Supplement of

Molecular Insights into Organic Aerosol Sources and Formation at a Regional Background Site in South China

Hongxing Jiang et al.,

Correspondence to: H. Guo (hai.guo@polyu.edu.hk)

Supplementary text

HR-TOF-AMS analysis

We conducted a comprehensive sampling campaign in the autumn during the COVID-19 lockdown period (September 29 to November 18, 2020) at a regional background site, i.e., Hok Tsui, in Hong Kong to study the influence of transport from mainland China on air quality of Hong Kong. During this period, we employed an online HR-TOF-AMS (Aerodyne Inc.) to measure the chemical compositions of NR-PM_1, including concentrations of total organics, sulfate, nitrate, ammonium, and chloride. The AMS operated alternately in V-mode and W-mode on a 2-minute cycle. The elemental composition of organics, determined using W-mode data due to its high mass resolution (~5000–6000), was then used for the source apportionment of OA. The collection efficiency (CE) for this study was determined to be 0.73 and was applied to all measured NR-PM_1 components.

Data processing was performed using the ToF-AMS Analysis Toolkit 1.59D and TofAMS HR Analysis 1.19D, both implemented in Igor Pro 6.37 software. We applied relative ionization efficiencies (RIE) of 1.4, 1.1, 1.2, 4.0, and 1.3 to calculate the concentrations of total organics, nitrate, sulfate, ammonium and chloride, respectively. Method detection limits (MDLs) for each species were determined by collecting background mass spectra (HEPA filtered air, 60 min every 2 days). The calculated MDLs for organics, nitrate, sulfate, ammonium, and chloride were 0.21μg m^{-3}, 0.020 μg m^{-3}, 0.022μg m^{-3}, 0.013 μg m^{-3}, and 0.013 μg m^{-3}, respectively. OA source apportionment was achieved using the Positive Matrix Factorization (PMF) Evaluation Toolkit (PET v2.05). Several steps were conducted, as described in our previous studies, to ensure robust results, including applying a minimum error value and ion filtrations(Huo et al., 2024a; Yao et al., 2022). The optimal 4-factor solution was selected based on Q/Q_{exp} values, residuals, and mass spectra(Zhang et al., 2011). Fig. S1 presents the mass spectra of the 4 OA factors resolved by PMF. The HOA factor was characterized by abundant alkyl fragments ($C_nH_{2n+1}^+$ and $C_nH_{2n-1}^+$), such as $C_3H_5^+$, $C_3H_7^+$, $C_4H_7^+$, and $C_4H_9^+$(Sun et al., 2011). The HOA factor identified in this study had O/C and H/C values of 0.25 and 1.72, respectively, which were the lowest and highest among the 4 factors. However, the O/C value was higher than typical ranges for HOA reported in previous studies (e.g., 0.05–0.25), indicating the influence of atmospheric aging processes(Huo et al., 2024a). The other 3 factors were characterized by high loadings of oxygenated species, such as m/z 44 (mainly CO_2^+), confirming their classification as oxygenated OA components. The 3 OOA factors were divided into two less oxidized OOAs (LO-OOA1 and LO-OOA2) and one more oxidized OOA (MO-OOA) based on their O/C values. The LO-OOA factors also contained many alkyl fragments, indicating closer associations with primary sources.

TAG-EI-TOF-MS analysis

During the sampling period, TAG-EI-TOF-AMS analysis was used to determine the concentrations of SOA tracers in the particle phase. Detailed descriptions of the instrument and its performance are provided in our previous studies (Huo et al., 2024b; Lyu et al., 2020). Briefly, air samples were collected using a collection

and thermal desorption (CTD) unit at 30 ℃ with a flow rate of 10 L min^{-1} over 90 min. Derivatization occurred in the CTD unit by purging it with a mixed flow of pure helium (20 mL/min) and saturated N-methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA)-helium (80 mL/min). During derivatization, the CTD unit was heated to 315 °C over 9 minutes and held at that temperature for 5 minutes. As the temperature increased, derivatives were desorbed and transferred to a focusing trap (FT) unit (~30 °C), where target compounds and their derivatives were trapped while excess MSTFA and most volatile organics were vented. The FT was then gradually heated to 315 °C over 12 minutes, with pure helium purging the target compounds from the FT to a mini-gas chromatography (GC) column. The GC column was heated according to a programmed temperature profile (0.75 mL/min), starting at ~40 °C and increasing to 45 °C in 1 minute, then rapidly rising to 330 °C over 10 minutes, and held at 330 °C for 12 minutes until the end of the GC analysis. Before sampling, external standards with known concentrations were used for locating and quantifying target compounds. Additionally, to address desensitization and peak drift during analysis, a mix of concentration-constant internal standards composed of 26 deuterated compounds was injected alongside each sample. A total of 62 compounds were identified, with details provided in Table S2.

PTR-TOF-MS analysis

An online proton transfer reaction quadrupole ion time-of-flight mass spectrometer (PTR-TOF-MS, IONICON Analytik GmbH, Innsbruck, Austria) was used to measure the concentrations of VOC species. Detailed information about the instrument is available in previous studies (Yuan et al., 2024). Ambient air was continuously pumped at a rate of 3.0 L/min through a Teflon tube (4.7 mm Teflon-membrane filter, Whatman Ihc. Clifton, NJ, USA). A small fraction of this air (0.2 L/min) was introduced into the TOF-MS for VOCs measurements. Certified standard gas mixtures (1 ppm, Linde Spectra Environmental Gases, USA) were used to quantify the target VOCs species, determine the transmission curve of the instrument, and assess the sensitivities of the corresponding species. Calibration of the PTR-TOF-MS was performed weekly using a liquid calibration unit (LCU, Ionicon) at room temperature.

Other ancillary measurements

In addition to the online instruments used for detecting atmospheric organics, a series of other online instruments were employed to measure the concentrations of $PM_{2.5}$, trace gases (CO, NO, NO_2, O_3 and SO_2), and meteorological parameters (temperature and relative humidity). $PM_{2.5}$ concentrations were measured using a Tapered Element Oscillating Microbalance (TEOM) (Thermo Scientific™ 1405). SO_2 concentrations were determined with a Teledyne Advanced Pollution Instrumentation (API) Trace-level UV Fluorescence SO_2 Analyzer (T100U), while nitrogen species were measured using the Teledyne True NO_2/NO/NO_x Analyzer (T200UP) Ozone (O_3) concentrations were measured With an Ecotech UV Absorption Ozone Analyzer (EC9810B/S10). The photolysis frequency of NO_2 (jNO_2) was measured using a MetCon Filter Radiometer.

Carbon analysis

The concentrations of organic carbon (OC) and elemental carbon (EC) in $PM_{2.5}$ were determined using a Sunset OC/EC analyzer with an enhanced thermal/optical reflectance protocol. A 1.5 cm^2 section of the $PM_{2.5}$ filters was placed in a quartz boat and subjected to stepwise heating in a quartz furnace to separate OC and EC. During the analysis, four OC peaks (OC1, OC2, OC3, and OC4) and three EC peaks (EC1, EC2, and EC3) were identified. Pyrolyzed organic carbon (PyC), formed by the coking of OC during the procedure, was monitored through changes in laser reflectance signals. The total contents of OC and EC were then calculated using the formulas: OC = OC1 + OC2 + OC3 + OC4+ PyC; EC= EC1 + EC2 + EC3- PyC.

HR-orbitrap-MS analysis

The molecular composition of OAs in offline $PM_{2.5}$ filters was analyzed using a high-resolution Q-Executive Orbitrap mass spectrometer (Thermo Electron, Inc.) coupled with an ultra-high performance liquid chromatography system (UHPLC, Dionex UltiMate 3,000, Thermo Electron, Inc.). Detailed descriptions of the analysis procedures and instrumental settings are available in a previous study (Zhang et al., 2024). Briefly, two pieces of $PM_{2.5}$ filters were punched using a stainless steel puncher (Φ=20 mm), and the dissolved organic matter was extracted with 6 mL of mix-solvents (2×3 times, methanol: toluene=1:1, v/v) in an ultrasonic cold-water bath for 20 minutes. The extracts were filtered through a 0.22 μm polytetrafluoroethylene filter membrane, combined, and evaporated to near dryness under a gentle stream of high-purity nitrogen. The residue was then redissolved in 150 μL of methanol and centrifuged, with the supernatant transferred for subsequent HR-MS analysis.

