

Supplement of

Seasonal regime shifts in atmospheric aging of vehicular water-soluble organic aerosols: evidence from functionalization transformation indices and molecular conformational modeling

Tantan Tan et al.

Correspondence to: Yanpeng Li (liyanp01@chd.edu.cn) and Jianxun Chen (chenjx1969@chd.edu.cn)

This PDF file includes:

Text S1 to S6

Table S1 to S7

Figure S1 to S14

List of Supplementary Text:

Supplementary Text S1. Representativeness of the sampling season.

Supplementary Text S2. Characteristics of the studied tunnels and detailed sampling methodology.

Supplementary Text S3. Quantitative assessment of potential functional group transformations between daytime and nighttime aerosol samples.

Supplementary Text S4. Molecular Modeling of Vehicular WSOA representative groups.

Supplementary Text S5. Description of the Vienna Soil Organic Matter Modeler 2 (VSOMM2).

Supplementary Text S6. Post-generation structural refinement and molecular descriptor calculation of WSOA models using Schrödinger suite.

List of tables:

Supplementary Table S1. The length, ventilation conditions, number of vehicles passing through the tunnel per hour, and the number of samples collected in each tunnel.

Supplementary Table S2. Weather conditions during the sampling period.

Supplementary Table S3. Input parameters of VSOMM2 modeling for summer samples.

Supplementary Table S4. Input parameters of VSOMM2 modeling for winter samples.

Supplementary Table S5. Distribution of the number of N among different chemical categories of core molecules in summer samples.

Supplementary Table S6. Distribution of the number of N among different chemical categories of core molecules in winter samples.

Supplementary Table S7. The number of molecules transformed through hydroxylation, carboxylation, and carbonylation from day to night across different seasons, as determined by their molecular formulas.

List of Figures:

Supplementary Fig. S1. Seasonal comparison of intensity-weighted DBE and Almod values.

Supplementary Fig. S2. Seasonal variations in the relationship between DBE and H/C.

Supplementary Fig. S3. Mass and corresponding relative intensity (RI) distribution of (a) Carb, (b) UnsathC, (c) aromatic, and (d) HOCs.

Supplementary Fig. S4. Seasonal variation in the proportion of core molecules.

Supplementary Fig. S5. The van Krevelen plots for core molecular communities of WSOA in the summer and winter. In comparison, the summer core molecular community has 5179 formulae, whereas the winter core molecular community has 3667 formulae.

Supplementary Fig. S6. Seasonal distribution of various compound classes and N-containing subgroups.

Supplementary Fig. S7. DBE/C and Almod/C distribution of UnsathC, aromatic and HOCs molecules of the core molecular community for summer and winter samples.

Supplementary Fig. S8. 3D visualizations of the 12 Model(WSOA) across seasons using VSOMM2 and Schrödinger, in which the edge length of the cubic box is 52 Å. 1, 2, and 3 denote molecular models that are positively correlated, uncorrelated, and negatively correlated with b₄₀₅, respectively. Green represents C atoms. White represents H atoms. Red represents O atoms. Blue represents N atoms. Yellow represents S atoms.

Supplementary Fig. S9. Model(WSOA) fits of the relative frequency distributions of molecular weight (MW) for different season samples. The RMSEs of the Model(WSOA) calculations range from 0.119 to 0.271. Red lines are FT-ICR-MS data and blue symbols with dashed lines are model calculations.

Supplementary Fig. S10. Model(WSOA) fits of the relative frequency distributions of DBE for different season samples. The RMSEs of the Model(WSOA) calculations range from 0.138 to 0.287. Red lines are FT-ICR-MS data and blue symbols with dashed lines are model calculations.

Supplementary Fig. S11. Model(WSOA) fits of the relative frequency distributions of O number for different season samples. The RMSEs of the Model(WSOA) calculations range from 0.083 to 0.245. Red lines are FT-ICR-MS data and blue symbols with dashed lines are model calculations.

Supplementary Fig. S12. Model(WSOA) fits of the relative frequency distributions of Almod for different season samples. The RMSEs of the Model(WSOA) calculations range from 0.137 to 0.234. Red lines are FT-ICR-MS data and blue symbols with dashed lines are model calculations.

Supplementary Fig. S13. Heatmap of molecular descriptors for summer and winter samples.

Supplementary Fig. S14. Seasonal and diurnal variations in molecular descriptor distributions.

Text S1. Representativeness of the Sampling Season

This study specifically contrasts summer and winter because these two seasons exhibit the most pronounced differences in photochemical intensity, ambient temperature, and relative humidity in northwestern China, thereby maximizing the signal separation between distinct aging pathways. In both cities, summer (May–July during sampling) is characterized by strong solar radiation (Average daytime irradiance $>300 \text{ W/m}^2$), Average temperature $>20 \text{ }^\circ\text{C}$, and elevated O_3 levels (Average = $98 \text{ }\mu\text{g/m}^3$), promoting photochemical oxidation as the dominant aging route. Winter (December), in contrast, features low solar radiation (Average daytime irradiance $< 240 \text{ W/m}^2$), low temperatures (Average $< 7 \text{ }^\circ\text{C}$), and reduced O_3 concentrations (Average = $31 \text{ }\mu\text{g/m}^3$), under which heterogeneous reactions and dark oxidation become relatively more important. Spring and autumn are transitional periods with rapidly fluctuating conditions that blur the boundary between aging mechanisms; therefore, they were excluded to maintain clear seasonal contrast.

To confirm the representativeness of the sampling periods, we compared meteorological parameters during the campaigns with local climatological normals (2015-2024). For example, the mean temperature during the Xi'an summer campaign ($22.9 \text{ }^\circ\text{C}$) was within $\pm 1.5 \text{ }^\circ\text{C}$ of the long-term June average, and the winter campaign mean ($4.7 \text{ }^\circ\text{C}$) similarly matched the December normal. Comparable consistency was observed for Yinchuan. These data demonstrate that the selected windows faithfully capture typical summer and winter conditions, enabling robust attribution of molecular evolution differences to contrasting aging environments.

All sampling durations were identical (7-8 days in each tunnel during each season) to ensure statistical comparability. Traffic flow rates and vehicle composition were monitored and showed no significant inter-seasonal variations within each tunnel, confirming that observed chemical differences primarily reflect environmental rather than emission changes.

Text S2. Characteristics of the studied tunnels and detailed sampling methodology

Field sampling was conducted in two urban tunnels in Northwestern China: the Jinhua Tunnel and the Yuehaihu Tunnel. The Jinhua Tunnel, with a length of approximately 500 m, is classified as a short tunnel and utilizes natural ventilation. It employs a separated main-and-service road design: the main section functions as a bidirectional six-lane arterial road, while the service road accommodates mixed motorized and non-motorized traffic. As a vital transportation artery linking the northern and southern parts of Xi'an, the tunnel experiences a high average daily traffic volume. Vehicles often undergo prolonged idling, particularly during morning and evening peak hours, leading to frequent traffic congestion. These characteristics establish it as a typical high-density urban mixed-use tunnel. The Yuehaihu Tunnel represents the first sub-aqueous tunnel in Ningxia traversing a water system, serving as a critical underground transportation hub in northern Yinchuan and a key component of the Shenyang Road Rapid Transit Corridor. Stretching approximately 1.8 km from Mancheng Avenue in the west to Qinshui Avenue in the east, the tunnel is configured as an eight-lane bidirectional urban arterial roadway. Although equipped with mechanical ventilation systems, it primarily relies on natural ventilation. Furthermore, the tunnel accommodates heavy-duty diesel trucks, thereby making it representative of real-world fleet emission characteristics in developing cities.

To capture both seasonal and diurnal variability, sampling campaigns were carried out in the summer (Jinhua: May 29 to Jun. 5, 2025; Yuehaihu: Jul. 7 to 14, 2025) and winter (Jinhua: Dec. 2 to 12, 2025; Yuehaihu: Dec. 16 to 23, 2025) of 2025. Within each tunnel, three sampling sites were established along the traffic direction: an entrance site (~50 m inward from the portal), an internal site (midpoint), and an exit site (~300 m inward from the exit portal), to investigate the spatial evolution of emissions. PM_{2.5} samples were collected using medium-volume aerosol samplers (ZR-3930) equipped with pre-baked (500 °C for 6 h) quartz microfiber filters at a flow rate of 100 L min⁻¹ and a sampling height of 1.2 m. Sampling was performed over consecutive 12-hour periods for daytime (08:30–20:30) and nighttime (20:30–08:30)

to resolve diurnal differences. After collection, all filter samples were wrapped in pre-baked aluminum foil, sealed in polyethylene bags, and stored at -20°C in the dark until laboratory analysis. Field blank filters were deployed at each site during both daytime and nighttime to correct for background contamination.

