

Linking Precursors and Volatility of Ambient Oxygenated Organic Aerosols Using Thermal Desorption Measurement and Machine Learning

Published as part of ACS ES&T Air *special issue* “Elevating Atmospheric Chemistry Measurements and Modeling with Artificial Intelligence”.

Xinyu Wang, Yongyi Zhao, Ke Hu, Jian Wang, Qiongqiong Wang, Nan Chen, Bo Zhu, Hong-Hai Zhang, and Huan Yu*



Cite This: ACS EST Air 2024, 1, 1239–1251



Read Online

ACCESS |

Metrics & More

Article Recommendations

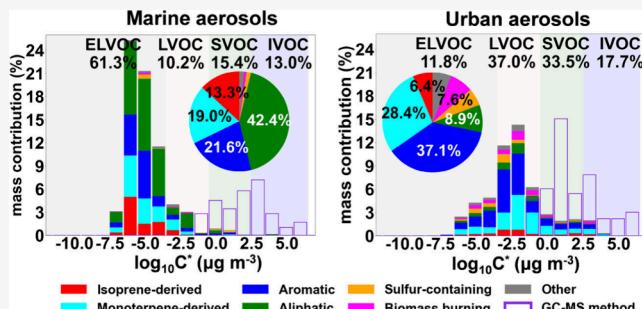
Supporting Information

ABSTRACT: We conducted thermal desorption measurements and machine learning analysis to investigate the volatility and precursors of ambient oxygenated organic aerosols (OOA), with a focus on the link between them, in a variety of urban and marine locations. We found that the OOA species measured by an iodide-based Chemical Ionization Mass Spectrometer equipped with a Filter Inlet for Gases and AEROSOL (FIGAERO-CIMS) accounted for $16 \pm 13\%$ of OA in those urban and marine locations and represented mostly the secondary and moderate-volatility portion of ambient OA. On average, 25.1% in species number and 26.8% in mass of the OOA species detected by the FIGAERO-CIMS in a winter campaign at an urban site in Wuhan, a megacity in central China, might be attributed to thermal decomposition fragments. Our results show that the volatility and precursor of ambient OOA differed systematically according to location, season, and pollution level. The OOA in the ocean atmosphere was characterized by high fractions of extremely low volatility organic compounds (ELVOC) and aliphatic species, while the inland urban OOA was characterized by aromatic species and fell primarily into the low volatility organic compounds (LVOCs) and semivolatile organic compounds (SVOCs) range. The volatilities of OOA in the inland urban atmosphere in summer, pollution days, and daytime were lower than those in winter, clean days, and nighttime. When the PM episode developed from clean to particle growth and then to pollution period, the OOA species shifted from Low-Mw OOA species to Median-Mw OOA species and then to highly nonvolatile species. The study of ambient OOA volatility in this work also provides important data for future closure studies of SOA formation and its precursors.

KEYWORDS: machine learning, volatility basis set, marine aerosols, urban aerosols, oxygenated organic aerosols

1. INTRODUCTION

The implementation of particulate matter (PM) control measures in China since 2016 has led to significantly declining PM concentrations. Oxygenated organic aerosols (OOA) contributes an important fraction of ambient aerosols and is often the bottleneck of further reducing PM pollution. OOA is mostly produced through gas-to-particle partitioning of oxidation products from volatile organic precursors or condensed-phase reactions in the atmosphere. One of the key physicochemical properties of atmospheric OOA is volatility, which reflects the contribution of OOA in particle formation and growth and also determines the evaporation potential and thus the lifetime of OOA. Knowing the volatility of ambient OOA is also useful for the closure studies of



secondary OA (SOA) formation and its precursors, combining the information about condensing vapor species.^{1–3}

Besides ambient temperature, aerosol mixing state, and atmospheric aging,⁴ it is well-known that the volatility of OOA depends strongly on the molecular structures of OOA species.^{5–8} However, it is unknown if there is any link between the OOA volatility and the key source or VOC precursor forming the OOA in the real atmosphere. For example, for an

Received: April 2, 2024

Revised: August 20, 2024

Accepted: August 20, 2024

Published: August 27, 2024



urban area typically influenced by both biogenic and anthropogenic VOC emissions, their relative importance in forming extremely low volatility organic compounds (EL-VOCs) is unclear. Therefore, quantitative knowledge about the precursors of ELVOCs in the real atmosphere is needed to know the key VOC precursor(s) driving particle formation and growth^{9–11} and to provide a scientific basis for formulating effective PM control measures.

Recently developed thermal desorption mass spectrometers, like Chemical Ionization Mass Spectrometer equipped with a Filter Inlet for Gases and AEROSOL (FIGAERO-CIMS), make it possible to determine the volatility and the molecular composition of OA simultaneously.¹² Equipped with a soft ionization source, FIGAERO-CIMS is selective for functionalized polar organic compounds and has been primarily applied in the study of SOA or OOA in laboratory investigations^{13,14} and field measurements at forest sites^{15,16} and urban sites.^{17,18} Most commonly, the thermal desorption temperature (e.g., T_{\max}) of an organic species is directly measured to deduce its saturation concentration. The species volatilities could be further binned to obtain bulk volatility.^{14,16,19,20}

After knowing the volatility of an OOA species, the next crucial step is to assign a precursor or source for the OOA species. Positive matrix factorization (PMF) is a common tool for addressing this issue based on continuous measurement of OOA species in the atmosphere.^{21,22} However, PMF is often insufficient in resolving the source of the OOA or factors because the temporal variations of the OOA species from different precursors in the atmosphere are often similar, considering that they are subject to the same atmospheric oxidation processes or their concentrations are dictated by the origin of the air mass. By mining a large amount of data sets, machine learning has the potential to uncover the relationships between molecular features and precursor type of an OOA species without relying on the concentration variation and thus can be a promising tool for assigning precursor or source.²³ Machine learning methods have been applied successfully in various data sets of atmospheric chemical composition measurements, aiming at compound identification,^{24,25} aerosol classification,^{26,27} precursor apportionment,^{28,29} and the prediction of properties.^{30,31}

In this work, we measured the volatilities of ambient OOA with a thermal desorption method using a FIGAERO-CIMS. Then the volatilities were linked to the precursor information on OOA obtained from supervised and unsupervised machine learning models.³² The approach was applied to a group of offline and online aerosol samples to assess the OOA volatilities and precursors in a variety of atmospheric environments. The volatility distributions measured with FIGAERO-CIMS were also compared to those measured with a GC-MS method and an EC/OC analyzer.

2. MATERIALS AND METHODS

2.1. Sample Collection. To study the precursor and volatility of OOA, a total of 52 24 h PM_{2.5} samples were collected at multiple sites in Wuhan, a megacity in central China, including 30 and 22 samples collected before (i.e., 2014–2015) and after (i.e., 2021–2022) the implementation of PM control measures in China, respectively. A total of 44 24 h total suspended particulate (TSP) samples were collected on the deck of a research vessel during an expedition over the Yellow and East China Seas (YES), the West Pacific Ocean (WPO), and the East Indian Ocean (EIO) in 2021. The

samples were immediately frozen in dark (-18°C) conditions after collection until chemical analysis. The potential effect of extended storage time on those samples collected in 2014–2015 is discussed in section 3.1.4. Because OC resided predominantly in $<2.5\ \mu\text{m}$ size in both urban^{33,34} and marine³⁵ atmosphere, the impact of different particle cutting size (that is, TSP from marine locations and PM_{2.5} from urban locations) can be considered negligible for the purpose of comparing organic compositions between marine and urban aerosols. Detailed information about the locations and numbers of these samples (referred to as “daily sample” hereafter) can be found in Text S1 of the Supporting Information. The concentrations of organic carbon (OC) and elemental carbon (EC) in the samples were measured using a Sunset laboratory carbon analyzer following the IMPROVE thermal-optical protocol.³⁶ Based on OC and EC, we further calculated primary OC (POC) and secondary OC (SOC) values using the minimum R squared method (MRS) proposed by Wu and Yu.³⁷ Details of the MRS method are listed in Text S2.

To obtain the precursor and volatility information with high time resolution, hourly measurement of the OOA was conducted in a winter campaign from 5 December 2022 to 8 January 2023 by deploying an iodide-based FIGAERO-CIMS (Aerodyne Research Inc., USA) at an urban site in Wuhan (114.6157°E , 30.4577°N), which is the only provincial supersite operated by the local environmental authority for monitoring urban air quality of Wuhan. The campaign lasted for a total of 695 h. Particle number size distributions from 14 to 615 nm every 4 min and hourly OC concentration were monitored routinely at this supersite.

2.2. Determination of Molecular Composition and Volatility of OOA. **2.2.1. OOA Molecular Composition.** The design of FIGAERO-CIMS for hourly OOA measurement was described by previous studies.^{12,38} Briefly, the FIGAERO operated in a measurement cycle of 1 h alternating between gas-phase and particle-phase modes. During the gas-phase mode, ambient air flow was pulled through a PM_{2.5} cyclone (URG-2000-30EN, URG Corp.) and then a PTFE filter (2 μm Zefluor, 25 mm, Pall Corp.), where particles smaller than 2.5 μm were deposited. During the particle-phase mode, the molecules on the PTFE filter underwent thermal desorption in a heated ultrahigh-purity (UHP) nitrogen flow, which was kept at room temperature for the first 2 min to remove gas species adsorbed on the filter. Subsequently, it was heated to 200°C in 15 min, held at 200°C for an additional 15 min to ensure the desorption of the majority of organic matter,¹² and then cooled to room temperature within 4 min. The desorbed molecules were directed into a turbulent flow ion–molecule reactor (IMR), where gaseous molecules were ionized and subsequently detected as adduct ions with the reagent ion I[−]. A field blank sample was collected every 24 h.

The OOA in the daily samples were analyzed with the FIGAERO-CIMS in so-called “offline mode”. Details about the “offline mode” procedure were reported by previous studies.^{17,39–41} In brief, we cut a 3 mm diameter punch from each sample filter, placed them between two prebaked Zefluor Teflon filters sized at 25 mm, and installed them into the FIGAERO filter holder (referred to as the “sandwich technique”). The molecules on the filter punch were thermally desorbed and detected in the same way described above. The loadings of aerosol mass and thus nitrate were usually high on the filter punches of daily samples. To avoid simultaneous desorption of nitrate that depleted reagent ion I[−] significantly,

a slower ramping procedure was applied:³⁹ (1) from room temperature ($\sim 25^\circ\text{C}$) to 60°C in 5 min, (2) from 60 to 110°C in 10 min, (3) from 110 to 200°C in 10 min, and (4) held at 200°C for an additional 15 min. Blank quartz fiber filters were processed by following the same procedure.