As noted in previous studies, potential intermolecular suppression effects can occur during the ionization process (Zhang et al., 2024; Thoma et al., 2022). To address these challenges, an UHPLC system was employed for compound separation. Samples (5 μL) were injected into the system, and separation was performed on an Acquity UPLC HSS T3 column (1.8 μm particle size, 100 mm × 2.1 mm; Waters, Milford, MA, USA) with a VanGuard pre-column (HSS T3, 1.8 μm) at a flow rate of 0.3 mL min^{-1}. A gradient elution procedure was used for compounds isolation, with eluent A/B consisting of 0.1% formic acid in ultrapure water/methanol. Eluent B was initially maintained at 10% for 2 min, increased to 54% over 15.2 min and held for 1 min, then increased to 90% over 7.5 min and held for 0.2 min, before returning to 1% within 1.8 min and held for 9.6 min before the next sample. Analytes were introduced to the heated electrospray ionization (ESI) source system and ionized in negative mode. The spray voltages were set as 3.0 kV for ESI−, with the capillary temperature at 320°C. The sheath gas flow, auxiliary gas flow, and sweep gas flow were set 35, 10, and 0 units, respectively. The mass spectrometer scanned a range of m/z 50-800 with a typical mass resolution of 140,000 at m/z 200.

Xcalibur software (V2.2; Thermo Scientific) was used to acquire the raw data, while further non-target compound analysis was performed using the open-source MZmine-2.37 software (http://mzmine.github.io). This analysis included raw data import, peak detection, shoulder peak filtering, chromatogram building,

chromatogram deconvolution, deisotoping, searching for adducts and peak complexes, alignment, gap filling, identification, and duplicate peak filtering. Detailed processing steps and settings are described in previous literature (Wang et al., 2017). Mass peaks were assigned to specific molecules with a mass tolerance of 2 ppm for ESI- mode. The atoms in the assigned molecular formulas were limited to $C_{1-40}H_{0-100}O_{0-40}N_{0-5}S_{0-3}$, with additional criteria on elemental ratios (e.g., H/C, O/C) and double bond equivalents (DBE) applied to eliminate chemically meaningless molecular formulas. The elemental ratios, DBE, and modified aromatic index were calculated based on the assigned formulas of $C_cH_hO_oN_nS_s$, where *c*, *h*, *o*, *n*, and *s* represent the number of carbon, hydrogen, oxygen, nitrogen, and sulfur atoms, respectively. All molecules reported in this study underwent blank subtraction, and those with an abundance ratio of less than 5:1 were eliminated (Ditto et al., 2018).

The chemical parameters, including DBE, nO_{eff}, (Nie et al., 2022), AI_{mod} (Koch and Dittmar, 2006), and volatility (LogC*) (Li et al., 2016), of the compounds are calculated using the following equations respectively:

$$DBE = (2\ n_C + 2 - n_H + n_N)/2$$

$$nO_{eff} = n_O - 2\ n_N - 3\ n_S$$

$$AI = (1+n_C-0.5n_O-n_S-0.5n_H)/(\ n_C - 0.5n_O - n_S - n_N)$$

$$Log_{10}C^* = (n_C^0 - n_C)b_C - n_O b_O - 2\frac{n_C n_O}{n_C + n_O}b_{CO} - n_N b_N - n_S b_S$$

where n_C, n_H, n_O, n_N, and n_S denote the numbers of carbon, hydrogen, oxygen, nitrogen, and sulfur atoms in the molecular formula, respectively. n_C^0 is the reference carbon number. The parameters b_C, b_O, b_N, and b_S denote the contribution of each atom to $log_{10}C^*$, respectively, while b_{CO} is the carbon–oxygen nonideality.

