Impact of natural NOM and AOM in Vacuum UV treatment

Flavia Visentin*[a], Siddharth Bhartia[b], Madjid Mohseni[b], Sigrid Peldszus[c] , Sarah Dorner[a] and Benoit Barbeau[a]

[a] Polytechnique de Montreal, NSERC- Industrial Chair in Drinking Water Treatment, Department of Civil, Geological and Mining Engineering , P.O. Box 6079, Downtown Station, Montreal, Quebec, Canada, H3C 3A7

[b] Department of Chemical and Biological Engineering, University of British Columbia, Vancouver, British Columbia, Canada

[c] Deptartment of Civil & Environmental Engineering University of Waterloo, 200 University Av. W Waterloo, Ontario, Canada N2L 3G1

Supplementary Information

SUVA values before and after VUV treatment

SUVA, defined as UVA_{254} normalized to DOC, is a good predictor of the aromaticity of NOM. In 1999, Edzwald and Tobiason proposed that a SUVA value greater than 4 indicates the predominance of aquatic humic matter, highly aromatic and hydrophobic character, and high molecular weight. While SUVA values between 2 and 4, are composed by a mixture of aquatic humic and non-humic matter, with a mix of aromatic and aliphatic character and low to high molecular weight. When SUVA lower than 2, then a high fraction of non-humic matter is present, aliphatic and hydrophobic character and low molecular weight [1].

Table SI. 1. SUVA, L m^{-1} mg C^{-1}, values before and after VUV treatment

		Lake			
		A	**B**	**C**	**C***
Raw Water	Untreated (retention time = 0 sec)	2.70	4.27	4.01	3.89
	VUV treated (retention time = 9.4 sec)	2.49	3.97	3.84	3.48
Filtered water	Untreated (retention time = 0 sec)	3.06	-	2.71	2.73
	VUV treated (retention time = 9.4 sec)	2.91	-	2.55	2.48

LC-OCD-OND chromatograph

Figure SI 1 (a, d, g and j) shows the ultraviolet detector (UVD) signal from the LC-OCD-OND for the four lakes. Without VUV treatment, in Lake A (where a bank filtration water system is responsible for NOM removal), the UVD signal in filtered water is 25% lower than in raw water. Among the four lakes, Lake B shows the higher UVD signal due to the high cell concentration (almost 4 times higher than Lake A and two times Lake C and C*). DOC measurement was stable and UVD declined, thus modest reductions of SUVA were obtained (about 0.2-0.4 units) (Table SI.1). In terms of TOC in Lake C and C*, it also remained stable (reductions: -1% to -6%). This was expected given the low fluence applied.

DOC characterization by LC-OCD-OND are depicted in Figure SI 1 (b, e, h, and k). The figure revealed a change in the molecular weight of NOM components after VUV treatment. While the overall reduction in DOC was small, humics were strongly bleached (strong reduction in UV_{254} response). An increase in biodegradability is observed through the increase in BDOC measurements (from +20% to +35%. Figure SI.2) as the HMW fraction decreases.

