Supporting Material to:

Nontarget Screening Exhibits a Seasonal Cycle of $PM_{2.5}$ Organic Aerosol Composition in Beijing

Jialiang Ma [†], *Florian Ungeheuer* [†], *Feixue Zheng* [‡], *Wei Du* [‡, §], *Yonghong Wang* [‡, §, ǁ], *Jing Cai* [‡, §], *Ying Zhou* [‡], *Chao Yan* [‡, §], *Yongchun Liu* [‡], *Markku Kulmala* [‡, §], *Kaspar R. Daellenbach* [‡, §, ⊥, *], *and Alexander L. Vogel* [†, *]

[†] Institute for Atmospheric and Environmental Sciences, Goethe-University Frankfurt, 60438 Frankfurt am Main, Germany

[‡]Aerosol and Haze Laboratory, Beijing Advanced Innovation Center for Soft Matter Science and Engineering, Beijing University of Chemical Technology, 100029 Beijing, China

[§] Institute for Atmospheric and Earth System Research/Physics, Faculty of Science, University of Helsinki, 00014 Helsinki, Finland.

[⊥] Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, 5232 Villigen, Switzerland

[ǁ] Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, 100085 Beijing, China

*(K.R.D.) E-mail: kaspar.daellenbach@psi.ch

*(A.L.V.) E-mail: vogel@iau.uni-frankfurt.de

Document Contains:

25 Pages

13 Figures

2 Tables

Contents

I, Aerosol Chemical Speciation Monitor Measurement

The Aerodyne Time-of-Flight Aerosol Chemical Speciation Monitor (ACSM) was deployed at Beijing University of Chemical Technology (BUCT, 39°58'N, 116°25'E), China. BUCT is an urban site which is located north-east of Beijing. A $PM_{2.5}$ cyclone was installed in front of the sampling line on the rooftop with a flow rate of 3 L/min and the aerosol was dried using a Nafion dryer (Perma Pure, MD-700-24F-3).

II, Figure S1, PM2.5 mass correlation

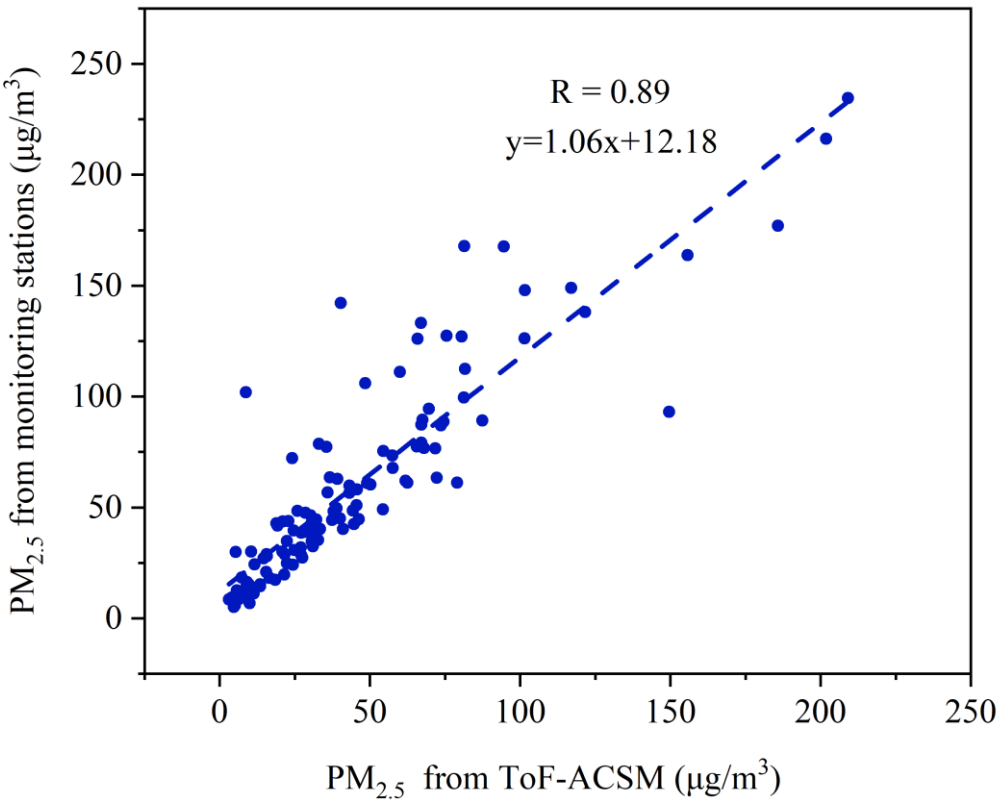


Figure S1. Concentration correlation of $PM_{2.5}$ mass concentrations measured by ToF-ACSM and the mean $PM_{2.5}$ mass concentrations measured at Wanliu, Gucheng, Wanshouxigong, and Guanyuan monitoring stations.