Text S3. Quantitative Assessment of Potential Functional Group Transformations Between Daytime and Nighttime Aerosol Samples

To elucidate the atmospheric aging processes occurring within the tunnel environment, we developed a computational framework to quantify potential molecular transformations between daytime and nighttime $\text{PM}_{2.5}$ samples analyzed by Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS). The approach was adapted from previously established methods for quantifying biochemical reaction networks and further modified to represent photochemical and dark oxidation processes relevant to urban aerosols.

The core assumption underlying this method is that certain molecules observed in nighttime samples may originate from precursor molecules present in daytime samples through oxidative functionalization reactions during transit or accumulation within the tunnel. Accordingly, for each unique molecular formula identified in the daytime dataset (denoted as D), we systematically screened all molecular formulas detected in the nighttime dataset (denoted as N) to identify potential transformation products based on net changes in elemental composition. Three canonical oxidation pathways were considered, defined by their stoichiometric constraints relative to the precursor molecule:

Hydroxylation: $\text{C}_c\text{H}_h\text{O}_o\text{N}_n\text{S}_s + x[\text{O}] \rightarrow \text{C}_c\text{H}_h\text{O}_{o+x}\text{N}_n\text{S}_s$ ($\Delta\text{C} = 0, \Delta\text{H} = 0, \Delta\text{O} = x$)

Carboxylation: $\text{C}_c\text{H}_h\text{O}_o\text{N}_n\text{S}_s + x[\text{O}] \rightarrow \text{C}_c\text{H}_{h-2x}\text{O}_{o+2x}\text{N}_n\text{S}_s$ ($\Delta\text{C} = 0, \Delta\text{H} = -2x, \Delta\text{O} = 2x$)

Carbonylation: $\text{C}_c\text{H}_h\text{O}_o\text{N}_n\text{S}_s + x[\text{O}] \rightarrow \text{C}_c\text{H}_{h-2x}\text{O}_{o+x}\text{N}_n\text{S}_s$ ($\Delta\text{C} = 0, \Delta\text{H} = -2x, \Delta\text{O} = +x$)

These pathways correspond to common atmospheric functionalization reactions (e.g., hydroxyl radical addition, carboxylic acid formation, ketone/aldehyde formation) and have been validated in prior studies on tunnel aerosols. For each daytime

molecule, the algorithm searched the entire nighttime molecular inventory for any molecule whose atomic differences exactly matched one of the three patterns above. When multiple nighttime candidates satisfied the same pattern for a given daytime molecule, all matches were retained to reflect the possibility of parallel transformation routes.

To reduce false positives arising from coincidental compositional similarity, an additional constraint was applied: the exact mass difference between the candidate pair must fall within ± 0.001 Da of the theoretical mass change predicted by the corresponding reaction. Only pairs meeting both the elemental and mass criteria were accepted as valid transformation events. The process was implemented in Python, iterating over all possible D–N combinations, and the output was a comprehensive table listing each matched pair along with the number of oxygen atoms gained (x) and the specific reaction type.

From these pairwise results, we derived three dimensionless indices, including Hydroxylation Index (HI), Carboxylation Index (CXI), and Carbonylation Index (CNI), to characterize the overall prevalence of each transformation pathway. Each index was calculated as the sum over all matched pairs. These indices provide a potential quantitative measure of the relative importance of different functionalization mechanisms during the daytime-to-nighttime transition, enabling comparison across seasons.

Text S4. Identification of WSOA representative molecules

To clarify the potential link between molecular composition and light absorption properties, Spearman correlation analysis was conducted between the relative intensity of individual molecular formulas obtained from FT-ICR MS and the b_{405} for all samples. As a non-parametric statistical approach, Spearman correlation is well-suited for characterizing monotonic associations in complex environmental organic datasets with non-normal and nonlinear distribution characteristics. Referring to previously reported classification criteria, all identified molecules were divided into three subgroups based on correlation coefficient and statistical significance.

Molecules with a significantly positive correlation ($r > 0$, $p \leq 0.05$) were classified as b(+). Those with a significantly negative correlation ($r < 0$, $p \leq 0.05$) were classified as b(-). Molecules showing no significant correlation ($p > 0.05$) were grouped as b(0). To reduce analytical noise and improve result stability, only molecules ubiquitously detected across all sampling batches (core molecules) were included in the final correlation calculation. It should be noted that this screening strategy may exclude sample-specific unique compounds, which could partially ignore their effects.

Text S5. Description of the Vienna Soil Organic Matter Modeler 2 (VSOMM2)

Atomistic condensed-phase models of Water-Soluble Organic Aerosols (WSOA) were generated using the Vienna Soil Organic Matter Modeler 2 (VSOMM2; <https://somm.boku.ac.at/>). The modeler was originally developed to construct humic substance models for soil dissolved organic matter (DOM). Its application to WSOA is motivated by the recognized chemical similarities between atmospheric water-soluble organic carbon and DOM. They both often share a broad functional group diversity and a complex, polydisperse mixture of compounds that cannot be represented by a single molecular structure. These common features make VSOMM2 a suitable framework for building representative WSOA models, while acknowledging that domain-specific adjustments may be required to capture the distinct formation and aging pathways of atmospheric organics.

The modeler constructs humic substances by assembling a user-defined number of molecular fragments, termed building blocks, into full DOM molecules. The building block library was extended in VSOMM2 to incorporate structural motifs derived from mass spectrometry studies (e.g., lignin-, lipid-, and peptide-like moieties) in addition to those based on previously proposed theoretical model compounds, thereby increasing the chemical diversity of the resulting models. The selection of building blocks is driven by a genetic algorithm (PyEvolve, 1000 generations, population size 100, mutation rate 2%, crossover rate 80%) that minimizes a fitness function representing the deviation between the experimental and modelled elemental (C, N mass fractions) and organic (^{13}C NMR-derived carbon fractions: carbonyl,

carboxyl, aromatic, acetal, heteroaliphatic, aliphatic) compositions. VSOMM2 prioritizes the satisfaction of the organic composition over the elemental composition, as the former more accurately reflects the chemistry of DOM. For each generated molecule, three-dimensional coordinates are produced by merging building-block PDB structures, followed by sequential energy minimisation (steepest descent, 1000 steps, electrostatics off to on, bond-length constraints) and a brief 2 ps equilibration at 300 K with a 2 fs integration timestep. The assembled molecules are placed in a periodic cubic box, solvated using the SPC water model, and neutralised by replacing water molecules experiencing the largest negative electrostatic potential with Na⁺ or Ca²⁺ ions. Protonation states of carboxyl groups are assigned according to a user-specified pKa (default 4.2, corresponding to benzoic acid). All intermolecular interactions are described using the united-atom GROMOS 54A7 force field,² and the output GROMOS coordinate and topology files, together with GROMACS-compatible input files, are provided for subsequent molecular dynamics simulations.

Text S6. Post-Generation Structural Refinement and Molecular Descriptor Calculation of WSOA Models Using Schrödinger Suite

All WSOA model structures generated as PDB coordinate files by VSOMM2 were further processed and analysed within the Schrödinger molecular modelling environment (Maestro 13.7, Schrödinger). The purpose of this post-generation workflow was to obtain a chemically consistent and energetically relaxed 3D geometry suitable for reliable descriptor calculation, and to derive quantitative molecular conformational and physicochemical parameters that enable comparative assessment of WSOA models across different sample categories.

Each PDB file was imported into a Maestro project. The structures were then treated using the Protein Preparation Wizard with the following optimised settings. In the Import and Process tab, bond orders were assigned using the CCD (Chemical Component Dictionary) database, hydrogen atoms were added while removing all original hydrogens, and disulfide bonds were created where applicable. Waters beyond 5.00 Å from hetero groups were deleted. In the H-bond

assignment section, sample water orientations were enabled, and the PROPKA method was used to assign protonation states at pH 7.0 (with pKa labelling active). The Minimize option was selected, applying restrained minimisation with the OPLS4 force field, heavy atoms constrained to an RMSD threshold of 0.30 Å (i.e., 0.030 nm), and hydrogens allowed to relax. Water molecules beyond 3.0 Å from hetero groups were additionally removed before minimisation. This protocol assigns chemically correct atom types, adds missing hydrogen atoms, and optimises the hydrogen-bonding network while preserving the global molecular architecture imposed by the VSOMM2 building-block assembly. All minimisations were performed in the gas phase, and the refined coordinates were saved for all subsequent quantitative analyses. The remaining descriptor calculation procedures (PSA, HBD/HBA, AlogP, rotatable bonds, and residue count) were executed as originally described, using the minimised structures as input.