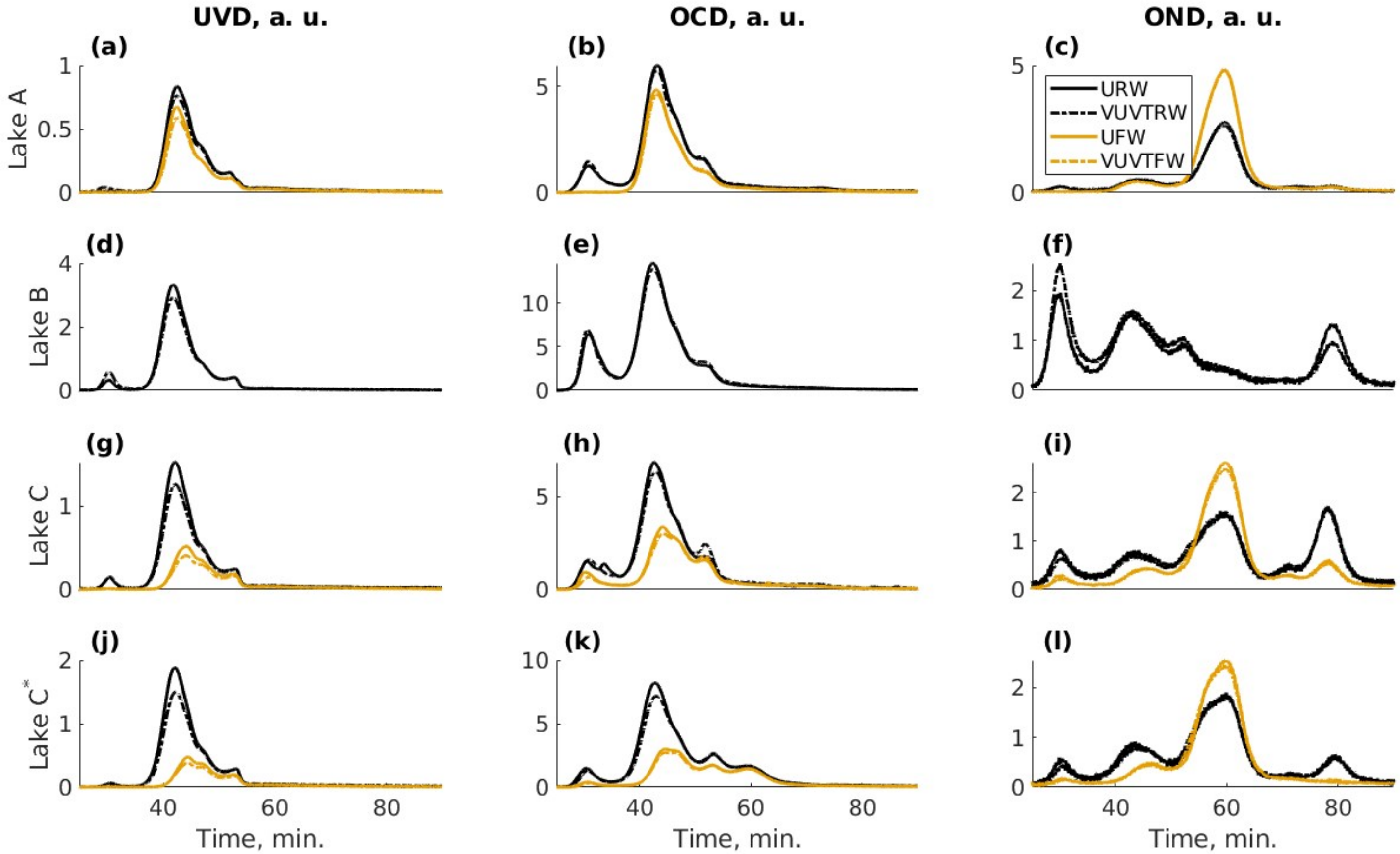

Figure SI. 1. LC-OCD-OND spectra (URW = Untreated Raw Water, VUVTRW = VUV Treated Raw Water, UFW = Untreated Filtered Water, UVTFW = VUV Treated Filtered Water, a.u. = arbitrary units)

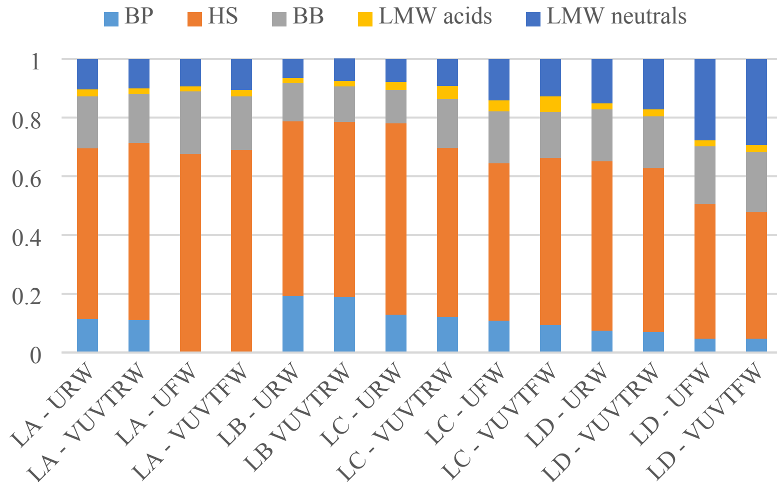

Figure SI. 2. Carbon fraction from LC-OCD-OND
Lake A, LB = Lake B, LC = Lake C, LD = Lake C, URW = untreated raw water, VUVRW = VUV treated raw water*