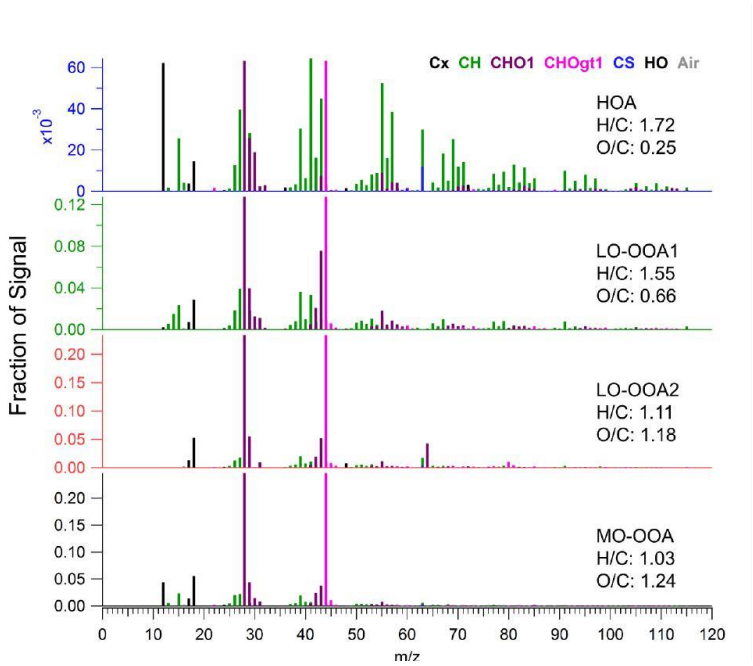


Figure S1. Mass spectra of OA components obtained from AMS-PMF analysis

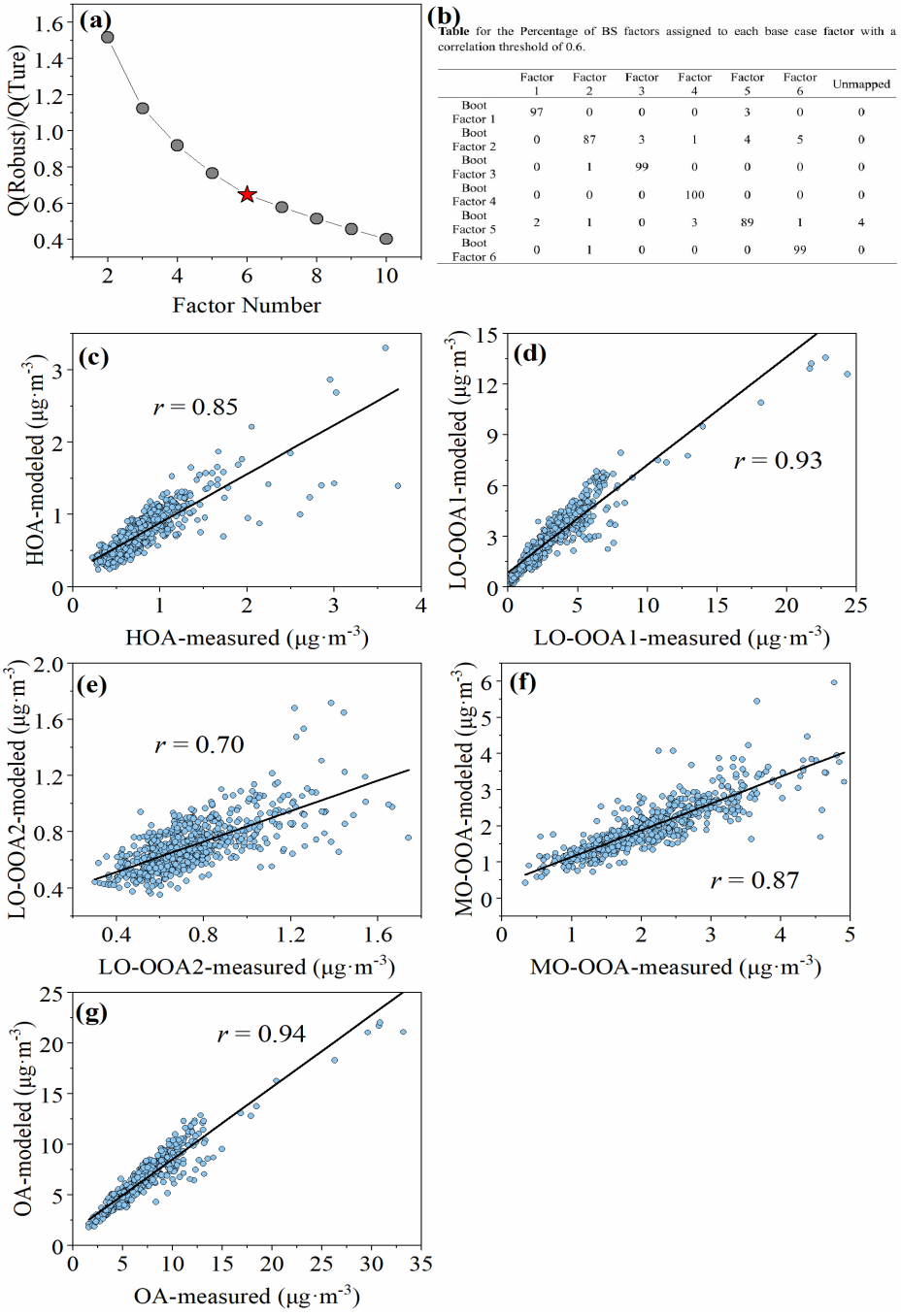


Table for the Percentage of BS factors assigned to each base case factor with a correlation threshold of 0.6.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Unmapped
Boot Factor 1	97	0	0	0	3	0	0
Boot Factor 2	0	87	3	1	4	5	0
Boot Factor 3	0	1	99	0	0	0	0
Boot Factor 4	0	0	0	100	0	0	0
Boot Factor 5	2	1	0	3	89	1	4
Boot Factor 6	0	1	0	0	0	99	0

Figure S2. Diagnostic results for tracer-based PMF analysis. (a) Variation of average Q/Q_{exp} with increasing factor numbers in PMF. (b)Table showing the percentage of bootstrap (BS) factors assigned to each base case factor with a correlation threshold of 0.6. (c-g) Correlation analysis between the concentrations of AMS-PMF derived OA components and the tracer-based PMF modeled values

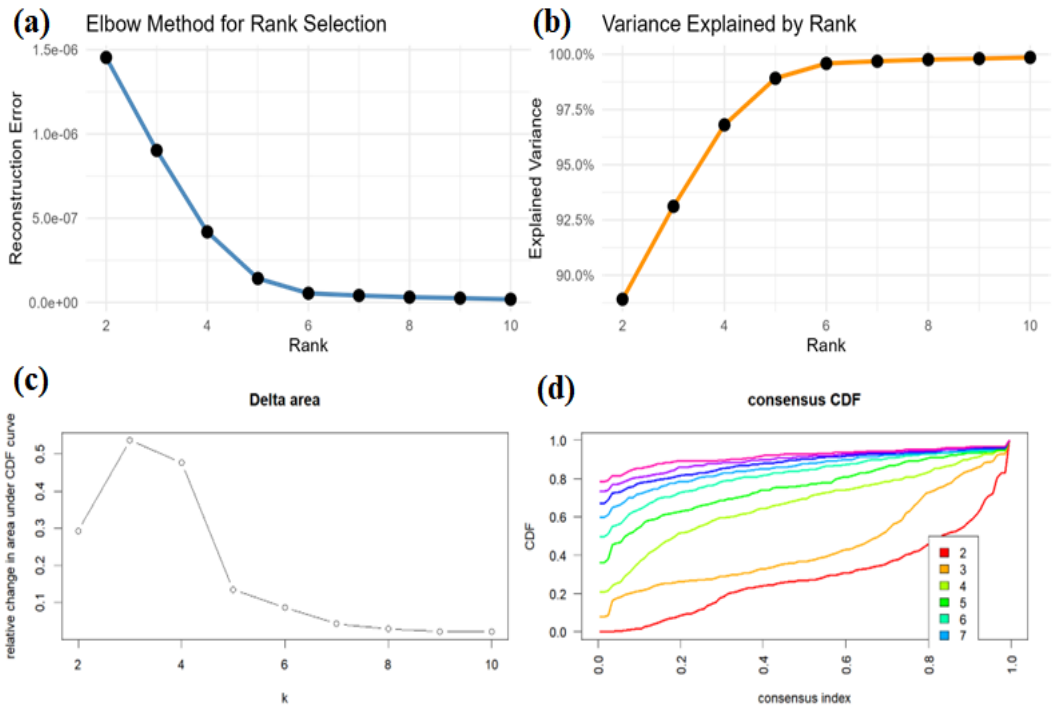
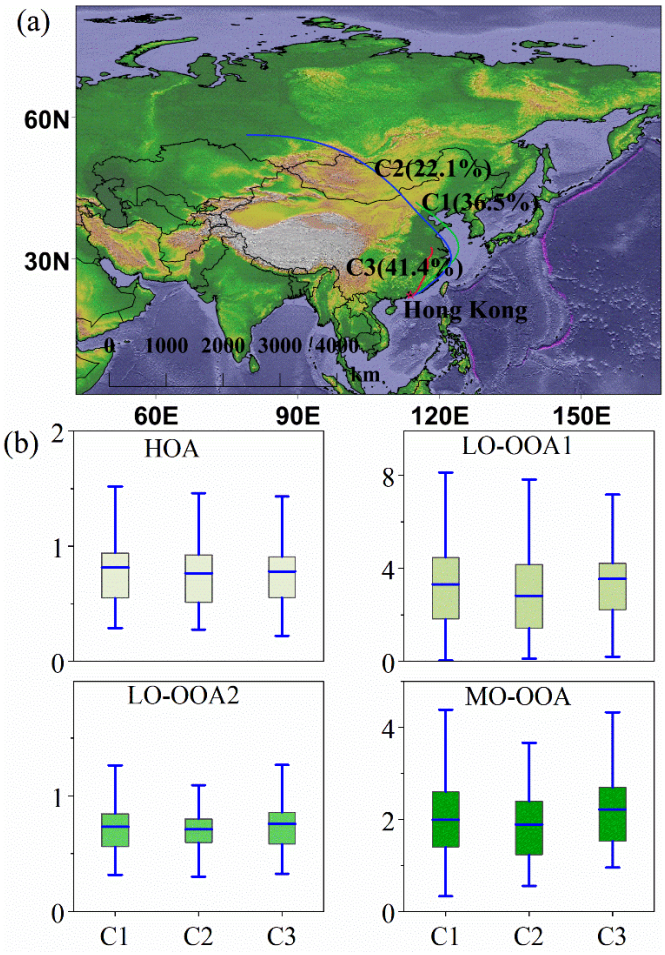


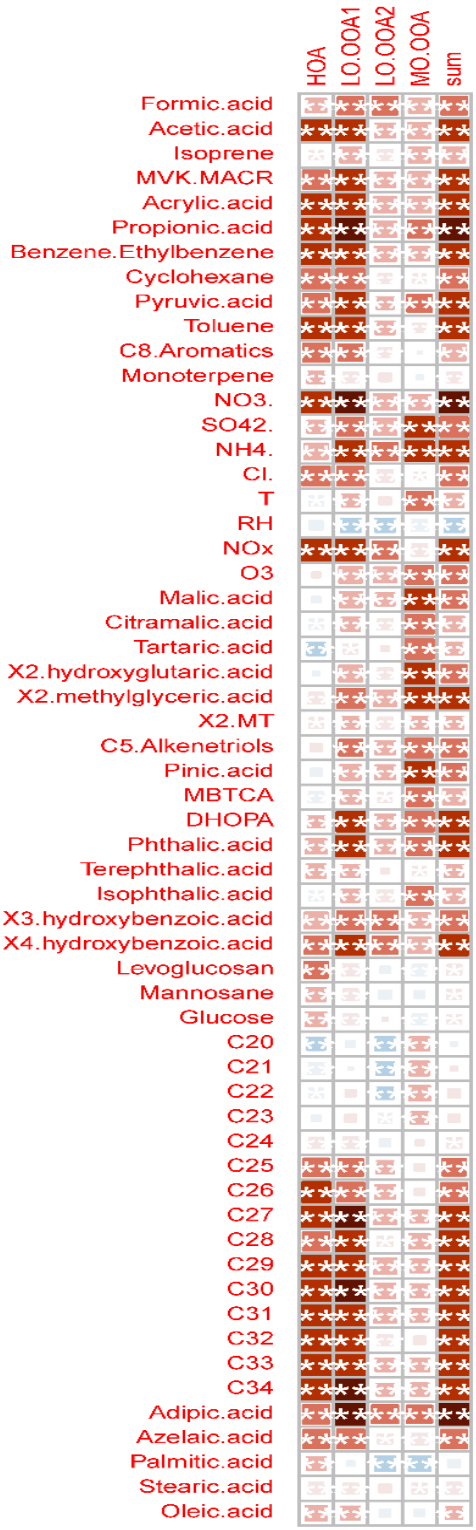
Figure S3. Diagnostic results for PMF results-constrained NMF analysis. (a) Variation of reconstructured error with increasing factor numbers. (b) Variation of explained total variances with increasing factor numbers. and (c) Variation of cumulative distribution function (CDF) delta area with increasing factor numbers. (d) Cumulative distribution functions of the consensus matrix for each factor number, estimated using a histogram with 100 bins



147

148 **Figure S4.** (a) Air mass clusters of the 120 - hr backward trajectories arriving at the sampling site during the
149 study period. (b) Box plot of the four OA components for different air mass influence periods: The blue line
150 in each box denotes the mean value, the upper (lower) boundary of the box represents the 75th (25th)
151 percentile, and the top and bottom whiskers indicate the 95th and 5th percentiles, respectively

152



153

154 **Figure S5.** Correlations between AMS-PMF derived OA components and various chemical species.