Table S1. The length, ventilation conditions, number of vehicles passing through the tunnel per hour, and the number of samples collected in each tunnel

Season	Tunnels	Length (km)	Ventilation conditions	Vehicles/h	Vehicle type and composition	Speed (km/h)	Number of samples
Summer	Jinhua	~0.5	Nature	1711 ± 563	Gasoline, new energy and light trucks (55%/42%/3%)	30-40	6+ 2 blanks
	Yuehaihu	~1.8	Nature	1142 ± 715	Gasoline, new energy and heavy-duty diesel vehicles (65%/24%/11%)	60-80	6+ 2 blanks
Winter	Jinhua	~0.5	Nature	1328 ± 621	Gasoline, new energy and light trucks (52%/46%/2%)	30-40	6+ 2 blanks
	Yuehaihu	~1.8	Nature	867 ± 757	Gasoline, new energy and heavy-duty diesel vehicles (61%/27%/12%)	60-80	6+ 2 blanks

Table S2. Weather conditions during the sampling period

Summer		Winter		
	Jinhua Tunnel (May 29 – Jun. 5)	Yuehaihu Tunnel (Jul. 7 – 14)	Jinhua Tunnel (Dec. 2 – 12)	Yuehaihu Tunnel (Dec. 16 – 23)
T (°C)	18.9–34.7	16.7–23.4	1.1–8.5	-5.6–5.2
	17.4–26.6	18.5–28.7	-2.4–7.5	-7.8–5.4
	16.8–25.3	20.7–31.8	1–10.2	-7.1–8
	18.1–21	13.7–28.3	-2.7–8.3	-4.7–3
	15.8–25.2	17.7–33.3	-1.1–11.5	-6–1
	16.3–29	18.1– 31.7	3.7–13.7	-9.9–3.2
	15.3–32.5	20–36.5	1.1–8.5	-8–4.3
	18.5–35.2	22.9–34.7	-0.3–8.7	-6.3–0.8
		23.5–35	1.7–13.5	
			0.2–14.3	
RH(%)	20–58	46–94	24–79	23–65
	33–98	30–72	35–75	32–74
	51–93	33–84	33–72	32–80
	76–95	37–98	44–93	40–72
	57–97	23–80	43–92	22–55
	35–81	24–91	27–71	37–75
	23–69	22–71	49–82	41–87
	23–74	32–73	47–90	30–66
		27–71	45–91	
			42–97	
		46–98		

Table S3. Input parameters of VSOMM2 modeling for summer samples

Models	Model D(+)	Model D(0)	Model D(-)	Model N(+)	Model N(0)	Model N(-)
Molecules	110	110	110	110	110	110
Building blocks	330	330	330	330	330	330
Building blocks per molecule	3	3	3	3	3	3
Water molecules	200	200	200	200	200	200
pH	6.8	7.1	6.8	6.7	6.5	6.5
Counter ion	Na ⁺	Na ⁺	Na ⁺	Na ⁺	Na ⁺	Na ⁺
Number of ions	198	207	206	190	148	165
C fraction	0.616	0.585	0.511	0.637	0.561	0.528
N fraction	0.003	0.008	0.029	0.002	0.010	0.015
NMR assignment	IHSS	IHSS	IHSS	IHSS	IHSS	IHSS
Carbonyl fraction	0.005	0.005	0.005	0.024	0.024	0.024
Carboxyl fraction	0.020	0.020	0.020	0.033	0.023	0.028
Aromatic fraction	0.055	0.050	0.050	0.073	0.070	0.073
Acetal fraction	0.067	0.060	0.060	0.025	0.025	0.025
Heteroaliphatic fraction	0.120	0.120	0.100	0.083	0.083	0.083
Aliphatic fraction	0.733	0.745	0.765	0.762	0.775	0.727
Forcefield	54a7	54a7	54a7	54a7	54a7	54a7
Random seed	111111	111111	111111	111111	111111	111111
Generate Coarse-Grained model of SOM	No	No	No	No	No	No

Table S4. Input parameters of VSOMM2 modeling for winter samples

Models	Model D(+)	Model D(0)	Model D(-)	Model N(+)	Model N(0)	Model N(-)
Molecules	110	110	110	110	110	110
Building blocks	330	330	330	330	330	330
Building blocks per molecule	3	3	3	3	3	3
Water molecules	200	200	200	200	200	200
pH	6.0	6.2	6.3	6.7	6.7	6.1
Counter ion	Na ⁺	Na ⁺	Na ⁺	Na ⁺	Na ⁺	Na ⁺
Number of ions	111	151	140	241	241	116
C fraction	0.540	0.570	0.610	0.546	0.581	0.599
N fraction	0.016	0.021	0.013	0.018	0.021	0.017
NMR assignment	IHSS	IHSS	IHSS	IHSS	IHSS	IHSS
Carbonyl fraction	0.005	0.014	0.025	0.006	0.006	0.005
Carboxyl fraction	0.022	0.041	0.030	0.058	0.058	0.022
Aromatic fraction	0.388	0.308	0.150	0.206	0.206	0.310
Acetal fraction	0.107	0.162	0.070	0.134	0.134	0.105
Heteroaliphatic fraction	0.122	0.012	0.135	0.018	0.018	0.033
Aliphatic fraction	0.356	0.463	0.590	0.578	0.578	0.525
Forcefield	54a7	54a7	54a7	54a7	54a7	54a7
Random seed	111111	111111	111111	111111	111111	111111
Generate Coarse-Grained model of SOM	No	No	No	No	No	No

Table S5. Distribution of the Number of N Among Different Chemical Categories of Core Molecules in Summer Samples

Category	Number of N atom				Sum
	0	1	2	3	
aliphatic/peptides	750 (14.48%)	457 (8.82%)	4 (0.08%)	0	1211 (23.28%)
aromatic	70 (1.35%)	24 (0.46%)	0	0	94 (1.82%)
carbohydrates	74 (1.43%)	113 (2.18%)	0	0	187 (3.61%)
CRAMs	1655 (31.96%)	1154 (22.28%)	164 (3.17%)	0	2973 (57.40%)
HOCs	4 (0.08%)	6 (0.12%)	1 (0.02%)	0	11 (0.21%)
lipids	345 (6.66%)	54 (1.04%)	1 (0.02%)	0	400 (7.72%)
unclassified	20 (0.39%)	34 (0.66%)	1 (0.02%)	0	55 (1.06%)
unsaturated hydrocarbons	143 (2.76%)	99 (1.91%)	5 (0.10%)	1 (0.02%)	248 (4.79%)
Sum	3061 (59.1%)	1941 (37.48%)	176 (3.40%)	1 (0.02%)	5179 (100%)

Table S6. Distribution of the Number of N Among Different Chemical Categories of Core Molecules in Winter Samples

Category	Number of N				Sum
	0	1	2	3	
aliphatic/peptides	55 (1.50%)	39 (1.06%)	0	0	94 (2.56%)
aromatic	156 (4.26%)	104 (2.84%)	28 (0.76%)	0	288 (7.86%)
CRAMs	1458 (39.77%)	1182 (32.24%)	470 (12.82%)	24 (0.65%)	3134 (85.49%)
HOCs	1 (0.03%)	5 (0.14%)	0	0	6 (0.16%)
lipids	47 (1.28%)	3 (0.08%)	0	0	50 (1.36%)
unclassified	0	2 (0.05%)	0	0	2 (0.05%)
unsaturated hydrocarbons	11 (0.30%)	73 (1.99%)	8 (0.22%)	0	92 (2.51%)
Sum	1728 (47.14%)	1408 (38.41%)	176 (13.80%)	1 (0.65%)	3666 (100%)

Table S7. The number of molecules transformed through hydroxylation, carboxylation, and carbonylation from day to night across different seasons, as determined by their molecular formulas.

Category		Summer			Winter		
Day	Night	Hydroxylation	Carboxylation	Carbonylation	Hydroxylation	Carboxylation	Carbonylation
CRAMs	CRAMs	81399	29666	48939	83557	31889	59003
	Aliphatic/ peptides	-	1827	7072	-	1625	6473
	UnsatHC	12198	2644	2472	10981	2997	2780
	Aromatics	-	699	2017	-	1570	5445
	Lipids	-	3104	9060	-	1646	5165
	Carbs	-	-	-	-	-	-
	Others	294	223	467	256	415	834
	HOCs	303	90	102	402	237	275
Aliphatic/ peptides	CRAMs	-	1833	7119	-	1601	6244
	Aliphatic/ peptides	15205	4742	8316	14469	4889	8927
	UnsatHC	-	1192	3076	-	1491	3739
	Aromatics	-	-	-	-	-	-
	Lipids	7010	2032	1619	4239	1240	1173
	Carbs	3570	802	743	1789	590	582
	Others	652	231	326	364	214	303
	HOCs	-	79	81	-	82	138

“-” indicates that no transformations were detected between the corresponding molecular classes.