III, Figure S2, Reproducibility of signal intensity

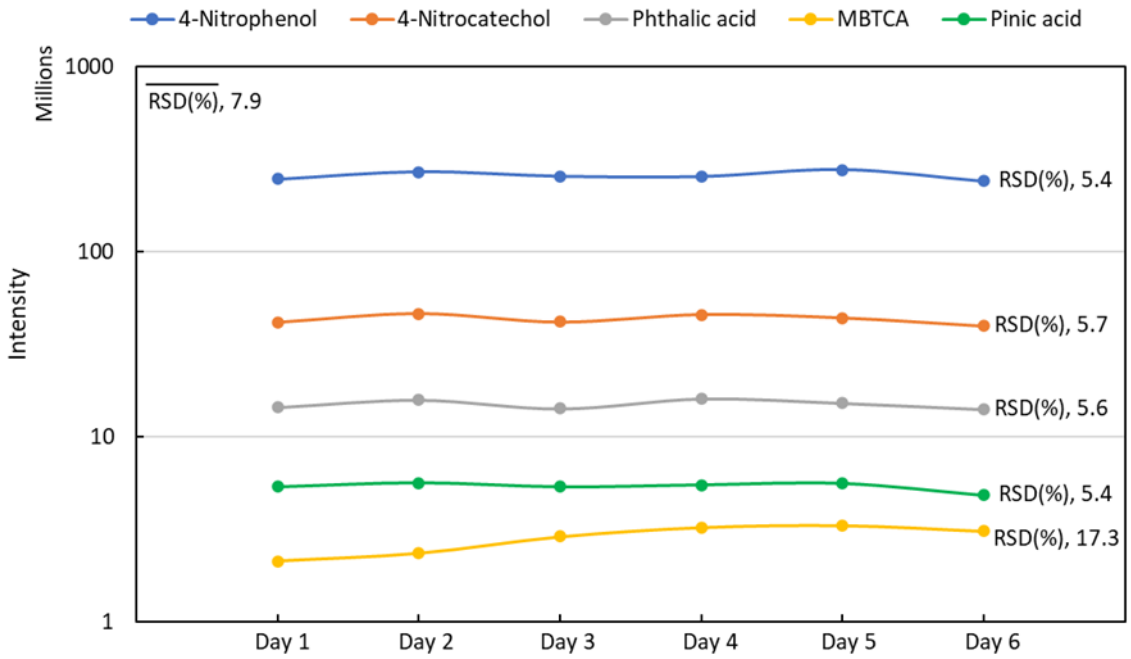


Figure S2. Intensity variation of the five quantified compounds (4-Nitrophenol, 4-Nitrocatechol, Phthalic acid, 3-methyl-1,2,3-butanetricarboxylic acid (MBTCA), Pinic acid) of the sample which was measured repeatedly on a daily basis to monitor the performance of the instrument (sample of 30.03.2019).

IV, Figure S3, Mass accuracy during measurement period

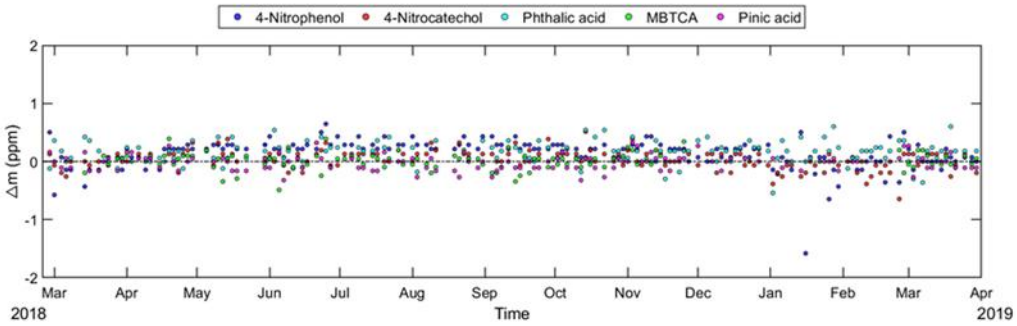


Figure S3. The mass accuracy (deviation between theoretical and measured mass) of 4-Nitrophenol, 4-Nitrocatechol, Phthalic acid, MBTCA, and Pinic acid in all measured samples.

V, Table S1, Detailed settings of the CD-workflow

Processing node 1: Select Spectra

---

1. General Settings:

- Precursor Selection: Use MS1 Precursor
- Use Isotope Pattern in Precursor Reevaluation: True
- Provide Profile Spectra: Automatic
- Store Chromatograms: False

2. Spectrum Properties Filter:

- Lower RT Limit: 0
- Upper RT Limit: 0
- First Scan: 0
- Last Scan: 0
- Ignore Specified Scans: (not specified)
- Lowest Charge State: 0
- Highest Charge State: 0
- Min. Precursor Mass: 0 Da
- Max. Precursor Mass: 5000 Da
- Total Intensity Threshold: 1000000
- Minimum Peak Count: 1

3. Scan Event Filters:

- Mass Analyzer: Is FTMS
- MS Order: Is MS2; MS1
- Activation Type: Is HCD
- Min. Collision Energy: 0
- Max. Collision Energy: 1000
- Scan Type: Any
- Polarity Mode: (not specified)

4. Peak Filters:

- S/N Threshold (FT-only): 1.5

5. Replacements for Unrecognized Properties:

- Unrecognized Charge Replacements: 1
- Unrecognized Mass Analyzer Replacements: ITMS
- Unrecognized MS Order Replacements: MS2
- Unrecognized Activation Type Replacements: CID
- Unrecognized Polarity Replacements: -
- Unrecognized MS Resolution@200 Replacements: 60000
- Unrecognized MSn Resolution@200 Replacements: 30000

Processing node 2: Align Retention Times

---

1. General Settings:

- Alignment Model: Adaptive curve
- Alignment Fallback: Use Linear Model
- Maximum Shift [min]: 0.5
- Shift Reference File: True
- Mass Tolerance: 4 ppm
- Remove Outlier: True