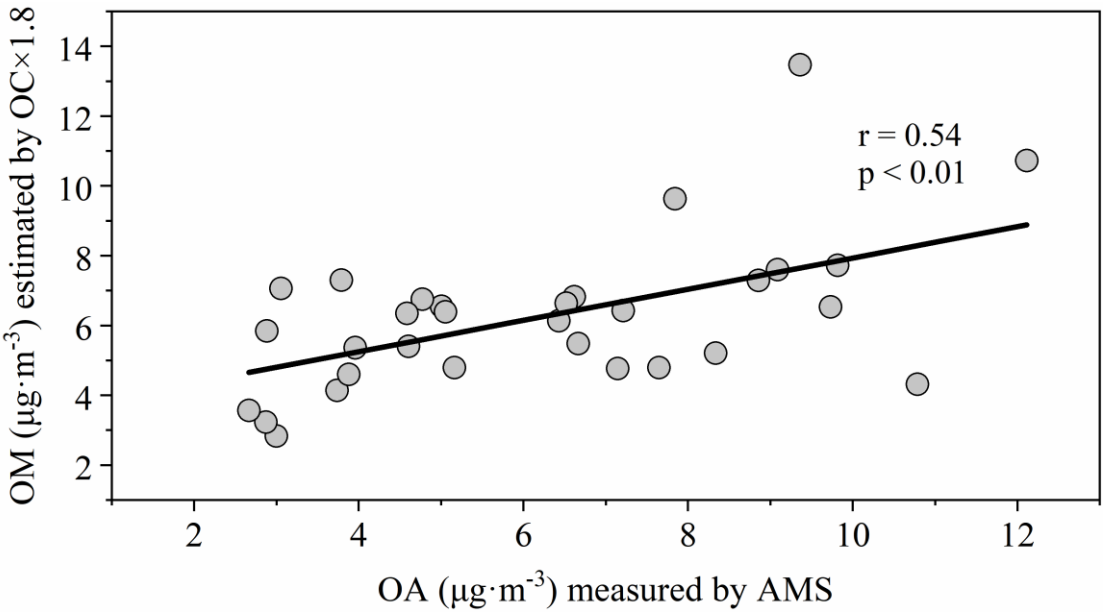


Figure S6. Correlations between AMS-derived OA concentrations and estimated OM (OC multiplied by a factor of 1.8). The correlation coefficient observed suggests that, despite the large uncertainty in OM estimations, errors between AMS and OC/EC analyzer, and the limited sample size, the OA in PM1 likely shares similar sources and formation pathways with OM in $PM_{2.5}$.

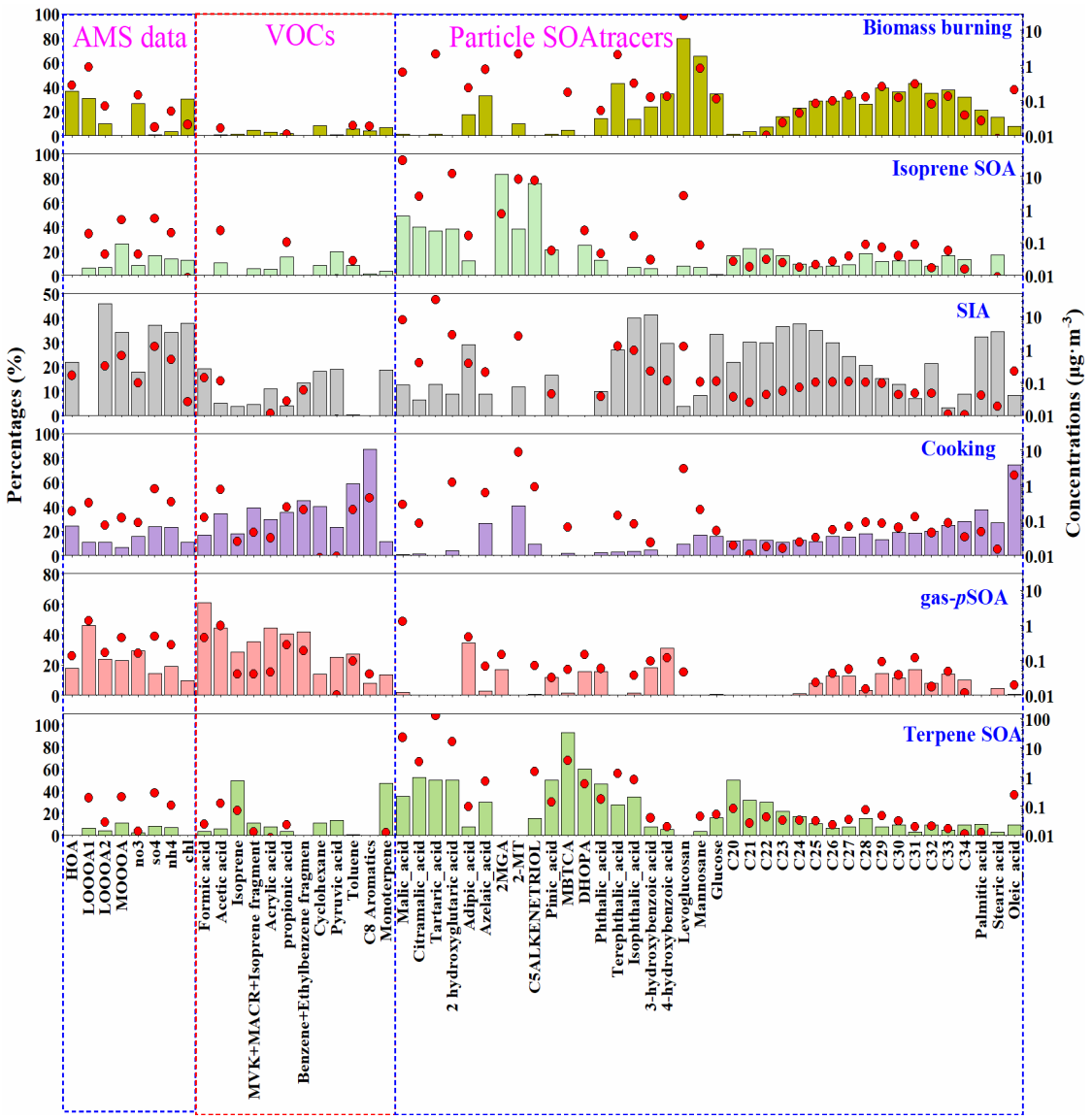


Figure S7. Source profiles of individual factors for the 6-factor solution resolved by tracer-based PMF