33

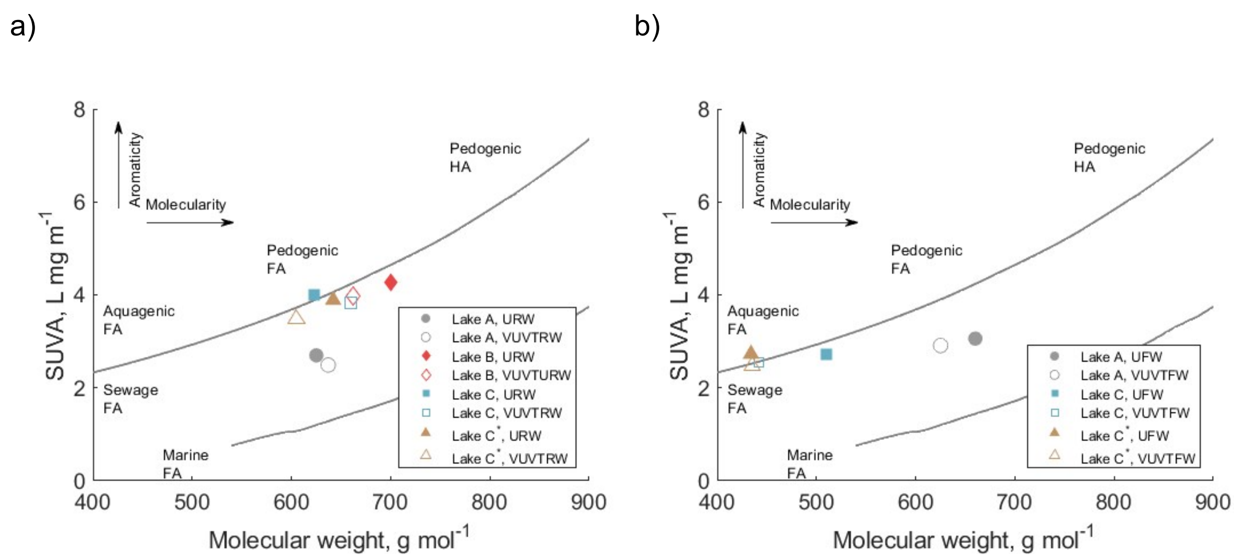

Figure SI. 3. Humic substances diagram (HS-diagram) before and after (retention time=9.4 sec) VUV treatment. a) Raw water: URW = untreated raw water, VUVTRW VUV treated raw water. b) Filtered water: UFW = untreated filtered water, VUVTFW VUV treated filtered water. Based on [2]

34 Biodegradability though biological dissolved organic carbon (BDOC) measurements

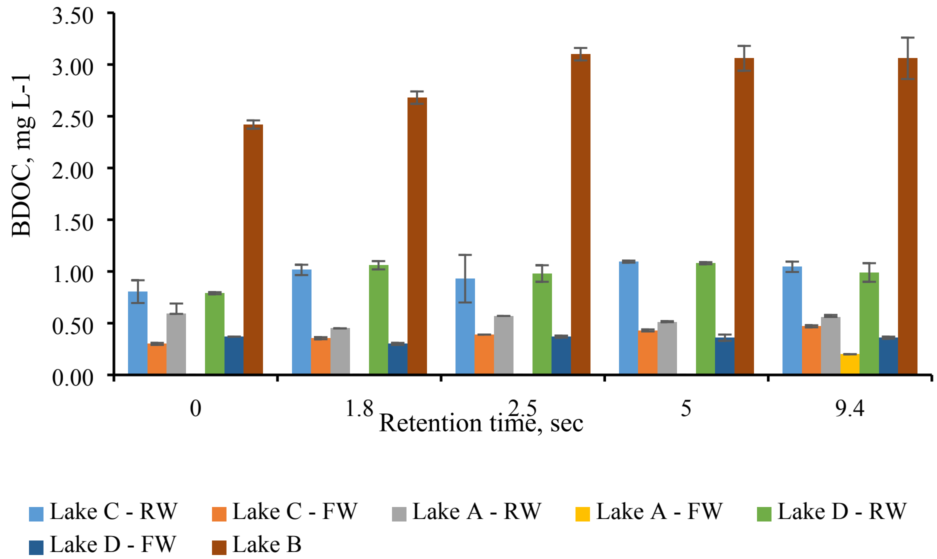

Figure SI. 4. BDCO before and after VUV treatment. RW = Raw Water, FW = Filtered Water

35 PARAFAC modeling

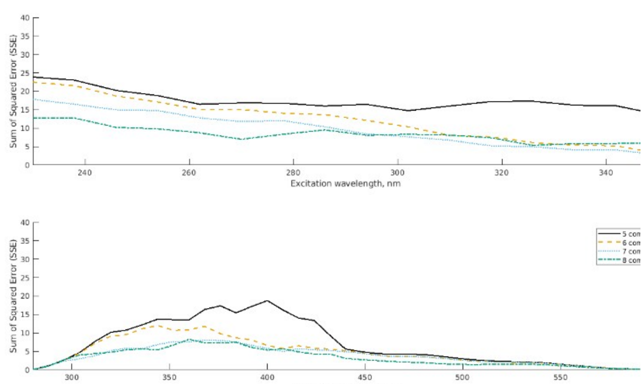

Figure SI. 5. SSE of PARAFAC models with different number of component (for supplementary information)

