

Supporting information

Molecular insight into aqueous-phase photolysis and photooxidation of water-soluble organic matter emitted from biomass burning and coal combustion

Tao Cao¹, Cuncun Xu^{1,2}, Hao Chen^{1,2}, Jianzhong Song^{1,3,*}, Jun Li^{1,3}, Haiyan Song⁴, Bin Jiang^{1,3}, Yin Zhong^{1,3}, Ping'an Peng^{1,2,3}

¹State Key Laboratory of Advanced Environmental Technology and Guangdong Provincial Key Laboratory of Environmental Protection and Resources Utilization, Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou 510640, China

²University of Chinese Academy of Sciences, Beijing 100049, China

³Guangdong-Hong Kong-Macao Joint Laboratory for Environmental Pollution and Control, Guangzhou 510640, China

⁴School of Chemistry, South China Normal University, Universities Town, Guangzhou 510006, China

* *Correspondence to:* Jianzhong Song, E-mail: songjzh@gig.ac.cn.

Contents:

1. Text S1. Optical properties analysis
2. Text S2. FT-ICR MS analysis
3. Table S1. The intensity-weighted average elemental ratios, DBE_w , DBE/C_w , $AI_{mod,w}$ and $OS_{c,w}$ of molecular formulas within RS and YL WSOM for fresh, after photolysis and after OH photooxidation.
4. Table S2. Relative abundances (%) of seven categories components within RS and YL WSOM for fresh, after photolysis and after OH photooxidation.
5. Table S3. The number and percentage of molecules resistant, degraded and produced during photolysis and OH photooxidation for RS and YL WSOM.
6. Table S4. The information of number and intensity-weighted average elemental ratios, DBE_w , DBE/C_w , $AI_{mod,w}$ and $OS_{c,w}$ for condensed aromatic compounds produced during OH photooxidation in RS and YL WSOM.
7. Table S5. The intensity-weighted average elemental ratios, DBE_w , DBE/C_w , $AI_{mod,w}$ and $OS_{c,w}$ for molecular resistant, degraded and produced after photolysis and OH photooxidation in RS and YL WSOM.
8. Figure S1. Van Krevelen diagrams of molecular formulas within RS WSOM (a: fresh, c: photolysis, e: OH photooxidation) and YL WSOM (b: fresh, d: photolysis, f: OH photooxidation). The seven regions represent: I lipids-like, II protein /aliphatic sugar, III carbohydrates, IV unsaturated hydrocarbon, V lignin/CRAM-like, VI condensed aromatic, VII tannins.
9. Figure S2. Van Krevelen diagrams for molecules degraded and produced during photolysis and OH photooxidation within RS (upper) and YL WSOM (bottom).

Text S1. Optical properties analysis

The absorption coefficient (α_λ , m^{-1}) was calculated to estimate the abundance of typical chromophores within WSOM and it can be calculated as following equation:

$$\alpha_\lambda = 2.303 \times \frac{A_\lambda}{l} \quad (1)$$

In this equation, the absorbances at 254nm and 365nm were used (A_λ), l represent light path (0.01m).

The absorption was normalized by organic carbon mass which defined as mass absorbance efficiency (MAE), the MAE at 365 nm (MAE_{365}) is an important optical parameter used to characterize the light absorbing capacity of WSOM. It was obtained using the following equation:

$$\text{MAE}_\lambda = \frac{A_\lambda}{c * l} * \ln(10) \quad (2)$$

where A_λ is the absorbance at λ nm, c is the carbon concentration of BrC in solution (mgC L^{-1}), and l is the absorbing path length (0.01m).

The absorption Ångström exponent (AAE) is a measure of the spectral dependence of the light absorption of WSOM solutions, which was calculated by the following equation:

$$A_\lambda = K \lambda^{-AAE} \quad (3)$$

where A_λ is the absorbance derived from the spectrophotometer at a given wavelength λ (330– 400 nm) and K is a constant.

The PARAFAC was computed using two to nine component models, with non-negativity constraints and a residual analysis; and split half analysis was used to validate the number of fluorescence components. According to the results of the split-half

and core consistency analysis, three component models were chosen for further analysis. The identified individual fluorophores were compared with online database OpenFluor (based on the identified fluorophores in nature organic matter and the similarity of results for both excitation and emission wavelength were set at 98%). The relative contribution of individual chromophores was estimated by calculating the maximum fluorescence intensities (F_{\max} : maximum fluorescence intensity of identified fluorescence components, relative content % = $F_{\max}/\Sigma F_{\max}$).