OOA species were identified using a nontarget strategy. Mass calibration was performed using ions such as NO_3^- , $\text{C}_2\text{F}_3\text{O}_2^-$, $\text{IC}_2\text{H}_2\text{O}_2^-$, $\text{IC}_2\text{F}_3\text{HO}_2^-$, $\text{IC}_3\text{F}_5\text{HO}_2^-$, and I_3^- , encompassing a mass range from 62 to 381 m/z . The spectral peaks were iteratively fitted with multiple peaks using a custom peak shape until the residual was reduced to less than 5%.^{38,42} Subsequently, the exact masses of these multiple peaks were matched with the most probable elemental formulas within the range $\text{C}_{1-30}\text{H}_{1-60}\text{O}_{0-20}\text{N}_{0-2}\text{S}_{0-2}\text{X}_{0-1}^-$, where X stands for halogen atoms, with mass errors smaller than 10 ppm (mass resolution of ~ 6000). Isotope distribution was inspected to match with theoretical isotope pattern. Elemental ratio and double bond equivalent (DBE) limits of the formulas were $0.3 \leq \text{H/C} \leq 3$, $\text{N/C} \leq 0.5$, $\text{O/C} \leq 3$, $\text{S/C} \leq 1$, and $0 \leq \text{DBE} \leq 20$.⁴³⁻⁴⁵

In terms of quantification, we essentially followed the method by Lee et al.,³⁸ Iyer et al.,⁴⁶ Lopez-Hilfiker et al.,⁴⁷ and Isaacman-VanWertz et al.,⁴⁸ which constrain the CIMS's sensitivity to a broader range of compounds in real atmosphere than the limited number of laboratory calibration standards. Briefly, the number of moles of a compound X on a quartz fiber filter punch (or a PTFE filter for hourly sample) was calculated as $n_X = \frac{\text{signal}}{S_{\max} \times \frac{1}{S_0} \times T_X}$, where "signal" is the integrated peak area of ion signal of compound X during the thermal desorption period of the filter (cps, counts per second normalized to million reagent ions). S_{\max} (cps mol^{-1}) is the maximum sensitivity of the instrument, corresponding to the sensitivity of compounds like levoglucosan that clusters with the reagent ion at the collision limit. S_{\max} was obtained by calibrating levoglucosan. $1/S_0$ is the sensitivity of compound X relative to maximum sensitivity due to the declustering of compound X with reagent ions I^- . $1/S_0$ was obtained using the declustering scanning procedure. T_X is the mass dependent transmission efficiency of compound X relative to the I^- ion during transmission through the ion optics of the mass spectrometer. Mass concentration of compound X in the atmosphere during sample collection (ng m^{-3}) was then calculated from the number of moles of X on the filter, molecular weight, and air sample volume. The determinations of S_{\max}/S_0 and T_X are detailed in [Text S3](#). We subtracted the instrument background of compound X in each sample, which was determined by scaling the ratio of the signals during the last 3 min of the soak period of the sample and the corresponding field blank.⁴⁰ Only those OOA species with a unit mass peak area ratio $> 20\%$ and a sample-to-blank ratio > 2 were accepted for further analysis. On average, the mass fraction of the selected OOA to the total OOA was 60% in all the samples. These criteria excluded those small peaks with higher uncertainty in peak fitting and formula identification.

2.2.2. OOA Volatility. An OOA species should exhibit reproducible thermograms with a well-defined shape during the thermal desorption period of aerosol samples. We examined the thermograms of each OOA species in the daily samples and the hourly samples. Only those OOA species with well-defined thermogram shapes and discernible T_{\max} (the temperature at which a desorbed compound's ion signal reaches a maximum) were retained. Due to the presence of

isomers, isobaric compounds, or decomposition products, some molecular formulas may exhibit two or more T_{\max} values. In this study, we considered the isomers with different T_{\max} values as different species of the OOA, despite the same chemical formula they have. For each species, their T_{\max} might differ from sample to sample, depending on instrument conditions, filter loading, aerosol viscosity, etc.,⁴⁹ but the variation was usually less than 10°C . Those OOA species with the same formula and T_{\max} difference less than 10°C were considered as the same species, while those OOA species with the same formula but T_{\max} difference more than 10°C were isomers. This strategy was adopted to minimize errors attributable to isomers or inaccuracies during the multiple peak fitting process. It should be noted that the alignment of the same species with T_{\max} less than 10°C was not mandatory between the daily samples and the hourly samples, because the filter material (quartz fiber filter vs PTFE filter) changed T_{\max} significantly. In the end, we identified a total of 1162 OOA species, which is a compilation of the species obtained from all of the samples in this study. In a next step, we transformed the T_{\max} of an OOA species into saturation concentration $\log C^*$.¹⁴ Following the method of Bannan et al.,⁵⁰ the calibration of the correlation between T_{\max} and $\log C^*$ was performed by measuring the T_{\max} of polyethylene glycol (PEG) series 4–8 standards with known $\log C^*$ spiked on blank filters. The calibration was performed on blank quartz fiber filters for the daily samples and PTFE filters for the hourly samples to account for the effect of filter material on the thermogram. The detailed calibration process is presented in [Text S4](#).

GC-MS can indirectly measure the volatility of an organic compound according to its retention time in GC elution. The GC-MS method has been mostly used for the volatility measurement of nonpolar or moderate-polar organic compounds in fresh source aerosols but rarely for ambient OA.⁵¹⁻⁵⁴ For comparison, the volatility of OA in the daily samples was also measured by GC-MS. The sample analysis process is shown in [Text S5](#). Following the method of Lu et al.⁵⁴ and Zhao et al.,⁵³ we segmented the GC elutes of each sample into 25 bins according to the retention times of C12 to C36 *n*-alkanes. The method assumes that the $\log C^*$ of an eluting compound in a bin can be approximated by the $\log C^*$ of the corresponding *n*-alkane in the same bin. The total mass of all the eluting compounds in a bin can be semiquantified using the total ion current (TIC) signal in the bin and the response factor of the *n*-alkane in the bin. The response factors of C12–C36 *n*-alkanes were calibrated using their respective authentic standards.

2.3. Assignment of OOA Species Using Machine Learning. 2.3.1. Assigning the Precursors of OOA Species.

To assign potential precursors of the OOA species in daily samples and hourly samples, we first singled out the OOA species containing sulfur atoms. These sulfur-containing OOA species are mostly organosulfates with $\text{O/S} \geq 4$ that could be formed from any VOC precursor, but SO_2 or sulfate is always involved in their formation. Second, a list of known biomass burning OOA species was created on the basis of a literature review of previous laboratory or field studies. This list excluded those formulas that were also observed in previous aromatic VOC oxidation laboratory or field studies and thus included exclusively those biomass burning tracers reported in the literature. [Tables S2 and S3](#) list the 107 sulfur-containing and 44 biomass burning OOA species, respectively.

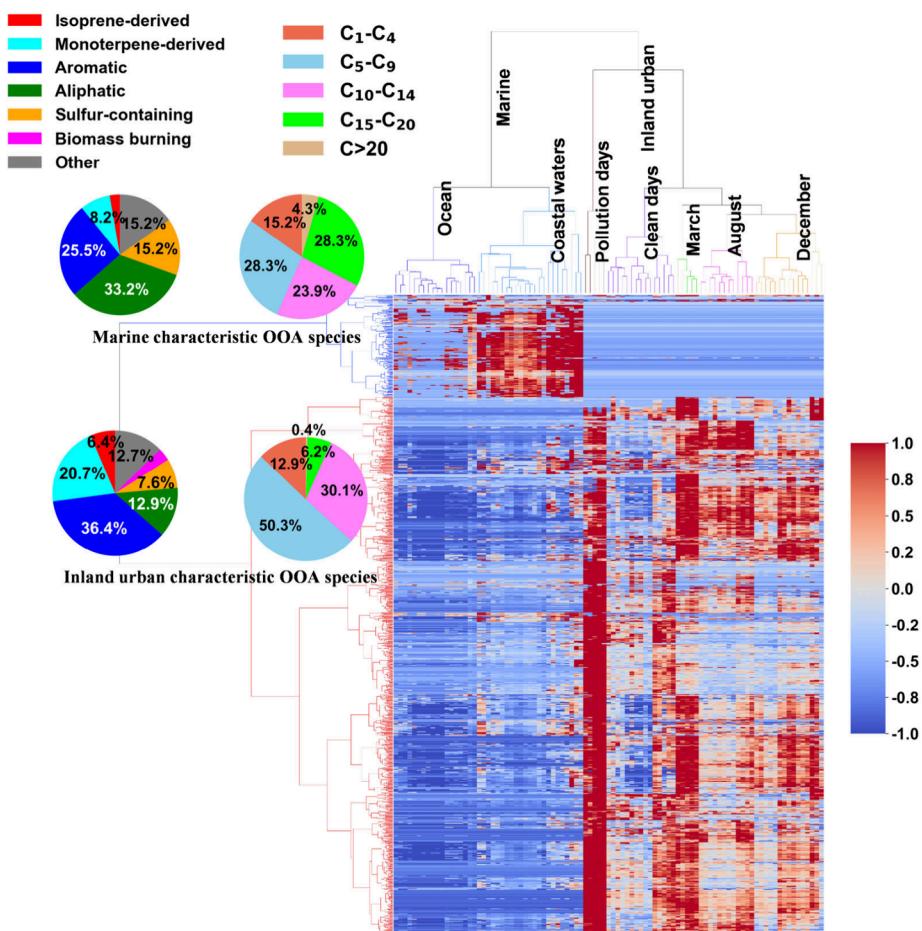


Figure 1. Hierarchical clustering analysis results for the daily samples. The columns correspond to the samples; the rows correspond to the OOA species. The colors in the hierarchical clustergram represent the values corresponding to the concentration of the OOA species after undergoing Z-transformation, which is a process of standardizing data by normalizing the difference of an OOA species between the mass concentration in an individual sample (ng m^{-3}) and the mean concentration in all 96 samples by the standard deviation. The values greater than 1 or less than -1 are rounded to 1 or -1, respectively. The pie chart displays the number percentages of precursor types or carbon number categories in the characteristic OOA species of inland urban and marine samples, respectively. If an OOA species was assigned to multiple precursor types, the number percentage of the species was equally distributed to each precursor type.

For the remaining 1011 OOA species, we employed a supervised machine learning method to predict their precursors. (1) We created a training data set including 4 precursor types of known oxidation products from isoprene (82 formulas), monoterpene (126 formulas), aliphatic (215 formulas), and aromatic (118 formulas) VOC precursors, respectively, on the basis of literature review of previous VOC oxidation laboratory or field studies using I-CIMS. Those OOA species with DBE < 2 were also regarded as the products from aliphatic precursors.⁵⁵ The 4 precursor types are named as “monoterpene-derived OOA” species, “isoprene-derived OOA” species, “aliphatic OOA” species, and “aromatic OOA” species, which are listed in Tables S4–S7, respectively. Please note that a known OOA species in the training data set may be derived from multiple precursors and thus belongs to multiple precursor types (Figure S7). The known species in the training data set jointly contributed 56.9% in mass to all the 1011 OOA species. (2) We selected 8 features for the machine learning of the OOA species. The first 4 features, the number of carbon atoms (n_C), hydrogen atoms (n_H), oxygen atoms (n_O), and nitrogen atoms (n_N), represent elemental composition; double bond equivalent (DBE) and H/C ratio indicate the carbon saturation state of the OOA species; oxygen-to-

carbon ratio (O/C) and oxidation state of carbon (OSc) reflect their carbon oxidation state. The average values of the 8 features of isoprene-derived OOA species, monoterpene-derived OOA species, aliphatic OOA species, and aromatic OOA species in the training data set are shown in Table S8. (3) A random forest algorithm, which is performed with the sklearn package (v1.2.2) in Python, was selected for assigning precursors for other unknown OOA species. (4) The model outputs the occurrence probability of an unknown OOA species in each precursor type. By setting an appropriate threshold of occurrence probability (0.6 in this study), an unknown OOA is assigned to one or more precursor types according to the probabilities of its occurrence in the four types. An OOA species is classified as “Others” if none of its occurrence probabilities in the four types exceed the threshold. For more details about assigning an OOA species to multiple precursor types, please refer to Text S6.