Processing node 3: Detect Compounds

---

1. General Settings:

- Mass Tolerance [ppm]: 4 ppm
- Intensity Tolerance [%]: 10
- S/N Threshold: 3
- Min. Peak Intensity: 500000
- Ions:
    [2M+FA-H]-1
    [2M-H]-1

[M+FA-H]-1

[M-H]-1

[M-H-H2O]-1

- Base Ions: [M-H]-1
- Min. Element Counts: C H
- Max. Element Counts: C90 H190 Br3 Cl4 N4 O20 P S3

2. Peak Detection:

- Filter Peaks: True
- Max. Peak Width [min]: 0.3
- Remove Singlets: True
- Min. # Scans per Peak: 5
- Min. # Isotopes: 2

Processing node 5: Group Compounds

---

1. Compound Consolidation:

- Mass Tolerance: 2 ppm
- RT Tolerance [min]: 0.1

2. Fragment Data Selection:

- Preferred Ions: [M-H]-1

Processing node 6: Fill Gaps

---

1. General Settings:

- Mass Tolerance: 2 ppm
- S/N Threshold: 1.5
- Use Real Peak Detection: True

Processing node 7: Mark Background Compounds

---

1. General Settings:

- Max. Sample/Blank: 3
- Max. Blank/Sample: 0
- Hide Background: False

Processing node 9: Assign Compound Annotations

---

1. General Settings:

- Mass Tolerance: 2 ppm

2. Data Sources:

- Data Source #1: mzCloud Search
- Data Source #2: Predicted Compositions
- Data Source #3: (not specified)
- Data Source #4: ChemSpider Search
- Data Source #5: (not specified)
- Data Source #6: (not specified)
- Data Source #7: (not specified)

3. Scoring Rules:

- Use mzLogic: True
- Use Spectral Distance: True
- SFit Threshold: 20
- SFit Range: 20

Processing node 10: Search mzCloud

---

1. General Settings:

- Compound Classes: All
- Precursor Mass Tolerance: 5 ppm
- FT Fragment Mass Tolerance: 5 ppm
- IT Fragment Mass Tolerance: 0.4 Da

- Library: Autoprocessed; Reference
- Post Processing: Recalibrated
- Max. # Results: 10
- Annotate Matching Fragments: False

2. DDA Search:
- Identity Search: HighChem HighRes
- Match Activation Type: True
- Match Activation Energy: Match with Tolerance
- Activation Energy Tolerance: 20
- Apply Intensity Threshold: True
- Similarity Search: Similarity Forward
- Match Factor Threshold: 60

3. DIA Search:
- Use DIA Scans for Search: False
- Max. Isolation Width [Da]: 500
- Match Activation Type: False
- Match Activation Energy: Any
- Activation Energy Tolerance: 100
- Apply Intensity Threshold: False
- Match Factor Threshold: 20

Processing node 8: Predict Compositions

---

1. Prediction Settings:
- Mass Tolerance: 2 ppm
- Min. Element Counts: C H
- Max. Element Counts: C90 H190 Br3 Cl4 N4 O20 P S3
- Min. RDBE: 0
- Max. RDBE: 40
- Min. H/C: 0.1

- Max. H/C: 3.5
- Max. # Candidates: 10
- Max. # Internal Candidates: 200

2. Pattern Matching:
- Intensity Tolerance [%]: 10
- Intensity Threshold [%]: 0.1
- S/N Threshold: 3
- Min. Spectral Fit [%]: 30
- Min. Pattern Cov. [%]: 90
- Use Dynamic Recalibration: True

3. Fragments Matching:
- Use Fragments Matching: True
- Mass Tolerance: 5 ppm
- S/N Threshold: 3

Processing node 11: Search ChemSpider

---

1. Search Settings:
- Database(s):
    EAWAG Biocatalysis/Biodegradation Database
    Nature Chemistry
    Sigma-Aldrich
- Search Mode: By Formula Only
- Mass Tolerance: 2 ppm
- Max. # of results per compound: 100
- Max. # of Predicted Compositions to be searched per Compound: 3
- Result Order (for Max. # of results per compound): Order By Reference Count (DESC)