Text S2. FT-ICR MS analysis

The samples were ionized in negative-ion mode using an Electrospray Ionization (ESI) source. The ion accumulation time was set to 0.65s and a total of 100 continuous 4M data FT-ICR transients were added to enhance the signal-to-noise ratio and the dynamic range. The detection mass range was set to m/z 100–800. Mass spectra were calibrated externally with arginine clusters in negative-ion mode using a linear calibration. The final spectrum was internally recalibrated with typical O_5 -class species peaks using quadratic calibration in DataAnalysis 4.4 (Bruker Daltonics). A typical mass-resolving power ($m/\Delta m_{50\%}$, where $\Delta m_{50\%}$ is the magnitude of the mass spectral peak full-width at half-maximum peak height) $> 450,000$ at m/z 319 with < 0.2 ppm absolute mass error was achieved.

The ultrahigh-resolution FT-ICR mass spectra were processed with custom software. All ions with relative abundances greater than 10 times the standard deviation of the baseline noise were explored. Mathematically possible formulae for these ions were then

calculated with a mass tolerance of ± 1 ppm. The formula CcHhOoNnSs was used to indicate the assigned compounds, in which C, H, O, N, S indicate carbon, hydrogen, oxygen, nitrogen, and sulfur, respectively, and c, h, o, n, s represent the number of atoms of the respective elements. The maximum numbers of atoms for the formula calculator were set as follows: 30 ^{12}C , 60 ^1H , 20 ^{16}O , 3 ^{14}N , 1 ^{32}S , 1 ^{13}C , 1 ^{18}O , and 1 ^{34}S . The identified formulae containing isotopomers (i.e., ^{13}C , ^{18}O , or ^{34}S) are not discussed. The calculated formulae were further characterized in terms of double-bond equivalents (DBE) and the modified aromaticity index (AImod), based on calculation of the elemental composition CcHhOoNnSs. The DBE value represents the number of rings plus double bonds, which reflects the degree of unsaturation (hydrogen deficiency) in a given compound. The DBE value can be calculated according to Equation (4):

$$\text{DBE} = \frac{2c+2-h+n}{2} \quad (4)$$

All calculated formulae with $\text{DBE} < 0$ and that disobey the nitrogen rule for even-number electron ions were excluded from the lists. The AImod value of each molecular formula was estimated based on the equation proposed by Koch and Dittmar:

$$AI_{mod} = \frac{1+c-0.5o-s-0.5h}{c-0.5o-s-n} \quad (5)$$

Based on the AI_{mod} values, the molecular formulae can be tentatively classified as aliphatic ($AI_{mod, w} = 0$), non-aromatic compounds ($0 < AI_{mod} < 0.5$), aromatic ($0.5 \leq AI_{mod} \leq 0.67$) and condensed aromatic ($0.67 < AI_{mod}$).

$$OSc \approx 2O/C - H/C \quad (6)$$

where O/C and H/C are the elemental ratio of oxygen-to-carbon and

hydrogen-to-carbon, respectively.

In the study, the overall properties of the WSOM fractions were assessed with relative abundance weighting because each molecule was present with different intensity. The relative-abundance-weighted molecular weight, elemental ratios, DBE, and AI_{mod} were calculated from the intensity (Int) of each assigned peak (i) based on the following equations:

$$MW_w = \Sigma(MW_i * Int_i) / \Sigma Int_i$$

$$OM/OC_w = \Sigma(OM/OC_i * Int_i) / \Sigma Int_i$$

$$O/C_w = \Sigma(O/C_i * Int_i) / \Sigma Int_i$$

$$O/N_w = \Sigma(O/N_i * Int_i) / \Sigma Int_i$$

$$O/S_w = \Sigma(O/S_i * Int_i) / \Sigma Int_i$$

$$DBE_w = \Sigma(DBE_i * Int_i) / \Sigma Int_i$$

$$AI_{mod,w} = \Sigma(AI_{mod,i} * Int_i) / \Sigma Int_i$$

$$OS_{C,w} = \Sigma(OS_C * Int_i) / \Sigma Int_i$$

where Int_i is the intensity for each individual molecular formula, i.