The accuracy of the random forest model was evaluated on the basis of 3 parameters: *Precision*, *Recall*, and *F1-score*. Taking monoterpene as an example, *Precision* characterizes how many OOA species that are predicted as monoterpene oxidation products are truly derived from monoterpene. *Recall* characterizes how many monoterpene-derived OOA species are

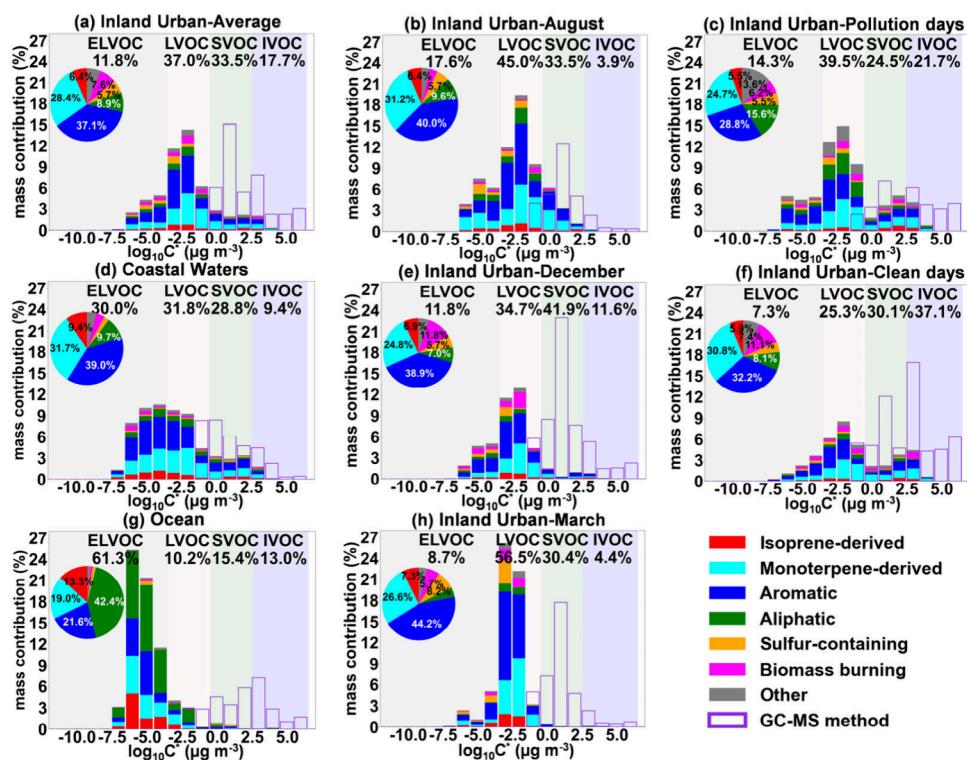


Figure 2. OOA volatility distribution and precursor contribution in the 7 scenarios clustered by the HCA. The volatility distribution of the OA obtained using the GC-MS technique is integrated with the OOA measured by FIGAERO-CIMS, assuming the two instruments measured different organic compounds. ELVOC: $C^* < 3 \times 10^{-4} \mu\text{g m}^{-3}$; LVOC: $3 \times 10^{-4} \mu\text{g m}^{-3} < C^* < 0.3 \mu\text{g m}^{-3}$; SVOC: $0.3 \mu\text{g m}^{-3} < C^* < 300 \mu\text{g m}^{-3}$; IVOC: $300 \mu\text{g m}^{-3} < C^* < 3 \times 10^6 \mu\text{g m}^{-3}$ according to Donahue et al.¹⁰

predicted as monoterpene-derived OOA species. *F1-score* is the harmonic mean of *Precision* and *Recall*, serving as a measure of the model's accuracy and robustness. In order to assess the performance of the machine learning model and determine the optimal number of decision trees, the training data set was randomly divided into two subsets: 70% of the data was allocated to the training subdata set for model training, and the remaining 30% formed the testing subdata set for assessing the performance of the trained model. As shown in Figure S8, the *F1-score* tends to stabilize as the number of decision trees increases. We ultimately selected 50 decision trees as the optimal number. Based on the evaluation of *F1-scores*, the highest *F1-score* of 0.95 was achieved for the aliphatic OOA species, while the lowest *F1-score* of 0.7 was achieved for the monoterpene-derived OOA species, showing the model's high accuracy in predicting the precursor types of the OOA species.

To reduce the biases introduced by random splitting of training and testing data sets, we further performed 10-fold cross-validation by dividing the training data set into ten groups. In each iteration, one group was used as the testing subdata set, while the remaining nine groups were used as the training subdata set. The iteration was repeated ten times. As shown in Figure S9, the average *F1-scores* for each iteration remain above 0.7. Therefore, it is reliable to train the model using the entire training data set that we have created for predicting precursor types of unknown OOA species.

2.3.2. Clustering OOA Groups According to Mw, T_{max} , n_C and concentration. Some studies, like Stark et al.,¹⁵ suggested that a large fraction of the compounds detected in ambient OA by FIGAERO-CIMS may be a result of thermal decomposition. To explore this issue, we employed a K-means algorithm,⁵⁶ which is an unsupervised machine learning

approach, to cluster the OOA species to High-Volatility Low-Mw OOA species, Low-Volatility Low-Mw OOA species, High-Volatility Median-Mw OOA species, Low-Volatility Median-Mw OOA species, High-Mw OOA species, and decomposition fragment on the basis of their T_{max} value, molecular weight (Mw), carbon number (n_C), and concentration measured at the urban supersite during the winter campaign. Here the concentration of an OOA species is the average concentration during the campaign. The input variables of the K-means algorithm were scaled to values between 0 and 1, based on their respective numerical ranges, to prevent any bias associated with the relative magnitude of each variable.

2.4. Hierarchical Clustering Analysis of Sample Groups. We employed hierarchical clustering analysis (HCA, Scipy and Seaborn package) to categorize the 96 daily samples according to the molecular composition measured by FIGAERO-CIMS. HCA is an unsupervised machine learning approach that merges individual observations into clusters of similar behavior through the determination of the similarity between every pair of objects in a data matrix.⁵⁷ The input for HCA is a matrix of the concentrations (ng m^{-3}) of OOA species in the 96 samples. HCA was performed separately on species rows and sample columns. The proximity between an observation pair was determined by the Euclidean distance for both species rows and sample columns. We used the Ward linkage method for merging different species and samples into species clusters and sample clusters, respectively.

3. RESULTS AND DISCUSSION

3.1. Volatility and Precursors of OOA in Different Atmospheric Environments. The HCA result shows that the molecular composition of ambient OOA differed systematically according to location, season, and pollution level. As depicted in the hierarchical clustergram (Figure 1), the samples primarily fell into 2 categories featuring distinct distributions of OOA species: inland urban vs marine. The marine samples were characterized by 186 OOA species, in which aliphatic OOA species accounted for 33.2% in species number and the OOA species with a carbon number of >15 accounted for 32.6% in species number. In contrast, the inland urban samples were characterized by 976 OOA species, in which aromatic OOA species accounted for 36.4% in species number and the OOA species with carbon number between 5 and 9 accounted for 50.3% in species number.

The HCA further clustered the marine samples into two subcategories: 18 “Ocean” samples ($OC: 1.9 \pm 0.8 \mu\text{g m}^{-3}$; $EC: 0.2 \pm 0.1 \mu\text{g m}^{-3}$) collected above open ocean and 23 “Coastal Waters” samples ($OC: 4.2 \pm 2.4 \mu\text{g m}^{-3}$; $EC: 0.8 \pm 0.6 \mu\text{g m}^{-3}$) collected near the coastline. The inland urban samples were first clustered into those collected before or after strict PM control measure implementation in China. Those collected after PM control were further clustered to two subcategories “clean days” (15 samples; $OC: 14.7 \pm 4.3 \mu\text{g m}^{-3}$; $EC: 5.5 \pm 2.3 \mu\text{g m}^{-3}$) and “pollution days” (5 samples; $OC: 27.0 \pm 5.0 \mu\text{g m}^{-3}$; $EC: 5.6 \pm 1.5 \mu\text{g m}^{-3}$) according to the OC level on the sampling days. The three subcategories of samples collected before PM control coincided with the months when they were collected: five March samples ($OC: 15.6 \pm 1.8 \mu\text{g m}^{-3}$; $EC: 5.0 \pm 1.3 \mu\text{g m}^{-3}$), 12 August samples ($OC: 8.4 \pm 2.2 \mu\text{g m}^{-3}$; $EC: 2.6 \pm 1.2 \mu\text{g m}^{-3}$), and 15 December samples ($OC: 20.6 \pm 4.6 \mu\text{g m}^{-3}$; $EC: 6.6 \pm 1.4 \mu\text{g m}^{-3}$). In sections 3.1.1–3.1.3 we discuss the variation of the OOA volatility and precursor according to the 7 sample subcategories (i.e., 7 scenarios) here.

The volatility distribution of ambient OOA is shown as volatility basis set (VBS) plots (Figure 2), which lumps the OOA species into evenly spaced volatility bins according to their $\log C^*$. The OOA measured with the FIGAERO-CIMS was predominantly distributed in the ELVOC and LVOC ranges (67.0%–98.5% in total). It should be noted that the OOA measured with the FIGAERO-CIMS stands for only those organic species being desorbed below 200 °C and accounted for only $16 \pm 13\%$ of the total OA ($OC \times 1.4$) measured with the thermal-optical method using the IMPROVE protocol. Moreover, we found the linear correlation between OOA and OC2 (evolve in 120–250 °C, Pearson correlation coefficient $r = 0.81$) was better than those between OOA and OC1 (room temperature to 120 °C, $r = 0.66$), OC3 (250–450 °C, $r = 0.51$), or OC4 (450–550 °C, $r = 0.65$). The OOA had a higher correlation with SOC ($r = 0.83$) than POC ($r = 0.37$). Apparently, the FIGAERO-CIMS cannot measure the full range of volatilities of ambient OA. The OOA measured by the FIGAERO-CIMS represented mostly the secondary and moderate-volatility portion of ambient OA.

