A PARAFAC-based long-term assessment of DOM in a multi-coagulant drinking water treatment scheme

ARTICLE *in* ENVIRONMENTAL SCIENCE & TECHNOLOGY · JANUARY 2014

Impact Factor: 5.33 · DOI: 10.1021/es4049384 · Source: PubMed

3 AUTHORS, INCLUDING:

Nancy P. Sanchez
Rice University
20 PUBLICATIONS **93** CITATIONS
SEE PROFILE

Christopher M Miller
University of Akron
43 PUBLICATIONS **465** CITATIONS
SEE PROFILE

Environmental Science & Technology

Article

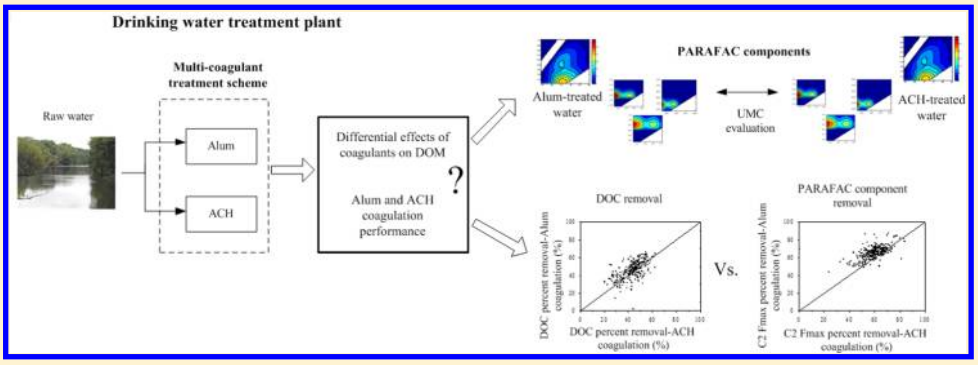
pubs.acs.org/est

A PARAFAC-Based Long-Term Assessment of DOM in a Multi-Coagulant Drinking Water Treatment Scheme

Nancy P. Sanchez,*[,†] Andrew T. Skeriotis, and Christopher M. Miller

Department of Civil Engineering, The University of Akron, Akron, Ohio 44325, United States

Supporting Information



ABSTRACT: A parallel factor (PARAFAC) analysis approach was used to study the character and composition of dissolved organic matter (DOM) in a multicoagulant (two aluminum-based coagulants) full scale drinking water treatment plant. A three year, long-term assessment was conducted based on deconstruction of the excitation−emission matrices (EEM) of over 1000 water samples collected before and after parallel coagulation treatment basins. Two humic moieties and a protein-like group were identified in the raw and treated waters. Apportionment of fluorophores was established using a novel approach based on the overall fluorescence intensity (OFI) of PARAFAC components. Uncorrected matrix correlation (UMC) revealed minimal changes of the fluorescence moieties after treatment (UMC > 0.98), and a comparable effect of both coagulants on the structure (UMC > 0.99) and distribution of these groups. Coagulation increased the proportion of the protein-like fluorophore and preferentially removed a humic-like group irrespective of the coagulant. Preference for this moiety was supported by a coagulant-affinity factor derived from the association between PARAFAC components after treatment. The suitability of a PARAFAC-based approach for coagulant evaluation/selection was demonstrated when compared to a dissolved organic carbon (DOC)-based criterion. This paper contributes to the understanding of the behavior of PARAFAC components in water treatment processes and presents several approaches for the future monitoring and control of coagulation at full scale treatment facilities.

INTRODUCTION

Coagulation followed by flocculation and clarification remains the preferred method for removal of natural organic matter (NOM) in conventional drinking water treatment systems.[1,2] The site-specific and complex character of NOM has been recognized as a major factor that determines the treatability of a specific water source,[3,4] and impacts the extent of formation of disinfection byproducts (DBPs) in chlorine-based disinfection.[5−8] Coagulation based on Al and Fe salts has been widely applied in drinking water treatment plants (DWTPs) in order to reduce turbidity, while removing DBP precursors associated with the dissolved fraction of NOM (DOM).[7,9,10] Coagulant dose and type, DOM nature, mixing conditions and pH have been identified as the main factors affecting the extent of NOM removal, which occurs by two main paths: adsorption to solid Al/Fe precipitates and charge neutralization/precipitation.[10−15]

