



Supplement of

The sources and diurnal variations of submicron aerosols in a coastal–rural environment near Houston, US

Jing Li et al.

Correspondence to: Lu Xu (xu1@wustl.edu) and Jian Wang (jian@wustl.edu)

The copyright of individual parts of the supplement might differ from the article licence.

Section 1. Detailed Procedures for PMF on ACSM Data

PMF analysis was conducted on the ToF-ACSM mass spectra to identify the key OA components and investigate their sources (Paatero, 1997; Paatero and Tapper, 1994). Briefly, in the bilinear mode used here, PMF describes the measured data matrix X as a product of two matrices, G and F , and the residual matrix E :

$$X = GF + E \quad (1)$$

Each column in matrix G represents the time series of a factor profile (mass spectrum) and each row in the matrix F represents the corresponding factor profile. In order to solve equation (1), quantity Q is minimized with respect to all model variables:

$$Q = \sum_{i=1}^n \sum_{j=1}^m \left(\frac{E_{ij}}{\sigma_{ij}} \right)^2 \quad (2)$$

Where σ corresponds to the matrix of measurement uncertainties of X .

In this study, we base the PMF approach on that described in Canonaco et al. (2021), in that we use the multilinear engine (ME-2, Paatero, 1999) algorithm, constrain certain profiles with reference mass spectra, and perform multiple PMF runs on a subset of data defined by a temporal window moved in 1 day increments across the entire dataset (“rolling PMF” strategy). This method allows for better capturing of seasonal variations in factor profiles (Canonaco et al., 2021). We carry out the PMF calculations using the Source Finder (SoFi) software package (Canonaco et al., 2013). However, the clustering and post-processing averaging steps were performed with custom routines written in the MATLAB environment (MATLAB 2020b).

S1.1 Unconstrained PMF

Before performing rolling PMF constrained with reference profiles, we carried out an unconstrained PMF run to identify a plausible set of reference profiles. The TRACER TOF-ACSM organic matrix and its associated error matrix generated by the TOFWerk TOFware software package were first averaged to a 90-minute sampling frequency. An unconstrained PMF run with random seeds was performed for four, five and six factors, 350 times each. Subsequently, each set of 350 PMF runs was then clustered using the k-means algorithm.

The optimized number of PMF profiles for each season is selected based on the calculated average silhouette value of the k-means clustering solution for 4, 5 and 6 factors. The silhouette value for each point is a measure of how similar that point is to other points in the same cluster, compared to points in other clusters. The silhouette value s_i for the i^{th} point is defined as

$$s_i = \frac{(b_i - a_i)}{\max(a_i, b_i)} \quad (3)$$

where a_i is the average distance from the i^{th} point to the other points in the same cluster as i , and b_i is the minimum average distance from the i^{th} point to points in a different cluster, minimized over the clusters. If the i^{th} point is the only point in its cluster, then the silhouette value s_i is set to 1. A high silhouette value indicates that the point is well matched to its own cluster, and poorly matched to other clusters. If most points have a high silhouette value, then the clustering solution is appropriate. If many points have a low or negative silhouette value, then the clustering solution might have too many or too few clusters. We select the appropriate seasonal PMF solution based on the highest silhouette value.

Fig. S1 shows average silhouette values for each unconstrained PMF solution. The five-factor solution has the highest silhouette value, which suggests that it captures the most reasonable number of factors. Figs. S2-S4 show the PMF factors for each solution and plots of key m/z markers for individual unconstrained PMF runs (f_{43} , f_{44} , f_{55} , f_{57} , f_{60} , f_{82} and f_{91}).

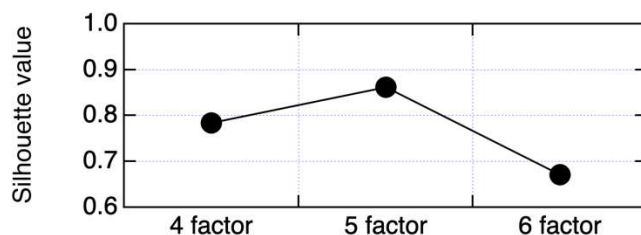


Figure S1. Silhouette values for 4-, 5- and 6-factor unconstrained PMF solutions.

Each solution includes a profile identified as HOA (black in Figs. S2-S4) based on a relatively high value of f_{55} and f_{57} . Each solution also includes a profile identified as MT-SOA (darkest green in Figs. S2-S4) identified based on relatively high f_{43} and f_{91} . Five and six factor solutions also include a factor identified as isoprene-SOA (blue in Figs. S3-S4), identified based on prominent f_{82} . Beyond this, all solutions include 1-3 aged OOA factors (lighter green in Figs. S2-S4) based on a high f_{44} to f_{43} ratio.

Based on the unconstrained runs, the rolling PMF solution was constrained with HOA, MT-SOA, and isoprene-SOA factors from the five-factor unconstrained solution. These reference factors are shown in Fig. S5

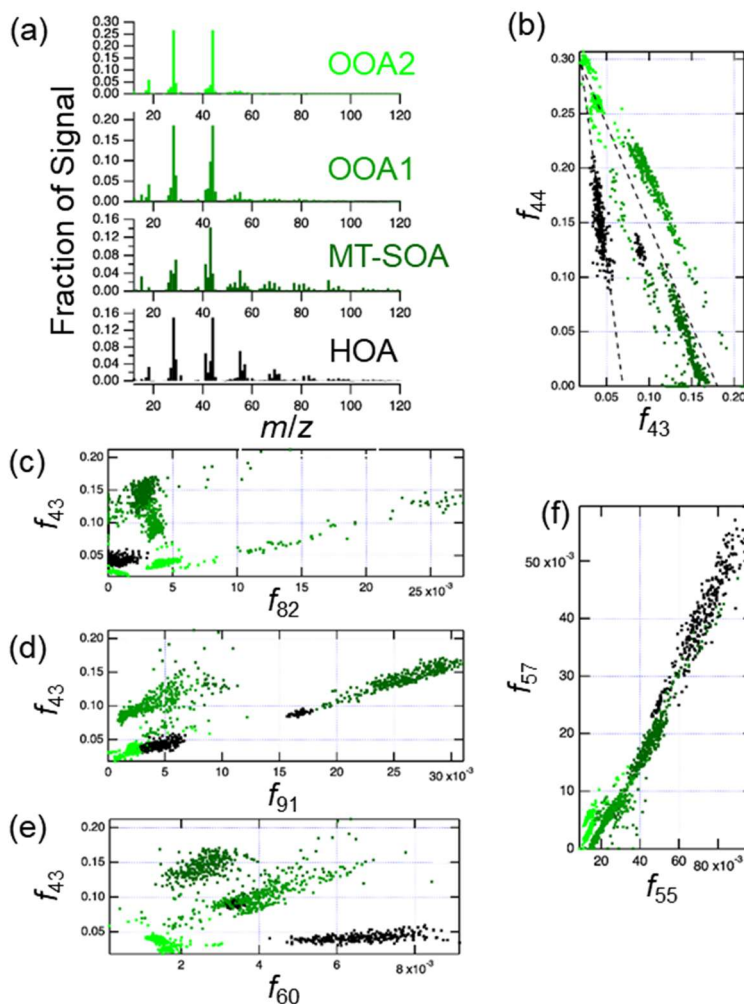


Figure S2. Four-factor unconstrained PMF solution. Clusters were defined using the k-means algorithm. (a) Factor profiles, represented by centroids of each k-means cluster. **(b)** Scatter plot of f_{44} vs. f_{43} . **(c)** Scatter plot of f_{43} vs. f_{82} . **(d)** Scatter plot of f_{43} vs. f_{91} . **(e)** Scatter plot of f_{43} vs. f_{60} . **(f)** Scatter plot of f_{57} vs. f_{55} . In (b)-(f) k-means clusters are identified by color corresponding to the profiles in (a).

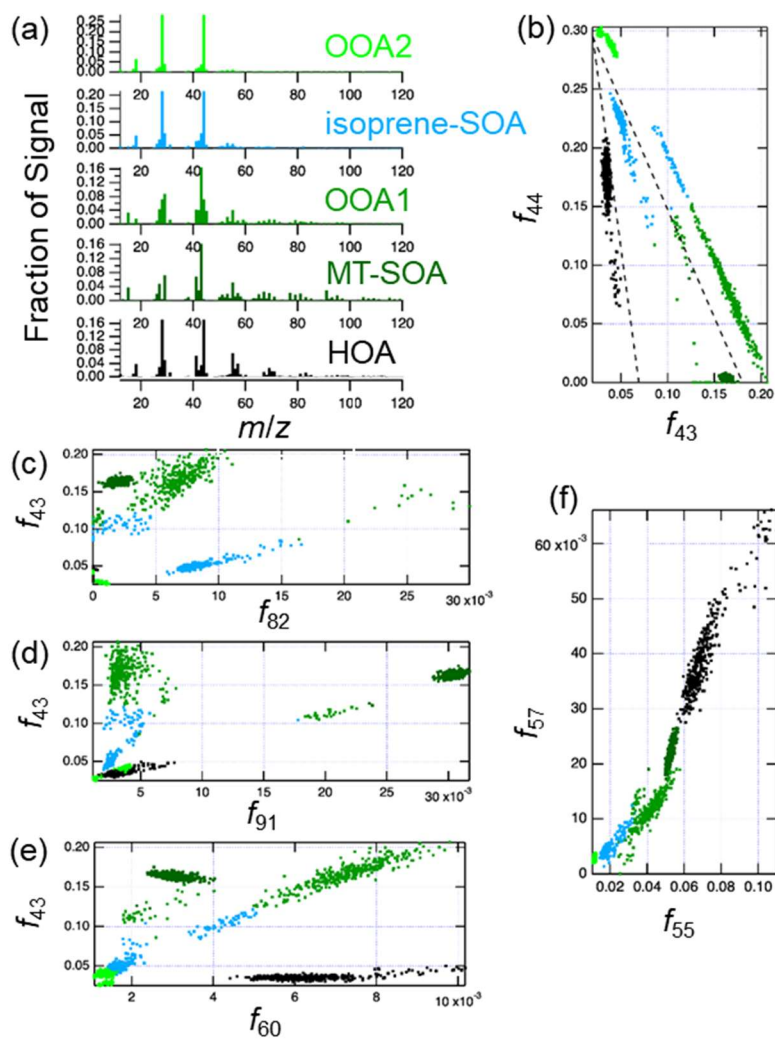


Figure S3. Five-factor unconstrained PMF solution. Clusters were defined using the k-means algorithm. (a) Factor profiles, represented by centroids of each k-means cluster. (b) Scatter plot of f_{44} vs. f_{43} . (c) Scatter plot of f_{43} vs. f_{82} . (d) Scatter plot of f_{43} vs. f_{91} . (e) Scatter plot of f_{43} vs. f_{60} . (f) Scatter plot of f_{57} vs. f_{55} . In (b)-(f) k-means clusters are identified by color corresponding to the profiles in (a).