On the other hand, the organic species measured by the GC-MS fell primarily into the SVOC and intermediate volatility organic compounds (IVOC) ranges (78.8%–91.7% in total) and accounted for $10 \pm 7\%$ of the total OA ($OC \times 1.4$). This portion of OA had a higher correlation with POC ($r = 0.77$)

than with SOC ($r = 0.58$). Therefore, the organic species measured by the GC-MS represented mostly the primary and high-volatility portion of ambient OA. To obtain a full range of volatility of ambient OA, other techniques are need to complement FIGAERO-CIMS.⁵⁸ The mass fractions of VBS reported in this study was calculated independently for the OOA concentrations measured by the FIGAERO-CIMS and the OA measured by the GC-MS, by assuming that there was minimal overlapping between them.

3.1.1. Inland Urban vs Marine. Figure 2a shows the average volatility distribution and precursor contribution of all urban Inland Urban samples. Inland urban OA fell primarily into the LVOC (37.0%) and SVOC (33.5%) range. In terms of precursor contribution, aromatic OOA species accounted for 37.1% of the OOA, which indicates a significant impact of anthropogenic aromatic precursors on the chemical composition of the OOA. In contrast, the OA in the Ocean samples (Figure 2g) was characterized by high fractions of ELVOCs (61.3%) and aliphatic OOA (42.4%). Studies have shown that marine surface microbial activities, especially during bubble bursting, can release aliphatics into the atmosphere.⁵⁹ These long carbon chain aliphatic compounds and their aging products might have contributed to the aliphatic OOA in ocean aerosols. The substantial amount of ELVOC might be due to the presence of these aliphatic compounds, as well as the aging process during the long-distance transport to those remote marine locations.⁴ In addition, the observed low volatility of marine OOA might also be partly attributed to the high content of sea salts in marine aerosols. Sea salt aerosols had stronger hygroscopicity in humid ocean conditions, leading to the formation of a liquid water phase, where aqueous phase production of ELVOC was enhanced.^{60,61} On the other hand, urban aerosols had more OC ($4.2 \pm 2.4 \mu\text{g m}^{-3}$) and EC ($0.8 \pm 0.6 \mu\text{g m}^{-3}$) than marine aerosols (OC: $1.9 \pm 0.8 \mu\text{g m}^{-3}$; EC: $0.2 \pm 0.1 \mu\text{g m}^{-3}$), partly enhancing the partitioning of more SVOCs to the particle phase.⁶²

The OOA precursor contribution in the Coastal Waters samples (Figure 2d) resembled that of inland urban samples (e.g., aromatic OOA: 39.0%), but the OA was less volatile than that of the inland urban samples with the fractions of ELVOCs 30.0%, LVOCs 31.8%, SVOCs 28.8%, and IVOCs 9.4%, respectively. This likely implies that the OOA in the atmosphere above coastal waters received significant input from continental emission sources but was more aged and thus less volatile than those in the urban atmosphere.

3.1.2. Summer vs Winter in the Inland Urban Atmosphere. The most remarkable difference of OOA precursors between summer and winter was monoterpene-derived OOA and biomass burning OOA. The fraction of monoterpene-derived OOA was higher in summer (31.2%) than that in winter (24.8%), while the fraction of biomass burning OOA in winter (11.8%) was >3 times of that in summer (3.4%). The overall volatility of OA was lower in the summer (August) samples than the winter (December) samples (ELVOC +LVOC: 62.6% in summer vs 46.5% in winter; Figure 2b,e). Several factors might have jointly contributed to such differences. (1) Low temperature in winter drove more volatile vapors from gas to particles.⁶² (2) The extended daylight hours, high temperature, and high O₃ concentration (176 $\mu\text{g m}^{-3}$ in summer vs 89 $\mu\text{g m}^{-3}$ in winter) in summer promoted the aging of aerosol particles.^{63,64} (3) More primary OA with high volatility was emitted into the atmosphere in winter.⁶⁵

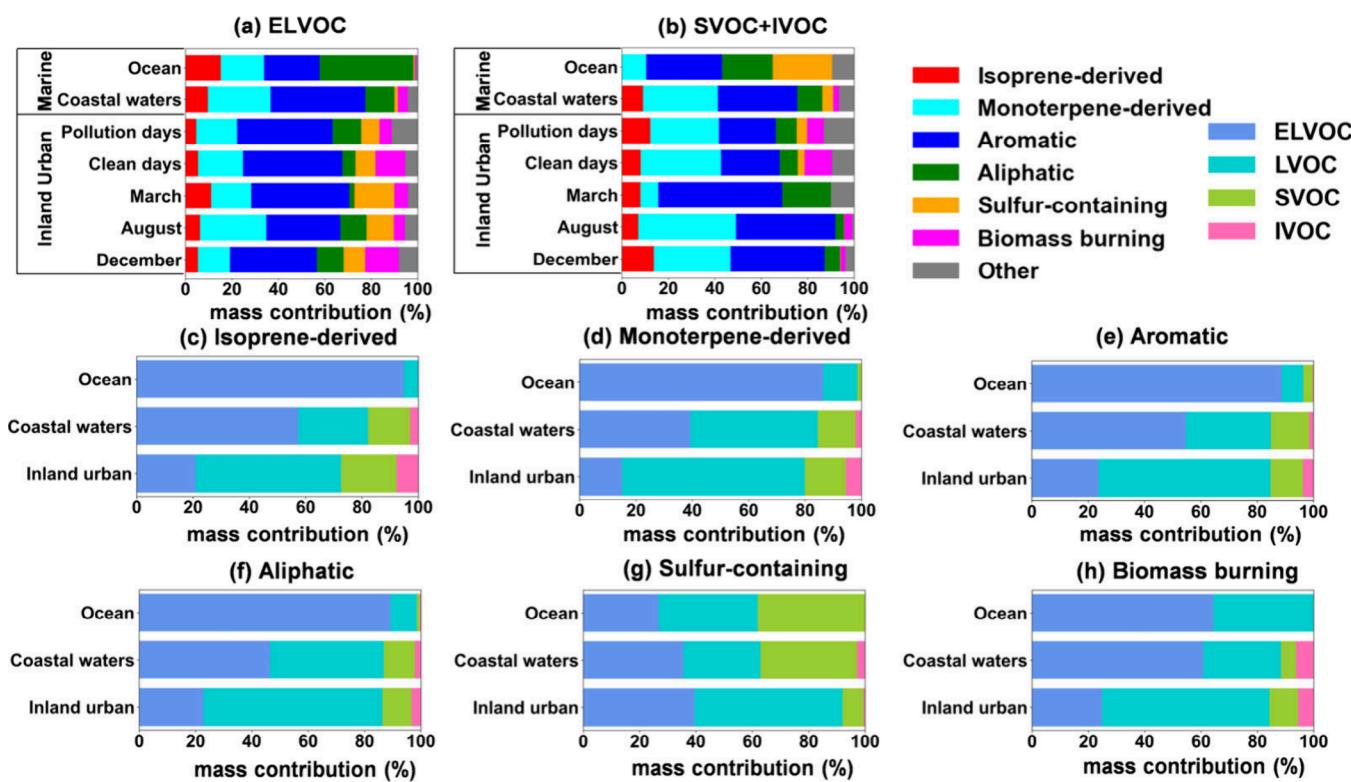


Figure 3. (a and b) The compositions of ELVOC and SVOC+IVOC, in terms of the precursors of the OOA precursors. (c–h) The volatilities of sulfur-containing OOA, biomass burning OOA, aromatic OOA, aliphatic OOA, isoprene-derived OOA, and monoterpane-derived OOA in the 3 sample categories.

3.1.3. Pollution Days vs Clean Days in the Inland Urban Atmosphere. OA was less volatile on the pollution days than the clean days (ELVOC+LVOC: 53.8% on the polluted days vs 32.6% on the clean days, **Figure 2c,f**), indicating OA was more aged and functionalized on the pollution days. We observed a notable increase of aliphatic and unidentified OOA species from 15.5% on the clean days to 29.2% on the polluted days. Previous measurements near traffic centers at urban sites of China observed high content of aliphatic compounds compared to other sites.⁴⁸ We expect these aliphatic OOA were associated with vehicle emissions.^{66–68}

3.1.4. Potential Effects of Extended Storage Time and PM Control Measures. In this study, we collected offline samples of urban aerosols in Wuhan during the winter season both before (December 2014) and after (December 2021) strict PM control measure implementation in China. By examining these 2 groups of samples, we investigated the potential effects of extended storage time and PM control measures on the OOA volatility distribution and precursor contribution. OC concentrations were $20.6 \pm 4.6 \mu\text{g m}^{-3}$ and $20.4 \pm 6.0 \mu\text{g m}^{-3}$ during the sampling periods in December 2014 and 2021, respectively. As shown in **Figure S10**, the year-2021 samples had higher mass fraction of IVOC+SVOC (22.5%) than the year-2014 samples (9.2%), which might be attributed to the volatilization of highly volatile species during the extended storage time. In terms of precursor, the year-2021 samples had less contribution from aromatic OOA species (31.8%) but more contribution from monoterpane-derived OOA species (30.0%) than the year-2014 samples (38.4% and 25.2%). The reduction of aromatic OOA species was likely caused by the anthropogenic VOC elimination efforts made in the past 10 years. Finally, it should be noted that other factors like

emission source, weather conditions, and even sample number might have also contributed to the observed differences of the OOA composition and volatility. Rigorous controlled experiments are needed to verify the above hypothesis.

3.1.5. The Link between Precursors and Volatility of OOA. To know the link between the precursors and the volatility of the OOA, we investigated (1) the compositions of ELVOC and SVOC+IVOC in terms of the OOA precursors (**Figure 3a,b**) and (2) the volatilities of sulfur-containing OOA, biomass burning OOA, aromatic OOA, aliphatic OOA, isoprene-derived OOA, and monoterpane-derived OOA (**Figure 3c–h**).

There is a significant difference of ELVOC precursors between marine and inland urban atmosphere (**Figure 3a**). In the ocean samples, the predominant OOA precursor in the ELVOC range was aliphatic OOA (40.4%). In contrast, aromatic OOA (39.1%) was the most abundant ELVOC in the inland urban samples, followed by monoterpane-derived OOA (19.3%) and sulfur-containing OOA (10.9%). Considering the important role of ELVOC in particle formation and growth, these results underscore the importance of controlling aromatic OOA in urban areas. The most abundant sulfur-containing OOA species in the inland urban samples were $\text{CH}_4\text{O}_3\text{S}$, $\text{C}_5\text{H}_8\text{O}_4\text{S}$, and $\text{C}_6\text{H}_{12}\text{O}_6\text{S}$. These species were clustered to thermal decomposition fragments in our study using the K-means algorithm due to their relatively low Mw but high thermal desorption temperature.