Considering the role of DOM in postcoagulation treatment stages, understanding of the specific effect of coagulant application on the DOM pool of a particular water source becomes fundamental for any DWTP. Key pieces of information such as the presence of recalcitrant structures to coagulation treatment, variation of DOM composition after coagulation, and preferred removal of a particular DOM fraction should be determined in order to optimize coagulation in a specific treatment facility. Approaches for evaluation of the DOM−coagulation link are commonly based on reduction of total organic carbon (TOC) and ultraviolet absorbance at 254 nm (UV_{254}).[16] Although these parameters provide general insight, they only offer limited information on the impact of coagulation on the DOM pool and might not reflect variations on its character and reactivity.[17] More specific analytical techniques such as fluorescence spectroscopy and high performance size exclusion chromatography (HPSEC) have been investigated as alternatives for coagulation assessment in DWTPs. HPSEC has been stated as a potential alternative to

Received: August 28, 2013
Revised: January 9, 2014
Accepted: January 13, 2014
Published: January 13, 2014

   dx.doi.org/10.1021/es4049384 | Environ. Sci. Technol. 2014, 48, 1582−1591

routine DWTP process monitoring but it has primarily remained as a research tool.[18]

Fluorescence spectroscopy, which can capture the specific character of fluorescent DOM as a 3D excitation−emission matrix (EEM), appears as a simple and promising technique for characterization and tracking of DOM in treatment processes.[19] Parallel factor (PARAFAC) analysis,[20,21] a multiway decomposition method, has been considered as one of the most comprehensive approaches for analysis of spectral information in the EEMs.[22,23] PARAFAC has been used for assessment of water treatment processes in a limited number of studies at full scale DWTPs[17,22,24−27] and at laboratory scale.[28] The main findings of these studies have been compiled in a recent review on the use of fluorescence-PARAFAC analysis in drinking water treatment processes.[29]

A few reports on application of PARAFAC to evaluate the coagulation process in DWTPs in the United States have been published. These studies involved short-term collection of raw water samples followed by laboratory-based coagulation assessment[28] or included a limited number of samples reflecting a short period of operation.[26] Moreover, little research on the long-term behavior of PARAFAC components in engineered systems has been reported.[19] The association between PARAFAC components and the coagulation process was recently presented by our research group based on the analysis of a single treatment facility with Fe-based coagulation.[27] Although this work provided some insight on the effect of coagulation on the fluorophores present in the DOM, further analyses considering multiple coagulants are necessary in order to (i) establish if a general trend in the behavior of PARAFAC components can be observed and (ii) determine differential effects of distinct coagulants on DOM.

Here, for the first time, a full-scale multicoagulant drinking water treatment configuration involving the side-by-side application of aluminum sulfate (alum) and aluminum chlorohydrate (ACH) on the same water source is studied. Assessment of the effect of these coagulants on DOM quality and quantity included two major aspects: (i) a structure-oriented study intended to examine potential different impacts on the spectral characteristics of the PARAFAC fluorophore groups, and (ii) a quantitative analysis comprising the study of the distribution of the fluorophore moieties before and after treatment with each coagulant. The capability of a PARAFAC-based methodology to complement the TOC/DOC removal-based criterion conventionally used for evaluation of coagulants in DWTPs was examined. This paper also explored two novel approaches derived from PARAFAC sample loadings for the analysis of the coagulation process: (i) the overall fluorescence intensity (OFI) of the PARAFAC components to determine the composition of the fluorescent DOM, and (ii) a correlation analysis of the PARAFAC moieties in order to generate a coagulant-affinity factor, which provides insight on the removal of specific fluorophore groups by a particular coagulant.

MATERIALS AND METHODS

Sample Collection. Raw and clarified water samples after coagulation treatment (treated-alum and treated-ACH water) were obtained one to three times a week from the Akron DWTP serving the city of Akron, Ohio. Sampling was performed from November 2009 to December 2012. Akron DWTP treats an average flow around 130 ML/d having Lake Rockwell as its water source. The treatment train includes a preoxidation stage (i.e., potassium permanganate and/or chlorine dioxide addition), rapid mix, clarification, filtration, and chlorination. Side-by-side coagulant application including ACH was introduced in 2009. This parallel coagulation train had complete functionality during most of the period of study, being temporarily interrupted from April to October 2012, when only alum was applied. Dose of alum varied between 0.11 and 0.52 mM Al^{3+}, while ACH dose ranged between and 0.10 and 0.54 mM Al^{3+}. Raw water samples were obtained directly from Lake Rockwell without any pretreatment, while clarified samples were collected from the sedimentation basis after coagulation by alum and ACH.