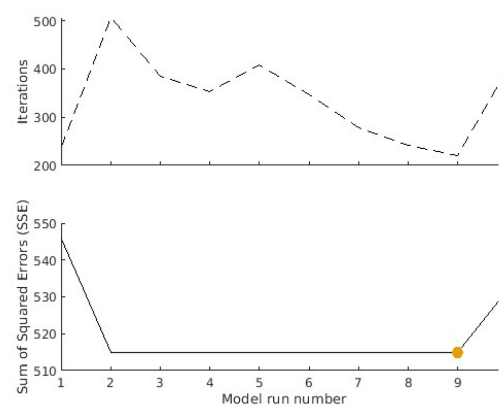

Figure SI. 6. Robustness of the 7-components PARAFAC model. Ten runs with random starting points (for supplementary information)

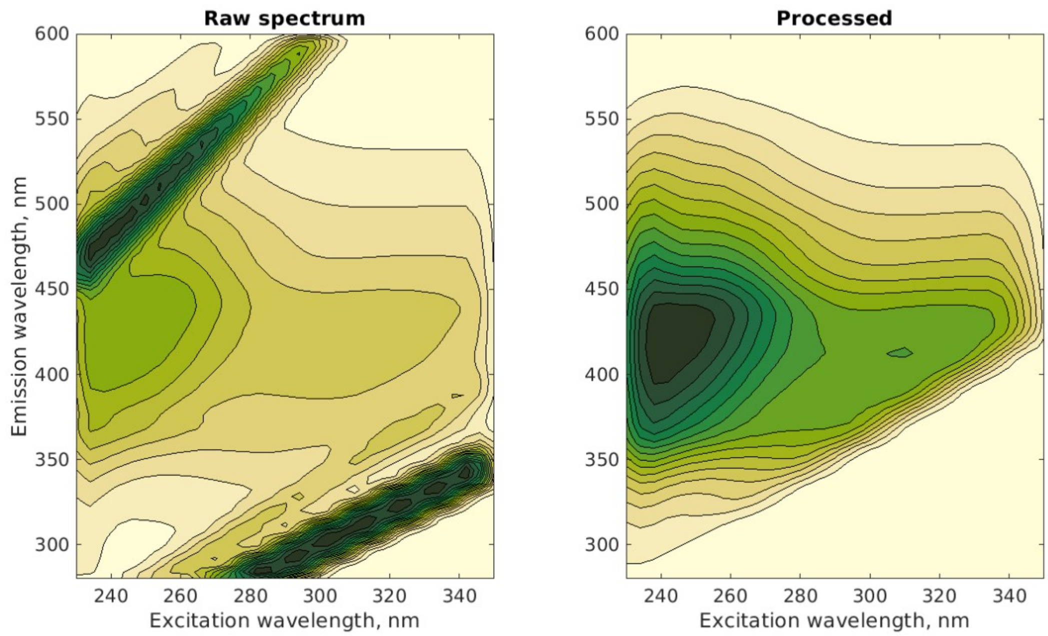

Figure SI. 7. Dataset pre-processing (for supplementary information)

