

Supporting Information

Deciphering Dissolved Organic Matter: Ionization, Dopant, and Fragmentation Insights via Fourier Transform-Ion Cyclotron Resonance Mass Spectrometry

Martin R. Kurek^{1*}, Brett A. Poulin^{2,3}, Amy M. McKenna⁴, and Robert G.M. Spencer¹

1. National High Magnetic Field Laboratory Geochemistry Group and Department of Earth, Ocean, and Atmospheric Science, Florida State University, Tallahassee, FL, 32306, USA.
2. U.S. Geological Survey, Water Mission Area, Boulder, Colorado, USA.
3. (Current Address) Department of Environmental Toxicology, University of California Davis, Davis, CA, 95616, USA
4. National High Magnetic Field Laboratory Ion Cyclotron Resonance Facility, Tallahassee, FL, USA.

*Corresponding author: Martin Kurek (mrk19f@my.fsu.edu)

Contains 6 tables, 5 figures, and 12 pages.

Supplemental Tables:

S2- Table S1

S3- Table S2

S4- Table S3

S5- Table S4

S6- Table S5

S7- Table S6

Supplemental Figures:

S8- Figure S1

S9- Figure S2

S10- Figure S3

S11- Figure S4

S12- Figure S5

Table S1. Molecular properties of assigned formulae for the Suwannee River DOM in toluene- and THF-doped APPI positive. Rows denote properties of formulae present in THF-doped APPI and toluene-doped APPI spectra, formulae common to both toluene- and THF-doped APPI spectra, formulae unique to THF-doped APPI spectra, formulae unique to toluene-doped APPI spectra, and formulae common between toluene- and THF-doped APPI as well as ESI negative spectra. H/C, O/C, Mass, AI_{mod} , DBE/C, formulaic N/C, and formulaic S/C are all computed means. % CHO, % CHON, % CHOS, and % CHONS values present the percentage of total formulae based on number averages, with intensity-weighted averages (relative abundance) presented in parentheses.

	Assigned Formulae	H/C	O/C	Mass (Da)	AI_{mod}	DBE/C	Formulaic N/C (x 1000)	Formulaic S/C (x 1000)	% CHO	% CHON	% CHOS	% CHONS
THF	7843	1.19	0.38	399.8608	0.31	0.49	48.9	6.6	46.8 (78.4)	44.1 (20.3)	5.0 (0.8)	4.1 (0.5)
Toluene	6687	1.14	0.39	350.9529	0.36	0.52	40.0	7.1	53.6 (85.2)	37.0 (13.4)	6.4 (1.0)	3.1 (0.4)
Common to THF and toluene	4593	1.16	0.40	354.7004	0.33	0.50	44.4	6.6	51.0	40.7	5.3	3.0
Unique to THF	3250	1.22	0.37	463.6828	0.28	0.47	55.2	6.5	40.9	49.0	4.4	5.7
Unique to toluene	2094	1.09	0.37	342.7331	0.42	0.56	30.5	8.3	59.3	28.8	8.7	3.2
Common to THF, toluene, and ESI neg	3101	1.08	0.44	382.7587	0.37	0.53	26.9	5.1	59.9	33.2	5.6	1.4

Table S2. Molecular properties of assigned formulae for the Kissimmee River DOM in toluene- and THF-doped APPI positive. Rows denote properties of formulae present in THF-doped APPI and toluene-doped APPI spectra, formulae common to both toluene- and THF-doped APPI spectra, formulae unique to THF-doped APPI spectra, formulae unique to toluene-doped APPI spectra, and formulae common between toluene- and THF-doped APPI as well as ESI negative spectra. H/C, O/C, Mass, AI_{mod} , DBE/C, formulaic N/C, and formulaic S/C are all computed means. % CHO, % CHON, % CHOS, and % CHONS values present the percentage of total formulae based on number averages, with intensity-weighted averages (relative abundance) presented in parentheses.