Water samples were collected in Teflon-lined caps amber glass bottles, filtered through 0.45 mm Whatman nylon membrane filters and refrigerated at 4 °C for a maximum of three days until analysis of DOC concentration, pH, UV_{254} and fluorescence excitation−emission matrices.

Water Sample Characterization and Fluorescence Analysis. Concentration of the nonpurgeable DOC was determined using a TOC-5000A total carbon analyzer (Shimadzu, Japan) after pH adjustment of the samples (pH ∼ 2.0, 1 M HCl). UV_{254} was determined using a 1601 UV−visible spectrophotometer (Shimadzu, Japan) with a 1 cm quartz cell. Water sample pH was determined using an Accumet Basic AB15 pH meter (Fisher Scientific, Hampton, NH). DOC, UV_{254} and pH levels during the period of monitoring are presented in Supporting Information (SI) Table S1.

Fluorescence EEMs were recorded using an F-7000 fluorescence spectrophotometer (Hitachi, Japan). Excitation and emission wavelength ranges were selected as 204−404 nm and 290−550 nm, respectively. Photomultiplier detector voltage and scan speed were fixed at 400 V and 60 000 nm/min, respectively. Excitation and emission intervals were set at 5 and 2 nm and spectral corrections were applied according to manufacturer instructions. Sample preparation for fluorescence analysis involved adjustment of pH (∼3.0) and ionic strength and was presented in detail elsewhere.[27] Although DOM fluorescence intensity is impacted by levels of these parameters,[30] previous results suggest that fluorophores observed at different conditions of analysis (e.g., pH levels) exhibit a highly similar spectral character, indicating that comparison of components determined in diverse studies is feasible.[22,27] Blank signal was subtracted from each water sample EEM in order to remove Rayleigh and Raman scatter. Raman peak area of the DI scan at an excitation wavelength of 369 nm (emission range: 376−420 nm) was used to monitor the variation in the lamp intensity and for normalization of the fluorescence intensity to Raman Units (R.U). Consistency of the lamp signal was also examined by monitoring the fluorescence intensity (excitation/emission wavelength: 310/450 nm) of a quinine sulfate solution (7 mg/L in 0.1 M H_2SO_4). Raman peak area and intensity of the quinine sulfate solution exhibited relative standard deviation below 5% during the period of analysis.

Data Preprocessing and PARAFAC Application. Prior to PARAFAC application, missing values were inserted in the regions of Rayleigh and Raman scatter in the EEMs. Missing values were included covering 20 nm beyond the regions of scatter. A triangle of zeros was inserted in the region of longer excitation and shorter emission in order to increase calculation speeds (SI Figure S1).[21,31] Signal at excitation wavelengths below 224 nm was removed from each EEM in order to exclude regions commonly associated with high levels of noise.[21,32]

PARAFAC modeling was performed using the N-way v.3.00 Toolbox[33] and DOMFluor v.1.7 Toolbox[34] for MATLAB. Non-negative constraints were applied to the excitation, emission, and sample modes. PARAFAC models were independently fit on each raw, treated and combined (raw and treated) water sample data sets. Two additional PARAFAC models were fit on the differential EEM data sets obtained from $EEM_{raw} - EEM_{treated}$ for the alum and ACH treated waters (DF-alum and DF-ACH models respectively).[27] After preliminary PARAFAC application, observations with high residual and leverage values were classified as outliers and removed from the data set. Number of samples included in each PARAFAC model after removal of outliers is presented in SI Table S2.