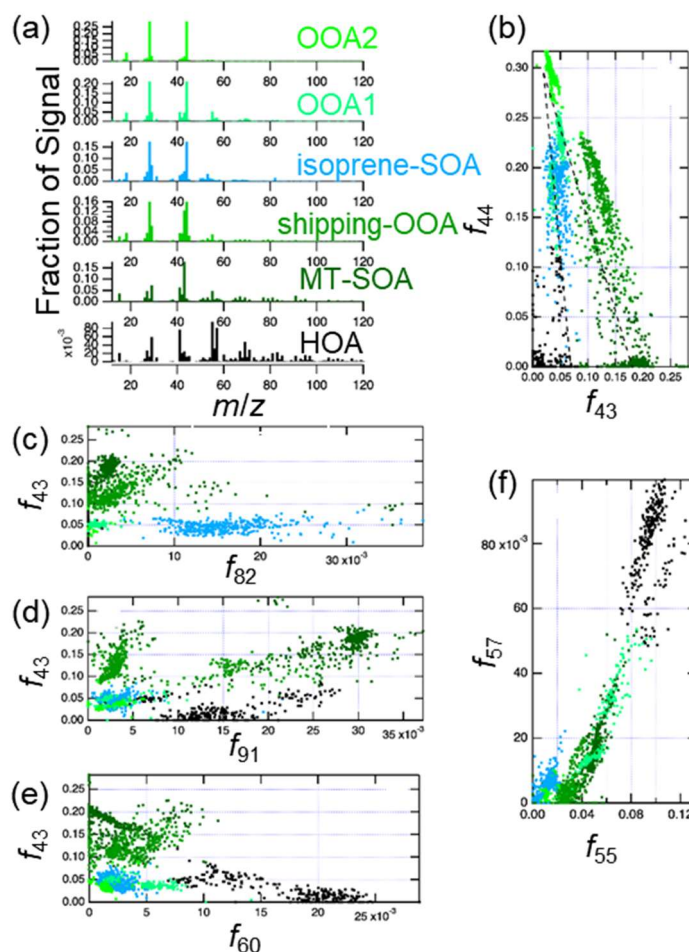


Figure S4. Six-factor unconstrained PMF solution. Clusters were defined using the k-means algorithm. (a) Factor profiles, represented by centroids of each k-means cluster. **(b)** Scatter plot of f_{44} vs. f_{43} . **(c)** Scatter plot of f_{43} vs. f_{82} . **(d)** Scatter plot of f_{43} vs. f_{91} . **(e)** Scatter plot of f_{43} vs. f_{60} . **(f)** Scatter plot of f_{57} vs. f_{55} . In (b)-(f) k-means clusters are identified by color corresponding to the profiles in (a).

55 S1.2 Constrained Rolling PMF

Rolling PMF is carried out on the ACSM organic matrix and its corresponding error matrix at the native sampling resolution of 10 minutes. The m/z 16, 17, 18 and 28, which are replicates of the variability in m/z 44 were removed from the PMF calculation and recalculated a posteriori as a function of the m/z 44 contribution attributed to each factor profile (Elser et al., 2016). A down-weighting function of the form of $1/S2N$ was applied for signal-to-noise (S2N) ratios lower than 1 and untouched otherwise (Visser et al., 2015) on each cell of the organic matrix separately. A rolling window of 7 days with a 1-day shift was used for the rolling strategy. Each rolling window run was initialized 35 times. The total number of individual PMF runs performed was 4375.

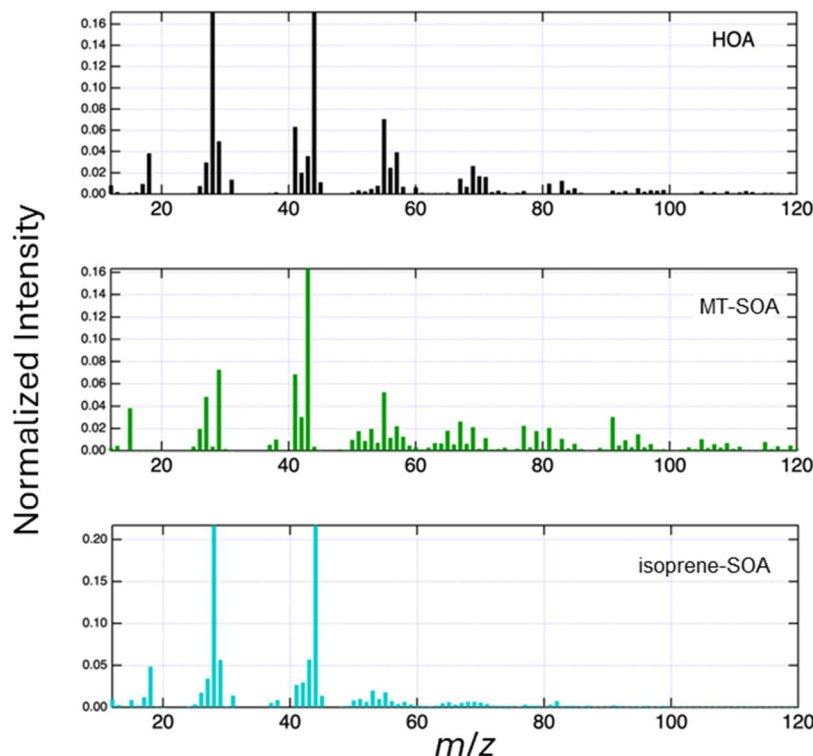


Figure S5. Reference profiles used to constrain the rolling PMF run.

From the unconstrained pre-tests shown in Figs. S3-S4, it is evident that the solution should contain HOA, MT-SOA and isoprene-SOA factors. We tested three different rolling solutions: (1) a 4-factor solution with constrained HOA, MT-SOA and isoprene-SOA factors and one unconstrained OOA factor, (2) a 5-factor solution with the same constrained factors and two unconstrained OOA factors, and (3) a 6-factor solution with the same constrained factors and three unconstrained OOA factors. In the 6-factor solution, the OOA factor with the lowest O:C ratio was identified and named as “shipping-OOA” based on the results of the source analysis, as discussed in the main manuscript section 3.2.2. For the constrained factors, we use the reference profiles shown in Fig. S5 and the α -value approach (Crippa et al., 2014) to allow variability in the constraints. Following Canonaco et al., (2021), we use a random α -value between 0.1 and 0.6 (with a step 0.01) for each rolling window.

Before averaging the individual rolling windows to arrive at a final solution, we eliminate outlier solution windows using one-class support vector machines (SVM), a type of unsupervised outlier detection algorithm (Hejazi and Singh, 2013). In this case, the one-class SVM is performed with a radial basis function kernel on each of the unconstrained OOA factors separately. If a given OOA factor is determined to be an outlier, the whole rolling PMF window is eliminated from the average solution. Approximately 5% of all PMF windows in each solution are determined to be outliers and excluded based on this analysis.

In rolling PMF, each time point is associated with ~ 50 PMF solutions, which are averaged to arrive at the final answer. In the sections below, we present the Quality metrics for the 4-, 5-, and 6-factor rolling PMF solutions.

S1.3 Quality metrics for the 4-factor rolling PMF solution

For the 4-factor solution, we computed the standard deviations at each time point to estimate the relative PMF error, as defined in Canonaco et al. (2021) as shown in Fig. S6. The centers of the log-normal fits are $\pm 25\%$, $\pm 11\%$, $\pm 23\%$, and $\pm 21\%$ for HOA, MT-SOA, isoprene-SOA and OOA respectively.

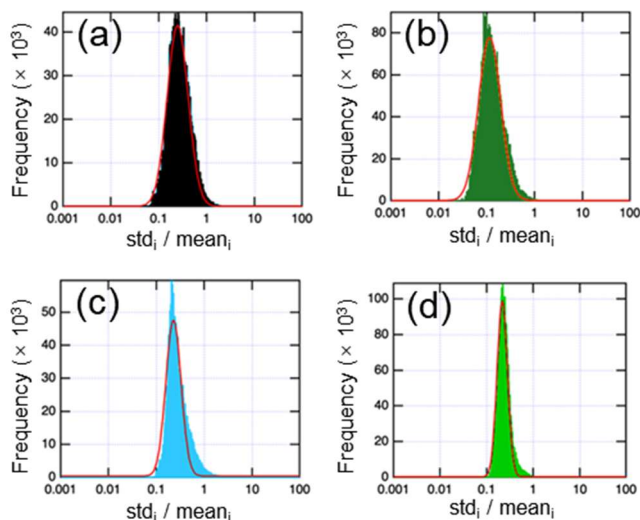
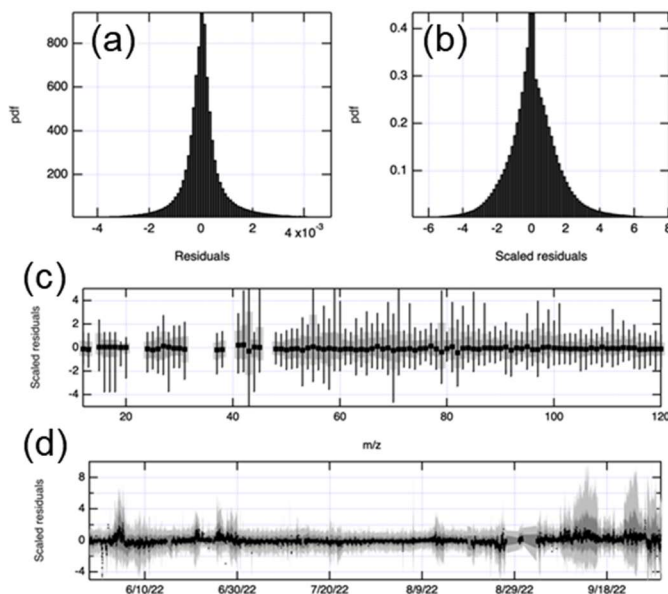


Figure S6. Uncertainty analysis for 4-factor rolling PMF solution. (a) HOA, (b) MT-SOA, (c) isoprene-SOA, (d) OOA.

85 The averaged time series and factor profiles can be used together with the organic matrix and its uncertainty matrix to calculate the residual matrix E (equation (1)) and scaled residuals (E_{ij}/σ_{ij}). Those are plotted as total histograms and statistics over the m/z and time in Fig. S7. The scaled residuals are distributed around 0, not revealing any systematic over- or under-estimation. The distribution of the scaled residuals almost always falls between ± 3 , with ± 3 being a reasonable range for scaled residuals as defined in Paatero and Hopke, (2009).