36 PARAFAC Modeling components

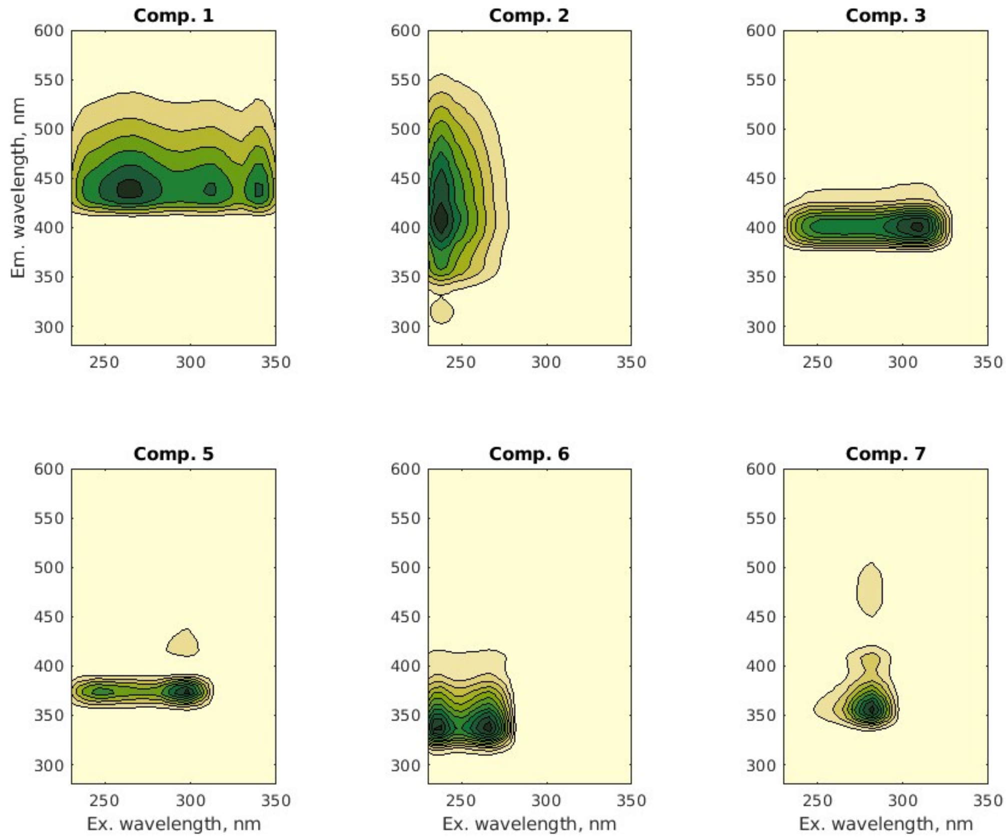

Figure SI. 8. Components obtained from PARAFAC modelling

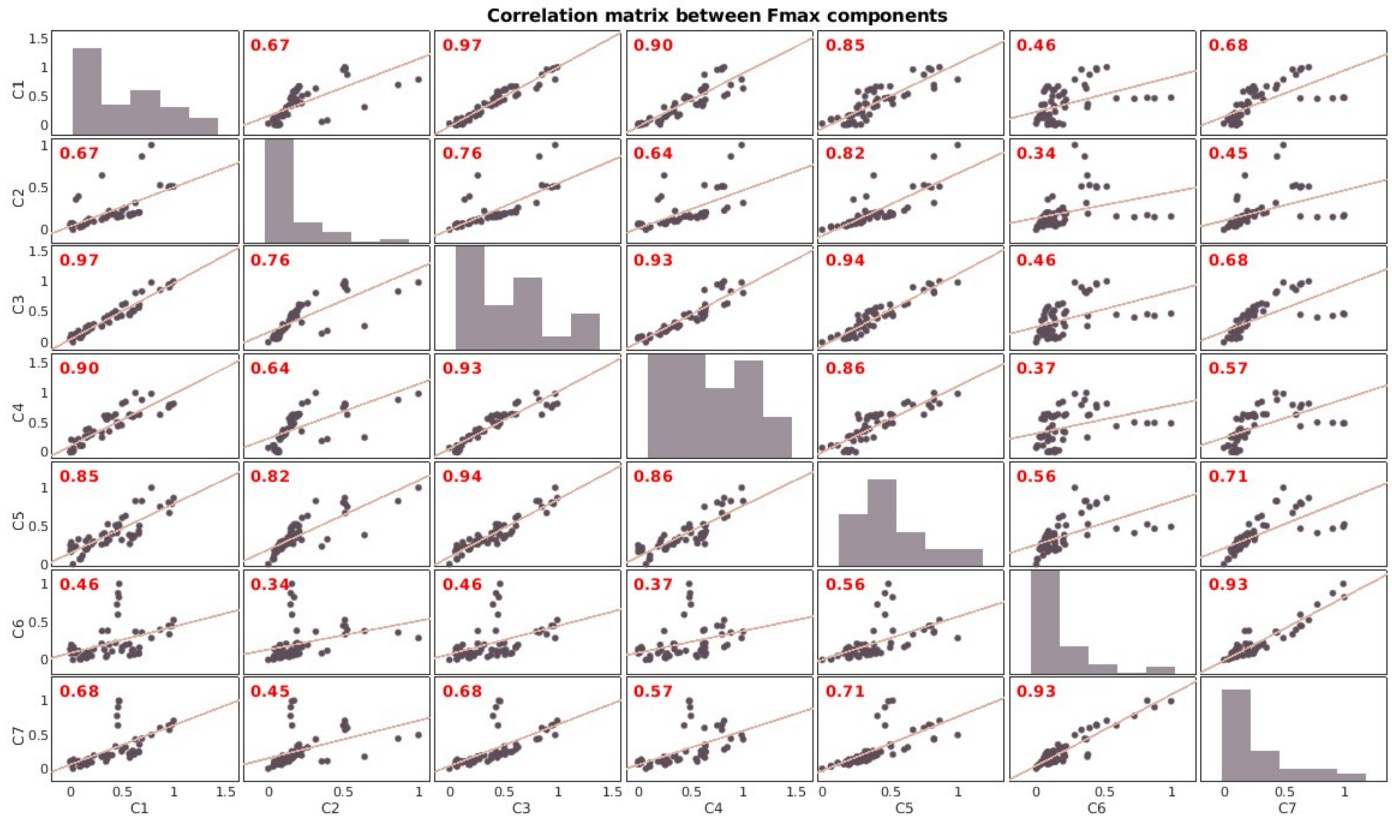

Figure SI. 9. F_{max} correlation between PARAFAC components

SOM modeling

SOM is a pattern recognition method. It clusterizes and reduces the dimensionality of input FEEM without make assumption about the data structure [3]. A dataset of 64 spectra were analyzed with the PARAllel FACtor analysis (PARAFAC) decomposition routines for Excitation Emission Matrices (drEEM, version 0.2.0 toolbox [4]) and self-organizing maps (SOM) with the SOM-Toolbox version 2.1 [5] both running on Matlab® R2018B software).

Figure SI 8 shows the U-matrix from the best SOM map selected along with the component planes of each of the components retrieved by PARAFAC. The U-matrix shows the distance between neighbouring neurons in the map, where large values (darker color) indicate highly dissimilar neurons. The component planes are useful to visualize the importance of different excitation-emission pairs or variables for each sample (neurons, from the SOM solution). We selected the pairs corresponding to the components proportions from PARAFAC (Figure SI.7).

The distribution of the samples obtained from the solution of SOM is shown through the best matching unit (BMU) graphics. In Figure SI 8 we show a combined image of the BMU overlapped with multiple hits histograms (colored hexagons) for each lake showing the importance of each neuron representing the lakes and the type of water (the larger the colored hexagon the higher number of hits in that neuron). The BMU is the best solution found by SOM and shows the distribution of the four lakes across the map for raw (R) and filtered (F) water. For example, samples associated to Lake B are located at different neurons in the