Six different criteria were integrated in a decision matrix and used to select the adequate number of PARAFAC components in each data set: (i) analysis of convergence to a repeated solution after initialization from different random values, (ii) core consistency diagnostic (CORCONDIA),[35,36] (iii) explained variance (EV), (iv) meaningfulness of excitation and emission loadings, (v) sum of squared error of the integrated spectra on the excitation and emission side,[37] and (vi) validation of the PARAFAC components based on split-half analysis.[38]

RESULTS AND DISCUSSION

Fluorescence Components in Raw and Treated Water. Analysis of the criteria included in the decision matrix led to the selection of three fluorophore groups in the model fit on the raw water data set. CORCONDIA and EV for models including between one and six components, and contour plots of the retained components along with results of the validation based on split-half analysis for the selected model are presented in SI Table S3 and Figure S2, respectively. Fluorophore groups in the raw water exhibit high resemblance with moieties previously identified in Lake Hodgson, a Northeast Ohio reservoir whose fluorescence characteristics were studied in a previous publication.[27]

Component 1 (C1) shows two excitation maxima at 234 and 319 nm with a single emission maximum at 400 nm. This component has been identified as a humic-like fluorophore commonly present in freshwater environments.[26,37,39] Component 2 (C2) exhibits excitation maximum at a wavelength below 224 nm and a secondary excitation peak at 354 nm with maximum emission at 468/470 nm. This component, extensively reported in the literature has been classified as a humic-like fluorophore with terrestrial/allochthonous origin.[21,37,39] C1 and C2 were identified as reoccurring PARAFAC components present in a wide range of ecosystems in a recent review examining more than 50 fluorescence-PARAFAC studies.[19] Component 3 (C3) shows excitation maxima below 224 nm and at 284 nm, with an emission peak at 342 nm and resembles protein-like structures similar to tryptophan.[32,40]

Independent PARAFAC modeling was conducted on the data sets of water samples collected after alum and ACH-based coagulation. CORCONDIA and EV levels for models containing from one to six components are presented in SI Table S3. According to the evaluation criteria, three fluorophore groups were retained in each data set. Contour plots of the principal fluorophore moieties identified in the water samples after coagulation are presented in Figure 1.

Coagulation Effect on PARAFAC Components: Multi-Coagulant Analysis. Changes introduced by coagulation treatment were evaluated using the measure of uncorrected

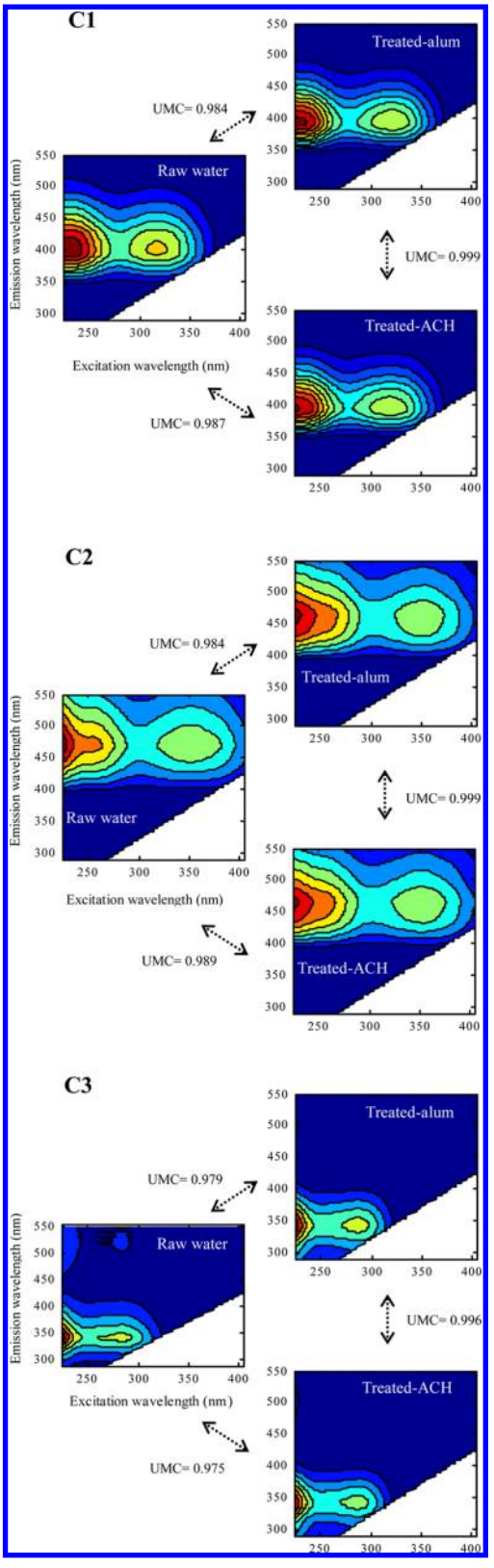


Figure 1. UMC-based evaluation of the impact of aluminum sulfate (alum) and aluminum chlorohydrate (ACH) on the PARAFAC components identified in Akron raw water.