	Assigned Formulae	H/C	O/C	Mass (Da)	AI_{mod}	DBE/C	Formulaic N/C (x 1000)	Formulaic S/C (x 1000)	% CHO	% CHON	% CHOS	% CHONS
THF	8012	1.19	0.38	373.3835	0.31	0.50	69.2	9.1	37.3 (60.5)	51.7 (37.4)	5.5 (1.0)	5.5 (1.2)
Toluene	7625	1.12	0.38	344.5269	0.37	0.54	51.0	9.9	46.1 (76.4)	41.1 (20.8)	8.7 (1.9)	4.2 (0.9)
Common to THF and toluene	4858	1.14	0.39	341.7572	0.34	0.52	59.0	9.2	43.7	45.4	6.1	4.8
Unique to THF	3154	1.25	0.36	422.0966	0.25	0.47	84.7	8.9	27.4	61.4	4.6	6.6
Unique to toluene	2767	1.09	0.37	349.3898	0.42	0.56	37.0	11.1	50.2	33.4	13.2	3.2
Common to THF, toluene, and ESI neg	3201	1.06	0.41	368.8398	0.40	0.55	33.2	5.9	54.7	37.3	6.4	1.6

Table S3. Molecular properties of assigned formulae for the Wakulla Springs DOM in toluene- and THF-doped APPI positive. Rows denote properties of formulae present in THF-doped APPI and toluene-doped APPI spectra, formulae common to both toluene- and THF-doped APPI spectra, formulae unique to THF-doped APPI spectra, formulae unique to toluene-doped APPI spectra, and formulae common between toluene- and THF-doped APPI as well as ESI negative spectra. H/C, O/C, Mass, AI_{mod} , DBE/C, formulaic N/C, and formulaic S/C are all computed means. % CHO, % CHON, % CHOS, and % CHONS values present the percentage of total formulae based on number averages, with intensity-weighted averages (relative abundance) presented in parentheses.

	Assigned Formulae	H/C	O/C	Mass (Da)	AI_{mod}	DBE/C	Formulaic N/C (x 1000)	Formulaic S/C (x 1000)	% CHO	% CHON	% CHOS	% CHONS
THF	5070	1.31	0.33	346.6571	0.25	0.45	83.6	6.0	32.7 (35.7)	59.3 (62.1)	2.8 (0.7)	5.2 (1.6)
Toluene	3641	1.29	0.33	304.7302	0.29	0.47	77.9	4.3	41.5 (69.0)	53.1 (29.9)	2.3 (0.3)	3.2 (0.8)
Common to THF and toluene	2554	1.30	0.34	298.9546	0.27	0.47	87.5	3.8	37.9	57.6	1.1	3.5
Unique to THF	2516	1.31	0.33	395.0800	0.23	0.44	79.7	8.3	27.4	61.1	4.5	7.0
Unique to toluene	1087	1.25	0.31	318.3006	0.35	0.50	55.2	5.5	50.0	42.4	5.2	2.5
Common to THF, toluene, and ESI neg	1173	1.32	0.36	346.5263	0.25	0.41	34.8	1.3	58.0	40.1	0.9	1.0

Table S4. Relative compositions of molecular classes from the IRMPD parent ions isolated from the 392-408 Da mass window and fragments from the 141 to 391 Da mass window ionized in THF-doped APPI positive for the Suwannee River and Kissimmee River DOM. % CHO, % CHON, % CHOS, and % CHONS values present the percentage of total formulae based on number averages, with intensity-weighted averages (relative abundance) presented in parentheses.

	% CHO	% CHON	% CHOS	% CHONS
Suwannee River parent ions	37.6 (80.7)	44.4 (16.8)	12.0 (1.9)	6.0 (0.6)
Kissimmee River parent ions	36.1 (69.5)	48.0 (27.2)	9.0 (2.0)	6.8 (1.2)
Suwannee River fragments	48.2 (85.7)	45.3 (13.6)	5.4 (0.6)	1.2 (0.1)
Kissimmee River fragments	40.5 (78.1)	47.2 (20.2)	9.2 (1.3)	3.1 (0.4)

Table S5. Molecular properties of assigned formulae from the organic isolates in ESI negative, THF-doped APPI positive, and from formulae common between both ESI negative and THF-doped APPI positive (common). H/C, O/C, Mass, AI_{mod} , DBE/C, formulaic N/C, and formulaic S/C are all computed means. % CHO, % CHON, % CHOS, and % CHONS values present the percentage of total formulae based on number averages, with intensity-weighted averages (relative abundance) presented in parentheses. Atomic N/C was calculated as ratios from the molar C and N compositions of the organic isolates (see methods).

