

Substantial contribution of transported emissions to organic aerosol in Beijing

Received: 2 February 2023

Accepted: 27 June 2024

Published online: 8 August 2024

 Check for updates

Kaspar R. Daellenbach ^{1,2,3}✉, Jing Cai ^{1,2}✉, Simo Hakala ^{1,2,4},
Lubna Dada ^{1,2,3}, Chao Yan ^{1,2,5}, Wei Du ^{1,2}, Lei Yao ^{1,6}, Feixue Zheng¹,
Jialiang Ma⁷, Florian Ungeheuer ⁷, Alexander L. Vogel ⁷,
Dominik Stolzenburg ^{1,2,8}, Yufang Hao³, Yongchun Liu ¹, Federico Bianchi ^{1,2},
Gaëlle Uzu ⁹, Jean-Luc Jaffrezo⁹, Douglas R. Worsnop ^{2,10},
Neil M. Donahue ¹¹ & Markku Kulmala ^{1,2}✉

Haze in Beijing is linked to atmospherically formed secondary organic aerosol, which has been shown to be particularly harmful to human health. However, the sources and formation pathways of these secondary aerosols remain largely unknown, hindering effective pollution mitigation. Here we have quantified the sources of organic aerosol via direct near-molecular observations in central Beijing. In winter, organic aerosol pollution arises mainly from fresh solid-fuel emissions and secondary organic aerosols originating from both solid-fuel combustion and aqueous processes, probably involving multiphase chemistry with aromatic compounds. The most severe haze is linked to secondary organic aerosols originating from solid-fuel combustion, transported from the Beijing–Tianjing–Hebei Plain and rural mountainous areas west of Beijing. In summer, the increased fraction of secondary organic aerosol is dominated by aromatic emissions from the Xi'an–Shanghai–Beijing region, while the contribution of biogenic emissions remains relatively small. Overall, we identify the main sources of secondary organic aerosol affecting Beijing, which clearly extend beyond the local emissions in Beijing. Our results suggest that targeting key organic precursor emission sectors regionally may be needed to effectively mitigate organic aerosol pollution.

Globally, air pollution is responsible for several million premature deaths, many of which occur where severe pollution meets a large population (in polluted megacities)¹. In China, despite the implementation of stringent mitigation strategies², a large proportion of the population is still impacted by poor air quality. A large fraction of fine particulate matter (that is, smaller than 2.5 μm , $\text{PM}_{2.5}$) is associated with atmospherically formed secondary inorganic (SIA) and organic aerosol (SOA)^{3–5}. The influence of chemical composition on $\text{PM}_{2.5}$ health effects remains uncertain; the health risk of $\text{PM}_{2.5}$ might not be driven by its major SIA constituents (ammonium, nitrate and sulfate), but rather on OA, dominated by SOA^{6–8}. Accordingly, the recent successful reduction in SIA,

especially in sulfate⁹, might not lead to the expected health benefits, and detailed knowledge about SOA sources is essential. However, the sources of SOA are uncertain, and the processes and pathways involved in its formation are not well understood^{3,4,10}. Without comprehensive information, designing efficient SOA mitigation strategies remains stymied. Therefore, a detailed identification of the sources of SOA (source sector, temporal variability and spatial origin), as well as its formation processes, is essential in devising targeted effective reduction strategies.

In the atmosphere, SOA is produced by the complex processing of multiple gaseous organic compounds (for example, aromatic and

A full list of affiliations appears at the end of the paper. ✉e-mail: kaspar.daellenbach@psi.ch; cai.jing@helsinki.fi; markku.kulmala@helsinki.fi

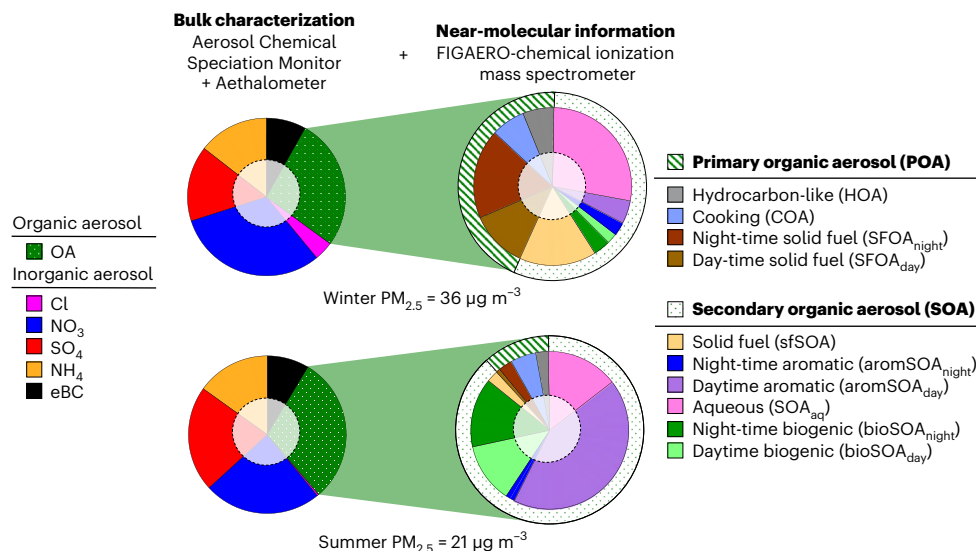


Fig. 1 | Sources of OA and their contribution to fine particle mass in winter (20 November 2019 to 25 January 2020) and summer (1 May 2020 to 2 July 2020). The bulk chemical composition of $PM_{2.5}$ from an Aerosol Chemical Speciation Monitor and Aethalometer (ToF-ACSM and AE33, left) shows that roughly one-third of the $PM_{2.5}$ mass is organic, without detailed information about its sources. Additional near-molecular information based on thermal desorption chemical ionization mass spectra (FIGAERO-CIMS, right) reveals that, even in winter, well over half of the OA is secondary, with a large contribution of aqueous processing. Solid-fuel sources with primary and secondary constituents

comprise almost half of wintertime OA. During summertime, half of the OA is secondary organic aerosol from aromatics, probably from mobile sources, and a quarter has biogenic origin. For the COVID lockdown period (26 January 2020 to 30 April 2020) see Extended Data Fig. 3. The holes in the pie charts represent constituents and sources not covered, which comprise less than 20% of the whole. A sensitivity assessment is presented in Extended Data Figs. 4 and 5 and shows that the FIGAERO-CIMS alone directly measures ~61% of the OA mass concentration based on MLR quantification, also including HOA and COA from the ToF-ACSM, presented in this figure.

biogenic). A cocktail of precursors together with atmospheric aging results in SOA with a chemical fingerprint that is similar, regardless of the original emission source, and this hinders identification of the emission sources for that SOA¹¹. Widely used mass spectrometers, such as the Aerosol Chemical Speciation Monitor, fragment the measured organic molecules present in the particles, further obscuring information on the precursor molecules¹². Only recently have newly developed field-deployable soft-ionization mass spectrometers offered semi-online characterization of OA combining substantial molecular speciation with high time resolution^{13,14} and thus greatly enhancing the potential to identify SOA sources^{15–18}.

Quantifying OA sources

In this Article we use quantitative OA aerosol mass spectrometry together with high-time resolution near-molecular OA characterization to identify and quantify SOA sources and their variability in Beijing by using advanced source apportionment techniques (positive matrix factorization)¹⁹. We combine the quantitative OA ToF-ACSM (time-of-flight Aerosol Chemical Speciation Monitor) analyses widely used for source apportionment with a factorization of time series of near-molecular organic aerosol mass spectra determined by Filter Inlet for Gas and Aerosols coupled to an iodide Chemical Ionization Mass Spectrometer (FIGAERO-CIMS)¹³. FIGAERO-CIMS uses soft chemical ionization, which allows for the detection of molecular ions—although some are affected by thermal decomposition during the measurement—and their chemical formulae, but not their structure. Overall, the FIGAERO-CIMS can detect a wide range of different anthropogenic and biogenic OA types^{20–27}. Although not all compounds in OA are detected, compared to tracer-based approaches, the FIGAERO-CIMS analyses represent a much larger OA mass fraction, estimated here to be ~61% (winter, 58%; COVID lockdown, 66%; summer, 59%), in line with previous studies^{13,28,29}, allowing for unprecedented assessment of the main sources of SOA. We can thus use the ToF-ACSM to determine the organic mass as well as the contribution of directly emitted primary organic aerosol (POA) from combustion (HOA) and cooking (COA), and use the FIGAERO-CIMS to

identify a set of SOA sources adding to POA from solid-fuel combustion. We assume that these FIGAERO-CIMS OA types constitute the ToF-ACSM OA, once HOA and COA have been accounted for, and we determine their mass loadings using multilinear regression (MLR), that is, fit the FIGAERO-CIMS OA types to OA minus (HOA + COA) from the ToF-ACSM acting as reference (Extended Data Figs. 1 and 2). To further support source identification, we rely on comparisons to laboratory SOA experiments and external tracers, for example, gas-phase oxidation products or particle-phase molecular organic source marker measurements based on $PM_{2.5}$ filters.

Beijing's $PM_{2.5}$ bulk composition is shown in Fig. 1 for a typical urban location in Beijing. Coinciding with the lunar new year in 2020 (25 January 2020), the global COVID pandemic led to a strong reduction in traffic density, coal consumption and general economic activity, all of which were restored to pre-COVID (2019) levels by the end of April 2020 (refs. 30,31). It is thus likely that emissions of anthropogenic $PM_{2.5}$ and precursors were reduced during this period but largely recovered to normal levels by the end of April. Despite those emission reductions, particulate pollution levels remained high during the COVID lockdown, although at lower concentrations compared to 2019 (Extended Data Fig. 3 and Supplementary Fig. 1). We observe a clear transition from more polluted winter conditions characterized by prominent pollution episodes, with a daily $PM_{2.5}$ mean concentration of $36 \mu\text{g m}^{-3}$, to cleaner summertime conditions, with a daily mean concentration of $21 \mu\text{g m}^{-3}$. In spite of this, the bulk chemical composition (measured by the ToF-ACSM) differs surprisingly little between seasons (Fig. 1). The $PM_{2.5}$ was composed of 61–65% SIA, 27–30% OA and 8–9% equivalent black carbon (eBC). The SIA formed from gaseous emissions, such as NO_x and SO_2 , from fossil-fuel combustion and NH_3 from diverse urban sources and agriculture. Throughout, the measured daily mean OA concentration was 3.3–3.5 times that of eBC, consistent with aged OA³² and in line with a large contribution of SOA found in previous studies^{33–36}.

There are four primary OA (POA) types: HOA (hydrocarbon organic aerosol) from liquid-fossil-fuel combustion, COA (cooking OA) and SFOA (solid-fuel combustion OA), which has contributions of biomass

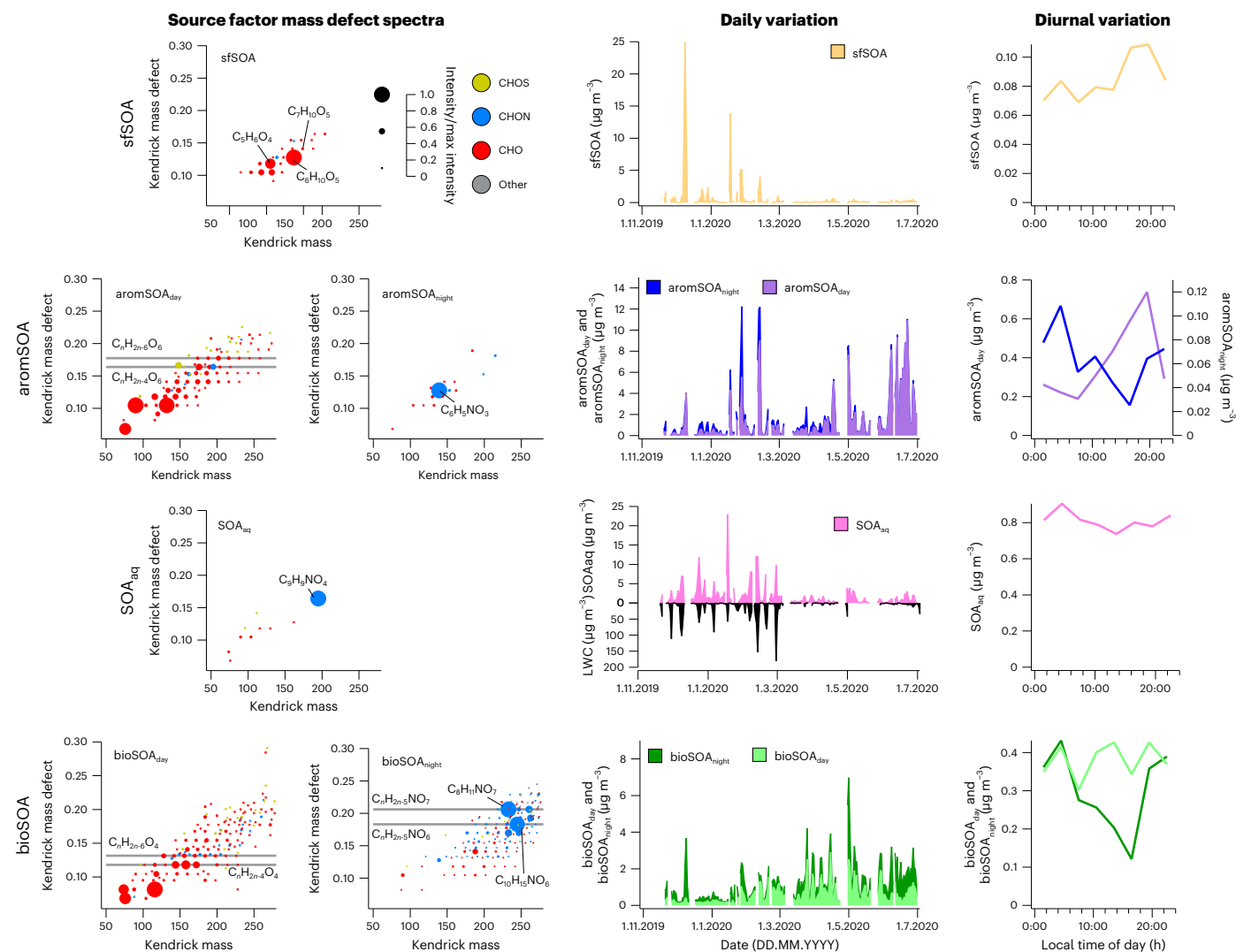


Fig. 2 | Chemical composition and temporal variability of the six secondary organic aerosol categories. Kendrick mass defect spectra show the characteristics of different sources, with near-molecular composition and relative intensity indicated by symbol colour and size, as shown in the legend (mass relative to $m(\text{CH}_2) = 14$; only compounds with an intensity of at least 10% of maximum compound intensity). The seasonal variation is shown as daily mean concentrations and the diurnal variation as median diel cycles. Solid-fuel as well as daytime aromatic-dominated and biogenic sources are principally $\text{C}_x\text{H}_y\text{O}_z$ oxidized organics. Wintertime solid-fuel SOA has prominent lignin and

cellulose-like constituents, $\text{C}_6\text{H}_{10}\text{O}_5$ and $\text{C}_7\text{H}_{10}\text{O}_5$. Daytime aromatic-dominated SOA has products consistent with trimethylbenzene laboratory SOA (grey lines)²³. Daytime biogenic SOA has constituents consistent with α -pinene laboratory SOA (grey lines)¹⁴. Nocturnal categories feature prominent nitrogen-containing species ($\text{C}_x\text{H}_y\text{O}_z\text{N}_n$), consistent with NO_3 radical oxidation. Aqueous SOA is clearly enhanced in the presence of high aerosol liquid water content (LWC), linking its formation to aqueous-phase processes. Night-time biogenic SOA has products consistent with D-limonene + NO_3 laboratory SOA (grey lines)²⁰.

burning OA (BBOA) and coal combustion OA (CCOA). There are also six secondary OA (SOA) types formed in the atmosphere: solid-fuel-SOA (sfSOA), aromatic-dominated-SOA (aromSOA_{day}, aromSOA_{night}), biogenic SOA (bioSOA_{day}, bioSOA_{night}) and aqueous SOA (SOA_{aq}).