matrix correlation (UMC). Unlike Tucker congruence coefficient (which is mainly formulated for factor comparison), UMC compares two matrices and establishes the degree of similarity between them, allowing a quantitative comparison of

the spectral characteristics of the components before and after treatment irrespective of variations in the intensity of their fluorescence signal. UMC values vary between 0 and 1 for matrices with null and complete spectral overlapping respectively. A detailed description of the mathematical formulation of UMC and its application for evaluation of EEMs has been presented in previous works.[27,41] UMC-based comparison (Figure 1) showed a highly similar character of the fluorescence structures before and after coagulation with alum and ACH. UMC values above 0.98 indicated that application of alum and ACH did not generate major changes in the structure of the humic-like fluorescence groups in the raw water DOM pool. UMC levels above 0.97 were obtained when the spectral characteristics of the component with protein-like nature were compared before and after coagulation.

Independent PARAFAC analysis on the raw and treated water data sets indicated that the same number of principal fluorophore groups could be validated in the DOM pool before and after treatment. This observation along with the high UMC values obtained for C1 to C3 demonstrates that new fluorescent moieties are not being formed after coagulant application and that a minimal structural change in the PARAFAC components is noticed when alum and ACH are used. These results, based on water samples collected over a three-year period strengthens previous conclusions reported by our research group after evaluation of a $FeCl_3$-based coagulation process.[27] Although results of this study are based on the analysis of a single DOM source/treatment facility and therefore might not represent a generalized phenomenon that can be extensive to DOM in multiple water sources, this is a three-year study and our results demonstrate: (i) minor variation in the structure of the PARAFAC components might be expected in the coagulation process, and (ii) decrease in the fluorescence signal of the DOM after coagulation can be potentially associated with physical removal of the fluorescence structures and not with chemical transformations being experienced by these moieties. Additionally, the results of this study further validate the suitability of application of PARAFAC on composite data sets comprising water samples before and after coagulation.

The specific effect of alum and ACH on the removal of the fluorescence structures in the parallel treatment train was also examined based on PARAFAC analysis of the differential EEMs. DF-alum and DF-ACH models were fit including from one to six components, and analysis of the different selection criteria indicated that three main fluorophore groups were adequate to describe the data in each model. EV and CORCONDIA values for the selected DF-alum and DF-ACH models were 99.1 and 82.5%, and 99.1 and 87.3%, respectively (SI Table S4). Contour plots of the retained components in the DF models are presented in SI Figure S3. PARAFAC models on differential EEMs can be considered as equivalent to a composition analysis of the DOM fluorescent fraction being removed by coagulation. Therefore, the PARAFAC components in the DF-alum and DF-ACH models correspond to the main fluorescent groups being impacted by each coagulant. A UMC-based comparison of the PARAFAC components identified in the DF models and the PARAFAC fluorophores in the raw water indicates a highly similar character for C1, C2, and C3 (UMC > 0.98). The presence of these three components in the DF models confirms that during the period of monitoring alum and ACH are removing all the fluorophore groups identified in the raw water. Despite potential differences in the removal levels attained for each component, both coagulants will affect the concentration of C1, C2, and C3. Highly preferential removal of a specific fluorophore group by alum or ACH would be indicated by the detection of only that particular moiety in the respective DF model. The absence of any of the raw water PARAFAC components in the DF-alum model would indicate a highly recalcitrant structure to be removed by alum, while the presence of the same fluorophore in the DF-ACH model would imply that when that moiety is present in the raw water, a selection of ACH as coagulant should be preferred. Average contribution of the components in the DF models to the total fluorescence maximum intensity ($\sum F_{max}$ of all the PARAFAC components identified in a sample) during the three-year monitoring was 34.3, 49.5, and 16.2%, and 32.9, 54.0, and 13.1% for the C1, C2, and C3 after alum and ACH treatment, respectively. This result indicates a predominance of C2 in the fraction of DOM being impacted by each coagulant, and shows this fluorophore group as the most amenable to be removed by both alum and ACH.

