnostics and the error estimates are provided in the supplementary materials (S3).

### 2.4 Other data collection and calculations

During the sampling periods, the meteorological parameters were simultaneously monitored using a HOBO meteorological station at our sampling site. The hourly concentrations of $O_3$ and CO were obtained from a nearby urban air monitoring station (3.4 km from the sampling site) via the website at http://www.bjmemc.com.cn. These data were transformed to 12-h averages corresponding to the sampling time.

The liquid water content (LWC) in inorganic aerosols was calculated by the ISORROPIA-II model (Fountoukis and Nenes, 2007), and the reverse mode was chosen in this study since the concentrations of gaseous pollutants such as HCl, $HNO_3$ and $NH_3$ were not available here. The total aerosol LWC was the summation of the water in both water-soluble ions and organic species, and the latter was calculated by the approach described by Cheng et al. (2016). To estimate the aerosol acidity, the approximate value of aerosol pH ($pH_F$) was also estimated by the ISORROPIA-II model (Pye et al., 2020). The molality of $H^+$ ($m_{H^+}$), which was calculated by: $m_{H^+} = 10^{-pH_F}$, was used for the correlation analysis.

The OC emission amounts from open biomass burning over the sampling periods in the four seasons of 2017 were obtained from the Fire Inventory (FINNv1.5), which provides daily estimates of the OC emissions from wildfire and agricultural fires with a resolution of 1 km (Wiedinmyer et al., 2011). These emission data were processed by the fire_emis utility provided on http://bai.acom.ucar.edu/Data/fire/. The values of OC emissions in the Beijing-Tianjin-Hebei region were extracted using the Geographic Information System (GIS).

## 3 Results and discussion

### 3.1 Temporal trends of carbonaceous species

The temporal variations of WSOC, OC, EC and $PM_{2.5}$ during the sampling periods in the four seasons of 2017 are shown in Figure 1, and their average concentrations are summarized in Table 1. In general, the carbonaceous species showed similar variation trends with that of $PM_{2.5}$ throughout the whole sampling period, which approximately varied in the periodic cycles of two to seven days (Guo et al., 2014). The average values of WSOC, OC, EC and $PM_{2.5}$ in January were all much higher than those during the sampling periods in other seasons. The mean levels of OC, EC and $PM_{2.5}$ were in the descending order of October > April > July, while WSOC showed comparable levels over the sampling periods in these three seasons. The fact that WSOC exhibited mild temporal variation with no sudden increase in October implied that short-term outdoor biomass burning after the harvest season in the surrounding areas of Beijing was well controlled over the sampling period in autumn. As shown in Table S5, the OC concentration in $PM_{2.5}$ in Beijing exhibited an overall declining trend in the past ten years, but the WSOC level showed no obvious change. Therefore, the control of WSOC seemed to be more challenging compared to the control of water-insoluble organic aerosols.

The temporal variation of the WSOC/OC ratio is also illustrated in Figure 1. WSOC/OC exhibited the highest average value during the sampling period in summer, followed by that in winter. The range of WSOC/OC in this study was similar to that previously reported in urban Beijing (Zhao et al., 2018; Yang et al., 2019). Previous studies showed that the aggravation of $PM_{2.5}$ pollution was usually accompanied by the elevated WSOC/OC ratio (Cheng et al., 2015; Yang et al., 2019), which was also confirmed by the significant positive correlations between $PM_{2.5}$ and the WSOC/OC ratio during the sampling periods in four seasons (winter: $r=0.68$, $p<0.01$; spring: $r=0.61$, $p<0.01$; summer: $r=0.36$, $p<0.05$; autumn: $r=0.91$, $p<0.01$) in this study, again underling the important role of WSOC during the haze evolution process. In this work, WSOC was further divided into its moderately and strongly hydrophilic fractions. As listed in Table 1, the moderately hydrophilic WSOC dominated in the total WSOC throughout the sampling period, and showed the highest proportion in July (0.84), followed by January (0.73), higher than the results previously reported in Beijing (Li et al., 2019a; Huang et al., 2020). Besides, the ratio of the strongly hydrophilic WSOC to the total WSOC showed significant positive correlations with $PM_{2.5}$ in January ($r=0.58$, $p<0.01$), July ($r=0.48$, $p<0.05$) and October ($r=0.44$, $p<0.05$).

### 3.2 Characteristics of organic tracers

#### 3.2.1 Seasonal variations and diurnal patterns

The average concentrations of the organic tracers identified in this study are summarized in Table 1. Levoglucosan is a tracer for biomass burning, while cholesterol is a good indicator of cooking, and both of them are known as primary organic tracers. Previous studies have shown that phthalic acid can be used as an aromatic SOA tracer (Al-Naiema and Stone, 2017; Huang et al., 2019), while 4-methyl-5-nitrocatechol can be used as a tracer for biomass burning SOA (Iinuma et al., 2010; Bertrand et al., 2018; Srivastava et al., 2018), and both of them serve as the anthropogenic SOA tracers. The biogenic SOA tracers in this study included the isoprene SOA tracer 2-methylerythritol, as well as *cis*-pinonic acid and 3-hydroxyglutaric acid, which are lower- and higher-generation oxidative products of monoterpenes respectively (Kourtchev et al., 2009). To correct for the different atmospheric diffusion conditions and better discuss the differences in emission strengths or secondary production rates in four seasons, the concentrations of the identified organic tracers were divided by the level of carbon monoxide (CO), which has relatively constant emission rate and is inert to chemical reactions. The CO-scaled concentration of these organic tracers during the sampling periods in four seasons are presented in Figure 2 (a), and the day to night ratios of the measured concentrations of these organic tracers (not normalized by CO) are shown in Figure 2 (b).

**Primary organic tracers.** The average CO-scaled concentration of levoglucosan in July was much lower than those during the study periods in other three seasons (only 15.4 % of that in January), suggesting that biomass burning might be not active during the study period in summer, similar to the findings in previous studies (Sun et al., 2018; Duan et al., 2020). The result of the variance analysis (ANOVA) indicated that levoglucosan exhibited similar CO-scaled concentrations in January, April and October ($p>0.05$). As shown in Figure 2 (c), the OC emissions from open biomass burning in the Beijing-Tianjin-Hebei region (provided by the FINN inventory) were 302 Mg (January), 1557 Mg (April), 1818 Mg (July) and 501 Mg (October), respectively, completely different from the seasonal variation trend of levoglucosan, suggesting that open biomass burning was not the major type of biomass burning in Beijing. Previous studies indicated that residential biofuel combustion might be the main source of levoglucosan in Beijing (Chen et al., 2017), and residential coal burning might also contribute a minor fraction of levoglucosan in winter (Yan et al., 2018). The CO-scaled concentration of cholesterol during the study periods in spring and autumn were significantly higher (ANOVA, $p<0.05$) than that in winter and summer. Some previous research found that both the regular cooking emissions and open barbecues might influence the organic aerosols in Beijing (Xu et al., 2015). The enhanced CO-scaled concentrations of cholesterol in spring and autumn were perhaps due to the enhanced open barbecues in spring and autumn when the weather was mild.

**Anthropogenic SOA tracers.** Phthalic acid showed the highest CO-scaled concentration in January, followed by July. The highest phthalic acid over the study period in winter might be due to stronger aromatic SOA formation. In the previous study, a significant increase of toluene-derived SOA in winter was universally observed in the northern cities of China (Ding et al., 2017). In winter, the enhanced emissions due to residential heating and the adverse atmospheric diffusion condition led to higher concentrations of PAHs (Feng et al., 2018; Sun et al., 2018), thus might facilitate the PAH-derived SOA formation. Although the concentrations of oxidants were usually lower in winter due to the weak solar radiation, a previous observation found that the ·OH concentration in Beijing was significantly higher than that in New York, Birmingham and Tokyo, and was nearly 1 order of magnitude larger than that predicted by global models in northern China in winter (Tan et al., 2018). Zhang et al. (2019) indicated that HONO, which was mainly from the heterogeneous reactions of $NO_2$ and traffic emissions, was the major precursor of OH radicals in winter. According to the WRF-Chem simulation, HONO resulted in a significant enhancement (5-25 $\mu g\ m^{-3}$) of SOA formation (most of which were from the aromatic precursors) during a haze episode in winter in the Beijing-Tianjin-Hebei region (Zhang et al., 2019). Besides, some recent studies suggested that brown carbon-derived singlet molecular oxygen ($^1O_2^*$) in aerosol liquid water could react rapidly with the electron-rich organics such as PAHs, thus facilitate the formation of aromatic SOA (Kaur et al., 2019; Manfrin et al., 2019). This process might be more significant during the sampling period in winter, when the HULIS concentration was much higher than that in other seasons. Furthermore, the high aerosol LWC during the sampling period in winter can also promote the heterogeneous or aqueous-phase formation of aromatic SOA. Except for stronger aromatic SOA formation, the enhanced gas-to-particle partition of phthalic acid due to low temperature, high relative humidity and high OA loadings in winter might also be one of the reasons for the highest phthalic acid in January (Huang et al., 2019). The second highest value of phthalic acid occurring in July was probably due to high temperature and relative humidity, strong solar radiation and abundant oxidants in summer, which was favorable for the secondary photochemical reactions (Zhao et al., 2018). As presented in Figure 2 (b), the day to night ratio of phthalic acid was obviously higher than 1 in April, July and October. Since the atmospheric diffusion condition was worse at night (Table S6), higher concentrations of phthalic acid in the daytime could be attributed to stronger secondary formation (Kawamura and Yasui, 2005). The day/night ratio of phthalic acid in July was much higher than those in other study periods, which was possibly due to more prominent effect of photochemical processes on the generation of aromatic SOA in summer (Kawamura and Yasui, 2005).

4-Methyl-5-nitrocatechol was usually below detection limit for the summer samples, again indicating that biomass burning was not active in July. It is notable that the CO-scaled concentrations of 4-methyl-5-nitrocatechol in January were much higher than those in April and October, similar to the result of Kahnt et al. (2013), while the primary emission intensity of biomass burning was relatively constant during the sampling periods in these seasons as revealed by the seasonal distribution pattern of levoglucosan. This phenomenon can be explained by the following reasons. First of all, the adverse atmospheric diffusion condition in winter favored the accumulation of its precursors such as phenols and $NO_x$ (Iinuma et al., 2010), thus might increase the formation rate of biomass burning SOA. Secondly, the low temperature and high aerosol LWC in winter were beneficial for the partitioning of the gaseous phenols to the aerosol phase. Besides, the high concentration of HONO in winter might also contribute to the formation of nitrocatechols in the condensed phase (Vidovic et al., 2018; Qu et al., 2019). Hence, the aqueous-phase or heterogeneous reactions were enhanced in winter (Li et al., 2014; Gilardon et al., 2016).

**Biogenic SOA tracers.** As presented in Figure 2, 2-methylerythritol and 3-hydroxyglutaric acid showed extremely high CO-scaled concentrations in July, followed by April and October, and showed the minimum concentrations in January. The mean concentration of 2-methylerythritol during the study period in summer was almost 50 times of that in winter. Such variations of biogenic SOA tracers indicated much stronger biogenic SOA formation during the sampling period in summer, which was attributable to the higher emissions of biogenic precursors and accelerated photochemical reactions (Shen et al., 2015; Qiu et al., 2020). The day/night ratios of 2-methylerythritol and 3-hydroxyglutaric acid were higher than 1 in July, and significantly higher than those in other seasons, indicating that photochemical oxidation played a more significant role in their formation processes during the sampling period in summer. Nevertheless, *cis*-pinonic acid, another monoterpene SOA tracer, showed slightly higher CO-scaled concentration in April than in July (ANOVA, $p<0.05$). This might be due to the active atmospheric oxidation processes in summer facilitated the transformation of *cis*-pinonic acid to its higher-generation oxidation products. Besides, *cis*-pinonic acid tends to evaporate into the gas phase in summer under high temperature (Li et al., 2013; Ding et al., 2016). Different from the other SOA tracers which showed the highest day/night ratios in July, *cis*-pinonic acid exhibited the lowest day/night ratio during the study period in summer. This might be also because *cis*-pinonic acid was rapidly transformed to its higher-generation oxidation products when the photochemical oxidation was the strongest in the daytime in summer.

### 3.2.2 Influencing factors for the formation of different SOA tracers

Spearman correlation analysis was conducted between the SOA tracers and meteorological parameters, $O_3$, aerosol acidity and aerosol LWC, and the results are shown in Table S7 and Figure 3. In fact, sometimes all the $PM_{2.5}$ components, even the primary components, can correlate well with RH and LWC (Sun et al., 2013). To solve this problem, the concentration of SOA was normalized by HOA (Sun et al., 2013), CO (Kleinman et al., 2008; Aiken et al., 2009), EC (Zheng et al., 2015) or OA (Xu et al., 2017) in previous studies, to better evaluate their secondary generation. The correlation coefficients between the ratio of SOA tracers to OC and RH as well as LWC are also listed in the supplementary materials (Table S8).

The ratio of WSOC/OC showed significant positive correlations with RH, LWC and the aerosol acidity in January, April and October, possibly because the acid-catalyzed heterogeneous or aqueous-phase reactions facilitated the formation of SOC in WSOC during the sampling periods in these seasons (Du et al., 2014; Yang et al., 2019). Besides, the increased aerosol LWC could facilitate the the partitioning of the gas-phase WSOC to the aerosol phase (Hennigan et al., 2009), resulting in a higher WSOC/OC ratio. No correlations between WSOC/OC and RH or LWC were found in July, likely due to enhanced gas-phase photochemical oxidation in the formation SOC in WSOC during the study period in summer. In January and October with lower temperature, WSOC/OC exhibited significant positive correlations with temperature (January: $r=0.50$, $p<0.01$; October: $r=0.47$, $p<0.01$), while in April and July when the weather became warmer, no significant correlations were found between them. Higher temperature in warm seasons might inhibit the gas-to-particle partition of semi-volatile or intermediate WSOC and related heterogeneous or aqueous-phase reactions (Qian et al., 2019; Lu et al., 2019), thus counteracting its promoting effect on the formation of SOC in WSOC.

The biomass burning SOA tracer, 4-methyl-5-nitrocatechol, showed strong positive relationships with RH, LWC and aerosol acidity, and strong negative relationships with the $O_3$ concentration and solar radiation in January and April. The correlations with the above parameters became weaker in October and no data was available in July because 4-methyl-5-nitrocatechol was below the detection limit. The ratio of 4-methyl-5-nitrocatechol to OC also showed strong correlations with RH (January: $r=0.73$, $p<0.01$; April: $r=0.86$, $p<0.01$) and LWC (January: $r=0.82$, $p<0.01$; April: $r=0.75$, $p<0.01$) during the study periods in winter and spring. Laboratory studies have revealed that phenolic compounds, which are massively emitted from biomass burning, could undergo rapid aqueous-phase oxidation and produce substantial amounts of SOA under either the simulated sunlight (Sun et al., 2010; Li et al., 2014; Yu et al., 2014) or the dark conditions (Hartikainen et al., 2018; Kristijan et al., 2018). Direct observations have also proved that aqueous-phase reactions of the precursors from biomass burning emissions contribute significantly to SOA formation (Gilardoni et al., 2016). Therefore, the correlation results above together with the literature findings suggested that the acid-catalyzed heterogeneous or aqueous-phase reactions might play a dominant role in the formation of 4-methyl-5-nitrocatechol over the sampling periods in winter and spring. Similar speculation was also made in the previous observation (Wang et al., 2019). $O_3$ can be used as a tracer to reflect the strength of photochemical oxidation (Herndon et al., 2008; Xu et al., 2017), and the significant negative correlations between 4-methyl-5-nitrocatechol and $O_3$ in all seasons implied that the photochemical reactions might not be the major formation pathway of 4-methyl-5-nitrocatechol. Instead, stronger solar radiation might result in the photolysis of 4-methyl-5-nitrocatechol (Wang et al., 2019). Besides, the reverse relationship with temperature during the whole sampling period and in April and October was probably because that the increasing temperature was conducive to the evaporation of phenolic species, thus inhibited the secondary production of 4-methyl-5-nitrocatechol through heterogeneous or aqueous-phase reactions. The study of Gilardoni et al. (2016) also indicated that lower temperature and higher RH were conducive to the formation of biomass burning SOA.

The aromatic SOA tracer, phthalic acid, showed different correlation patterns over the sampling periods in different seasons. Chamber studies have revealed that the enhanced aerosol LWC could significantly increase the yields of aromatic SOA (Jia and Xu, 2018; Lu et al., 2019; Zhou et al., 2019). In this study, both phthalic acid and the ratio of phthalic acid/OC showed significant positive relationships with RH and LWC in January and April, suggesting that the acid-catalyzed heterogeneous or aqueous-phase reactions significantly contributed to the formation of phthalic acid during the study periods in winter and spring. In contrast, phthalic acid showed significant positive relationship with $O_3$ ($r=0.65$, $p<0.01$) and negative correlation ($r=0.60$, $p<0.01$) with RH in July, which might be due to the enhanced photochemical processes on the formation of phthalic acid during the study period in summer (Kawamura and Yasui, 2005). Different from 4-methyl-5-nitrocatechol, phthalic acid showed significantly positive correlations with temperature except in April, likely due to the higher temperature accelerated the production rates of phthalic acid (Kawamura and Yasui, 2005). Though the lower temperature might facilitate the gas-to-particle partition of phthalic acid, this process seemed not to play a dominant role during the study periods in four seasons.

The isoprene SOA tracer, 2-methylerythritol, could be formed through the gas-phase reaction with ·OH radical (Claeys et al., 2004a) and acid-catalyzed heterogeneous oxidation with $H_2O_2$ (Claeys et al., 2004b). In addition, it can also be formed by reactive uptake of the isoprene-derived epoxydiols (IEPOX) generated in the gaseous phase and subsequent aqueous-phase processing (Surratt et al., 2010; Xu et al., 2015). As shown in Table S7, during the whole study period, 2-methylerythritol exhibited significant positive relationships with temperature ($r=0.60$, $p<0.01$), RH ($r=0.55$, $p<0.01$), $O_3$ ($r=0.33$, $p<0.01$), the aerosol acidity ($r=0.79$, $p<0.01$) and LWC ($r=0.56$, $p<0.01$), suggesting that both the gas-phase photooxidation and aqueous-phase reactions played significant roles in the formation of 2-methylerythritol. From Figure 3 and Table S7, 8, it seems that the heterogeneous or aqueous-phase were more closely associated with the formation of 2-methylerythritol during the study periods in autumn and winter, while the photochemical oxidation played an enhancing role over the study period in summer.

2-Methylerythritol did not correlate significantly ($p>0.05$) with temperature in January, April and October, however, when the temperature was above 25°C, it grew rapidly as temperature increased, similar to the result found in the previous research (Liang et al., 2012).

The monoterpene SOA tracer, 3-hydroxyglutaric acid also showed significant positive relationships with temperature ($r=0.63$, $p<0.01$), RH ($r=0.53$, $p<0.01$), $O_3$ ($r=0.37$, $p<0.01$), aerosol acidity ($r=0.82$, $p<0.01$) and LWC ($r=0.58$, $p<0.01$) during the whole study period (Table S7). Besides, 3-hydroxyglutaric acid correlated strongly with 2-methylerythritol ($r=0.94$, $p<0.01$), implying similar influencing factors for the formation of these two biogenic SOA tracers. The correlation pattern of another monoterpene SOA tracer, *cis*-pinonic acid, was different from that of 3-hydroxyglutaric acid. Chamber studies showed that *cis*-pinonic acid could be produced through gas-phase reactions of monoterpenes (Yu et al., 1999; Larsen et al., 2001), which significantly contributed to the newly nucleated particles (Zhang et al., 2012). Previous field observations also found that *cis*-pinonic acid was closely associated with the nucleation processes as the first step in the SOA formation from organic vapors (Alier et al. 2013; van Drooge et al., 2018). On the whole year scale, both *cis*-pinonic acid ($r=0.35$, $p<0.01$) and the ratio of *cis*-pinonic acid/OC ($r=0.64$, $p<0.01$) indeed exhibited significant correlations with $O_3$. However, no significant correlation was found between *cis*-pinonic acid and $O_3$ in July, possibly because it is an unstable intermediate and might more easily further generate higher-generation products such as 3-hydroxyglutaric acid (Kourtchev et al., 2009). And this speculation was supported by the significant positive correlation between the ratio of 3-hydroxyglutaric acid to *cis*-pinonic acid and $O_3$ in July ($r=0.52$, $p<0.01$). Besides, *cis*-pinonic acid showed significant positive correlations with temperature in January ($r=0.37$, $p<0.05$) and October ($r=0.37$, $p<0.05$), while the correlation became weaker in April and even negative in July. At lower temperature, increasing temperature might facilitate the emission of monoterpenes and the formation rate of *cis*-pinonic acid. However, when the weather became warmer, the enhanced temperature might facilitate the transformation of *cis*-pinonic acid to higher-generation products or its evaporation into the gas phase (Li et al., 2013; Ding et al., 2016).

## 3.3 Primary sources and secondary generation of WSOC

### 3.3.1 Source apportionment of WSOC

Source apportionment of PMF was conducted to investigate the source contributions of WSOC as well as its moderately and strongly hydrophilic fractions. Nine types of sources were identified in this study as shown in Figure 4. Factor 1 showed high levels of levoglucosan and EC, thus was interpreted as the direct emissions from biomass burning. Factor 2 had a high level of cholesterol, hence was identified as cooking. Factor 3 exhibited a large fraction of EC which can not be explained by the direct emissions of biomass burning, suggesting that it was the direct emissions from other combustion sources, such as coal combustion, traffic and waste burning, etc. Factor 4 was characterized by high loadings of $Mg^{2+}$ and $Ca^{2+}$, thus was regarded as dust. No significant EC but high proportions of 4-methyl-5-nitrocatechol and phthalic acid were observed in Factor 5 and Factor 6, respectively, which were identified as SOC from biomass burning (biomass burning SOC) and aromatic precursors (aromatic SOC), respectively. Factor 7 exhibited a high level of *cis*-pinonic acid, thus was explained as the freshly generated biogenic SOC. Factor 8 was featured by high fractions of 2-methylerythritol and 3-hydroxyglutaric acid, which are the end oxidation products from isoprene and monoterpenes respectively, hence was recognized to be the aged biogenic SOC. Note that 3-hydroxyglutaric acid and *cis*-pinonic acid were not grouped in one factor though they are both tracers of monoterpenes, due to their different oxidation degree as discussed above. Factor 9 covered the secondary components (such as $SO_4^{2-}$, $NO_3^-$, $NH_4^+$ and $C_2O_4^{2-}$) that can not be well explained by the identified sources above, thus was considered to be SOA from other sources. More detailed discussion can be found in the supplementary materials (S3).

Source contributions to the total WSOC as well as its moderately and strongly hydrophilic fractions are illustrated in Figure 5. During the whole sampling period, the primary emissions of biomass burning contributed 23.0 % to the total WSOC, with a higher contribution to its strongly hydrophilic fraction (37.9 %) than to the moderately hydrophilic portion (15.2 %), which was probably due to the large amounts of saccharides with high water-solubility from biomass burning (Yan et al., 2019; Xu et al., 2020). The total contribution of primary and secondary biomass burning to WSOC (35.1 %) was slightly lower than those previously reported in Beijing (Cheng et al., 2013; Li et al., 2018; Duan et al., 2020), likely due to the effective control of the open biomass burning activities in the surrounding areas of Beijing in recent years. Other primary combustion sources (Factor 3) also contributed significantly to WSOC (13.0 %), most of which contributed to its moderately hydrophilic fraction (16.5 %). A recent source apportionment based on CMAQ model in North China reported that coal combustion contributed 15.1 % to water-soluble HULIS (HULISws) annually, which is the major component of moderately hydrophilic WSOC (Li et al., 2019a). High concentrations of HULISws were also observed in the coal combustion smoke, again suggesting that coal combustion might be a significant source of HULISws (Fan et al., 2016). The primary emission strength of coal combustion was the strongest in winter among four seasons, since the domestic heating required extra amounts of coal combustion in this season. However, the contribution of Factor 3 to WSOC during the study period in winter was not the highest among four seasons, implying that there could be other sources beyond coal combustion included in Factor 3. Previous studies using the PMF model also found that traffic and waste burning both contributed more than 15 % to HULISws in Beijing (Ma et al., 2018; Li et al., 2019c). Therefore, the mixed primary sources in Factor 3 possibly consisted of coal combustion, traffic emission, waste incineration, etc. Previous AMS studies in Beijing indicated that cooking contributed more than 10 % to total organic aerosols (Hu et al., 2016; Sun et al., 2018; Duan et al., 2020). However, the contribution of cooking to WSOC was quite low (2.5 %) in this study, probably because it contributed more significantly to the water-insoluble fraction of organic aerosols (Zhao et al., 2007). Our results also showed that cooking only contributed to the moderately hydrophilic WSOC. Dust contributed 7.2 % to WSOC, with 4.6 % to its moderately hydrophilic fraction and 12.1 % to strongly hydrophilic fraction. The different contributions of dust between the moderately and strongly hydrophilic fractions might be explained by the strongly hydrophilic fulvic acids from soil resuspension and the strongly hydrophilic saccharides carried by dust (Li et al., 2018).