2. Predicted Composition Annotation:
- Check All Predicted Compositions: True

Processing node 12: Apply mzLogic

---

1. Search Settings:

- FT Fragment Mass Tolerance: 10 ppm

- IT Fragment Mass Tolerance: 0.4 Da

- Max. # Compounds: 0

- Max. # mzCloud Similarity Results to consider per Compound: 10

- Match Factor Threshold: 30

Processing node 13: Apply Spectral Distance

---

1. Pattern Matching:

- Mass Tolerance: 5 ppm

- Intensity Tolerance [%]: 30

- Intensity Threshold [%]: 0.1

- S/N Threshold: 3

- Use Dynamic Recalibration: True

Processing node 4: Merge Features

---

1. Peak Consolidation:

- Mass Tolerance: 2 ppm

- RT Tolerance [min]: 0.2

VI, Figure S4, Data preparation for hierarchical cluster analysis

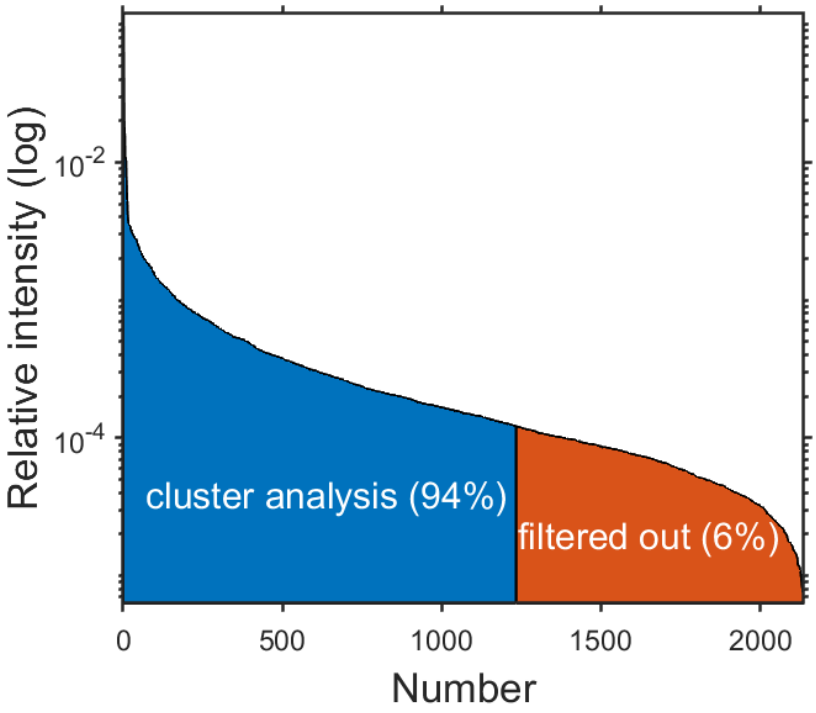


Figure S4. Data preparation for hierarchical cluster analysis.

VII, Figure S5, Molecular fingerprints (average intensity)

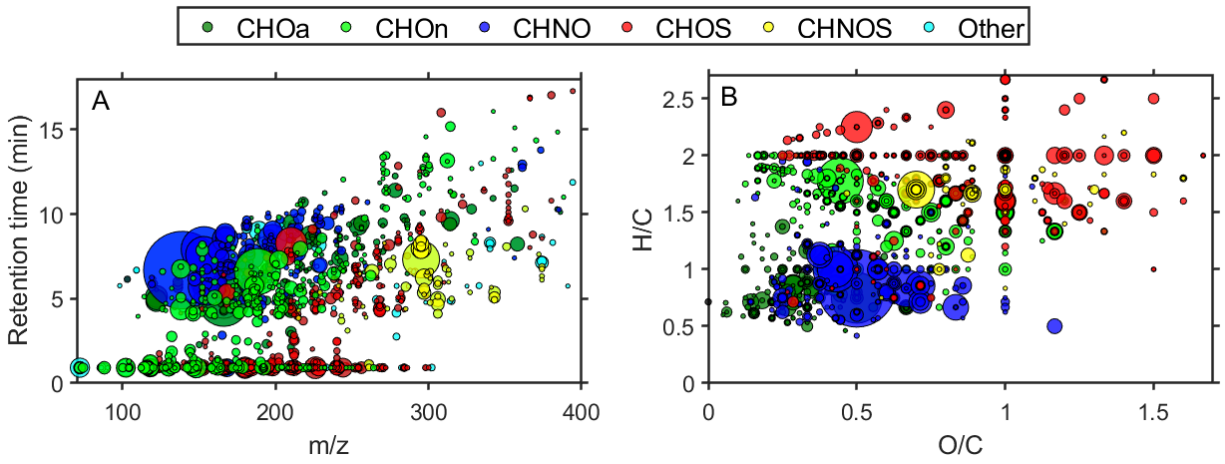


Figure S5. Molecular fingerprints of average intensity for 13 months' samples (after data filtering). (A), *m/z* vs RT plot. (B), Van Krevelen-diagram.