Table S7 (continued). The number of molecules transformed through hydroxylation, carboxylation, and carbonylation from day to night across different seasons, as determined by their molecular formulas.

Category		Summer			Winter		
Day	Night	Hydroxylation	Carboxylation	Carbonylation	Hydroxylation	Carboxylation	Carbonylation
UnsatHC	CRAMs	11528	2594	2444	10838	2901	2636
	Aliphatic/ peptides	-	1200	3091	-	1518	3794
	UnsatHC	2681	459	1562	2775	660	1902
	Aromatics	-	80	242	-	45	90
	Lipids	-	751	502	-	490	349
	Carbs	-	86	319	-	63	324
	Others	99	57	156	169	97	219
	HOCs	142	28	37	253	77	148
Aromatics	CRAMs	-	758	2057	-	1450	5211
	Aliphatic/ peptides	-	-	-	-	-	-
	UnsatHC	-	102	188	-	65	107
	Aromatics	4378	674	1415	7837	1859	2901
	Lipids	-	-	89	-	-	33
	Carbs	-	-	-	-	-	-
	Others	652	14	10	705	98	69
	HOCs	103	-	-	319	-	-

“-” indicates that no transformations were detected between the corresponding molecular classes.

Table S7 (continued). The number of molecules transformed through hydroxylation, carboxylation, and carbonylation from day to night across different seasons, as determined by their molecular formulas.

Category		Summer			Winter		
Day	Night	Hydroxylation	Carboxylation	Carbonylation	Hydroxylation	Carboxylation	Carbonylation
Lipids	CRAMs	-	3009	8575	-	1394	4244
	Aliphatic/ peptides	6765	1870	1574	3627	1037	958
	UnsatHC	-	754	483	-	417	312
	Aromatics	-	-	69	-	-	40
	Lipids	3313	939	2296	1299	264	837
	Carbs	1102	12	-	268	11	-
	Others	183	35	118	47	12	78
	HOCs	-	8	33	-	8	33
Carbs	CRAMs	-	-	-	-	-	-
	Aliphatic/ peptides	3829	819	781	1723	574	563
	UnsatHC	-	91	346	-	54	284
	Aromatics	-	-	-	-	-	-
	Lipids	1249	15	-	307	12	-
	Carbs	1765	527	1322	573	136	455
	Others	449	223	231	191	129	151
	HOCs	-	27	106	-	25	90

“-” indicates that no transformations were detected between the corresponding molecular classes.

Table S7 (continued). The number of molecules transformed through hydroxylation, carboxylation, and carbonylation from day to night across different seasons, as determined by their molecular formulas.

Category		Summer			Winter		
Day	Night	Hydroxylation	Carboxylation	Carbonylation	Hydroxylation	Carboxylation	Carbonylation
Others	CRAMs	310	186	417	247	358	739
	Aliphatic/ peptides	686	206	225	307	194	258
	UnsatHC	94	51	125	153	77	186
	Aromatics	142	10	6	540	72	59
	Lipids	199	23	92	54	13	87
	Carbs	417	209	226	171	125	141
	Others	344	70	183	273	45	162
	HOCs	32	12	18	44	10	43
HOCs	CRAMs	11528	103	115	388	177	210
	Aliphatic/ peptides	-	77	77	-	86	107
	UnsatHC	2681	29	43	242	65	119
	Aromatics	-	-	-	253	-	-
	Lipids	-	13	36	-	15	26
	Carbs	-	29	93	-	21	96
	Others	99	17	27	42	13	53
	HOCs	142	-	-	56	-	-

“-” indicates that no transformations were detected between the corresponding molecular classes.