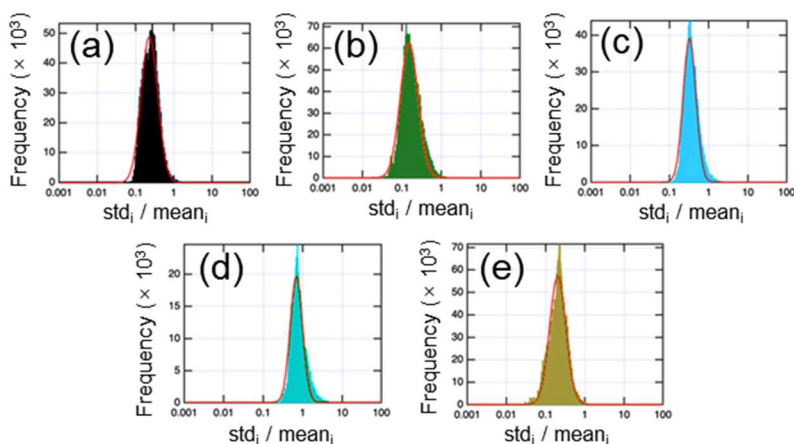


90

Figure S7. Residuals for the 4-factor rolling PMF solution. (a) Histogram of the residual matrix (E). (b) Histogram of the scaled residual matrix (E_{ij}/σ_{ij}). (c) Residual statistics over m/z . (d) Residual statistics over time. In (c) and (d), the whiskers are 90% confidence level, boxes are 75% confidence level.

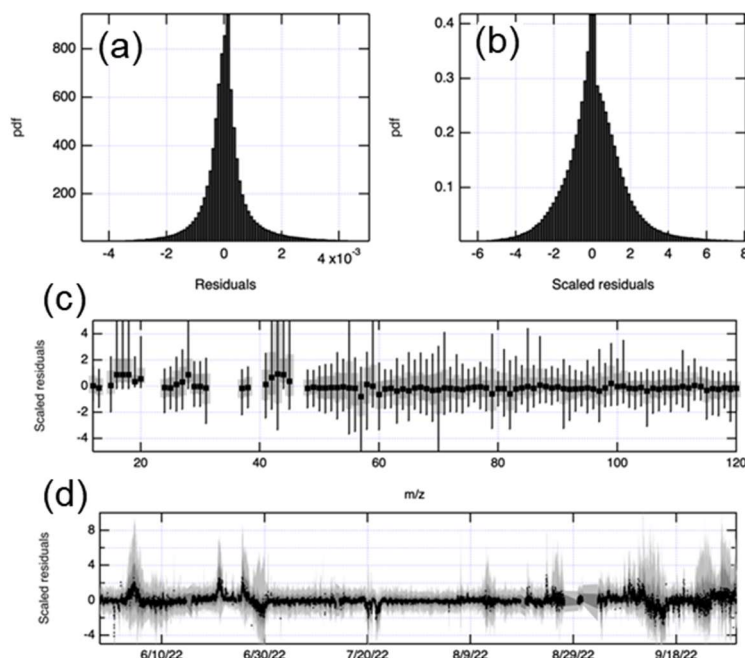
S1.4 Quality metrics for the 5-factor rolling PMF solution

95 In the 5-factor solution, the two OOA factors are ordered prior to averaging so that OOA1 always has lower f_{44} (less aged) than OOA2. Similarly as for the 4-factor solution, we compute the standard deviations at each time point to estimate the relative PMF error, as defined in Canonaco et al. (2021) as shown in Fig. S8. The centers of the log-normal fits are $\pm 23\%$, $\pm 15\%$, $\pm 32\%$, $\pm 68\%$ and $\pm 20\%$ for HOA, MT-SOA, isoprene-SOA, OOA1 and OOA2 respectively.



100 **Figure S8. Uncertainty analysis for 5-factor rolling PMF solution.** (a) HOA, (b) MT-SOA, (c) isoprene-SOA, (d) OOA1, (e) OOA2.

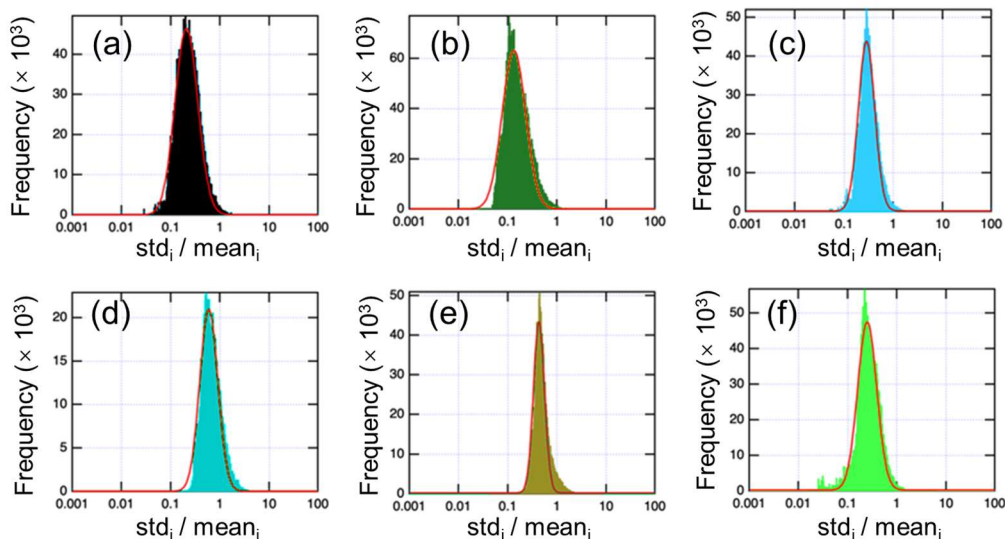
Fig. S9 shows the residual analysis for the 5-factor solution. As with the 4-factor solution, the distribution of the scaled residuals tends to fall between ± 3 .



105 **Figure S9. Residuals for the 5-factor rolling PMF solution.** (a) Histogram of the residual matrix (E). (b) Histogram of the scaled residual matrix (E_{ij}/σ_{ij}). (c) Residual statistics over m/z . (d) Residual statistics over time. In (c) and (d), the whiskers are 90% confidence level, boxes are 75% confidence level.

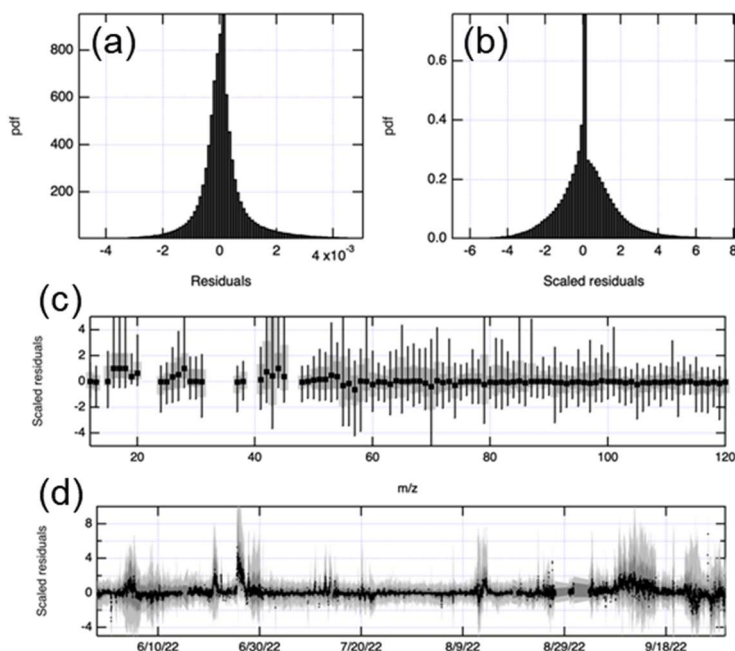
S1.5 Quality metrics for the 6-factor rolling PMF solution

110 In the 6-factor solution, the three OOA factors are ordered prior to averaging so that shipping-OOA always has the lowest f_{44} (least aged), OOA2 always has the highest f_{44} (most aged), and OOA1 has intermediate f_{44} . Similarly as for the 4-factor and 5-factor solutions, we compute the standard deviations at each time point to estimate the relative PMF error, as defined in Canonaco et al. (2021) as shown in Fig. S10. The centers of the log-normal fits are $\pm 21\%$, $\pm 13\%$, $\pm 28\%$, $\pm 59\%$, $\pm 43\%$, and $\pm 25\%$ for HOA, MT-SOA, isoprene-SOA, shipping-OOA, OOA1 and OOA2 respectively.



115 **Figure S10. Uncertainty analysis for 6-factor rolling PMF solution.** (a) HOA, (b) MT-SOA, (c) isoprene-SOA, (d) shipping-OOA, (e) OOA1, (f) OOA2.

Fig. S11 shows the residual analysis for the 6-factor solution. As with the 4-factor and 5-factor solutions, the distribution of the scaled residuals tends to fall between ± 3 .



120 **Figure S11. Residuals for the 6-factor rolling PMF solution.** (a) Histogram of the residual matrix (E). (b) Histogram of the scaled residual matrix (E_{ij}/σ_{ij}). (c) Residual statistics over m/z . (d) Residual statistics over time. In (c) and (d), the whiskers are 90% confidence level, boxes are 75% confidence level.

S1.6 Comparison of 3 rolling PMF solutions

125 We consider the 6-factor solution optimal based on the mass spectral profiles and the correlations of the components with time series for tracer species. First, the mass spectral profiles of HOA are different in 3 solutions. The HOA factor in 6-factor solution exhibits a low O:C ratio (0.14, Fig. 5), consistent with previous studies (Docherty et al., 2011; Mohr et al., 2012). In contrast, the HOA factors in the 4- and 5-factor solutions have much higher O:C ratio of 0.61 and 0.64 (Fig. S12a and Fig. S13a), respectively. Second, compared to 4-factor solutions, the 6-factor solution provides a more refined separation of OOA factors, yielding three

distinct factors with different O:C ratios. Specifically, it resolves one factor associated with shipping emissions (shipping-OOA) and two more oxidized secondary organic aerosol factors (OOA1 and OOA2). In contrast, the OOA factor in the 4-factor solution does not exhibit the characteristic signal as shipping-OOA in the 6-factor solution. Third, in the 6-factor solution, the correlation coefficient R^2 between isoprene-SOA and sulfate is 0.36, higher than those in the 4- and 5-factor solutions (0.25 and 0.27, respectively). Therefore, we consider the 6-factor solution to be the most physically meaningful and interpretable result for characterizing the organic aerosol sources in this study.