VIII, Figure S6, The PM2.5 mass distribution of clusters SH, SL, WL, and WH

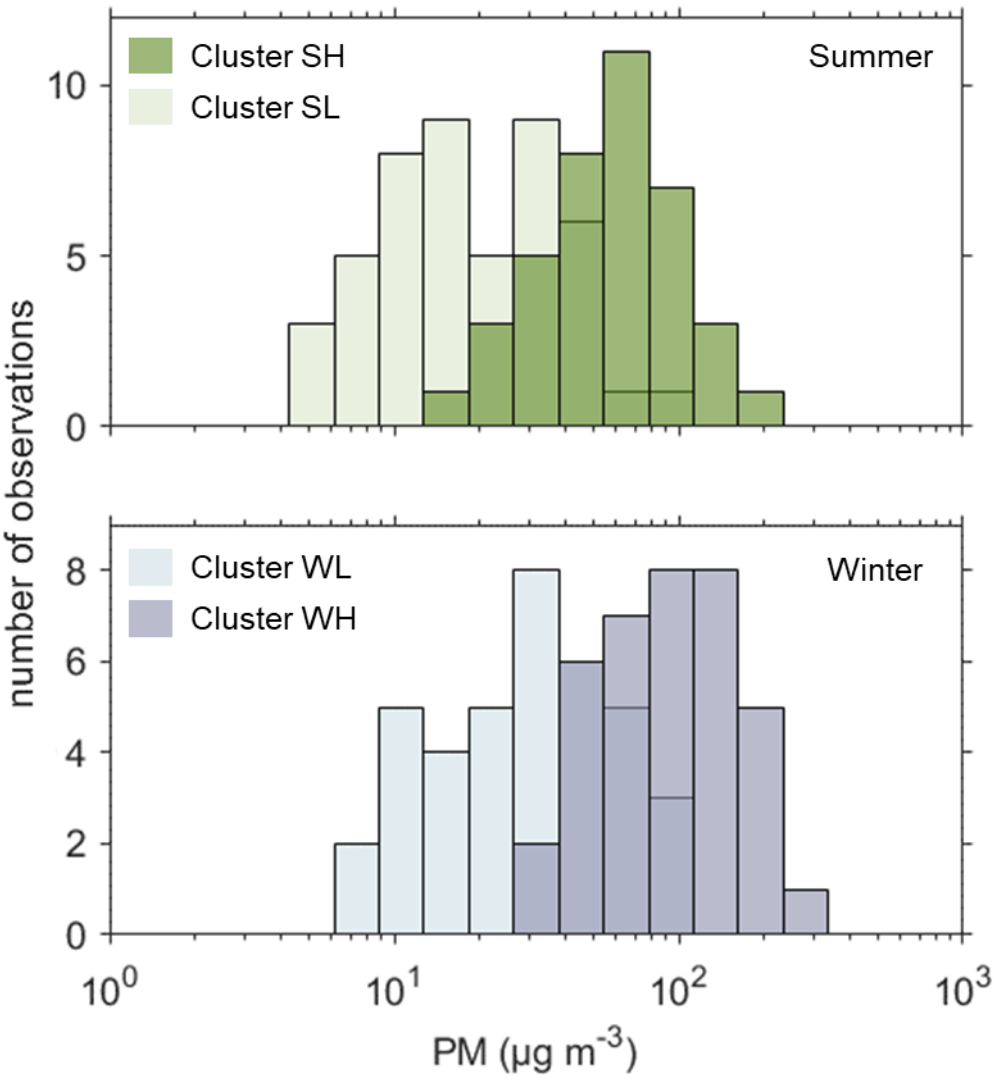


Figure S6. The $PM_{2.5}$ mass distribution of clusters SH, SL, WL, and WH. The standard deviations of the four clusters are 37, 16, 26, and 52 μg m^{-3}, respectively.

IX, Figure S7, Meteorological conditions during the sampling period

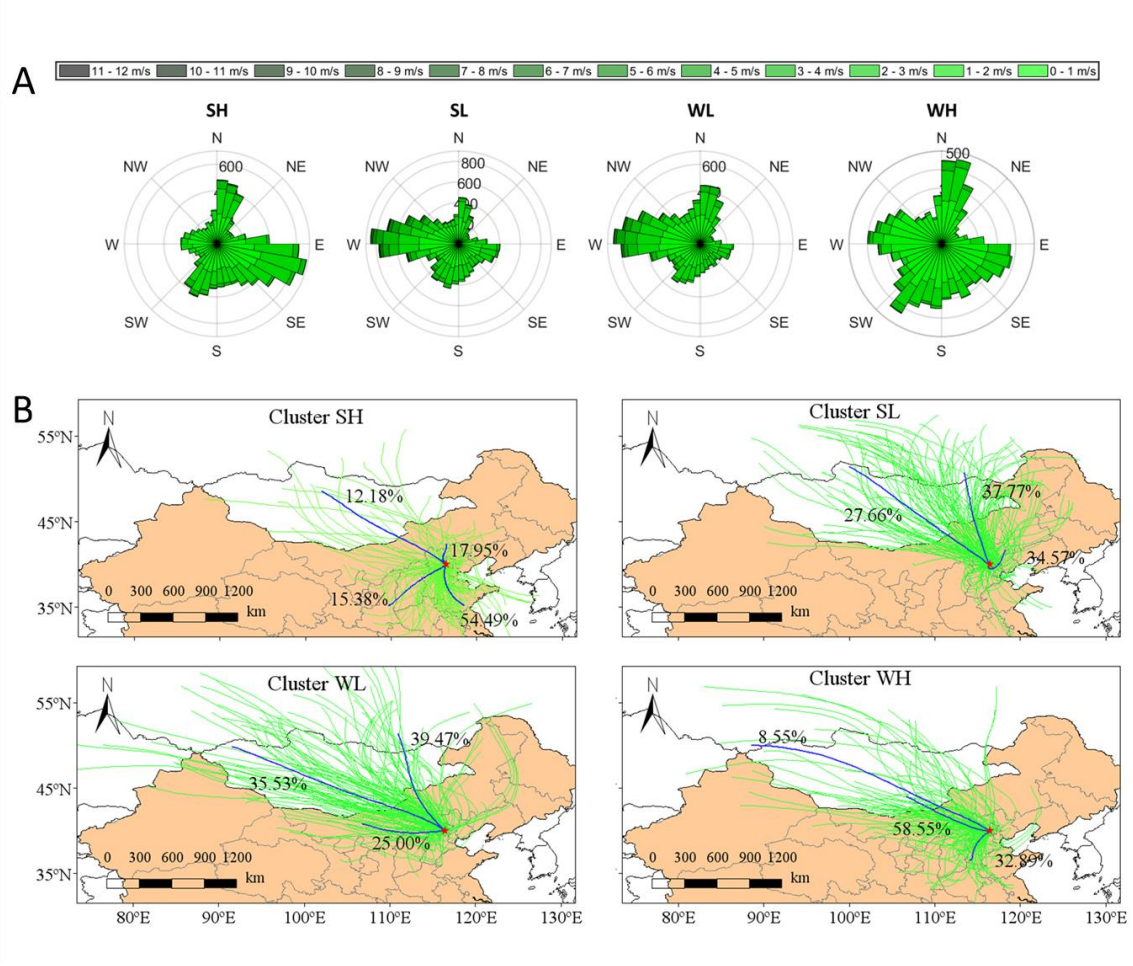
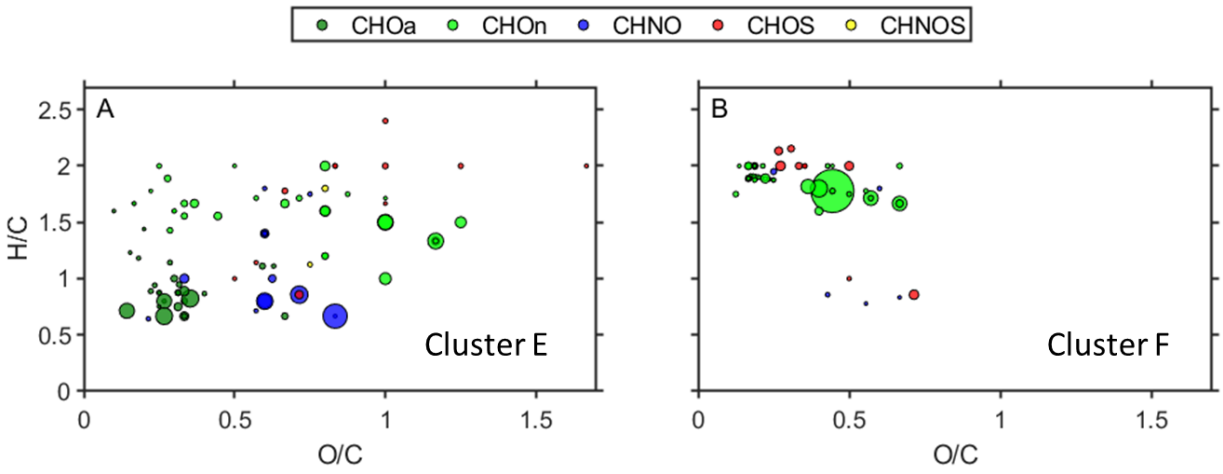