Although the negative ESI- FT-ICR MS provided detailed information of WSOM, it should be noted that the composition information is incomplete. According to previous studies, different ionization sources are favorable for different compounds. In fact, BrC is a complex of light-absorbing compounds, the solvent-extractable BrC light absorption is not only attributed to polar compounds but also due to the water-insoluble and less polar compounds such as PAHs and their derivatives. Therefore, other ionization techniques such

as positive-ion ESI and APPI are complementary to the negative-ion analysis in further studies.

Table S1. The intensity-weighted average elemental ratios, DBE_w, DBE/C_w, AI_{mod,w} and OS_{c,w} of molecular formulas in RS and YL WSOM for fresh, after photolysis and after OH photooxidation.

Samples	Composition	MW _w	H/C _w	O/C _w	O/N _w	O/S _w	DBE _w	DBE/C _w	AI _{mod,w}	OS _{c,w}	
RS WSOM	Fresh	CHO	236	1.19	0.38		5.98	0.50	0.43	-0.43	
		CHON	298	1.17	0.36	4.24		7.88	0.55	0.50	-0.44
		CHOS	292	1.28	0.55		6.25	5.31	0.45	0.22	-0.19
		CHONS	327	1.29	0.56	6.50	6.88	6.16	0.48	0.23	-0.17
		Total	252	1.19	0.38	1.05	0.19	6.40	0.51	0.44	-0.43
	Photolysis	CHO	267	1.17	0.44			6.63	0.50	0.41	-0.29
		CHON	342	1.15	0.39	5.14		8.89	0.54	0.47	-0.37
		CHOS	310	1.39	0.48		5.73	5.16	0.39	0.21	-0.42
		CHONS	310	1.33	0.59	6.51	6.63	5.51	0.47	0.20	-0.16
		Total	288	1.17	0.43	1.38	0.10	7.20	0.51	0.42	-0.31
	OH Photooxidation	CHO	302	1.16	0.60			6.73	0.50	0.35	0.045
		CHON	355	1.15	0.57	7.45		8.55	0.55	0.41	-0.0045
		CHOS	389	1.09	0.44		6.16	11.73	0.52	0.37	-0.22
		CHONS	362	1.34	0.68	7.14	7.14	8.53	0.47	0.20	0.020
		Total	319	1.16	0.59	1.88	0.26	7.36	0.52	0.36	0.023

Continued Table S1:

Samples	Composition	MW _w	H/C _w	O/C _w	O/N _w	O/S _w	DBE _w	DBE/C _w	AI _{mod,w}	OS _{c,w}	
YL WSOM	Fresh	CHO	218	0.95	0.39		7.07	0.62	0.60	-0.17	
		CHON	227	1.03	0.51	4.39		6.59	0.66	0.66	-0.011
		CHOS	255	1.23	0.56		5.34	5.14	0.49	0.28	-0.11
		CHONS	361	1.19	0.54	7.32	7.32	7.72	0.52	0.32	-0.11
		Total	231	1.03	0.46	1.62	1.07	6.60	0.61	0.56	-0.11
	Photolysis	CHO	255	0.98	0.43			7.65	0.60	0.55	-0.13
		CHON	326	1.00	0.42	6.02		9.79	0.62	0.58	-0.17
		CHOS	279	1.31	0.52		5.65	5.25	0.44	0.26	-0.27
		CHONS	326	1.22	0.53	6.59	6.59	6.77	0.52	0.32	-0.15
		Total	268	1.02	0.43	0.89	0.60	7.70	0.59	0.52	-0.15
	OH Photooxidation	CHO	293	1.02	0.57			7.54	0.57	0.46	0.12
		CHON	357	1.00	0.54	8.01		9.93	0.61	0.52	0.074
		CHOS	342	1.16	0.51		6.38	8.64	0.50	0.32	-0.14
		CHONS	460	0.97	0.39	5.23	5.23	17.95	0.62	0.51	-0.19
		Total	303	1.02	0.57	1.10	0.14	7.90	0.58	0.47	0.11

Table S2. Relative abundances (%) of seven categories components within RS and YL WSOM for fresh, after photolysis and after $\cdot\text{OH}$ photooxidation.