map (the top-left and bottom-left in Figure SI 7), indicating the differences in NOM composition from raw and filtered. Basically, raw water contains NOM and AOM (particulate + dissolved) and filtered water only contains the dissolved fraction of NOM and AOM. Thus, SOM maps locate raw and filtered water in different region (Figure SI 7). Therefore, SOM allows discriminating differences in cell concentration between samples. We identify two clusters for Lake A: one (bottom-right) belongs to raw water from the lake while the other (middle) corresponds to filtered water through the well (Lake A supplies a bank filtration water system). Lakes C and C* show share similar neurons in the map, because they represent the same lake without (Lake C) and with (Lake C*) bloom conditions. For a given lake is seen that the SOM solution properly discriminate between raw and filtered water, particularly those with bloom condition (Lakes B and C*).

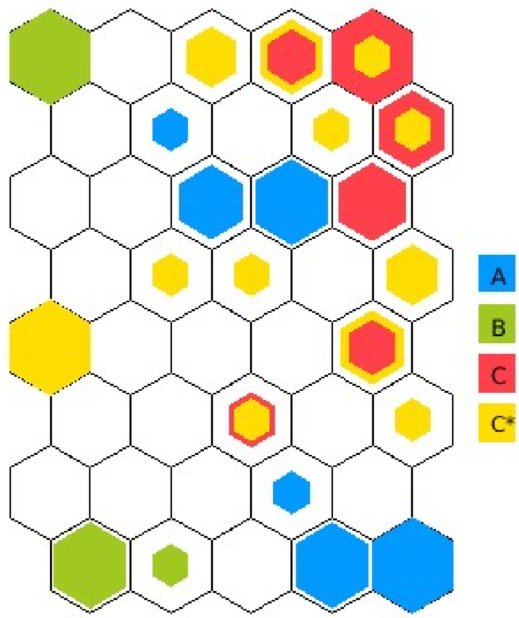

Figure SI. 10. SOM map with multiple hits overlapped showing the sites distribution for raw and filtered water.

By comparing Figure SI 7 and each component plane from Figure SI 8 is possible to correlate that PARAFAC components C1(1), C6(1), C6(2) and C7(1) characterize Lake B. Furthermore, C6(1), C6(2) and C7(1) are essentially due to samples from Lake B. This allows us to conclude that Lake B is mostly characterized by humic acids-like and (Region V, Figure 3) and soluble microbial by-product-like (Region IV, Figure 3). Component planes show that raw water in Lake A is associated to C1(1), C5(1), C5(2) and C7(2), while filtered water correlates to C1(2), C3(1), C4(1) and C7(2). Here also the regions that characterize the lake are V and IV (Figure 3). Lakes C and C* are characterized by C1(2), C1(3), C3(2), C4(1) and C7(2), lying in regions V and IV.