Figure S7. Meteorological conditions during the sampling period. (A) Wind direction and wind speed at Beijing University of Chemical Technology (BUCT); (B) Comparison of air mass trajectories for the different clusters based on 48 h backward HYSPLIT-trajectories.[1]

X, Figure S8, Van Krevelen-diagram of clusters E and F



315

Figure S8. Van Krevelen-diagram (average intensity for 13 months' samples) of (A) cluster E and (B) cluster F from HCA.

XI, Figure S9, MS/MS fragmentation patterns

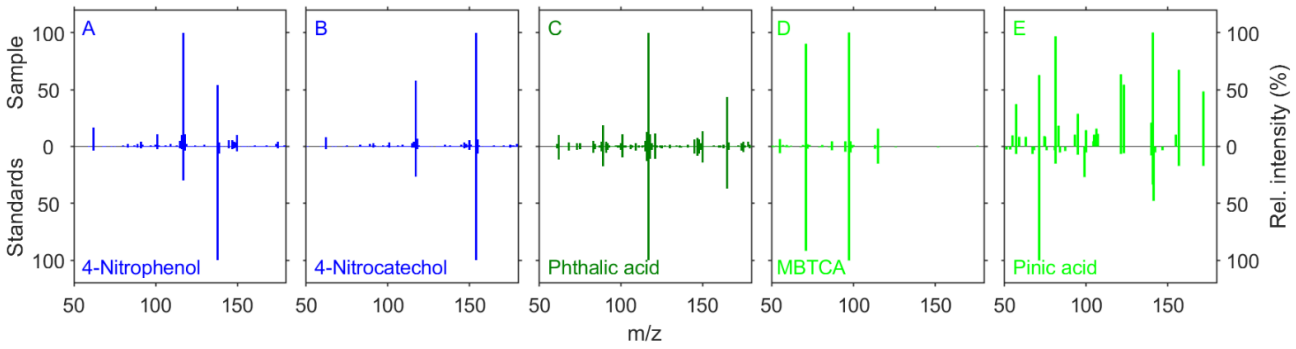


Figure S9. Comparison of the MS/MS fragmentation patterns of (A) 4-nitrophenol, (B) 4-nitrocatechol, (C) phthalic acid, (D) MBTCA, and (E) pinic acid measured in ambient samples (upward spectra) and their standards (downward spectra).