	Penobscot River			Suwannee River			Everglades			Lake Fryxell			Pacific Ocean		
	ESI neg	APPI THF	Common	ESI neg	APPI THF	Common	ESI neg	APPI THF	Common	ESI neg	APPI THF	Common	ESI neg	APPI THF	Common
Assigned formulae	8876	6013	4170	7726	4484	3170	10025	6723	4956	9110	6453	5098	5650	4298	2703
H/C	1.00	1.17	1.08	0.97	1.21	1.10	1.02	1.17	1.09	1.22	1.29	1.25	1.26	1.28	1.27
O/C	0.44	0.38	0.40	0.44	0.39	0.41	0.44	0.39	0.40	0.44	0.40	0.41	0.45	0.37	0.40
DBE/C	0.56	0.49	0.53	0.57	0.47	0.51	0.55	0.50	0.53	0.47	0.43	0.45	0.42	0.43	0.42
AI_{mod}	0.40	0.33	0.38	0.42	0.31	0.37	0.38	0.32	0.37	0.25	0.22	0.25	0.22	0.26	0.25
Mass (Da)	481.7525	386.434	397.7703	481.6426	385.541	400.5448	459.0104	386.0038	389.7462	459.7179	426.6683	412.3246	556.0078	428.3332	441.0678
Formulaic N/C (x1000)	23.2	38.4	22.6	19.1	29.6	18.2	34.1	47.6	36.5	43.8	49.2	45.5	24.9	37.6	26.9
Formulaic S/C (x1000)	8.1	6.7	4.7	8.3	2.6	2.1	13.7	11.8	9.4	19.1	16.3	14.8	3.7	2.9	0.8
%CHO	52.9 (89.4)	55.0 (89.0)	65.1	55.7 (91.4)	63.1 (91.6)	71.6	35.6 (75.1)	40.0 (74.6)	44.5	28.1 (62.8)	30.2 (65.3)	33.3	52.4 (86.3)	50.4 (83.2)	59.8
%CHON	31.9 (7.1)	35.7 (9.9)	28.2	28.6 (5.8)	33.0 (7.8)	25.2	39.4 (17.6)	44.4 (22.3)	41.3	42.1 (22.3)	47.3 (27.8)	44.8	39.5 (11.6)	45.4 (16.2)	38.8
%CHOS	13.5 (3.3)	6.8 (0.8)	6.6	14.7 (2.7)	3.4 (0.4)	3.2	18.8 (6.3)	10.3 (2.1)	10.5	21.6 (12.6)	15.7 (5.2)	15.4	7.8 (2.0)	3.6 (0.1)	1.4
%CHONS	1.7 (0.2)	2.5 (0.3)	0.1	1.0 (0.1)	0.5 (0.1)	0.0	6.1 (1.2)	5.3 (1.0)	3.7	8.2 (2.3)	6.7 (1.8)	6.5	0.4 (0.1)	0.6 (0.1)	0.0
Atomic N/C (x1000)		15.4			12.9			25.9			44.7			17.2	

Table S6. Correlations and percent contributions of variables for principal components 1 and 2 from the PCA in figure 3. CHO, CHON, CHOS, and CHONS were computed as percent relative abundances.

	PCA1		PCA2	
	Correlation	Contribution	Correlation	Contribution
m/z	-0.74	9.41	0.25	1.91
H/C	0.78	10.31	0.03	0.03
O/C	-0.85	13.51	0.34	3.63
N/C	0.92	14.29	0.21	1.44
S/C	-0.25	1.10	0.81	20.27
CHO index	-0.93	14.51	0.14	0.61
DBE	-0.94	14.80	0.08	0.19
CHO	-0.63	6.91	-0.70	15.38
CHON	0.79	10.45	0.48	7.07
CHOS	-0.47	3.67	0.82	21.11
CHONS	0.24	1.03	0.95	28.37