We find that the main sources of OA vary substantially depending on the season (Fig. 1 and Extended Data Fig. 3). In winter, less than half of the OA pollution is primary, with primary solid-fuel emissions predominating (SFOA), and much of the SOA is related to solid-fuel SOA, along with aqueous SOA produced from aqueous particle-phase chemistry. In summer, SOA predominates, and solid-fuel OA (primary or secondary) almost vanishes. Although biogenic SOA is present, it remains a relatively small contributor to OA, even in summer (in northern China). Instead, we find that SOA, in summer, is dominated by emissions from anthropogenic activities. The most prominent are aromatic emissions unrelated to solid-fuel combustion, forming almost half of the OA. Other sources, such as liquid-fuel POA (HOA; winter 6% of OA, summer 3% of OA) and cooking emissions

(COA; winter 7% of OA, summer 5% of OA) contribute to a lesser extent during both seasons.

SOA emission sources and formation

SOA is diverse and governed by a variety of emissions and atmospheric formation processes. To identify the sources, we rely on the near-molecular composition of each SOA component, which we compare to laboratory SOA (Fig. 2 and Extended Data Figs. 6 and 7). We also rely on the temporal variability of each SOA component along with additional parameters (for example, organic marker compounds, gas-phase measurements; Fig. 2 and Extended Data Figs. 8 and 9). In Fig. 3, we determine the geographical origin of the SOA components.

In winter, solid-fuel combustion emissions contribute substantially to OA. We find that primary solid-fuel OA emissions are represented by distinct daytime and night-time components (SFOA_{day}, SFOA_{night}; Extended Data Fig. 6). Both SFOA_{day} and SFOA_{night} are characterized by a large contribution of $\text{C}_6\text{H}_{10}\text{O}_5$ —plausibly levoglucosan.

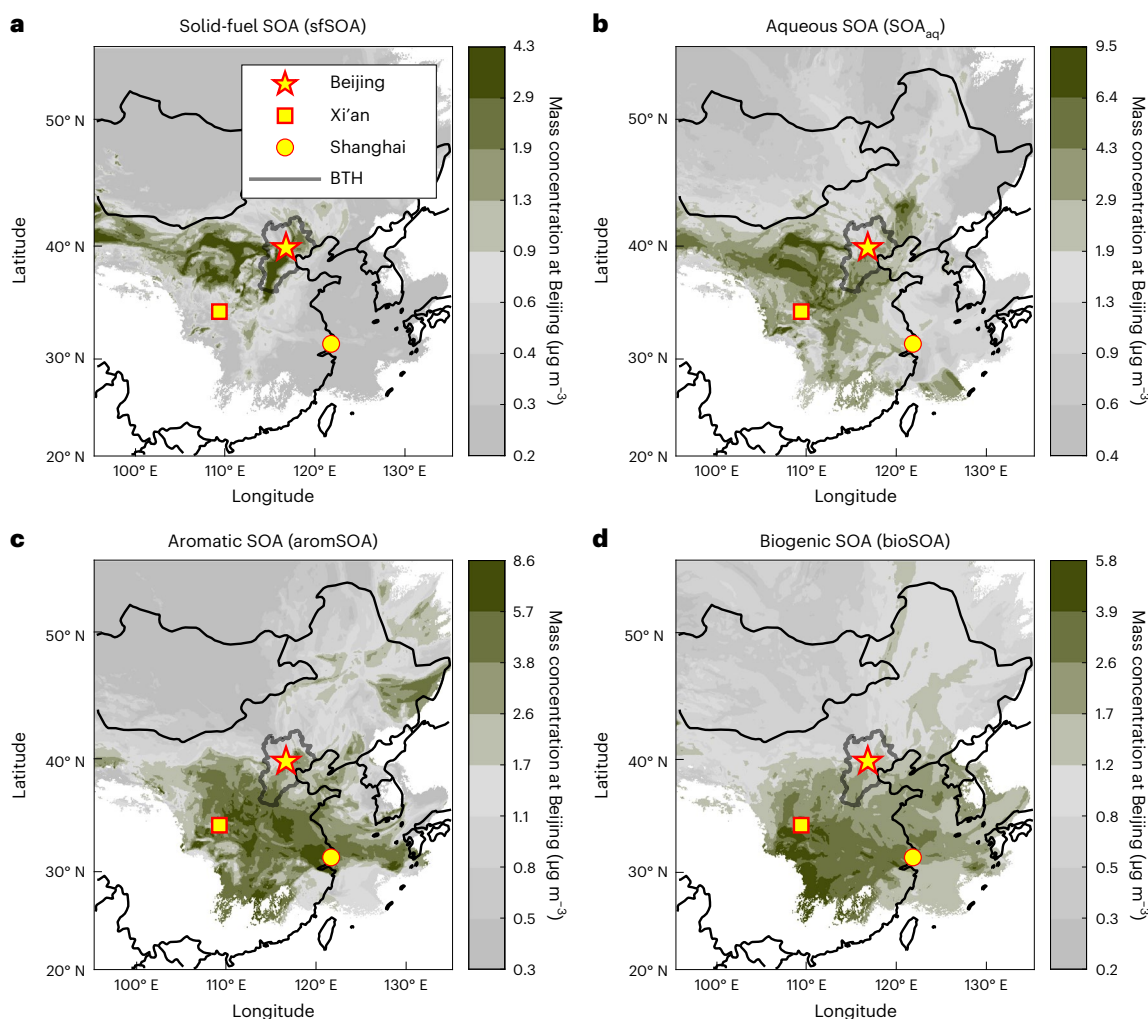


Fig. 3 | Geographical origin of SOA categories. a–d, Three-day back dispersion maps of air observed at the Beijing site (yellow star) in contact with the surface, coloured by component concentration observed at the site: solid-fuel SOA (a), aqueous SOA (b), aromatic-dominated SOA (c) and biogenic SOA (d). The average for the whole measurement period is shown. Areas contributing to above-average concentrations at the Beijing site are indicative of high emissions of the specific SOA precursors and shown in olive (below-average concentrations are shown in grey). Wintertime air masses from the Beijing–Tianjing–Hebei

(BTH) region and from regions west of Beijing show high solid-fuel SOA concentrations. Aqueous SOA also associates largely with air from regions with high SO_2 emissions, typically with very high relative humidity and thus high LWC. Summertime aromatic-dominated SOA shows regional origins from throughout the region south of Beijing, delimited roughly by a triangle defined by Beijing (yellow star), Xi'an (yellow square) and Shanghai (yellow circle). Summertime biogenic SOA also originates from the same direction but principally to the south of this region where biogenic emissions are largest.

Levoglucosan is emitted during solid-fuel combustion—predominantly from biomass but also from coal³⁷—and during this study period the concentrations are comparable to previous years (Supplementary Fig. 2). $\text{SFOA}_{\text{night}}$ is largely dominated by $\text{C}_6\text{H}_{10}\text{O}_5$, with minor contributions of $\text{C}_5\text{H}_8\text{O}_4$ (possibly glutaric acid). SFOA_{day} shows a higher contribution of other compounds than $\text{C}_6\text{H}_{10}\text{O}_5$, such as $\text{C}_7\text{H}_{10}\text{O}_5$ (found in laboratory biomass burning SOA³⁸) and $\text{C}_8\text{H}_{12}\text{O}_5$, as well as nitroaromatics such as $\text{C}_6\text{H}_5\text{NO}_4$ and $\text{C}_7\text{H}_7\text{NO}_4$. This chemical composition of SFOA_{day} is in line with aged emissions²⁸. Together with the observed daytime maximum concentrations, this suggests that SFOA undergoes rapid photochemical transformation. The sum of SFOA_{day} and $\text{SFOA}_{\text{night}}$ shows a similar temporal behaviour as solid-fuel POA (sum of biomass burning, BBOA and coal combustion, CCOA, emissions) quantified by the ToF-ACSM (Extended Data Fig. 6). Although during clean winter conditions coal combustion contributes between 65% and 96% to SFOA, more polluted episodes are strongly affected or even dominated (48–90%) by biomass burning (Extended Data Fig. 6). Additionally, an aged solid-fuel component (solid-fuel SOA, sfSOA) has substantial $\text{C}_6\text{H}_{10}\text{O}_5$, but also a prominent influence of low-molecular-weight

compounds ($\text{C}_{2-5}\text{H}_{2-8}\text{O}_4$)—plausibly related to small dicarboxylic acids (Fig. 2). The solid-fuel SOA (as well as $\text{SFOA}_{\text{night}}$ and SFOA_{day}) is clearly enhanced during cold-period haze episodes and decreases substantially towards the warm season (winter mean, 15% of OA; summer mean, 2% of OA). Solid-fuel SOA and SFOA show high concentrations in air masses arriving from the Beijing–Tianjing–Hebei region, but also from the rural mountainous regions west and northeast of Beijing (Fig. 3a, Extended Data Fig. 10 and Supplementary Fig. 3), suggesting strong precursor emissions in these regions that are transported to Beijing.

Many activity sectors, including industry, energy and transportation, contribute to the emission of SOA precursors. Here we identify SOA related to emissions in the highly populated Xi'an–Shanghai–Beijing region (Fig. 3c), which show chemical characteristics indicative of aromatic precursor emissions (aromSOA; Fig. 2 and Extended Data Fig. 7). Daytime aromatic-dominated SOA ($\text{aromSOA}_{\text{day}}$) in general shows a chemical fingerprint resembling laboratory SOA from aromatic precursors (here trimethylbenzene, TMB)²³ reacting with OH, with prominent contributions of, for example, $\text{C}_{7-9}\text{H}_{8-12}\text{O}_6$, $\text{C}_{6-7}\text{H}_{8-10}\text{O}_6$ and $\text{C}_{6-7}\text{H}_{8-10}\text{O}_5$ (Fig. 2 and Extended Data Fig. 7c). Consistent with such a

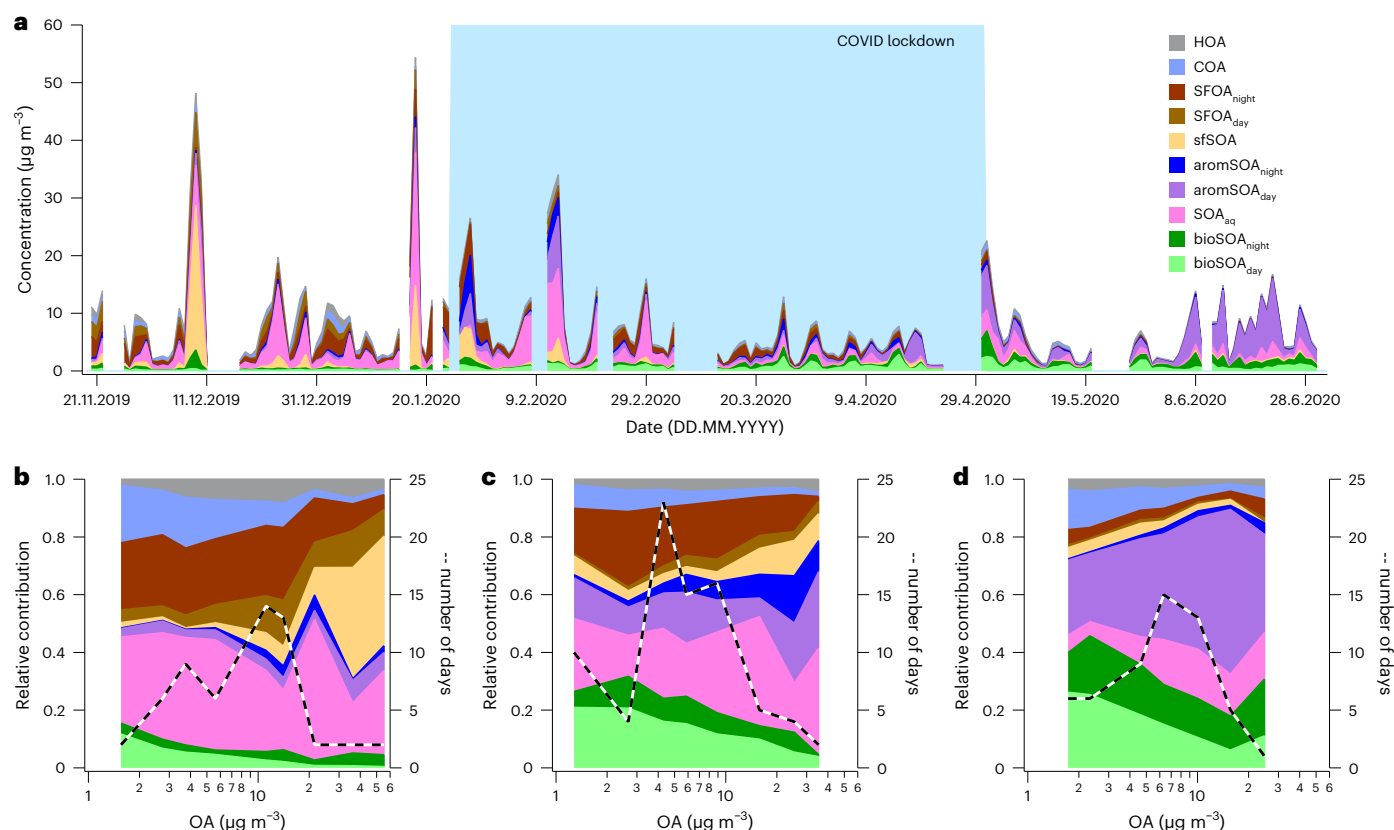


Fig. 4 | Impact of emission sources on OA air pollution. a–d. Concentration time series (a) and relative contribution of sources to OA at different pollution levels (b–d, daily averages). The measurement period is separated into winter (20 November 2019 to 25 January 2020, b), COVID lockdown (26 January 2020 to 30 April 2020, c), summer (1 May 2020 to 2 July 2020, d). The most prominent

OA sources are disproportionately important during severe haze episodes, with solid fuel (especially secondary SFOA) comprising more than half of OA during wintertime haze events and aromatic-dominated SOA playing a major role during summertime haze events.

formation pathway, aromSOA_{day} concentrations build up throughout the photochemically active hours of the day. Nevertheless, we do not rule out that other emissions (predominantly anthropogenic), including non-aromatic (such as alkanes from traffic, for example), contribute to this SOA category. On the other hand, night-time aromatic-dominated SOA (aromSOA_{night}) is characterized by a dominant C₆H₅NO₃ signal—plausibly nitrophenol peaking at night. Although aromSOA_{night} is highest during the spring, aromSOA_{day} concentrations are highest during the summer, consistent with higher irradiation and photochemical activity. In addition, the East Asian monsoon circulation probably contributes to the seasonal variability in aromatic-dominated SOA. During winter, the transport of pollution to Beijing is mainly influenced by north China, but in summer the influence extends further south throughout the Xi'an–Shanghai–Beijing region (Fig. 3c, Extended Data Fig. 10 and Supplementary Figs. 4 and 5)^{39,40}. Thus, SOA from precursor sources that are mainly located in these southern regions are expected to show higher concentrations in Beijing during the summer. Air masses from regions associated with aromatic-dominated SOA have an age of up to two to three days (Supplementary Fig. 6), so long-range transport over extended time periods further facilitated by high oxidant concentrations^{41–43} enabling atmospheric processing could plausibly explain why aromSOA_{day} is strongly oxidized when arriving at the measurement site.