As shown in Figure 5 (a), the secondary sources contributed 54.4 % to total WSOC, slightly higher than the results of Cheng et al. (2013) and Tao et al. (2016), and similar to that of Du et al. (2014). Secondary sources showed a higher contribution to the moderately hydrophilic WSOC (58.5 %) than the strongly hydrophilic WSOC (44.6 %). The aromatic SOC was the most abundant secondary source of WSOC (28.8 %) as well as its moderately (25.9 %) and strongly hydrophilic fractions (33.5 %) in urban Beijing (Tang et al., 2018). The biogenic SOC (11.0 %) and biomass burning SOC (12.1 %) contributed comparable proportions to total WSOC, and both contributed more to the moderately hydrophilic WSOC. Biomass burning SOC, which is mainly generated from the phenolic compounds, has been widely recognized to contribute notably to water-soluble brown carbon ($\mathrm{HULIS_{WS}}$) (Smith et al., 2016; Pang et al., 2020). Biogenic SOC in the strongly hydrophilic WSOC was more aged than that in moderately hydrophilic WSOC, possibly because the atmospheric aging processes would generally increase the polarity of organic aerosols (Baduel et al., 2011; Wu et al., 2016; Kuang et al., 2020).

### 3.3.2 Temporal variation of the source contributions

The seasonal variation of the source contributions to total WSOC and its moderately and strongly hydrophilic fractions are shown in Figure 5 (b). WSOC were mainly from secondary sources during the study periods in summer (75.1 %) and winter (67.4 %), but dominated by primary emissions over the sampling periods in spring (65.2 %) and autumn (71.5 %), similar to the result of Qiu et al. (2020). During the sampling period in winter, WSOC mainly originated from anthropogenic sources (Li et al., 2018; Zhang et al., 2018), including aromatic SOC (38.3 %), biomass burning SOC (27.6 %), and primary biomass burning (21.6 %). The significant contributions of anthropogenic SOC during the study period in winter probably resulted from the stronger emissions of aromatic precursors and phenolic compounds from the domestic heating activities such as the household combustion of solid fuels (Liu et al., 2016; Ding et al., 2017). The meteorological conditions over the study period in winter (slow wind speed, low temperature and high relative humidity) also favored the accumulation and heterogeneous uptake of the SOC precursors, leading to stronger SOC production. It is therefore crucial to control biomass burning and the aromatic precursors to reduce WSOC in winter. During the sampling period in summer, the largest contributor to WSOC was biogenic SOC (39.8 %). A significant fraction of the biogenic SOC was highly oxidized during the study period in summer, in consistent with the higher O/C ratio found in summer in the previous research (Hu et al., 2016; Xu et al., 2017; Qiu et al., 2020). Aromatic SOC (31.3 %) also contributed significantly to WSOC during the sampling period in summer, as reported previously in the summer of Beijing (Guo et al., 2012). Unlike the case in winter with stronger anthropogenic emissions and stagnant and humid meteorological conditions, the high fraction of aromatic SOC in summer was largely associated with the stronger photooxidation capacity. Both primary and secondary contributions from biomass burning were negligible during the study period in summer. The direct emissions of biomass burning contributed higher portions to WSOC during the study periods in spring (27.0 %) and autumn (41.8 %), with much greater contributions to the strongly hydrophilic WSOC than the moderately hydrophilic WSOC in both seasons. Besides, during the sampling period in spring when the weather was windy and dusty, dust showed the highest contribution (23.3 %) to WSOC among four seasons. The contribution of fresh biogenic SOC (13.0 %) was significantly enhanced during the sampling period in spring compared to that in winter as the temperature gradually increased, while the contribution of aromatic SOC (19.3 %) decreased dramatically compared to that in winter due to the cessation of heating activities (Yu et al., 2019).

As shown in Figure 1 as well as in previous studies, winter and autumn in North China Plain are the two seasons that endure more severe air pollution and show higher concentrations of $PM_{2.5}$. Therefore, the source contributions to WSOC during the sampling periods in winter and autumn under different pollution levels were also investigated, as shown in Figure 6. During the study period in winter, both primary and secondary contributions from biomass burning increased significantly from the clean days to moderate hazy days, while aromatic SOC and biomass burning SOC dominated during the severe hazy days (Zhang et al., 2018), highlighting the important roles of biomass burning and aromatic SOC over the hazy periods in winter (Elser et al., 2016; Li et al., 2017a; Huang et al., 2019; Yu et al., 2019). During the study period in autumn, the contribution of aromatic SOC gradually increased as the haze conditions aggravated, while the fresh biogenic SOC became less important with the aggravation of haze conditions, which might be due to the inhibited gas-phase photochemical oxidation as discussed above. Again, the change of the source contributions during the study periods in winter and autumn under different pollution levels suggested that the control of biomass burning and reduction of the aromatic precursors would be of great significance for controlling WSOC in hazy days.

### 3.3.3 Implications for the formation of moderately and strongly hydrophilic SOC

As secondary sources contributed significantly to both the moderately hydrophilic WSOC (58.5 %) and strongly hydrophilic WSOC (44.6 %), the key influencing factors for the formation of SOC in moderately and strongly hydrophilic WSOC were explored and compared. Figure 7 (a) shows the RH versus $O_3$ dependence of the ratios of moderately hydrophilic SOC to OC (MH-SOC/OC), strongly hydrophilic SOC to OC (SH-SOC/OC) and strongly hydrophilic SOC to moderately hydrophilic SOC (SH-SOC/MH-SOC) over the whole sampling period. Both the ratios of MH-SOC/OC and SH-SOC/OC increased with $O_3$ and RH, suggesting that both gas-phase photooxidation and heterogeneous or aqueous-phase reactions might play critical roles in the formation of moderately hydrophilic and strongly hydrophilic SOC. Compared to strongly hydrophilic SOC, the ratio of MH-SOC/OC was more dependent on $O_3$ ($r=0.35$, $p<0.01$), and showed slightly higher values in the daytime (paired t test, $p<0.05$), implying a more significant role of the photooxidation in the generation of moderately hydrophilic SOC. The laboratory studies showed that the photo-induced auto-oxidation of PAHs could lead to the formation of HULISws that can not be formed under dark conditions (Haynes et al., 2019). Besides, the photo-induced oligomerization of several phenolic compounds could also form HULISws (Vione et al., 2019). Field observations also reported an obviously higher contribution of the ultrafine aerosol mode to the moderately hydrophilic WSOC in summer, indicating that the formation of moderately hydrophilic SOC was tightly associated with the gaseous phase nucleation (Frka et al., 2018). In comparison, SH-SOC/OC ratio was more sensitive to RH ($r=0.37$, $p<0.01$) and LWC ($r=0.46$, $p<0.01$) than to $O_3$ ($p>0.05$), implying a more significant impact of the heterogeneous or aqueous-phase processing on the formation of strongly hydrophilic SOC. The SH-SOC/MH-SOC ratio exhibited strong correlations with RH ($r=0.61$, $p<0.01$) and LWC ($r=0.60$, $p<0.01$), which also indicated that the heterogeneous or aqueous-phase oxidation might facilitate the transformation of moderately hydrophilic SOC to the strongly hydrophilic SOC. The quantum chemical calculations in a recent study suggested that the carbenium ion-mediated reactions which involve highly hydrophilic organic species occurred efficiently in the weakly acidic aerosols and cloud/fog droplets, contributing significantly to the SOC generation (Ji et al., 2020). The hygroscopicity of organic aerosols increased with their aging degree during their evolution processes (Jimenez et al., 2009; Wu et al., 2016; Kuang et al., 2020), therefore, the ratio of SH-SOC/MH-SOC was also employed to indicate the aging degree of SOC. The significant positive correlations between the ratio of SH-SOC/MH-SOC and RH, similar to a previous AMS study in Beijing that observed higher O/C ratios with higher RH (Xu et al., 2017), indicated that aqueous-phase oxidation might play a major role in the aging processes of SOA. Research on the aging processes in fog droplets and aerosols also supported the finding that the aqueous-phase reactions might facilitate the formation of more oxidized and strongly hydrophilic SOA (Brege et al., 2018). Previous observation suggested that a large fraction of ambient SOA was more oxidized than those formed in the dry smog chambers, in while SOA could only be produced through gas-phase oxidation (Aiken et al., 2008). There have been some hypotheses for the difference between the chamber SOA and ambient SOA, such as the losses of vapors to the walls and the autoxidation in the chamber (McVay et al., 2016; Thornto et al., 2020). The results of this study also indicated that the aqueous-phase processing, which can produce more hydrophilic SOA, may be one of the reasons for the discrepancy in the oxidation degrees of ambient SOA and chamber SOA (Ervens et al., 2011).

To further investigate the key influencing factors for the formation of moderately and strongly hydrophilic SOC during the sampling periods in four seasons, the correlation coefficients between the above ratios and some meteorological parameters, $O_3$, $PM_{2.5}$, aerosol acidity and aerosol LWC in fours seasons are illustrated in Figure 7 (b). In January, both the MH-SOC/OC and SH-SOC/OC ratios exhibited strong positive relationships with RH, aerosol acidity and LWC, suggesting that the acid-catalyzed heterogeneous or aqueous-phase reactions might be the dominant formation mechanisms of both moderately and strongly hydrophilic SOC during the sampling period in winter. Such result was consistent with the previous finding that the heterogeneous or aqueous-phase oxidation dominated in winter when the gaseous photochemical oxidation was usually weak (Duan et al., 2016; Wu et al., 2019; Yu et al., 2019). However, in July, SOC/OC exhibited significant positive correlations with $O_3$ ($r=0.44$, $p<0.05$) and temperature ($r=0.54$, $p<0.01$), implying that the photooxidation processing might be the major formation pathway of SOC in WSOC, possibly due to the stronger solar radiation in summer (Tang et al., 2016; Duan et al., 2020). However, heterogeneous or aqueous-phase oxidation might still play a critical role in transforming less oxidized MH-SOC to more oxidized SH-SOC based on the significant correlation between the SH-SOC/MH-SOC ratio and LWC ($r=0.60$, $p<0.01$). During the sampling periods in spring and autumn, similarly based on the correlation pattern of the respective ratio with the influencing factors, the production of moderately hydrophilic SOC showed a stronger linkage to the photooxidation process, while the aqueous-phase reactions might play a more critical role in the formation of strongly hydrophilic SOC and the aging processes of SOA. Besides, the SOC/OC ratio showed strong positive relationships with $PM_{2.5}$ in January ($r=0.81$ $p<0.01$), reflecting enhanced SOC formation as the pollution aggravated during the sampling period in winter (Zhang et al., 2014; Li et al., 2019b). However, no such relationship was found in July ($r=0.02$, $p>0.05$) when the gaseous photooxidation dominated in SOC formation.

## 4 Conclusions

Based on the WSOC and related SOA tracer analysis for the $PM_{2.5}$ samples collected in downtown Beijing in four seasons of 2017, the moderately hydrophilic fraction of WSOC dominated in WSOC throughout the sampling period, which showed the highest proportion to WSOC during the sampling period in summer. However, the ratio of WSOC/OC increased as pollution aggravated and the ratio of strongly hydrophilic WSOC to total WSOC increased with $PM_{2.5}$ during the sampling periods in seasons other than summer. Compared to the previous studies in Beijing over the past decades, the reduction of WSOC in this study was not as obvious as that of OC, indicating that the control of WSOC is a more challenging task than the control of water-insoluble organics.

The secondary sources contributed more than 50 % to WSOC, with higher contributions during the study periods in summer (75.1 %) and winter (67.4 %) than in spring (34.8 %) and autumn (28.5 %). Aromatic SOC (28.8 %) was the most abundant secondary source over the entire study period. Biomass burning SOC played a significant role (27.6 %) during the sampling period in winter, while biogenic SOC showed the highest contribution (39.8 %) during the sampling period in summer. The direct emissions from biomass burning and other primary combustion sources were the major primary sources of moderately hydrophilic WSOC, while strongly hydrophilic WSOC was largely affected by the direct emissions of biomass burning and dust. The total contribution of primary and secondary biomass burning to WSOC was slightly lower than those previously reported in Beijing due to the effective control of open burning in the surrounding areas. The contributions of aromatic SOC and biomass burning SOC to WSOC obviously increased as pollution aggravated during the haze period in winter. Besides, the contribution of aromatic SOC increased as $PM_{2.5}$ increased in autumn. Therefore, the control of biomass burning and reduction of aromatic precursors would be of great significance for controlling WSOC during the severe haze episodes in winter and autumn.

According to the correlation patterns with the key influencing factors of the gas phase and aqueous phase reactions, the acid-catalyzed heterogeneous or aqueous-phase processing was suggested as the major formation pathway of SOC over the study period in winter, while photooxidation played a critical role during the study period in summer. The photooxidation played a more prominent role in the formation of moderately hydrophilic SOC, whereas heterogeneous or aqueous-phase processing posed more profound effects on the formation of strongly hydrophilic SOC and aging processes of SOC. The SOA modeling based on the chemical transport models has been a powerful tool for SOA study in regional scale. However, SOA modeling remains a challenge and the model studies have shown systematically underestimated SOA concentrations compared to the observed values. The findings of this study would help to reduce the model-measurement discrepancies of SOA by underlining the importance of the SOA properties and the contributions of heterogeneous formation processes in different seasons.

**Data availability.**

The data used in this article are available from the authors upon request (jingchen@bnu.edu.cn).

**Author contribution**

The corresponding author, JC, provided the ideas and funding, discussed the results, and revised the paper, QY conducted the sampling and chemical analysis, analyzed the data, drawn the picture, and wrote the manuscript. WhQ, SmC, YpZ, YwS, KX, MA contributed to the field sampling and put forward suggestions on the discussion. WhQ and YpZ helped with the data processing from FINNv1.5.

**Competing interests.**

The authors declare that they have no conflict of interest.

## References

Aiken, A. C., Decarlo, P. F., Kroll, J. H., Worsnop, D. R., Huffman, J. A., Docherty, K. S., Ulbrich, I. M., Mohr, C., Kimmel, J. R., Sueper, D., Sun, Y., Zhang, Q., Trimborn, A., Northway, M. J., Ziemann, P., Canagaratna, M. R., Onasch, T. B., Alfarra, M. R., Prevot, A. S. H., Dommen, J., Duplissy, J., Metzger, A., Baltensperger, U., and Jimenez, J. L.: O/C and OM/OC ratios of primary, secondary, and ambient organic aerosols with high-resolution time-of-flight aerosol mass spectrometry, Environ. Sci. Technol., 42, 4478-4485, doi:10.1021/es703009q, 2008.

Aiken, A. C., Salcedo, D., Cubison, M. J., Huffman, J. A., DeCarlo, P. F., Ulbrich, I. M., Docherty, K. S., Sueper, D., Kimmel, J. R., Worsnop, D. R., Trimborn, A., Northway, M., Stone, E. A., Schauer, J. J., Volkamer, R. M., Fortner, E., Foy, B., Wang, J., Laskin, A., Shutthanandan, V., Zheng, J., Zhang, R., Gaffney, J., Marley, N. A., Paredes-Miranda, G., Arnott, W. P., Molina, L. T., Sosa, G., and Jimenez, J. L.: Mexico City aerosol analysis during MILAGRO using high resolution aerosol mass spectrometry at the urban supersite (T0) - Part 1: Fine particle composition and organic source apportionment, Atmos. Chem. Phys. 9, 6633-6653, 2009.

Alier, M., van Drooge, B. L., Dallosto, M., Querol, X., Grimalt, J. O., and Tauler, R.: Source apportionment of submicron organic aerosol at an urban background and a road site in Barcelona (Spain) during SAPUSS, Atmos. Chem. Phys., 13, 10353-10371, doi:10.5194/acp-13-10353-2013, 2013.

Al-Naiema, I., and Stone, E.: Evaluation of anthropogenic secondary organic aerosol tracers from aromatic hydrocarbons, Atmos. Chem. Phys., 17, 2053-2065, doi:10.5194/acp-17-2053-2017, 2017.

Andreae, M. O., and Gelencser, A.: Black carbon or brown carbon? The nature of light-absorbing carbonaceous aerosols, Atmos. Chem. Phys., 6, 3131-3148, doi:10.5194/acp-6-3131-2006, 2006.

Baduel, C., Monge, M. E., Voisin, D., Jaffrezo, J., George, C., Haddad, I. E., Marchand, N., and Danna, B.: Oxidation of atmospheric humic like substances by ozone: a kinetic and structural analysis approach, Environ. Sci. Technol., 45, 5238-5244, doi:10.1021/es200587z, 2011.

Bertrand, A., Stefenelli, G., Jen, C. N., Pieber, S. M., Bruns, E. A., Ni, H., Temimeroussel, B., Slowik, J. G., Goldstein, A. H., Haddad, I. E., Baltensperger, U., Prevot, A. S. H., Wortham, H., and Marchandet, N.: Evolution of the chemical fingerprint of biomass burning organic aerosol during aging, Atmos. Chem. Phys., 18, 7607-7624, doi:10.5194/acp-18-7607-2018, 2018.

Bozzetti, C., Sosedova, Y., Xiao, M., Daellenbach, K. R., Ulevicius, V., Dudoitis, V., Mordas, G., Bycenkiene, S., Plauskaite, K., Vlachou, A., Golly, B., Chazeau, B., Besombes, J. L., Baltensperger, U., Jaffrezo, J. L., Slowik, J. G., Haddad, I. E., and Prevot, A. S. H.: Argon offline-AMS source apportionment of organic aerosol over yearly cycles for an urban, rural, and marine site in northern Europe, Atmos. Chem. Phys., 17, 117-141, doi:10.5194/acp-17-117-2017, 2017.

Brege, M., Paglione, M., Gilardoni, S., Decesari, S., Facchini, M. C., and Mazzoleni, L.: Molecular insights on aging and aqueous-phase processing from ambient biomass burning emissions-influenced Po Valley fog and aerosol, Atmos. Chem. Phys., 18, 13197-13214, doi:10.5194/acp-18-13197-2018, 2018.

Chen, J., Qiu, S., Shang, J., Wilfrid, O. M., Liu, X., Tian, H., and Boman, J.: Impact of relative humidity and water soluble constituents of $PM_{2.5}$ on visibility impairment in Beijing, China, Aerosol Air Qual. Res., 14, 260-268, doi: 10.4209/aaqr.2012.12.0360, 2014.

Chen, J., Li, C., Ristovski, Z., Milic, A., Gu, Y., Islam, M. S., Wang, S., Hao, J., Zhang, H., He, C., Guo, H., Fu, H., Miljevic, B., Morawska, L., Thai, P. K., Lam, Y. F., Pereira, G., Ding, A., Huang, X., and Dumka, U. C.: A review of biomass burning: Emissions and impacts on air quality, health and climate in China, Sci. Total Environ., 579, 1000-1034, doi: 10.1016/j.scitotenv.2016.11.025, 2017.

Chen, P., Kang, K., Tripathee, L., Ram, K., Rupakheti, M., Panday, A. K., Zhang, Q., Guo, J., Wang, X., Pu, T., and Li, C.: Light absorption properties of elemental carbon (EC) and water-soluble brown carbon (WS-BrC) in the Kathmandu Valley, Nepal: a 5-year study, Environ. Pollut., 261, 114239, doi:10.1016/j.envpol.2020.114239, 2020.

Chen, Q., Wang, M., Wang, Y., Zhang, L., Li, Y., and Han, Y.: Oxidative potential of water-soluble matter associated with chromophoric substances in $PM_{2.5}$ over Xi'an, China, Environ. Sci. Technol., 53, 8574-8584, doi:10.1021/acs.est.9b01976, 2019.

Cheng, J., Su, J., Cui, T., Li, X., Dong, X., Sun, F., Yang, Y., Tong, D., Zheng, Y., Li, Y., Li, J., Zhang, Q., and He, K.: Dominant role of emission reduction in $PM_{2.5}$ air quality improvement in Beijing during 2013–2017: a model-based decomposition analysis, Atmos. Chem. Phys, 19, 6125-6146., doi:10.5194/acp-19-6125-2019, 2019.

Cheng, X., Li, H., Zhang, Y., Li, Y., Zhang, W., Wang, X., Bi, F., Zhang, H., Gao, J., Chai, F., Lun, X., Chen, Y., Gao, J., and Lu, J.: Atmospheric isoprene and monoterpenes in a typical urban area of Beijing: pollution characterization, chemical reactivity and source identification, J. Environ. Sci., 71, 150-167, doi:10.1016/j.jes.2017.12.017, 2018.

Cheng, Y., He, K., Duan, F., Zheng, M., Du, Z., Ma, Y., and Tan, J.: Ambient organic carbon to elemental carbon ratios: influences of the measurement methods and implications, Atmos. Environ., 45, 2060-2066, doi:10.1016/j.atmosenv.2011.01.064, 2011.

Cheng, Y., Engling, G., He, K., Duan, F., Ma, Y., Du, Z., Liu, J., Zheng, M., and Weber, R. J.: Biomass burning contribution to Beijing aerosol, Atmos. Chem. Phys., 13, 7765-7781, doi:10.5194/acp-13-7765-2013, 2013.

Cheng, Y., He, K., Du, Z., Zheng, M., Duan, F., and Ma, Y.: Humidity plays an important role in the $PM_{2.5}$ pollution in Beijing, Environ. Pollut., 197, 68-75, doi:10.1016/j.envpol.2014.11.028, 2015.

Cheng, Y., Zheng, G., Wei, C., Mu, Q., Zheng, B., Wang, Z., Gao, M., Zhang, Q., He, K., Carmichael, G. R., Poschl, U., and Su, H.: Reactive nitrogen chemistry in aerosol water as a source of sulfate during haze events in China, Sci. Adv., 2, e1601530, doi:10.1126/sciadv.1601530, 2016.

Claeys, M., Graham, B., Vas, G., Wang, W., Vermeylen, R., Pashynska, V., Cafmeyer, J., Guyon, P., Andreae, M. O., Artaxo, P., and Maenhaut, W.: Formation of secondary organic aerosols through photooxidation of isoprene, Science, 303, 1173-1176, doi:10.1126/science.1092805, 2004a.

Claeys, M., Wang, W., Ion, A. C., Kourtchev, I., Gelencser, A., and Maenhaut, W.: Formation of secondary organic aerosols from isoprene and its gas-phase oxidation products through reaction with hydrogen peroxide, Atmos. Environ., 38, 4093-4098, doi:10.1016/j.atmosenv.2004.06.001, 2004b.

Ding, X., Zheng, M., Yu, L., Zhang, X., Weber, R. J., Yan, B., Russell, A. G., Edgerton, E. S., and Wang, X.: Spatial and seasonal trends in biogenic secondary organic aerosol tracers and water-soluble organic carbon in the southeastern United States, Environ. Sci. Technol., 42, 5171-5176, doi:10.1021/es7032636, 2008.

Ding, X., Zhang, Y., He, Q., Yu, Q., Shen, R., Zhang, Y., Zhang, Z., Lyu, S., Hu, Q., Wang, Y., Li, L., Song, W., and Wang, X.: Spatial and seasonal variations of secondary organic aerosol from terpenoids over China, J. Geophys. Res. Atmos., 121, 14661-14678, doi:10.1002/2016JD025467, 2016.

Ding, X., Zhang, Y., He, Q., Yu, Q., Wang, J., Shen, R., Song, W., Wang, Y., and Wang, X.: Significant increase of aromatics-derived secondary organic aerosol during fall to winter in China, Environ. Sci. Technol., 51, 7432-7441, doi:10.1021/acs.est.6b06408, 2017.

Du, Z., He, K., Cheng, Y., Duan, F., Ma, Y., Liu, J., Zhang, X., Zheng, M., and Weber, R.: A yearlong study of water-soluble organic carbon in Beijing I: Sources and its primary vs. secondary nature, Atmos. Environ., 92, 514-521, doi:10.1016/j.atmosenv.2014.04.060, 2014.

Duan, F., He, K., Ma, Y., Ihozaki, T., Kawasaki, H., Arakawa, R., Kitayama, S., Tujimoto, K., Huang, T., Kimoto, T., Furutani, H., and Toyoda, M.: High molecular weight organic compounds (HMW-OCs) in severe winter haze: Direct observation and insights on the formation mechanism, Environ. Pollut., 218, 289-296, doi:10.1016/j.envpol.2016.07.004, 2016.

Duan, J., Huang, R., Li, Y., Chen, Q., Zheng, Y., Chen, Y., Lin, C., Ni, H., Wang, M., Ovadnevaite, J., Ceburnis, D., Chen, C., Worsnop, D. R., Hoffmann, T., Odowd, C. D., and Cao, J.: Summertime and wintertime atmospheric processes of secondary aerosol in Beijing, Atmos. Chem. Phys., 20, 3793–3807, doi:10.5194/acp-20-3793-2020, 2020.

Elser, M., Huang, R. J., Wolf, R., Slowik, J. G., Wang, Q., Canonaco, F., Li, G., Bozzetti, C., Daellenbach, K. R., Huang, Y., Zhang, R., Li, Z., Cao, J., Baltensperger, U., El-Haddad, I., and Prevot, A. S. H.: New insights into $PM_{2.5}$ chemical composition and sources in two major cities in China during extreme haze events using aerosol mass spectrometry, Atmos. Chem. Phys., 16, 3207-3225, doi:10.5194/acp-16-3207-2016, 2016.