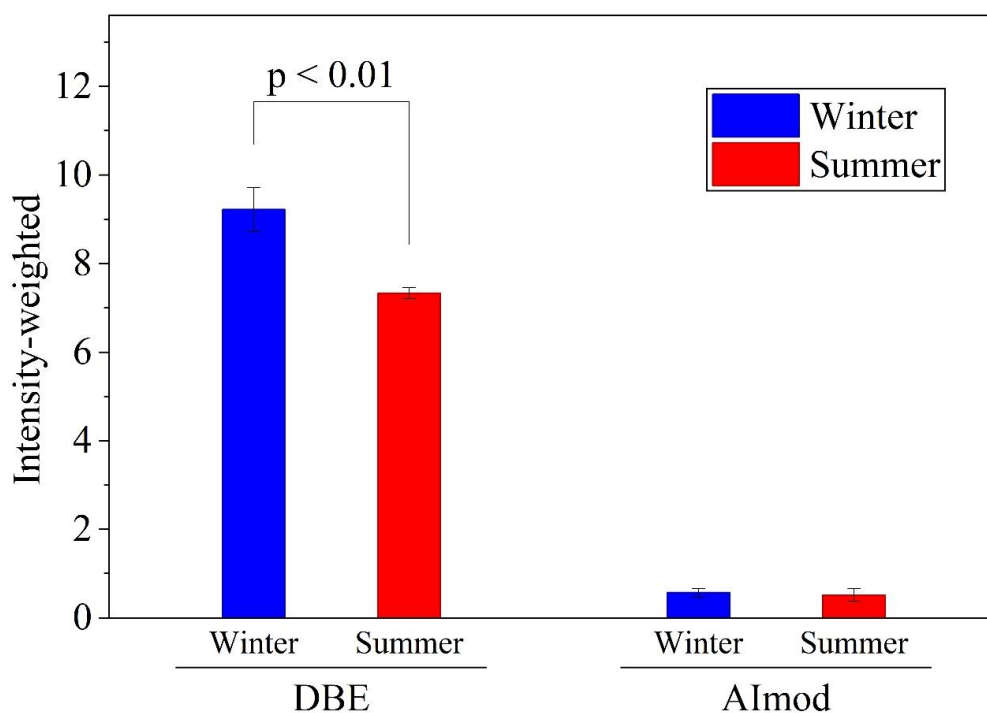


Figure S1. Seasonal comparison of intensity-weighted DBE and Almod values

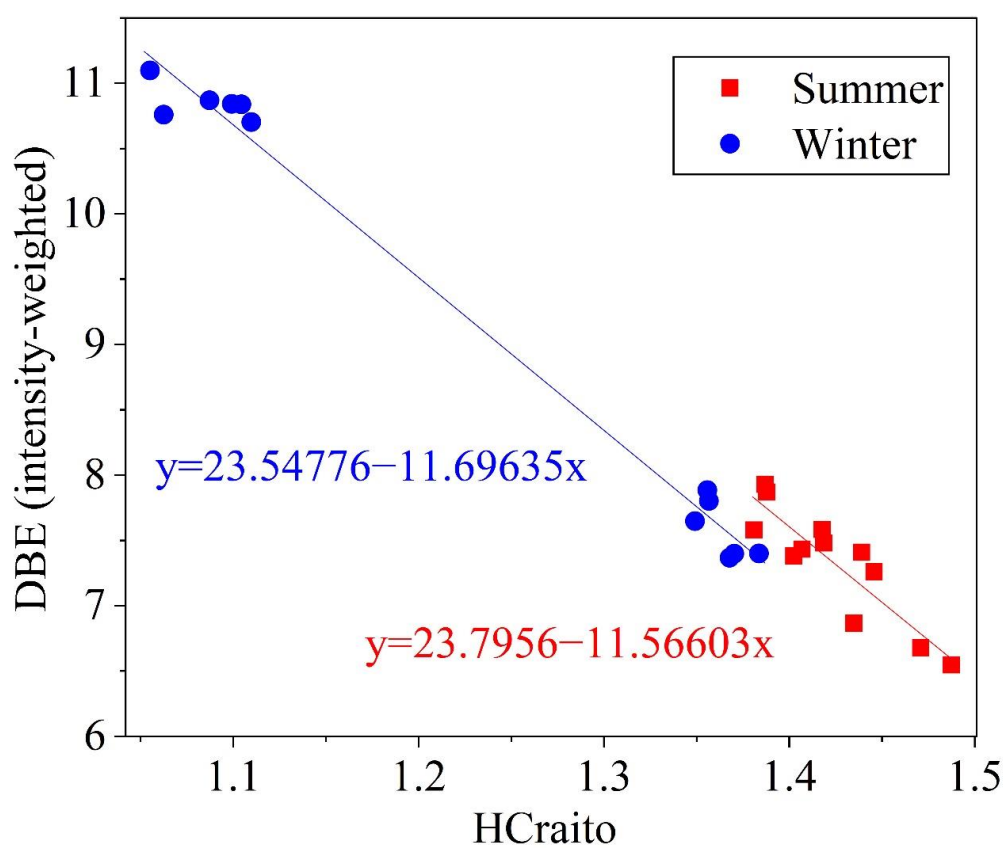


Figure S2. Seasonal variations in the relationship between DBE and H/C

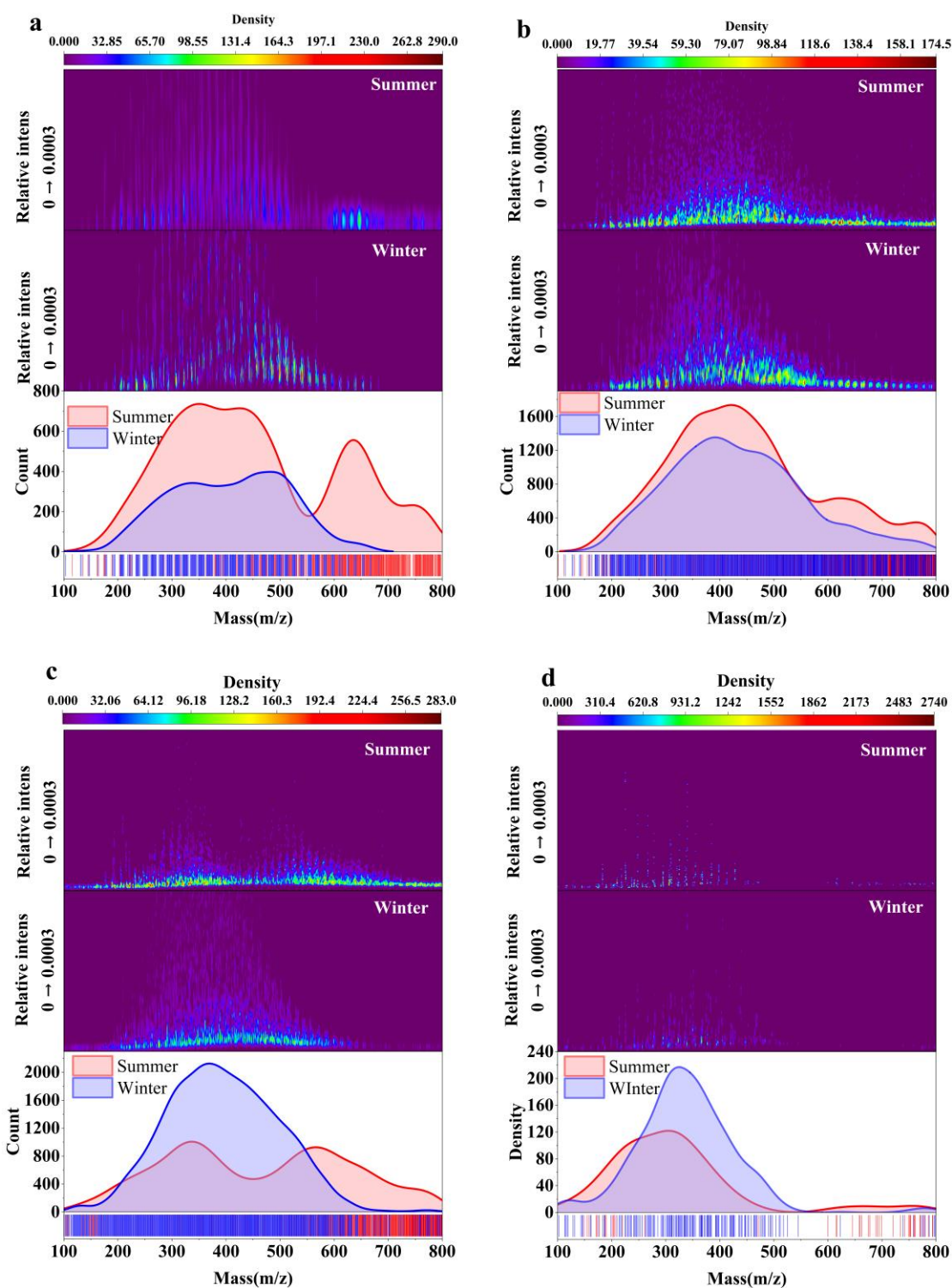


Figure S3. Mass and corresponding relative intensity (RI) distribution of (a) Carb, (b) UnsathC, (c) aromatic, and (d) HOCs

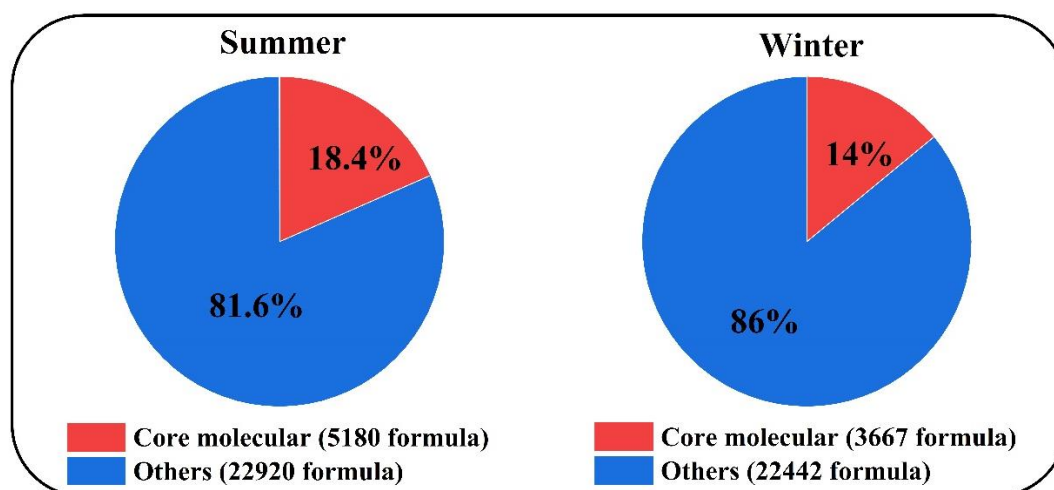
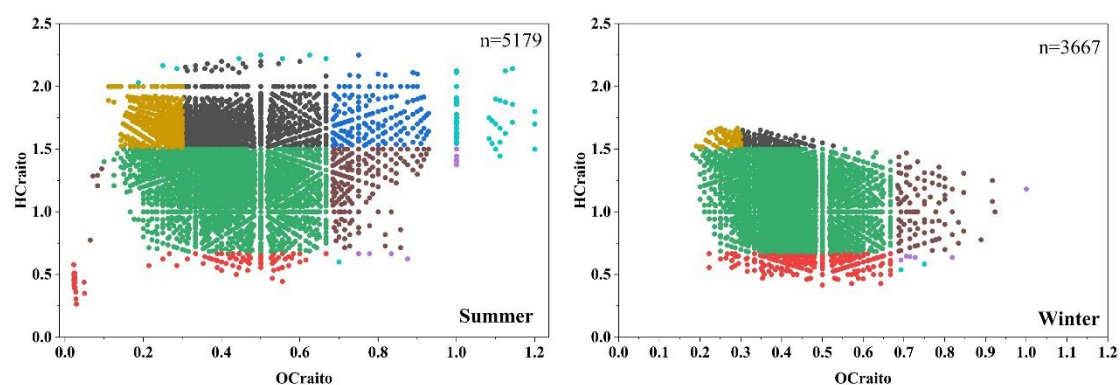


Figure S4. Seasonal variation in the proportion of core molecules



	aliphatic/ peptides	aromatic	carbohydrates	CRAMs	HOCs	lipids	unclassified	unsaturated hydrocarbons
Summer	23.4%	1.8%	3.6%	57.4%	0.2%	7.7%	1.1%	4.8%
Winter	2.6%	7.9%	-	85.5%	0.2%	1.4%	0.1%	2.5%

Figure S5. The van Krevelen plots for core molecular communities of WSOA in the summer and winter. In comparison, the summer core molecular community has 5179 formulae, whereas the winter core molecular community has 3667 formulae.