With increasing temperatures, biogenic-SOA precursor emissions increase, driving increased gas-phase concentrations of oxygenated organic molecules (OOM) from monoterpene and isoprene oxidation⁴⁴. We observe an increasing biogenic-SOA concentration during the transition from winter to summer (Fig. 2). In the summer, biogenic

SOA contributes an average of 27% to OA. Its concentration increases as the temperature rises (0.6 µg m⁻³ at 0 °C; 2.1 µg m⁻³ at 25–30 °C). The biogenic-SOA concentrations found here are similar to an estimate based on the concentration–temperature relation of methylbutanetricarboxylic acid (MBTCA) and pinic acid⁷, oxidation products of α-pinene (Extended Data Fig. 8). In addition, biogenic SOA correlates better than aromatic-dominated SOA with gas-phase OOM from isoprene and monoterpene oxidation (Extended Data Fig. 9). Biogenic SOA is sensitive to emissions from the forested areas in southern China (Fig. 3d and Extended Data Fig. 10), where large biogenic emission fluxes are expected⁴⁵. Daytime biogenic SOA (bioSOA_{day}) is characterized by compounds that exhibit similarities with laboratory α-pinene ozonolysis SOA such as C₈H₁₂O₄ or C₈H₁₀O₅ (Fig. 2 and Extended Data Fig. 7)¹⁴. In addition to compounds indicative of biogenic SOA from terpenes, smaller-molecular-weight compounds were also substantial contributors. These may be related to enhanced atmospheric fragmentation in the urban atmosphere or in part to other biogenic SOA precursors such as isoprene subjected to NO_x (C₂H₄O₃, C₄H₈O₃, C₄H₇NO₅ and C₅H₉NO₅)²¹. Some of the smaller-molecular-weight compounds (for example, C₂H₄O₃ and C₄H₈O₃) could also be fragmentation products of larger compounds from thermal decomposition in FIGAERO-CIMS during thermal desorption (Supplementary Fig. 7). Accordingly, we use the entire chemical fingerprint (including small- and large-molecular-weight compounds) to interpret the SOA sources. Although, chemically, daytime biogenic SOA shows some similarity to daytime aromatic-dominated SOA, compounds found in laboratory aromatic SOA are clearly less abundant in bioSOA_{day} than in aromSOA_{day} (Extended Data Fig. 7). During the night, biogenic SOA (bioSOA_{night}) is

dominated by compounds ($C_8H_{11}NO_7$, $C_{10}H_{15}NO_6$, $C_{10}H_{17}NO_6$, $C_9H_{15}NO_7$, $C_{10}H_{15}NO_7$, $C_{10}H_{17}NO_7$ and $C_{10}H_{15}NO_8$) that have been identified as dominant in laboratory SOA from limonene reacting with nitrate radicals—a typical reaction pathway during the night²⁰ (Fig. 2 and Extended Data Fig. 7).

SOA can also be formed by the multiphase chemistry of condensing vapours on particles, or in fog droplets. It has often been hypothesized that a substantial fraction of haze SOA in Beijing is formed in the aqueous phase^{10,46–50}. However, recent estimates suggest that in Beijing during the winter, a major fraction of SOA is formed through oxidation and subsequent condensation of gas-phase precursors^{51,52}. Based on our measurements, we observe an SOA type strongly associated with high particle liquid water content (LWC) concentrations (winter, 28% of OA; summer, 15% of OA), indicating that it is SOA formed in the aqueous phase (aqueous SOA, SOA_{aq} ; Fig. 2, $R = 0.68$). In comparison, aromatic-dominated SOA is only weakly correlated with particle LWC ($R_{aromSOA_{night}} = 0.32$; $R_{aromSOA_{day}} = 0.17$) and is thus apparently not related to aqueous formation pathways. Solid-fuel SOA is enhanced in the presence of high particle LWC, although the association between solid-fuel SOA and particle LWC is quite scattered ($R = 0.48$), suggesting that other formation pathways play an important role. Air-mass backward dispersion analysis further supports the identification of aqueous SOA. Air masses with high aqueous SOA and sulfate—known to be strongly influenced by aqueous formation⁴—pass over similar regions characterized by high SO_2 emissions (Fig. 3b and Supplementary Figs. 8 and 9). In addition, wintertime air masses with high aqueous-SOA loadings are also influenced by transport over the Bohai Sea where the air masses can take up water vapour (Extended Data Fig. 10). During the winter, aqueous SOA contributes 49% to SOA, highlighting the important role of multiphase chemistry. This is consistent with estimates that 38% of SOA during winter in Beijing is formed through condensing oxygenated organic molecules^{51,52}, leaving 62% of SOA formed via other unaccounted-for formation processes, such as multiphase pathways. Aqueous SOA is dominated by $C_9H_9NO_4$ (found in ambient cloud water⁵³, possibly dimethylnitrobenzoic acid), indicating a strong influence from anthropogenic aromatic emissions. Additionally, aqueous SOA contains small-molecular-weight compounds ($C_{2-5}H_{2-8}O_{3-5}$) consistent with small mono- and dicarboxylic acids, further supporting our assignment of aqueous SOA²⁶.

Sources governing OA during pollution episodes

In the winter, during clean conditions, ~50% of OA consists of POA from traffic exhaust (HOA), cooking (COA), but especially SFOA, dominated by CCOA at low concentrations (Fig. 4a,b and Extended Data Fig. 6g). SOA, dominated by aqueous-phase formation processes (SOA_{aq}), contributes ~50%. During pollution episodes (daily mean OA concentrations reaching $>35 \mu g m^{-3}$), the contribution of SOA driven by solid-fuel SOA (38–39% of OA) increases substantially, reaching up to 80% during severe haze episodes (Fig. 4a,b). Because biomass-burning emissions dominate the primary SFOA, such emissions could also be the main driver of solid-fuel SOA during these events. Other sources contribute to winter SOA during haze episodes, including SOA related to the aqueous particle phase, aqueous SOA (17–29% of OA) and aromatic-dominated SOA (8–9% of OA). Interestingly, precursor emission sources driving increased SOA are located outside Beijing, with a substantial contribution from the Beijing–Tianjin–Hebei region (Fig. 3a and Extended Data Fig. 10).

In contrast, during clean summer conditions, OA is dominated by SOA, with POA (HOA, COA and SFOA) comprising only 19% of the OA (Fig. 4a,d). In the summer, aromatic-dominated SOA is the main driver of SOA (61%), and biogenic SOA remains a substantially smaller contributor (36%). During pollution episodes in the summer, the contribution of aromatic-dominated SOA is enhanced, contributing 38–58% to OA. Even if biogenic SOA is a relatively small contributor to OA during pollution episodes, it is clearly influenced by interactions with

anthropogenic NO_x emissions. There is more $bioSOA_{night}$ compared to $bioSOA_{day}$ at high OA levels than at lower OA levels, indicating that during polluted conditions, biogenic SOA is preferentially produced through interactions between biogenic SOA precursors and anthropogenic NO_x . This is in line with observations elsewhere showing that biogenic SOA is enhanced when interacting with urban anthropogenic emissions³⁴. SOA precursors driving SOA formation are mainly emitted outside Beijing, as air masses reside over the Xi'an–Shanghai–Beijing region (Fig. 3 and Extended Data Fig. 10). Overall, SOA drives OA pollution episodes in summer and winter Beijing. However, SOA precursor emissions differ in summer and winter, with distinct sources and geographical origins outside Beijing.

Implications

To fully understand haze in highly polluted megacities and to design targeted effective mitigation strategies, detailed molecular information is needed to identify the sources of OA, which could be particularly harmful to human health^{6–8}. SOA in summer and winter is driven by chemically and geographically different precursor emission sources arriving from outside Beijing. This shows that, although our focus is on pollution within Beijing, haze is a large-scale regional phenomenon, with transport of different SOA sources over hundreds of kilometres before the particles are removed. This is in line with previous observations and model simulations^{55–58}. Accordingly, to achieve substantial OA reductions, coordinated and stringent large-scale air-quality policies are required across one of the most populated regions (Xi'an–Shanghai–Beijing)⁵⁹. Our conclusions are consistent with observations during the COVID-19 lockdown, during which reductions in traffic density, coal consumption and general economic activity did not fully mitigate pollution in Beijing driven by secondary $PM_{2.5}$ formed in the atmosphere^{60,61}. As a starting point for globally improving air quality via targeted mitigation strategies, our framework based on detailed near-molecular chemical characterization of particulate air pollution opens new research avenues for identifying aerosol sources and assessing their toxicity as well as their impact on public health.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41561-024-01493-3>.

References

- Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D. & Pozzer, A. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* **525**, 367–371 (2015).
- Li, X. et al. Responses of gaseous sulfuric acid and particulate sulfate to reduced SO_2 concentration: a perspective from long-term measurements in Beijing. *Sci. Total Environ.* **721**, 137700 (2020).
- Huang, R.-J. et al. High secondary aerosol contribution to particulate pollution during haze events in China. *Nature* **514**, 218–222 (2014).
- Elser, M. et al. 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 (2016).
- Sun, Y. et al. 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 (2018).
- Zhou, J. et al. Predominance of secondary organic aerosol to particle-bound reactive oxygen species activity in fine ambient aerosol. *Atmos. Chem. Phys.* **19**, 14703–14720 (2019).

7. Daellenbach, K. R. et al. Sources of particulate-matter air pollution and its oxidative potential in Europe. *Nature* **587**, 414–419 (2020).
8. Pye, H. O. T., Ward-Caviness, C. K., Murphy, B. N., Appel, K. W. & Seltzer, K. M. Secondary organic aerosol association with cardiorespiratory disease mortality in the United States. *Nat. Commun.* **12**, 7215 (2021).
9. Zheng, B. et al. Trends in China's anthropogenic emissions since 2010 as the consequence of clean air actions. *Atmos. Chem. Phys.* **18**, 14095–14111 (2018).
10. Tong, Y. et al. Quantification of solid fuel combustion and aqueous chemistry contributions to secondary organic aerosol during wintertime haze events in Beijing. *Atmos. Chem. Phys.* **21**, 9859–9886 (2021).
11. Kroll, J. H. et al. Carbon oxidation state as a metric for describing the chemistry of atmospheric organic aerosol. *Nat. Chem.* **3**, 133–139 (2011).
12. Canagaratna, M. et al. Chemical and microphysical characterization of ambient aerosols with the aerodyne aerosol mass spectrometer. *Mass Spectrom. Rev.* **26**, 185–222 (2007).
13. Lopez-Hilfiker, F. et al. A novel method for online analysis of gas and particle composition: description and evaluation of a Filter Inlet for Gases and AEROSols (FIGAERO). *Atmos. Meas. Tech.* **7**, 983–1001 (2014).
14. Lopez-Hilfiker, F. D. et al. An extractive electrospray ionization time-of-flight mass spectrometer (EESI-TOF) for online measurement of atmospheric aerosol particles. *Atmos. Meas. Tech.* **12**, 4867–4886 (2019).
15. Stefenelli, G. et al. Organic aerosol source apportionment in Zurich using an extractive electrospray ionization time-of-flight mass spectrometer (EESI-TOF-MS) – Part 1: Biogenic influences and day-night chemistry in summer. *Atmos. Chem. Phys.* **19**, 14825–14848 (2019).
16. Qi, L. et al. Organic aerosol source apportionment in Zurich using an extractive electrospray ionization time-of-flight mass spectrometer (EESI-TOF-MS) – Part 2: Biomass burning influences in winter. *Atmos. Chem. Phys.* **19**, 8037–8062 (2019).
17. Chen, Y. et al. Chemical characterization of secondary organic aerosol at a rural site in the southeastern US: insights from simultaneous high-resolution time-of-flight aerosol mass spectrometer (HR-ToF-AMS) and FIGAERO chemical ionization mass spectrometer (CIMS) measurements. *Atmos. Chem. Phys.* **20**, 8421–8440 (2020).
18. Mehra, A. et al. Using highly time-resolved online mass spectrometry to examine biogenic and anthropogenic contributions to organic aerosol in Beijing. *Faraday Discuss.* **226**, 382–408 (2021).
19. Paatero, P. & Tapper, U. Positive matrix factorization: a non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics* **5**, 111–126 (1994).
20. Faxon, C., Hammes, J., Le Breton, M., Pathak, R. K. & Hallquist, M. Characterization of organic nitrate constituents of secondary organic aerosol (SOA) from nitrate-radical-initiated oxidation of limonene using high-resolution chemical ionization mass spectrometry. *Atmos. Chem. Phys.* **18**, 5467–5481 (2018).
21. D'Ambro, E. L. et al. Molecular composition and volatility of isoprene photochemical oxidation secondary organic aerosol under low- and high-NO_x conditions. *Atmos. Chem. Phys.* **17**, 159–174 (2017).
22. Dada, L. et al. Role of sesquiterpenes in biogenic new particle formation. *Sci. Adv.* **9**, eadi5297 (2023).
23. Mehra, A. et al. Evaluation of the chemical composition of gas- and particle-phase products of aromatic oxidation. *Atmos. Chem. Phys.* **20**, 9783–9803 (2020).
24. Masoud, C. G. et al. Molecular composition and gas-particle partitioning of indoor cooking aerosol: insights from a FIGAERO-CIMS and kinetic aerosol modeling. *Aerosol Sci. Technol.* **56**, 1156–1173 (2022).
25. Gaston, C. J. et al. Online molecular characterization of fine particulate matter in Port Angeles, WA: evidence for a major impact from residential wood smoke. *Atmos. Environ.* **138**, 99–107 (2016).
26. Zheng, Y. et al. Precursors and pathways leading to enhanced secondary organic aerosol formation during severe haze episodes. *Environ. Sci. Technol.* **55**, 15680–15693 (2021).
27. Cai, J. et al. Influence of organic aerosol molecular composition on particle absorptive properties in autumn Beijing. *Atmos. Chem. Phys.* **22**, 1251–1269 (2022).
28. Ye, C. et al. Chemical characterization of oxygenated organic compounds in the gas phase and particle phase using iodide CIMS with FIGAERO in urban air. *Atmos. Chem. Phys.* **21**, 8455–8478 (2021).
29. Lee, B. H. et al. Resolving ambient organic aerosol formation and aging pathways with simultaneous molecular composition and volatility observations. *ACS Earth Space Chem.* **4**, 391–402 (2020).
30. Cao, X. et al. Emission variations of primary air pollutants from highway vehicles and implications during the COVID-19 pandemic in Beijing. *China Int. J. Environ. Res. Public Health* **18**, 4019 (2021).
31. ECB. China's path to normalisation in the aftermath of the COVID-19 pandemic. *ECB Economic Bulletin*, 6/2020 (2020); https://www.ecb.europa.eu/pub/economic-bulletin/articles/2020/html/ecb.ebart202006_01-853248c2d2.en.html
32. Chen, G. et al. European aerosol phenomenology—8: Harmonised source apportionment of organic aerosol using 22 year-long ACSM/AMS datasets. *Environ. Int.* **166**, 107325 (2022).
33. Sun, Y. et al. A chemical cocktail during the COVID-19 outbreak in Beijing, China: insights from six-year aerosol particle composition measurements during the Chinese New Year holiday. *Sci. Total Environ.* **742**, 140739 (2020).
34. Gu, Y. et al. Multiple pathways for the formation of secondary organic aerosol in the North China Plain in summer. *Atmos. Chem. Phys.* **23**, 5419–5433 (2023).
35. Duan, J. et al. Summertime and wintertime atmospheric processes of secondary aerosol in Beijing. *Atmos. Chem. Phys.* **20**, 3793–3807 (2020).
36. Zheng, Y. et al. Secondary formation of submicron and supermicron organic and inorganic aerosols in a highly polluted urban area. *J. Geophys. Res. Atmos.* **128**, e2022JD037865 (2023).
37. Yan, C. et al. Residential coal combustion as a source of levoglucosan in China. *Environ. Sci. Technol.* **52**, 1665–1674 (2018).
38. Yee, L. D. et al. Secondary organic aerosol formation from biomass burning intermediates: phenol and methoxyphenols. *Atmos. Chem. Phys.* **13**, 8019–8043 (2013).
39. Yancheva, G. et al. Influence of the intertropical convergence zone on the East Asian monsoon. *Nature* **445**, 74–77 (2007).
40. Kaboth-Bahr, S. et al. A tale of shifting relations: East Asian summer and winter monsoon variability during the Holocene. *Sci. Rep.* **11**, 6938 (2021).
41. Nan, J., Wang, S., Guo, Y., Xiang, Y. & Zhou, B. Study on the daytime OH radical and implication for its relationship with fine particles over megacity of Shanghai, China. *Atmos. Environ.* **154**, 167–178 (2017).
42. Lu, K. et al. Fast photochemistry in wintertime haze: consequences for pollution mitigation strategies. *Environ. Sci. Technol.* **53**, 10676–10684 (2019).