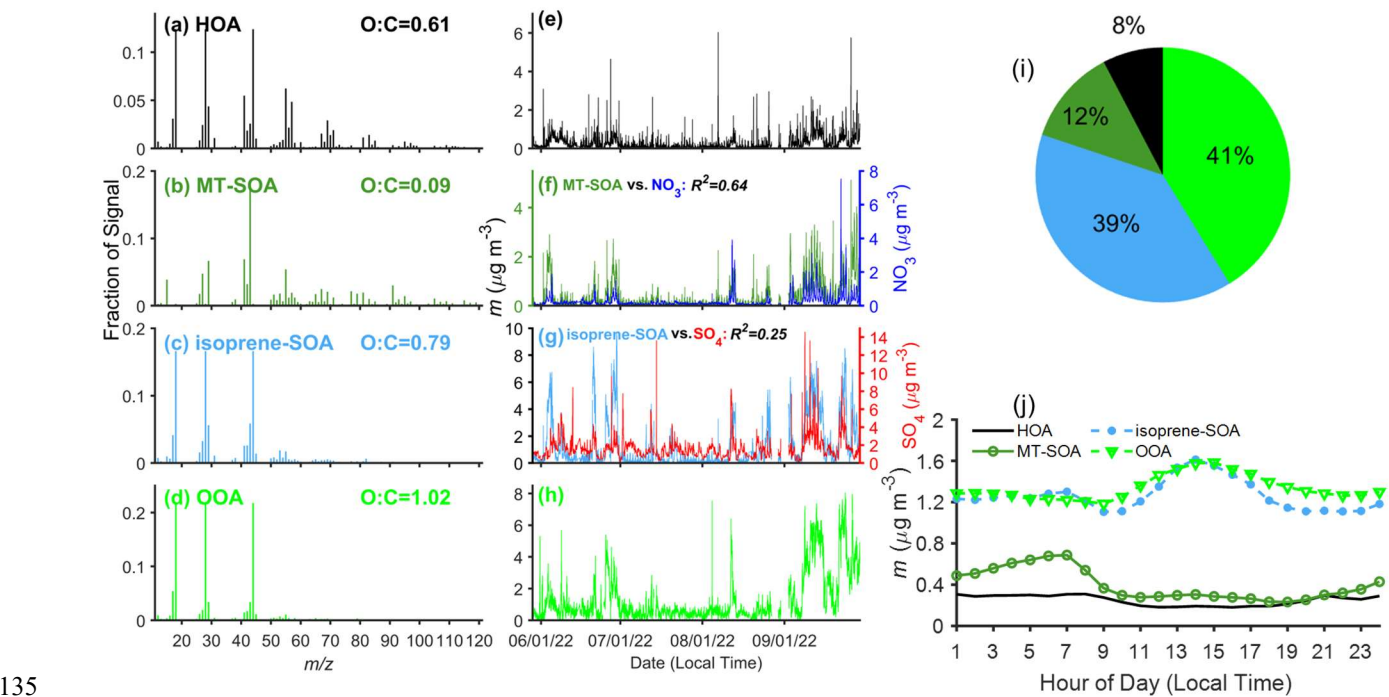
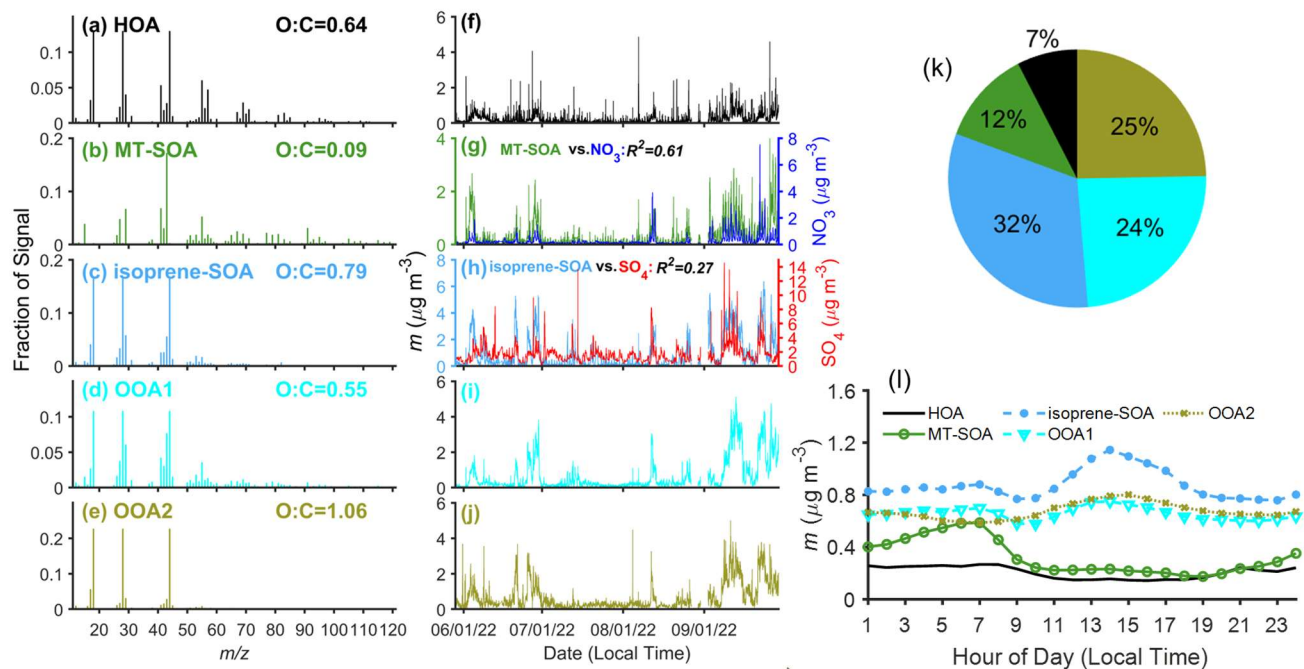


Figure S122. For 4-factor solution. (a-d) Mass spectra of PMF OA factors and (e-h) Time series of OA factors. The correlations between OA factors and relevant species are also shown (f, g). (i) Mass fraction of PMF OA factors over IOP. (j) Diurnal variations of OA factors mass concentrations during the IOP.



140

Figure S133. For 5-factor solution. (a-e) Mass spectra of PMF OA factors and (f-j) Time series of OA factors. The correlations between OA factors and relevant species are also shown (g, h). (k) Mass fraction of PMF OA factors over IOP. (l) Diurnal variations of OA factors mass concentrations during the IOP.

Section 2. Figures

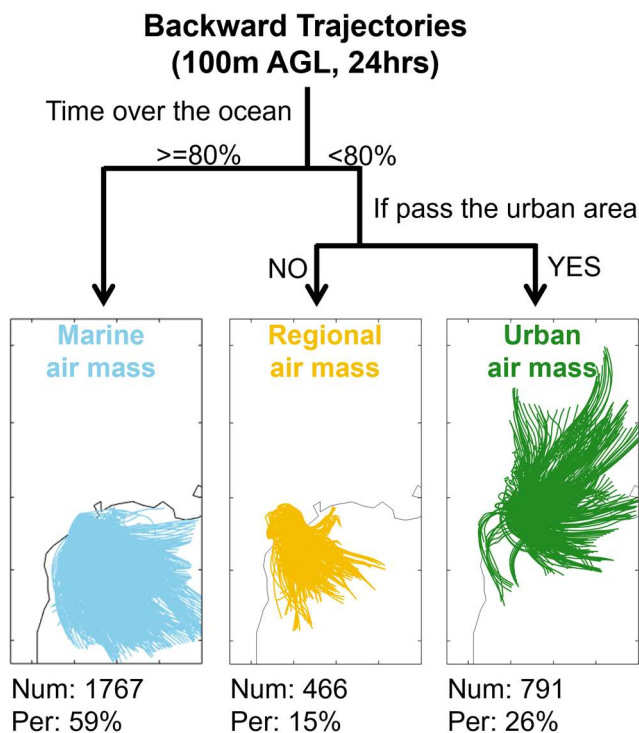


Figure S14. Criteria of the air masses classification.

145

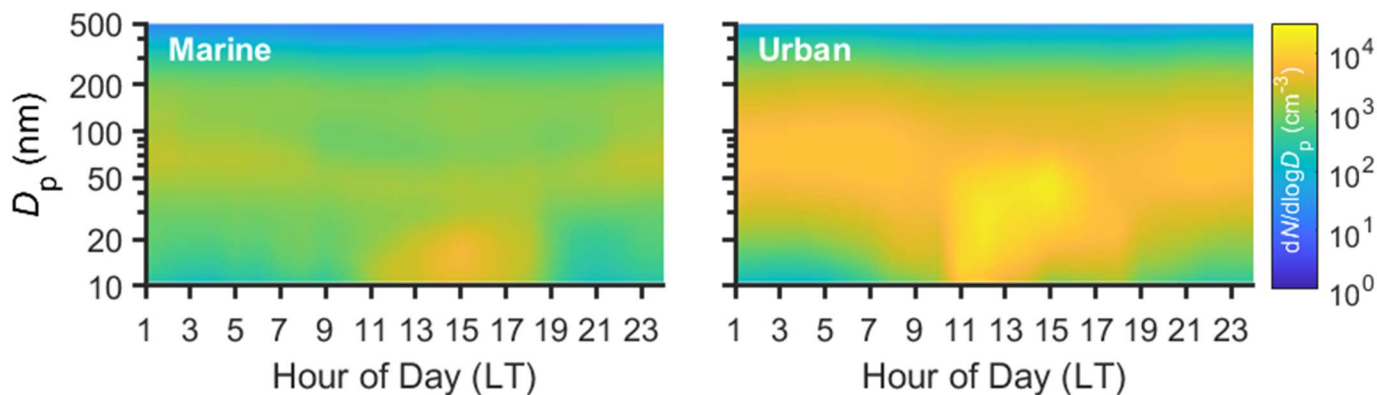


Figure S15. Diurnal variations of particle size distributions in marine and urban air masses observed at the ANC site during the IOP.

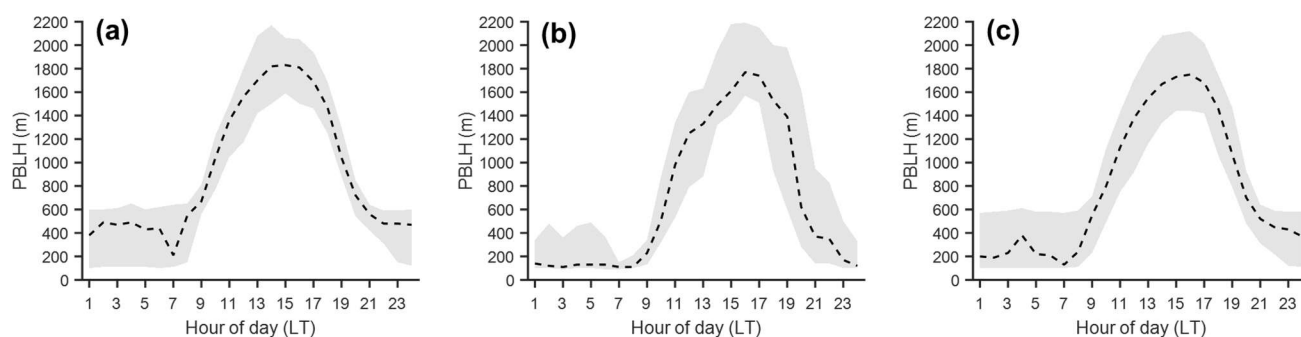


Figure S16. Diurnal variations of planetary boundary layer height (PBLH) of (a) marine, (b) urban, and (c) unseparated (all) air masses. The dashed line represents the median value, while the gray shaded area shows the interquartile range (25th to 75th percentile).

