Supplement material

Urban organic aerosol composition in Eastern China differs from North to South: Molecular insight from a liquid chromatography-Orbitrap mass spectrometry study

Kai Wang et al.

Corresponding Author: Ru-Jin Huang (rujin.huang@ieecas.cn) and Thorsten Hoffmann (t.hoffmann@uni-mainz.de)
The mass spectra were processed using a custom software named SIEVE, which was developed by Thermo Scientific in Germany. The detailed processing steps and settings in the software are presented as follows: A threshold peak abundance of $1 \times 10^5$ arbitrary units in the two-dimensional space of the retention time window from 0–11.05 min and m/z window from 50–500 was applied to all ions. The software automatically searched the ions with their peak abundance above the threshold and only ions with peak abundance in the ambient samples 10 times higher than those in the blank samples were retained. Subsequently, all mathematically possible formulas for these ions were calculated with a mass tolerance of ±2 ppm with the elemental number ranges of $1–39$ ($^{12}$C), $1–72$ ($^1$H), $0–20$ ($^{16}$O), $0–7$ ($^{14}$N), $0–4$ ($^{32}$S) and $0–2$ ($^{35}$Cl) (Kind and Fiehn, 2007; Lin et al., 2012; Wang et al., 2018). In the positive mode (ESI+), $0–1$ of Na was also included in the formula calculation because of the high tendency of sodium to form adducts with polar organic molecules. To remove the chemically unreasonable formulas, the identified formulas were further constraint by setting H/C, O/C, N/C, S/C and Cl/C ratios in the ranges of 0.3–3, 0–3, 0–1.3, 0–0.8 and 0–0.8, respectively (Kind and Fiehn, 2007; Lin et al., 2012; Rincón et al., 2012; Wang et al., 2018; Zielinski et al., 2018). Meanwhile, the resulting neutral formulas with a non-integer or negative double bond equivalent (DBE) or elemental composition which disobey the nitrogen rule for even electron ions were also removed. It should be noted that only formulas detected in the three repetitions and observed in all three filter samples for each city were discussed in this study. The peak abundance of a compound in each city sample refers to the average area of its chromatographic peak in the three filter samples and was blank-corrected. After that, the arbitrary abundances of all isomers for a given formula were added up.

To reflect the degree of unsaturation of a compound, the double-bond equivalence (DBE) was calculated for chemical formula $\text{C}_c\text{H}_h\text{O}_o\text{N}_n\text{S}_s\text{Cl}_x$ by the equation: $\text{DBE} = (2c + 2 - h - x + n) / 2$, where $c$, $h$, $x$ and $n$ represent the numbers of atoms of carbon, hydrogen, chlorine and nitrogen, respectively.

Additionally, the aromaticity equivalent ($X_c$) was suggested to improve the identification and characterization of monoaromatic and polyaromatic compounds (Yassine et al., 2014; Kourtchev et al., 2016; Wang et al., 2017). $X_c$ of the formula $\text{C}_c\text{H}_h\text{O}_o\text{N}_n\text{S}_s\text{Cl}_x$ was calculated as follows:

$$X_c = [3(\text{DBE} - (p \times o + q \times n)) - 2]/[\text{DBE} - (p \times o + q \times n)]$$

where $p$ and $q$, respectively, refer to the fraction of oxygen and sulfur atoms involved in $\pi$-bond structure of a compound that varies based on the category of the compound. For example,
carboxylic acids and esters are characterized using $p = q = 0.5$, while $p = q = 1$ and $p = q = 0$ are used for carbonyl and hydroxyl, respectively. Since it is impossible to identify the structures of the hundreds of formulas observed in this study, we cannot know the exact values of $p$ and $q$ in an individual compound. Therefore, in this study, $p = q = 0.5$ was applied for compounds detected in ESI− due to carboxylic compounds are prone to be ionized in negative mode. However, because of the high complexity of compounds detected in ESI+, $p = q = 1$ was used in ESI+ to avoid an overestimation of the amount of aromatics. Moreover, if $\text{DBE} \leq (p \times o + q \times n)$ or $X_c \leq 0$, then $X_c$ was defined as zero. When using $p = q = 0.5$ and $(p \times o + q \times n)$ gave an odd number, the value of $(p \times o + q \times n)$ was rounded down to the lower integer (Yassine et al., 2014). $X_c \geq 2.50$ and $X_c \geq 2.71$ are suggested to unambiguous minimum criteria for the presence of monoaromatics and polyaromatics, respectively (Yassine et al., 2014).