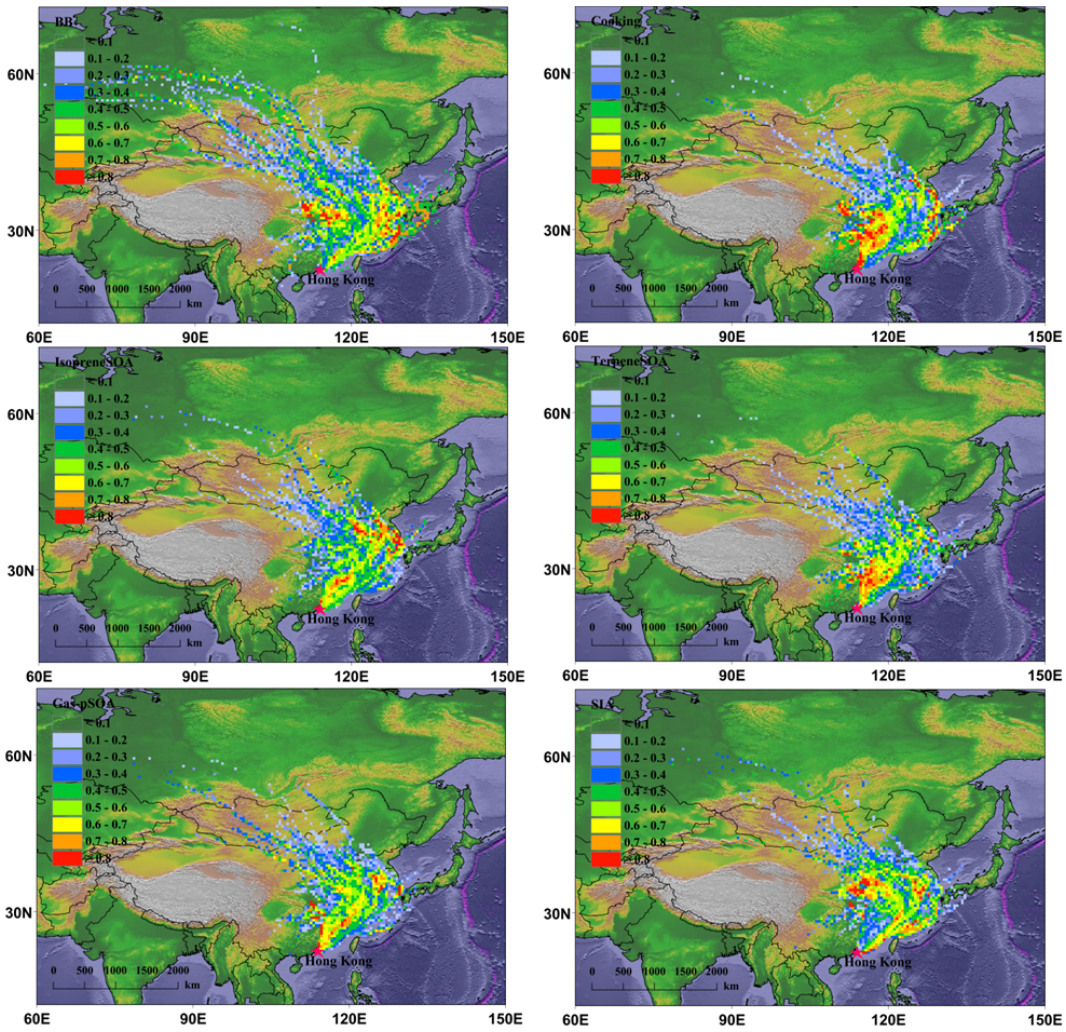


Figure S8. PSCF maps of the 6 PMF-resolved sources during the sampling period.

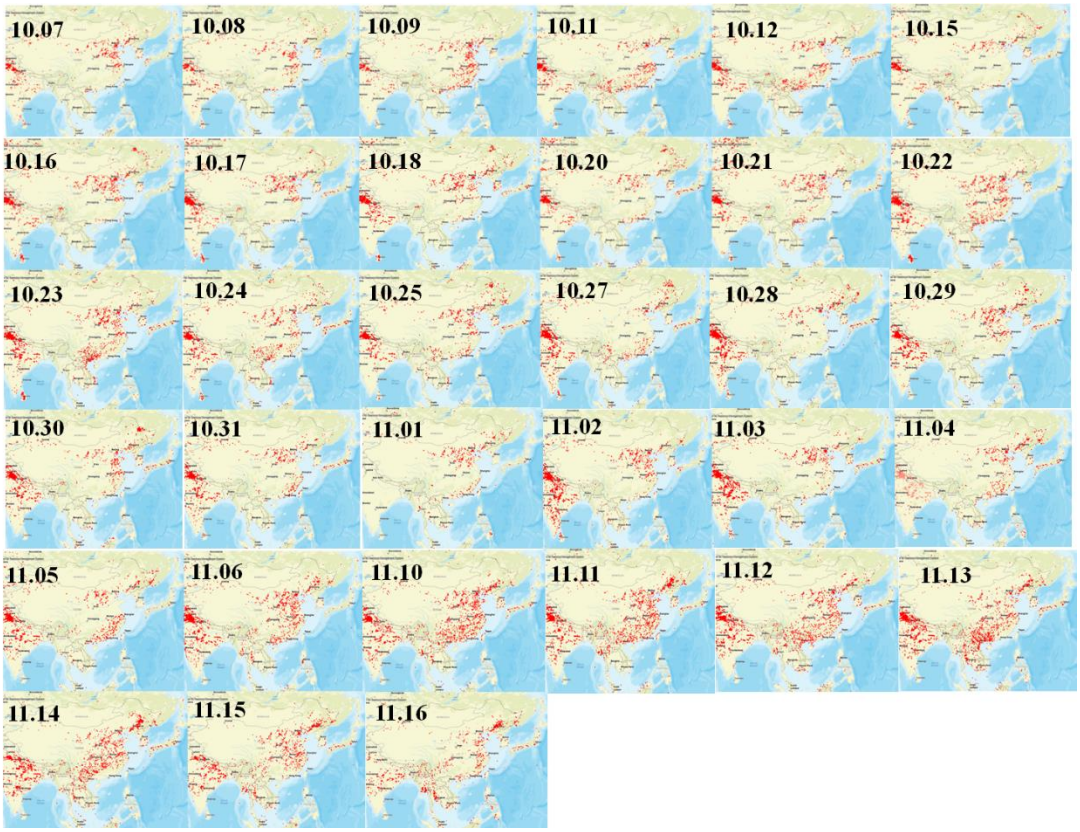
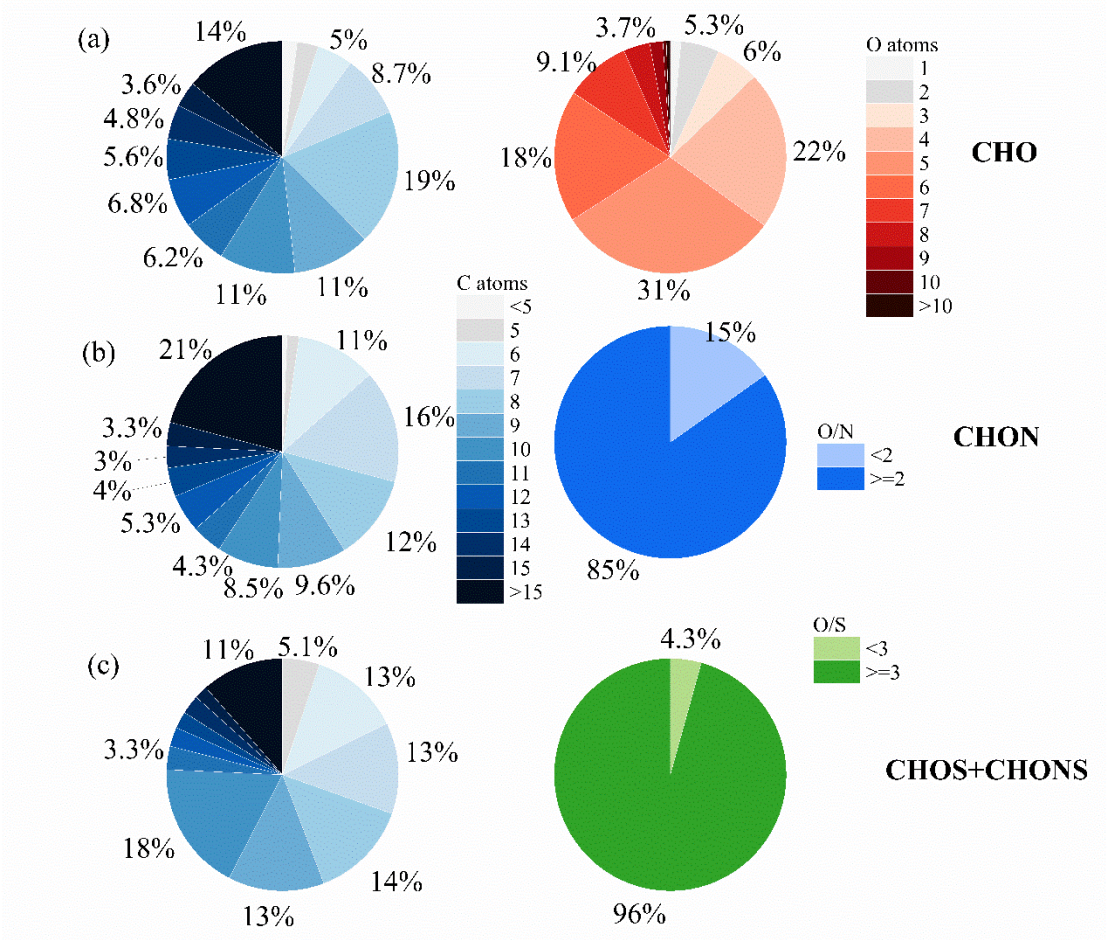
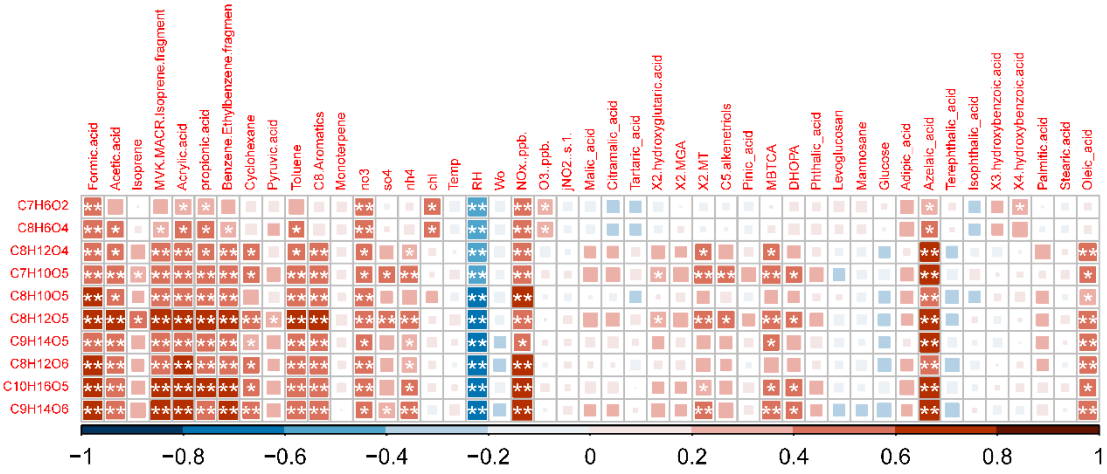