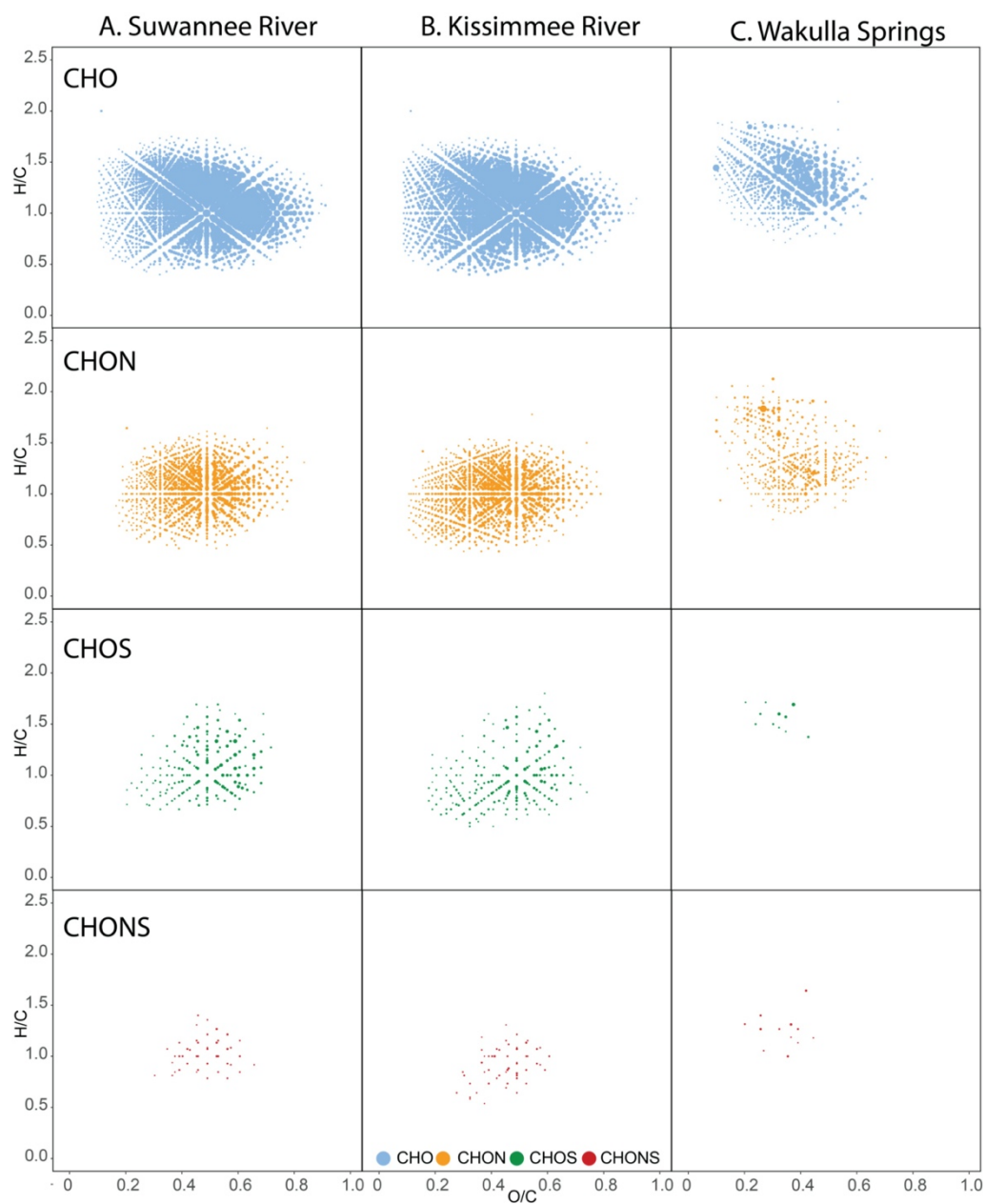


Figure S1. van Krevelen diagrams of the Suwannee River, Kissimmee River, and Wakulla Springs common DOM formulae observed between in ESI negative, APPI positive with THF doping, and APPI positive with toluene doping. Molecular formulae are colored by heteroatom composition (blue = CHO, top row; orange = CHON, 2nd row; green = CHOS, 3rd row; red = CHONS, bottom row) with the sizes scaled to relative abundances.