In comparison with ELVOC, SVOC and IVOC in the inland urban samples (**Figure 3b**) had more contribution from monoterpane-derived OOA (29.4%), less sulfur-containing OOA (1.6% in SVOC and IVOC vs 10.9% in ELVOC), and comparable contribution from aromatic OOA (37.3%). In the

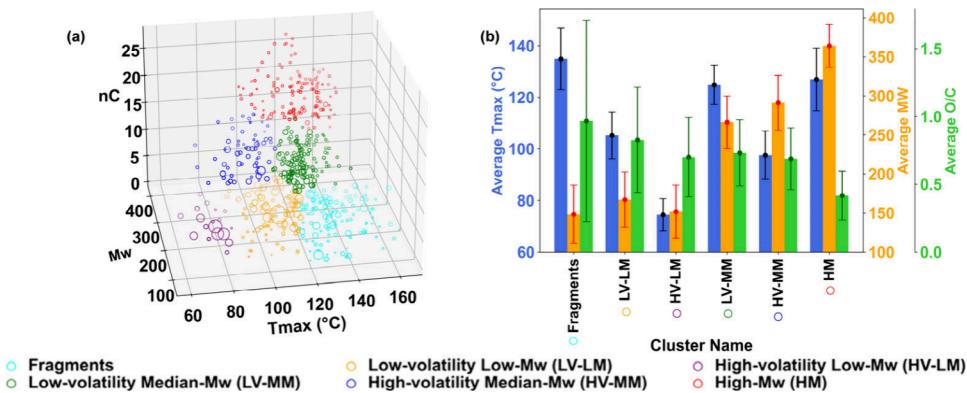


Figure 4. Groups of OOA species using K-means clustering analysis. (a) Three-dimensional distribution of OOA species by their T_{\max} , molecular weight, carbon number, and concentration. The size of the circle is proportional to the concentration of the OOA species. (b) Average and standard deviation of molecular weight, T_{\max} and O/C ratios of the OOA species in each group.

ocean samples, we observed that 25.6% of SVOC+IVOC was sulfur-containing OOA species (Figure 3b), among which the most abundant species are $C_{13}H_{18}O_8S$, $C_{14}H_{20}O_8S$, $C_{13}H_{18}O_9S$, and $C_{18}H_{14}O_5S_2$. The four species accounted for 14.1% of SVOC+IVOC and were most likely organosulfates formed from ocean-origin long-chain organic compounds.

The volatility of the OOA from each precursor type varied with sample locations. In the ocean atmosphere, the OOA from the 5 precursor types fell primarily into the ELVOC range (64.4%–94.9%, Figure 3c–f,h) except S-containing OOA (26.7%, Figure 3g). In the urban locations, LVOC was the most important volatility set in the OOA from all precursor types, followed by ELVOC (Figure 3c–h). In total, the ranking of OOA volatility in urban aerosols is sulfur-containing OOA < aliphatic OOA < aromatic OOA < biomass burning OOA < monoterpene-derived OOA < isoprene-derived OOA.

3.2. Temporal Variations of OOA Volatility and Precursor in the Winter Campaign. Online measurements, with hourly time resolution, allowed the studies of the volatility and precursor of the OOA species over short periods, as well as their diurnal variations, which cannot be achieved with offline measurement. Before looking into the online samples, we examined the consistency between the online samples and offline samples. As shown in Figure S10, OOA volatility distribution and precursor contribution are roughly consistent between the online samples and offline samples collected in winter at the urban supersite, except that the VBS of the offline samples showed an additional minor peak in the SVOC and IVOC range. We are not surprised to see such difference, considering the two groups of samples were collected using different filter substrates (PTFE filter vs quartz fiber filter) in different years (the winter of 2022 vs the winter of 2021).

For the winter campaign, we used the K-means clustering algorithm to group the OOA species into 6 groups based on their Mw, T_{\max} , n_C and concentration (Figure 4a). The OOA species with low Mw (157.5 ± 32.6) and n_C (5.9 ± 1.6) were grouped into high-volatility low-Mw OOA species ($T_{\max}: 74.4 \pm 6.3$ °C) and low-volatility low-Mw OOA species ($T_{\max}: 101.2 \pm 6.3$ °C). The OOA species with medium Mw (269.4 ± 38.3) and n_C (10.4 ± 2.1) were divided into high-volatility median-Mw OOA species ($T_{\max}: 97.7 \pm 9.2$ °C) and low-volatility median-Mw OOA species ($T_{\max}: 124.0 \pm 8.3$ °C). The OOA species with high Mw (362.6 ± 29.0), high n_C (18.1 ± 3.0), high T_{\max} (126.9 ± 12.0 °C), and low O/C (0.4 ± 0.2) were grouped to high-Mw OOA species. The OOA species

with low molecular weight (149.1 ± 36.4), low n_C (5.4 ± 1.8), but high O/C ratios (0.9 ± 0.7) and high T_{\max} (132.4 ± 12.9 °C), represented thermal decomposition fragments from less volatile substances.⁵⁶ On average, the fragments accounted for 25.1% in species number and 26.8% of the mass of the OOA species during the winter campaign.

The winter campaign experienced four PM_{2.5} episodes, with a prolonged haze event from December 26 to January 3. Figure 5a shows the volatility, precursor, and group of OOA in 1 h time resolution. Total OOA concentration correlated closely with OC ($r = 0.82$) and accounted for $26 \pm 8\%$ of OA by assuming OA = OC × 1.4 during the winter campaign. We selected five time periods for our further study. Details can be found in Table S9.

From a diurnal perspective (Figure 5b), the OOA in the daytime had a higher fraction of ELVOC + LVOC (99.5%) and decomposition fragments (41.0%) than those in the nighttime (68.5% and 17.0%). On the other hand, the OOA in the daytime OOA had a higher fraction of aromatic OOA (42.9%), while that in the nighttime had a higher fraction of aliphatic OOA (19.0%). In summary, this evidence suggests that (1) active photochemistry in the daytime promoted the formation of less volatile aromatic OOA species and (2) low temperature at night enhanced the partitioning of more volatile aliphatic species to particle phase.

We compare the precursor and volatility of OOA in typical clean, particle growth, and pollution periods during the winter campaign (Figure 5b). When the PM episode developed from clean to particle growth and then to pollution period, the OOA species shifted from Low-Mw OOA species to Median-Mw OOA species and then to decomposition fragments. In the end, the mass fractions of ELVOC (43.0%) and decomposition fragments (51.1%) were highest during the pollution period, suggesting the enhanced formation of highly nonvolatile species. The OOA precursors showed relatively little variation among the three periods, except that a relatively higher contribution from biomass burning OOA was observed in the clean period. This is consistent with the observations by Cai et al.⁴⁰ who found that biomass burning emissions played an important role during winter clean periods in Beijing, China.

4. CONCLUSION

Accounting for $16 \pm 13\%$ of OA, the OOA measured by the FIGAERO-CIMS represented mostly the secondary and moderate-volatility portion of ambient OA in urban and

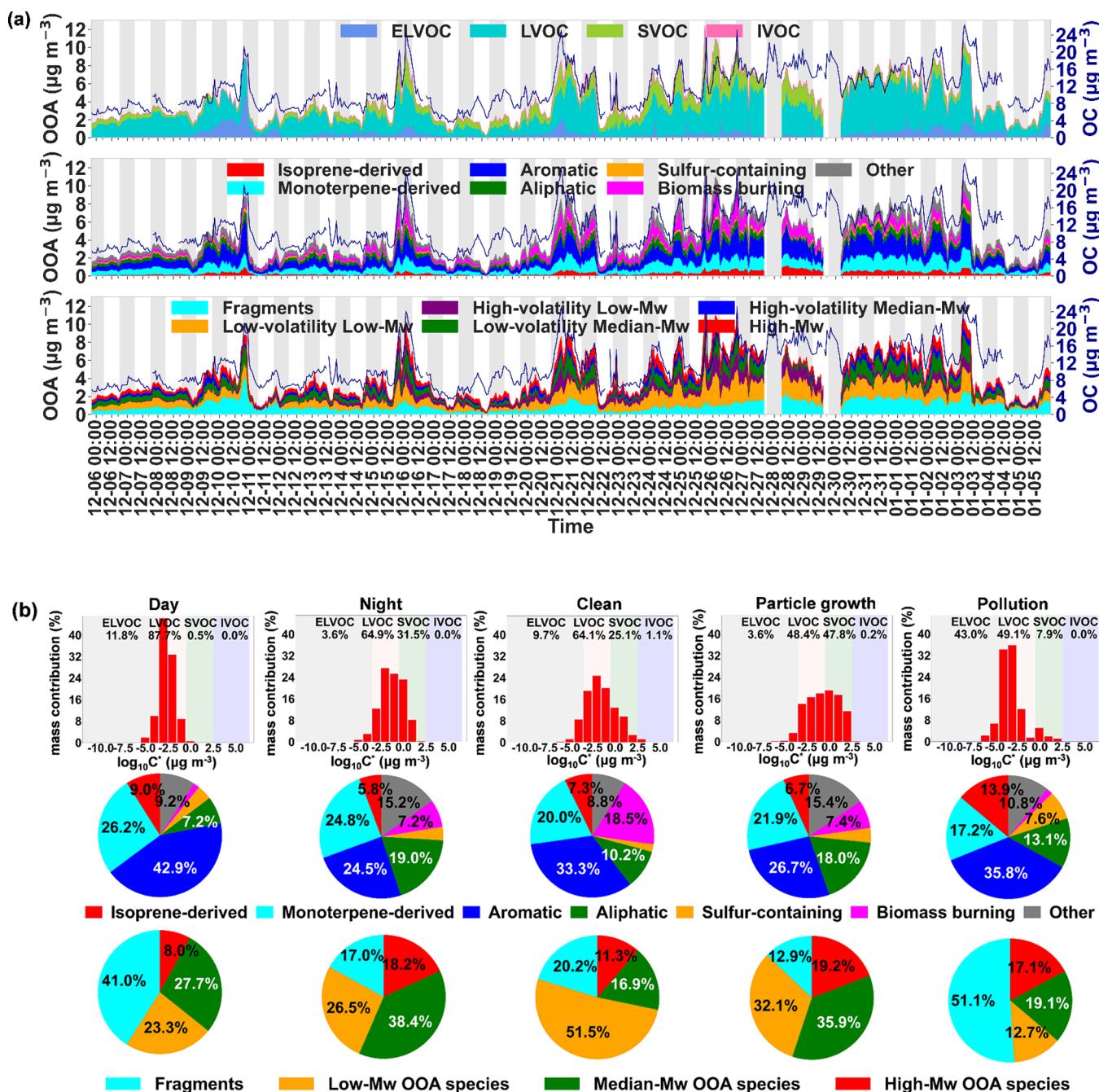


Figure 5. (a) Time series of volatility, precursor, and group of the OOA during the winter campaign. Also shown is OC concentration measured independently by Sunset ECOC analyzer. (b) The OOA volatility distribution, OOA precursor contribution, and OOA group fraction in selected atmospheric conditions.

marine atmosphere. To obtain a full range of volatilities of ambient OA, other techniques, like GC-MS, are needed to complement FIGAERO-CIMS. Our study shows that the molecular composition of ambient OOA differed systematically according to location, season, and pollution level. The OOA in the ocean samples was characterized by high fractions of ELVOC and aliphatic species, while the inland urban OOA is characterized by aromatic species and falls primarily within the LVOC and SVOC range. These results underscore the importance of controlling aromatic OOA precursors in urban areas. In terms of temporal variation, the volatility of the OOA in the inland urban atmosphere was lower in summer, pollution days, and daytime than that in winter, clean days,

and nighttime. During the winter campaign at the urban site, our study suggested that on average 25.1% in species number and 26.8% in mass of the OOA species detected by the FIGAERO-CIMS might be attributed to thermal decomposition fragments. When the PM episode developed from clean to particle growth and then to pollution period, the OOA species shifted from Low-Mw OOA species to Median-Mw OOA species and then to decomposition fragments, resulting in the enhanced formation of highly nonvolatile species during the pollution period.