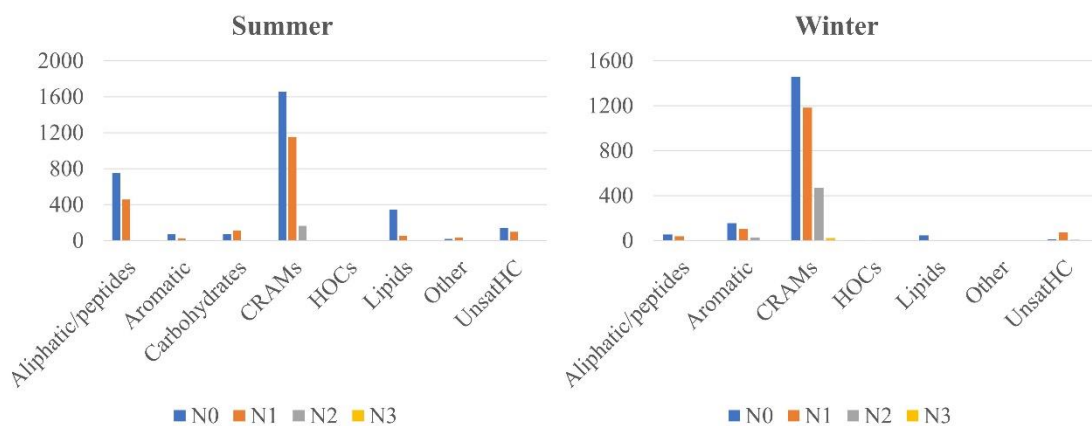


Figure S6. Seasonal distribution of various compound classes and N-containing subgroups

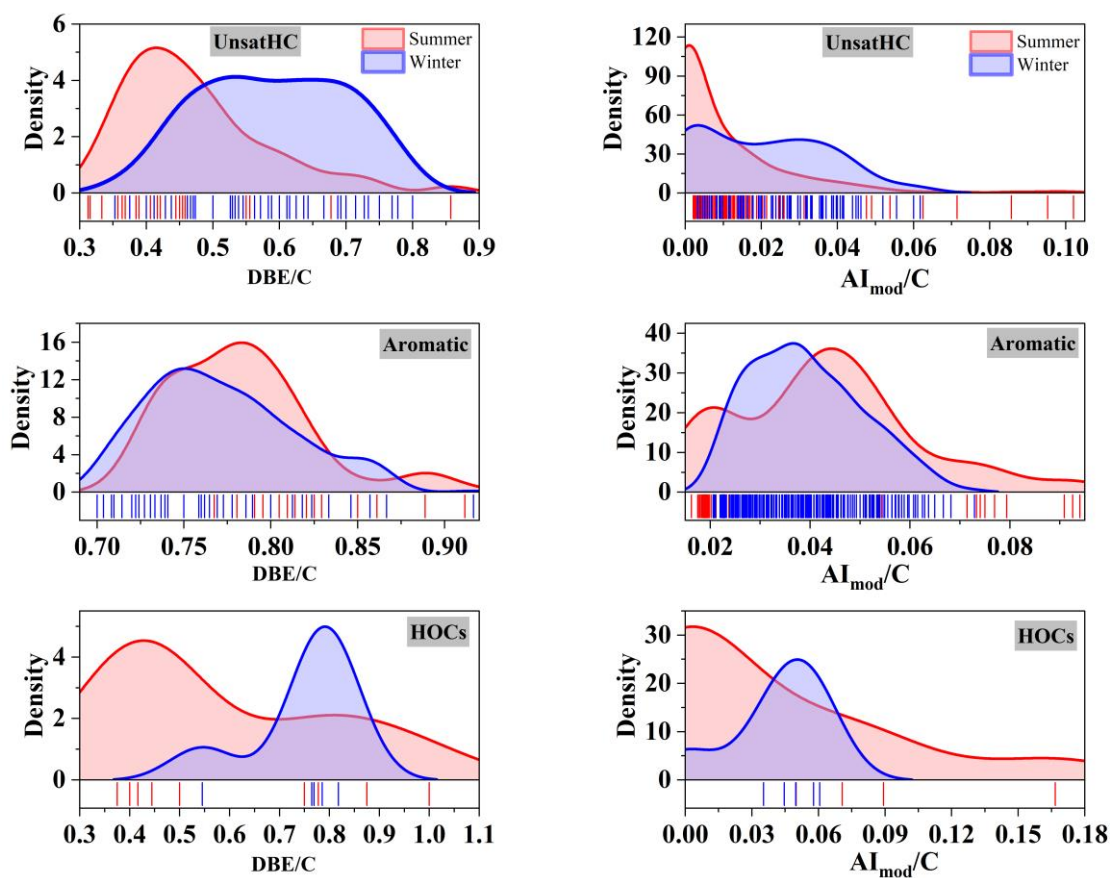


Figure S7. DBE/C and AImod/C distribution of UnsathHC, aromatic, and HOCs molecules of the core molecular community for summer and winter samples

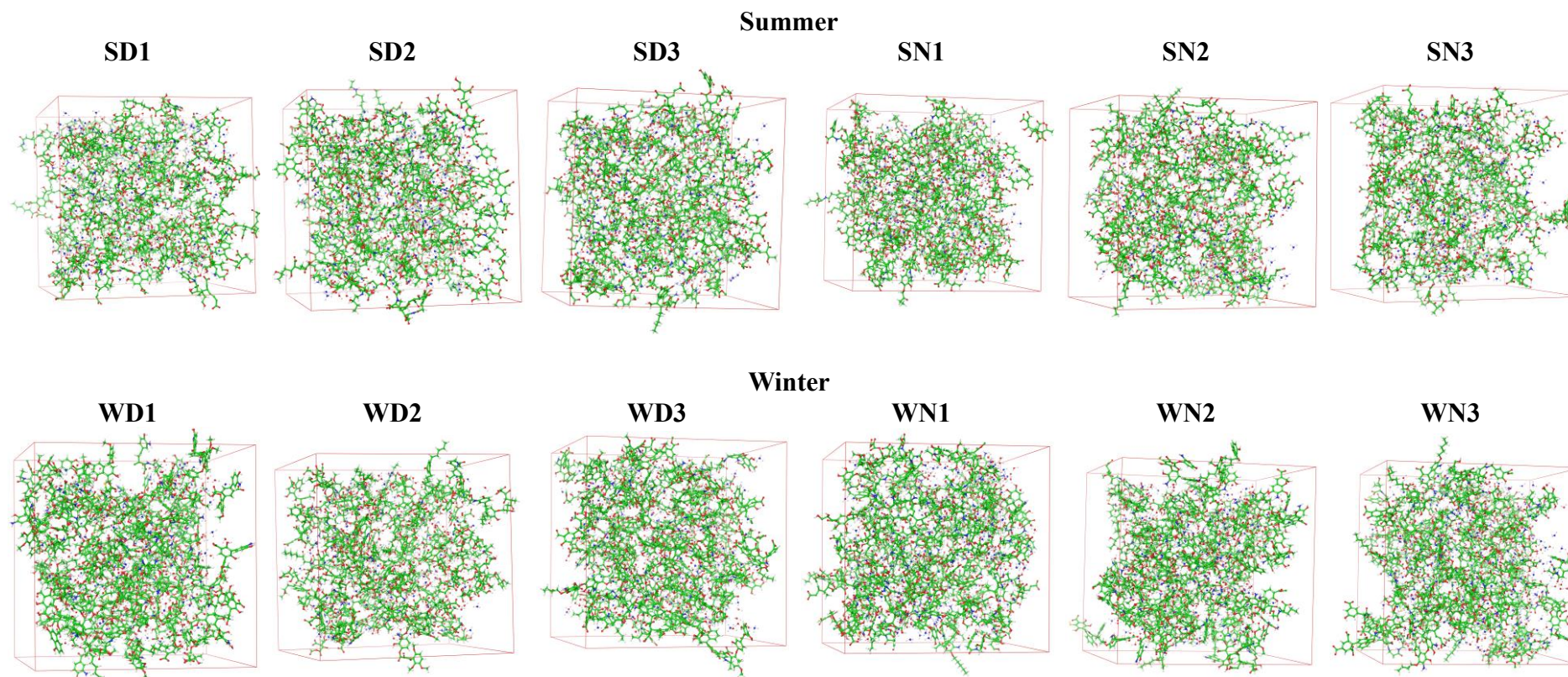


Figure S8. 3D visualizations of the 12 Model(WSOA) across seasons using VSOMM2 and Schrödinger, in which the edge length of the cubic box is 52 Å. 1, 2, and 3 denote molecular models that are positively correlated, uncorrelated, and negatively correlated with b₄₀₅, respectively. Green represents C atoms. White represents H atoms. Red represents O atoms. Blue represents N atoms. Yellow represents S atoms.