43. Tan, Z. et al. Radical chemistry at a rural site (Wangdu) in the North China Plain: observation and model calculations of OH, HO₂ and RO₂ radicals. *Atmos. Chem. Phys.* **17**, 663–690 (2017).
44. Guo, Y. et al. Seasonal variation of oxygenated organic molecules in urban Beijing and their contribution to secondary organic aerosol. *Atmos. Chem. Phys. Discuss.* **22**, 10077–10097 (2022).
45. Yin, L. et al. Estimation of biogenic volatile organic compound (BVOC) emissions in China using WRF-CLM-MEGAN coupled model. *Biogeosci. Discuss.* **2020**, 1–30 (2020).
46. Wang, J. et al. Aqueous production of secondary organic aerosol from fossil-fuel emissions in winter Beijing haze. *Proc. Natl Acad. Sci. USA* **118**, e2022179118 (2021).
47. Xu, W. et al. Effects of aqueous-phase and photochemical processing on secondary organic aerosol formation and evolution in Beijing, China. *Environ. Sci. Technol.* **51**, 762–770 (2017).
48. Duan, J. et al. Distinctions in source regions and formation mechanisms of secondary aerosol in Beijing from summer to winter. *Atmos. Chem. Phys.* **19**, 10319–10334 (2019).
49. Sun, Y. et al. Primary and secondary aerosols in Beijing in winter: sources, variations and processes. *Atmos. Chem. Phys.* **16**, 8309–8329 (2016).
50. Hu, W. et al. Chemical composition, sources and aging process of submicron aerosols in Beijing: contrast between summer and winter. *J. Geophys. Res. Atmos.* **121**, 1955–1977 (2016).
51. Nie, W. et al. Secondary organic aerosol formed by condensing anthropogenic vapours over China's megacities. *Nat. Geosci.* **15**, 255–261 (2022).
52. Wang, Y. et al. Molecular composition of oxygenated organic molecules and their contributions to organic aerosol in Beijing. *Environ. Sci. Technol.* **56**, 770–778 (2022).
53. Desyaterik, Y. et al. Speciation of 'brown' carbon in cloud water impacted by agricultural biomass burning in eastern China. *J. Geophys. Res. Atmos.* **118**, 7389–7399 (2013).
54. Shrivastava, M. et al. Urban pollution greatly enhances formation of natural aerosols over the Amazon rainforest. *Nat. Commun.* **10**, 1046 (2019).
55. An, Z. et al. Severe haze in northern China: a synergy of anthropogenic emissions and atmospheric processes. *Proc. Natl Acad. Sci. USA* **116**, 8657–8666 (2019).
56. Wu, J. et al. Contributions of trans-boundary transport to summertime air quality in Beijing, China. *Atmos. Chem. Phys.* **17**, 2035–2051 (2017).
57. Wang, Z. et al. Modeling study of regional severe hazes over mid-eastern China in January 2013 and its implications on pollution prevention and control. *Sci. China Earth Sci.* **57**, 3–13 (2013).
58. Xing, J. et al. Quantifying nonlinear multiregional contributions to ozone and fine particles using an updated response surface modeling technique. *Environ. Sci. Technol.* **51**, 11788–11798 (2017).
59. Huang, X. F. et al. Highly time-resolved chemical characterization of atmospheric submicron particles during 2008 Beijing Olympic Games using an Aerodyne High-Resolution Aerosol Mass Spectrometer. *Atmos. Chem. Phys.* **10**, 8933–8945 (2010).
60. Huang, X. et al. Enhanced secondary pollution offset reduction of primary emissions during COVID-19 lockdown in China. *Natl Sci. Rev.* **8**, nwaa137 (2021).
61. Zhou, W. et al. Unexpected increases of severe haze pollution during the post COVID-19 period: effects of emissions, meteorology and secondary production. *J. Geophys. Res. Atmos.* **127**, e2021JD035710 (2022).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024

¹Aerosol and Haze Laboratory, Beijing Advanced Innovation Center for Soft Matter Science and Engineering, Beijing University of Chemical Technology, Beijing, China. ²Institute for Atmospheric and Earth System Research, Faculty of Science, University of Helsinki, Helsinki, Finland. ³PSI Center for Energy and Environmental Sciences, Paul Scherrer Institute, Villigen, Switzerland. ⁴Department of Meteorology (MISU) and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden. ⁵Nanjing-Helsinki Institute in Atmospheric and Earth System Sciences, Nanjing University, Suzhou, China. ⁶Shanghai Key Laboratory of Atmospheric Particle Pollution and Prevention (LAP3), Department of Environmental Science and Engineering, Fudan University, Shanghai, China. ⁷Institute for Atmospheric and Environmental Sciences, Goethe University Frankfurt, Frankfurt, Germany. ⁸Institute of Materials Chemistry, TU Wien, Vienna, Austria. ⁹Centre National de la Recherche Scientifique (CNRS), Institut de Recherche pour le Développement (IRD), Institute of Engineering and Management Univ. Grenoble Alpes (Grenoble INP), Institut des Géosciences de l'Environnement (IGE), Université Grenoble Alpes, Grenoble, France. ¹⁰Aerodyne Research Inc., Billerica, MA, USA. ¹¹Center for Atmospheric Particle Studies, Carnegie Mellon University, Pittsburgh, PA, USA. ✉e-mail: kaspar.daellenbach@psi.ch; cai.jing@helsinki.fi; markku.kulmala@helsinki.fi

Methods

Measurement site

The sampling site in Beijing is located near the west 3rd Ring Road situated on the west campus of Beijing University of Chemical Technology (BUCT; 39° 56' 31" N, 116° 17' 50" E). The observatory is located on the top floor of a five-storey building (~20 m above ground level). The station is surrounded by residential areas with possible local emissions. Overall, the station represents a typical urban residential location in Beijing^{62,63}. Between November 2019 and July 2020, a detailed chemical characterization of PM_{2.5} was performed.

Chemical characterization of PM_{2.5}

A ToF-ACSM set-up equipped with a PM_{2.5} lens and a standard vaporizer^{64,65} was used to quantitatively characterize the dry (Nafion dryer Perma Pure, MD-700-24F-3) non-refractory PM_{2.5} content and its bulk constituents (organic aerosol (OA), nitrate (NO₃), sulfate (SO₄), chloride (Cl) and ammonium (NH₄)). The aerosol mass spectrometer and ToF-ACSM provides chemical fingerprints of OA that are widely used for source apportionment^{3,66–68}. However, the measurement principle (vaporization at 600 °C, electron impact ionization) causes strong fragmentation of the organic compounds, leading to the detection of small fragment ions instead of molecular ions. Thus, information, in particular on SOA, is largely lost, although information on POA sources can be well retrieved. The data analysis flow is detailed in ref. 62. The relative ionization efficiencies (RIE) were 4.0 for NH₄, 0.86 for SO₄ and 1.5 for Cl (for OA a default of 1.4 was used). A composition-dependent collection efficiency (CE) was determined and used to correct the data⁶⁹. A collocated seven-wavelength, dual-spot aethalometer (AE33, Magee Scientific Corp.) was used to measure the concentration of eBC⁷⁰. The chemically resolved PM_{2.5} was compared to total PM_{2.5} data from the surrounding monitoring stations, and was found to be in good agreement (Supplementary Fig. 10). During the ToF-ACSM downtime, NH₄, SO₄, NO₃ and Cl were gapfilled using measurements from a monitor for aerosols and gases in ambient air (MARGA, 2060R, Metrohm Process Analytics). Based on the bulk chemical composition, particle LWC values were computed using ISORROPIA⁷¹. In addition, PM_{2.5} values were collected on preheated quartz fibre filter samples using a HiVol sampler (24 h) between February 2018 and March 2019 (stored at –20 °C). Based on the water extracts of these filters, levoglucosan was quantified using high-performance liquid chromatography with the pulsed amperometric detector method^{72,73}. MBTCA and pinic acid were quantified based on an external calibration with self-synthesized standards. Extraction was carried out with acetonitrile/water (50/50 vol/vol) in an orbital shaker for 20 min (in two steps with 250 µl and 150 µl). The combined extracts were separated by ultrahigh performance liquid chromatography (UHPLC, Vanquish Flex, Thermo Fisher Scientific) on a C18 column (Accucore 150 × 2.1 mm, 2.6-µm particle size, Thermo Fisher Scientific), and the compounds were ionized by heated electrospray ionization (operated in negative polarity) on an Orbitrap mass spectrometer (Q Exactive Focus hybrid mass spectrometer, Thermo Fisher Scientific). The eluents used were as follows: A, ultrapure water with 0.1% formic acid (vol/vol); B, acetonitrile with 0.1% formic acid (vol/vol). The gradient was set as follows: starting with 1% B for 2 min, increasing to 99% B within 13 min, holding for 2 min, decreasing to 1% B within 1 min, followed by 2 min for re-equilibration. Extracted ion chromatograms (±4 ppm) of the molecular ions ([M – H][–]) were used for peak integration.

Near-molecular OA characterizations using FIGAERO-CIMS

The near-molecular composition of PM_{2.5} was characterized online with FIGAERO-CIMS¹³. FIGAERO-CIMS uses soft chemical ionization, allowing for the detection of molecular ions and their chemical formulae, but not their structure. With this approach, it is able to detect a wide range of organic compounds in widely different OA types including biogenic laboratory SOA (isoprene, IEPOX, different monoterpenes

and sesquiterpenes)^{20,22,74,75}, laboratory anthropogenic SOA from pure components (such as catechol, trimethylbenzene, methylbenzene and toluene)^{23,76}, cooking POA in indoor settings^{24,77} and ambient biomass smoke²⁵, as well as complex indoor⁷⁸ and ambient OA^{26,27,79–83}. Although not all compounds in OA are detected^{13,28,29}, compared to previously used tracer-based approaches, a much larger OA mass fraction is represented, allowing for unprecedented assessment of the main sources of SOA. Thermal evaporation can result in the fragmentation of labile organic compounds, which need then to be interpreted with caution. The FIGAERO-CIMS alternates automatically between collecting PM_{2.5} on a polytetrafluoroethylene filter (Zefon International, 25-mm diameter, 1-µm pore size) and chemically analysing the collected PM_{2.5}. Post collection, the filter is moved to the desorption port, where it remains for 2 min before initiating the heating phase (dry synthetic air from a pure air generator, desorption flow = 2.3 l min^{–1}). Subsequently, the filter is thermally desorbed at a heating rate of 11 °C min^{–1} for 15 min from room temperature (25–27 °C) to the maximum temperature of 190–194 °C. Afterwards, the filter is soaked at the highest set temperature for 15 min and then finally cooled to room temperature. After this first heating cycle, the filter was directly (without additional exposure) subjected to an identical second desorption cycle to measure the background signal without added particles. The vapour resulting from the desorption was ionized by the addition of iodide (generated from methyl iodide subjected to an X-ray source, with an ion-molecule reaction region (IMR) pressure of 300 mbar). Finally, the ions were analysed by a long ToF mass spectrometer (LToF-MS, *m/z* calibration within 2 ppm, mass resolving power *m/Δm* – 9,000 – 11,000, Supplementary Fig. 11). Field blank measurements were performed by removing the particles from the PM_{2.5} collection stream using a high-efficiency particulate air filter.

Data processing. FIGAERO-CIMS data were analysed by Tofware 3.1.0. The mass spectral data [X][–] were corrected for fluctuations in the reagent ion ([I][–]) as suggested by ref. 84:

$$[X]_{I_c}^{-} = \ln \left(1 + \frac{[X]^{-}}{[I]^{-}} \right) \quad (1)$$

In the presence of a high total ion current from ions other than the reagent ion ([I][–]), the signal of the contamination peak (C₄H₅ClO₃I[–], [ContI][–]) appears to be suppressed (that is, a lower signal during the first desorption cycle than during the second; Supplementary Fig. 12). Thus, in the second stage, the analyte concentrations were further corrected by this peak's signal ratio of the desorption cycles $\left(\frac{[ContI]_{I_c, des2}^{-}}{[ContI]_{I_c, des1}^{-}} \right)$:

$$[X]_{I+c}^{-} = [X]_{I_c, des1}^{-} \times \frac{[ContI]_{I_c, des2}^{-}}{[ContI]_{I_c, des1}^{-}} \quad (2)$$

Furthermore, a thermal baseline was computed for each peak for the first and second desorption cycle, which was subtracted from the respective desorption cycle:

$$[X]_{I+c, bsl_c}^{-} = [X]_{I+c}^{-} - [X]_{I+c, bsl}^{-} \quad (3)$$

The thermal baseline was computed based on an algorithm presented in ref. 85. An example of a thermal baseline is presented in Supplementary Fig. 13.

The difference between the first and second desorption was computed:

$$[X]_{I+c, bsl_c, bkg_c}^{-} = [X]_{I+c, bsl_c, des1}^{-} - [X]_{I+c, bsl_c, des2, avg}^{-} \quad (4)$$

where $[X]_{I+c, bsl_c, des2, avg}^{-}$ is the average of the second desorption before and after $[X]_{I+c, bsl_c, des1}^{-}$.