Ervens, B., Turpin, B. J., and Weber, R. J.: Secondary organic aerosol formation in cloud droplets and aqueous particles (aqSOA): a review of laboratory, field and model studies, Atmos. Chem. Phys., 11, 11069-11102, doi:10.5194/acp-11-11069-2011, 2011.

Fan, X., Song, J., and Peng, P.: Comparison of isolation and quantification methods to measure humic-like substances (HULIS) in atmospheric particles, Atmos. Environ., 60, 366-374, doi:10.1016/j.atmosenv.2012.06.063, 2012.

Fan, X., Wei, S., Zhu, M., Song, J., and Peng, P.: Comprehensive characterization of humic-like substances in smoke $PM_{2.5}$ emitted from the combustion of biomass materials and fossil fuels, Atmos. Chem. Phys., 6, 13321-13340, doi: 10.5194/acp-16-13321-2016, 2016.

Feng, B., Li, L., Xu, H., Wang, T., Wu, R., Chen, J., Zhang, Y., Liu, S., Ho, S. S. H., Cao, J., and Huang, W.: $PM_{2.5}$-bound polycyclic aromatic hydrocarbons (PAHs) in Beijing: seasonal variations, sources, and risk assessment, J. Environ. Sci., 77, 11-19, doi:10.1016/j.jes.2017.12.025, 2018.

Feng, J., Li, M., Zhang, P., Gong, S., Zhong, M., Wu, M., Zheng, M., Chen, C., Wang, H., and Lou, S.: Investigation of the sources and seasonal variations of secondary organic aerosols in $PM_{2.5}$ in Shanghai with organic tracers, Atmos. Environ., 79, 614-622, doi:10.1016/j.atmosenv.2013.07.022, 2013.

Feng, T., Zhao, S., Bei, N., Wu, J., Liu, S., Li, X., Liu, L., Qian, Y., Yang, Q., Wang, Y., Zhou, W., Cao, J., and Li, G.: Secondary organic aerosol enhanced by increasing atmospheric oxidizing capacity in Beijing–Tianjin–Hebei (BTH), China, Atmos. Chem. Phys., 19, 7429-7443, doi:10.5194/acp-19-7429-2019, 2019.

Fountoukis, C. and Nenes, A.: ISORROPIA II: a computationally efficient thermodynamic equilibrium model for $K^+$-$Ca^{2+}$-$Mg^{2+}$-$NH_4^+$-$Na^+$-$SO_4^{2-}$-$NO_3^-$-$Cl^-$-$H_2O$ aerosols, Atmos. Chem. Phys., 7, 4639-4659, 2007.

Frka, S., Grgic, I., Tursic, J., Gini, M. I., and Eleftheriadis, K.: Seasonal variability of carbon in humic-like matter of ambient size-segregated water soluble organic aerosols from urban background environment, Atmos. Environ., 173, 239-247, doi:10.1016/j.atmosenv.2017.11.013, 2018.

Geng, X., Mo, Y., Li, J., Zhong, C., Tang, J., Jiang, H., Ding, X., Malik, R. N., and Zhang, G.: Source apportionment of water-soluble brown carbon in aerosols over the northern South China Sea: Influence from land outflow, SOA formation and marine emission, Atmos. Environ., 229, 117484, doi:10.1016/j.atmosenv.2020.117484, 2020.

George, C., Ammann, M., Danna, B., Donaldson, D. J., and Nizkorodov, S. A.: Heterogeneous photochemistry in the atmosphere, Chem. Rev., 115, 4218-4258, doi:10.1021/cr500648z, 2015.

Gilardoni, S., Massoli, P., Paglione, M., Giulianelli, L., Carbone, C., Rinaldi, M., Decesari, S., Sandrini, S., Costabile, F., Gobbi, G. P., Pietrogrande, M. C., Visentin, M., Scotto, F., Fuzzi, S., and Facchini, M. C.: Direct observation of aqueous secondary organic aerosol from biomass-burning emissions, P. Natl. Sci. USA, 113, 10013-10018, doi:10.1073/pnas.1602212113, 2016.

Guo, S., Hu, M., Guo, Q., Zhang, X., Zheng, M., Zheng, J., Chang, C. C., Schauer, J. J., and Zhang, R.: Primary sources and secondary formation of organic aerosols in Beijing, China, Environ. Sci. Technol., 46, 9846-9853, doi:10.1021/es2042564, 2012.

Guo, S., Hu, M., Zamora, M. L., Peng, J., Shang, D., Zheng, J., Du, Z., Wu, Z., Shao, M., Zeng, L., Molina, M. J., and Zhang, R.: Elucidating severe urban haze formation in China, P. Natl. Sci. USA, 2014, 111, 17373-17378, doi:10.1073/pnas.1419604111, 2014.

Hartikainen, A., Ylipirila, P., Tiitta, P., Leskinen, A., Kortelainen, M., Orasche, J., Schnellekreis, J., Lehtinen, K. E. J., Zimmermann, R., Jokiniemi, J., and Sippula, O.: Volatile organic compounds from logwood combustion: emissions and transformation under dark and photochemical aging conditions in a smog chamber, Environ. Sci. Technol., 52, 4979-4988, doi:10.1021/acs.est.7b06269, 2018.

Haynes, J. P., Miller, K. E., and Majestic, B. J.: Investigation into photoinduced auto-oxidation of polycyclic aromatic hydrocarbons resulting in brown carbon production, Environ. Sci. Technol., 53, 682-691, doi:10.1021/acs.est.8b05704, 2019.

Hennigan, C. J., Bergin, M. H., Russell, A. G., Nenes, A., and Weber, R. J.: Gas/particle partitioning of water-soluble organic aerosol in Atlanta, Atmos. Chem. Phys., 9, 3613-3628, doi:10.5194/acp-9-3613-2009, 2009.

Hu, W., Hu, M., Hu, W., Jimenez, J. L., Yuan, B., Chen, W., Wang, M., Wu, Y., Chen, C., Wang, Z., Peng, J., Zeng, L., and Shao, M.: Chemical composition, sources, and aging process of submicron aerosols in Beijing: Contrast between summer and winter, J. Geophys. Res. Atmos., 121, 1955-1977, doi:10.1002/2015JD024020, 2016.

Hu, W., Hu, M., Hu, W., Zheng, J., Chen, C., Wu, Y., and Guo, S.: Seasonal variations in high time-resolved chemical compositions, sources, and evolution of atmospheric submicron aerosols in the megacity Beijing, Atmos. Chem. Phys., 17, 9979-10000, doi:10.5194/acp-17-9979-2017, 2017.

Huang, G., Liu, Y., Shao, M., Li, Y., Chen, Q., Zheng, Y., Wu, Z., Liu, Y., Wu, Y., Hu, M., Li, X., Lu, S., Wang, C., Liu, J., Zheng, M., and Zhu, T.: Potentially important contribution of gas-phase oxidation of naphthalene and methylnaphthalene to secondary organic aerosol during haze events in Beijing, Environ. Sci. Technol., 53, 1235-1244, doi:10.1021/acs.est.8b04523, 2019.

Huang, R. J., Zhang, Y., Bozzetti, C., Ho, K., Cao, J., Han, Y., 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., Schnellekreis, J., Zimmermann, R., An, Z., Szidat, S., Baltensperger, U., Haddad, I. E., and Prevot, A. S. H.: High secondary aerosol contribution to particulate pollution during haze events in China, Nature, 2014, 514, 218-222, doi:10.1038/nature13774, 2014.

Huang, R. J., Yang, L., Shen, J., Yuan, W., Gong, Y., Guo, J., Cao, W., Duan, J., Ni, H., Zhu, C., Dai, W., Li, Y., Chen, Y., Chen, Q., Wu, Y., Zhang, R., Dusek, U., O'Dowd, C., and Hoffmann, T.: Water-insoluble organics dominate brown carbon in wintertime urban aerosol of China: chemical characteristics and optical properties, Environ. Sci. Technol., doi: 10.1021/acs.est.0c01149, 2020.

Huang, X., Yu, J., He, L., and Yuan, Z.: Water-soluble organic carbon and oxalate in aerosols at a coastal urban site in China: Size distribution characteristics, sources, and formation mechanisms, J. Geophys. Res. Atmos., 111, D22212, doi: 10.1029/2006JD007408, 2006.

Iinuma, Y., Boge, O., Grafe, R., and Herrmann, H.: Methyl-nitrocatechols: atmospheric tracer compounds for biomass burning secondary organic aerosols, Environ. Sci. Technol., 44, 8453-8459, doi:10.1021/es102938a, 2010.

Islam, M. R., Jayarathne, T., Simpson, I. J., Werden, B., Maben, J. R., Gilbert, A., Praveen, P. S., Adhikari, S., Panday, A. K., Rupakheti, M., Blake, D. R., Yokelson, R. J., Decarlo, P. F., Keene, W. C., and Stone, E. A.: Ambient air quality in the Kathmandu Valley, Nepal, during the pre-monsoon: concentrations and sources of particulate matter and trace gases, Atmos. Chem. Phys., 20, 2927-2951, doi:10.5194/acp-20-2927-2020, 2020.

Ji, Y., Shi, Q., Li, Y., An, T., Zheng, J., Peng, J., Gao, Y., Chen, J., Li, G., Wang, Y., Zhang, F., Zhang, A. L., Zhao, J., Molina, M. J., and Zhang, R.: Carbenium ion-mediated oligomerization of methylglyoxal for secondary organic aerosol formation, P. Natl. Sci. USA, doi:10.1073/pnas.1912235117, 2020.

Jia, L., and Xu, Y.: Different roles of water in secondary organic aerosol formation from toluene and isoprene, Atmos. Chem. Phys., 18, 8137-8154, doi:10.5194/acp-18-8137-2018, 2018.

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

Kahnt, A., Behrouzi, S., Vermeylen, R., Shalamzari, M. S., Vercauteren, J., Roekens, E., Claeys, M., and Maenhaut, W.: One-year study of nitro-organic compounds and their relation to wood burning in $PM_{10}$ aerosol from a rural site in Belgium, Atmos. Environ., 81, 561-568, doi:10.1016/j.atmosenv.2013.09.041, 2013.

Kamens, R. M., Zhang, H., Chen, E. H., Zhou, Y., Parikh, H. M., Wilson, R. L., Galloway, K. E., and Rosen, E. P.: Secondary organic aerosol formation from toluene in an atmospheric hydrocarbon mixture: Water and particle seed effects, Atmos. Environ., 45, 2324-2334, doi:10.1016/j.atmosenv.2010.11.007, 2011.

Kang, M., Fu, P., Kawamura, K., Yang, F., Zhang, H., Zang, Z., Ren, H., Ren, L., Zhao, Y., Sun, Y., and Wang, Z.:. Characterization of biogenic primary and secondary organic aerosols in the marine atmosphere over the East China Sea, Atmos. Chem. Phys., 18, 13947-13967, doi:10.5194/acp-18-13947-2018, 2018a.

Kang, M., Ren, L., Ren, H., Zhao, Y., Kawamura, K., Zhang, H., Wei, L., Sun, Y., Wang, Z., and Fu, P.: Primary biogenic and anthropogenic sources of organic aerosols in Beijing, China: Insights from saccharides and n-alkanes, Environ. Pollut., 243, 1579-1587, doi:10.1016/j.envpol.2018.09.118, 2018b.

Kaur, R., Labins, J. R., Helbock, S. S., Jiang, W., Bein, K. J., Zhang, Q., and Anastasio, C.: Photooxidants from brown carbon and other chromophores in illuminated particle extracts, Atmos. Chem. Phys., 19, 6579-6594, doi:10.5194/acp-19-6579-2019, 2019.

Kawamura, K. and Kaplan, I. R.: Motor exhaust emissions as a primary source for dicarboxylic acids in Los-Angeles ambient air, Environ. Sci. Technol., 21, 105-110, 1987.

Kawamura, K., and Yasui, O.: Diurnal changes in the distribution of dicarboxylic acids, ketocarboxylic acids and dicarbonyls in the urban Tokyo atmosphere, p 39, 1945-1960, doi:10.1016/j.atmosenv.2004.12.014, 2005.

Kiss, G., Varga, B., Galambos, I., and Ganszky, I.: Characterization of water-soluble organic matter isolated from atmospheric fine aerosol, J. Geophys. Res. Atmos., 107(D21), 8339, doi:10.1029/2001JD000603, 2002.

Kleindienst, T. E., Jaoui, M., Lewandowski, M., Offenberg, J. H., Lewis, C. W., Bhave, P. V., and Edney, E. O.: Estimates of the contributions of biogenic and anthropogenic hydrocarbons to secondary organic aerosol at a southeastern US location, Atmos. Environ., 41, 8288-8300, doi:10.1016/j.atmosenv.2007.06.045, 2007.

Kleinman, L. I., Springston, S. R., Daum, P. H., Lee, Y. N., Nunnermacker, L. J., Senum, G. I., Wang, J., Weinstein-Lloyd, J., Alexander, M. L., Hubbe, J., Ortega, J., Canagaratna, M. R., and Jayne, J.: The time evolution of aerosol composition over the Mexico City plateau, Atmos. Chem. Phys. 8, 1559-1575, 2008.

Kourtchev, I., Copolovici, L., Claeys, M., and Maenhaut, W.: Characterization of atmospheric aerosols at a forested site in Central Europe, Environ. Sci. Technol., 43, 4665-4671, doi:10.1021/es803055w, 2009.

Kuang, B., Lin, P., Huang, X., and Yu, J.: Sources of humic-like substances in the Pearl River Delta, China: positive matrix factorization analysis of $PM_{2.5}$ major components and source markers, Atmos. Chem. Phys., 15, 1995-2008, doi: 10.5194/acp-15-1995-2015, 2015.

Kuang, Y., He, Y., Xu, W., Zhao, P., Cheng, Y., Zhao, G., Tao, J., Ma, N., Su, H., Zhang, Y., Sun, J., Cheng, P., Yang, W., Zhang, S., Wu, C., Sun, Y., and Zhao, C.: Distinct diurnal variation in organic aerosol hygroscopicity and its relationship with oxygenated organic aerosol, Atmos. Chem. Phys., 20, 865-880, doi:10.5194/acp-20-865-2020, 2020.

Larsen, B. R., Bella, D. D., Glasius, M., Winterhalter, R., Jensen, N. R., and Hjorth, J.: Gas-phase OH oxidation of monoterpenes: gaseous and particulate products, J. Atmos. Chem., 38, 231-276, doi:10.1023/A:1006487530903, 2001.

Li, H., Zhang, Q., Zhang, Q., Chen, C., Wang, L., Wei, Z., Zhou, S., Parworth, C. L., Zheng, B., Canonaco, F., Prevot, A. S. H., Chen, P., Zhang, H., Wallington, T. J., and He, K.: Wintertime aerosol chemistry and haze evolution in an extremely polluted city of the North China Plain: significant contribution from coal and biomass combustion, Atmos. Chem. Phys., 17, 4751-4768, doi:10.5194/acp-17-4751-2017, 2017a.

Li, J., Wang, G., Cao, J., Wang, X., and Zhang, R.: Observation of biogenic secondary organic aerosols in the atmosphere of a mountain site in central China: temperature and relative humidity effects, Atmos. Chem. Phys., 13, 11535-11549, doi:10.5194/acp-13-11535-2013, 2013.

Li, J., Du, H., Wang, Z., Sun, Y., Yang, W., Li, J., Tang, X., and Fu, P.: Rapid formation of a severe regional winter haze episode over a mega-city cluster on the North China Plain, Environ. Pollut., 223, 605-615, doi:10.1016/j.envpol.2017.01.063, 2017b.

Li, L., Ren, L., Ren, H., Yue, S., Xie, Q., Zhao, W., Kang, M, Li, J., Wang, Z., Sun, Y., and Fu, P.: Molecular characterization and seasonal variation in primary and secondary organic aerosols in Beijing, China, J. Geophys. Res. Atmos., 123, 12394-12412, doi:10.1029/2018JD028527, 2018.

Li, X., Han, J., Hopke, P. K, Hu, J., Shu, Q., Chang, Q., and Ying, Q.: Quantifying primary and secondary humic-like substances in urban aerosol based on emission source characterization and a source-oriented air quality model, Atmos. Chem. Phys., 19, 2327-2341, doi:10.5194/acp-19-2327-2019, 2019a.

Li, X., Wu, J., Elser, M., Tong, S., Liu, S., Li, X., Liu, L., Cao, J., Zhou, J., Elhaddad, I., Huang, R., Ge, M., Tie, X., Prevot, A. S. H., and Li, G.: Wintertime secondary organic aerosol formation in Beijing-Tianjin-Hebei (BTH): contributions of HONO sources and heterogeneous reactions, Atmos. Chem. Phys., 19, 2343-2359, doi:10.5194/acp-19-2343-2019, 2019b.

Li, X., Yang, K., Han, J., Ying, Q., and Hopke, P. K.: Sources of humic-like substances (HULIS) in $PM_{2.5}$ in Beijing: receptor modeling approach, Sci. Total Environ., 671, 765-775, doi:10.1016/j.scitotenv.2019.03.333, 2019c.

Li, Y., Huang, D., Cheung, H. Y., Lee, A. K. Y., and Chan, C. K.: Aqueous-phase photochemical oxidation and direct photolysis of vanillin - a model compound of methoxy phenols from biomass burning, Atmos. Chem. Phys., 14, 27641-27675, doi:10.5194/acp-14-2871-2014, 2014.

Liang, L., Engling, G., Duan, F., Cheng, Y., and He, K.: Characteristics of 2-methyltetrols in ambient aerosol in Beijing, China, Atmos. Environ., 59, 376-381, doi:10.1016/j.atmosenv.2012.05.052, 2012.

Lin, P., Huang, X., He, L., and Yu, J.: Abundance and size distribution of HULIS in ambient aerosols at a rural site in South China, J. Aerosol. Sci., 41, 74-87, doi: 10.1016/j.jaerosci.2009.09.001, 2010.

Liu, J., Mauzerall, D. L., Chen, Q., Zhang, Q., Song, Y., Peng, W., Klimont, Z., Qiu, X., Zhang, S., Hu, M., Lin, W., Smith, K. R., and Zhu, T.: Air pollutant emissions from Chinese households: a major and underappreciated ambient pollution source, P. Natl. Sci. USA, 113, 7756-7761, doi:10.1073/pnas.1604537113, 2016.

Lu, J., Ge, X., Liu, Y., Chen, Y., Xie, X., Ou, Y., Ye, Z., and Chen, M.: Significant secondary organic aerosol production from aqueous-phase processing of two intermediate volatility organic compounds, Atmos. Environ., 211, 63-68, doi:10.1016/j.atmosenv.2019.05.014, 2019.

Ma, Y., Cheng, Y., Qiu, X., Gao, G., Fang, Y., Wang, J., Zhu, T., Yu, J. Z., and Hu, D.: Sources and oxidative potential of water-soluble humic-like substances ($HULIS_{WS}$) in fine particulate matter ($PM_{2.5}$) in Beijing, Atmos. Chem. Phys., 18, 5607-5617, doi:10.5194/acp-18-5607-2018, 2018.

Ma, T., Furutani, H., Duan, F., Kimoto, T., Jiang, J., Zhang, Q., Xu, X., Wang, Y., Gao, J., Geng, G., Li, M., Song, S., Ma, Y., Che, F., Wang, J., Zhu, L., Huang, T., Toyoda, M., and He, K.: Contribution of hydroxymethanesulfonate (HMS) to severe winter haze in the North China Plain, Atmos. Chem. Phys., 1-17, doi:10.5194/acp-2020-113, 2020.

Manfrin, A., Nizkorodov, S. A., Malecha, K. T., Getzinger, G. J., Mcneill, K., and Borduasdedekind, N.: Reactive oxygen species production from secondary organic aerosols: the importance of singlet oxygen, Environ. Sci. Technol., 53, 8553-8562, doi: 10.1021/acs.est.9b01609, 2019.

McVay, R. C., Zhang, X., Aumont, B., Valorso, R., Camredon, M., La, Y. S., Wennberg, P. O., and Seinfeld, J. H.: SOA formation from the photooxidation of α-pinene: systematic exploration of the simulation of chamber data, Atmos. Chem. Phys., 16, 2785-2802, doi:10.5194/acp-16-2785-2016, 2016.

Miyazaki, Y., Kondo, Y., Shiraiwa, M., Takegewa, N., Miyakawa, T., Han, S., Kita, K., Hu, M., Deng, Z., Zhao, Y., Sugimoto, N., Blake, D. R., and Weber, R. J.: Chemical characterization of water-soluble organic carbon aerosols at a rural site in the Pearl River Delta, China, in the summer of 2006, J. Geophys. Res. Atmos., 114, D12408, doi:10.1029/2009JD011736, 2009.

Norris, G., Duvall, R., Brown, S., and Bai, S.: EPA positive matrix factorization (PMF) 5.0 fundamentals and user guide, US Environmental Protection Agency, Office of Research and Development, Washington, DC, 2014.

Pang, H., Zhang, Q., Lu, X., Li, K., Chen, H., Chen, J., Yang, X., Ma, Y., Ma, J., and Cheng, H.: Nitrite-mediated photooxidation of vanillin in the atmospheric aqueous phase, Environ. Sci. Technol., 53, 14253-14263, doi:10.1021/acs.est.9b03649, 2020.

Paatero, P., Eberly, S., Brown, S. G., and Norris, G. A.: Methods for estimating uncertainty in factor analytic solutions, Atmos. Meas. Tech., 7, 781-797, doi: 10.5194/amt-7-781-2014, 2014.

Pye, H. O. T., Nenes, A., Alexander, B., Ault, A. P., Barth, M. C., Clegg, S. L., Collett, J. L., Fahey, K. M., Hennigan, C. J., Herrmann, H., Kanakidou, M., Kelly, J. T., Ku, I. T., McNeill, V. F., Riemer, N., Schaefer, T., Shi, G., Tilgner, A., Walker, J. T., Wang, T., Weber, R., Xing, J., Zaveri, R. A., and Zuend, A.: The acidity of atmospheric particles and clouds, Atmos. Chem. Phys., 20, 4809-4888, doi:10.5194/acp-20-4809-2020, 2020.

Qian, X., Shen, H., and Chen, Z.: Characterizing summer and winter carbonyl compounds in Beijing atmosphere, Atmos. Environ., 214, 116845, doi:10.1016/j.atmosenv.2019.116845, 2019.

Qiu, Y., Xu, W., Jia, L., He, Y., Fu, P., Zhang, Q., Xie, Q., Hou, S., Xie, C., Xu, Y., Wang, Z., Worsnop, D. R., and Sun, Y.: Molecular composition and sources of water-soluble organic aerosol in summer in Beijing, Chemosphere, 225, 126850, doi:10.1016/j.chemosphere.2020.126850, 2020.

Qu, Y., Chen, Y., Liu, X., Zhang, J., Guo, Y., and An, J.: Seasonal effects of additional HONO sources and the heterogeneous reactions of $N_2O_5$ on nitrate in the North China Plain, Sci. Total Environ., 690, 97-107, doi:10.1016/j.scitotenv.2019.06.436, 2019.

Shen, R., Ding, X., He, Q., Cong, Z., Yu, Q., and Wang, X.: Seasonal variation of secondary organic aerosol tracers in Central Tibetan Plateau, Atmos. Chem. Phys., 15, 8781-8793, doi:10.5194/acp-15-8781-2015, 2015.

Shen, X., Vogel, H., Vogel, B., Huang, W., Mohr, C., Ramisetty, R., Leisner, T., Prevot, A. S. H., and Saathoff, H.: Composition and origin of $PM_{2.5}$ aerosol particles in the upper Rhine valley in summer, Atmos. Chem. Phys., 19, 13189-13208, doi:10.5194/acp-19-13189-2019, 2019.

Smith, J. D., Kinney, H., and Anastasio, C.: Phenolic carbonyls undergo rapid aqueous photodegradation to form low-volatility, light-absorbing products, Atmos. Environ., 126, 36-44, doi:10.1016/j.atmosenv.2015.11.035, 2016.

Srivastava, D., Favez, O., Bonnaire, N., Lucarelli, F., Lucarelli, M., Perraudin, E., Gros, V., Villenave, E., and Albinet, A.: Speciation of organic fractions does matter for aerosol source apportionment. Part 2: Intensive short-term campaign in the Paris area (France), Atmos. Environ., 634, 267-278, doi: 10.1016/j.scitotenv.2018.03.296, 2018.

Sun, J., Wu, F., Hu, B., Tang, G., Zhang, J., and Wang, Y.: VOC characteristics, emissions and contributions to SOA formation during hazy episodes, Atmos. Environ., 2016, 141, 560-570, doi:10.1016/j.atmosenv.2016.06.060, 2016.

Sun, Y., Zhang, Q., Anastasio, C., and Sun, J.: Insights into secondary organic aerosol formed via aqueous-phase reactions of phenolic compounds based on high resolution mass spectrometry, Atmos. Chem. Phys., 10, 4809-4822, doi:10.5194/acp-10-4809-2010, 2010.

Sun, Y., Wang, Z., Fu, P., 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, doi:10.1016/j.atmosenv.2013.06.019, 2013.