It should be noted that different organic compounds have different signal response in the mass spectrometer, so uncertainties exist when comparing the peak areas between compounds. In this work, we assume that all organic compounds observed in this study have the same abundance response in the mass spectrometer. And the abundance-weighted average molecular mass, elemental ratios, DBE, and $X_c$ for formula $C_{c}H_{h}O_{o}N_{n}S_{s}Cl_{x}$ were calculated using following equations:

\[
MM_{\text{avg}} = \frac{\sum (M_{M} \times A_{i})}{\sum A_{i}}
\]

\[
O/C_{\text{avg}} = \frac{\sum (O/C_{i} \times A_{i})}{\sum A_{i}}
\]

\[
H/C_{\text{avg}} = \frac{\sum (H/C_{i} \times A_{i})}{\sum A_{i}}
\]

\[
\text{DBE}_{\text{avg}} = \frac{\sum (\text{DBE}_{i} \times A_{i})}{\sum A_{i}}
\]

\[
X_{c\text{avg}} = \frac{\sum (X_{c_{i}} \times A_{i})}{\sum A_{i}}
\]

where $A_{i}$ was the peak abundance for each individual formula $i$.  

84
Table S1: The daily average concentrations of PM$_{2.5}$, SO$_2$, NO$_2$, CO and O$_3$, and average temperature (T) in Changchun, Shanghai and Guangzhou during filter sampling dates.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Sampling Date</th>
<th>PM$_{2.5}$ (µg m$^{-3}$)</th>
<th>SO$_2$ (µg m$^{-3}$)</th>
<th>NO$_2$ (µg m$^{-3}$)</th>
<th>CO (mg m$^{-3}$)</th>
<th>O$_3$ (µg m$^{-3}$)</th>
<th>T (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Changchun</td>
<td>04–01–2014</td>
<td>222</td>
<td>72</td>
<td>63</td>
<td>1.5</td>
<td>43</td>
<td>−14</td>
</tr>
<tr>
<td></td>
<td>24–01–2014</td>
<td>162</td>
<td>77</td>
<td>51</td>
<td>1.4</td>
<td>70</td>
<td>−11</td>
</tr>
<tr>
<td></td>
<td>29–01–2014</td>
<td>185</td>
<td>70</td>
<td>29</td>
<td>0.9</td>
<td>58</td>
<td>−9</td>
</tr>
<tr>
<td>Shanghai</td>
<td>01–01–2014</td>
<td>171</td>
<td>63</td>
<td>99</td>
<td>1.7</td>
<td>61</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>19–01–2014</td>
<td>159</td>
<td>33</td>
<td>61</td>
<td>1.3</td>
<td>74</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>20–01–2014</td>
<td>172</td>
<td>59</td>
<td>76</td>
<td>1.8</td>
<td>42</td>
<td>6</td>
</tr>
<tr>
<td>Guangzhou</td>
<td>05–01–2014</td>
<td>152</td>
<td>39</td>
<td>89</td>
<td>1.4</td>
<td>113</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>06–01–2014</td>
<td>138</td>
<td>42</td>
<td>109</td>
<td>1.7</td>
<td>117</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>11–01–2014</td>
<td>138</td>
<td>24</td>
<td>80</td>
<td>1.6</td>
<td>69</td>
<td>16</td>
</tr>
</tbody>
</table>

The data was supported by an open air quality database (www.aqistudy.cn).
Figure S1. The van Krevelen diagram for CHO− compounds in Changchun, Shanghai and Guangzhou samples. The area of the circles is proportional to the fourth root of the peak abundance of an individual compound and the color bar denotes the aromaticity equivalent (gray with XC < 2.50, purple with 2.50 ≤ XC < 2.70 and red with XC ≥ 2.70).
Figure S2. Classification of CHON+ compounds into different subgroups according to O/N ratios in their formulas. The y-axis indicates the relative contribution of each subgroup to the sum of peak abundance of CHON+ compounds.
Figure S3. Double bond equivalent (DBE) vs C number for all CHON+ compounds of all samples. The molecular formula represents the abundance-weighted average CHON+ formula and the area of the circles is proportional to the fourth root of the peak abundance of an individual compound. The color bar denotes the aromaticity equivalent (gray with $X_C < 2.50$, purple with $2.50 \leq X_C < 2.70$ and red with $X_C \geq 2.70$). The pie charts show the percentage of each $X_C$ category (i.e., gray color-coded compounds, purple color-coded compounds and red color-coded compounds) in each sample in terms of peak abundance.
References


Wang, K., Zhang, Y., Huang, R.-J., Cao, J., and Hoffmann, T.: UHPLC-Orbitrap mass spectrometric characterization of organic aerosol from a central European city (Mainz, Germany) and a Chinese megacity (Beijing), Atmos. Environ., 189, 22-29, 10.1016/j.atmosenv.2018.06.036, 2018.