■ ASSOCIATED CONTENT

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsestair.4c00076>.

Sample collection details (Text S1, Tables S1 and S9, Figure S1), measurements and analysis methods (Text S2–S5, Figures S2–S6), machine learning details (Text S6, Tables S2–S8, Figures S7–S9), and the comparisons of OOA volatility distribution and precursor contribution in certain samples (Figure S10) ([PDF](#))

■ AUTHOR INFORMATION

Corresponding Author

Huan Yu — School of Environmental Studies, China University of Geosciences, Wuhan 430074, China; Research Centre for Complex Air Pollution of Hubei Province, Wuhan 430074, China; orcid.org/0000-0001-6078-8192; Email: yuhuan@cug.edu.cn

Authors

Xinyu Wang — School of Environmental Studies, China University of Geosciences, Wuhan 430074, China

Yongyi Zhao — School of Environmental Studies, China University of Geosciences, Wuhan 430074, China

Ke Hu — Wuhan Environmental Monitoring Center, Wuhan 430022, China

Jian Wang — Frontiers Science Center for Deep Ocean Multispheres and Earth System, and Key Laboratory of Marine Chemistry Theory and Technology, Ministry of Education, Ocean University of China, Qingdao 266100, China

Qiongqiong Wang — School of Environmental Studies, China University of Geosciences, Wuhan 430074, China; Research Centre for Complex Air Pollution of Hubei Province, Wuhan 430074, China

Nan Chen — Research Centre for Complex Air Pollution of Hubei Province, Wuhan 430074, China; Hubei Ecological Environment Monitoring Center Station, Wuhan 430070, China

Bo Zhu — Research Centre for Complex Air Pollution of Hubei Province, Wuhan 430074, China; Hubei Ecological Environment Monitoring Center Station, Wuhan 430070, China

Hong-Hai Zhang — Frontiers Science Center for Deep Ocean Multispheres and Earth System, and Key Laboratory of Marine Chemistry Theory and Technology, Ministry of Education, Ocean University of China, Qingdao 266100, China; orcid.org/0000-0001-9576-4035

Complete contact information is available at:

<https://pubs.acs.org/doi/10.1021/acsestair.4c00076>

Author Contributions

H.Y. and H.-H.Z. designed the study. X.W., Y.Z., K.H., J.W., N.C., and B.Z. contributed to sample collection and chemical analysis. X.W. and H.Y. analyzed the data and wrote the manuscript. Q.W. reviewed and revised the manuscript.

Funding

This research was supported by the National Natural Science Foundation of China (42175131) and the Fundamental

Research Funds (No. G1323523063) for the Central Universities, China University of Geosciences (Wuhan).

Notes

The authors declare no competing financial interest.

■ REFERENCES

- (1) Lu, Q.; Murphy, B. N.; Qin, M.; Adams, P. J.; Zhao, Y.; Pye, H. O. T.; Efstratiou, C.; Allen, C.; Robinson, A. L. Simulation of organic aerosol formation during the CalNex study: updated mobile emissions and secondary organic aerosol parameterization for intermediate-volatility organic compounds. *Atmos. Chem. Phys.* **2020**, *20* (7), 4313–4332.
- (2) Zhao, B.; Wang, S.; Donahue, N. M.; Jathar, S. H.; Huang, X.; Wu, W.; Hao, J.; Robinson, A. L. Quantifying the effect of organic aerosol aging and intermediate-volatility emissions on regional-scale aerosol pollution in China. *Sci. Rep.* **2016**, *6* (1), No. 28815.
- (3) Stolzenburg, D.; Wang, M.; Schervish, M.; Donahue, N. M. Tutorial: Dynamic organic growth modeling with a volatility basis set. *J. Aerosol Sci.* **2022**, *166*, No. 106063.
- (4) Feng, T.; Wang, Y.; Hu, W.; Zhu, M.; Song, W.; Chen, W.; Sang, Y.; Fang, Z.; Deng, W.; Fang, H.; Yu, X.; Wu, C.; Yuan, B.; Huang, S.; Shao, M.; Huang, X.; He, L.; Lee, Y. R.; Huey, L. G.; Canonaco, F.; Prevot, A. S. H.; Wang, X. Impact of aging on the sources, volatility, and viscosity of organic aerosols in Chinese outflows. *Atmos. Chem. Phys.* **2023**, *23* (1), 611–636.
- (5) Mohr, C.; Thornton, J. A.; Heitto, A.; Lopez-Hilfiker, F. D.; Lutz, A.; Riipinen, I.; Hong, J.; Donahue, N. M.; Hallquist, M.; Petäjä, T.; Kulmala, M.; Yli-Juuti, T. Molecular identification of organic vapors driving atmospheric nanoparticle growth. *Nat. Commun.* **2019**, *10* (1), 4442.
- (6) Donahue, N. M.; Epstein, S. A.; Pandis, S. N.; Robinson, A. L. A two-dimensional volatility basis set: 1. organic-aerosol mixing thermodynamics. *Atmos. Chem. Phys.* **2011**, *11* (7), 3303–3318.
- (7) Stolzenburg, D.; Fischer, L.; Vogel, A. L.; Heinritzi, M.; Schervish, M.; Simon, M.; Wagner, A. C.; Dada, L.; Ahonen, L. R.; Amorim, A.; Baccarini, A.; Bauer, P. S.; Baumgartner, B.; Bergen, A.; Bianchi, F.; Breitenlechner, M.; Brilke, S.; Buenrostro Mazon, S.; Chen, D.; Dias, A.; Draper, D. C.; Duplissy, J.; El Haddad, I.; Finkenzeller, H.; Frege, C.; Fuchs, C.; Garmash, O.; Gordon, H.; He, X.; Helm, J.; Hofbauer, V.; Hoyle, C. R.; Kim, C.; Kirkby, J.; Kontkanen, J.; Kürten, A.; Lampilahti, J.; Lawler, M.; Lehtipalo, K.; Leimingher, M.; Mai, H.; Mathot, S.; Mentler, B.; Molteni, U.; Nie, W.; Nieminen, T.; Nowak, J. B.; Ojdanic, A.; Onnela, A.; Passananti, M.; Petäjä, T.; Quéléver, L. L. J.; Rissanen, M. P.; Sarnela, N.; Schallhart, S.; Tauber, C.; Tomé, A.; Wagner, R.; Wang, M.; Weitz, L.; Wimmer, D.; Xiao, M.; Yan, C.; Ye, P.; Zha, Q.; Baltensperger, U.; Curtius, J.; Dommen, J.; Flagan, R. C.; Kulmala, M.; Smith, J. N.; Worsnop, D. R.; Hansel, A.; Donahue, N. M.; Winkler, P. M. Rapid growth of organic aerosol nanoparticles over a wide tropospheric temperature range. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115* (37), 9122–9127.
- (8) Pankow, J. F.; Asher, W. E. SIMPOL.1: a simple group contribution method for predicting vapor pressures and enthalpies of vaporization of multifunctional organic compounds. *Atmos. Chem. Phys.* **2008**, *8* (10), 2773–2796.
- (9) Schervish, M.; Donahue, N. M. Peroxy radical chemistry and the volatility basis set. *Atmos. Chem. Phys.* **2020**, *20* (2), 1183–1199.
- (10) Donahue, N. M.; Kroll, J. H.; Pandis, S. N.; Robinson, A. L. A two-dimensional volatility basis set – Part 2: Diagnostics of organic-aerosol evolution. *Atmos. Chem. Phys.* **2012**, *12* (2), 615–634.
- (11) Bianchi, F.; Kurtén, T.; Riva, M.; Mohr, C.; Rissanen, M. P.; Roldin, P.; Berndt, T.; Crounse, J. D.; Wennberg, P. O.; Mentel, T. F.; Wildt, J.; Junninen, H.; Jokinen, T.; Kulmala, M.; Worsnop, D. R.; Thornton, J. A.; Donahue, N.; Kjaergaard, H. G.; Ehn, M. Highly Oxygenated Organic Molecules (HOM) from Gas-Phase Autoxidation Involving Peroxy Radicals: A Key Contributor to Atmospheric Aerosol. *Chem. Rev.* **2019**, *119* (6), 3472–3509.
- (12) Lopez-Hilfiker, F. D.; Mohr, C.; Ehn, M.; Rubach, F.; Kleist, E.; Wildt, J.; Mentel, T. F.; Lutz, A.; Hallquist, M.; Worsnop, D.;