Sun, Y., Xu, W., Zhang, Q., Jiang, Q., Canonaco, F., Prevot, A. S. H., Fu, P., Li, J., Jayne, J. T., Worsnop, D. R., and Wang, Z.: Source apportionment of organic aerosol from 2-year highly time-resolved measurements by an aerosol chemical speciation monitor in Beijing, China, Atmos. Chem. Phys., 18, 8469-8489, doi:10.5194/acp-18-8469-2018, 2018.

Surratt, J. D., Chan, A. W., Eddingsaas, N. C., Chan, M., Loza, C. L., Kwan, A. J., Hersey, S. P., Flagan, R. C., Wennberg, P. O., and Seinfeld, J. H.: Reactive intermediates revealed in secondary organic aerosol formation from isoprene, P. Natl. Sci. USA, 107, 6640-6645, doi:10.1073/pnas.0911114107, 2010.

Tan, Z., Rohrer, F., Lu, K., Ma, X., Bohn, B., Broch, S., Dong, H., Fuchs, H., Gkatzelis, G. I., Hofzumahaus, A., Holland, F., Li, X., Liu, Y., Liu, Y., Novelli, A., Shao, M., Wang, H., Wu, Y., Zeng, L., Hu, M., Kiendler-Scharr, A., Wahner, A., and Zhang, Y.: Wintertime photochemistry in Beijing: observations of $RO_x$ radical concentrations in the North China Plain during the BEST-ONE campaign, Atmos. Chem. Phys., 18, 12391-12411, doi:10.5194/acp-18-12391-2018, 2018.

Tang, R., Wu, Z., Li, X., Wang, Y., Shang, D., Xiao, Y., Li, M., Zeng, L., Wu, Z., Hallquist, M., Hu, M., and Guo, S.: Primary and secondary organic aerosols in summer 2016 in Beijing, Atmos. Chem. Phys., 18, 4055-4068, doi:10.5194/acp-18-4055-2018, 2018.

Tang, X., Zhang, X., Wang, Z., and Ci, Z.: Water-soluble organic carbon (WSOC) and its temperature-resolved carbon fractions in atmospheric aerosols in Beijing, Atmos. Res., 181, 200-210, doi:10.1016/j.atmosres.2016.06.019, 2016.

Tao, J., Zhang, L., Zhang, R., Wu, Y., Zhang, Z., Zhang, X., Tang, Y., Cao, J., and Zhang, Y.: Uncertainty assessment of source attribution of $PM_{2.5}$ and its water-soluble organic carbon content using different biomass burning tracers in positive matrix factorization analysis-a case study in Beijing, China, Sci. Total Environ., 543, 326-335, doi:10.1016/j.scitotenv.2015.11.057, 2016.

Thornton, J. A., Shilling, J. E., Shrivastava, M., D'Ambro, E. L., Zawadowicz, M. A., and Liu, J.: A near-explicit mechanistic evaluation of isoprene photochemical secondary organic aerosol formation and evolution: simulations of multiple chamber experiments with and without added $NO_x$, ACS Earth Space Chem., 4, 1161-1181, doi: 10.1021/acsearthspacechem.0c00118, 2020.

van Drooge, B. L., Fontal, M., Fernández, P., Fernández, M. A., Muñoz-Arnanz, J., Jiménez, B., and Grimalt, J. O.: Organic molecular tracers in atmospheric $PM_1$ at urban intensive traffic and background sites in two high-insolation European cities, Atmos. Environ., 188, 71-81, doi:10.1016/j.atmosenv.2018.06.024, 2018.

Varga, B., Kiss, G., Ganszky, I., Gelencser, A., and Krivacsy, Z.: Isolation of water soluble organic matter from atmospheric aerosol, Talanta, 55, 561-572, 2011.

Verma, V., Ricomartinez, R., Kotra, N., King, L. E., Liu, J., Snell, T. W., and Weber, R. J.: Contribution of water-soluble and insoluble components and their hydrophobic/hydrophilic subfractions to the reactive oxygen species-generating potential of fine ambient aerosols, Environ. Sci. Technol., 46, 11384-11392, doi:10.1021/es302484r, 2012.

Verma, V., Fang, T., Xu, L., Peltier, R. E., Russell, A. G., Ng, N. L., and Weber, R. G.: Organic aerosols associated with the generation of reactive oxygen species (ROS) by water-soluble $PM_{2.5}$, Environ. Sci. Technol., 49, 4646-4656, doi:10.1021/es505577w, 2015.

Vidovic, K., Jurkovic, D. L., Sala, M., Kroflic, A., and Grgic, I.: Nighttime aqueous-phase formation of nitrocatechols in the atmospheric condensed phase, Environ. Sci. Technol., 52, 9722-9730, doi:10.1021/acs.est.8b01161, 2018.

Vione, D., Albinet, A., Barsotti, F., Mekic, M., Jiang, B., Minero, C., Brigante, M., and Gligorovski, S.: Formation of substances with humic-like fluorescence properties, upon photoinduced oligomerization of typical phenolic compounds emitted by biomass burning, Atmos. Environ., 206, 197-207, doi:10.1016/j.atmosenv.2019.03.005, 2019.

Wang, Y., Zhang, X., and Draxler, R. R.: TrajStat: GIS-based software that uses various trajectory statistical analysis methods to identify potential sources from long-term air pollution measurement data, Environ. Modell. Softw., 24, 938-939, doi:10.1016/j.envsoft.2009.01.004, 2009.

Wang, Y., Hu, M., Wang, Y., Zheng, J., Shang, D., Yang, Y., Liu, Y., Li, X., Tang, R., Zhu, W., Du, Z., Wu, Y., Guo, S., Wu, Z., Lou, S., Hallquist, M., and Yu, J.: The formation of nitro-aromatic compounds under high $NO_X$ and anthropogenic VOC conditions in urban Beijing, China, Atmos. Chem. Phys., 19, 7649-7665, doi:10.5194/acp-19-7649-2019, 2019.

Wang, Y., Wang, M., Li, S., Sun, H., Mu, Z., Zhang, L., Li, Y., and Chen, Q.: Study on the oxidation potential of the water-soluble components of ambient $PM_{2.5}$ over Xi'an, China: pollution levels, source apportionment and transport pathways, Environ. Int., 136, 105515, doi:10.1016/j.envint.2020.105515, 2020.

Wiedinmyer, C., Akagi, S. K., Yokelson, R. J., Emmons, L. K., Al-Saadi, J. A., Orlando, J. J., and Soja, A. J.: The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning, Geosci. Model. Dev. 4, 625-641, doi:10.5194/gmd-4-625-2011, 2011.

Wu, J., Bei, N., Hu, B., Liu, S., Zhou, M., Wang, Q., Li, X., Liu, L., Feng, T., Liu, Z., Wang, Y., Cao, J., Tie, X., Wang, J., Molina, L. T., and Liu, G.: Is water vapor a key player of the wintertime haze in North China Plain, Atmos. Chem. Phys., 19, 8721-8739, doi:10.5194/acp-19-8721-2019, 2019.

Wu, Z., Zheng, J., Shang, D., Du, Z., Wu, Y., Zeng, L., Wiedensohler, A., and Hu, M.: Particle hygroscopicity and its link to chemical composition in the urban atmosphere of Beijing, China, during summertime, Atmos. Chem. Phys., 16, 1123-1138, doi:10.5194/acp-16-1123-2016, 2016.

Xu, L., Guo, H., Boyd, C. M., Klein, M., Bougiatioti, A., Cerully, K. M., Hite, J. R., Isaacman-VanWertz, G., Kreisberg, N. M., Knote, C., Olson, K., Koss, A., Goldstein, A. H., Hering, S. V., de Gouw, J., Baumann, K., Lee, S. H., Nenes, A., Weber, R. J., and Ng, N. L.: Effects of anthropogenic emissions on aerosol formation from isoprene and monoterpenes in the southeastern United States, P. Natl. Sci. USA, 112, 37-42, doi:10.1073/pnas.1417609112, 2015.

Xu, S., Ren, L., Lang, Y., Hou, S., Ren, H., Wei, L., Wu, L., Deng, J., Hu, W., Pan, X., Sun, Y., Wang, Z., Su, H., Cheng, Y., and Fu, P.: Molecular markers of biomass burning and primary biological aerosols in urban Beijing: size distribution and seasonal variation, Atmos. Chem. Phys., doi:10.5194/acp-2019-841, 2020.

Xu, W., Han, T., Du, W., Wang, Q., Chen, C., Zhao, J., Zhang, Y., Li, J., Fu, P., Wang, Z., Worsnop, D. R., and Sun, Y.: Effects of aqueous-phase and photochemical processing on secondary organic aerosol formation and evolution in Beijing, China, Environ. Sci. Technol., 51, 762-770, doi:10.1021/acs.est.6b04498, 2017.

Yan, C., Zheng, M., Sullivan, A. P., Bosch, C., Desyaterik, Y., Andersson, A., Li, X., Guo, X., Zhou, T., Gustafsson, O., and Collett, J. L.: Chemical characteristics and light-absorbing property of water-soluble organic carbon in Beijing: Biomass burning contributions, Atmos. Environ., 121, 4-12, doi:10.1016/j.atmosenv.2015.05.005, 2015.

Yan, C., Zheng, M., Sullivan, A. P., Shen, G., Chen, Y., Wang, S., Zhao, B., Cai, S., Desyaterik, Y., Li, X., Zhou, T., Gustafsson, O., and Collett, J. L.: Residential coal combustion as a source of levoglucosan in China, Environ. Sci. Technol., 52, 1665-1674, doi:10.1021/acs.est.7b05858, 2018.

Yan, C., Sullivan, A. P., Cheng, Y., Zheng, M., Zhang, Y., Zhu, T., and Collett, J. L.: Characterization of saccharides and associated usage in determining biogenic and biomass burning aerosols in atmospheric fine particulate matter in the North China Plain, Sci. Total Environ., 650, 2939-2950, doi:10.1016/j.scitotenv.2018.09.325, 2019.

Yang, S., Duan, F., Ma, Y., He, K., Zhu, L., Ma, T., Ye, S., Li, H., Huang, T., and Kimoto, T.: Haze formation indicator based on observation of critical carbonaceous species in the atmosphere, Environ. Pollut., 244, 84-92, doi:10.1016/j.envpol.2018.10.006, 2019.

Yu, H., Wei, J., Cheng, Y., Subedi, K., and Verma, V.: Synergistic and antagonistic interactions among the particulate matter components in generating reactive oxygen species based on the dithiothreitol assay, Environ. Sci. Technol., 52, 2261-2270, doi:10.1021/acs.est.7b04261, 2018.

Yu, J. Z., Cocker, D. R., Griffin, R. J., Flagan, R. C., and Seinfeld, J. H.: Gas-phase ozone oxidation of monoterpenes: Gaseous and particulate products, J. Atmos. Chem., 34, 207-258, doi:10.1023/A:1006487530903, 1999.

Yu, L., Smith, J., Laskin, A., Anastasio, C., Laskin, J., and Zhang, Q.: Chemical characterization of SOA formed from aqueous-phase reactions of phenols with the triplet excited state of carbonyl and hydroxyl radical, Atmos. Chem. Phys., 14, 13801-13816, doi:10.5194/acp-14-13801-2014, 2014.

Yu, Q., Chen, J., Qin, W., Cheng, S., Zhang, Y., Ahmad, M., and Ouyang, W.: Characteristics and secondary formation of water-soluble organic acids in $PM_1$, $PM_{2.5}$ and $PM_{10}$ in Beijing during haze episodes, Sci. Total Environ., 669, 175-184, doi:10.1016/j.scitotenv.2019.03.131, 2019.

Zhang, C., Zou, Z., Chang, Y., Zhang, Y., Wang, X., and Yang, X.: Source assessment of atmospheric fine particulate matter in a Chinese megacity: insights from long-term, high-time resolution chemical composition measurements from Shanghai flagship monitoring supersite, Chemosphere, 251, 126598, doi:10.1016/j.chemosphere.2020.126598, 2020.

Zhang, J., Sun, Y., Liu, Z., Ji, D., Hu, B., Liu, Q., and Wang, Y.: Characterization of submicron aerosols during a month of serious pollution in Beijing, 2013, Atmos. Chem. Phys., 14, 2887-2903, doi:10.5194/acp-14-2887-2014, 2014.

Zhang, J., Chen, J., Xue, C., Chen, H., Zhang, Q., Liu, X., Mu, Y., Guo, Y., Wang, D., Chen, Y., Li, J., Qu, Y., and An, J.: Impacts of six potential HONO sources on $HO_x$ budgets and SOA formation during a wintertime heavy haze period in the North China Plain, Sci. Total Environ., 681, 110-123, doi:10.1016/j.scitotenv.2019.05.100, 2019.

Zhang, R., Khalizov, A., Wang, L., Hu, M., and Xu, W.: Nucleation and growth of nanoparticles in the atmosphere, Chem. Rev., 112, 1957-2011, doi:10.1021/cr2001756, 2012.

Zhang, Y., Shao, M., Zhang, Y., Zeng, L., He, L., Zhu, B., Wei, J., and Zhu, X.: Source profiles of particulate organic matters emitted from cereal straw burnings, J. Environ. Sci., 19, 167-175, doi:10.1016/S1001-0742(07)60027-8, 2007.

Zhang, Y., El-Haddad, I., Huang, R., Ho, K., Cao, J., Han, Y., Zotter, P., Bozzetti, C., Daellenbach, K. R., Slowik, J. G., Salazar, G., Prevot, A. S. H., and Szidat, S.: Large contribution of fossil fuel derived secondary organic carbon to water soluble organic aerosols in winter haze in China, Atmos. Chem. Phys., 18, 4005-4017, doi:10.5194/acp-18-4005-2018, 2018.

Zhao, W., Kawamura, K., Yue, S., Wei, L., Ren, H., Yan, Y., Kang, M., Li, L., Ren, L., Lai, S., Li, J., Sun, Y., Wang, Z., and Fu, P.: Molecular distribution and compound-specific stable carbon isotopic composition of dicarboxylic acids, oxocarboxylic acids and α-dicarbonyls in $PM_{2.5}$ from Beijing, China, Atmos. Chem. Phys., 18, 2749-2767, doi:10.5194/acp-18-2749-2018, 2018.

Zhao, Y., Hu, M., Slanina, S., and Zhang, Y.: Chemical compositions of fine particulate organic matter emitted from Chinese cooking, Environ. Sci. Technol., 41, 99-105, doi:10.1021/es0614518, 2007.

Zheng, G., Duan, F., Su, H., Ma, Y., Cheng, Y., Zheng, B., Zhang, Q., Huang, T., Kimoto, T., Chang, D., Poschl, U., Cheng, Y., and He, K.: Exploring the severe winter haze in Beijing: the impact of synoptic weather, regional transport and heterogeneous reactions, Atmos. Chem. Phys, 15, 2969-2983, doi:10.5194/acp-15-2969-2015, 2015.

Zhou, C., Jang, M., and Yu, Z.: Simulation of SOA formation from the photooxidation of monoalkylbenzenes in the presence of aqueous aerosols containing electrolytes under various $NO_X$ levels, Atmos. Chem. Phys., 19, 5719-5735, doi:10.5194/acp-19-5719-2019, 2019.

**Table 1** The average concentrations and standard deviations of the identified carbonaceous species in $PM_{2.5}$ during the sampling periods in four seasons.

| **Compounds (μg $m^{-3}$)** | **Winter** | | | **Spring** | | | **Summer** | | | **Autumn** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Daytime** | **Nighttime** | **Mean** | **Daytime** | **Nighttime** | **Mean** | **Daytime** | **Nighttime** | **Mean** | **Daytime** | **Nighttime** | **Mean** |
| CO | 1.7 ± 1.7 | 2.2 ± 2.8 | 1.9 ± 2.3 | 0.6 ± 0.4$^*$ | 0.8 ± 0.4$^*$ | 0.7 ± 0.4 | 1.1 ± 0.4 | 0.9 ± 0.2 | 1.0 ± 0.3 | 1.1 ± 0.4$^*$ | 1.4 ± 0.4$^*$ | 1.2 ± 0.4 |
| $PM_{2.5}$ | 120 ± 107 | 147 ± 154 | 133 ± 131 | 60.6 ± 36.2 | 64.5 ± 34.8 | 62.5 ± 34.9 | 59.8 ± 28.6 | 51.9 ± 20.6 | 55.8 ± 24.8 | 75.2 ± 58.1 | 81.1 ± 50.8 | 78.2 ± 53.7 |
| OC | 20.1 ± 19.2 | 21.0 ± 24.8 | 20.6 ± 21.9 | 7.9 ± 2.6 | 9.5 ± 3.4 | 8.7 ± 3.1 | 8.7 ± 3.4$^*$ | 6.8 ± 4.3$^*$ | 7.8 ± 3.9 | 9.4 ± 3.8 | 10.1 ± 3.7 | 9.7 ± 3.7 |
| EC | 3.9 ± 3.1 | 4.7 ± 5.8 | 4.3 ± 4.6 | 1.9 ± 1.1$^*$ | 2.7 ±1.4$^*$ | 2.3 ± 1.3 | 1.4 ± 1.0 | 1.3 ± 1.0 | 1.3 ± 1.0 | 2.4 ± 1.4$^*$ | 3.4 ± 1.7$^*$ | 2.9 ± 1.6 |
| OC/EC | 4.6 ± 1.1 | 4.3 ± 1.3 | 4.5 ± 1.2 | 5.2 ± 2.1 | 4.4 ± 2.5 | 4.8 ± 2.3 | 6.7 ± 3.9 | 5.3 ± 4.2 | 6.1 ± 4.1 | 4.4 ± 1.6$^*$ | 3.3 ± 0.9$^*$ | 3.8 ± 1.4 |
| WSOC | 11.4 ± 11.3 | 12.0 ± 16.4 | 11.7 ± 13.9 | 4.1 ± 2.0 | 4.7 ± 2.6 | 4.4 ± 2.3 | 5.3 ± 2.1$^*$ | 4.0 ± 2.7$^*$ | 4.7 ± 2.5 | 4.7 ± 3.0 | 4.9 ± 2.8 | 4.8 ± 2.8 |
| WSOC/OC | 0.53 ± 0.08 | 0.51 ± 0.08 | 0.52 ± 0.08 | 0.50 ± 0.10 | 0.47 ± 0.14 | 0.49 ± 0.12 | 0.62 ± 0.11 | 0.59 ± 0.10 | 0.60 ± 0.11 | 0.47 ± 0.12 | 0.46 ± 0.12 | 0.46 ± 0.12 |
| MH-WSOC | 7.9 ± 7.6 | 8.0 ± 10.3 | 8.0 ± 8.9 | 2.8 ± 1.3 | 2.9 ± 1.6 | 2.9 ± 1.5 | 4.1 ± 1.2 | 3.4 ± 1.6 | 3.8 ± 1.5 | 2.9 ± 1.6 | 2.9 ± 1.3 | 2.9 ± 1.5 |
| SH-WSOC | 3.2 ± 3.8 | 4.0 ± 6.1 | 3.6 ± 5.0 | 1.3 ± 0.9 | 1.8 ± 1.1 | 1.6 ± 1.0 | 1.2 ± 1.0$^*$ | 0.7 ± 1.1$^*$ | 1.0 ± 1.1 | 1.8 ± 1.4 | 2.0 ± 1.5 | 1.9 ± 1.4 |
| **Organic tracers (ng $m^{-3}$)** | | | | | | | | | | | | |
| Levoglucosan | 307 ± 300 | 388 ± 394 | 349 ± 348 | 100 ± 87.8$^*$ | 194 ± 175$^*$ | 147 ± 144 | 23.6 ± 11.0 | 34.2 ± 24.2 | 28.9 ± 19.3 | 136 ± 102$^*$ | 234 ± 125$^*$ | 185 ± 123 |
| Cholesterol | 5.0 ± 3.0 | 4.9 ± 3.3 | 4.9 ± 3.1 | 3.9 ± 1.9 | 4.8 ± 2.5 | 4.3 ± 2.3 | 4.1 ± 2.4 | 3.0 ± 1.1 | 3.6 ± 1.9 | 6.1 ± 4.4 | 6.3 ± 3.1 | 6.2 ± 3.8 |
| Phthalic acid | 88.7 ± 84.8 | 90.8 ± 121 | 89.8 ± 103 | 27.3 ± 20.8 | 21.9 ± 14.0 | 24.6 ± 17.7 | 55.9 ± 22.0$^*$ | 17.6 ± 9.1$^*$ | 36.8 ± 25.5 | 27.6 ± 21.8 | 19.9 ± 13.3 | 23.8 ± 18.2 |
| 4-Methyl-5-nitrocatechol | 24.7 ± 26.4 | 35.2 ± 41.0 | 30.1 ± 34.5 | 1.8 ± 1.9 | 3.3 ± 2.7 | 2.6 ± 2.4 | 0.1 ± 0.3 | 0.0 ± 0.0 | 0.1 ± 0.2 | 1.6 ± 1.2$^*$ | 4.4 ± 3.6$^*$ | 3.0 ± 3.0 |
| 2-Methylerythritol | 2.1 ± 2.3 | 2.2 ± 3.5 | 2.2 ± 2.9 | 1.2 ± 0.6 | 1.5 ± 0.8 | 1.4 ± 0.7 | 55.4 ± 48.5 | 41.6 ± 34.6 | 48.5 ± 42.0 | 2.3 ± 1.1 | 2.6 ± 1.2 | 2.5 ± 1.1 |
| 3-Hydroxyglutaric acid | 4.4 ± 3.9 | 4.2 ± 5.0 | 4.3 ± 4.5 | 4.2 ± 2.8 | 4.9 ± 5.1 | 4.6 ± 4.0 | 37.1 ± 22.7$^*$ | 27.3 ± 18.5$^*$ | 32.2 ± 20.9 | 7.5 ± 4.6 | 7.0 ± 4.4 | 7.2 ± 4.5 |
| *cis*-Pinonic acid | 3.3 ± 2.4 | 3.0 ± 2.1 | 3.2 ± 2.2 | 9.0 ± 6.0$^*$ | 6.9 ± 3.6$^*$ | 7.9 ± 5.0 | 7.3 ± 4.2 | 10.1 ± 6.0 | 8.7 ± 5.3 | 7.3 ± 3.0$^*$ | 3.6 ± 0.8$^*$ | 5.5 ± 2.9 |

$^*$ According to paired t test, the values with $^*$ showed statistically significant differences ($p<0.05$) between day and night.