Finally, $[XI^-]_{I+C_c, bsl_c, bkg_c}$ was integrated through the entire thermogram and normalized to the amount of air sampled during the preceding collection period (V_{air}):

$$[XI^-]_{I+C_c, bsl_c, bkg_c}^s = \sum_{n_{thermo}} [XI^-]_{I+C_c, bsl_c, bkg_c} dt \times \frac{1}{V_{air}} \quad (5)$$

The field blank measurements were interpolated based on a parametrization for each peak that links the ratio of the filter measurement to the previous and following ambient measurements ('blank fraction') to the filter loading of the respective peak (proportional to $[XI^-] \times V_{air}$) through an exponential relationship (example in Supplementary Fig. 14):

$$\frac{[XI^-]_{I+C_c, bsl_c, bkg_c}^s(\text{blank})}{[XI^-]_{ambient}} = y_0 + A \times \exp\left(-\frac{[XI^-]_{ambient} \times V_{air}}{\tau}\right) \quad (6)$$

with $[XI^-]_{ambient}$ estimated as the average of the ambient sample before and after the blank measurement. This fit was bootstrapped, leading to 100 estimates of y_0 , A and τ . Using these 100 combinations of the fitting parameters, 100 blank concentrations were computed for each peak $[XI^-]$ at each point in time, with the average used as the best estimate of the blank concentration:

$$[XI^-]_{I+C_c, bsl_c, bkg_c, bl}^s = y_0 + A \times \exp\left(-\frac{[XI^-]_{I+C_c, bsl_c, bkg_c} \times V_{air}}{\tau}\right) \times [XI^-]_{I+C_c, bsl_c, bkg_c}^s \quad (7)$$

This best estimate of the field blank concentration at each time point was subsequently subtracted:

$$[XI^-]_{I+C_c, bsl_c, bkg_c, bl_c}^s = [XI^-]_{I+C_c, bsl_c, bkg_c}^s - \text{avg}([XI^-]_{I+C_c, bsl_c, bkg_c, bl}^s) \quad (8)$$

Previous publications suggest that estimates based on the counting uncertainty underestimate the measurement uncertainty of FIGAERO-CIMS^{13,17}. Accordingly, we based the uncertainty estimates here on the repeatability measure for raw FIGAERO-CIMS thermograms (10%)¹³. After that, the uncertainty introduced by any additional computational step was estimated and propagated to the uncertainty of the raw thermograms. Peaks with poor signal-to-noise ratios were excluded from further analysis. The uncertainty of the blank was further propagated and estimated as the quadratic sum of the uncertainty related to the measurement of $[XI^-]_{I+C_c, bsl_c, bkg_c}^s$ scaled with the 'blank fraction' and the uncertainty of the 'blank fraction' scaled with $[XI^-]_{I+C_c, bsl_c, bkg_c}^s$.

Finally, the mass spectral matrix and the uncertainty matrix were multiplied by the molecular weight of the ion (excluding the weight of I^-). For data mining, all peaks that were not associated to a $C_xH_yO_zN_sI$ analyte iodide cluster were discarded. Finally, the signal of each organic compound (and related uncertainty) was expressed as levoglucosan-equivalent concentration by multiplication with a levoglucosan calibration factor (levoglucosan spike on a filter; Supplementary Fig. 2). We thus assume that all peaks have the same response factor as that of levoglucosan. Note that the FIGAERO-CIMS OA sources are quantified through an MLR approach relying on the ToF-ACSM OA concentrations (Supplementary Section 'Quantification of FIGAERO-CIMS PMF factors'). Using the data analysis framework described here, the FIGAERO-CIMS OA correlates well with OA (here OA minus (HOA + COA)) measured by the ToF-ACSM ($R = 0.87$). In addition, tracers of nitrate (HNO_3I , $R = 0.91$) and sulfate (SO_3I , $R = 0.85$) also correlate well with the quantities measured by ToF-ACSM. This is in line with a previous offline filter-based study showing similarities in the temporal behaviour between FIGAERO-CIMS and ToF-ACSM^{27,86}.

Source apportionment analysis

Source apportionment was performed using the positive matrix factorization algorithm (PMF)¹⁹ as implemented in the multilinear engine 2 (ME-2)⁸⁷ and controlled by the Source Finder interface (SoFi)^{88,89}. PMF is a statistical unmixing model widely used in atmospheric aerosol science. In this study, OA source apportionment analyses were performed independently for the ToF-ACSM and FIGAERO-CIMS OA characterizations.

ToF-ACSM OA source apportionment. ToF-ACSM OA source apportionment analysis relied on the in situ observations at BUCT (1-h averages) as well as on a priori information from the literature. Previous ToF-ACSM-based studies highlight that OA in Beijing is affected by a multitude of sources^{3-5,62,90}: HOA (related to traffic exhaust emissions and in general liquid-fossil-fuel combustion), COA, BBOA, CCOA, and a varying number of OOA components related to SOA. Exploratory analyses of the present dataset showed mixed sources, so we used the mass spectral signatures of specific POA sources to improve their separation (HOA, COA, BBOA and CCOA). In practice, the chemical OA fingerprints of hydrocarbon-like POA⁹¹, COA⁹¹, BBOA⁴ and CCOA⁴ were constrained as a priori information. In a preliminary PMF run with six factors covering data from 2018 to 2020, HOA and COA were constrained with an α -value, that is, a tolerated relative deviation from the anchor, of 0.1 and BBOA as well as CCOA with 0.2 up to m/z 85 (ramping up to an α -value of 1 at m/z 102; at higher m/z values, any value between 0 and 0.014 for BBOA and 0.016 for CCOA, respectively, was allowed) (initial guesses being the maximum divided by two). This was used to gapfill missing information on the BBOA and CCOA chemical composition at $m/z > 115$. In further analyses, these resulting mass spectra for BBOA and CCOA were used as constraints (α -value varied between 0 and 0.4 with an increment of 0.2) in addition to the OA chemical fingerprints for HOA and COA from ref. 91. (The α -value varied between 0 to 0.2 with an increment of 0.1.) We assessed the mathematical quality of the PMF solution based on the PMF residuals (res) normalized to the measurement uncertainty (σ)⁸⁸. First, an overview parameter Q/Q_{exp} was computed for each solution ($Q = \sum_{i=1}^m \sum_{j=1}^n \left(\frac{res_{i,j}}{\sigma_{i,j}}\right)^2$) and $Q_{exp} = n \times m - p \times (m + n)$, where n is the number of time points, m the number of ions, and p the number of factors). Although Q/Q_{exp} decreased by 32% when increasing the number of factors from 5 to 6, the decrease was clearly smaller when including a seventh factor (18%). This is in line with the more temporal structure in the measurement error-weighted PMF residuals for five than six or seven factors (Supplementary Fig. 15). Although the six-factor solution resolved two OOA components with different chemical fingerprints, these factors were mathematically further split when including a seventh factor. Because the results were not improved by including a seventh factor, here we present a six-factor solution (HOA, COA, BBOA, CCOA and two free factors). For the final source apportionment analysis, 28 day chunks of data (for each 81 PMF runs with randomly chosen α -values, see above) were used at a time (shifted by two days) until the entire study period was covered (rolling PMF^{32,89,92-94}). The free/unconstrained factors were sorted based on their fractional content of m/z 44 (f44). For each window, the correlation coefficient ($R_{pearson}$) between eBC and the combustion-related factors was computed as $eBC = a \times HOA + b \times CCOA + c \times BBOA$. The lower threshold for solution acceptance was based on three median absolute deviations from the median $R_{pearson}$. For periods during which the FIGAERO-CIMS was operating but not the ToF-ACSM, we interpolated the approximate HOA concentrations based on the eBC concentration ($HOA_{approx} = 0.37 \times eBC$), and COA was approximated based on a parametrization of the COA/HOA ratio ($COA_{approx} = (\text{base} + \frac{(\text{max} - \text{base})}{1 + (\frac{x_{half}}{HOA_{approx}})^{rate})} \times HOA_{approx}$; base = 174.9, max = 1.1, rate = 0.9, xhalf = 0.0023; Supplementary Fig. 16). In a last step, we assumed, based on ref. 95, that HOA and COA have a response factor

(RIE $\times$ CE) between two and three times larger than the default RIE of OA (RIE_{OA, default} = 1.4, CE = 0.5) leading to an RIE $\times$ CE of 1.5–2 (RIE of HOA and COA of 3.5). The chemical fingerprints and time series of the resolved OA components, as well as their markers, are presented in Supplementary Figs. 17–19 (scatterplots between selected OA components and their markers are displayed in Supplementary Fig. 20). Both ToF-ACSM OOA components correlate with all FIGAERO-CIMS SOA factors (Supplementary Fig. 21). However, one OOA component correlates better with FIGAERO-CIMS sfSOA, aromSOA_{night} and SOA_{aq}, whereas the other correlates better with bioSOA_{night}.

Particle-phase FIGAERO-CIMS OA source apportionment. Given the lack of well-understood near-molecular source compositions, the FIGAERO-CIMS source apportionment analysis did not rely on a priori information. Q/Q_{exp} is reduced when increasing the number of factors and the selected eight-factor solution has a Q/Q_{exp} of 0.9. When introducing more than nine factors, Q/Q_{exp} does not decrease substantially more (<5%). With more detailed assessments based on the change in the measurement uncertainty-weighted PMF residuals, we found that the PMF explains the data increasingly better when increasing the number of factors up to eight, but there is no further substantial improvement beyond nine factors (Supplementary Fig. 22). We thus examined PMF solutions with up to nine factors, presented here though an eight-factor solution given its best environmental interpretability. We performed a sensitivity analysis based on 200 bootstrapping PMF runs. The factors were identified based on their time-series correlation (R_{pearson}) with a base case. The lower threshold for solution acceptance was based on three median absolute deviations from the median R_{pearson} . The chemical compositions of the factors are presented in Fig. 2 and Extended Data Figs. 6 and 7, and the temporal variation in Fig. 2, Extended Data Fig. 6 and Supplementary Fig. 23.

Quantification of FIGAERO-CIMS PMF factors. Because the sensitivity of FIGAERO-CIMS depends on the chemical composition of a molecule, the response factors for the identified FIGAERO-CIMS factors cannot be assumed to be the same. We aimed to determine a response factor for each FIGAERO-CIMS factor for improved quantification. To that end, we performed MLR relating the FIGAERO-CIMS factor time series to the ToF-ACSM data (equation (9)). Plausibly, the FIGAERO-CIMS cannot efficiently measure the hydrocarbon components that constitute the major parts of traffic and cooking emissions. Yet, it has been shown that other constituents of cooking emissions can be detected by FIGAERO-CIMS²⁴. In the present dataset we did not identify an HOA or COA factor based on the FIGAERO-CIMS data; for example, no FIGAERO-CIMS factor shows a lunch- and dinner-time peak characteristic of COA (ToF-ACSM COA; Supplementary Fig. 19). On the other hand, HOA and COA can be well quantified by the ToF-ACSM data (section ‘ToF-ACSM OA source apportionment’). We thus subtracted HOA and COA (estimated via ToF-ACSM OA analyses) from OA and used this difference (OA(t) minus (HOA(t) – COA(t))) as the reference concentration ($\mu\text{g m}^{-3}$) for the MLR to quantify the FIGAERO-CIMS factors (absolute signal), instead of using the entire ToF-ACSM OA:

$$\begin{aligned} & \text{OA}(t) - (\text{HOA}(t) + \text{COA}(t)) \\ &= \text{normal} \left(\sum_k \frac{\text{OA}_k^{\text{FIGAERO}}(t)}{R_k^{\text{FIGAERO}}}, (\text{OA}(t) - \text{HOA}(t) - \text{COA}(t)) \times \text{rel}_{\text{err}} \right) \end{aligned} \quad (9)$$

$\text{OA}_k^{\text{FIGAERO}}(t)$ represents the time series of FIGAERO-CIMS OA factors, R_k^{FIGAERO} their response factors, and rel_{err} the approximate uncertainty of OA minus (HOA + COA), which is assumed to be 10%. MLR was performed with the advanced statistical software STAN⁹⁶ via Hamiltonian Markov chain Monte Carlo sampling from the posterior distribution of our model given the data (equation (9)). Thereby, R_k combinations were drawn in proportion to their posterior probability. Their variability thus provides direct uncertainty estimates of R_k .

In a preliminary analysis, the sum of SFOA_{day} and SFOA_{night} from the FIGAERO-CIMS (FIGAERO-SFOA) correlates well with the sum of BBOA and CCOA from the ToF-ACSM (ToF-ACSM-SFOA). We thus assume that FIGAERO-SFOA is the same as ToF-ACSM-SFOA. However, the preliminary quantification approach overall results in higher FIGAERO-SFOA concentrations compared to the ToF-ACSM-SFOA ($\text{SFOA}_{\text{day}}^{\text{FIGAERO}} + \text{SFOA}_{\text{night}}^{\text{FIGAERO}} = 1.6 \times [\text{BBOA} + \text{CCOA}]$). During the summer, SFOA from FIGAERO-CIMS is considerably lower than BBOA + CCOA from ToF-ACSM. During this period, BBOA dominates the sum of BBOA and CCOA. Plausibly, summertime BBOA is mixed with other OA components because in previous studies BBOA could not be identified during summer⁶². We thus optimized SFOA_{FIGAERO} such that FIGAERO-SFOA was similar to the ToF-ACSM-SFOA concentration for November 2019 to March 2020 (with an assumed relative error of 5%– $\text{rel}_{\text{err}2}$; equation (10)):

$$\begin{aligned} & \text{mean}_{\text{Nov-Mar}} (\text{BBOA}(t) + \text{CCOA}(t)) \\ &= \text{normal} \left(\text{mean}_{\text{Nov-Mar}} \left(\frac{\text{SFOA}_{\text{day}}^{\text{FIGAERO}}(t)}{R_{\text{SFOA day}}^{\text{FIGAERO}}} + \frac{\text{SFOA}_{\text{night}}^{\text{FIGAERO}}(t)}{R_{\text{SFOA night}}^{\text{FIGAERO}}} \right), \sigma \right), \quad (10) \\ & \sigma = \text{mean}_{\text{Nov-Mar}} (\text{BBOA}(t) + \text{CCOA}(t)) \times \text{rel}_{\text{err}2} \end{aligned}$$

In Supplementary Fig. 24 we compare the response factors of the FIGAERO-CIMS OA components to a bulk OA response factor (R_{bulk}) assuming that all components have the same response factor (via linear regression between the sum of all FIGAERO-CIMS OA components and OA from the ToF-ACSM). OA minus (HOA + COA) modelled by the FIGAERO factors, corrected with their respective response factors, reconstructs OA minus (HOA + COA) measured by the ToF-ACSM (slope = 0.87, $R = 0.87$; Supplementary Fig. 25), and the MLR residuals do not depend on the (HOA + COA) fraction of OA. With that approach we found the FIGAERO-CIMS response factors to the eight OA sources vary with a factor of ~ 10 (Supplementary Fig. 24). We note that this is an approach to estimate the concentrations of the different OA sources/factors, not of the single molecules therein. The range of estimated bioSOA concentrations is similar to estimates from a tracer-based approach using MBTCA and pinic acid (oxidation products of terpenes) from offline filter analyses from 2018/2019 (assuming a bioSOA/[MBTCA + pinic acid] ratio of 0.126 $\mu\text{g bioSOA}/\text{ng} [\text{MBTCA} + \text{pinic acid}]$ ⁷ (Extended Data Fig. 8).

We performed a Monte Carlo sensitivity assessment on the RIEs of HOA, COA and SFOA (BBOA, CCOA), as well as the relative response factors of the FIGAERO OA factors, with a total of 3,600 runs (Supplementary Fig. 24). We designed three RIE scenarios and accounted for the uncertainties of both the respective RIEs (POAs) and response factors (all FIGAERO OA factors; Supplementary Fig. 24):

1. Default RIEs for POA: RIE_{HOA} = 1.4 ± 0.3 , RIE_{COA} = 1.4 ± 0.3 , RIE_{SFOA} = 1.4 ± 0.3
2. Adapted RIEs for HOA and COA: RIE_{HOA} = 3.5 ± 0.5 , RIE_{COA} = 3.5 ± 0.5 , RIE_{SFOA} = 1.4 (ref. 95)
3. Adapted RIEs for POA: RIE_{HOA} = 5.74 ± 1.95 , RIE_{COA} = 4.55 ± 1.55 , RIE_{SFOA} = (4.44 ± 1.51 or 5.55 ± 1.89) (ref. 97).

The results illustrate that the relative contribution of the different SOA components is subject to uncertainty, but that this does not affect our findings and conclusions (Extended Data Figs. 4 and 5).