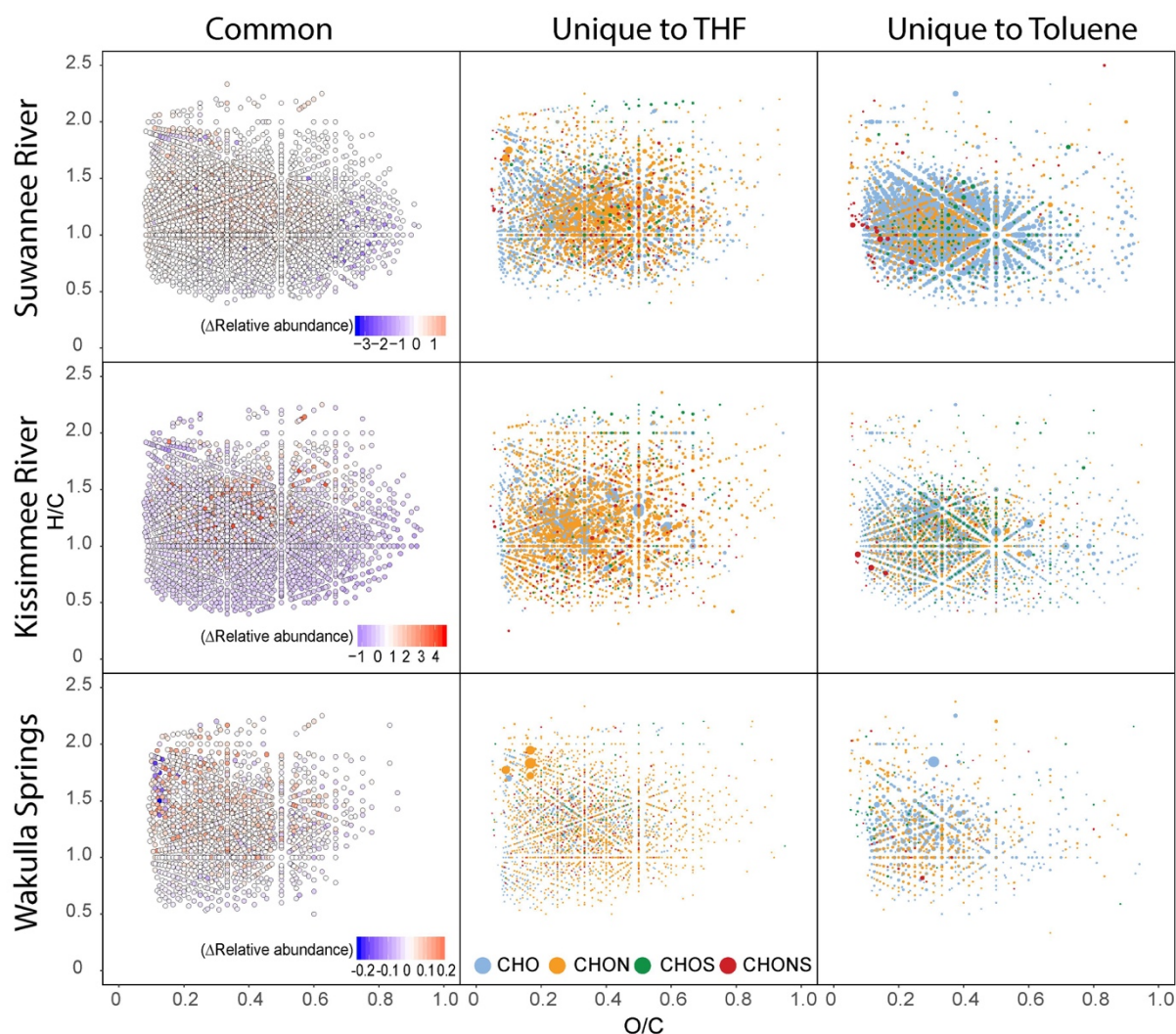


Figure S2. van Krevelen diagrams of assigned formulae by THF-doped and toluene-doped APPI measurements of the Suwannee River DOM, (top row), Kissimmee River DOM (middle row), and Wakulla Springs DOM (bottom row). Common formulae between THF- and toluene-doped measurements are colored by their change in relative abundance (90% quantile) where positive values (red) indicate greater relative abundance in THF-doped APPI and negative values (blue) indicate greater relative abundance in toluene-doped APPI. Formulae unique to both THF and toluene are colored by their molecular class (blue = CHO, orange = CHON, green = CHOS, red = CHONS) with their sizes scaled to relative abundances.