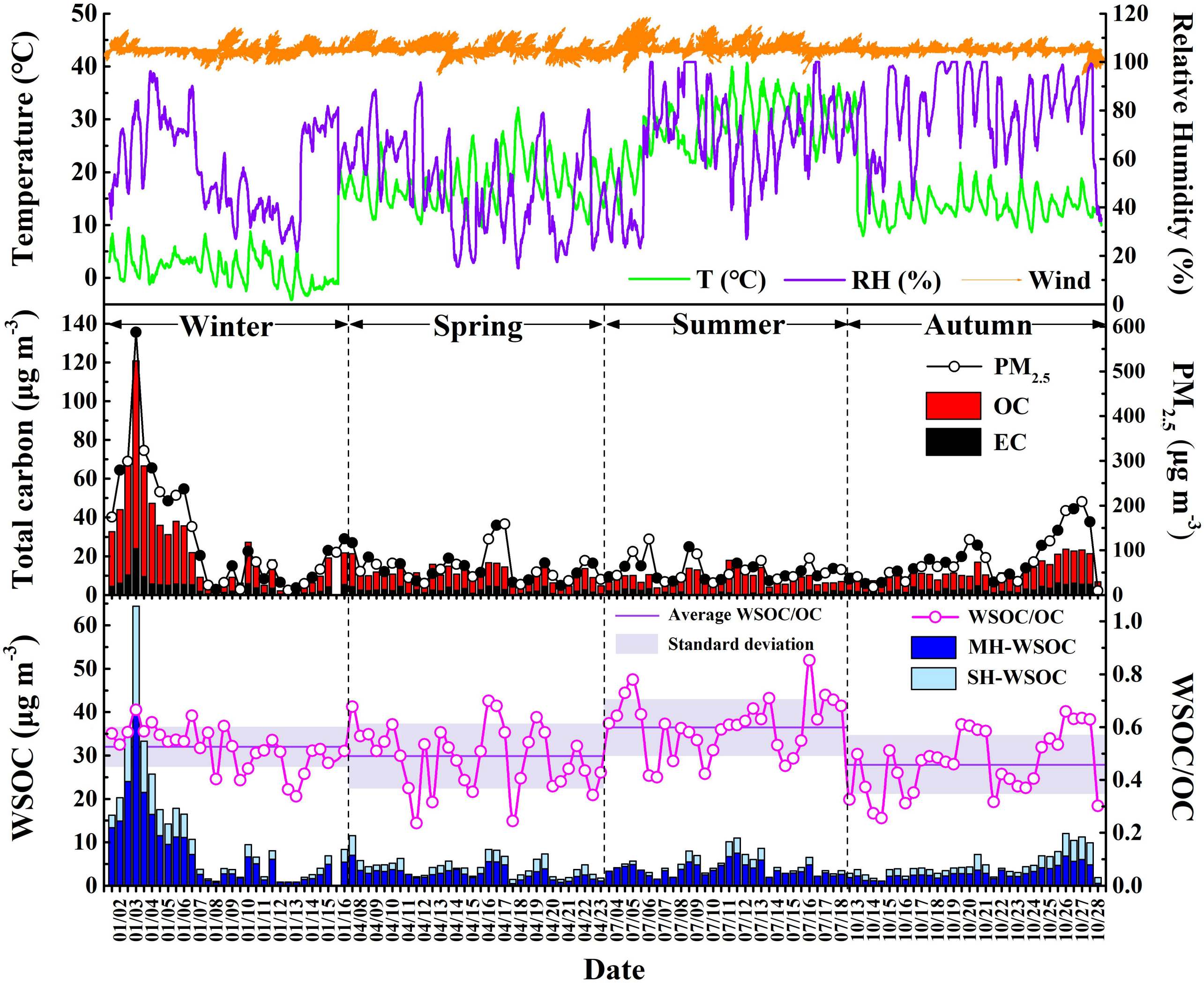
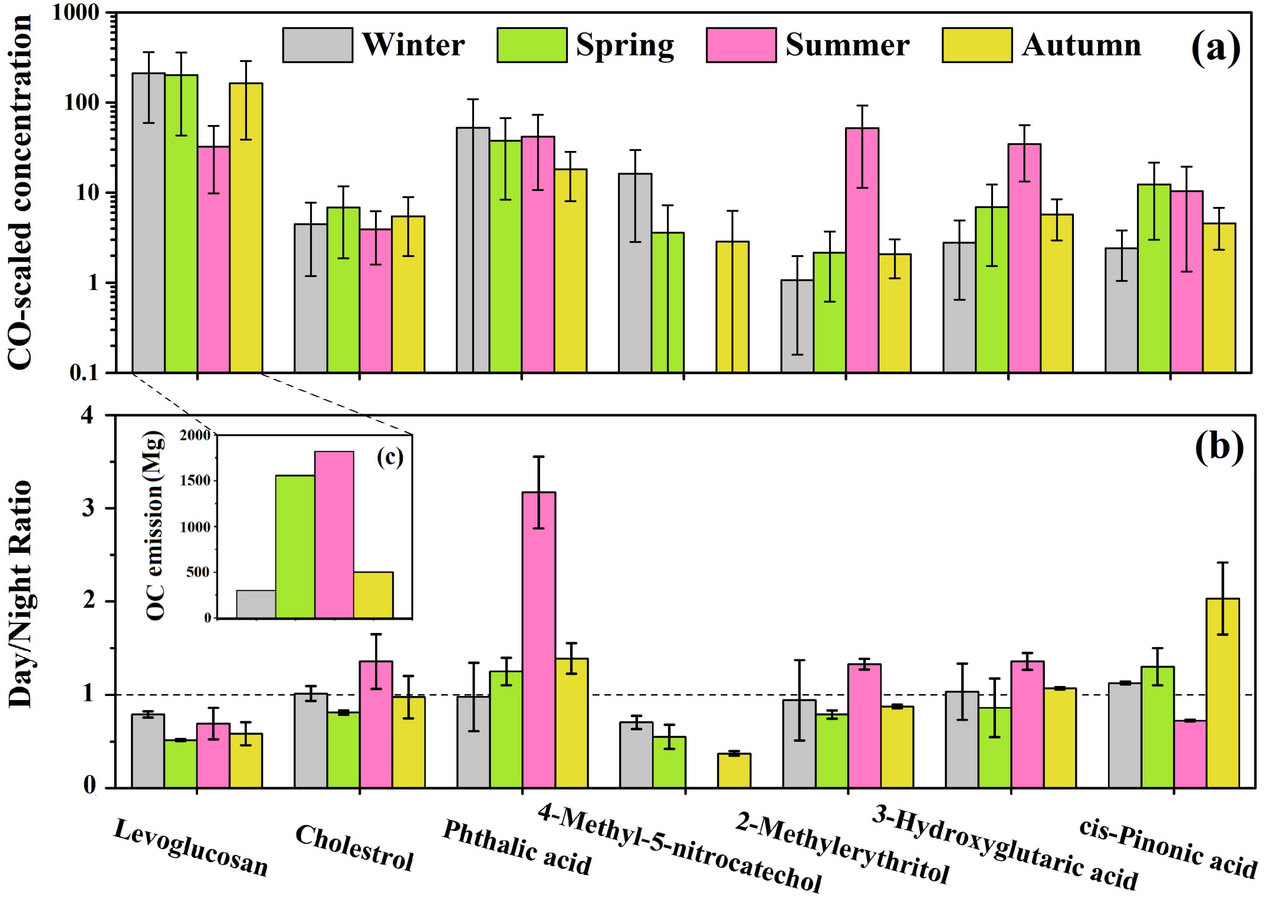
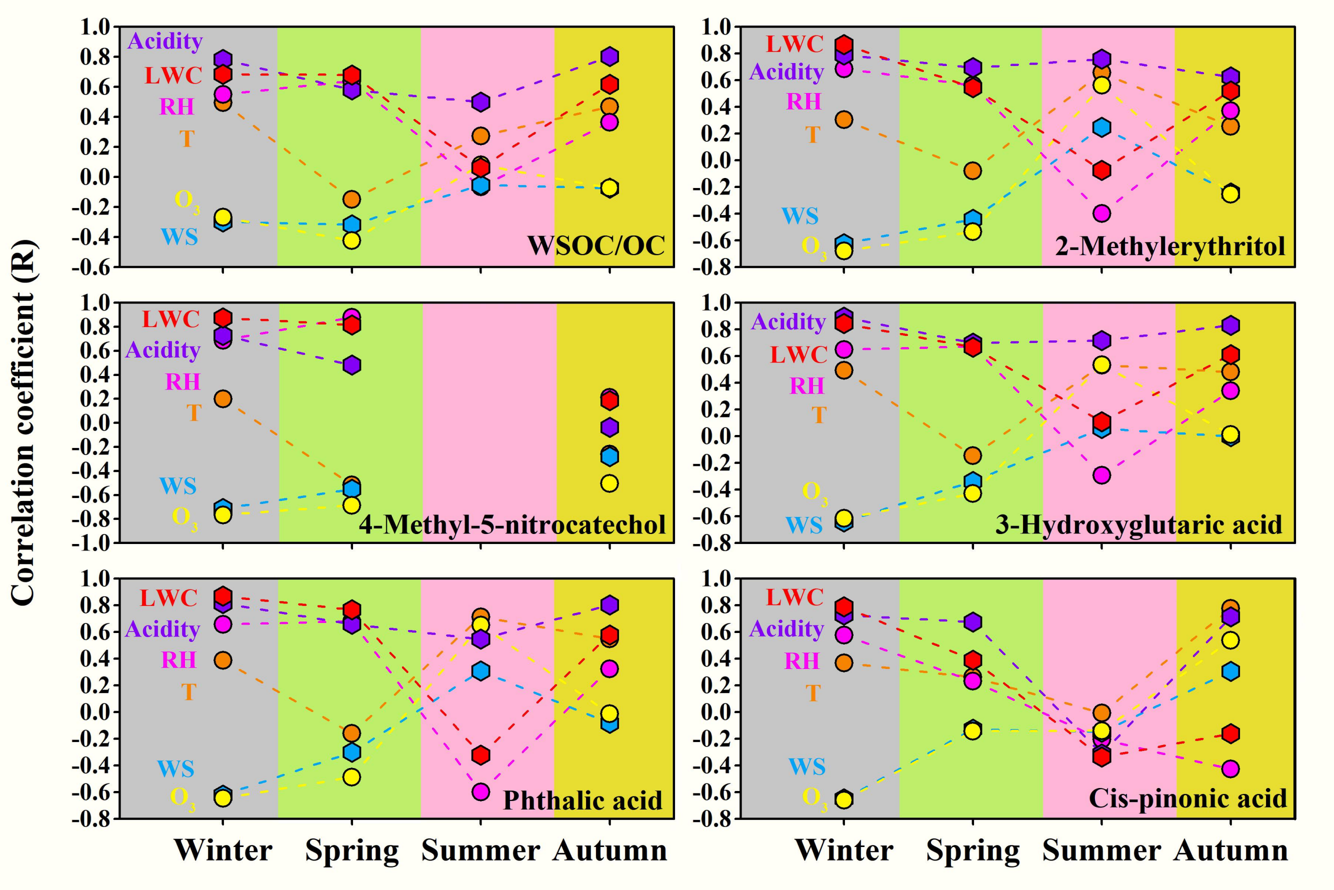


**Figure 1.** Temporal variations of meteorological parameters, the mass concentrations of PM$_{2.5}$, OC, EC, WSOC and WSOC/OC ratio in Beijing during the sampling periods in four seasons of 2017.



5 **Figure 2.** (a) The CO-scaled concentration of the identified organic tracers; (b) the day to night ratios of the measured concentrations of the organic tracers; (c) the OC emission amounts from open biomass burning provided by the Fire Inventory (FINN) in Beijing during the sampling periods in four seasons of 2017.



10 **Figure 3**. Correlation coefficients between SOA tracers and WSOC/OC and the meteorological parameters, $O_3$ concentration, aerosol acidity and LWC during the sampling periods in four seasons of 2017.

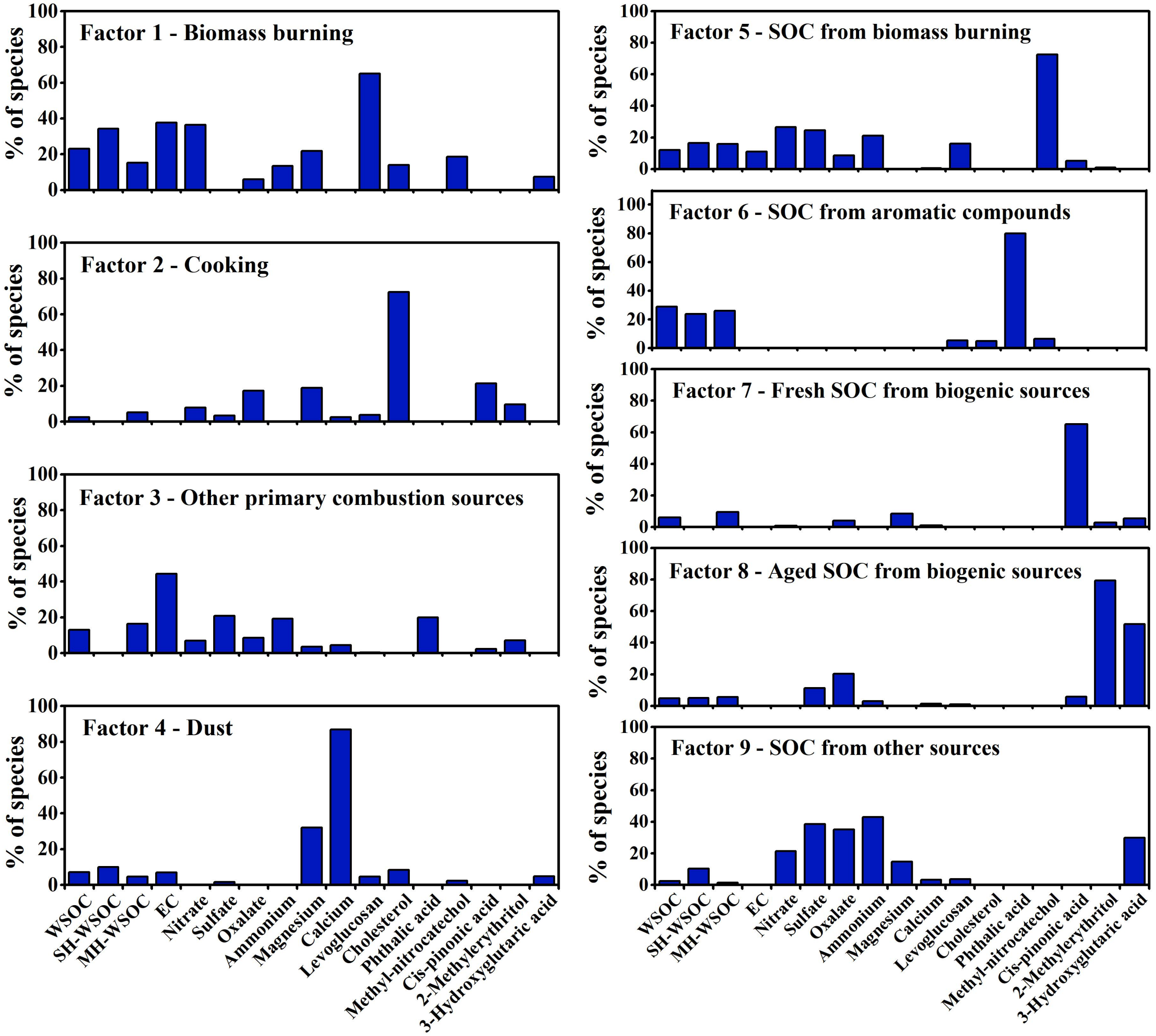


**Figure 4**. The constrained 9-factor solution resolved by the PMF model.

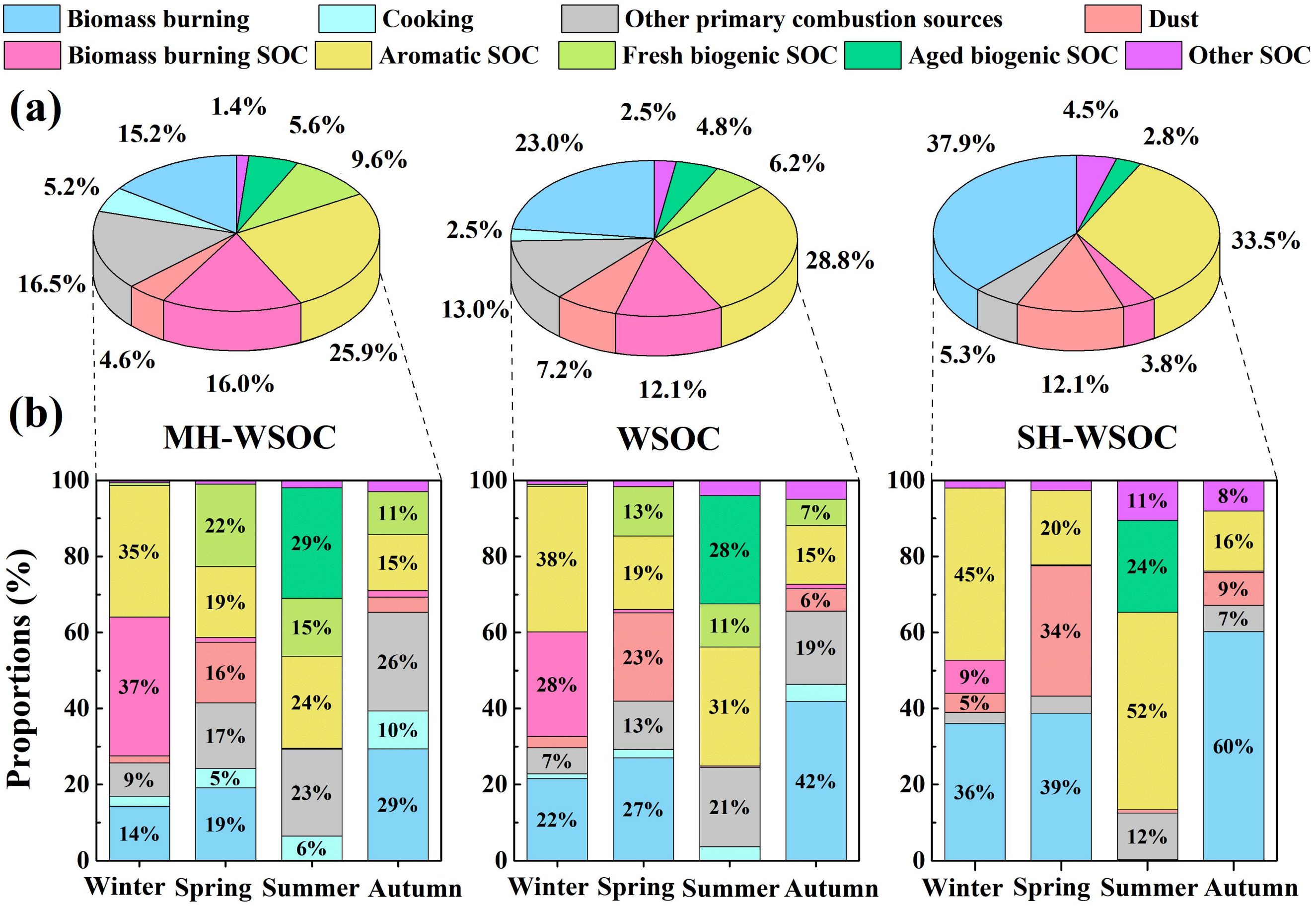


**Figure 5**. Source contributions of WSOC as well as its moderately hydrophilic and strongly hydrophilic fractions in Beijing: (a) comparison of the source contributions to total WSOC, moderately hydrophilic WSOC and strongly hydrophilic WSOC on a whole year scale; (b) seasonal variation of the source contributions to total, moderately and strongly hydrophilic WSOC during the sampling periods in four seasons.

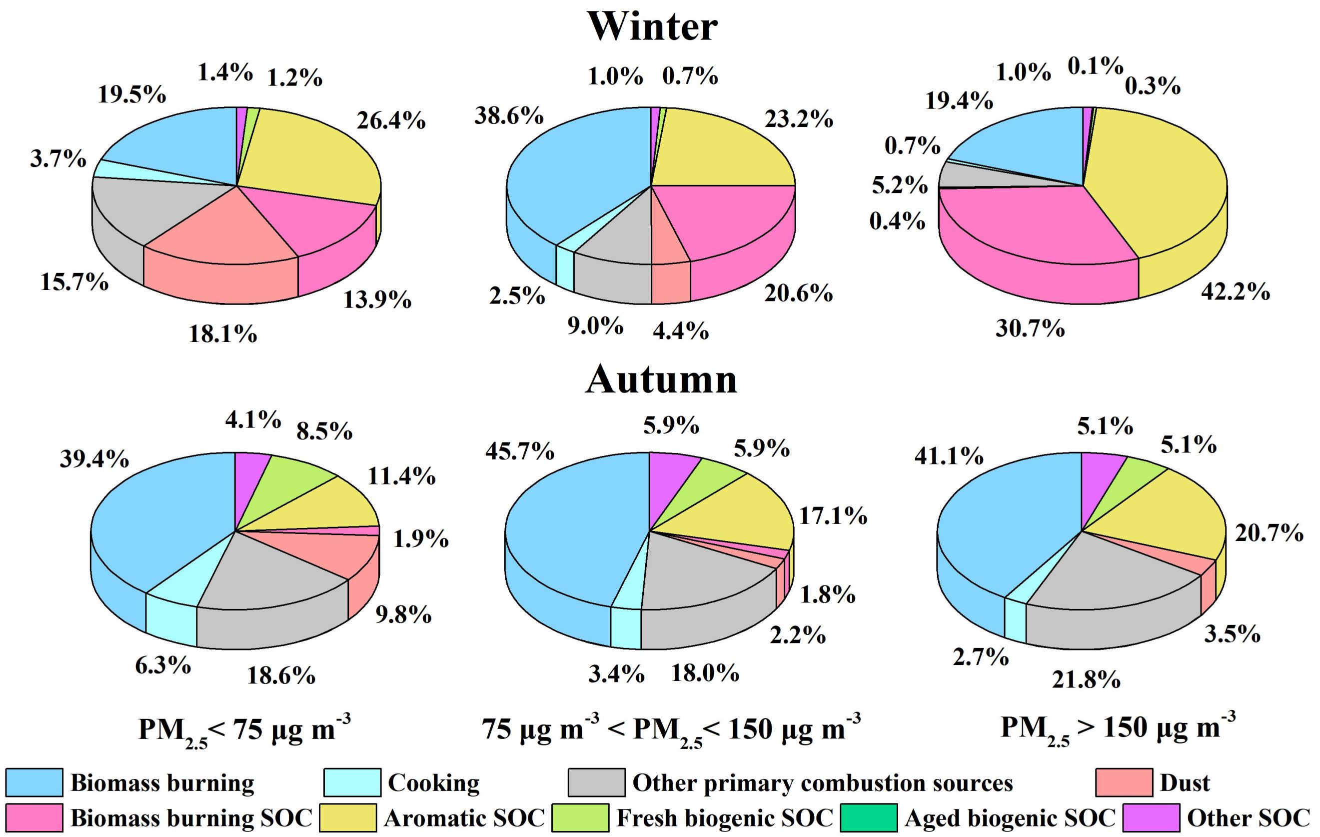


**Figure 6**. Source contributions to WSOC in $PM_{2.5}$ at different pollution levels in winter and autumn.

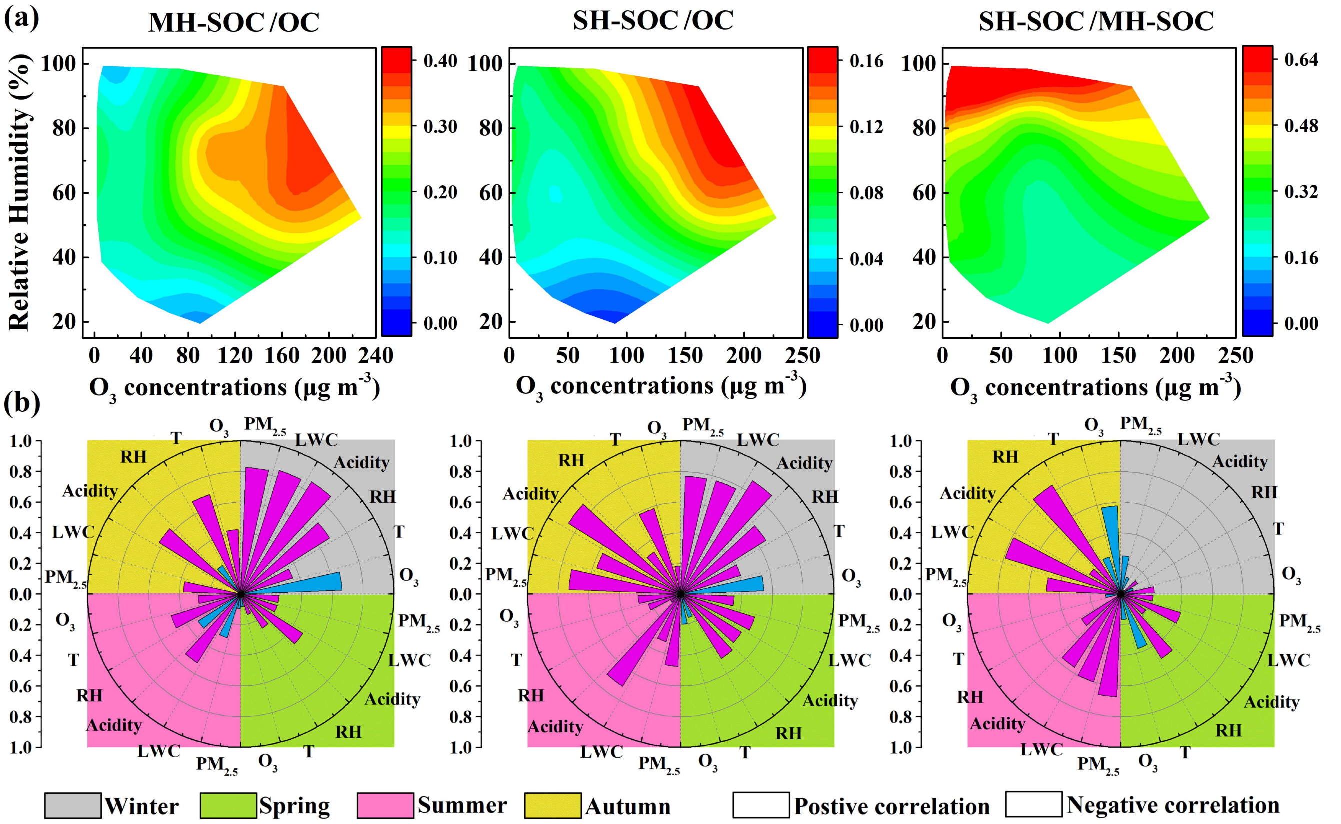


**Figure 7**. (a) The RH versus $O_3$ dependence of MH-SOC/OC, SH-SOC/OC and SH-SOC/MH-SOC ratios during the whole sampling period, and (b) correlation coefficients between the above ratios and meteorological parameters, $O_3$, $PM_{2.5}$, aerosol acidity and LWC in the four seasons.

**Figure: The logo of Copernicus Publications.**

# Characteristics, sources and formation mechanisms of water soluble organic compounds in atmospheric aerosols in downtown Beijing

## Supporting Information

Qing Yu[1,2], Jing Chen[1,2,3], Weihua Qin[1,2], Yuewei Sun[1,2], Siming Cheng[1,2], Ke Xin[1,2], Yuepeng Zhang[1,2]

1. State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, Beijing Normal University, Beijing 100875, China.

2. Center of Atmospheric Environmental Studies, Beijing Normal University, Beijing 100875, China.

3. State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Nanjing Hydraulic Research Institute, Nanjing 210029, China.

Correspondence to: Jing Chen (jingchen@bnu.edu.cn)

Tel.: +86-10-58801848;

Fax: +86-10-58801848.