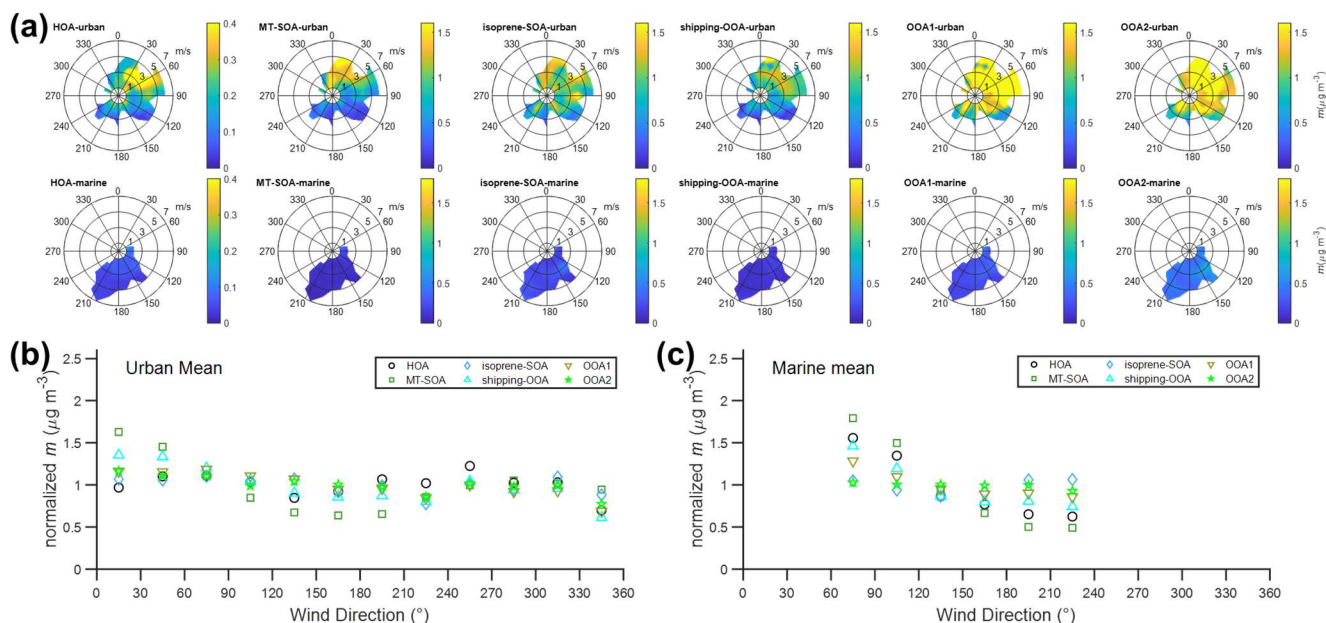


Figure S17. Wind pattern dependence of OA factors. (a) Wind-rose plots of OA factors in urban and marine air masses. Top row: urban. Bottom row: marine. (b-c) Normalized average mass concentration of OA factors as a function of wind direction in (b) urban and (c) marine air masses. OOA1 and OOA2 mass concentrations are largely independent of wind direction in urban and marine air masses.

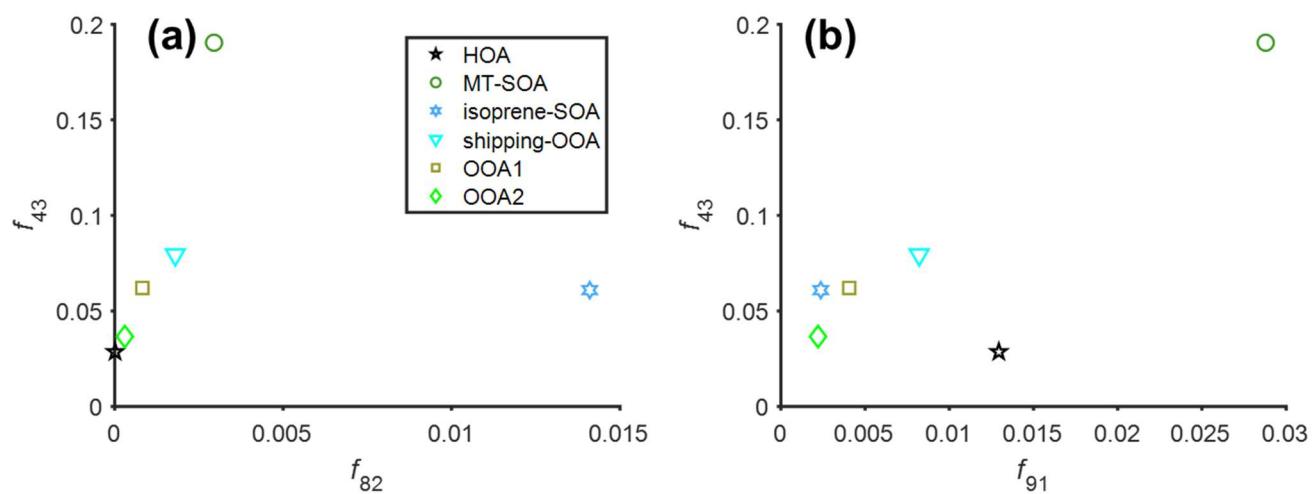
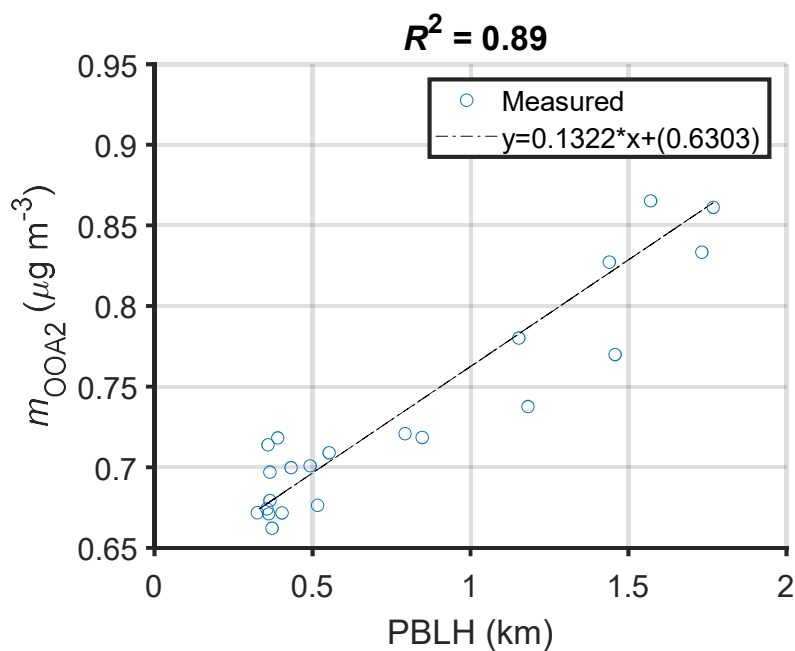


Figure S18. (a-b) f_{43} vs. f_{82} and f_{43} vs. f_{91} for all the PMF factors.



165 Figure S19. Diurnal variations of hourly averaged OOA2 mass concentrations vs. PBLH for unseparated air masses. The dashed line represents the least square linear regression.

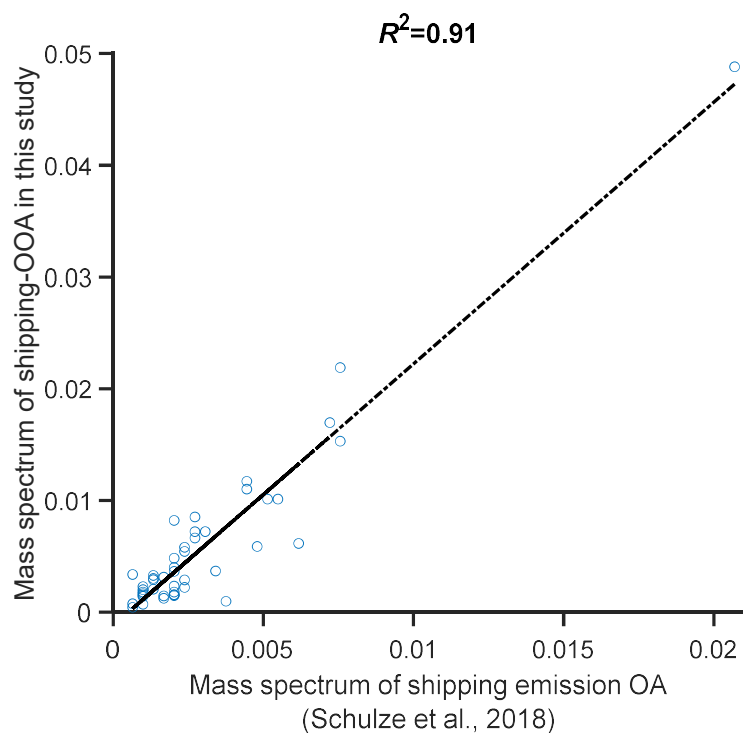


Figure S20. Correlation of the mass spectrum of shipping-OOA in this study with that of heavy shipping emission organics from Schulze et al. (2018).