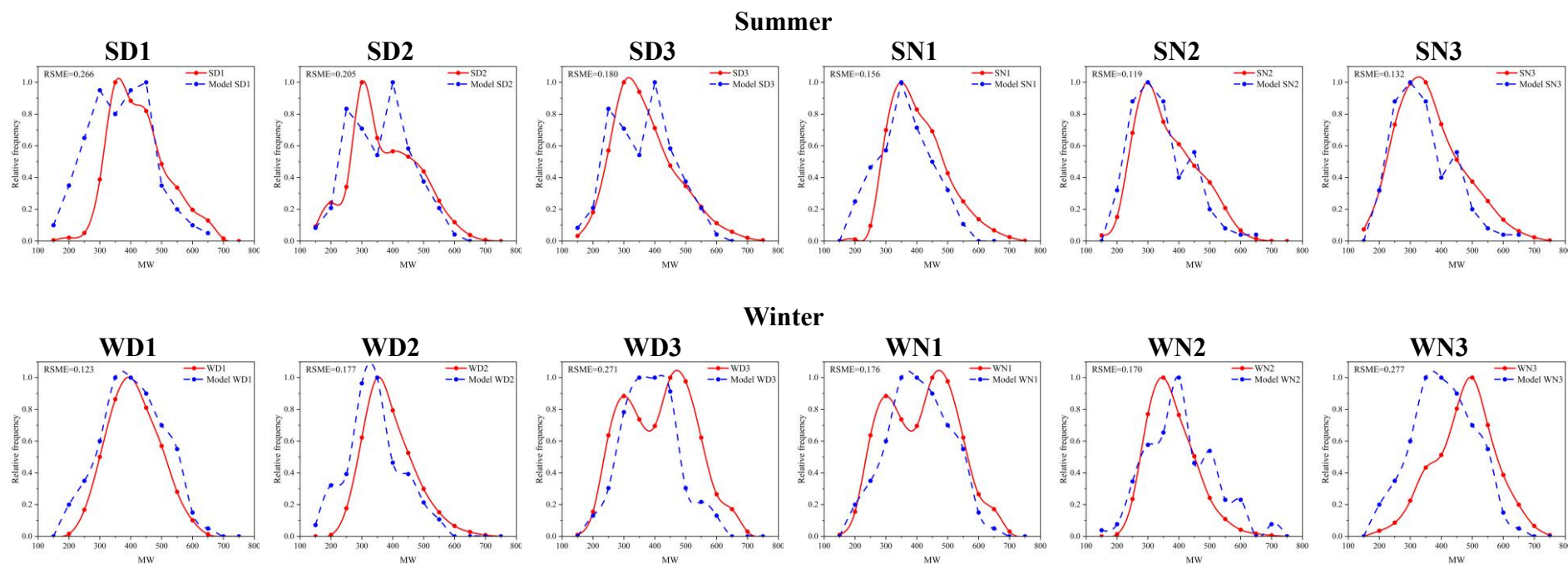


Figure S9. Model(WSOA) fits of the relative frequency distributions of molecular weight (MW) for different season samples. The RMSEs of the Model(WSOA) calculations range from 0.119 to 0.271. Red lines are FT-ICR-MS data and blue symbols with dashed lines are model calculations

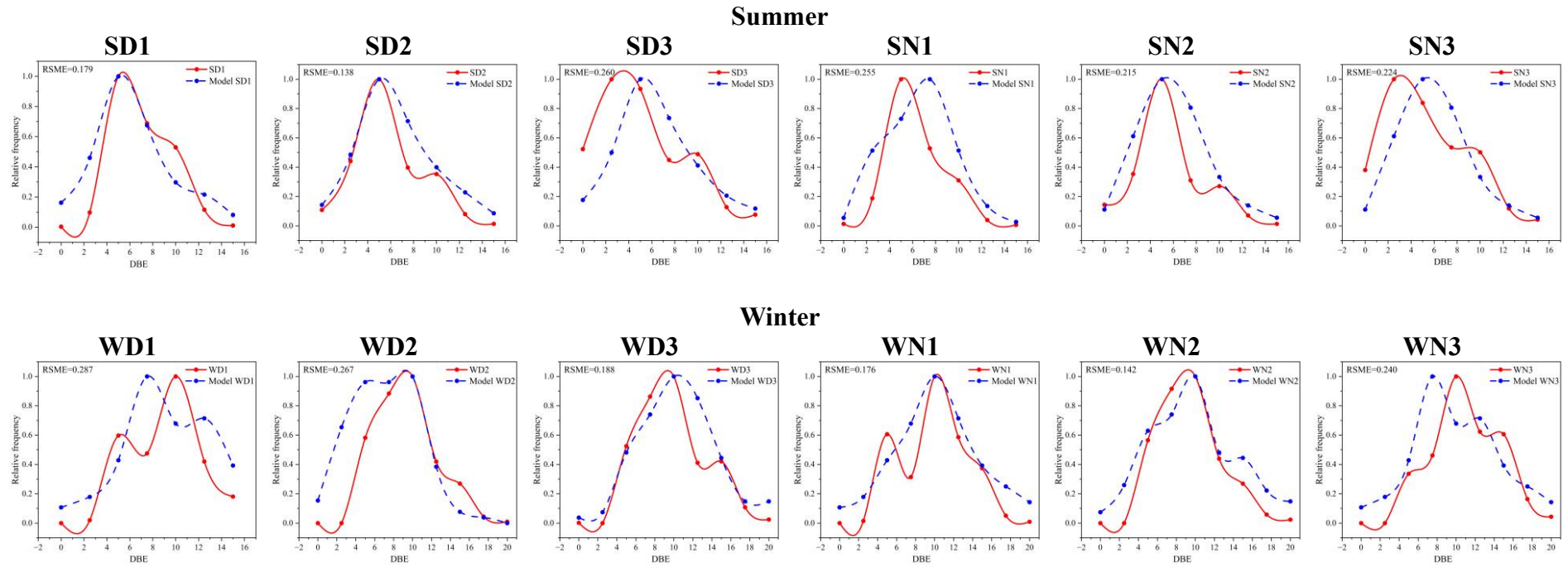


Figure S10. Model(WSOA) fits of the relative frequency distributions of DBE for different season samples. The RMSEs of the Model(WSOA) calculations range from 0.138 to 0.287. Red lines are FT-ICR-MS data and blue symbols with dashed lines are model calculations

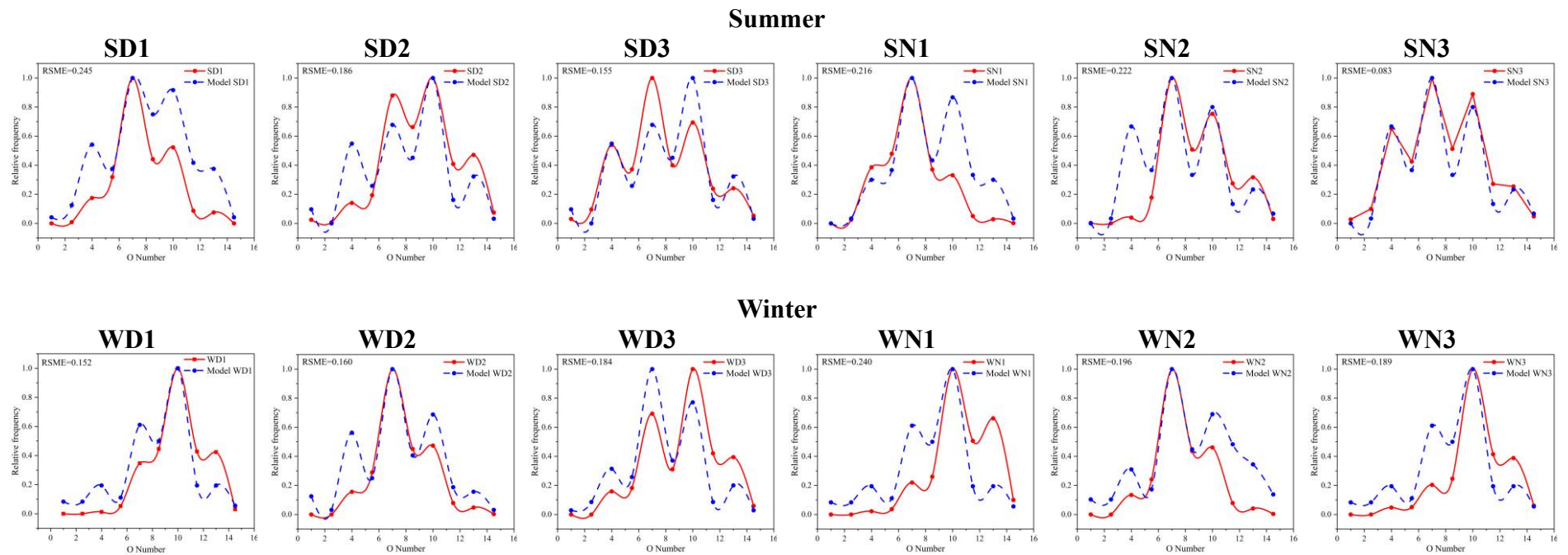


Figure S11. Model(WSOA) fits of the relative frequency distributions of O number for different season samples. The RMSEs of the Model(WSOA) calculations range from 0.083 to 0.245. Red lines are FT-ICR-MS data and blue symbols with dashed lines are model calculations