XII, Figure S10, Kroll-diagram of clusters B, C, and D

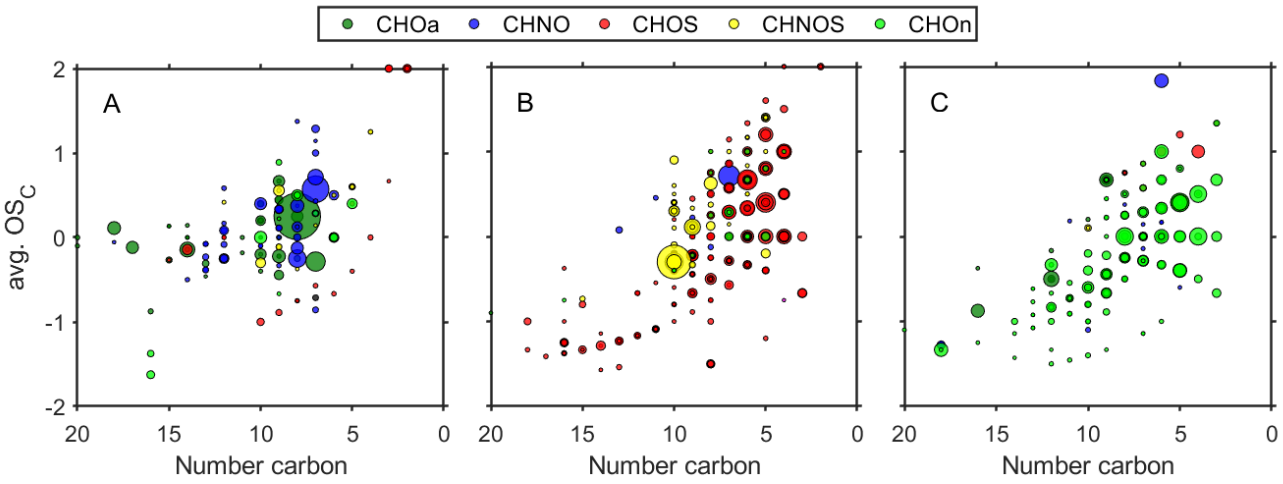


Figure S10. Kroll-diagram of (A) cluster B, (B) cluster C, and (E) cluster D from HCA.

XIII, Figure S11, Calibration curves

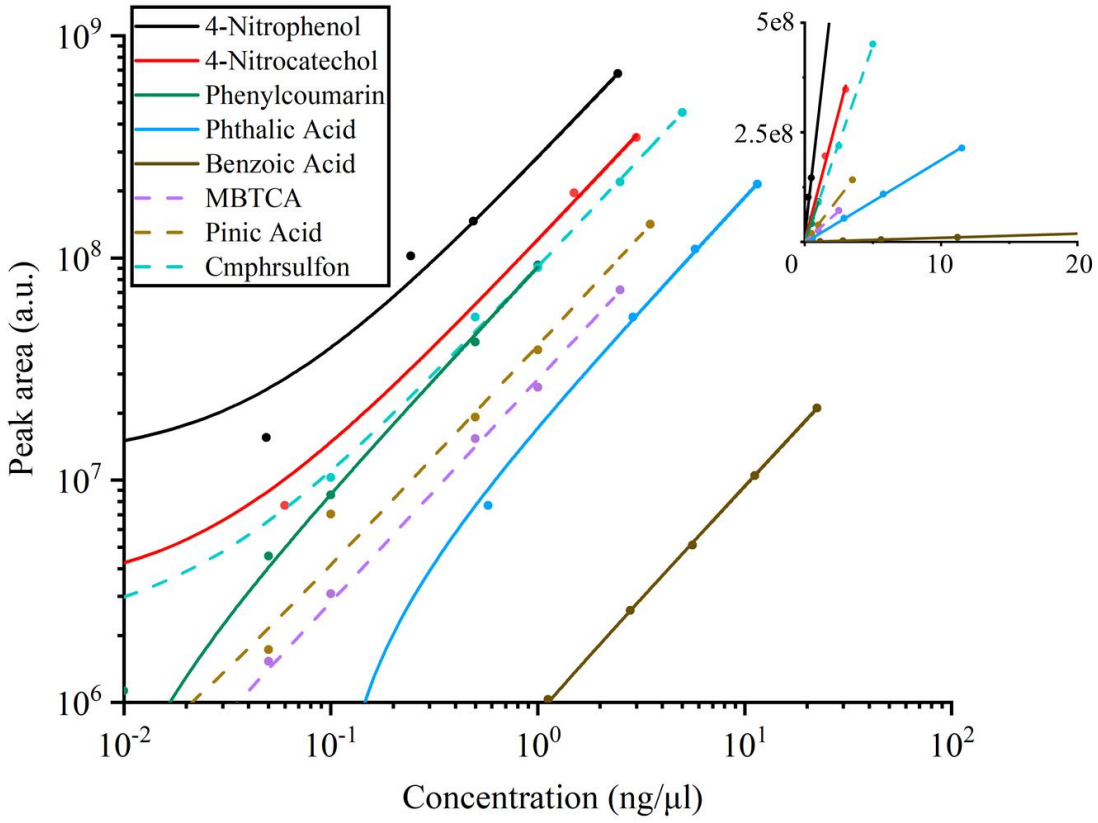


Figure S11. Calibration curves of selected compounds detected by (−)ESI, solid lines represent aromatic compounds while dashed lines non-aromatic compounds. The inset shows the calibration curves on a linear scale. Abbreviations: Phenylcoumarin (5, 7-Dihydroxy-4-phenylcoumarin), MBTCA (3-methyl-1, 2, 3-butanetricarboxylic acid), and Cmphrsulfon (camphor-10-sulfonic acid).