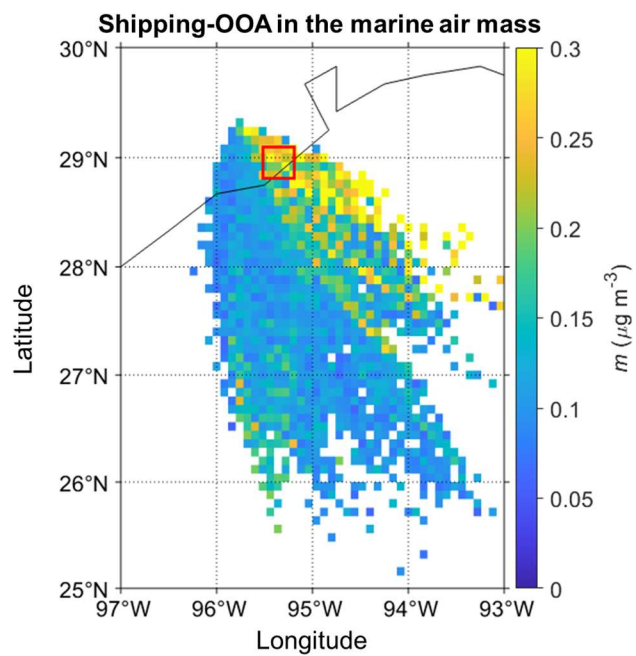
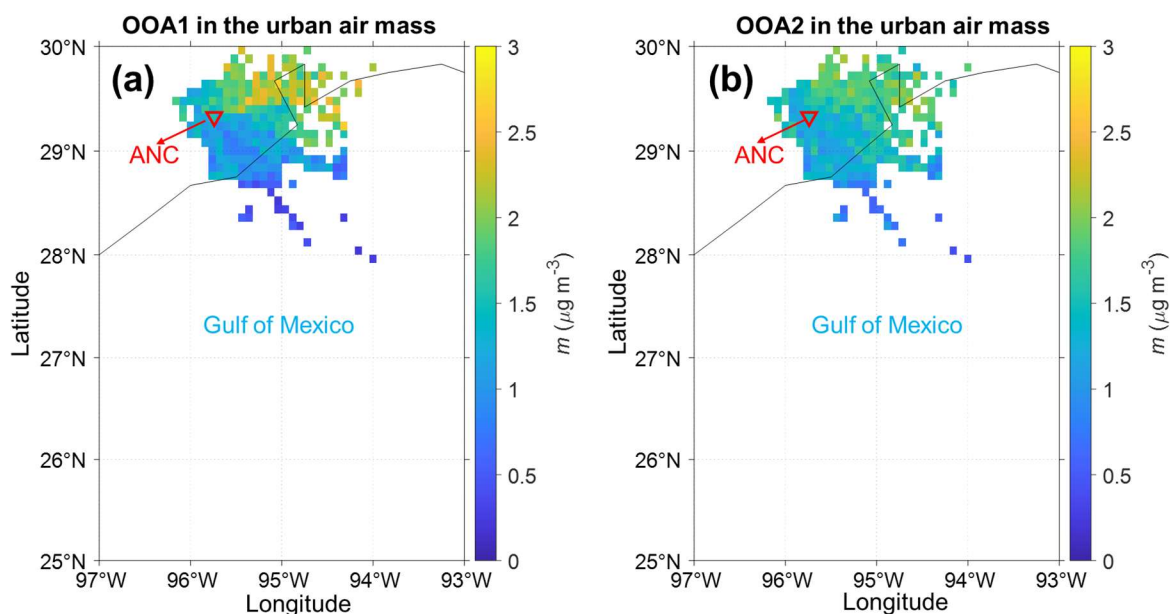
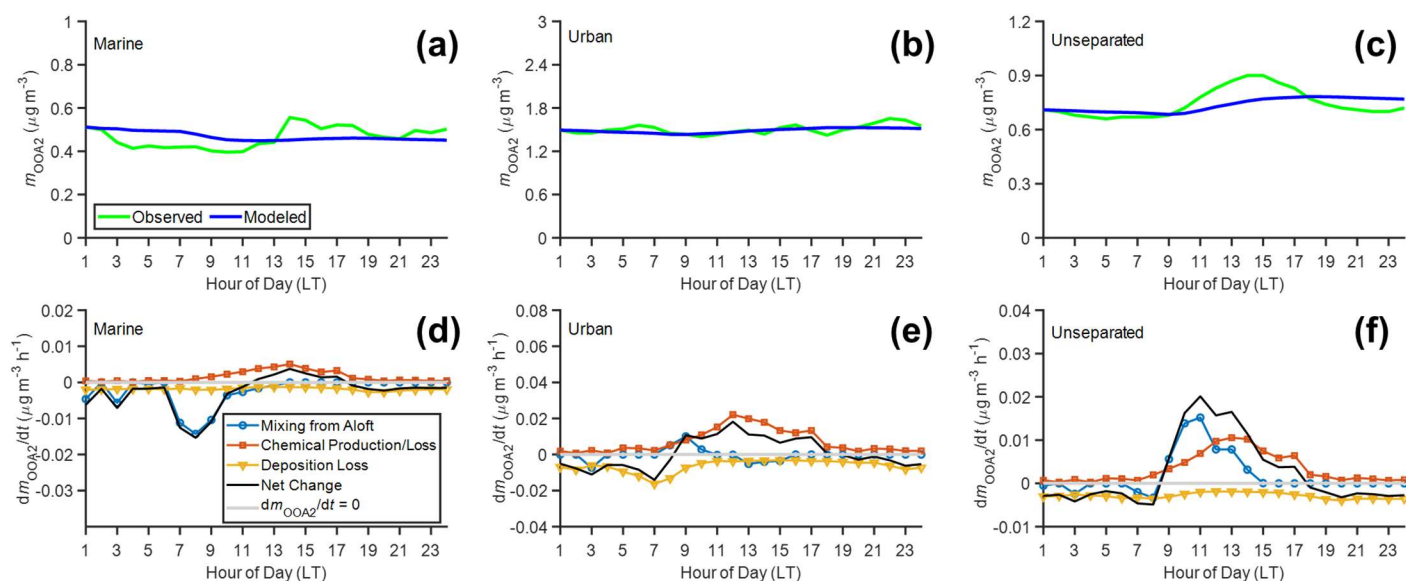


Figure S21. CWT analysis of shipping-OOA in marine air masses. The red square marks Freeport.



175 **Figure S22.** CWT analysis in the urban air mass of (a) OOA1 and (b) OOA2. The red triangle marks the sampling site.



180 **Figure S23.** The sensitivity test of the 1-D box model (Same as Fig. 9 in the main text except that k_1 , k_2 , and k_3 are set to 2.5×10^{-12} , 5×10^{-13} , and 5×10^{-14} $\text{cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$, which are 50% lower than base values). (a, b, c) Diurnal variations of observed and modeled OOA2 mass concentrations in marine (a), urban (b), and unseparated (c) air masses from left to right. (d, e, f) Simulated contributions from different processes (mixing from aloft, chemical production/loss, deposition loss) and the net change rate of OOA2 within the PBLH in the marine (d), urban (e), and unseparated (f) air mass from left to right.

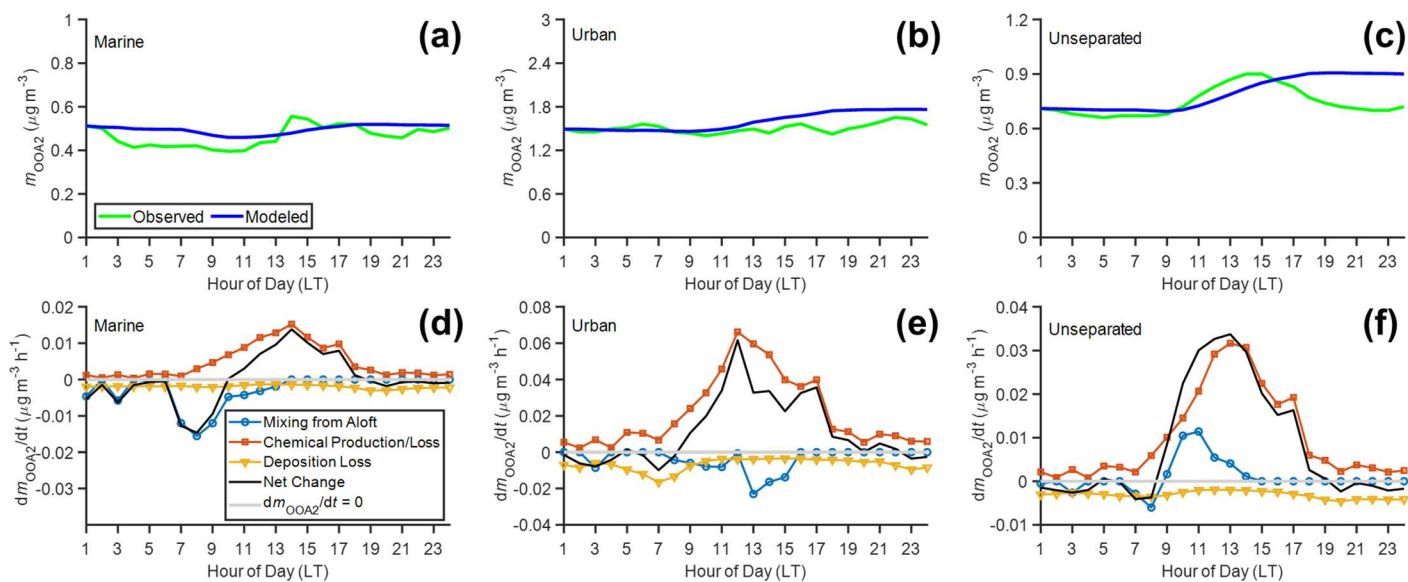


Figure S24. The sensitivity test of the 1-D box model (Same as Fig. 9 in the main text except that k_1 , k_2 , and k_3 are set to 7.5×10^{-12} , 1.5×10^{-12} , and $1.5 \times 10^{-13} \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$, which are 50% higher than base values). (a, b, c) Diurnal variations of observed and modeled OOA2 mass concentrations in marine (a), urban (b), and unseparated (c) air masses from left to right. (d, e, f) Simulated contributions from different processes (mixing from aloft, chemical production/loss, deposition loss) and the net change rate of OOA2 within the PBLH in the marine (d), urban (e), and unseparated (f) air mass from left to right.