Figure S9. Daily fire maps corresponding to the offline PM2.5 filter samples.



166

167 **Figure S10.** Intensity distribution of different compound groups



168

169 **Figure S11.** Correlations between the intensity of several highly abundant CHO compounds and various

170 chemical species.

171 **Table S1.** Sampling information for offline PM2.5 samples.

Sampling ID	Sampling date	$PM_{2.5}$ ($\mu g \cdot m^{-3}$)	Temp (℃)	RH (%)	OC ($\mu g \cdot m^{-3}$)	EC ($\mu g \cdot m^{-3}$)
HT01	2020/10/07	16.4	24.3	72	5.41	1.63
HT02	2020/10/08	16.8	25.0	68	5.05	1.60
HT03	2020/10/09	22.0	25.6	68	6.73	2.10
HT04	2020/10/10	13.8	26.7	76	4.92	1.97
HT05	2020/10/11	13.5	26.6	82	4.36	1.13
HT06	2020/10/12	14.5	25.9	77	5.99	2.56
HT07	2020/10/15	17.0	25.7	78	4.63	1.31
HT08	2020/10/16	16.0	24.6	77	4.25	1.26
HT09	2020/10/17	14.8	24.0	77	3.97	1.31
HT10	2020/10/18	18.3	23.4	70	2.56	1.27
HT11	2020/10/19	21.0	24.4	62	2.78	1.16
HT12	2020/10/20	16.0	25.6	61	2.87	0.96
HT13	2020/10/21	20.8	23.5	54	3.70	1.19
HT14	2020/10/22	23.2	23.5	70	5.45	1.04
HT15	2020/10/23	13.0	24.2	76	1.70	0.55
HT16	2020/10/24	11.4	24.0	81	1.67	0.60
HT17	2020/10/25	10.7	23.3	84	1.60	0.59
HT18	2020/10/27	19.7	23.9	84	2.11	1.19
HT19	2020/10/28	20.2	23.5	76	2.81	0.86
HT20	2020/10/29	12.5	23.2	72	2.08	0.76
HT21	2020/10/30	20.9	23.9	72	2.65	0.95
HT22	2020/10/31	26.7	23.5	67	3.57	1.45
HT23	2020/11/01	26.8	23.3	69	4.01	1.35
HT24	2020/11/02	16.6	24.2	60	2.15	0.83
HT25	2020/11/03	11.7	22.8	77	1.48	0.74
HT26	2020/11/04	34.8	26.1	59	5.20	2.05
HT27	2020/11/05	22.0	22.4	66	3.68	1.18
HT28	2020/11/06	16.4	21.7	67	2.20	0.71
HT29	2020/11/10	19.4	22.6	68	3.62	1.43
HT30	2020/11/11	21.3	22.0	80	2.55	1.03
HT31	2020/11/13	19.0	23.0	77	1.60	0.61

172

173

Table S2. Quantified compounds using TAG-EI-TOF-MS.

Compounds	Abbreviations	IS
Malic acid	MA	^{13}C-Pentaerythritol
Citramalic acid	CA	^{13}C-Pentaerythritol
Tartaric acids	TA	^{13}C-Pentaerythritol
2-hydroxyglutaric acid	2-HGA	^{13}C-Pentaerythritol
2-methylglyceric acid	2-MGA	^{13}C-Pentaerythritol
cis-2-methyl-1,3,4trihydroxy-1-butene		^{13}C-Pentaerythritol
3-methyl-2,3,4trihydroxy-1-butene	C5-alkenetriols	^{13}C-Pentaerythritol
trans-2-methy-1,3,4trihydroxy-1-butene		^{13}C-Pentaerythritol
2-methylthreitol	2-MTs	^{13}C-Pentaerythritol
2-methylerythritol		^{13}C-Pentaerythritol
Pinic acid	PA	1-Dodecan-D25-ol
3-methyl-1,2,3butanetricarboxylic acid	MBTCA	^{13}C-Pentaerythritol
2,3-dihydroxy-4oxopentanoic acid	DHOPA	^{13}C-Pentaerythritol
Phthalic acid	PhA	D-phthalic acid
Terephthalic acid	TPA	D-phthalic acid
Isophthalic acid	IPA	D-Pentadecanol
Levoglucosan	Lev	^{13}C-Pentaerythritol
Mannosane	Man	^{13}C-Pentaerythritol
Glucose	Glu	^{13}C-Pentaerythritol
Adipic acid	AdiA	D-adipic acid
Azelaic acid	AzeA	D-Pentadecanol
3-hydroxybenzoic acid	3-HBA	1-Dodecan-d25-ol
4-hydroxybenzoic acid	4-HBA	1-Dodecan-d25-ol
Palmitic acid	PalA	1-Octadeca-d37-nol
Stearic acid	StA	Stearic-d35 acid
Oleic acid	OleA	Stearic-d35 acid
C20		Eicosane-d42
C21		Eicosane-d42
C22		Docosane-d46
C23		Docosane-d46
C24		Tetracosane-d50
C25		Tetracosane-d50
C26		Hexacosane-d54
C27		Hexacosane-d54
C28		Octacosane-d58
C29		Octacosane-d58
C30		Triacontane-d62
C31		Triacontane-d62
C32		Dotriacontane-d66
C33		Dotriacontane-d66
C34		Tetratriacontane-d70