Samples		Lipids	Protein	Carbohydrate	Unsaturated	Lignin	Condensed	Tannins
RS WSOM	Fresh	5.69	7.19	0.67	0.36	83.1	1.08	1.91
	Photolysis	2.81	7.18	0.86	0.18	84.6	1.55	2.77
	$\cdot\text{OH}$ Photooxidation	0.074	4.94	3.69	0.13	63.3	4.86	23.0
YL WSOM	Fresh	0.52	1.67	0.36	0.31	88.4	2.86	5.86
	Photolysis	0.51	3.11	0.53	0.16	88.7	4.08	2.92
	$\cdot\text{OH}$ Photooxidation	0.052	2.32	0.96	0.032	73.9	5.38	17.3

Table S3. The number and percentage of molecules resistant, degraded and produced during photolysis and OH photooxidation for RS and YL WSOM.

Samples	Aging processes	Molecules	CHO	CHON	CHOS	CHONS	Total
RS WSOM	Photolysis	Resistant (n)	1731	2624	183	187	4725
		Degraded (n)	200	337	122	170	829
		Produced (n)	393	1165	82	18	1658
		Percentage of degraded in fresh (%)	3.60	6.07	2.20	3.06	14.9
		Percentage of produced in aged (%)	6.16	18.3	1.28	0.38	26.0
	OH photooxidation	Resistant (n)	919	1030	116	70	2135
		Degraded (n)	1012	1930	189	287	3418
		Produced (n)	904	1813	460	160	3337
		Percentage of degraded in fresh (%)	18.2	34.8	3.40	5.17	61.6
		Percentage of produced in aged (%)	16.5	33.1	8.41	6.97	61.0
YL WSOM	Photolysis	Resistant (n)	1480	1347	717	389	3933
		Degraded (n)	219	408	285	269	1181
		Produced (n)	703	868	215	37	1823
		Percentage of degraded in fresh (%)	4.28	7.98	5.57	5.26	23.1
		Percentage of produced in aged (%)	12.2	15.1	3.74	0.93	31.7
	OH photooxidation	Resistant (n)	779	702	279	31	1791
		Degraded (n)	920	1053	723	627	3323
		Produced (n)	1010	1097	233	31	2371
		Percentage of degraded in fresh (%)	18.0	20.6	14.1	12.3	65.0
		Percentage of produced in aged (%)	24.3	26.4	5.60	1.70	57.0

Table S4. The information of number and intensity-weighted average elemental ratios, DBE_w, DBE/C_w, AI_{mod,w} and OS_{c,w} for condensed aromatic compounds produced during OH photooxidation in RS and YL WSOM.

Samples	Composition	Number	MW _w	H/C _w	O/C _w	O/N _w	O/S _w	DBE _w	DBE/C _w	AI _{mod,w}	OS _{c,w}
RS WSOM	CHO	127	394	0.58	0.21			18.7	0.75	0.75	-0.16
	CHON	380	601	0.61	0.16	2.25		28.4	0.77	0.80	-0.55
	CHOS	111	426	0.60	0.14		3.69	19.7	0.74	0.73	-0.40
	CHONS	41	524	0.64	0.14	4.34	4.34	23.5	0.73	0.72	-0.53
	Total	659	483	0.60	0.16	0.95	1.37	22.6	0.75	0.76	-0.39
YL WSOM	CHO	55	377	0.53	0.18			19.3	0.78	0.78	-0.18
	CHON	199	670	0.55	0.15	2.67		33.4	0.78	0.80	-0.44
	CHOS	64	443	0.57	0.12		3.34	21.3	0.75	0.74	-0.41
	CHONS	25	721	0.58	0.10	4.73	4.73	35.3	0.74	0.74	-0.49
	Total	343	529	0.56	0.14	1.26	1.27	26.2	0.77	0.77	-0.37

Table S5. The intensity-weighted average elemental ratios, DBE_w, DBE/C_w, AI_{mod,w} and OS_{c,w} for molecular resistant, degraded and produced after photolysis and OH photooxidation in RS and YL WSOM.