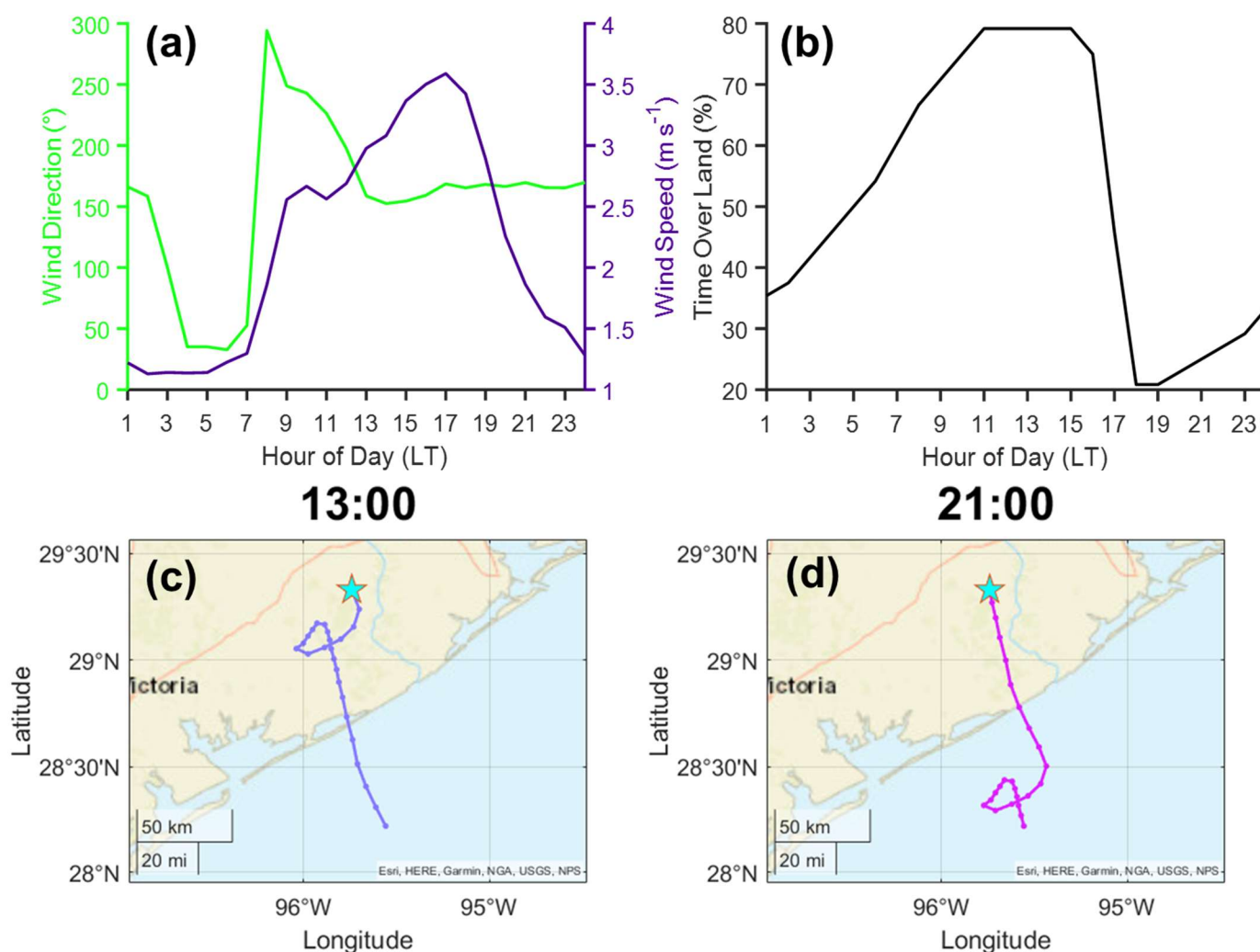


Figure S25. The circulation of land-sea/bay breezes increases the time of air mass over land and the fraction of urban air mass type observed at the ANC site midday. (a) Averaged diurnal variations of wind direction and wind speed over days with air mass changes during the IOP. **(b)** Percentage of time spent by the air mass over the land during the 24-hour period before arriving at the ANC site. The backward trajectories are derived from the diurnal variations of wind direction and speed averaged over days with air mass changes during IOP (i.e. Fig. S25a). **(c-d)** Derived backward trajectories at local times **(c)** 13:00 and **(d)** 21:00. ANC site is marked by the cyan star. .

Section 3. Tables

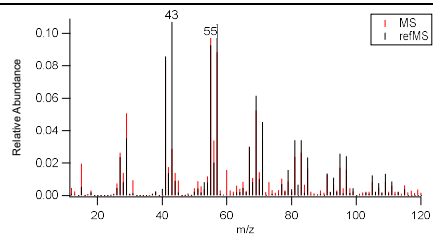
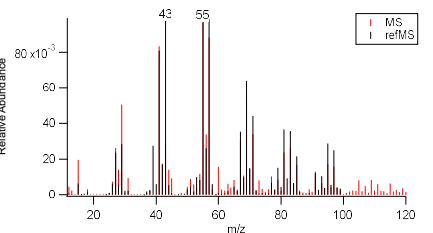
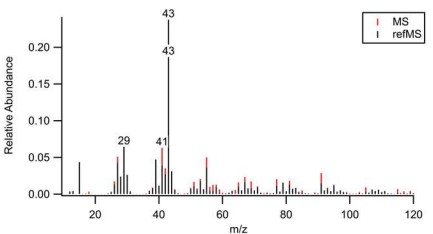
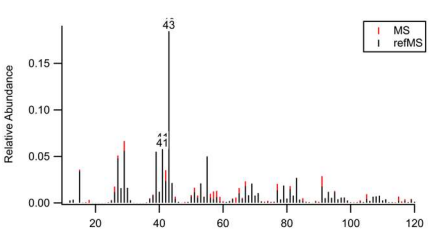
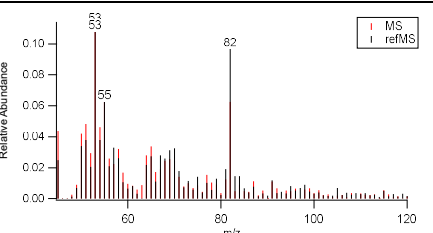
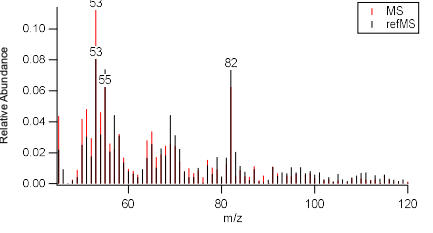
Table S1. Instruments deployed at the ANC site during the TRACER IOP and measurements that are related to this study.

Measurement type	Instruments	Variable	Data available time (Local Time)	Time resolution
Meteorology	Vaisala automatic weather station	Surface wind speed, wind direction, air temperature, relative humidity, air pressure	05/29/2022-09/29/2022	1 min
Clouds	Ceilometer	Planetary boundary layer heights	06/30/2022-09/29/2022	16 s
Aerosol	Aerosol Chemical Speciation Monitor (ACSM)	Chemical compositions of aerosol particles	05/29/2022-09/29/2022	10 min
	Scanning Mobility Particle Sizer (SMPS)	Aerosol size distribution	05/29/2022-09/29/2022	5 min
	Condensation Particle Counter (CPC)	Total particle number concentration	05/29/2022-09/29/2022	1 min

Table S2. Correlation coefficients R^2 of NR-PM₁ chemical compositions.

	NO ₃	Org	SO ₄	NH ₄
NO ₃	1.00	0.54	0.09	0.24
Org		1.00	0.23	0.33
SO ₄			1.00	0.83
NH ₄				1.00

Table S3. Comparison of OA mass spectra with those from earlier laboratory and field studies.

PMF Factors	Mass Spectra Comparison	Potential Sources	Potential Formation Pathways	ACSM Database numbers and References
HOA		Road traffic (Barcelona, Spain)	Primary emission: Fossil fuel combustion dominated by diesel exhaust	256_DAURE campaign_2009_HOA; (Mohr et al., 2012)
		Road traffic (Riverside, USA)	Primary emission: Fuel and lubricating oil combustion	242_SOAR-1_Campaign_2005_HOA (Docherty et al., 2011)
MT-SOA		Lab experiment	Oxidation product of limonene by nitrate	380_Chamber limonene_Nitrate_AS_LIM-2 (Boyd et al., 2015)
		Lab experiment	Sequentially oxidation product of a-pinene and limonene by nitrate	382_Chamber α-pinene+limonene_Nitrate_AS_dry_SEQ-1 (Takeuchi et al., 2022)
isoprene-SOA		Tropical rainforest (Danum Valley, Borneo, Malaysia)	Isoprene oxidation	93_Borneo Rainforest_Fac82 (Robinson et al., 2011)
		Forest (SE US forest)	isoprene oxidation	124_SOAS campaign_2013_IEP OX_SOA (Hu et al., 2015)

PMF Factors	Mass Spectra Comparison	Potential Sources	Potential Formation Pathways	ACSM Database numbers and References
shipping- OOA		Urban area (Pasadena, CA, USA)	Photochemical formation of traffic emissions	292_CalNex campaign_2010_SV OOA (Hayes et al., 2013)
		Urban area (Beijing, China)	VOC photochemical product	302_Beijing urban area_2011_LOOOA (Hu et al., 2016)
OOA1		Urban area [Average of spectra in Beijing, Tokyo (summer), Pittsburgh, Riverside, New York City (summer), and Zurich]	VOC photochemical product	79_LVVOA_avg (Ng et al., 2011)
		Downwind of pollution sources (Changdao island, China)	VOC photochemical product	299_Changdao island_2011_LVOO A (Hu et al., 2013)
OOA2		Urban area (Zurich, Switzerland)	Highly aged OOA; mass spectrum similar to fulvic acid	65_Zurich_summer_ 2005_OOA_I (Lanz et al., 2007)
		Urban area (Pittsburgh)	Highly aged OOA; well correlated to aged rural organic aerosols	64_Pittsburgh_OOA (Zhang et al., 2005)

Table S4. Correlation coefficients R^2 of OA factors with NO_3 and SO_4 .

	NO_3	SO_4
HOA	0.28	0.07
MT-SOA	0.58	0.05
isoprene-SOA	0.40	0.36
shipping-OOA	0.42	0.17
OOA1	0.42	0.24
OOA2	0.36	0.25

205 **Table S5. Comparison of the mass concentrations of key aerosol components with those reported by prior studies in the Houston region.**

	Location	Air mass	Total ($\mu\text{g}\cdot\text{m}^{-3}$)	Org ($\mu\text{g}\cdot\text{m}^{-3}$)	SO₄ ($\mu\text{g}\cdot\text{m}^{-3}$)	NH₄ ($\mu\text{g}\cdot\text{m}^{-3}$)	NO₃ ($\mu\text{g}\cdot\text{m}^{-3}$)	HOA ($\mu\text{g}\cdot\text{m}^{-3}$)	SOA ($\mu\text{g}\cdot\text{m}^{-3}$)
This study	Guy	Marine	3.55	1.42	1.47	0.50	0.16	0.07	1.29
		Urban	9.96	6.58	2.15	0.74	0.49	0.24	5.87
Yoon 2021	Manvel Croix							0.41	3.74
Dai 2019	Sugar Land		3.58	1.7	1.3	0.5	0.08	0.2	1.0
Wallace 2018	Manchester St.		10.8	5.5	2.5	1.3	1.5	0.67	1.93
Schulze 2018	Southwest of Galveston	Marine	3.82	0.7	2.4	0.7	0.02	0.05	0.64
		Urban	9.8	7.2	1.9	0.6	0.1	0.16	7.06
Al-Naiema 2018	Houston Ship Channel		2.86	1.14	1.29	0.4	0.03	0.42	0.72
Cleveland 2012	University of Houston		10.9	5.5	4.1	0.9	0.4	1.7	3.7