Table S3. Most intense compounds detected in different compound groups

Formula	MW	DBE	AI_{mod}	VOC class	Structural	Possible name or Precursors
CHO compounds						
$C_6H_{10}O_1$	98.19	2	0.27	VOC	Lipids	Fatty acids (Sun et al., 2011)
$C_7H_6O_2$	122.17	5	0.67	IVOC	CRAMS	Benzoic acid and hydroxybenzaldehyde
$C_7H_{10}O_2$	126.20	3	0.33	IVOC	CRAMS	Trimethyl benzene isomers (Mehra et al., 2020)
$C_7H_{10}O_4$	158.20	3	0.20	IVOC	CRAMS	Trimethyl benzene isomers (Mehra et al., 2020)
$C_8H_6O_4$	166.19	6	0.67	IVOC	CRAMS	Phthalic acid
$C_8H_{12}O_4$	172.24	3	0.17	IVOC	CRAMS	cis-Norpinic acid/ terpenylic acid
$C_7H_{10}O_5$	174.20	3	0.11	IVOC	HOC	Bicyclic hydroperoxide
$C_8H_{14}O_4$	174.25	2	0.00	IVOC	Protein	Suberic acid
$C_7H_{12}O_5$	176.22	2	0.00	IVOC	Carbohydrates	Monoterpenes
$C_8H_6O_5$	182.19	6	0.64	IVOC	CRAMS	Hydroxyphthalic acid
$C_8H_{10}O_5$	186.22	4	0.27	IVOC	CRAMS	Trimethyl benzene isomers/ monoterpenes
$C_9H_{14}O_4$	186.27	3	0.14	IVOC	Protein	Pinic acid
$C_8H_{12}O_5$	188.23	3	0.09	IVOC	CRAMS	Monoterpene / hydroxyterpenylic acid
$C_9H_{16}O_4$	188.29	2	0.00	IVOC	Protein	Azelaic acid / monoterpene
$C_8H_{14}O_5$	190.25	2	0.00	IVOC	Protein	Diaterpenylic acid/ monoterpene
$C_9H_{12}O_5$	200.25	4	0.23	IVOC	CRAMS	Aromatics (Molteni et al., 2018)
$C_8H_{10}O_6$	202.22	4	0.20	SVOC	HOC	Trimethyl benzene isomers/ isoprene (Nguyen et al., 2011)
$C_9H_{14}O_5$	202.27	3	0.08	IVOC	Protein	Monoterpene
$C_8H_{12}O_6$	204.23	3	0.00	SVOC	HOC	Monoterpene/ Trimethylbenzene
$C_{10}H_{14}O_5$	214.29	4	0.20	IVOC	CRAMS	Monoterpene
$C_{10}H_{16}O_5$	216.30	3	0.07	IVOC	Protein	Monoterpene
$C_9H_{14}O_6$	218.27	3	0.00	SVOC	Protein	Monoterpene/ Trimethylbenzene
$C_{12}H_{14}O_4$	222.32	6	0.40	IVOC	CRAMS	Monoterpene/biomass burning
$C_{10}H_{14}O_6$	230.29	4	0.14	SVOC	CRAMS	Monoterpene
$C_{11}H_{18}O_5$	230.34	3	0.06	SVOC	Protein	Monoterpene
$C_{10}H_{16}O_6$	232.30	3	0.00	SVOC	Protein	Monoterpene
$C_{12}H_{18}O_5$	242.35	4	0.16	SVOC	CRAMS	Monoterpene
$C_{11}H_{16}O_6$	244.32	4	0.13	SVOC	CRAMS	Biogenic
$C_{12}H_{20}O_5$	244.37	3	0.05	SVOC	Protein	Biogenic
$C_{13}H_{20}O_5$	256.39	4	0.14	SVOC	Protein	β-Caryophyllene
CHON compounds						
$C_6H_5NO_3$	139.15	5	0.86	IVOC	CRAMS	Nitrophenol/catechol

$C_5H_4N_2O_3$	140.13	5	1.67	IVOC	CRAMS	Methylglyoxal +Ammonium Sulfate
$C_7H_7NO_3$	153.19	5	0.67	IVOC	CRAMS	Nitrocresol
$C_6H_5NO_4$	155.15	5	0.83	IVOC	CRAMS	4-Nitrocatechol
$C_8H_6N_2O_2$	162.20	7	1.00	SVOC	CRAMS	3-Nitroindole
$C_7H_4N_2O_3$	164.17	7	1.29	SVOC	Condensed Aromatics	\
$C_8H_7NO_3$	165.20	6	0.73	IVOC	CRAMS	Nitroacetophenone or methyl-nitrobenzaldehyde
$C_8H_9NO_3$	167.22	5	0.55	IVOC	CRAMS	Dimethyl-nitrophenol
$C_7H_7NO_4$	169.18	5	0.63	SVOC	CRAMS	2-Methyl-4-nitroresorcinol
$C_7H_5NO_5$	183.17	6	0.86	SVOC	HOC	2-Methyl-5-nitrobenzoic acid
$C_9H_{17}NO_3$	187.30	2	0.00	SVOC	Protein	Biomass burning
$C_{10}H_7NO_3$	189.24	8	0.80	SVOC	CRAMS	2-Nitro-1-naphthol
$C_9H_7NO_4$	193.22	7	0.75	SVOC	CRAMS	Biomass burning
$C_8H_7NO_5$	197.20	6	0.67	SVOC	CRAMS	Methyl-hydroxy-nitrobenzoate
$C_8H_9NO_5$	199.22	5	0.44	SVOC	CRAMS	Dimethoxy-nitrophenol
$C_7H_4N_2O_7$	228.17	7	1.67	LVOC	Others	Toluene/ 3,5-dinitrosalicylic acid
$C_6H_3N_3O_7$	229.15	7	0.00	LVOC	Others	Picric acid
$C_{23}H_{17}N_3O_2$	367.56	17	0.76	ELVOC	Unsaturated Hydrocarbons	\
$C_{23}H_{49}NO_4$	403.80	0	0.00	LVOC	Others	\
$C_{27}H_{17}N_5O_{16}$	667.64	22	0.82	ULVOC	Condensed Aromatics	\
CHOS+CHONS compounds						
$C_6H_{10}O_6S_1$	210.25	2	0	SVOC	Others	Green leaf volatiles
$C_6H_{12}O_6S_1$	212.26	1	0	SVOC	Others	Monoterpene/ olefinic acid
$C_5H_{10}O_7S_1$	214.23	1	0	LVOC	Others	Isoprene/ olefinic acid
$C_7H_{12}O_6S_1$	224.28	2	0	SVOC	Carbohydrates	Monoterpene
$C_6H_{10}O_7S_1$	226.25	2	0	LVOC	Others	Isoprene/ green leaf volatiles
$C_7H_{14}O_6S_1$	226.30	1	0	SVOC	Carbohydrates	Olefinic acid
$C_6H_{12}O_7S_1$	228.26	1	0	LVOC	Others	Isoprene
$C_7H_{12}O_7S_1$	240.28	2	0	LVOC	Others	Isoprene
$C_6H_{10}O_8S_1$	242.25	2	0	LVOC	Others	\
$C_9H_{16}O_6S_1$	252.35	2	0	SVOC	Protein	Monoterpene
$C_8H_{14}O_7S_1$	254.32	2	0	LVOC	Carbohydrates	Monoterpene
$C_7H_{12}O_8S_1$	256.28	2	0	LVOC	Others	Isoprene
$C_8H_{12}O_8S_1$	268.30	3	0	LVOC	HOC	Monoterpene
$C_{10}H_{16}O_7S_1$	280.37	3	0	LVOC	Carbohydrates	Monoterpene
C9H14O8S1	282.33	3	0	LVOC	Carbohydrates	Monoterpene
C10H18O7S1	282.38	2	0	LVOC	Carbohydrates	Monoterpene
C8H12O9S1	284.30	3	0	LVOC	Others	\
C9H16O8S1	284.35	2	0	LVOC	Carbohydrates	Monoterpene
C10H17O7N1S1	295.38	3	0	ELVOC	Carbohydrates	Monoterpene
C10H16O8S1	296.37	3	0	LVOC	Carbohydrates	Monoterpene

Table S4. Intensity-weighted molecular characteristics of organic compounds associated with NMF-derived OA factors