In addition, we compared the MLR-based quantified FIGAERO-CIMS OA factors (equations (9) and (10)) to (1) levoglucosan-equivalent concentrations assuming the same response factor for all compounds and (2) a direct quantification of the FIGAERO-CIMS OA factors (Extended Data Figs. 4 and 5). For the latter direct quantification approach, we relied on our own levoglucosan calibration together with a parametrization of the response factors relative to levoglucosan (rRF) as a function of m/z , ($\text{rRF}(m/z) = a + b \times (\exp(-(\frac{m/z-c}{d})^2))$), based on data

from the literature²⁸ (central estimate: $a = 0.3$, $b = 0.66$, $c = 249.17$, $d = 87.95$; upper limit: $a = 0.33$, $b = 0.86$, $c = 235.10$, $d = 115.37$; lower limit: $a = 0.23$, $b = 0.54$, $c = 259.89$, $d = 63.77$).

Characterization of gas-phase OOM

The gas-phase oxygenated organic molecules (OOM) were characterized by a nitrate-based CIMS (NO_3 -CIMS, here equipped with an LTOF mass spectrometer with a mass resolution of 8,000–12,000)^{98,99}. The configuration and calibration of this instrument have been described previously^{100,101}. In brief, the NO_3 -CIMS was calibrated with a known amount of sulfuric acid¹⁰² and the OOM assumed to have the same response because their structures are unknown and thus cannot be calibrated for. The OOM concentrations were then computed by normalization to the reagent ions and subsequent scaling with the response factor.

The OOM related to isoprene oxidation and monoterpene oxidation were identified based on a decision tree designed using atmospheric and laboratory experiments^{44,51}. Essentially, the isoprene oxidation products were identified based on a list of compounds presented in the literature^{44,51}, and the monoterpene oxidation products were identified as compounds with a carbon number of 10, an equivalent oxygen number of at least 4, and a double bond equivalent between 2 and 4.

Spatial distribution of emissions

We used a concentration-weighted trajectory (CWT) method to study the spatial distribution of the precursor sources for the different factors. Instead of trajectories, we used potential emission sensitivity (PES) fields, which were calculated using a Lagrangian particle dispersion model FLEXPART version 9.02 (ref. 103) with European Centre for Medium-Range Weather Forecasts (ECMWF) operational forecast data (0.15° horizontal resolution, 137 vertical levels and 1-h temporal resolution) as the meteorological input. In the FLEXPART model simulations, 50,000 tracer particles were initially distributed evenly between 0 and 100 m above the measurement site and then followed backwards in time for 72 h. The output PES fields (domain: 20–60° N 95–135° E; horizontal resolution, 0.05°) contain the residence times of the air mass (tracer particles) above the simulation grid cells. In the CWT method, we assigned each grid cell with a concentration value (C_{ij}), which represents the expected concentration at the measurement site if an air mass passes over said grid cell upon its arrival to the station. The C_{ij} value was calculated based on the observed concentrations (C_t) and the air mass residence times (τ):

$$C_{ij} = \frac{\sum_t C_t \tau_{tij}}{\sum_t \tau_{tij}} \quad (11)$$

where C_t is the observed concentration at time t , and τ_{tij} is the residence time of the air mass over the ij th grid cell obtained from the PES field for a tracer release at time t . Because the precursor emissions for the different factors are ground-based, we used PES fields that only include the residence times of the tracer particles residing within 500 m above ground level. In addition, C_{ij} values are only shown for grid cells containing data from more than ten different observation times.

Data availability

The full dataset used in the figures is publicly available at <https://doi.org/10.5281/zenodo.10977390> (ref. 104).

Code availability

Source Finder (SoFi) is available via its distributor Datalystica Ltd (www.datalystica.com). Tofware is available via its distributors Tofwerk AG and Aerodyne Inc. FLEXPART is available via flexpart.eu. Mapping Tools for MATLAB, used for plotting the maps for the positioning data, are available from MathWorks File Exchange (<https://www.mathworks.com/>).

References

62. Cai, J. et al. Size-segregated particle number and mass concentrations from different emission sources in urban Beijing. *Atmos. Chem. Phys.* **20**, 12721–12740 (2020).
63. Liu, Y. et al. Continuous and comprehensive atmospheric observations in Beijing: a station to understand the complex urban atmospheric environment. *Big Earth Data* **4**, 295–321 (2020).
64. Fröhlich, R. et al. The ToF-ACSM: a portable aerosol chemical speciation monitor with TOFMS detection. *Atmos. Meas. Tech.* **6**, 3225–3241 (2013).
65. Xu, W. et al. Laboratory characterization of an aerosol chemical speciation monitor with $\text{PM}_{2.5}$ measurement capability. *Aerosol Sci. Technol.* **51**, 69–83 (2017).
66. Crippa, M. et al. Organic aerosol components derived from 25 AMS data sets across Europe using a consistent ME-2 based source apportionment approach. *Atmos. Chem. Phys.* **14**, 6159–6176 (2014).
67. Jimenez, J. L. et al. Evolution of organic aerosols in the atmosphere. *Science* **326**, 1525–1529 (2009).
68. Zhang, Q. et al. Understanding atmospheric organic aerosols via factor analysis of aerosol mass spectrometry: a review. *Anal. Bioanal. Chem.* **401**, 3045–3067 (2011).
69. Middlebrook, A. M., Bahreini, R., Jimenez, J. L. & Canagaratna, M. R. Evaluation of composition-dependent collection efficiencies for the aerodyne aerosol mass spectrometer using field data. *Aerosol Sci. Technol.* **46**, 258–271 (2012).
70. Drinovec, L. et al. The ‘dual-spot’ Aethalometer: an improved measurement of aerosol black carbon with real-time loading compensation. *Atmos. Meas. Tech.* **8**, 1965–1979 (2015).
71. Nenes, A., Pandis, S. N. & Pilinis, C. ISORROPIA: a new thermodynamic equilibrium model for multiphase multicomponent inorganic aerosols. *Aquat. Geochem.* **4**, 123–152 (1998).
72. Waked, A. et al. Source apportionment of PM_{10} in a north-western Europe regional urban background site (Lens, France) using positive matrix factorization and including primary biogenic emissions. *Atmos. Chem. Phys.* **14**, 3325–3346 (2014).
73. Samaké, A. et al. Polyols and glucose particulate species as tracers of primary biogenic organic aerosols at 28 French sites. *Atmos. Chem. Phys.* **19**, 3357–3374 (2019).
74. Wu, C. et al. Photolytically induced changes in composition and volatility of biogenic secondary organic aerosol from nitrate radical oxidation during night-to-day transition. *Atmos. Chem. Phys.* **21**, 14907–14925 (2021).
75. Bell, D. M. et al. Particle-phase processing of α -pinene NO_3 secondary organic aerosol in the dark. *Atmos. Chem. Phys.* **22**, 13167–13182 (2022).
76. Wang, M. et al. Photo-oxidation of aromatic hydrocarbons produces low-volatility organic compounds. *Environ. Sci. Technol.* **54**, 7911–7921 (2020).
77. Reyes-Villegas, E. et al. Online chemical characterization of food-cooking organic aerosols: implications for source apportionment. *Environ. Sci. Technol.* **52**, 5308–5318 (2018).
78. Farmer, D. K. et al. Overview of HOMEChem: house observations of microbial and environmental chemistry. *Environ. Sci. Process. Impacts* **21**, 1280–1300 (2019).
79. Huang, W. et al. Seasonal characteristics of organic aerosol chemical composition and volatility in Stuttgart, Germany. *Atmos. Chem. Phys.* **19**, 11687–11700 (2019).
80. Huang, W. et al. Chemical characterization of highly functionalized organonitrates contributing to night-time organic aerosol mass loadings and particle growth. *Environ. Sci. Technol.* **53**, 1165–1174 (2019).
81. Mishra, S. et al. Rapid night-time nanoparticle growth in Delhi driven by biomass-burning emissions. *Nat. Geosci.* **16**, 224–230 (2023).

82. Thornton, J. A. et al. Evaluating organic aerosol sources and evolution with a combined molecular composition and volatility framework using the filter inlet for gases and aerosols (FIGAERO). *Acc. Chem. Res.* **53**, 1415–1426 (2020).
83. Lee, B. H. et al. Highly functionalized organic nitrates in the southeast United States: contribution to secondary organic aerosol and reactive nitrogen budgets. *Proc. Natl Acad. Sci. USA* **113**, 1516–1521 (2016).
84. Heinritzi, M. et al. Characterization of the mass-dependent transmission efficiency of a CIMS. *Atmos. Meas. Tech.* **9**, 1449–1460 (2016).
85. Wang, J. M. et al. Near-road air pollutant measurements: accounting for inter-site variability using emission factors. *Environ. Sci. Technol.* **52**, 9495–9504 (2018).
86. Cai, J. et al. Characterization of offline analysis of particulate matter with FIGAERO-CIMS. *Atmos. Meas. Tech.* **16**, 1147–1165 (2023).
87. Paatero, P. The multilinear engine—a table-driven, least squares program for solving multilinear problems, including the n-way parallel factor analysis model. *J. Comput. Graph. Stat.* **8**, 854–888 (1999).
88. Canonaco, F., Crippa, M., Slowik, J. G., Baltensperger, U. & Prévôt, A. S. H. SoFi, an IGOR-based interface for the efficient use of the generalized multilinear engine (ME-2) for the source apportionment: ME-2 application to aerosol mass spectrometer data. *Atmos. Meas. Tech.* **6**, 3649–3661 (2013).
89. Canonaco, F. et al. A new method for long-term source apportionment with time-dependent factor profiles and uncertainty assessment using SoFi Pro: application to 1 year of organic aerosol data. *Atmos. Meas. Tech.* **14**, 923–943 (2021).
90. Kulmala, M. et al. Is reducing new particle formation a plausible solution to mitigate particulate air pollution in Beijing and other Chinese megacities? *Faraday Discuss.* **226**, 334–347 (2021).
91. Crippa, M. et al. Wintertime aerosol chemical composition and source apportionment of the organic fraction in the metropolitan area of Paris. *Atmos. Chem. Phys.* **13**, 961–981 (2013).
92. Via, M. et al. Rolling vs. seasonal PMF: real-world multi-site and synthetic dataset comparison. *Atmos. Meas. Tech.* **15**, 5479–5495 (2022).
93. Chen, G. et al. Time-dependent source apportionment of submicron organic aerosol for a rural site in an alpine valley using a rolling positive matrix factorisation (PMF) window. *Atmos. Chem. Phys.* **21**, 15081–15101 (2021).
94. Chen, G. et al. Real-time source apportionment of organic aerosols in three European cities. *Environ. Sci. Technol.* **56**, 15290–15297 (2022).
95. Katz, E. F. et al. Quantification of cooking organic aerosol in the indoor environment using aerodyne aerosol mass spectrometers. *Aerosol Sci. Technol.* **55**, 1099–1114 (2021).
96. Carpenter, B. et al. Stan: a probabilistic programming language. *J. Stat. Softw.* **76**, 1–32 (2017).
97. Nault, B. A. et al. Laboratory evaluation of organic aerosol relative ionization efficiencies in the aerodyne aerosol mass spectrometer and aerosol chemical speciation monitor. *Aerosol Sci. Technol.* **57**, 981–997 (2023).
98. Eisele, F. L. & Tanner, D. J. Measurement of the gas phase concentration of H_2SO_4 and methane sulfonic acid and estimates of H_2SO_4 production and loss in the atmosphere. *J. Geophys. Res.* **98**, 9001–9010 (1993).
99. Jokinen, T. et al. Atmospheric sulphuric acid and neutral cluster measurements using CI-API-TOF. *Atmos. Chem. Phys.* **12**, 4117–4125 (2012).
100. Yan, C. et al. The synergistic role of sulfuric acid, bases, and oxidized organics governing new-particle formation in Beijing. *Geophys. Res. Lett.* **48**, e2020GL091944 (2021).
101. Yan, C. et al. The effect of COVID-19 restrictions on atmospheric new particle formation in Beijing. *Atmos. Chem. Phys.* **22**, 12207–12220 (2022).
102. Kürten, A., Rondo, L., Ehrhart, S. & Curtius, J. Calibration of a chemical ionization mass spectrometer for the measurement of gaseous sulfuric acid. *J. Phys. Chem. A* **116**, 6375–6386 (2012).
103. Stohl, A., Forster, C., Frank, A., Seibert, P. & Wotawa, G. Technical note: the Lagrangian particle dispersion model FLEXPART version 6.2. *Atmos. Chem. Phys.* **5**, 2461–2474 (2005).
104. Daellenbach, K. R. Data related to article: Substantial contribution of transported emissions to organic aerosol in Beijing [Data set]. Zenodo <https://doi.org/10.5281/zenodo.10977390> (2024).

Acknowledgements

We thank C. Mohr for fruitful scientific discussions. K.R.D. acknowledges support by SNSF mobility grant P2EZP2_181599 and SNSF Ambizione grant PZPGP2_201992. L.D. acknowledges support by SNSF Ambizione grant PZOP2_216181. S.H. acknowledges the Doctoral Programme in Atmospheric Sciences at the University of Helsinki (ATM-DP). J.C. acknowledges support by the Wihuri Foundation. Y.H. acknowledges support by the Swiss Agency for Development and Cooperation (grant no. 7F-09802.01.03) and SNSF grant 200020_188624. We acknowledge the following projects: ACCC Flagship funded by the Academy of Finland grant no. 337549, Academy professorship funded by the Academy of Finland (grant no. 302958), Academy of Finland projects nos 1325656, 311932, 316114, 332547 and 325647, the ‘Quantifying carbon sink, CarbonSink+ and their interaction with air quality’ INAR project funded by the Jane and Aatos Erkko Foundation, European Research Council (ERC) project ATM-GTP contract no. 742206 and the Wihuri Foundation, Technical and scientific staff at BUCT/AHL are acknowledged. Y.L. acknowledges support by the National Natural Science Foundation of China (42275117). N.M.D. acknowledges support from US NSF grant no. AGS-2132089. D.S. received funding from the Vienna Science and Technology Fund (WWTF) through project VRG22-003.

Author contributions

K.R.D. designed the study and analysed the data. M.K., C.Y., Y.L., F.B., K.R.D., L.D. and D.R.W. conceptualized the comprehensive long-term observations. K.R.D., J.C., C.Y., L.Y. and W.D. performed the FIGAERO-CIMS measurements. K.R.D., J.C., F.Z. and W.D. performed the ToF-ACSM measurements. J.-L.J., G.U., F.U., J.M. and A.L.V. performed offline filter measurements. Y.H. performed statistical multilinear regression modelling. S.H. performed Flexpart simulations. K.R.D., S.H. and L.D. interpreted the results. K.R.D., L.D., S.H., N.M.D. and D.S. wrote the original manuscript. All authors participated in relevant scientific discussions and commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

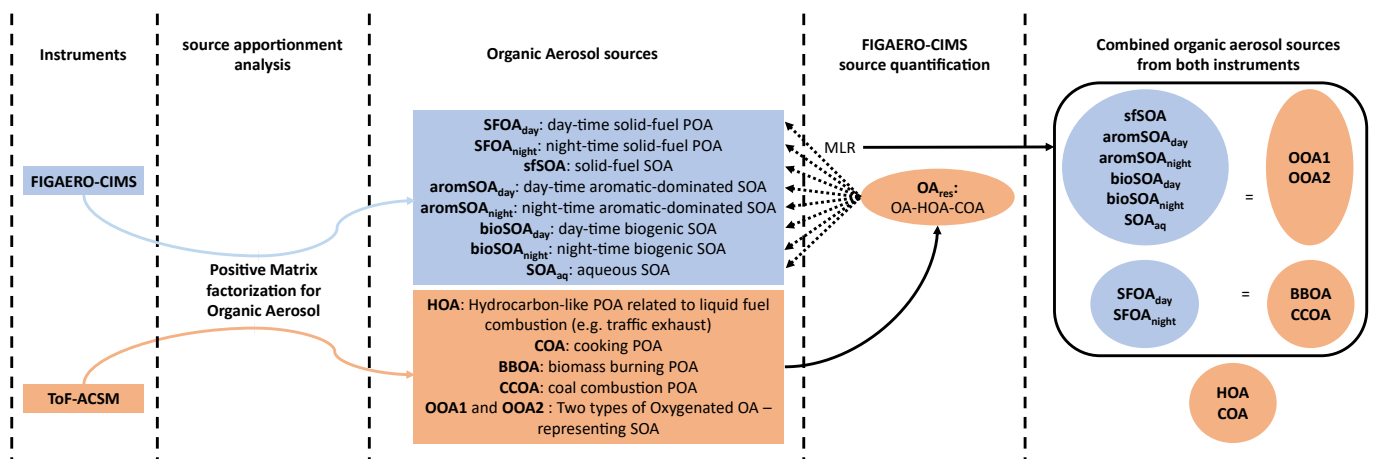
Extended data is available for this paper at <https://doi.org/10.1038/s41561-024-01493-3>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41561-024-01493-3>.