## S1 Estimation for the sampling artifacts of organic aerosols.

Sampling of organic carbon is accompanied by both positive and negative artifacts. The positive artifact is due to the adsorption of gaseous organics to the sampling filter, and the negative artifact is caused by the evaporation of collected particulate organic carbon. To eliminate the positive artifact, a denuder can be placed upstream the sample filter to remove the gaseous organics by diffusion to the adsorbent surface (Cheng et al., 2009). The use of a denuder in the sampling system has been reported in previous studies (Eatough et al., 1993, 1999; Mader et al., 2001; Matsumoto et al., 2003; Viana et al., 2006; Cheng et al., 2009, 2010, 2012; Kristensen et al., 2016). The use of a denuder may induce a larger negative artifact, however, as the removal of gaseous organics can enhance the evaporation of particulate OC. Thus a backup filter should also be included in the sampling system (Cheng et al., 2009). Besides, the flow rate passing through the denuder was very low in most studies (Matsumoto et al., 2003; Viana et al., 2006; Cheng et al., 2009, 2010, 2012; Kristensen et al., 2016). This might be due to the significantly decreased removal efficiency of the denuder as the air flow rate increased (Cui et al., 1998; Ding et al., 2002). To collect enough samples for the accurate measurement of trace organic species, the flow rate of 1.05 $m^3$ $min^{-1}$ was chosen in this study. The air flow rate of about 1.05 $m^3$ $min^{-1}$ has been frequently used in the field sampling of organic aerosols (Kawamura et al., 2013; Verma et al., 2012, 2015; Li et al., 2018; Ma et al., 2018; Huang et al., 2020). At this flow rate, a denuder with a high removal efficiency is hardly commercially available.

We estimated the sampling artifact of OC based on the literature results. Firstly, different OC fractions which have distinct volatility show different adsorption behavior. Besides, the adsorption behavior of the same OC fraction also vary with meteorological conditions. Cheng et al. (2015) compared the concentrations of different OC fractions (OC1, OC2, OC3, OC4) on bare quartz filters with those on denuded quartz filters in four seasons of Beijing, and the results are summarized in Table S1. The contributions of different OC fractions measured in this study are also shown in Table S1.

In addition, the positive artifact of OC also depends on the sampling procedure. McDow (1986) systematically investigated the effect of sampling procedure on the OC measurement. The adsorption of organic vapors on bare quartz filters (Cpostive artifact) was a function of the sampling duration (t) multiplied by the face velocity (v) as follows:

$$C_{postive\ artifact} = \sum_i \rho_i \frac{1-e^{-\varepsilon_i vt}}{\varepsilon_i vt} \quad (1)$$

where the face velocity (v, $cm \cdot s^{-1}$) is the ratio of the flow rate ($cm^3 \cdot s^{-1}$) to the sampling area of the filter ($cm^2$), $\rho_i$ is the concentration of adsorptive vapor $i$ ($g \cdot cm^{-3}$), and $\varepsilon_i$ is a constant which can be defined as:

$$\varepsilon_i = \frac{1}{l}\left[\beta f(t) t_o \exp\left(\frac{Q_A}{RT}\right) + 1\right] \quad (2)$$

where $l$ is the effective filter thickness. The average thickness of the quartz filter used in this study was 463 µm. The other parameters are all constants.

Therefore, it can be calculated that $\varepsilon_i > 1/l > 20\ cm^{-1}$, and $1-e^{-\varepsilon vt} \approx 1$. Hence, the positive artifact ($C_{positive\ artifact}$) is inversely proportional to the product of the sampling duration and the face velocity (v×t). The face velocity of Cheng et al. (2015) was 9.8 $cm \cdot s^{-1}$, while that in our study was 47.3 $cm \cdot s^{-1}$. The sampling duration of Cheng et al. (2015) was 24 h, while that in our study was 12 h. That is to say, the positive artifact of Cheng et al. (2015) was about 2.4 times higher than that in our study.

Based on the literature results and taking into account the above factors (seasons, OC fractions, sampling procedure), the contribution of positive artifact to the measured OC was estimated to be 2.3 %, 1.4 %, 9.9 %, and 2.2 % during the sampling periods in winter, spring, summer and autumn respectively in this study, which is roughly acceptable.

To further estimate the impact of gas-particle partitioning and potential reactions occurring on filters, we overlapped two quartz filters and took samples at a flow rate of 1.05 $m^3 \cdot min^{-1}$ for a duration of 12 h. The organic tracers selected in this study were measured. The organic tracers on the backup filters typically originate from three sources: (1) adsorption of the gas-phase organic species; (2) adsorption of the semi-volatile species evaporated from the front filter; (3) secondary formation from the adsorbed organic vapors on the backup filter. Except for *cis*-pinonic acid, the tracer concentrations on the backup filter were all less than 5 % of those on the front filters, while the concentration of *cis*-pinonic acid on the backup filter was 21.6 % of that on the front filter. This result suggested that the sampling procedure in this study might bring some uncertainties for the measurement of *cis*-pinonic acid, and the sampling artifact was not significant for the other organic tracers.

## S2 Detailed information for the chemical analysis.

### S2-1 Chemical analysis of water-soluble ions and water-soluble organic carbon (WSOC)

To analyze the concentrations of water-soluble ions and water-soluble organic carbon (WSOC), a punch of each sampled filter was cut into pieces and extracted with 40 mL ultrapure water (>18.2 MΩ) for 30 min, then passed through a 0.45 μm PTFE filter. Five cations ($Na^+$, $NH_4^+$, $K^+$, $Mg^{2+}$, $Ca^{2+}$) and four anions ($Cl^-$, $NO_3^-$, $SO_4^{2-}$, $C_2O_4^{2-}$) were measured using the ion chromatography (Dionex 600), with the methanesulfonic acid (MSA) solution as cationic eluent and the potassium hydroxide (KOH) solution as anionic eluent. The concentration of WSOC was measured by a TOC analyzer (Shimazdu TOC-L CPN). The standard solution of total carbon (TC) was prepared by potassium acid phthalate ($C_8H_5KO_4$), and that of inorganic carbon (IC) was made by sodium carbonate ($Na_2CO_3$) and sodium bicarbonate ($NaHCO_3$). Total organic carbon (TOC) was calculated as total carbon minus inorganic carbon.

### S2-2 The parameter settings of GC/MS/MS for analyzing organic tracers

The derivatives were immediately analyzed by a Shimadzu TQ8040 gas chromatography triple quadrupole mass spectrometry (GC/MS/MS). A JA-5MS capillary column (30 m × 0.25 mm i.d., film thickness 0.25 μm) was used as the GC column and helium was used as the carrier gas (1.0 mL $min^{-1}$). The injector was set splitless at a temperature of 290 °C. The programmed oven temperature increased from 70 °C to 150 °C at 2 °C $min^{-1}$, then to 200 °C at 5 °C $min^{-1}$, then to 300 °C at 25 °C $min^{-1}$, and stayed at 300 °C for 6 min. The MS was operated in EI mode at 70 eV with a scan range of 50-650 amu.

## S3 The detailed procedures of the PMF source apportionment

### 1. Uncertainties of the input data

According to the User Guide of PMF5.0 (Norris et al., 2014), the uncertainties of the target species can be calculated as follow:

$$\mathrm{Unc} = 5/6 \times \mathrm{MDL}\ (\mathrm{c} \leq \mathrm{MDL}) \tag{3}$$

$$Unc = \sqrt{(P \times c)^2 + (0.5 \times MDL)^2}\ \ (c > MDL) \tag{4}$$

where *Unc* is the data uncertainty, c is the concentration of the target species, MDL is the method detection limit, and P is the error fraction. Since the User Guide did not give the calculation method for the error fraction (P), we estimated the P values referring to the measured relative standard deviations (RSD) of the target species. The RSD values were calculated by measuring six identical portions of an ambient sample. P was set as 10 % when RSD < 10 %, and 15 % or 20 % when RSD > 10 %.

### 2. Selection of base solutions

The chemical components input into the PMF model were selected based on our understanding of the possible WSOC sources (Norris et al., 2014). Interpretability was usually considered to be the most important factor for selecting the optimum PMF solution (Shrivastava et al., 2007; Huang et al., 2014). The interpretable solutions are those which group tracers from different sources into distinct factors, while those grouping tracers from multiple sources into the same factor, distributing tracers for one source across multiple factors, or including factors with no distinct grouping of species are judged less interpretable (Shrivastava et al., 2007; Sowlat et al., 2016). In some previous literature, the optimal solution was defined as that with the maximum number of factors which had distinctive groupings of species, and explained at least 90 % of the total variable (Shrivastava et al., 2007). In this study, PMF was run repeatedly by changing the number of factors and the start seed numbers. The base solution was selected based on: (1) the interpretability of the derived factor profiles and the temporal variations of source contributions; (2) the reconstruction of the total variable and $R^2$ of input organic tracers ($R^2$>0.90); (3) the scaled residuals of the input species.

As presented in Figure S1, the 7-factor solution separated cholesterol (the tracer for cooking) into multiple sources. It was difficult to explain why cholesterol appeared in the factor profiles of biomass burning, dust and fresh biogenic SOC. Besides, this solution led to poor fits for cholesterol

($R^2$ = 0.28) and *cis*-pinonic acid ($R^2$ = 0.32), which were the key tracers selected in this study. Therefore, the 7-factor solution was not selected. As shown in Figure S2, the 8-factor solution also distributed cholesterol into multiple factors. This solution also resulted in a poor fit ($R^2$ = 0.28) for cholesterol. Therefore, the 8-factor solution was not chosen in this study. As shown in Figures S5-7, the solutions with 4 to 6 factors all showed poor interpretability for the derived factor profiles and poor fits for the key organic tracers. The 10-factor solution involved a factor without any tracer of high loading to indicate a specific source, thus could not be explained. By comparing the results with different factor numbers, the solution with 9 factors (Figure S3) was thought to be the most interpretable one.

## 3. Diagnostics for the base model run

The selected 9-factor solution was converged, and Q(Robust) was similar to Q(True). As shown in Figure S4, most of the input species showed normally distributed residuals between -2.0 and +2.0, indicating that these species were well modeled. The $R^2$ of WSOC and HULIS were 0.94 and 0.93, respectively, and the $R^2$ for all the organic tracers were higher than 0.96, again suggesting that these species were well modeled.

## 4. Error estimation

The selected base solution was subjected to displacement (DISP) and bootstrap (BS) tests for error estimation. For the DISP test, the percent change in Q (%dQ) was less than 0.1 %, indicating that this solution was the global minimum (Paatero et al., 2014). No factor swapped for any value of dQmax, indicating little rotational ambiguity in this solution (Paatero et al., 2014). For the BS test, the factor of “cooking” was mapped 79 % of the runs, the factor of “other primary combustion sources” was mapped 69 % of the runs, while other factors were mapped more than 91 % of the runs. The BS results indicated some uncertainties for the factors of cooking and other primary combustion sources, while the other factors were relatively stable.

Brown et al. (2015) indicated that the unstable PMF solution might be due to too many factors involved. To investigate the effect of factor number on the stability of solutions, the BS results for solutions with different factor numbers were compared and shown in Table S3. As shown in Table S3, reducing the number of factors did not significantly increase the successful rates of BS mapping, but decreased the interpretability of the derived factor profiles. As recommended by the previous studies (Norris et al., 2014; Paatero et al., 2014), some constraints can be defined based on the priori information of the sources to reduce the variability of the solution.

## 5. Constrained model run

Bozzetti et al. (2017) exploited the markers' source specificity to set constraints for the profiles, so as to solve the problem of large mixtures of PMF factors associated with contributions of markers from different sources. They treated the contribution of unrelated source-specific markers as zero for each source, while non-source-specific variables were freely apportioned by the PMF algorithm. In addition, they set constraints for primary markers and combustion-related markers that can be seen as negligible in the secondary factors.

In the constrained model run, we set the constraints similar to those of Bozzetti et al. (2017), with a slight difference that we set the constraints by "soft pulling" so as to obtain a stable solution with a minimal change in the Q-value (dQ). The constraints were set as follows: (1) Levoglucosan was pulled up maximally with a limit of 0.25 % dQ for the factor of "primary biomass burning"; (2) Cholesterol was pulled up maximally with a limit of 0.50 % dQ for the factor of "cooking"; (3) Sulfate, *cis*-pinonic acid and 2-methylerythritol were pulled down maximally with limits of 0.25 % dQ for the factor of "other primary combustion sources"; (4) Phthalic acid was pulled up maximally with a limit of 0.25 % dQ for the factor of "aromatic SOA". The dQ(Robust) for all the constraints were 0.93 % in the final constrained model run, which was acceptable (below 1 %) as recommended by the PMF user guide (Norris et al., 2014). As shown in Table S4, all the factors were mapped more than 94 % of the runs, suggesting that this solution was stable. Thus the constrained 9-factor solution was chosen as the final solution.

## 6. Factor identification

The source profiles of the final solution are shown in Figure 4. Factor 1 showed high levels of levoglucosan and EC, thus was interpreted as the direct emissions from biomass burning. Factor 2 exhibited a high level of cholesterol, thus was regarded as cooking. Factor 3 showed a large fraction of EC that could not be explained by the direct emissions from biomass burning, suggesting that it was the direct emissions from other combustion sources, such as coal combustion, traffic emissions and waste burning. Factor 4 was featured by high loadings of $Mg^{2+}$ and $Ca^{2+}$, thus was considered as dust. No significant EC but high fractions of 4-methyl-5-nitrocatechol and phthalic acid were found in Factor 5 and Factor 6, respectively, which were regarded as SOC from biomass burning (biomass burning SOC) and aromatic precursors (aromatic SOC), respectively. Factor 7 exhibited a high level of *cis*-pinonic acid, thus was explained as fresh biogenic SOC. Factor 8 was characterized by high fractions of 2-methylerythritol and 3-hydroxyglutaric acid, which are the end products from isoprene and monoterpenes respectively, thus was identified as aged biogenic SOC. Note that *cis*-pinonic acid and 3-hydroxyglutaric acid were not grouped into the same factor though they are both SOA tracers of monoterpenes, owing to their different oxidation degree as discussed above. Factor 9 covered the secondary components (such as $SO_4^{2-}$, $NO_3^-$, $NH_4^+$ and $C_2O_4^{2-}$) that can not be well explained by the identified sources above, thus was considered to be SOC from other sources.

## 6-1 The reason why we separated the primary and secondary biomass burning SOA.

Levoglucosan correlated strongly with 4-methyl-5-nitrocatechol ($r=0.87$, $p<0.01$) in this study. It seems that levoglucosan and 4-methyl-5-nitrocatechol should be distributed into the same factor. Nevertheless, in fact, even though we reduced the factor number from nine to five, levoglucosan and 4-methyl-5-nitrocatechol could not be merged into one factor (Figure S1, 2, 3, 5, 6). When the factor number decreased to four, levoglucosan and 4-methyl-5-nitrocatechol were merged into one factor (Figure S8). However, this solution was less interpretable, and resulted in poorer fits for most of the input species (cholesterol: $R^2=0.17$; *cis*-pinonic acid: $R^2=0.25$; $Ca^{2+}$: $R^2=0.67$; $Mg^{2+}$: $R^2=0.73$; $NO_3^-$: $R^2=0.75$; etc). Furthermore, the slope of the fitting equation for the observed and predicted values of 4-methyl-5-nitrocatechol was even only 0.31, that is, the high values of 4-methyl-5-nitrocatechol in winter were not reproduced by the 4-factor solution. Hence, the 4-factor solution was also excluded in this study.

It was indeed interesting that levoglucosan and 4-methyl-5-nitrocatechol were not distributed in the same factor, though they showed strong correlation with each other. We attempted to explain this phenomenon as follows. The ratio of 4-methyl-5-nitrocatechol to levoglucosan showed significantly higher values ($p<0.01$) in winter ($0.071 \pm 0.029$) than in other seasons ($0.010 \pm 0.009$), which implied different types of biomass burning sources (primary and secondary). If they were merged into one factor, the ratio of 4-methyl-5-nitrocatechol to levoglucosan was regarded to be constant throughout the year, which was not the truth. According to the uncertainty estimation method for the input species (Equation 2), the data with lower concentrations usually have lower uncertainties, thus may have a larger impact on the Q value. Taking the 4-factor solution (Figure S7) as an example, when these two tracers were merged into one factor, to minimize the Q value, the algorithm in the PMF model tended to assign a low value for the ratio of 4-methyl-5-nitrocatechol to levoglucosan in the factor profile of biomass burning (i.e. 0.024 in Factor 4). As shown in Figure S8, this ratio (orange line) was closer to the regression slope in other seasons (0.017), but much lower than that in winter (0.096). As a consequence, the high concentration of 4-methyl-5-nitrocatechol in winter could not be reproduced at all by such PMF solution. In conclusion, the solution which merged these two tracers into the same factor might bring about large uncertainties, and fail to reproduce the peak values of 4-methyl-5-nitrocatechol over the study period in winter.

Fresh biomass burning emissions show a high fraction of anhydrosugar, such as levoglucosan. The relative intensity of anhydrosugar decreased due to the degradation or oxidation reactions (Gilardon et al., 2016; Sengupta et al., 2020). Chamber studies indicated that substantial amounts of nitrogen-containing organic compounds, such as nitrophenols and methyl-nitrocatechols, were generated during aging (Bertrand et al., 2018; Hartikainen et al., 2018). 4-Methyl-5-nitrocatechol was recommended to be the secondary tracer for aged biomass burning SOA (Bertrand et al., 2018). In the previous studies (Gilardon et al., 2016; Zhou et al., 2017; Srivastava et al., 2018), the biomass burning source was also separated into primary and secondary fractions with the PMF model. In Gilardon et al. (2016), the factor profile of primary biomass burning was featured by a high loading of anhydrosugar (signal at m/z=60) but a low level of the aged OA (signal at m/z=44), while the factor profile of biomass burning SOA was featured by a high fraction of the aged OA (signal at m/z=44) but a low fraction of anhydrosugar (signal at m/z=60) (Gilardon et al., 2016). Srivastava et al. (2018) also used levoglucosan and 4-methyl-5-nitrocatechol as tracers to differentiate the primary biomass burning and biomass burning SOA using the PMF model (92 samples, which was less than that in our study). In this study, as shown in Figure 4, Factor 1 had high fractions of levoglucosan and EC, but a low fraction of 4-methyl-5-nitrocatechol, thus was considered as the direct emission from biomass burning. The concentration ratio of levoglucosan to WSOC in this factor was 0.085 $\mu g \cdot \mu g^{-1}$, similar to that measured in the primary combustion of crop straws (0.097 $\mu g \cdot \mu g^{-1}$), wood (0.081 $\mu g \cdot \mu g^{-1}$) and leaves (0.095 $\mu g \cdot \mu g^{-1}$) in North China (Yan et al., 2018). Factor 5 showed a high level of 4-methyl-5-nitrocatechol, but low loadings of EC and levoglucosan, thus was identified as biomass burning SOA.

### 6-2 The interpretation for Factor 7, Factor 8, and Factor 9.

As presented in Figures S1-3 and Figures S5-7, even if we reduced the factor number from nine to four, 2-methylerythritol, 3-hydroxyglutaric acid and *cis*-pinonic acid could not be merged into the same factor. Large fractions of 3-hydroxyglutaric acid and 2-methylerythritol were usually grouped into one factor, since they strongly correlated with each other ($r=0.94$, $p<0.01$). *Cis*-pinonic acid could not be distributed in this factor since it correlated less strongly with 2-methylerythritol ($r=0.51$, $p<0.01$) and 3-hydroxyglutaric acid ($r=0.58$, $p<0.01$). As stated in Section 3.2, *cis*-pinonic acid is a lower-generation oxidative product from monoterpenes, while 2-methylerythritol and 3-hydroxyglutaric acid are more aged products from isoprene and monoterpenes, respectively (Kourtchev et al., 2009). Hence, Factor 7 with a high level of cis-pinonic acid was interpreted as the fresh biogenic SOC, and Factor 8 with high loadings of 2-methylerythritol and 3-hydroxyglutaric acid was interpreted as aged biogenic SOC. As shown in Figure 5, the seasonal variation of their source contributions also supported this interpretation.

Since the major fraction of 3-hydroxyglutaric acid was distributed in Factor 8, it was not proper to interpret Factor 9 as monoterpene SOC. In fact, as shown in Figure S1-3 and Figure S5-7, a minor fraction of 3-hydroxyglutaric acid was always distributed in factors other than the biogenic SOC. It was more interpretable when this minor fraction of 3-hydroxyglutaric acid was distributed in the same factor together with $SO_4^{2-}$, $NO_3^-$, $NH_4^+$ and $C_2O_4^{2-}$. In this case, Factor 9 of the selected 9-factor solution could be interpreted as a mixed secondary source and explain the secondary species that were not well fitted by other identified secondary sources. Similar factor profile has also been resolved in the literature, and was usually interpreted as the “inorganic-rich SOA” (Huang et al., 2014).

### 6-3 The interpretation for Factor 3 (Other primary combustion sources).

As shown in Figure 4, for the constrained 9-factor solution, Factor 3 showed a significant level of EC that could not be explained by direct emissions of biomass burning, implying that it could be associated with the primary emissions from other combustion sources, such as coal combustion, traffic emissions, and waste burning, etc. Indeed, a minor fraction of $SO_4^{2-}$ (20.8 %), $NH_4^+$ (19.3 %) and phthalic acid (20.0 %) were also distributed in Factor 3. However, in fact, previous studies have indicated that $SO_4^{2-}$, $NO_3^-$, $NH_4^+$ and phthalic acid can also be directly emitted from coal combustion

(Zhang et al., 2008; Dai et al., 2019) and traffic emissions (Al-Naiema and Stone, 2017; Hao et al., 2019). Dai et al. (2019) suggested that primary $SO_4^{2-}$ accounted for 38.9 % and 16.9 % to the total $SO_4^{2-}$ in $PM_{2.5}$ in the heating and non-heating seasons respectively. Accordingly, such distribution of $SO_4^{2-}$ (20.8 %) in Factor 3 was acceptable. Similar loadings of $SO_4^{2-}$ and $NH_4^+$ were also found in the factor profile of coal combustion in the previous source apportionment study (Huang et al., 2014). To sum up, it was reasonable that a minor fraction of $SO_4^{2-}$, $NH_4^+$ and phthalic acid presented in the factor profile of Factor 3, i.e., other primary combustion sources.