By overlapping PARAFAC and SOM results, it is possible to distinguish components that characterise each lake. All four lake could be characterize by a reduced number of regions (Figure 3): region IV(humic acidlike) and V (soluble microbial by-product-like). By focusing on this FEEM region, it could be possible to make a relationship with NOM / AOM and DBP's formation and reactivity.

Table SI. 2. SOM and PARAFAC. Characteristic component of each lake

Lake	PARAFAC Component
A	C1(1), C5(1), C5(2) and C7(2) (Raw water)
	C1(2), C3(1), C4(1) and C7(2) (filtered water)
B	C1(1), C6(1), C6(2) and C7(1)
C and C*	C1(2), C1(3), C3(2), C4(1) and C7(2)

Combination of PARAFAC components and SOM allowed to characterize each lake with fewer components:

- Components C1(1), C6(1), C6(2) and C7(1) characterize Lake B
- Raw water in Lake A is associated to C1(1), C5(1), C5(2) and C7(2), while filtered water correlates to C1(2), C3(1), C4(1) and C7(2)

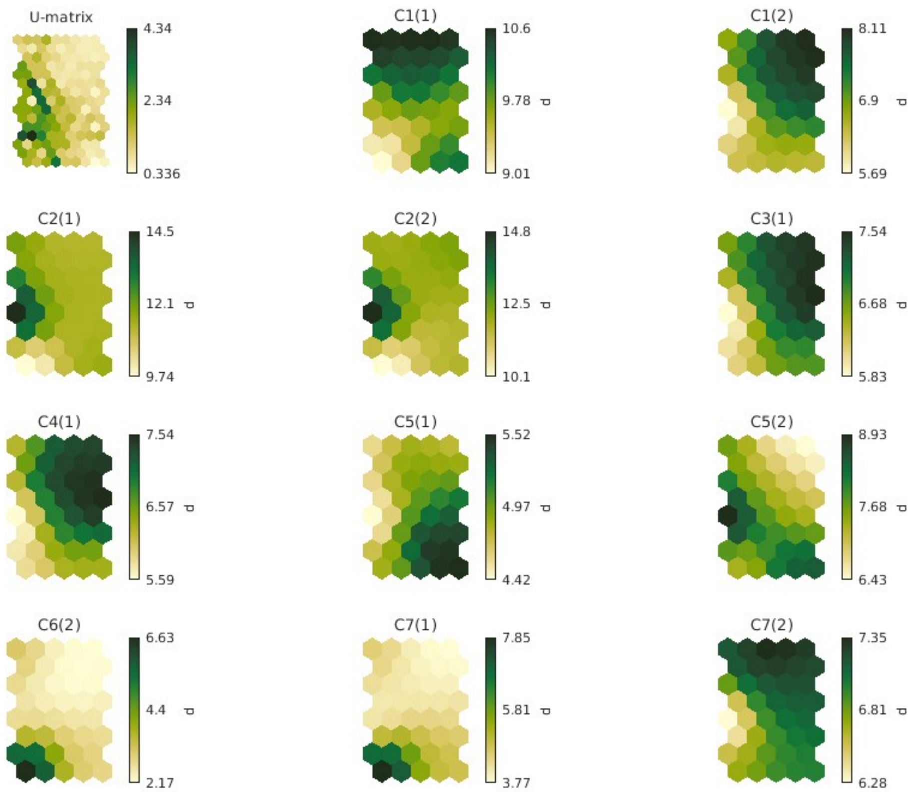

Figure SI. 11. U-matrix and component planes for the proportions of PARAFAC components resulted from SOM analysis