Correspondence and requests for materials should be addressed to Kaspar R. Daellenbach, Jing Cai or Markku Kulmala.

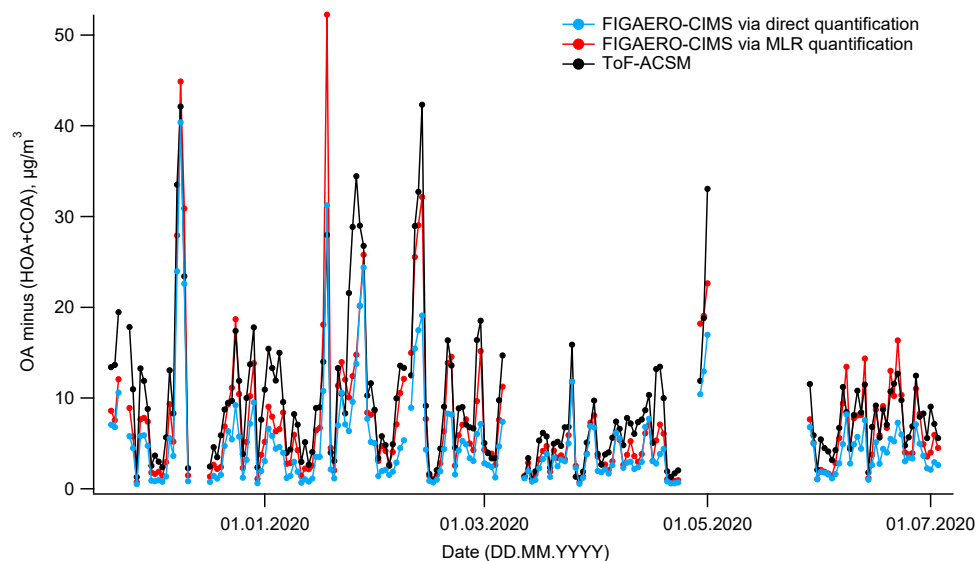
Peer review information *Nature Geoscience* thanks Hugh Coe and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Xujia Jiang, in collaboration with the *Nature Geoscience* team.

Reprints and permissions information is available at www.nature.com/reprints.

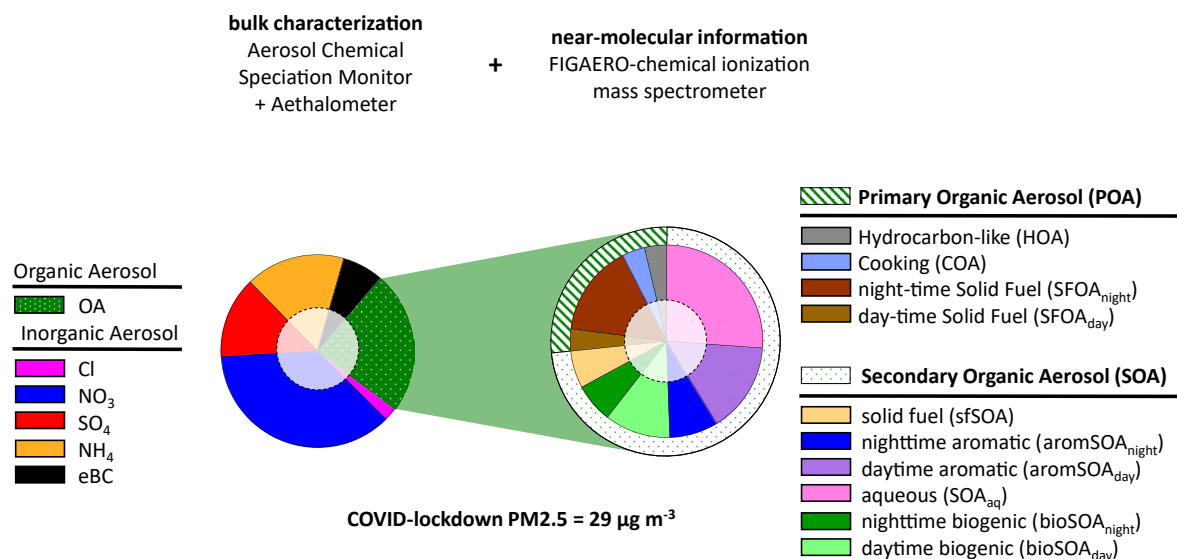


Extended Data Fig. 1 | Data analysis stream used in this study. Chemical characterizations of ToF-ACSM and FIGAERO-CIMS are used to quantify the sources of primary and secondary OA. We combine the quantitative OA ToF-ACSM analyses with a factorization of time series of near-molecular organic aerosol mass spectra determined by FIGAERO-CIMS. The ToF-ACSM is used to determine the organic mass and the contribution of primary organic aerosol (POA) from liquid fossil fuel combustion (HOA) and cooking (COA); the FIGAERO-CIMS is used to produce a set of SOA sources besides POA from

solid-fuel combustion, which are converted into mass loadings using a multilinear regression to the ToF-ACSM OA mass loadings, once some of the POA sources (HOA, COA) have been subtracted from it. Thereby, we obtain response factors of the FIGAERO-CIMS to the different SOA sources. Were these to be single compounds (instead of SOA sources), the response factors would be calibration factors but here they were applied across a wide range of compounds contained in a specific SOA source in this instance.

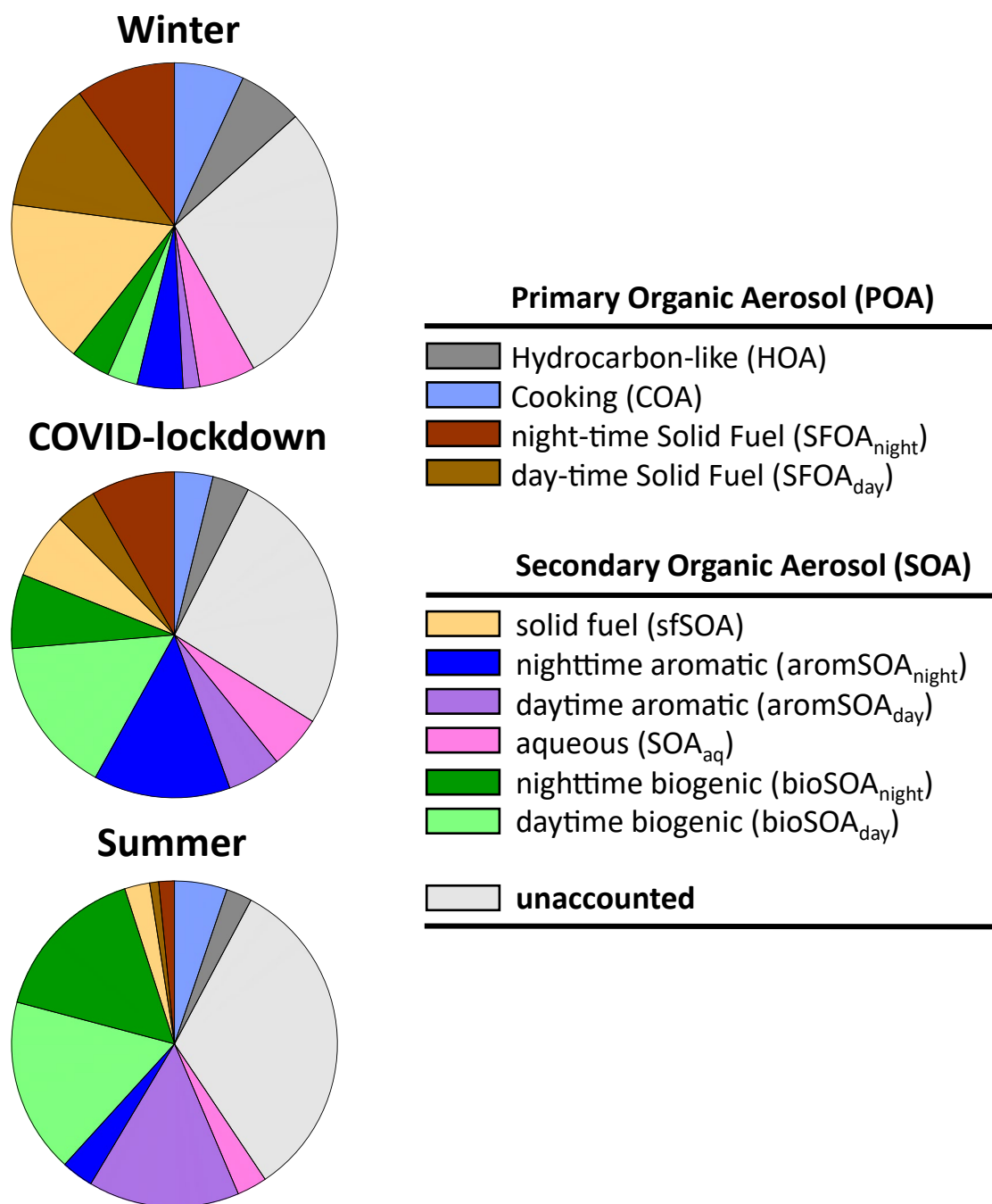


Extended Data Fig. 2 | Comparison between FIGAERO-CIMS and ToF-ACSM. Daily mean concentration time series of OA minus (HOA+COA) from FIGAERO-CIMS via MLR quantification and direct quantification (see methods) and ToF-ACSM.



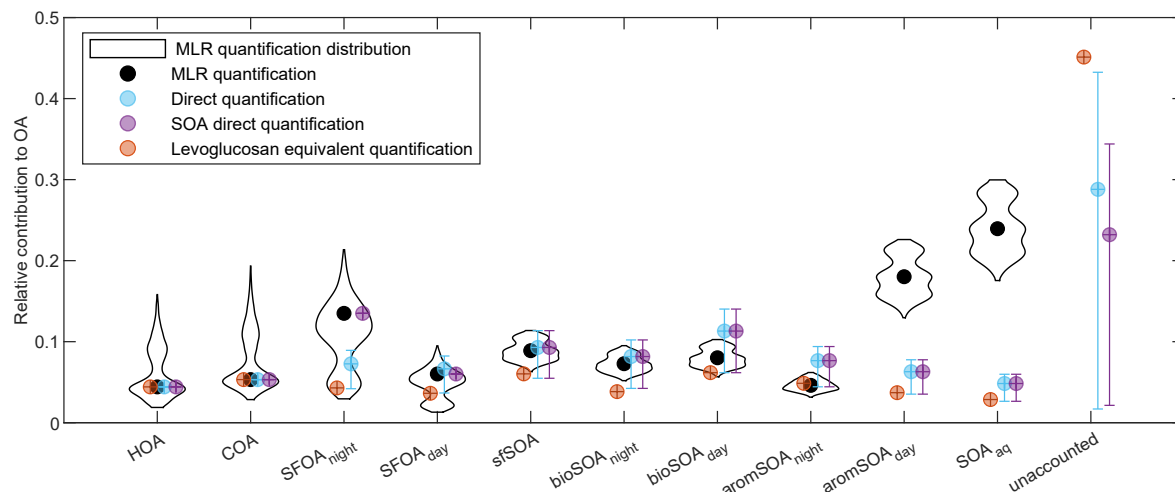
Extended Data Fig. 3 | Sources of organic aerosol and their contribution to fine particle mass during the COVID lockdown period. Bulk chemical composition of particulate matter from an Aerosol Chemical Speciation Monitor and Aethalometer (ToF-ACSM and AE33, left) shows that roughly one-quarter of the particulate matter mass is organic without detailed information on the

sources. Additional near-molecular information based on thermal desorption chemical ionization mass spectra (FIGAERO-CIMS, right) reveals that three-quarter of the organic aerosol is secondary, with a roughly equal contributions of biogenic SOA, aromatic-dominated-SOA, and aqueous SOA. The holes in the pie charts are constituents and sources not covered, which comprise less than 20%.



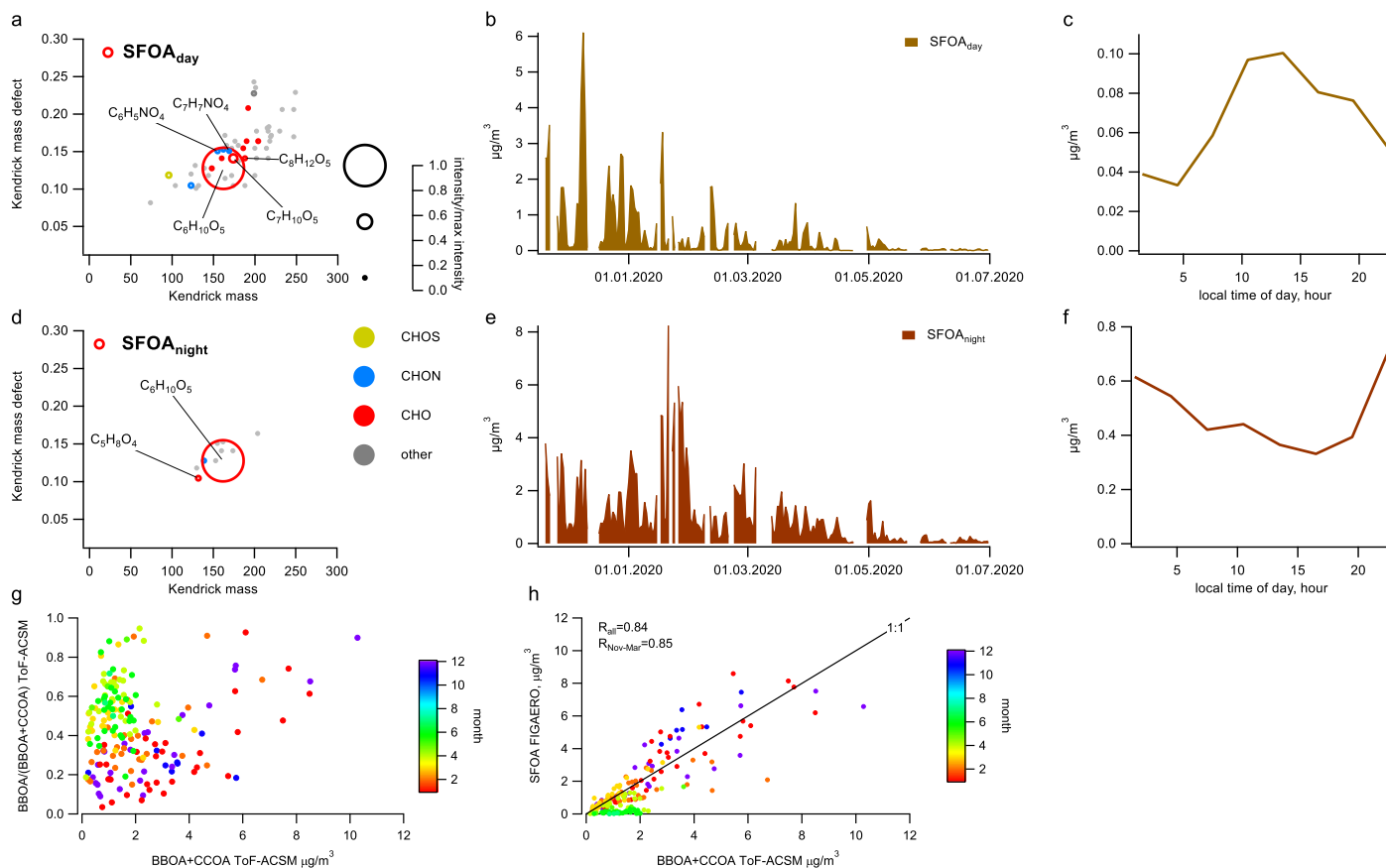
Extended Data Fig. 4 | Relative contribution of POA and SOA types to OA via direct quantification of the FIGAERO-CIMS. The concentrations of FIGAERO-CIMS OA types are determined via the direct quantification approach (see methods and Extended Data Fig. 5) and combined with Hydrocarbon-like OA

(HOA) and Cooking OA (COA) from the ToF-ACSM. This approach leaves the OA mass fraction that is neither quantified by the ToF-ACSM (HOA + COA) nor directly measured by the FIGAERO-CIMS as unaccounted (winter 29%, COVID lockdown 26%, summer 33%).