- Thornton, J. A. 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. Technol.* **2014**, *7* (4), 983–1001.
- (13) Lopez-Hilfiker, F. D.; Mohr, C.; D'Ambro, E. L.; Lutz, A.; Riedel, T. P.; Gaston, C. J.; Iyer, S.; Zhang, Z.; Gold, A.; Surratt, J. D.; Lee, B. H.; Kurten, T.; Hu, W. W.; Jimenez, J.; Hallquist, M.; Thornton, J. A. Molecular Composition and Volatility of Organic Aerosol in the Southeastern U.S.: Implications for IEPOX Derived SOA. *Environ. Sci. Technol.* **2016**, *50* (5), 2200–2209.
- (14) Wang, M. Y.; Chen, D. X.; Xiao, M.; Ye, Q.; Stolzenburg, D.; Hofbauer, V.; Ye, P. L.; Vogel, A. L.; Mauldin, R. L.; Amorim, A.; Baccarini, A.; Baumgartner, B.; Brilke, S.; Dada, L.; Dias, A.; Duplissy, J.; Finkenzeller, H.; Garmash, O.; He, X. C.; Hoyle, C. R.; Kim, C.; Kvashnin, A.; Lehtipalo, K.; Fischer, L.; Molteni, U.; Petaja, T.; Pospisilova, V.; Quelever, L. L. J.; Rissanen, M.; Simon, M.; Tauber, C.; Tome, A.; Wagner, A. C.; Weitz, L.; Volkamer, R.; Winkler, P. M.; Kirkby, J.; Worsnop, D. R.; Kulmala, M.; Baltensperger, U.; Dommen, J.; El Haddad, I.; Donahue, N. M. Photo-oxidation of Aromatic Hydrocarbons Produces Low-Volatility Organic Compounds. *Environ. Sci. Technol.* **2020**, *54* (13), 7911–7921.
- (15) Stark, H.; Yatavelli, R. L. N.; Thompson, S. L.; Kang, H.; Krechmer, J. E.; Kimmel, J. R.; Palm, B. B.; Hu, W.; Hayes, P. L.; Day, D. A.; Campuzano-Jost, P.; Canagaratna, M. R.; Jayne, J. T.; Worsnop, D. R.; Jimenez, J. L. Impact of Thermal Decomposition on Thermal Desorption Instruments: Advantage of Thermogram Analysis for Quantifying Volatility Distributions of Organic Species. *Environ. Sci. Technol.* **2017**, *51* (15), 8491–8500.
- (16) Barreira, L. M. F.; Ylisirnio, A.; Pullinen, I.; Buchholz, A.; Li, Z. J.; Lipp, H.; Junninen, H.; Horrak, U.; Noe, S. M.; Krasnova, A.; Krasnov, D.; Kask, K.; Talts, E.; Niinemets, U.; Ruiz-Jimenez, J.; Schobesberger, S. The importance of sesquiterpene oxidation products for secondary organic aerosol formation in a springtime hemiboreal forest. *Atmos. Chem. Phys.* **2021**, *21* (15), 11781–11800.
- (17) Huang, W.; Saathoff, H.; Shen, X. L.; Ramisetty, R.; Leisner, T.; Mohr, C. Seasonal characteristics of organic aerosol chemical composition and volatility in Stuttgart, Germany. *Atmos. Chem. Phys.* **2019**, *19* (18), 11687–11700.
- (18) Gaston, C. J.; Lopez-Hilfiker, F. D.; Whybrey, L. E.; Hadley, O.; McNair, F.; Gao, H.; Jaffe, D. A.; Thornton, J. A. Online molecular characterization of fine particulate matter in Port Angeles, WA: Evidence for a major impact from residential wood smoke. *Atmos. Environ.* **2016**, *138*, 99–107.
- (19) Ye, Q.; Wang, M. Y.; Hofbauer, V.; Stolzenburg, D.; Chen, D. X.; Schervish, M.; Vogel, A.; Mauldin, R. L.; Baalbaki, R.; Brilke, S.; Dada, L.; Dias, A.; Duplissy, J.; El Haddad, I.; Finkenzeller, H.; Fischer, L.; He, X. C.; Kim, C.; Kurten, A.; Lamkaddam, H.; Lee, C. P.; Lehtipalo, K.; Leiminger, M.; Manninen, H. E.; Marten, R.; Mentler, B.; Partoll, E.; Petaja, T.; Rissanen, M.; Schobesberger, S.; Schuchmann, S.; Simon, M.; Tham, Y. J.; Vazquez-Pufleau, M.; Wagner, A. C.; Wang, Y. H.; Wu, Y. S.; Xiao, M.; Baltensperger, U.; Curtius, J.; Flagan, R.; Kirkby, J.; Kulmala, M.; Volkamer, R.; Winkler, P. M.; Worsnop, D.; Donahue, N. M. Molecular Composition and Volatility of Nucleated Particles from alpha-Pinene Oxidation between -50 degrees C and +25 degrees C. *Environ. Sci. Technol.* **2019**, *53* (21), 12357–12365.
- (20) Ylisirnio, A.; Barreira, L. M. F.; Pullinen, I.; Buchholz, A.; Jayne, J.; Krechmer, J. E.; Worsnop, D. R.; Virtanen, A.; Schobesberger, S. On the calibration of FIGAERO-ToF-CIMS: importance and impact of calibrant delivery for the particle-phase calibration. *Atmos. Meas. Technol.* **2021**, *14* (1), 355–367.
- (21) Yan, C.; Nie, W.; Äijälä, M.; Rissanen, M. P.; Canagaratna, M. R.; Massoli, P.; Junninen, H.; Jokinen, T.; Sarnela, N.; Häme, S. A. K.; Schobesberger, S.; Canonaco, F.; Yao, L.; Prévôt, A. S. H.; Petäjä, T.; Kulmala, M.; Sipilä, M.; Worsnop, D. R.; Ehn, M. Source characterization of highly oxidized multifunctional compounds in a boreal forest environment using positive matrix factorization. *Atmos. Chem. Phys.* **2016**, *16* (19), 12715–12731.
- (22) Liu, Y.; Liu, C.; Nie, W.; Li, Y.; Ge, D.; Chen, L.; Zhu, C.; Wang, L.; Zhang, Y.; Liu, T.; et al. Exploring condensable organic vapors and their co-occurrence with PM 2.5 and O₃ in winter in Eastern China. *Environ. Sci.: Atmos.* **2023**, *3* (2), 282–297.
- (23) Qiao, X.; Li, X.; Yan, C.; Sarnela, N.; Yin, R.; Guo, Y.; Yao, L.; Nie, W.; Huang, D.; Wang, Z.; et al. Precursor apportionment of atmospheric oxygenated organic molecules using a machine learning method. *Environ. Sci.: Atmos.* **2023**, *3*, 230–237.
- (24) Franklin, E. B.; Yee, L. D.; Aumont, B.; Weber, R. J.; Grigas, P.; Goldstein, A. H. Ch3MS-RF: a random forest model for chemical characterization and improved quantification of unidentified atmospheric organics detected by chromatography–mass spectrometry techniques. *Atmos. Meas. Technol.* **2022**, *15* (12), 3779–3803.
- (25) Boiko, D. A.; Kozlov, K. S.; Burykina, J. V.; Ilyushenkova, V. V.; Ananikov, V. P. Fully Automated Unconstrained Analysis of High-Resolution Mass Spectrometry Data with Machine Learning. *J. Am. Chem. Soc.* **2022**, *144* (32), 14590–14606.
- (26) Christopoulos, C. D.; Garimella, S.; Zawadowicz, M. A.; Möhler, O.; Cziczo, D. J. A machine learning approach to aerosol classification for single-particle mass spectrometry. *Atmos. Meas. Technol.* **2018**, *11* (10), 5687–5699.
- (27) Bland, G. D.; Battifarano, M.; Pradas del Real, A. E.; Sarret, G.; Lowry, G. V. Distinguishing Engineered TiO₂ Nanomaterials from Natural Ti Nanomaterials in Soil Using spICP-TOFMS and Machine Learning. *Environ. Sci. Technol.* **2022**, *56* (5), 2990–3001.
- (28) Pande, P.; Shrivastava, M.; Shilling, J. E.; Zelenyuk, A.; Zhang, Q.; Chen, Q.; Ng, N. L.; Zhang, Y.; Takeuchi, M.; Nah, T.; Rasool, Q. Z.; Zhang, Y.; Zhao, B.; Liu, Y. Novel Application of Machine Learning Techniques for Rapid Source Apportionment of Aerosol Mass Spectrometer Datasets. *ACS Earth Space Chem.* **2022**, *6* (4), 932–942.
- (29) Wang, F.; Yu, H.; Wang, Z.; Liang, W.; Shi, G.; Gao, J.; Li, M.; Feng, Y. Review of online source apportionment research based on observation for ambient particulate matter. *Sci. Total Environ.* **2021**, *762*, No. 144095.
- (30) Gong, X.; Wex, H.; Müller, T.; Henning, S.; Voigtlander, J.; Wiedensohler, A.; Stratmann, F. Understanding aerosol microphysical properties from 10 years of data collected at Cabo Verde based on an unsupervised machine learning classification. *Atmos. Chem. Phys.* **2022**, *22* (8), 5175–5194.
- (31) Ruiz-Jimenez, J.; Okuljar, M.; Sietiö, O. M.; Demaria, G.; Liangsupree, T.; Zagatti, E.; Aalto, J.; Hartonen, K.; Heinonsalo, J.; Bäck, J.; Petäjä, T.; Riekola, M. L. Determination of free amino acids, saccharides, and selected microbes in biogenic atmospheric aerosols – seasonal variations, particle size distribution, chemical and microbial relations. *Atmos. Chem. Phys.* **2021**, *21* (11), 8775–8790.
- (32) Mahesh, B. Machine learning algorithms-a review. *International Journal of Science and Research (IJSR).* **2020**, *9* (1), 381–386.
- (33) Offenberg, J.; Baker, J. Aerosol size distributions of elemental and organic carbon in urban and over-water atmosphere. *Atmos. Environ.* **2000**, *34*, 1509–1517.
- (34) Chen, S.-J.; Liao, S.-H.; Jian, W.-J.; Lin, C.-C. Particle size distribution of aerosol carbons in ambient air. *Environ. Int.* **1997**, *23* (4), 475–488.
- (35) Liu, F.; Peng, L.; Dai, S.; Bi, X.; Shi, M. Size-Segregated Characteristics of Organic Carbon (OC) and Elemental Carbon (EC) in Marine Aerosol in the Northeastern South China Sea. *Atmosphere* **2023**, *14* (4), 661.
- (36) Chow, J. C.; Watson, J. G.; Crow, D.; Lowenthal, D. H.; Merrifield, T. Comparison of IMPROVE and NIOSH Carbon Measurements. *Aerosol Sci. Technol.* **2001**, *34* (1), 23–34.
- (37) Wu, C.; Yu, J. Z. Determination of primary combustion source organic carbon-to-elemental carbon (OC/EC) ratio using ambient OC and EC measurements: secondary OC-EC correlation minimization method. *Atmos. Chem. Phys.* **2016**, *16* (8), 5453–5465.
- (38) Lee, B. H.; Lopez-Hilfiker, F. D.; Mohr, C.; Kurten, T.; Worsnop, D. R.; Thornton, J. A. An Iodide-Adduct High-Resolution Time-of-Flight Chemical-Ionization Mass Spectrometer: Application