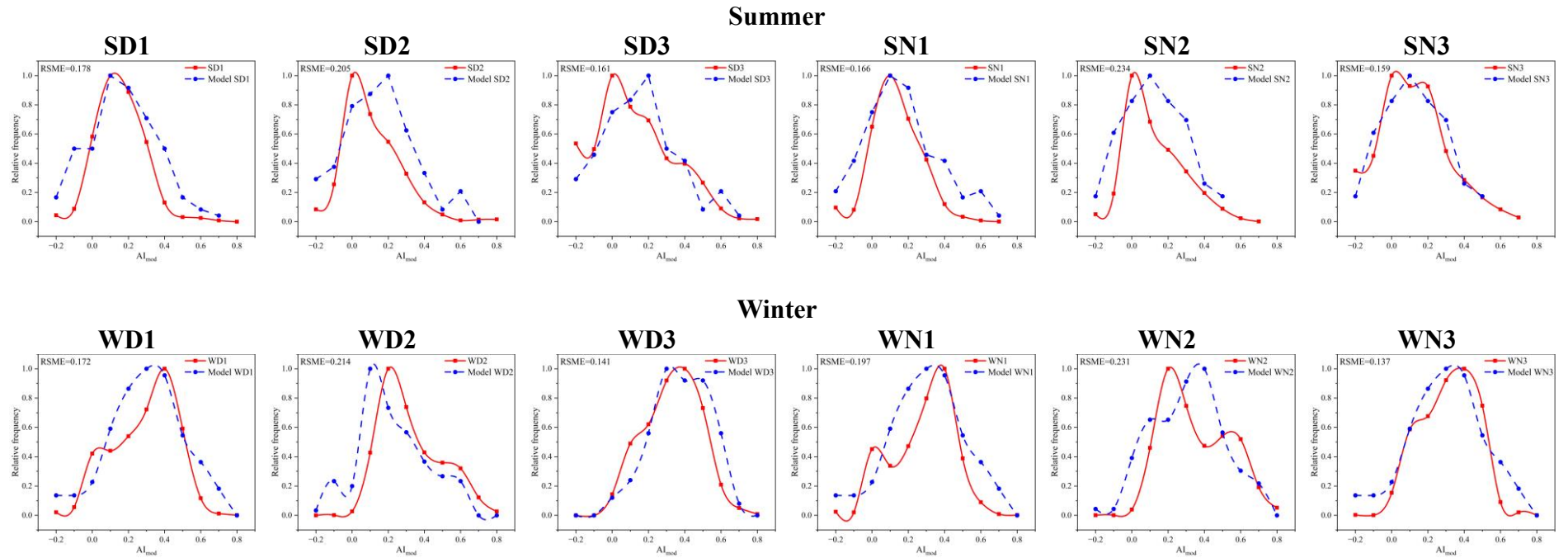


Figure S12. Model(WSOA) fits of the relative frequency distributions of AI_{mod} for different season samples. The RMSEs of the Model(WSOA) calculations range from 0.137 to 0.234. Red lines are FT-ICR-MS data and blue symbols with dashed lines are model calculations

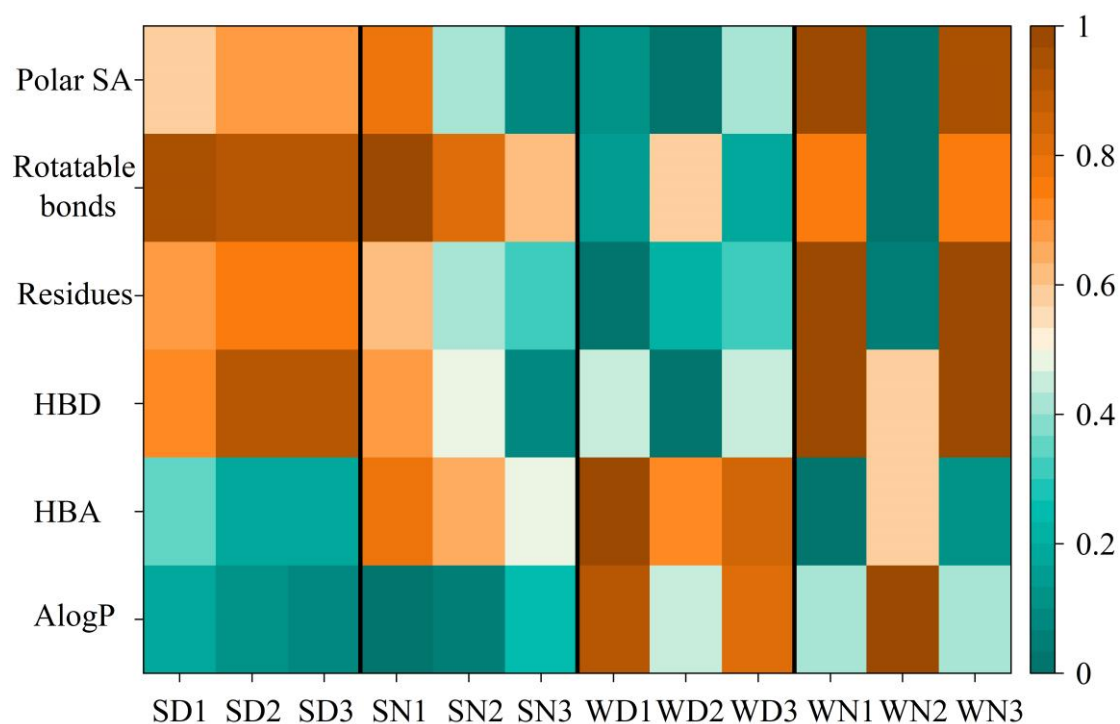


Figure S13. Heatmap of molecular descriptors for summer and winter samples

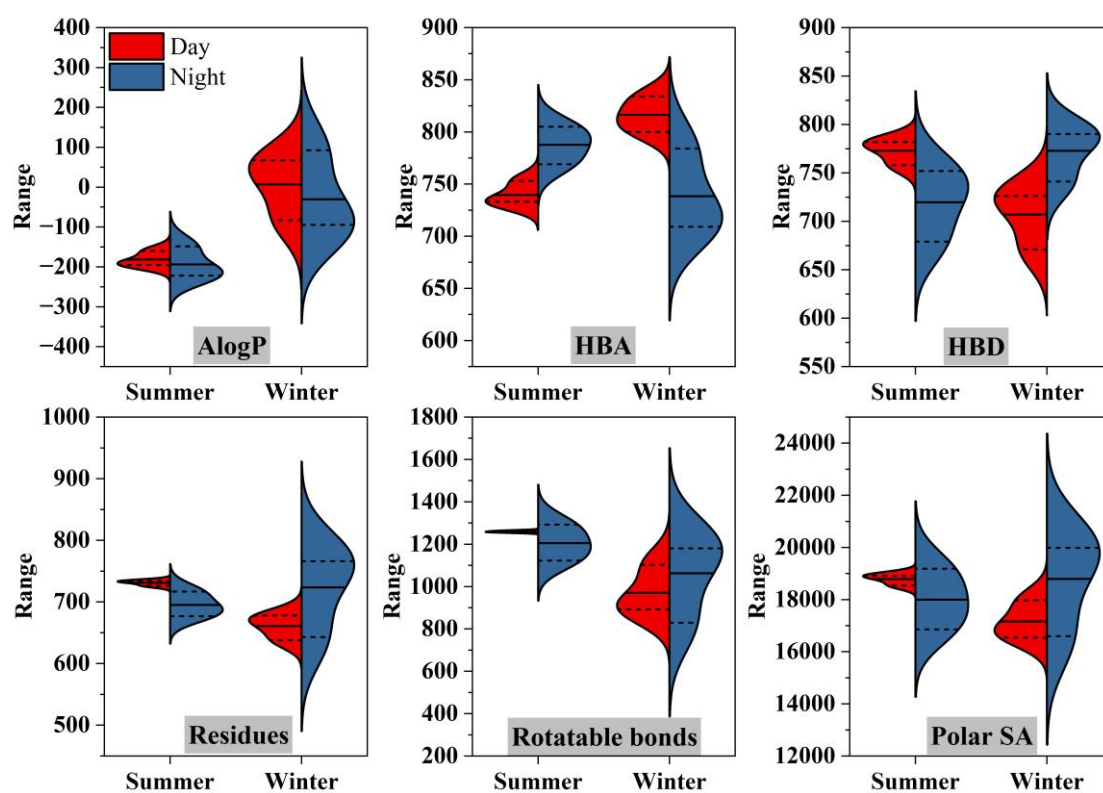


Figure S14. Seasonal and diurnal variations in molecular descriptor distributions