Samples	Aging processes	MW _w	H/C _w	O/C _w	O/N _w	O/S _w	DBE _w	DBE/C _w	AI _{mod,w}	OS _{c,w}	
RS WSOM	Photolysis	Resistant in fresh	249	1.19	0.38	1.02	0.14	6.36	0.51	0.45	-0.42
		Resistant in aged	278	1.17	0.43	1.24	0.09	6.95	0.51	0.42	-0.39
		Degraded	392	1.29	0.33	2.43	2.51	8.83	0.44	0.33	-0.62
		Produced	472	1.07	0.49	4.23	0.28	12.0	0.54	0.43	-0.10
	OH photooxidation	Resistant in fresh	226	1.16	0.41	0.83	0.11	5.93	0.53	0.47	-0.34
		Resistant in aged	278	1.18	0.58	1.27	0.09	6.22	0.51	0.37	-0.018
		Degraded	351	1.31	0.27	1.89	0.50	8.24	0.42	0.36	-0.77
		Produced	423	1.10	0.61	3.43	0.71	10.30	0.53	0.35	0.13
YLWSOM	Photolysis	Resistant in fresh	227	1.03	0.46	1.59	1.03	6.49	0.61	0.56	-0.10
		Resistant in aged	256	1.02	0.43	0.78	0.60	7.39	0.59	0.53	-0.15
		Degraded	403	1.09	0.38	2.94	2.86	11.18	0.55	0.50	-0.33
		Produced	458	1.03	0.46	2.60	0.67	12.51	0.55	0.44	-0.12
	OH Photooxidation	Resistant in fresh	208	1.01	0.48	1.53	0.84	6.02	0.63	0.57	-0.053
		Resistant in aged	254	1.02	0.55	0.80	0.12	6.82	0.59	0.50	0.087
		Degraded	355	1.10	0.35	2.16	2.37	9.88	0.54	0.47	-0.40
		Produced	425	1.03	0.60	1.84	0.18	10.59	0.55	0.39	0.16