References

- Boyd, C. M., Sanchez, J., Xu, L., Eugene, A. J., Nah, T., Tuet, W. Y., Guzman, M. I., and Ng, N. L.: Secondary organic aerosol formation from the β -pinene+NO₃ system: effect of humidity and peroxy radical fate, *Atmospheric Chemistry and Physics*, 15, 7497–7522, <https://doi.org/10.5194/acp-15-7497-2015>, 2015.
- Canonaco, F., Crippa, M., Slowik, J. G., Baltensperger, U., and 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, <https://doi.org/10.5194/amt-6-3649-2013>, 2013.
- Canonaco, F., Tobler, A., Chen, G., Sosedova, Y., Slowik, J. G., Bozzetti, C., Daellenbach, K. R., El Haddad, I., Crippa, M., Huang, R.-J., Furger, M., Baltensperger, U., and Prévôt, A. S. H.: 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, <https://doi.org/10.5194/amt-14-923-2021>, 2021.
- Crippa, M., Canonaco, F., Lanz, V. A., Äijälä, M., Allan, J. D., Carbone, S., Capes, G., Ceburnis, D., Dall’Osto, M., Day, D. A., DeCarlo, P. F., Ehn, M., Eriksson, A., Freney, E., Hildebrandt Ruiz, L., Hillamo, R., Jimenez, J. L., Junninen, H., Kiendler-Scharr, A., Kortelainen, A.-M., Kulmala, M., Laaksonen, A., Mensah, A. A., Mohr, C., Nemitz, E., O’Dowd, C., Ovadnevaite, J., Pandis, S. N., Petäjä, T., Poulain, L., Saarikoski, S., Sellegri, K., Swietlicki, E., Tiitta, P., Worsnop, D. R., Baltensperger, U., and Prévôt, A. S. H.: 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, <https://doi.org/10.5194/acp-14-6159-2014>, 2014.
- Docherty, K. S., Aiken, A. C., Huffman, J. A., Ulbrich, I. M., DeCarlo, P. F., Sueper, D., Worsnop, D. R., Snyder, D. C., Peltier, R. E., Weber, R. J., Grover, B. D., Eatough, D. J., Williams, B. J., Goldstein, A. H., Ziemann, P. J., and Jimenez, J. L.: The 2005 Study of Organic Aerosols at Riverside (SOAR-1): instrumental intercomparisons and fine particle composition, *Atmospheric Chemistry and Physics*, 11, 12387–12420, <https://doi.org/10.5194/acp-11-12387-2011>, 2011.
- Elser, M., Bozzetti, C., El-Haddad, I., Maasikmets, M., Teinemaa, E., Richter, R., Wolf, R., Slowik, J. G., Baltensperger, U., and Prévôt, A. S. H.: Urban increments of gaseous and aerosol pollutants and their sources using mobile aerosol mass spectrometry measurements, *Atmos. Chem. Phys.*, 16, 7117–7134, <https://doi.org/10.5194/acp-16-7117-2016>, 2016.
- Hayes, P. L., Ortega, A. M., Cubison, M. J., Froyd, K. D., Zhao, Y., Cliff, S. S., Hu, W. W., Toohey, D. W., Flynn, J. H., Lefer, B. L., Grossberg, N., Alvarez, S., Rappenglück, B., Taylor, J. W., Allan, J. D., Holloway, J. S., Gilman, J. B., Kuster, W. C., de Gouw, J. A., Massoli, P., Zhang, X., Liu, J., Weber, R. J., Corrigan, A. L., Russell, L. M., Isaacman, G., Worton, D. R., Kreisberg, N. M., Goldstein, A. H., Thalman, R., Waxman, E. M., Volkamer, R., Lin, Y. H., Surratt, J. D., Kleindienst, T. E., Offenberg, J. H., Dusanter, S., Griffith, S., Stevens, P. S., Brioude, J., Angevine, W. M., and Jimenez, J. L.: Organic aerosol composition and sources in Pasadena, California, during the 2010 CalNex campaign, *Journal of Geophysical Research: Atmospheres*, 118, 9233–9257, <https://doi.org/10.1002/jgrd.50530>, 2013.
- Hejazi, M. and Singh, Y. P.: ONE-CLASS SUPPORT VECTOR MACHINES APPROACH TO ANOMALY DETECTION, *Applied Artificial Intelligence*, 27, 351–366, <https://doi.org/10.1080/08839514.2013.785791>, 2013.
- Hu, W., Hu, M., Hu, W., Jimenez, J. L., Yuan, B., Chen, W., Wang, M., Wu, Y., Chen, C., Wang, Z., Peng, J., Zeng, L., and Shao, M.: Chemical composition, sources, and aging process of submicron aerosols in Beijing: Contrast between summer and winter, *Journal of Geophysical Research: Atmospheres*, 121, 1955–1977, <https://doi.org/10.1002/2015JD024020>, 2016.
- Hu, W. W., Hu, M., Yuan, B., Jimenez, J. L., Tang, Q., Peng, J. F., Hu, W., Shao, M., Wang, M., Zeng, L. M., Wu, Y. S., Gong, Z. H., Huang, X. F., and He, L. Y.: Insights on organic aerosol aging and the influence of coal combustion at a regional receptor site of central eastern China, *Atmospheric Chemistry and Physics*, 13, 10095–10112, <https://doi.org/10.5194/acp-13-10095-2013>, 2013.
- Hu, W. W., Campuzano-Jost, P., Palm, B. B., Day, D. A., Ortega, A. M., Hayes, P. L., Krechmer, J. E., Chen, Q., Kuwata, M., Liu, Y. J., de Sá, S. S., McKinney, K., Martin, S. T., Hu, M., Budisulistiorini, S. H., Riva, M., Surratt, J. D., St. Clair, J. M., Isaacman-Van Wertz, G., Yee, L. D., Goldstein, A. H., Carbone, S., Brito, J., Artaxo, P., de Gouw, J. A., Koss, A., Wisthaler, A., Mikoviny, T., Karl, T., Kaser, L., Jud, W., Hansel, A., Docherty, K. S., Alexander, M. L., Robinson, N. H., Coe, H., Allan, J. D., Canagaratna, M. R., Paulot, F., and Jimenez, J. L.: Characterization of a real-time tracer for isoprene epoxydiols-derived

- secondary organic aerosol (IEPOX-SOA) from aerosol mass spectrometer measurements, *Atmospheric Chemistry and Physics*, 15, 11807–11833, <https://doi.org/10.5194/acp-15-11807-2015>, 2015.
- 255 Lanz, V. A., Alfarra, M. R., Baltensperger, U., Buchmann, B., Hueglin, C., and Prévôt, A. S. H.: Source apportionment of submicron organic aerosols at an urban site by factor analytical modelling of aerosol mass spectra, *Atmospheric Chemistry and Physics*, 7, 1503–1522, <https://doi.org/10.5194/acp-7-1503-2007>, 2007.
- 260 Mohr, C., DeCarlo, P. F., Heringa, M. F., Chirico, R., Slowik, J. G., Richter, R., Reche, C., Alastuey, A., Querol, X., Seco, R., Peñuelas, J., Jiménez, J. L., Crippa, M., Zimmermann, R., Baltensperger, U., and Prévôt, A. S. H.: Identification and quantification of organic aerosol from cooking and other sources in Barcelona using aerosol mass spectrometer data, *Atmospheric Chemistry and Physics*, 12, 1649–1665, <https://doi.org/10.5194/acp-12-1649-2012>, 2012.
- Ng, N. L., Canagaratna, M. R., Jimenez, J. L., Zhang, Q., Ulbrich, I. M., and Worsnop, D. R.: Real-Time Methods for Estimating Organic Component Mass Concentrations from Aerosol Mass Spectrometer Data, *Environ. Sci. Technol.*, 45, 910–916, <https://doi.org/10.1021/es102951k>, 2011.
- 265 Paatero, P.: Least squares formulation of robust non-negative factor analysis, *Chemometrics and Intelligent Laboratory Systems*, 37, 23–35, [https://doi.org/10.1016/S0169-7439\(96\)00044-5](https://doi.org/10.1016/S0169-7439(96)00044-5), 1997.
- Paatero, P. and Hopke, P. K.: Rotational tools for factor analytic models, *Journal of Chemometrics*, 23, 91–100, <https://doi.org/10.1002/cem.1197>, 2009.
- Paatero, P. and Tapper, U.: Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values, *Environmetrics*, 5, 111–126, <https://doi.org/10.1002/env.3170050203>, 1994.
- 270 Robinson, N. H., Hamilton, J. F., Allan, J. D., Langford, B., Oram, D. E., Chen, Q., Docherty, K., Farmer, D. K., Jimenez, J. L., Ward, M. W., Hewitt, C. N., Barley, M. H., Jenkin, M. E., Rickard, A. R., Martin, S. T., McFiggans, G., and Coe, H.: Evidence for a significant proportion of Secondary Organic Aerosol from isoprene above a maritime tropical forest, *Atmospheric Chemistry and Physics*, 11, 1039–1050, <https://doi.org/10.5194/acp-11-1039-2011>, 2011.
- 275 Schulze, B. C., Wallace, H. W., Bui, A. T., Flynn, J. H., Erickson, M. H., Alvarez, S., Dai, Q., Usenko, S., Sheesley, R. J., and Griffin, R. J.: The impacts of regional shipping emissions on the chemical characteristics of coastal submicron aerosols near Houston, TX, *Atmospheric Chemistry and Physics*, 18, 14217–14241, <https://doi.org/10.5194/acp-18-14217-2018>, 2018.
- Takeuchi, M., Berkemeier, T., Eris, G., and Ng, N. L.: Non-linear effects of secondary organic aerosol formation and properties in multi-precursor systems, *Nat Commun*, 13, 7883, <https://doi.org/10.1038/s41467-022-35546-1>, 2022.
- 280 Visser, S., Slowik, J. G., Furger, M., Zotter, P., Bukowiecki, N., Canonaco, F., Flechsig, U., Appel, K., Green, D. C., Tremper, A. H., Young, D. E., Williams, P. I., Allan, J. D., Coe, H., Williams, L. R., Mohr, C., Xu, L., Ng, N. L., Nemitz, E., Barlow, J. F., Halios, C. H., Fleming, Z. L., Baltensperger, U., and Prévôt, A. S. H.: Advanced source apportionment of size-resolved trace elements at multiple sites in London during winter, *Atmos. Chem. Phys.*, 15, 11291–11309, <https://doi.org/10.5194/acp-15-11291-2015>, 2015.
- 285 Zhang, Q., Alfarra, M. R., Worsnop, D. R., Allan, J. D., Coe, H., Canagaratna, M. R., and Jimenez, J. L.: Deconvolution and Quantification of Hydrocarbon-like and Oxygenated Organic Aerosols Based on Aerosol Mass Spectrometry, *Environ. Sci. Technol.*, 39, 4938–4952, <https://doi.org/10.1021/es048568l>, 2005.