Further insight on the specific effect of alum and ACH on the structure of the PARAFAC moieties can be obtained by comparison of the PARAFAC components identified in the treated water samples (Figure 1). UMC values of ~1 indicate complete overlapping of the spectral character of the humic and protein-like fluorophores in the treated-alum and treated-ACH data sets. This observation, along with UMC values above 0.99 when components in the DF-alum and DF-ACH models are compared, demonstrates that in terms of structural variation, both coagulants have a comparable impact on the PARAFAC fluorophore groups in the raw water. This comparable and minor impact on the character of the components indicates that any decrease in the fluorescence intensity observed after alum and ACH treatment should be attributed to preferential removal of a particular fluorophore group, and not to variation/transformation of the fluorescence structures introduced by coagulant application. The observations presented here contribute to addressing documented research needs related to the association between water treatment processes and DOM fluorescence[19] and provide further insight on the behavior of PARAFAC components in engineered systems.

Coagulation Effect on DOM Pool. The effect of coagulant application on the fluorescent DOM was examined based on the variation in the contribution of the PARAFAC components to the total F_{max} and OFI in the water samples. F_{max} has been generally used as a surrogate to estimate the relative concentration of the PARAFAC components in the DOM pool, and the variation in the contribution of the PARAFAC components to the total F_{max} has been reported in order to evaluate the effect of coagulation and disinfection treatment on the fluorescent DOM.[17,22,28,42] As F_{max} does not reflect the multiple excitation maxima present in the configuration of the fluorophore groups identified by PARAFAC and does not account for the whole fluorescence signal exhibited by a fluorophore group, it is of interest to evaluate the contribution of the PARAFAC components before and after coagulation treatment based on different metrics that capture other spectral information. A volume-based distribution of the PARAFAC components was presented in a previous publication as an alternative to F_{max}-based contribution.[27] Although this methodology considers the entire fluorescence signal of the PARAFAC component; its calculation involves some additional steps to those required for F_{max} evaluation, and its units (i.e., nm) might not be meaningful. Here, we introduce the use of the OFI of