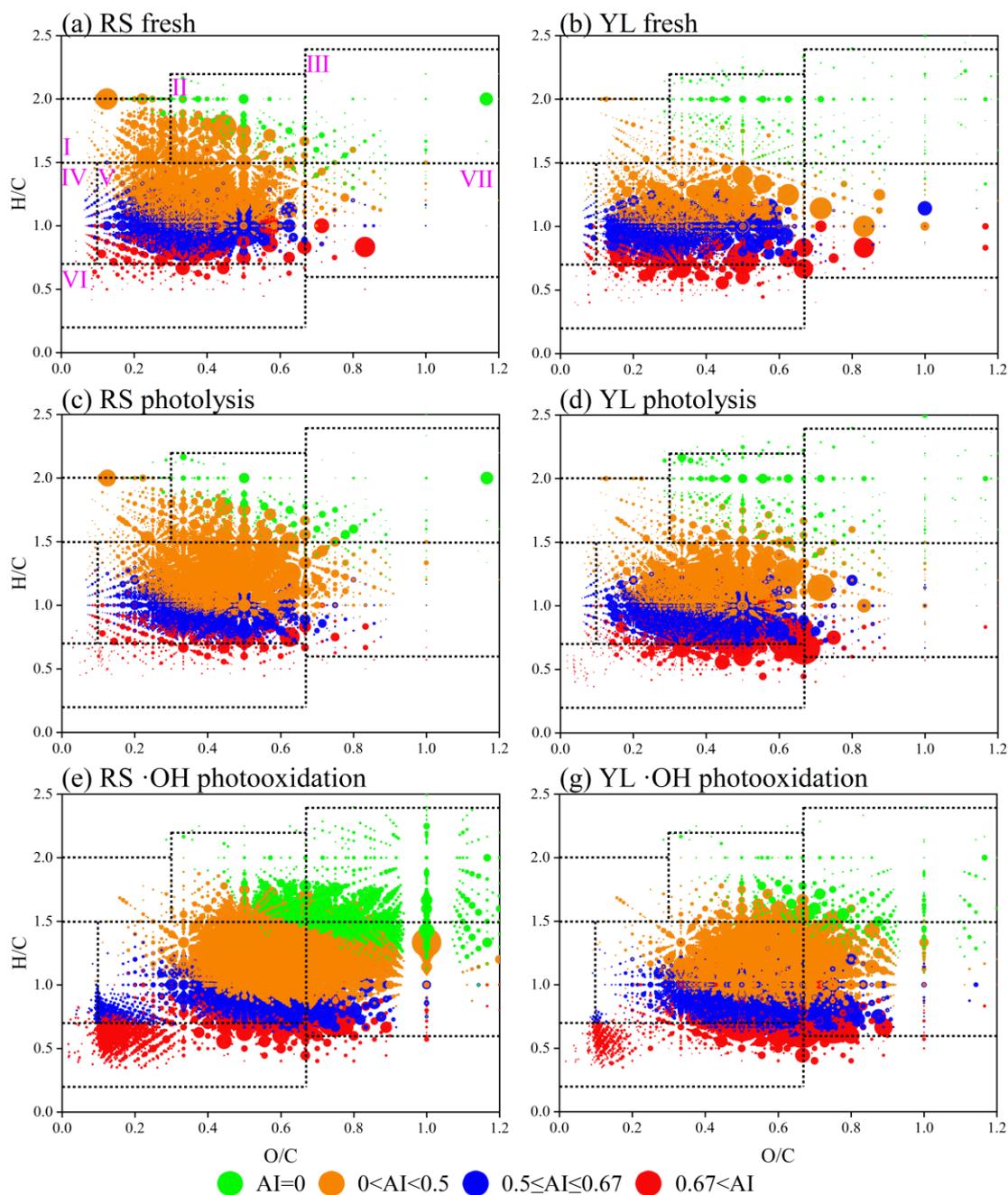


Figure S1. Van Krevelen diagrams of molecular formulas in RS WSOM (a: fresh, c: photolysis, e: OH photooxidation) and YL WSOM (b: fresh, d: photolysis, f: OH photooxidation). The seven regions represent: I lipids-like, II protein /aliphatic sugar, III carbohydrates, IV unsaturated hydrocarbon, V lignin/CRAM-like, VI condensed aromatic, VII tannins.

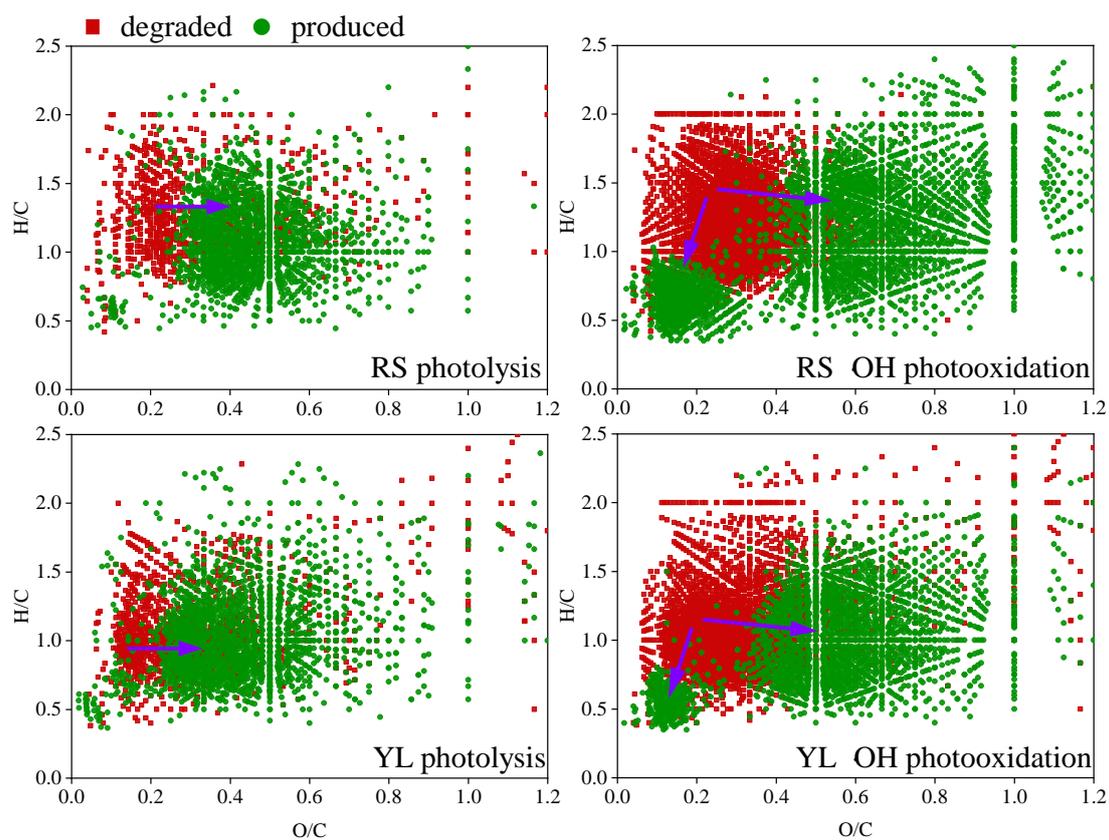


Figure S2. Van Krevelen diagrams for molecules degraded and produced during photolysis and OH photooxidation within RS (upper) and YL WSOM (bottom).