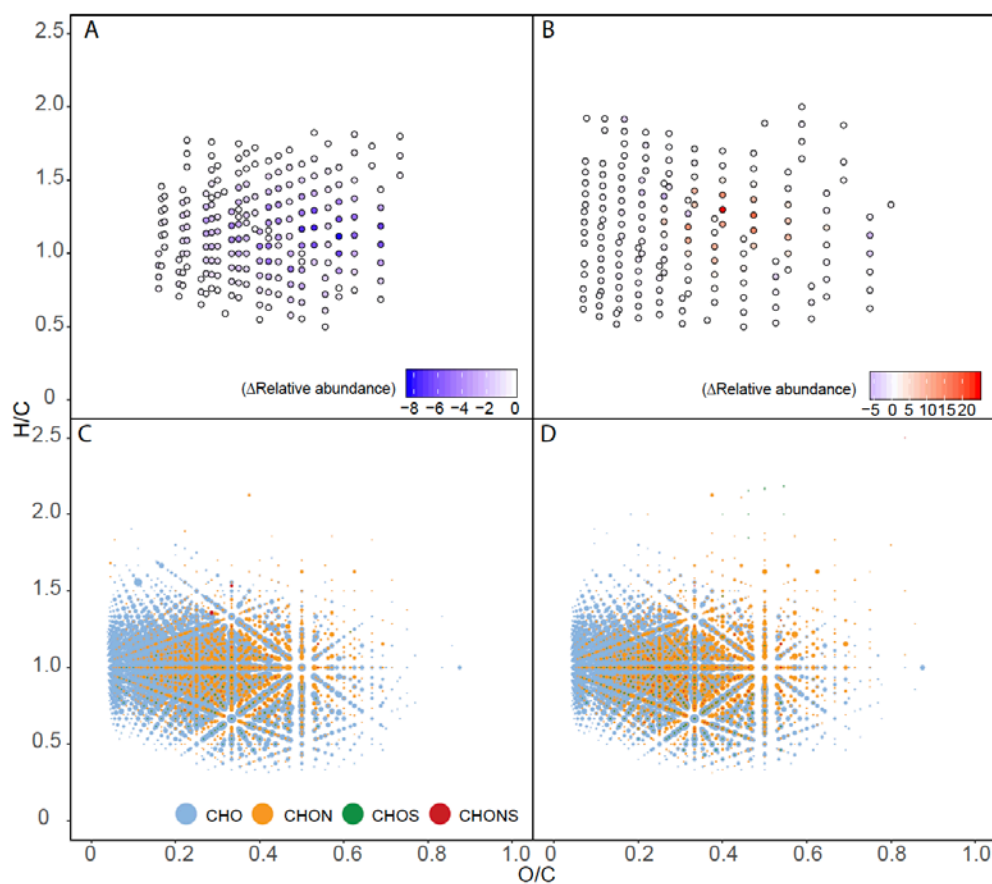


Figure S3. van Krevelen diagrams of assigned formulae in the isolated mass range of 392-408 Da from IRMPD experiments in both the Suwannee and Kissimmee River DOM samples. Common formulae observed in both DOM samples are colored by their change in relative abundance where positive values (red) are more abundant in Suwannee and negative values (blue) are more abundant in Kissimmee for (A) CHON-containing and (B) CHO-containing formulae. Fragments from 141-391 Da in the (C) Suwannee and (D) Kissimmee River are colored by their molecular classes (blue = CHO, orange = CHON, green = CHOS, red = CHONS) with their sizes scaled to relative abundances.