**Table S1** The ratio of the OC concentrations on the bare quartz filters to those on the denuded quartz filters in Cheng et al. (2015), as well as the contribution of different OC fractions measured in this study.

| | The ratio of OC on bare quartz filters to denuded quartz filters (Cheng et al., 2015) | | | | The contribution of different OC fractions measured in this study | | | |
|---|---|---|---|---|---|---|---|---|
| | OC1 | OC2 | OC3 | OC4 | OC1 | OC2 | OC3 | OC4 |
| Winter | 1.27 | 1.03 | 1.02 | 1.05 | 10.8 % | 19.6 % | 24.7 % | 44.8 % |
| Spring | 2.05 | 1.05 | 1.00 | 1.01 | 3.9 % | 27.2 % | 43.1 % | 25.7 % |
| Summer | 2.45 | 1.60 | 1.17 | 1.08 | 4.4 % | 37.6 % | 36.0 % | 22.0 % |
| Autumn | 2.08 | 1.05 | 0.99 | 1.01 | 7.9 % | 26.5 % | 40.2 % | 25.3 % |

**Table S2** The detailed information of the authentic standards used in this study.

| **Authentic standard** | **Molecular formula** | **CAS number** | **Company** | **Purity** |
|---|---|---|---|---|
| Levoglucosan | $C_6H_{10}O_5$ | 498-07-7 | Sigma-Aldrich | 99% |
| Cholesterol | $C_{27}H_{46}O$ | 57-88-5 | Sigma-Aldrich | 93% |
| Phthalic acid | $C_8H_6O_4$ | 88-99-3 | Sigma-Aldrich | 99% |
| 4-Methyl-5-nitrocatechol | $C_7H_7NO_4$ | 68906-21-8 | Toronto Research Chemicals | 98% |
| 2-Methylerythritol | $C_5H_{12}O_4$ | 58698-37-6 | Sigma-Aldrich | 90% |
| 3-Hydroxyglutaric acid | $C_5H_8O_5$ | 638-18-6 | Sigma-Aldrich | 95% |
| cis-Pinonic acid | $C_{10}H_{16}O_3$ | 61826-55-9 | Sigma-Aldrich | 98% |

**Table S3** The successful rates of BS mapping for the solutions with different numbers of factors. The values no more than 85 % were shown in bold.

| BS mapping | 3-factor solution | 4-factor solution | 5-factor solution | 6-factor solution | 7-factor solution | 8-factor solution | 9-factor solution |
|---|---|---|---|---|---|---|---|
| Factor 1 | 93 % | 93 % | 95 % | 95 % | **82 %** | 99 % | 99 % |
| Factor 2 | **83 %** | **85 %** | 99 % | 98 % | 88 % | 95 % | **79 %** |
| Factor 3 | 100 % | 95 % | **78 %** | 99 % | **74 %** | 100 % | 100 % |
| Factor 4 | | 87 % | **74 %** | **75 %** | **79 %** | 98 % | **69 %** |
| Factor 5 | | | 94 % | 91 % | 92 % | 92 % | 96 % |
| Factor 6 | | | | **81 %** | 99 % | **67 %** | 96 % |
| Factor 7 | | | | | 98 % | 100 % | 100 % |
| Factor 8 | | | | | | 98 % | 91 % |
| Factor 9 | | | | | | | 99 % |

**Table S4** The successful rates of BS mapping for the constrained 9-factor solution.

| BS Mapping: | Base Factor 1 | Base Factor 2 | Base Factor 3 | Base Factor 4 | Base Factor 5 | Base Factor 6 | Base Factor 7 | Base Factor 8 | Base Factor 9 | Unmapped |
|---|---|---|---|---|---|---|---|---|---|---|
| Boot Factor 1 | 97 % | 0 | 0 | 0 | 3 % | 0 | 0 | 0 | 0 | 0 |
| Boot Factor 2 | 1 % | 95 % | 0 | 0 | 0 | 0 | 0 | 1 % | 3 % | 0 |
| Boot Factor 3 | 0 | 0 | 100 % | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Boot Factor 4 | 4 % | 0 | 0 | 94 % | 0 | 1 % | 0 | 1 % | 0 | 0 |
| Boot Factor 5 | 3 % | 0 | 0 | 0 | 97 % | 0 | 0 | 0 | 0 | 0 |
| Boot Factor 6 | 0 | 0 | 0 | 1 % | 0 | 99 % | 0 | 0 | 0 | 0 |
| Boot Factor 7 | 0 | 0 | 0 | 0 | 0 | 0 | 100 % | 0 | 0 | 0 |
| Boot Factor 8 | 0 | 0 | 0 | 0 | 0 | 1 % | 0 | 97 % | 2 % | 0 |
| Boot Factor 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 100 % | 0 |

**Table S5** WSOC, OC concentrations and WSOC/OC ratios in $PM_{2.5}$ Beijing in recent years.

| Sampling period | WSOC (μg $m^{-3}$) | OC (μg $m^{-3}$) | WSOC/OC | OC method[*] | Reference |
|---|---|---|---|---|---|
| 2017 Winter | 11.71 ± 13.92 | 20.56 ± 21.89 | 0.52 ± 0.08 | IMPROVE (TOR) | This study |
| 2017 Spring | 4.40 ± 2.34 | 8.70 ± 3.12 | 0.49 ± 0.12 | | |
| 2017 Summer | 4.66 ± 2.46 | 7.75 ± 3.91 | 0.60 ± 0.11 | | |
| 2017 Autumn | 4.77 ± 2.83 | 9.71 ± 3.69 | 0.46 ± 0.12 | | |
| 2016 Spring | 3.8 ± 3.8[a] | 7.9 ± 7.4 | 0.48 | Optical transmittance | Yang et al. (2019) |
| 2016 Summer | 2.3 ± 0.9[a] | 3.4 ± 1.0 | 0.68 | | |
| 2016 Autumn | 4.8 ± 4.1[a] | 7.9 ± 7.6 | 0.61 | | |
| 2016 Winter | 10.0 ± 10.1[a] | 20.7 ± 16.2 | 0.49 | | |
| 2013 Autumn | Not mentioned | 11.4 | 0.70 ± 0.27 | IMPROVE (TOR) | Zhao et al. (2018) |
| 2013 Winter | | 19.4 | 0.49 ± 0.11 | | |
| 2014 Spring | | 8.53 | 0.56 ± 0.07 | | |
| 2014 Summer | | 5.29 | 0.58 ± 0.10 | | |
| 2013-2014 Winter | 12.8 | 29.1 | 0.46 | IMPROVE-A | Huang et al. (2020) |
| 2013 Winter | 10.8 ± 3.1 | 32.9 ± 16.8 | 0.39 ± 0.16 | IMPROVE-A (TOT) | Yan et al. (2015) |
| 2013 Summer | 6.4 ± 2.2 | 9.7 ± 2.9 | 0.66 ± 0.06 | | |
| 2012 Summer | 4.4 ± 3.6 | 8.5 ± 5.2 | 0.52 | IMPROVE-A (TOR) | Li et al. (2019a) |
| 2012 Autumn | 5.2 ± 4.0 | 10.3 ± 7.4 | 0.50 | | |
| 2013 Winter | 10.3 ± 9.8 | 28.9 ± 22.0 | 0.36 | | |
| 2013 Spring | 5.9 ± 4.9 | 14.6 ± 10.8 | 0.40 | | |
| 2011-2012 Winter | Not mentioned | Not mentioned | 0.36 ± 0.05[b]<br>0.44 ± 0.05[c]<br>0.47 ± 0.05[d] | IMPROVE-A (TOT) | Cheng et al. (2015) |
| 2011 Summer | 4.48 | 13.55 | 0.33 | IMPROVE (TOR) | Xiang et al. (2017) |
| 2011 Autumn | 5.82 | 25.42 | 0.25 | | |
| 2011 Winter | 5.53 | 28.16 | 0.20 | | |
| 2012 Spring | 3.90 | 16.57 | 0.27 | | |
| 2012 Summer | 5.81 | 16.54 | 0.34 | | |
| 2011 Summer | 7.8 ± 4.4 | 12.0 ± 6.3 | 0.65 | IMPROVE-A (TOT) | Cheng et al. (2013) |
| 2011-2012 Winter | 11.2 ± 8.2 | 24.6 ± 17.1 | 0.46 | | |
| 2010 Fall | 8.6 ± 6.4 | 20.4 ± 15.4 | 0.42 | IMPROVE (TOT) | Du et al. (2014) |
| 2010 Winter | 8.0 ± 6.7 | 20.6 ± 16.1 | 0.39 | | |
| 2011 Spring | 4.7 ± 3.1 | 10.2 ± 6.8 | 0.46 | | |
| 2011 Summer | 6.7 ± 4.4 | 10.7 ± 6.2 | 0.61 | | |
| 2011 Fall | 8.6 ± 6.1 | 19.7 ± 15.4 | 0.44 | | |
| 2009 Spring | 6.7 ± 1.8 | 13.7 ± 4.4 | 0.49 | Not mentioned | Tao et al.(2016) |
| 2009 Summer | 3.2 ± 1.1 | 11.1 ± 1.8 | 0.29 | | |
| 2009 Autumn | 7.7 ± 5.0 | 17.8 ± 5.6 | 0.43 | | |

| | | | | | |
|---|---|---|---|---|---|
| 2010 Winter | 7.7 ± 3.6 | 24.9 ± 15.6 | 0.31 | | |
| 2009 Winter | 7.28 | 27.7 ± 15.4[e] | 0.26 | IMPROVE-A (TOR) | Cheng et al. (2011) |
| | | 30.9 ± 16.3[f] | 0.24 | | |
| | | 32.6 ± 18.6[e] | 0.22 | IMPROVE-A (TOT) | |
| | | 36.1 ± 19.5[f] | 0.20 | | |
| 2009 Summer | 3.36 | 7.2 ± 2.4[e] | 0.48 | IMPROVE-A (TOR) | |
| | | 9.4 ± 2.7[f] | 0.36 | | |
| | | 8.8 ± 3.3[e] | 0.38 | IMPROVE-A (TOT) | |
| | | 11.4 ± 3.6[f] | 0.30 | | |

[*] The thermal-optical reflectance (TOR) method and thermal-optical transmittance (TOT) method are two different charring correction methods to determine the split of OC and EC. The transmittance-defined EC is the carbon measured after the filter transmittance returns to its initial value in the $He/O_2$ atmosphere, whereas the reflectance-defined EC is the carbon measured after the filter reflectance returns to its initial value (Cheng et al., 2011).

[a] In Yang et al. (2019), the concentrations of WSOC were measured by UV/VIS absorption (at wavelengths of about 250 nm)

[b,c,d] In Cheng et al. (2015), "b" refers to the constructed $PM_{2.5}$ below 30 µg $m^{-3}$, "c" between 30 µg $m^{-3}$ and 90 µg $m^{-3}$, and "d" above 90 µg $m^{-3}$.

[e,f] In Cheng et al. (2011), "e" was measured using the denuded quartz filter and "f" was measured using the un-denuded (bare) quartz filter.

**Table S6** The mean values and standard deviations (SD) of planetary boundary layer (PBL) heights and wind speeds in Beijing during the study periods in four seasons.

| | | **Winter** | | **Spring** | | **Summer** | | **Autumn** | |
|---|---|---|---|---|---|---|---|---|---|
| | | **Day** | **Night** | **Day** | **Night** | **Day** | **Night** | **Day** | **Night** |
| **PBL (m)** | **Mean** | 309.7 | 283.2 | 1149.4 | 303.9 | 871.6 | 210.5 | 423.2 | 90.9 |
| | **SD** | 236.7 | 259.1 | 841.1 | 363.2 | 405.6 | 164.3 | 278.9 | 90.4 |
| **WS (m $s^{-1}$)** | **Mean** | 1.14 | 0.83 | 1.75 | 0.83 | 1.03 | 0.63 | 0.72 | 0.39 |
| | **SD** | 0.79 | 0.61 | 0.71 | 0.75 | 0.68 | 0.44 | 0.62 | 0.25 |

**Table S7** Spearman correlation coefficients between SOA tracers as well as the WSOC/OC ratio and meteorological parameters, $O_3$, aerosol acidity, and aerosol liquid water content (LWC) in Beijing during the sampling periods in four seasons.

| Compounds | Sampling periods | T | RH | WS | SR | $O_3$ | Acidity | LWC |
|---|---|---|---|---|---|---|---|---|
| WSOC/OC | Whole | 0.296$^{**}$ | 0.290$^{**}$ | -0.142 | 0.021 | 0.152 | 0.444$^{**}$ | 0.387$^{**}$ |
| | January | 0.495$^{**}$ | 0.550$^{**}$ | -0.298 | -0.508 | -0.267 | 0.783$^{**}$ | 0.684$^{**}$ |
| | April | -0.149 | 0.640$^{**}$ | -0.317 | -0.762$^{**}$ | -0.424$^{*}$ | 0.580$^{**}$ | 0.680$^{**}$ |
| | July | 0.273 | -0.066 | -0.053 | -0.082 | 0.081 | 0.499$^{**}$ | 0.063 |
| | October | 0.469$^{**}$ | 0.365$^{*}$ | -0.074 | -0.068 | -0.073 | 0.803$^{**}$ | 0.619$^{**}$ |
| 4-Methyl-5-nitrocatechol | Whole | -0.780$^{**}$ | 0.051 | -0.229$^{*}$ | -0.691$^{**}$ | -0.841$^{**}$ | -0.247$^{*}$ | 0.227$^{*}$ |
| | January | 0.201 | 0.688$^{**}$ | -0.712$^{**}$ | -0.618$^{*}$ | -0.766$^{**}$ | 0.730$^{**}$ | 0.875$^{**}$ |
| | April | -0.516$^{**}$ | 0.880$^{**}$ | -0.553$^{**}$ | -0.932$^{**}$ | -0.685$^{**}$ | 0.482$^{**}$ | 0.816$^{**}$ |
| | July | - | - | - | - | - | - | - |
| | October | -0.259 | 0.211 | -0.281 | -0.044 | -0.504$^{**}$ | -0.037 | 0.186 |
| Phthalic acid | Whole | 0.048 | 0.212$^{*}$ | -0.198$^{*}$ | -0.137 | -0.146 | 0.181 | 0.452$^{**}$ |
| | January | 0.389$^{*}$ | 0.657$^{**}$ | -0.621$^{**}$ | -0.648$^{*}$ | -0.645$^{**}$ | 0.815$^{**}$ | 0.869$^{**}$ |
| | April | -0.158 | 0.681$^{**}$ | -0.300 | -0.865$^{**}$ | -0.486$^{**}$ | 0.660$^{**}$ | 0.768$^{**}$ |
| | July | 0.714$^{**}$ | -0.597$^{*}$ | 0.308 | 0.186 | 0.653$^{**}$ | 0.547$^{**}$ | -0.321 |
| | October | 0.549$^{**}$ | 0.325 | -0.082 | -0.103 | -0.013 | 0.803$^{**}$ | 0.579$^{**}$ |
| 2-Methylerythritol | Whole | 0.595$^{**}$ | 0.545$^{**}$ | -0.343$^{**}$ | 0.211 | 0.333$^{**}$ | 0.786$^{**}$ | 0.562$^{**}$ |
| | January | 0.304 | 0.686$^{**}$ | -0.624$^{**}$ | -0.530 | -0.679$^{**}$ | 0.789$^{**}$ | 0.868$^{**}$ |
| | April | -0.079 | 0.563$^{**}$ | -0.440$^{*}$ | -0.524$^{*}$ | -0.535$^{**}$ | 0.694$^{**}$ | 0.545$^{**}$ |
| | July | 0.657$^{**}$ | -0.399$^{*}$ | 0.249 | 0.343 | 0.563$^{**}$ | 0.759$^{**}$ | -0.074 |
| | October | 0.255 | 0.371$^{*}$ | -0.243 | 0.109 | -0.253 | 0.627$^{**}$ | 0.520$^{**}$ |
| 3-Hydroxyglutaric acid | Whole | 0.626$^{**}$ | 0.534$^{**}$ | -0.299$^{**}$ | 0.212 | 0.372$^{**}$ | 0.822$^{**}$ | 0.576$^{**}$ |
| | January | 0.495$^{**}$ | 0.649$^{**}$ | -0.640$^{**}$ | -0.653$^{*}$ | -0.615$^{**}$ | 0.893$^{**}$ | 0.844$^{**}$ |
| | April | -0.146 | 0.672$^{**}$ | -0.338 | -0.715$^{**}$ | -0.429$^{*}$ | 0.699$^{**}$ | 0.668$^{**}$ |
| | July | 0.533$^{**}$ | -0.293 | 0.059 | 0.186 | 0.536$^{**}$ | 0.718$^{**}$ | 0.108 |
| | October | 0.482$^{**}$ | 0.340 | -0.002 | 0.018 | 0.012 | 0.832$^{**}$ | 0.613$^{**}$ |
| *cis*-Pinonic acid | Whole | 0.591$^{**}$ | 0.032 | -0.092 | 0.178 | 0.348$^{**}$ | 0.586$^{**}$ | 0.111 |
| | January | 0.368$^{*}$ | 0.577$^{**}$ | -0.650$^{**}$ | -0.473 | -0.660$^{**}$ | 0.728$^{**}$ | 0.790$^{**}$ |
| | April | 0.263 | 0.232 | -0.128 | -0.785$^{**}$ | -0.141 | 0.675$^{**}$ | 0.390$^{*}$ |
| | July | -0.007 | -0.204 | -0.149 | 0.236 | -0.139 | -0.314 | -0.336 |
| | October | 0.778$^{**}$ | -0.426$^{*}$ | 0.306 | 0.344 | 0.537$^{**}$ | 0.714$^{**}$ | -0.161 |

T: temperature; RH: relative humidity; WS: wind speed; SR: solar radiation; LWC: liquid water content.

Level of significance: $^{*}$: $p<0.05$; $^{**}$: $p<0.01$.

**Table S8** Correlations between the ratios of SOA tracers to OC and RH as well as LWC.

| Compounds | Sampling periods | RH | LWC |
|---|---|---|---|
| 4-Methyl-5-nitrocatechol/OC | Whole | 0.192 | $0.248^{*}$ |
| | January | $0.733^{**}$ | $0.819^{**}$ |
| | April | $0.856^{**}$ | $0.750^{**}$ |
| | July | - | - |
| | October | 0.220 | 0.105 |
| Phthalic acid/OC | Whole | 0.105 | $0.220^{*}$ |
| | January | $0.380^{*}$ | $0.666^{**}$ |
| | April | $0.536^{**}$ | $0.599^{**}$ |
| | July | -0.272 | -0.156 |
| | October | 0.265 | $0.499^{**}$ |
| 2-Methylerythritol/OC | Whole | $0.506^{**}$ | $0.385^{**}$ |
| | January | $0.593^{**}$ | $0.690^{**}$ |
| | April | 0.244 | 0.174 |
| | July | -0.079 | 0.081 |
| | October | $0.399^{*}$ | $0.438^{**}$ |
| 3-Hydroxyglutaric acid/OC | Whole | $0.439^{**}$ | $0.323^{**}$ |
| | January | -0.119 | -0.168 |
| | April | $0.396^{*}$ | $0.366^{*}$ |
| | July | 0.136 | 0.356 |
| | October | 0.327 | $0.590^{**}$ |
| *cis*-Pinonic acid/OC | Whole | -0.076 | $-0.208^{*}$ |
| | January | $-0.668^{**}$ | $-0.788^{**}$ |
| | April | -0.214 | -0.126 |
| | July | 0.292 | -0.016 |
| | October | $-0.547^{**}$ | $-0.389^{*}$ |

Level of significance: $^{*}$: $p<0.05$; $^{**}$: $p<0.01$.