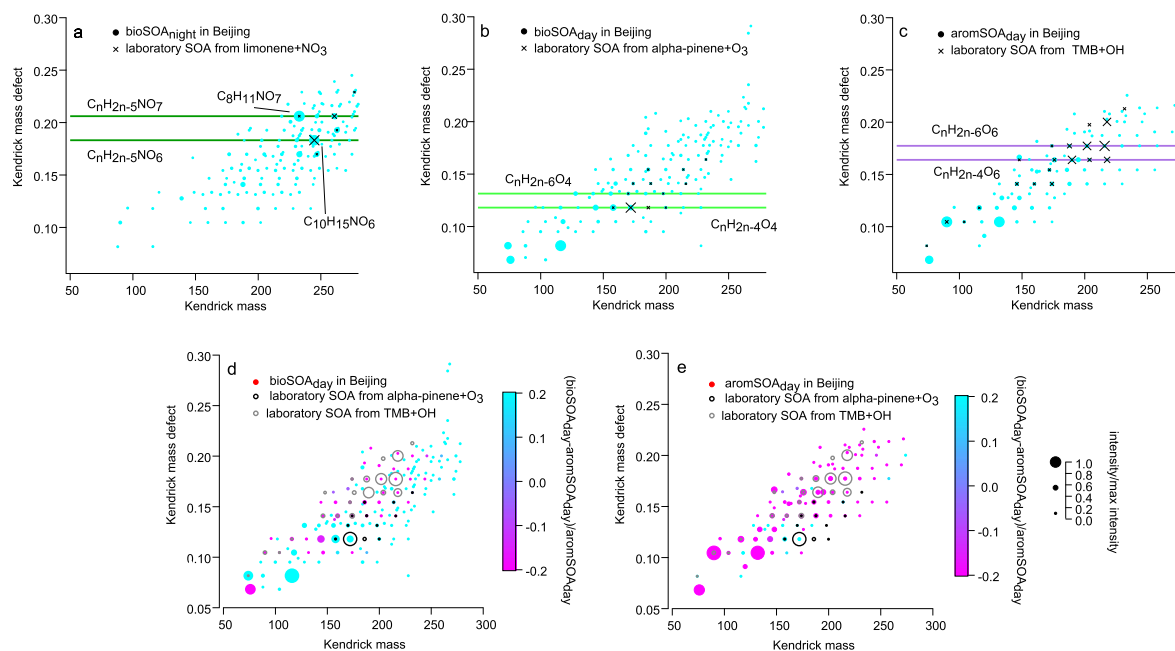
Extended Data Fig. 5 | Sensitivity assessment of relative contribution of POA and SOA types to OA during the measurement period. Results from the sensitivity assessment also accounting for uncertainties in the FIGAERO-CIMS response factors are shown as violin plots. Values based on the MLR quantification setting used in the main text are shown as a black dot (violin plots show the range from 1st to 99th percentile from a Monte Carlo simulation with three scenarios totalling 3600 runs). Results reported as levoglucosan equivalent concentrations are reported as orange dots. In addition to the levoglucosan-equivalent quantification, the direct quantification approach combines our own

levoglucosan calibration with response factors relative to levoglucosan. The results from direct quantification of all FIGAERO-CIMS factors are displayed in blue and of all FIGAERO-CIMS SOA factors in violet (the uncertainty displayed is based on 1 standard deviation of response factors relative to $C_6H_{10}O_5$ in the data presented by Ye et al.²⁸, parametrized based on a moving average across the mass spectrum using 40 ions at a time). With our MLR-based approach combining FIGAERO-CIMS and ToF-ACSM, we quantify which SOA types are behind the OA mass not accounted for by the directly quantified FIGAERO-CIMS OA mass.



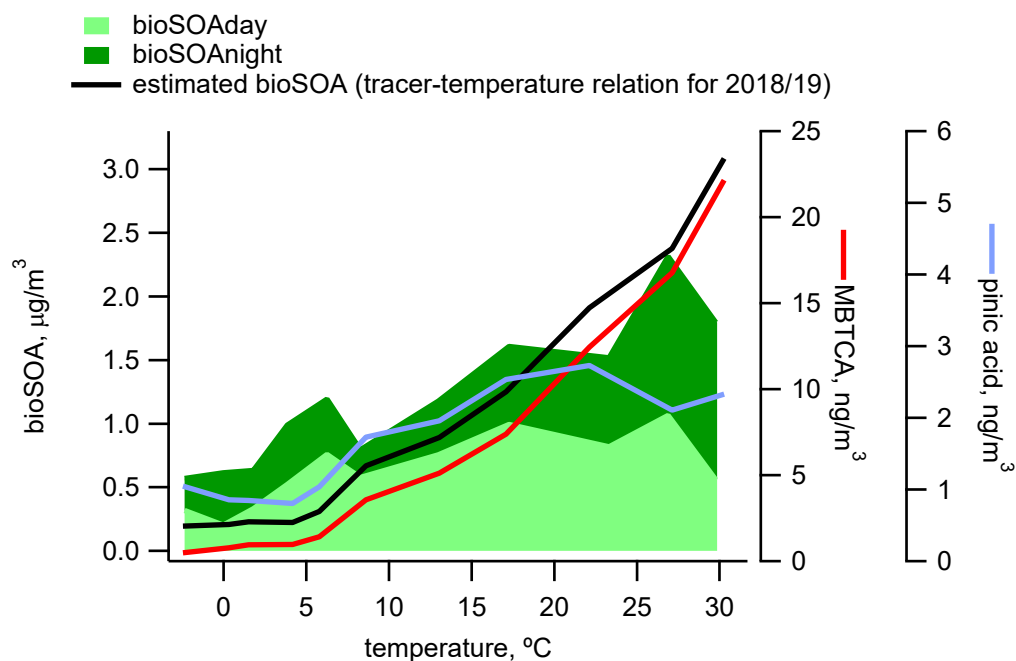
Extended Data Fig. 6 | Characteristics of solid-fuel POA and the impact of biomass burning and coal combustion. a–h, Comparison between solid-fuel POA ($SFOA_{day}+SFOA_{night}$) (FIGAERO-CIMS OA data) to the sum of biomass burning and coal combustion POA resolved (ToF-ACSM) (daily means). During winter CCOA dominates SFOA at low and BBOA dominates at high SFOA (or BBOA+CCOA) concentrations. During the warm season, the ToF-ACSM based estimate (BBOA+CCOA, related to mostly BBOA) is larger than the FIGAERO-CIMS based estimated (SFOA). The chemical composition is displayed as Kendrick mass

defect plots of the neutral compound with CH_2 base. Compounds that are at least 10% of the maximum compound intensity are size-coded with the compound's intensity normalized to the highest intensity compound and are color-coded with compound families: $C_xH_yO_z$: red, $C_xH_yO_2N_z$: blue, $C_xH_yO_2S_z$: yellow, other: grey. Compounds with an intensity between 5 and 10% of the maximum compound intensity are shown as small unsized solid grey dots (others are not displayed). The seasonal variation is shown as daily mean concentrations and the diurnal variation as median diel cycles.



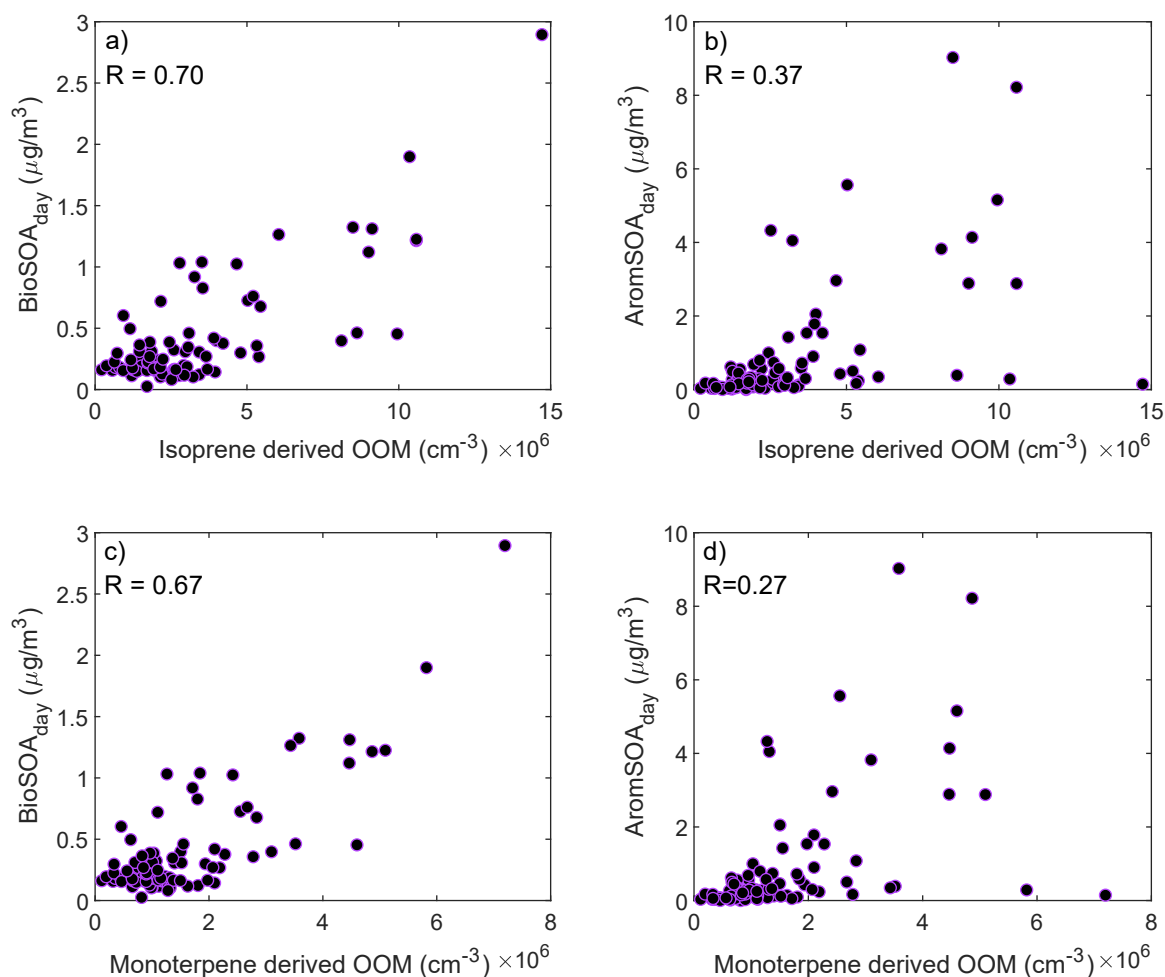
Extended Data Fig. 7 | Chemical composition of biogenic-SOA and daytime aromatic-dominated-SOA. a–e. The chemical composition of biogenic-SOA (bioSOA_{day}, bioSOA_{night}) and daytime aromatic-dominated-SOA (aromSOA_{day}) is displayed as Kendrick mass defect plots of the neutral compound with CH₂ base. For comparison the composition of laboratory SOA of trimethylbenzene (TMB) - (reacted with OH), α -pinene (reacted with ozone), and limonene (reacted with NO₃ radicals) is displayed. Top panel: Symbols are size-coded with a compounds intensity normalized to the highest intensity compound, compounds with intensity < 10% of maximum compound intensity are discarded, the symbols are color-coded with compound families). The violet line refers to common trimethylbenzene SOA compounds from reaction with OH, the dark green line refers to common SOA compounds from limonene reacted with NO₃ radicals,

and the light green line refers to common α -pinene ozonolysis SOA compounds. Bottom panel: Symbols are size-coded with a compounds intensity normalized to the highest intensity compound and color-coded with relative difference in the normalized intensity between bioSOA_{day} and aromSOA_{day} (compounds with intensity < 10% of maximum compound intensity are discarded). This comparison illustrates that 1) bioSOA_{night} has a distinct chemical composition similar to limonene reacted with NO₃ radicals (an important reaction pathway during the night), 2) different compounds are characteristic for bioSOA_{day} and aromSOA_{day} and that 3) many of the peaks characteristic for bioSOA_{day} are found to be important for laboratory α -pinene - biogenic - SOA and similarly for aromSOA_{day} and TMB - anthropogenic - SOA.



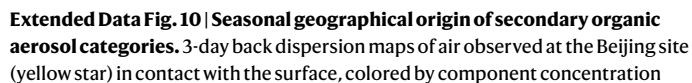
Extended Data Fig. 8 | Biogenic-SOA and concentrations at different temperatures. bioSOAday and bioSOAnight are displayed as stacked green shaded areas. They are compared to biogenic-SOA concentrations at different

temperatures estimated using offline filter analysis of MBTCA and pinic acid for 2018/19 (assuming a bioSOA/(MBTCA+pinic acid) ratio of 0.126 μg bioSOA per ng MBTCA+pinic acid⁷).



Extended Data Fig. 9 | Relation between biogenic and aromatic-dominated-SOA to gas-phase Oxygenated Organic Molecules from the oxidation of isoprene and monoterpene. a–d. Biogenic-SOA correlates better than aromatic-dominated-SOA with gas-phase Oxygenated Organic Molecules (OOM) from isoprene (a,b) and monoterpene (c,d) oxidation. Correlation coefficients shown in the figure. Here the period 20.11.2019–27.04.2020 is displayed during which

the daily mean temperature did not exceed 20 °C. Up to 20 °C pinic acid (more volatile) and MBTCA (less volatile) - molecules widely used as bioSOA markers - correlate well but at higher temperatures the correlation deteriorates, likely because of gas-to-particle partitioning (Supplementary Fig. 26). Such effects also impact the gas-phase to particle-phase relation and thus only days with a mean temperature lower than 20 °C are used for this comparison.



Nature Geoscience