- to Atmospheric Inorganic and Organic Compounds. *Environ. Sci. Technol.* **2014**, *48* (11), 6309–6317.
- (39) Cai, J.; Daellenbach, K. R.; Wu, C.; Zheng, Y.; Zheng, F.; Du, W.; Haslett, S. L.; Chen, Q.; Kulmala, M.; Mohr, C. Characterization of offline analysis of particulate matter with FIGAERO-CIMS. *Atmos. Meas. Technol.* **2023**, *16* (5), 1147–1165.
- (40) Cai, J.; Wu, C.; Wang, J.; Du, W.; Zheng, F.; Hakala, S.; Fan, X.; Chu, B.; Yao, L.; Feng, Z.; Liu, Y.; Sun, Y.; Zheng, J.; Yan, C.; Bianchi, F.; Kulmala, M.; Mohr, C.; Daellenbach, K. R. Influence of organic aerosol molecular composition on particle absorptive properties in autumn Beijing. *Atmos. Chem. Phys.* **2022**, *22* (2), 1251–1269.
- (41) Siegel, K.; Karlsson, L.; Zieger, P.; Baccarini, A.; Schmale, J.; Lawler, M.; Salter, M.; Leck, C.; Ekman, A. M.; Riipinen, I.; et al. Insights into the molecular composition of semi-volatile aerosols in the summertime central Arctic Ocean using FIGAERO-CIMS. *Environmental science: atmospheres* **2021**, *1* (4), 161–175.
- (42) Stark, H.; Yatavelli, R. L. N.; Thompson, S. L.; Kimmel, J. R.; Cubison, M. J.; Chhabra, P. S.; Canagaratna, M. R.; Jayne, J. T.; Worsnop, D. R.; Jimenez, J. L. Methods to extract molecular and bulk chemical information from series of complex mass spectra with limited mass resolution. *Int. J. Mass Spectrom.* **2015**, *389*, 26–38.
- (43) Kind, T.; Fiehn, O. Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinformatics* **2007**, *8* (1), 105.
- (44) Lee, B.; Lopez-Hilfiker, F. D.; D'Ambro, E. L.; Zhou, P. T.; Boy, M.; Petaja, T.; Hao, L. Q.; Virtanen, A.; Thornton, J. A. Semi-volatile and highly oxygenated gaseous and particulate organic compounds observed above a boreal forest canopy. *Atmos. Chem. Phys.* **2018**, *18* (15), 11547–11562.
- (45) Kind, T.; Fiehn, O. Metabolomic database annotations via query of elemental compositions: Mass accuracy is insufficient even at less than 1 ppm. *BMC Bioinformatics* **2006**, *7* (1), 234.
- (46) Iyer, S.; Lopez-Hilfiker, F.; Lee, B. H.; Thornton, J. A.; Kurten, T. Modeling the Detection of Organic and Inorganic Compounds Using Iodide-Based Chemical Ionization. *J. Phys. Chem. A* **2016**, *120* (4), 576–587.
- (47) Lopez-Hilfiker, F. D.; Iyer, S.; Mohr, C.; Lee, B. H.; D'Ambro, E. L.; Kurten, T.; Thornton, J. A. Constraining the sensitivity of iodide adduct chemical ionization mass spectrometry to multifunctional organic molecules using the collision limit and thermodynamic stability of iodide ion adducts. *Atmos. Meas. Technol.* **2016**, *9* (4), 1505–1512.
- (48) Isaacman-VanWertz, G.; Massoli, P.; O'Brien, R.; Lim, C.; Franklin, J. P.; Moss, J. A.; Hunter, J. F.; Nowak, J. B.; Canagaratna, M. R.; Misztal, P. K.; Arata, C.; Roscioli, J. R.; Herndon, S. T.; Onasch, T. B.; Lambe, A. T.; Jayne, J. T.; Su, L.; Knopf, D. A.; Goldstein, A. H.; Worsnop, D. R.; Kroll, J. H. Chemical evolution of atmospheric organic carbon over multiple generations of oxidation. *Nat. Chem.* **2018**, *10* (4), 462–468.
- (49) Wang, D. S.; Hildebrandt Ruiz, L. Chlorine-initiated oxidation of n-alkanes under high-NO_x conditions: insights into secondary organic aerosol composition and volatility using a FIGAERO-CIMS. *Atmos. Chem. Phys.* **2018**, *18* (21), 15535–15553.
- (50) Bannan, T. J.; Le Breton, M.; Priestley, M.; Worrall, S. D.; Bacak, A.; Marsden, N. A.; Mehra, A.; Hammes, J.; Hallquist, M.; Alfara, M. R.; Krieger, U. K.; Reid, J. P.; Jayne, J.; Robinson, W.; McFiggans, G.; Coe, H.; Percival, C. J.; Topping, D. A method for extracting calibrated volatility information from the FIGAERO-HR-ToF-CIMS and its experimental application. *Atmos. Meas. Technol.* **2019**, *12* (3), 1429–1439.
- (51) Fraser, M. P.; Cass, G. R.; Simoneit, B. R. T.; Rasmussen, R. A. Air Quality Model Evaluation Data for Organics. 4. C₂–C₃₆ Non-Aromatic Hydrocarbons. *Environ. Sci. Technol.* **1997**, *31* (8), 2356–2367.
- (52) Zhao, Y.; Hennigan, C. J.; May, A. A.; Tkacik, D. S.; de Gouw, J. A.; Gilman, J. B.; Kuster, W. C.; Borbon, A.; Robinson, A. L. Intermediate-Volatility Organic Compounds: A Large Source of Secondary Organic Aerosol. *Environ. Sci. Technol.* **2014**, *48* (23), 13743–13750.
- (53) Zhao, Y.; Nguyen, N. T.; Presto, A. A.; Hennigan, C. J.; May, A. A.; Robinson, A. L. Intermediate Volatility Organic Compound Emissions from On-Road Diesel Vehicles: Chemical Composition, Emission Factors, and Estimated Secondary Organic Aerosol Production. *Environ. Sci. Technol.* **2015**, *49* (19), 11516–11526.
- (54) Lu, Q.; Zhao, Y.; Robinson, A. L. Comprehensive organic emission profiles for gasoline, diesel, and gas-turbine engines including intermediate and semi-volatile organic compound emissions. *Atmos. Chem. Phys.* **2018**, *18* (23), 17637–17654.
- (55) Nie, W.; Yan, C.; Huang, D. D.; Wang, Z.; Liu, Y.; Qiao, X.; Guo, Y.; Tian, L.; Zheng, P.; Xu, Z.; et al. Secondary organic aerosol formed by condensing anthropogenic vapours over China's megacities. *Nat. Geosci.* **2022**, *15* (4), 255–261.
- (56) 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.* **2018**, *18* (8), 5467–5481.
- (57) Ma, J.; Ungeheuer, F.; Zheng, F.; Du, W.; Wang, Y.; Cai, J.; Zhou, Y.; Yan, C.; Liu, Y.; Kulmala, M.; Daellenbach, K. R.; Vogel, A. L. Nontarget Screening Exhibits a Seasonal Cycle of PM_{2.5} Organic Aerosol Composition in Beijing. *Environ. Sci. Technol.* **2022**, *56* (11), 7017–7028.
- (58) Thompson, S. L.; Yatavelli, R. L. N.; Stark, H.; Kimmel, J. R.; Krechmer, J. E.; Day, D. A.; Hu, W.; Isaacman-VanWertz, G.; Yee, L.; Goldstein, A. H.; Khan, M. A. H.; Holzinger, R.; Kreisberg, N.; Lopez-Hilfiker, F. D.; Mohr, C.; Thornton, J. A.; Jayne, J. T.; Canagaratna, M.; Worsnop, D. R.; Jimenez, J. L. Field intercomparison of the gas/particle partitioning of oxygenated organics during the Southern Oxidant and Aerosol Study (SOAS) in 2013. *Aerosol Sci. Technol.* **2017**, *51* (1), 30–56.
- (59) Schmitt-Kopplin, P.; Liger-Belair, G.; Koch, B. P.; Flerus, R.; Kattner, G.; Harir, M.; Kanawati, B.; Lucio, M.; Tziotis, D.; Hertkorn, N.; Gebefugi, I. Dissolved organic matter in sea spray: a transfer study from marine surface water to aerosols. *Biogeosciences* **2012**, *9* (4), 1571–1582.
- (60) Rasmussen, B. B.; Nguyen, Q. T.; Kristensen, K.; Nielsen, L. S.; Bilde, M. What controls volatility of sea spray aerosol? Results from laboratory studies using artificial and real seawater samples. *J. Aerosol Sci.* **2017**, *107*, 134–141.
- (61) Xu, W.; Ovadnevaite, J.; Fossum, K. N.; Lin, C.; Huang, R. J.; O'Dowd, C.; Ceburnis, D. Aerosol hygroscopicity and its link to chemical composition in the coastal atmosphere of Mace Head: marine and continental air masses. *Atmos. Chem. Phys.* **2020**, *20* (6), 3777–3791.
- (62) Donahue, N. M.; Robinson, A. L.; Stanier, C. O.; Pandis, S. N. Coupled Partitioning, Dilution, and Chemical Aging of Semivolatile Organics. *Environ. Sci. Technol.* **2006**, *40* (8), 2635–2643.
- (63) Hallquist, M.; Wenger, J. C.; Baltensperger, U.; Rudich, Y.; Simpson, D.; Claeys, M.; Dommen, J.; Donahue, N. M.; George, C.; Goldstein, A. H.; Hamilton, J. F.; Herrmann, H.; Hoffmann, T.; Iinuma, Y.; Jang, M.; Jenkin, M. E.; Jimenez, J. L.; Kiendler-Scharr, A.; Maenhaut, W.; McFiggans, G.; Mentel, T. F.; Monod, A.; Prévôt, A. S. H.; Seinfeld, J. H.; Surratt, J. D.; Szmigielski, R.; Wildt, J. The formation, properties and impact of secondary organic aerosol: current and emerging issues. *Atmos. Chem. Phys.* **2009**, *9* (14), 5155–5236.
- (64) Zhang, J.; Huff Hartz, K. E.; Pandis, S. N.; Donahue, N. M. Secondary Organic Aerosol Formation from Limonene Ozonolysis: Homogeneous and Heterogeneous Influences as a Function of NO_x. *J. Phys. Chem. A* **2006**, *110* (38), 11053–11063.
- (65) Xu, W.; Chen, C.; Qiu, Y.; Li, Y.; Zhang, Z.; Karnezi, E.; Pandis, S. N.; Xie, C.; Li, Z.; Sun, J.; Ma, N.; Xu, W.; Fu, P.; Wang, Z.; Zhu, J.; Worsnop, D. R.; Ng, N. L.; Sun, Y. Organic aerosol volatility and viscosity in the North China Plain: contrast between summer and winter. *Atmos. Chem. Phys.* **2021**, *21* (7), 5463–5476.
- (66) Wang, K.; Huang, R. J.; Brüggemann, M.; Zhang, Y.; Yang, L.; Ni, H.; Guo, J.; Wang, M.; Han, J.; Bilde, M.; Glasius, M.; Hoffmann, T. Urban organic aerosol composition in eastern China differs from

north to south: molecular insight from a liquid chromatography–mass spectrometry (Orbitrap) study. *Atmos. Chem. Phys.* **2021**, *21* (11), 9089–9104.

(67) Tao, S.; Lu, X.; Levac, N.; Bateman, A. P.; Nguyen, T. B.; Bones, D. L.; Nizkorodov, S. A.; Laskin, J.; Laskin, A.; Yang, X. Molecular Characterization of Organosulfates in Organic Aerosols from Shanghai and Los Angeles Urban Areas by Nanospray-Desorption Electrospray Ionization High-Resolution Mass Spectrometry. *Environ. Sci. Technol.* **2014**, *48* (18), 10993–11001.

(68) Wang, X.; Hayeck, N.; Brüggemann, M.; Yao, L.; Chen, H.; Zhang, C.; Emmelin, C.; Chen, J.; George, C.; 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.* **2017**, *122* (21), 11703–11722.