XIV, Figure S12, Relationship between clusters C and D and sulfate

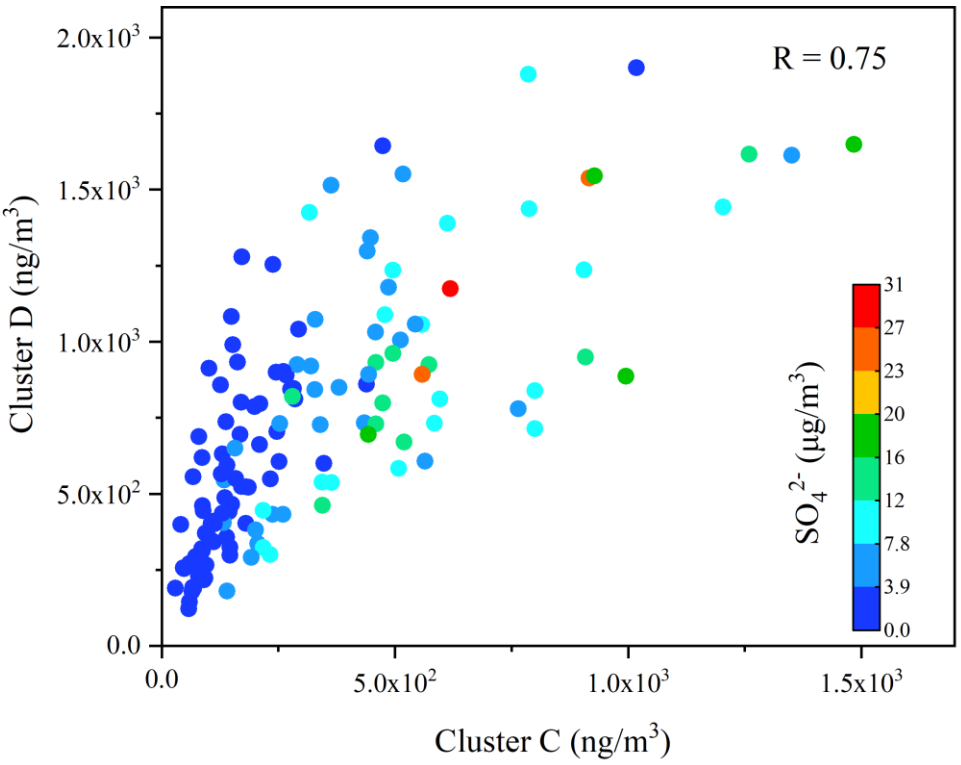


Figure S12. Correlation coefficient of total compound concentrations in cluster C and cluster D. The circles are colored according to the SO_4^{2-} mass concentration (data from ACSM).

XV, Figure S13, Fraction of the four clusters to the total OA

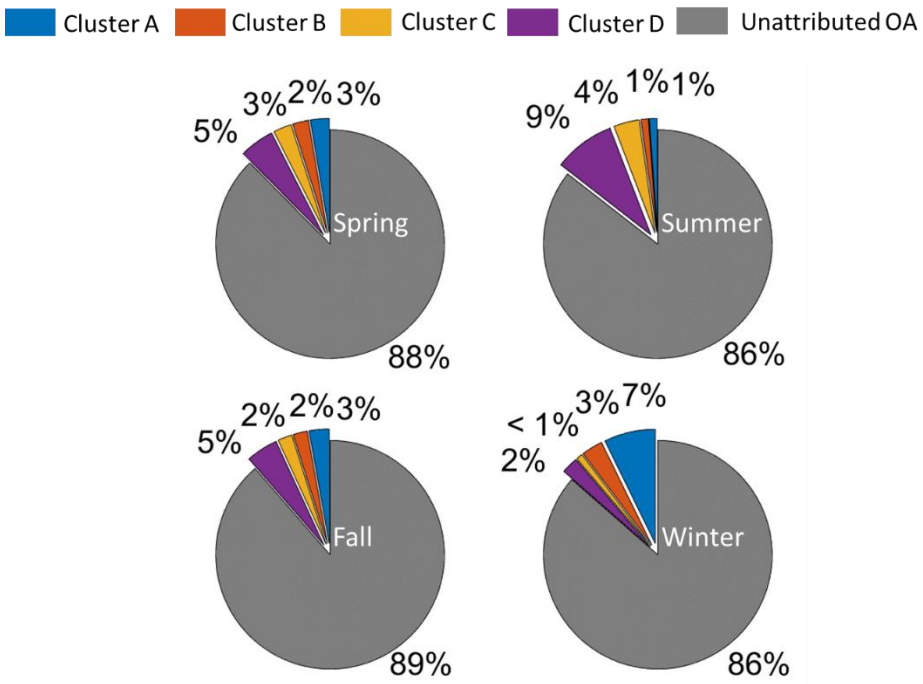
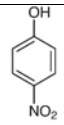
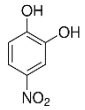
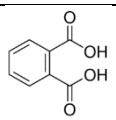
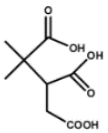
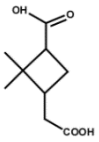


Figure S13. The concentration fraction of the four clusters to the total OA (data from ACSM) in spring, summer, autumn, and winter.

XVI, Table S2, Detailed information for identified molecules

Common name	Formula	Structure	Theoretical $[M-H]^-$	△Mass (ppm)
4-Nitrophenol	$C_6H_5NO_3$		138.0197	-0.72
4-Nitrocatechol	$C_6H_5NO_4$		154.0146	-0.39
Phthalic acid	$C_8H_6O_4$		165.0193	-0.76
MBTCA	$C_8H_{12}O_6$		203.0561	-0.14
Pinic acid	$C_9H_{14}O_4$		185.0819	-0.52

△Mass (ppm): Difference between measured and theoretical mass in ppm.

Reference:

(1) Stein, A. F.; Draxler, R. R.; Rolph, G. D.; Stunder, B. J. B.; Cohen, M. D.; Ngan, F. NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System. *Bull. Am. Meteorol. Soc.* **2015**, *96* (12), 2059–2077. https://doi.org/10.1175/BAMS-D-14-00110.1.