	Cooking	IsopreneSOA	Biomass burning	Gas-pSOA	SIA
MW	216	213	217	221	214
C	8.91	9.44	10.00	9.25	9.76
H	12.75	13.38	12.92	12.78	13.76
O	5.25	5.02	4.81	5.42	4.83
N	0.22	0.11	0.33	0.20	0.17
S	0.28	0.12	0.07	0.23	0.10
H/C	0.63	0.59	0.52	0.63	0.56
O/C	1.44	1.42	1.28	1.38	1.40
N/C	0.03	0.01	0.04	0.03	0.02
S/C	0.04	0.01	0.01	0.03	0.01
O/N	0.88	0.30	1.09	0.71	0.40
O/S	1.84	0.75	0.43	1.59	0.56
N/S	0.09	0.04	0.03	0.04	0.03
DBE	3.64	3.81	4.71	3.96	3.97
DBE/C	0.41	0.42	0.49	0.44	0.43
nO_{eff}	3.97	4.43	3.94	4.34	4.18
AI_{mod}	0.20	0.22	0.34	0.24	0.24
Xc	1.40	1.28	1.70	1.46	1.40
LogC*	1.42	2.12	1.90	1.39	2.09
Compounds classes					
DBE=0	0.68	1.76	0.30	0.65	1.2
DBE=1	5.8	4.6	4.0	4.8	3.9
DBE=2	19	17	13	18	14
DBE=3	29	28	22	25	22
DBE>=4	45	49	61	51	58
Condensed Aromatics	3.7	3.3	5.6	3.0	3.1
CRAMS	36	34	43	36	38
Lipids	4.7	7.4	10.6	6.1	9.1
Protein	18	25	20	21	19
HOC	12	11	10	12	13
Carbohydrates	12	5	4	8	2
Unsaturated Hydrocarbons	0.88	1.25	0.72	0.56	2.7
Others	13	14	5	13	13

References

Ditto, J. C., Barnes, E. B., Khare, P., Takeuchi, M., Joo, T., Bui, A. A. T., Lee-Taylor, J., Eris, G., Chen, Y., Aumont, B., Jimenez, J. L., Ng, N. L., Griffin, R. J., and Gentner, D. R.: An omnipresent diversity and variability in the chemical composition of atmospheric functionalized organic aerosol, Communications Chemistry, 1, 75, doi:10.1038/s42004-018-0074-3, 2018.

Huo, Y., Yao, D., and Guo, H.: Differences in aerosol chemistry at a regional background site in Hong Kong before and during the COVID-19 pandemic, Sci Total Environ, 926, 171990, doi:10.1016/j.scitotenv.2024.171990, 2024a.

Huo, Y., Lyu, X., Yao, D., Zhou, B., Yuan, Q., Lee, S. c., and Guo, H.: Exploring the Formation of High Levels of Hydroxyl Dicarboxylic Acids at an Urban Background Site in South China, Journal of Geophysical Research: Atmospheres, 129, e2023JD040096, doi:10.1029/2023jd040096, 2024b.

Koch, B. P. and Dittmar, T.: From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter, Rapid Communications in Mass Spectrometry, 20, 926–932, doi:10.1002/rcm.2386, 2006.

Li, Y., Pöschl, U., and Shiraiwa, M.: Molecular corridors and parameterizations of volatility in the chemical evolution of organic aerosols, Atmospheric Chemistry and Physics, 16, 3327-3344, doi:10.5194/acp-16-3327-2016, 2016.

Lyu, X., Guo, H., Yao, D., Lu, H., Huo, Y., Xu, W., Kreisberg, N., Goldstein, A. H., Jayne, J., Worsnop, D., Tan, Y., Lee, S. C., and Wang, T.: In Situ Measurements of Molecular Markers Facilitate Understanding of Dynamic Sources of Atmospheric Organic Aerosols, Environ Sci Technol, 54, 11058-11069, doi:10.1021/acs.est.0c02277, 2020.

Mehra, A., Wang, Y., Krechmer, J. E., Lambe, A., Majluf, F., Morris, M. A., Priestley, M., Bannan, T. J., Bryant, D. J., Pereira, K. L., Hamilton, J. F., Rickard, A. R., Newland, M. J., Stark, H., Croteau, P., Jayne, J. T., Worsnop, D. R., Canagaratna, M. R., Wang, L., and Coe, H.: Evaluation of the chemical composition of gas- and particle-phase products of aromatic oxidation, Atmospheric Chemistry and Physics, 20, 9783-9803, doi:10.5194/acp-20-9783-2020, 2020.

Molteni, U., Bianchi, F., Klein, F., El Haddad, I., Frege, C., Rossi, M. J., Dommen, J., and Baltensperger, U.: Formation of highly oxygenated organic molecules from aromatic compounds, Atmospheric Chemistry and Physics, 18, 1909-1921, doi:10.5194/acp-18-1909-2018, 2018.

Nguyen, T. B., Roach, P. J., Laskin, J., Laskin, A., and Nizkorodov, S. A.: Effect of humidity on the composition of isoprene photooxidation secondary organic aerosol, Atmospheric Chemistry and Physics, 11, 6931-6944, doi:10.5194/acp-11-6931-2011, 2011.

Nie, W., Yan, C., Huang, D. D., Wang, Z., Liu, Y., Qiao, X., Guo, Y., Tian, L., Zheng, P., Xu, Z., Li, Y., Xu, Z., Qi, X., Sun, P., Wang, J., Zheng, F., Li, X., Yin, R., Dallenbach, K. R., Bianchi, F., Petäjä, T., Zhang, Y., Wang, M., Schervish, M., Wang, S., Qiao, L., Wang, Q., Zhou, M., Wang, H., Yu, C., Yao, D., Guo, H., Ye, P., Lee, S., Li, Y. J., Liu, Y., Chi, X., Kerminen, V.-M., Ehn, M., Donahue, N. M., Wang, T., Huang, C., Kulmala, M., Worsnop, D., Jiang, J., and Ding, A.: Secondary organic aerosol formed by condensing anthropogenic vapours over China's megacities, Nature Geoscience, 15, 255-261, doi:10.1038/s41561-022-00922-5, 2022.

Sun, Y. L., Zhang, Q., Schwab, J. J., Demerjian, K. L., Chen, W. N., Bae, M. S., Hung, H. M., Hogrefe, O., Frank, B., Rattigan, O. V., and Lin, Y. C.: Characterization of the sources and processes of organic and inorganic aerosols in New York city with a high-resolution time-of-flight aerosol mass apectrometer, Atmospheric Chemistry and Physics, 11, 1581-1602, doi:10.5194/acp-11-1581-2011, 2011.

Thoma, M., Bachmeier, F., Gottwald, F. L., Simon, M., and Vogel, A. L.: Mass spectrometry-based Aerosolomics: a new approach to resolve sources, composition, and partitioning of secondary organic aerosol, Atmospheric Measurement Techniques, 15, 7137-7154, doi:10.5194/amt-15-7137-2022, 2022.

Wang, X., Hayeck, N., Brüggemann, M., Yao, L., Chen, H., Zhang, C., Emmelin, C., Chen, J., George, C., and Wang, L.: Chemical Characteristics of Organic Aerosols in Shanghai: A Study by Ultrahigh-Performance Liquid Chromatography Coupled With Orbitrap Mass Spectrometry, J. Geophys. Res. Atmos., 122, 11703–11722, doi:10.1002/2017jd026930, 2017.

Yao, D., Guo, H., Lyu, X., Lu, H., and Huo, Y.: Secondary organic aerosol formation at an urban background site on the coastline of South China: Precursors and aging processes, Environmental Pollution, 309, 119778, doi:10.1016/j.envpol.2022.119778, 2022.

Yuan, Q., Zhang, Z., Chen, Y., Hui, L., Wang, M., Xia, M., Zou, Z., Wei, W., Ho, K. F., Wang, Z., Lai, S., Zhang, Y., Wang, T., and Lee, S.: Origin and transformation of volatile organic compounds at a regional background site in Hong Kong: Varied photochemical processes from different source regions, Sci Total Environ, 908, 168316, doi:10.1016/j.scitotenv.2023.168316, 2024.
Zhang, M., Cai, D., Lin, J., Liu, Z., Li, M., Wang, Y., and Chen, J.: Molecular characterization of atmospheric organic aerosols in typical megacities in China, npj Climate and Atmospheric Science, 7, 230, doi:10.1038/s41612-024-00784-1, 2024.
Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Ulbrich, I. M., Ng, N. L., Worsnop, D. R., and Sun, Y.: Understanding atmospheric organic aerosols via factor analysis of aerosol mass spectrometry: a review, Anal Bioanal Chem, 401, 3045-3067, doi:10.1007/s00216-011-5355-y, 2011.