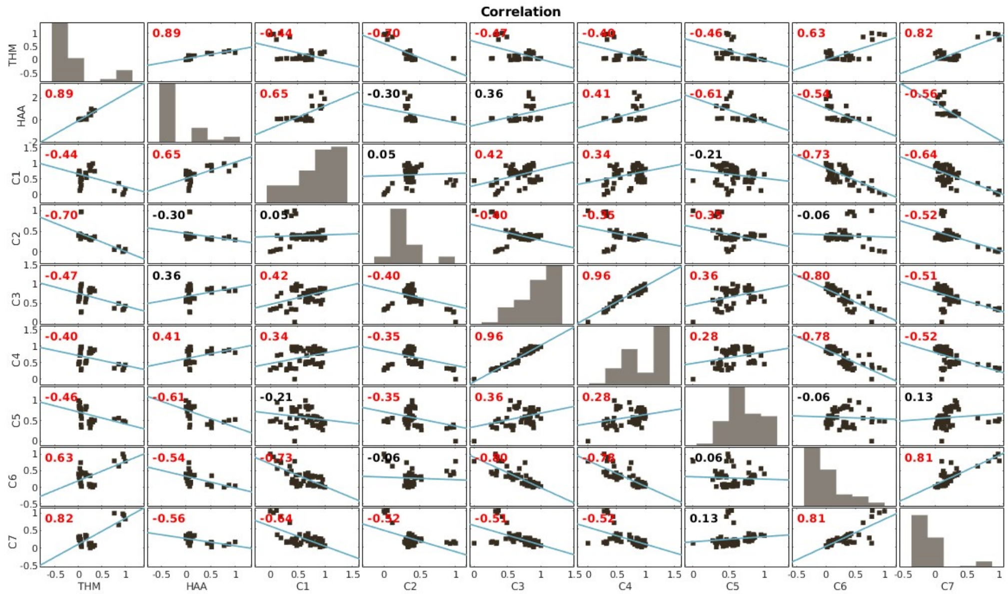

Figure SI. 12. THM and HAA correlation with PARAFAC compoments

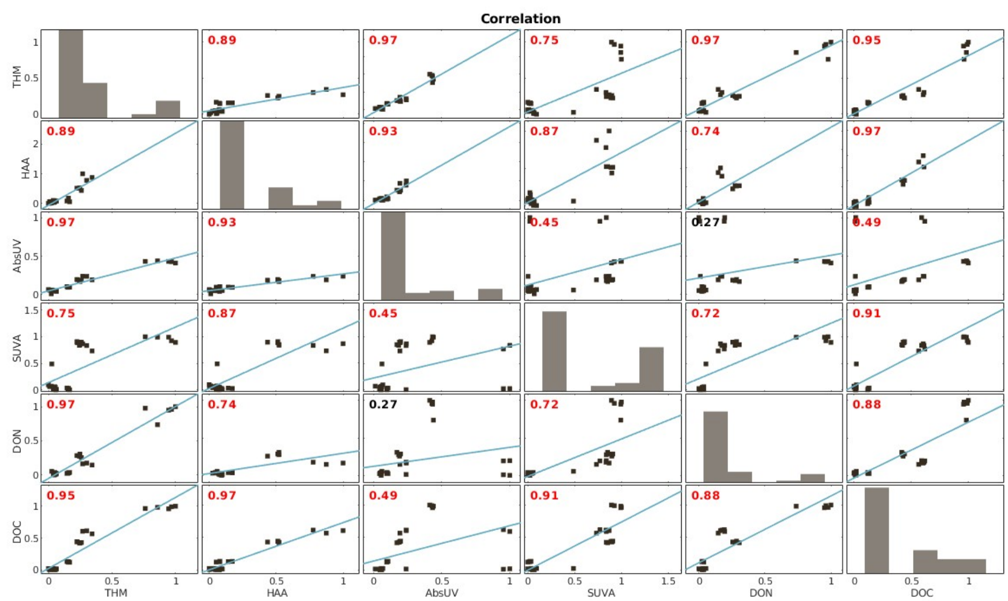

Figure SI. 13. THM and HAA correlation with UVA_{254}, SUVA, DON and DOC

Table SI. 4. Carbon fractions and yields values from DBP's formation correlation

	Yields, μg_F / $\mu g_{THM\ or\ HAA}$	
C or N fraction (F)	**THM**	**HAA**
DOC	27	43
DON	0.9	0.6
HS	20	28
BP	6.9	4.7
BB	3.8	7.2
LMW Ac	0.51	1.2
LMW Neutral	1.9	6

References

1. Edzwald, J.K. and J.E. Tobiason, *Enhanced coagulation: us requirements and a broader view.* Water Science and Technology, 1999. **40**(9): p. 63-70.
2. Huber, S.A., et al., *Characterisation of aquatic humic and non-humic matter with size-exclusion chromatography--organic carbon detection--organic nitrogen detection (LC-OCD-OND).* Water Res, 2011. **45**(2): p. 879-85.
3. Ejarque-Gonzalez, E. and A. Butturini, *Self-organising maps and correlation analysis as a tool to explore patterns in excitation-emission matrix data sets and to discriminate dissolved organic matter fluorescence components.* PLoS One, 2014. **9**(6): p. e99618.
4. Murphy, K.R., et al., *Fluorescence spectroscopy and multi-way techniques. PARAFAC.* Analytical Methods, 2013. **5**(23).
5. Vatanen, T., et al., *Self-organization andmissingvaluesinSOMandGTM.* Neurocomputing, 2015. **147**: p. 60-70.