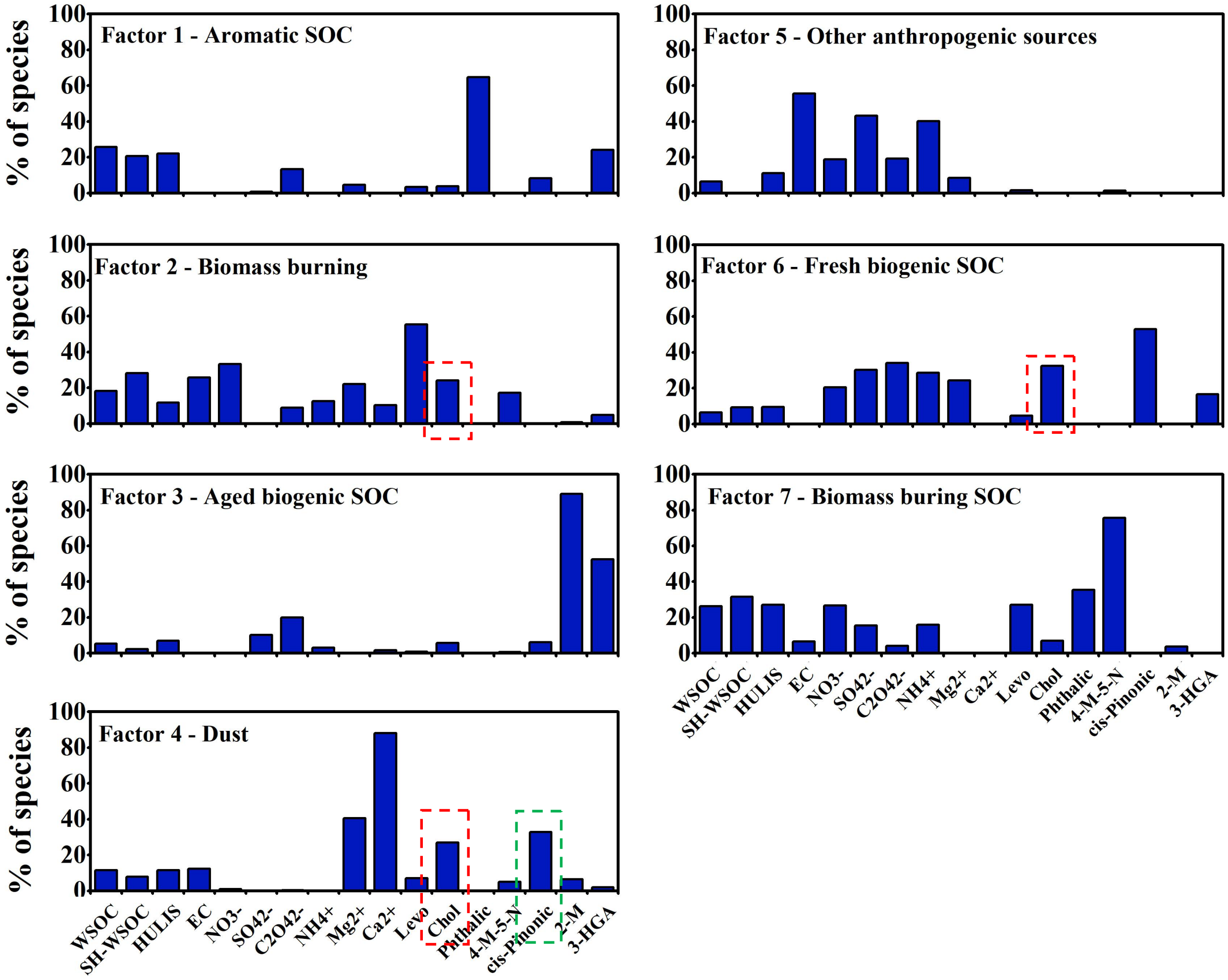


**Figure S1**. A 7-factor solution resolved by the PMF model.

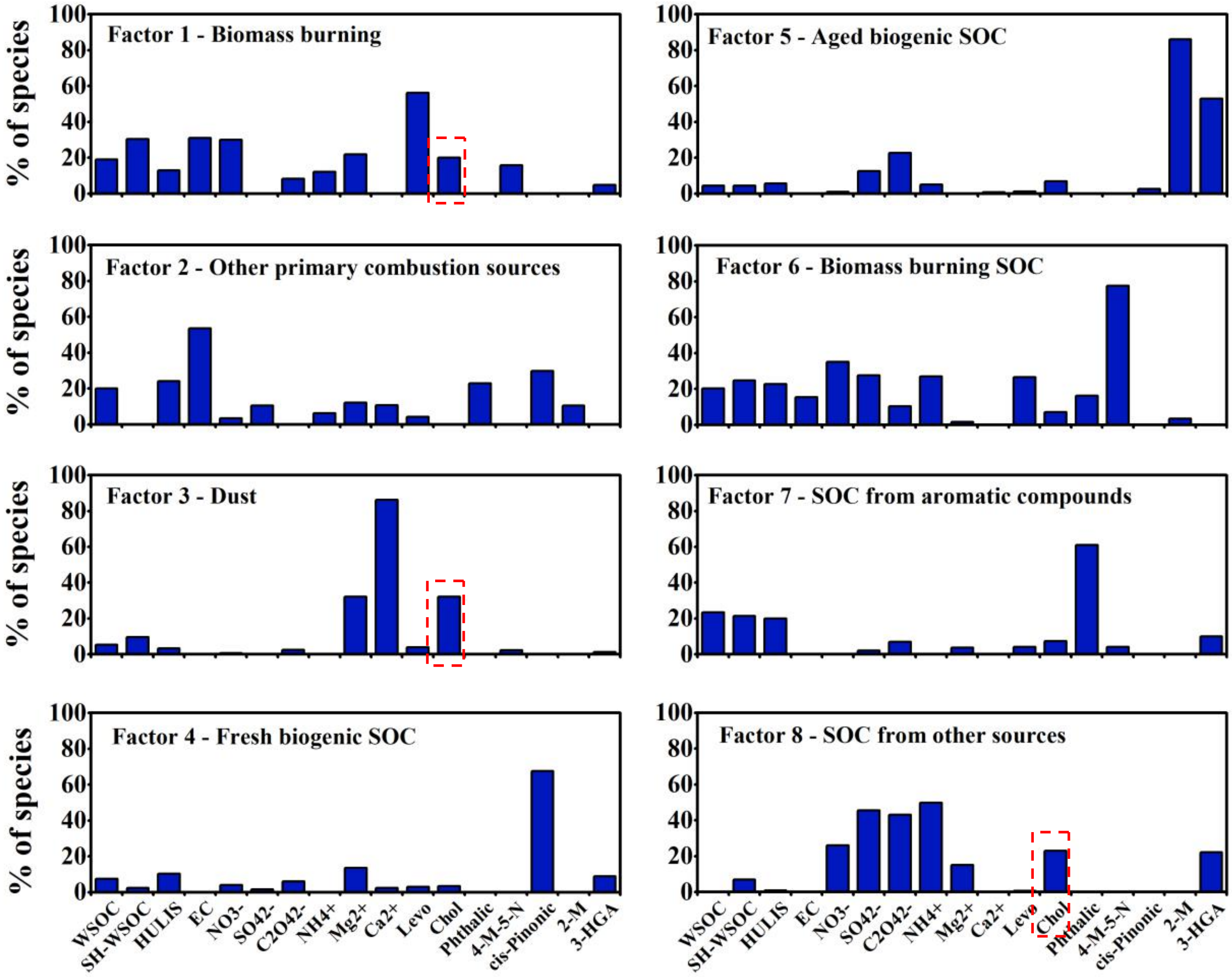


**Figure S2**. An 8-factor solution resolved by the PMF model.

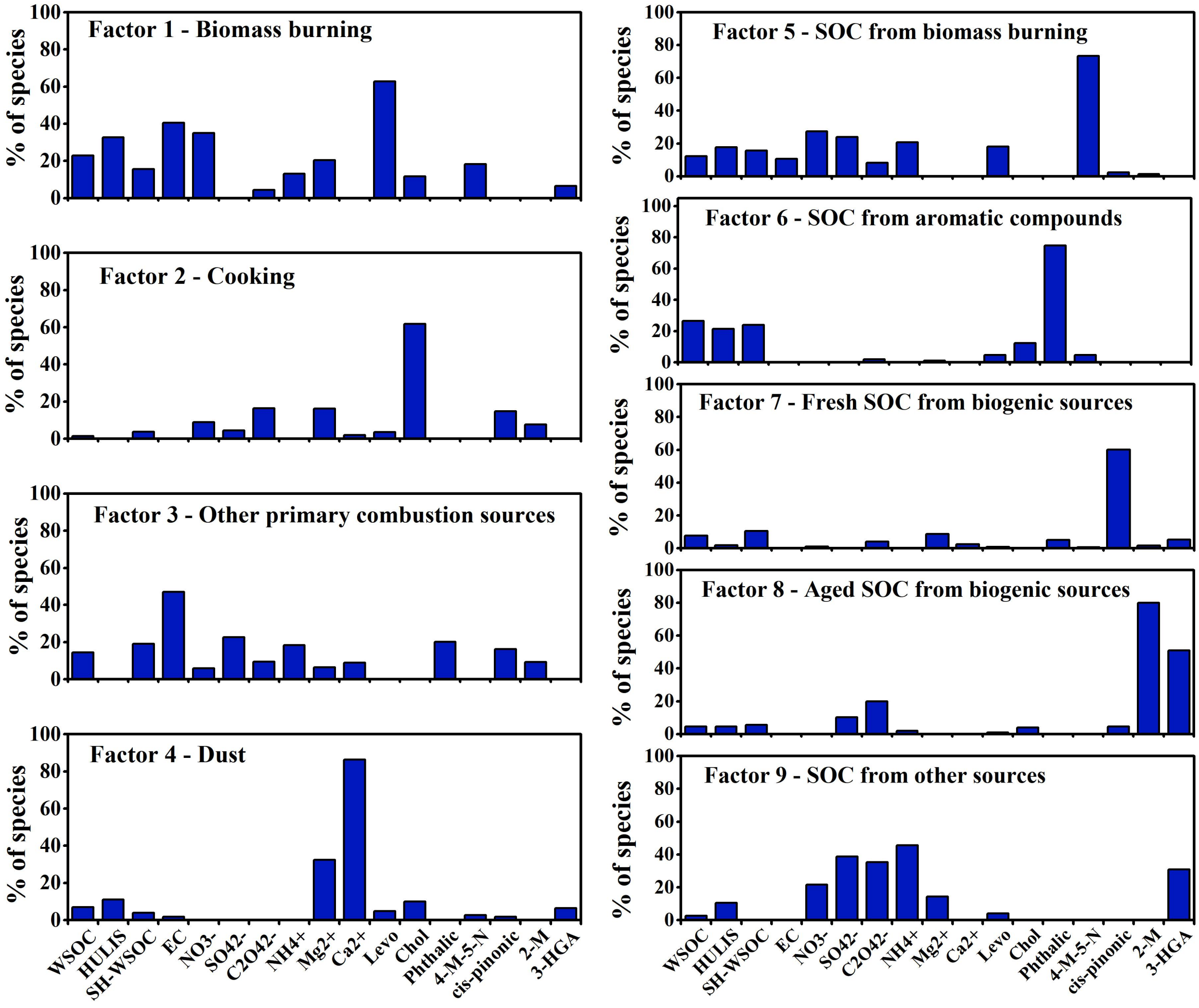


**Figure S3**. A 9-factor solution resolved by the PMF model.

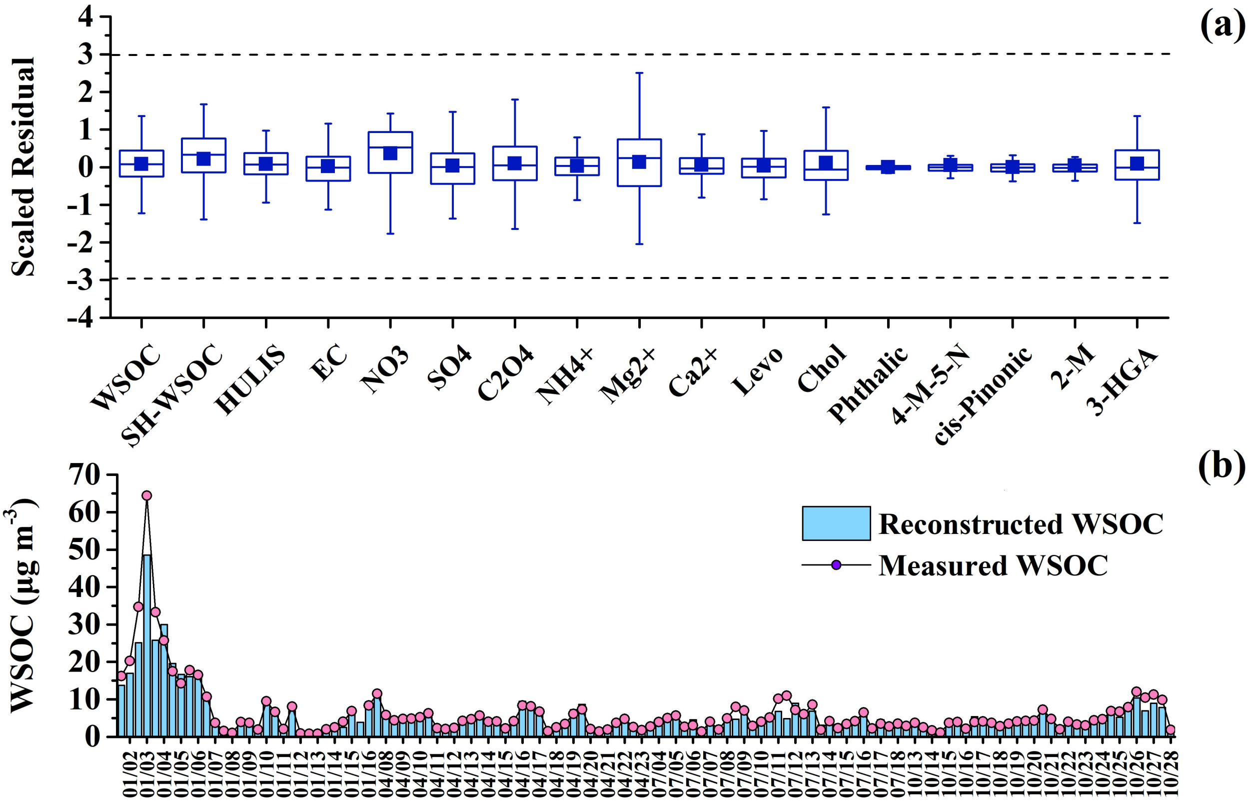


**Figure S4.** (a) The box plots showing the distributions of the scaled residuals for each species; (b) The time series of the measured WSOC and the reconstructed WSOC based on the 9-factor solution.

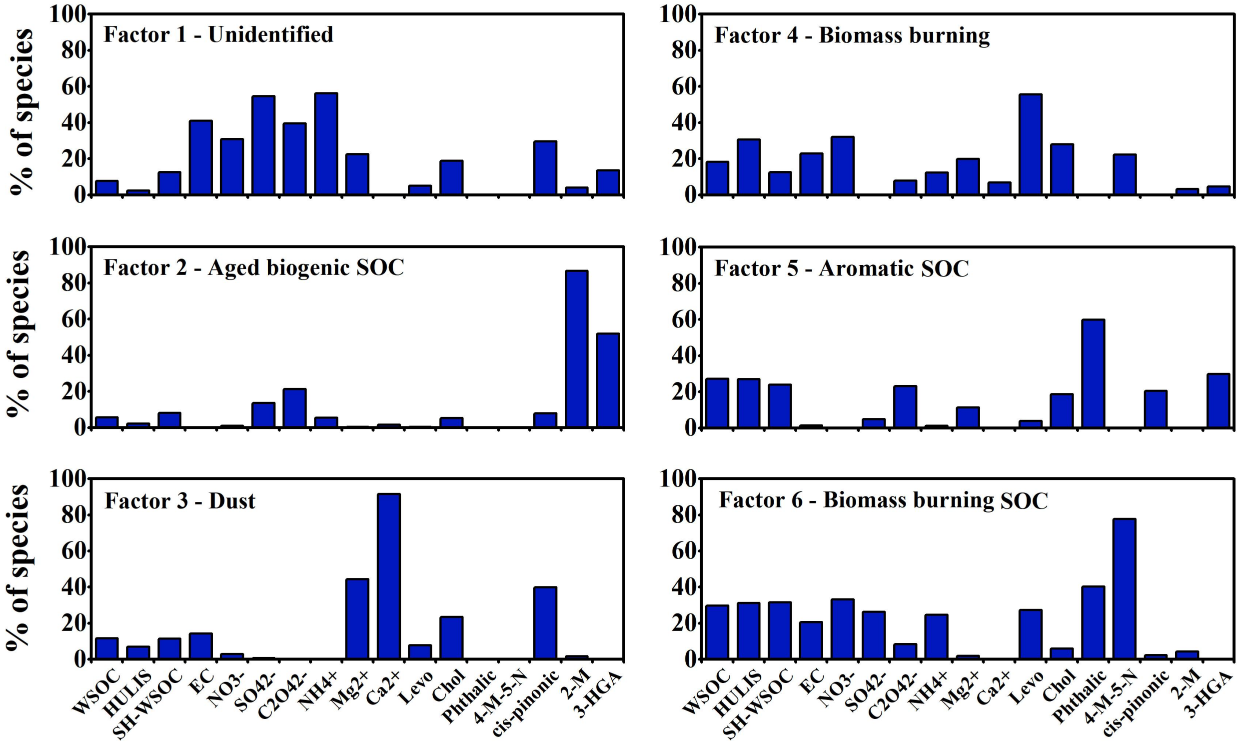


**Figure S5.** A 6-factor solution resolved by the PMF model.

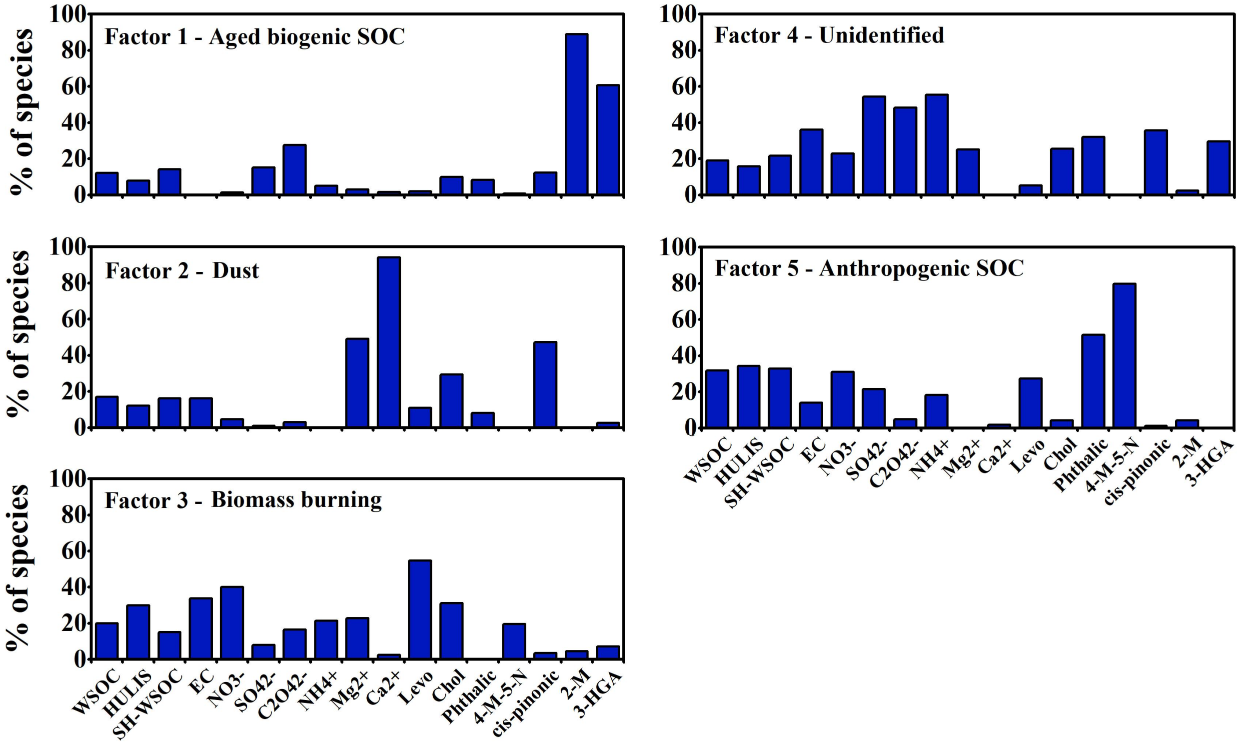


**Figure S6.** A 5-factor solution resolved by the PMF model.

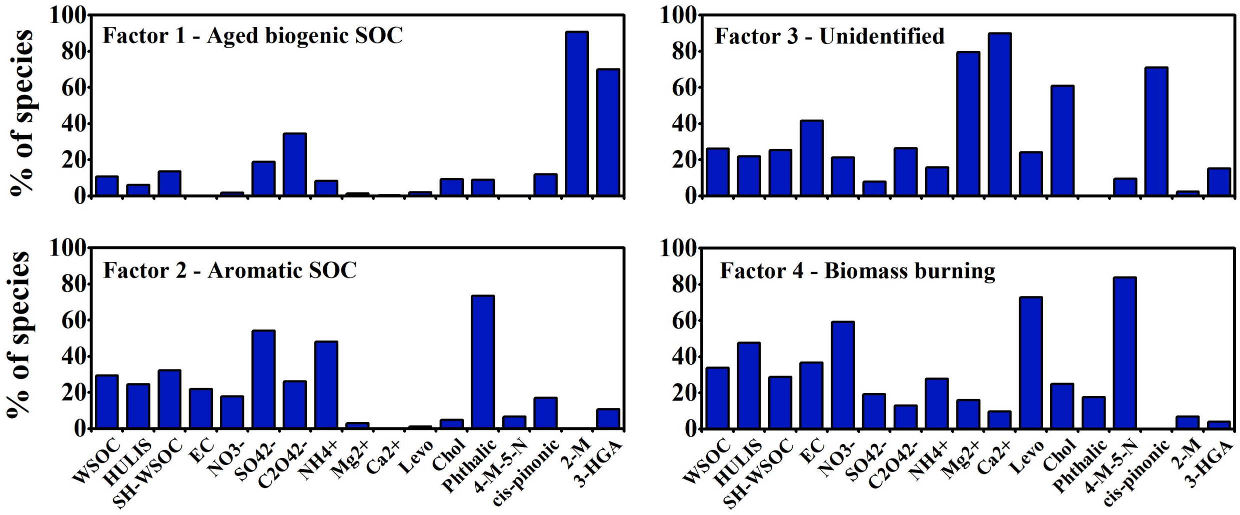


**Figure S7.** A 4-factor solution resolved by the PMF model.

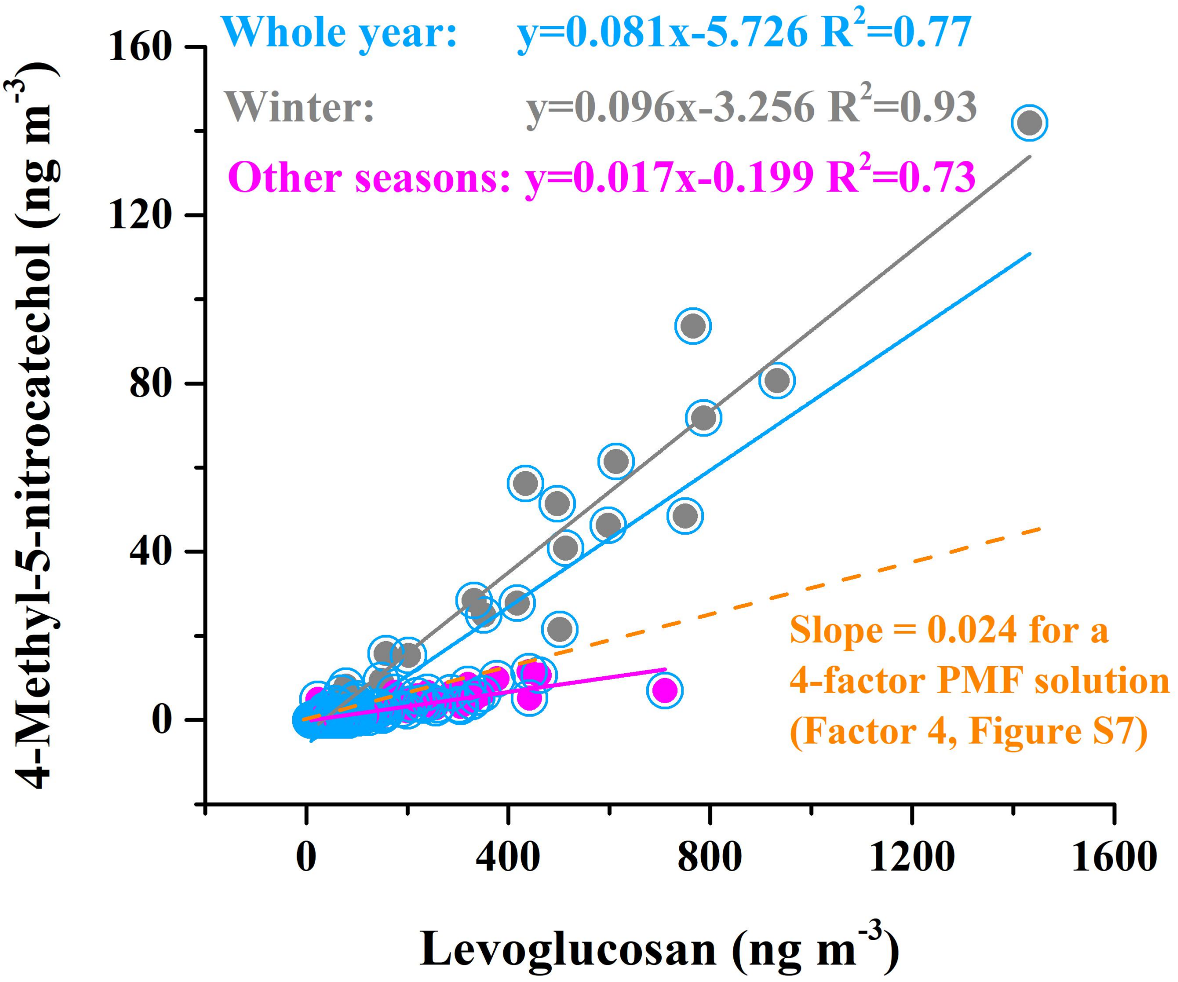


**Figure S8** The relationship between levoglucosan and 4-methyl-5-nitrocatechol.

## Reference

Brown, S. G., Eberly, S., Paatero, P., and Norris, G. A.: Methods for estimating uncertainty in PMF solutions: Examples with ambient air and water quality data and guidance on reporting PMF results, Sci. Total Environ., 518-519, 626-635, doi:10.1016/j.scitotenv.2015.01.022, 2015.

Cheng, Y., He, K., Duan, F., Zheng, M., Ma, Y., and Tan, J.: Positive sampling artifact of carbonaceous aerosols and its influence on the thermal-optical split of OC/EC, Atmos. Chem. Phys., 9, 7243-7256, 2009.

Cheng, Y., He, K., Duan, F., Zheng, M., Ma, Y., Tan, J., and Du, Z.: Improved measurement of carbonaceous aerosol: evaluation of the sampling artifacts and inter-comparison of the thermal-optical analysis methods, Atmos. Chem. Phys., 10, 8533-8548, doi:10.5194/acp-10-8533-2010, 2010.

Cheng, Y., He, K., Duan, F., Zheng, M., Du, Z., Ma, Y., and Tan, J.: Ambient organic carbon to elemental carbon ratios: Influences of the measurement methods and implications, Atmos. Environ., 45, 2060-2066, doi:10.1016/j.atmosenv.2011.01.064, 2011.

Cheng, Y., Duan, F., He, K., Du, Z., Zheng, M., and Ma, Y.: Sampling artifacts of organic and inorganic aerosol: Implications for the speciation measurement of particulate matter, Atmos. Environ., 55, 229-233, doi:10.1016/j.atmosenv.2012.03.032, 2012.

Cheng, Y., and He, K.: Uncertainties in observational data on organic aerosol: An annual perspective of sampling artifacts in Beijing, China, Environ. Pollut., 206, 113-121, doi:10.1016/j.envpol.2015.06.012, 2015.

Cui, W., Eatough, D. J., and Eatough, N. L.: Fine particulate organic material in the Los Angeles Basin-I: Assessment of the high-volume Brigham Young University Organic Sampling System, BIG BOSS, J. Air Waste Manage Assoc., 48, 1024-1037, doi: 10.1080/10473289.1998.10463760, 1998.

Dai, Q., Bi, X., Song, W., Li, T., Liu, B., Ding, J., Xu, J., Song, C., Yang, N., Schulze, B. C., Zhang, Y., Feng, Y., and Hopke, P. K.: Residential coal combustion as a source of primary sulfate in Xi'an, China, Atmos. Environ., 196, 66-76, doi: 10.1016/j.atmosenv.2018.10.002, 2019.

Ding, Y. M., Pang, Y. B., and Eatough, D. J.: High-volume diffusion denuder sampler for the routine monitoring of fine particulate matter: I. design and optimization of the PC-BOSS, Aerosol Sci. Tech., 36, 369–382, 2002.

Eatough, D. J., Wadsworth, A., Eatough, D. A., Crawford, J. W., Hansen, L. D., and Lewis, E. A.: A multiple-system, multichannel diffusion denuder sampler for the determination of fine particulate organic material in the atmosphere, Atmos. Environ., 27A, 1213–1219, 1993.

Eatough, D. J., Obeidi, F., Pang, Y. B., Ding, Y. M., Eatough, N. L., and Wilson, W. E.: Integrated and real-time diffusion denuder sampler for $PM_{2.5}$, Atmos. Environ., 33, 2835–2844, 1999.

Hao, Y., Gao, C., Deng, S., Yuan, M., Song, W., Lu, Z., and Qiu, Z.: Chemical characterisation of $PM_{2.5}$ emitted from motor vehicles powered by diesel, gasoline, natural gas and methanol fuel, Sci. Total Environ., 674, 128-139, doi: 10.1016/j.scitotenv.2019.03.410, 2019.

Kawamura, K., Tachibana, E., Okuzawa, E., Aggarwal, S. G., Kanaya, Y., and Wang, Z.: High abundances of water-soluble dicarboxylic acids, ketocarboxylic acids and α-dicarbonyls in the mountaintop aerosols over the North China Plain during wheat burning season, Atmos. Chem. Phys., 13, 8285-8302, doi: 10.5194/acp-13-8285-2013, 2013.

Kristensen, K., Bilde, M., Aalto, P. P., Petaja, T., and Glasius, M.: Denuder/filter sampling of organic acids and organosulfates at urban and boreal forest sites: Gas/particle distribution and possible sampling artifacts, Atmos. Environ., 130, 36-53, doi:10.1016/j.atmosenv.2015.10.046, 2016.

Mader, B. T., Flagan, R. C., and Seinfeld, J. H.: Sampling atmospheric carbonaceous aerosols using a particle trap impactor/denuder sampler, Environ. Sci. Technol., 35, 4857-4867, 2001.

Matsumoto, K., Hayano, T., and Uematsu, M: Positive artifact in the measurement of particulate carbonaceous substances using an ambient carbon particulate monitor, Atmos. Environ., 37, 4713-4717, doi: 10.1016/j.atmosenv.2003.07.006, 2003.

McDow, S. R.: The effect of sampling procedures on organic aerosol measurement, Ph.D. Dissertation, Oregon Graduate Center, 1987.

Sengupta, D., Sengupta, V., Bhattarai, C., Watts, A. C., Moosmuller, H., and Khlystov, A. Y.: Polar semivolatile organic compounds in biomass-burning emissions and their chemical transformations during aging in an oxidation flow reactor, Atmos. Chem. Phys., 20, 8227-8250, doi:10.5194/acp-20-8227-2020, 2020.

Shrivastava, M. K., Subramanian, R., Rogge, W. F., and Robinsen, A. L.: Sources of organic aerosol: Positive matrix factorization of molecular marker data and comparison of results from different source apportionment models, Atmos. Environ., 41, 9353-9369, doi:10.1016/j.atmosenv.2007.09.016, 2007.

Sowlat, M. H., Hasheminassab, S., and Sioutas, C.: Source apportionment of ambient particle number concentrations in central Los Angeles using positive matrix factorization (PMF), Atmos. Chem. Phys., 16, 4849-4866, doi: 10.5194/acp-16-4849-2016, 2016.

Viana, M., Chi, X., Maenhaut, W., Cafmeyer, J., Querol, X., Alastuey, A., Mikuska, P., and Vecera, Z.: Influence of sampling artifacts on measured PM, OC, and EC levels in carbonaceous aerosols in an urban area, Aerosol Sci. Tech., 40, 107-117, 2006.

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, doi:10.1021/es7022576, 2008.

Zhou, S., Collier. S., Jaffe, D. A., Briggs, N. L., Hee, J., Sedlacek III, A. J., Kleinman, L., Onasch, T. B., and Zhang, Q.: Regional influence of wildfires on aerosol chemistry in the western US and insights into atmospheric aging of biomass burning organic aerosol, Atmos. Chem. Phys., 17, 2477-2493, doi:10.5194/acp-17-2477-2017, 2017.