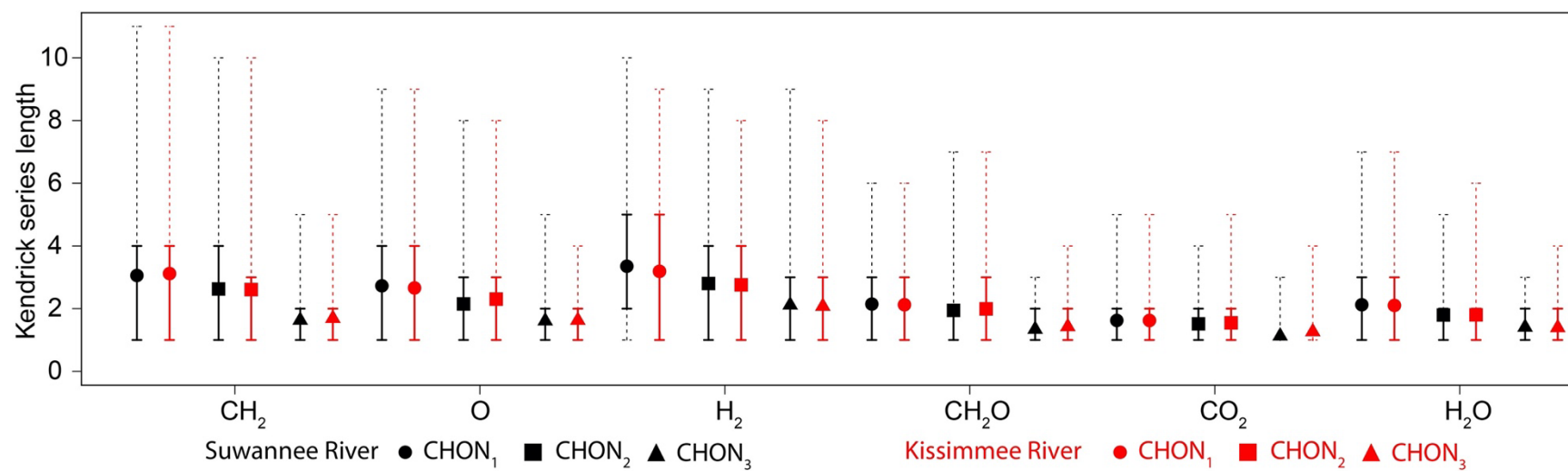


Figure S4. Boxplots of CH₂, O, H₂, CH₂O, CO₂, and H₂O-based Kendrick series lengths of CHON formulae from the Suwannee (black) and Kissimmee (red) River. N-containing classes are subdivided by CHON₁ (circles), CHON₂ (squares), and CHON₃ (triangles).

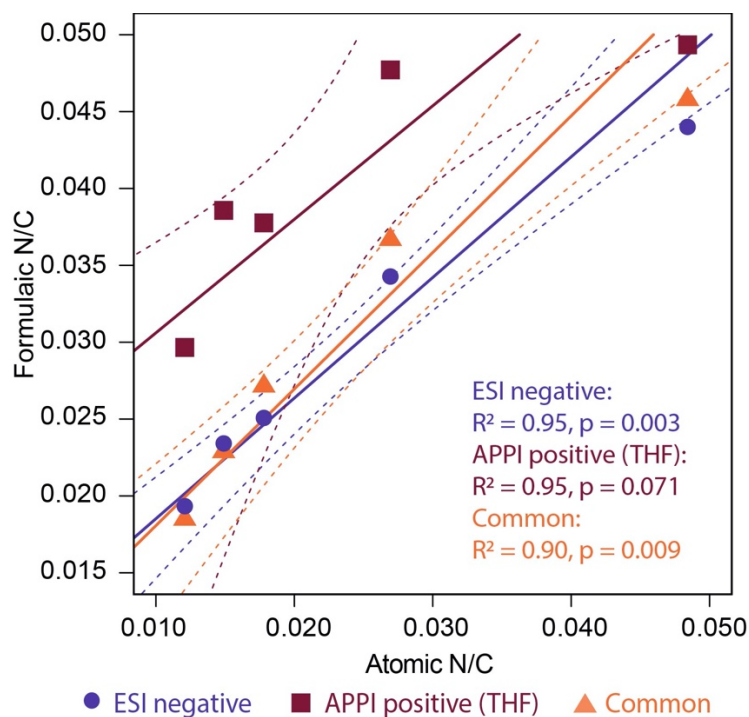


Figure S5. Relationship between the measured atomic N/C of the DOM isolates and the number-averaged formulaic N/C from the FT-ICR mass spectra measured by ESI negative (blue circles), THF-doped APPI positive (magenta squares), and formulae common to both ionization modes (orange triangles). Linear regressions (colored solid lines) and 95% confidence intervals (colored dashed lines) from each ionization were modeled with all DOM isolates.