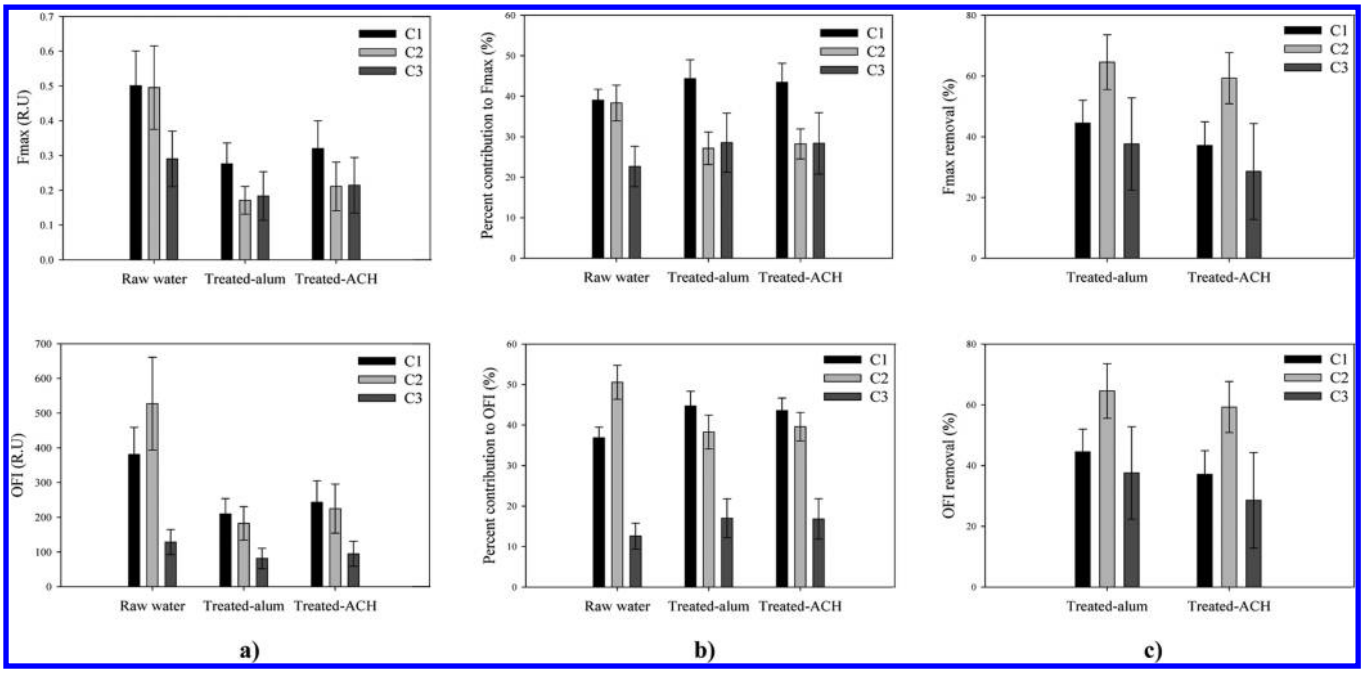


Figure 2. Contribution of the PARAFAC components to the total fluorescence maximum intensity (F_{max}) and overall fluorescence intensity (OFI), and average fluorescence removal after coagulation with alum and ACH during the sampling period comprised between November 2009 and December 2012 (N = 892). (a) F_{max} and OFI, (b) percent contribution to F_{max} and OFI, and (c) Percent F_{max} and OFI removal.

each PARAFAC component as an alternative approach. OFI corresponds to the total summation of the fluorescence intensities in an EEM expressed as R.U,[43] and has been used to determine the total fluorescence signal of a sample. In this study, OFI of a specific component in a particular sample is calculated as the summation of the intensity values of the matrix representing the PARAFAC component in the sample. A detailed description of the calculation of this parameter and the computation of the contribution of the PARAFAC components to the total OFI and F_{max} is presented in the SI (pages S15 to S17). Use of the OFI of PARAFAC components offers some advantages compared with the volumetric approach, including a rapid calculation and direct comparability with F_{max} (i.e., F_{max} and OFI are expressed in R.U).

Figure 2 shows the average F_{max} and OFI levels of the raw and treated waters during the three-year monitoring period. Average contribution of the PARAFAC components to F_{max} and OFI in the water samples before and after coagulation and average removal of these fluorescence structures during the sampling period are also shown. According to Figure 2, fluorophores with humic nature (C1 and C2) have a dominant contribution to the total F_{max} in the raw and treated water data sets. The aggregated average contribution of the humic-like PARAFAC components to F_{max} and OFI in the raw water was 77 and 87%, respectively. Unlike F_{max}-based results, contribution of C2 to OFI is predominant in the raw water samples (average ∼51%).

F_{max} and OFI-based analyses show a variation in the specific contribution of the components to the fluorescence signal after coagulation. Although key differences are observed in the relative contribution of the fluorophore groups when these approaches are used, both methodologies indicate a highly similar contribution of the PARAFAC components after alum and ACH-based treatment. This observation suggests that despite differences in the coagulant concentration and levels of fluorescence signal being removed by each coagulant, the impact of alum and ACH on the composition of the DOM pool is similar. This result concurs with our previous observations regarding similar effects of both coagulants on the fluorescence structures in the raw water.

F_{max}-based results show C1 as the fluorophore group with the highest contribution in the water after alum and ACH application (average ∼44% for C1 vs ∼28% for C2), while contribution to OFI indicates only a slightly higher presence of C1 in the water after treatment with both coagulants (average of 44 and 38%, and 44 and 40% for C1 and C2 in the alum and ACH-treated water, respectively). Contribution of PARAFAC components to F_{max} and OFI indicates a more relevant role of the protein-like moiety in the treated waters. F_{max}-based results show a higher contribution of this group with an average of 28% compared with 17% observed in an OFI-based analysis.

In general, differences between contribution of the PARAFAC components to F_{max} and OFI can be summarized as: (i) while C1 has the highest average contribution to the total F_{max}, C2 represents the highest percentage of OFI (ii) in terms of OFI, the protein-like fluorophore group shows a significantly lower contribution to the fluorescence signal in the raw and treated waters, leading to an aggregate contribution of the humic-like components to the OFI of around 90% in the water before and after coagulation, and (iii) C1 and C2 present a similar contribution to OFI in the treated waters, while F_{max}-based results indicate a predominant contribution of C1 in the water after treatment.

These results show that the evaluation of the contribution of PARAFAC components based on alternative metrics such as OFI can lead to fundamentally different observations regarding the effect of coagulation on the fluorescent DOM. Although the consistency and statistical robustness of OFI should be further evaluated, an important finding in this paper is that this divergence in the results generates questions regarding the suitability of only using F_{max} to this end,[28,42] and suggests that

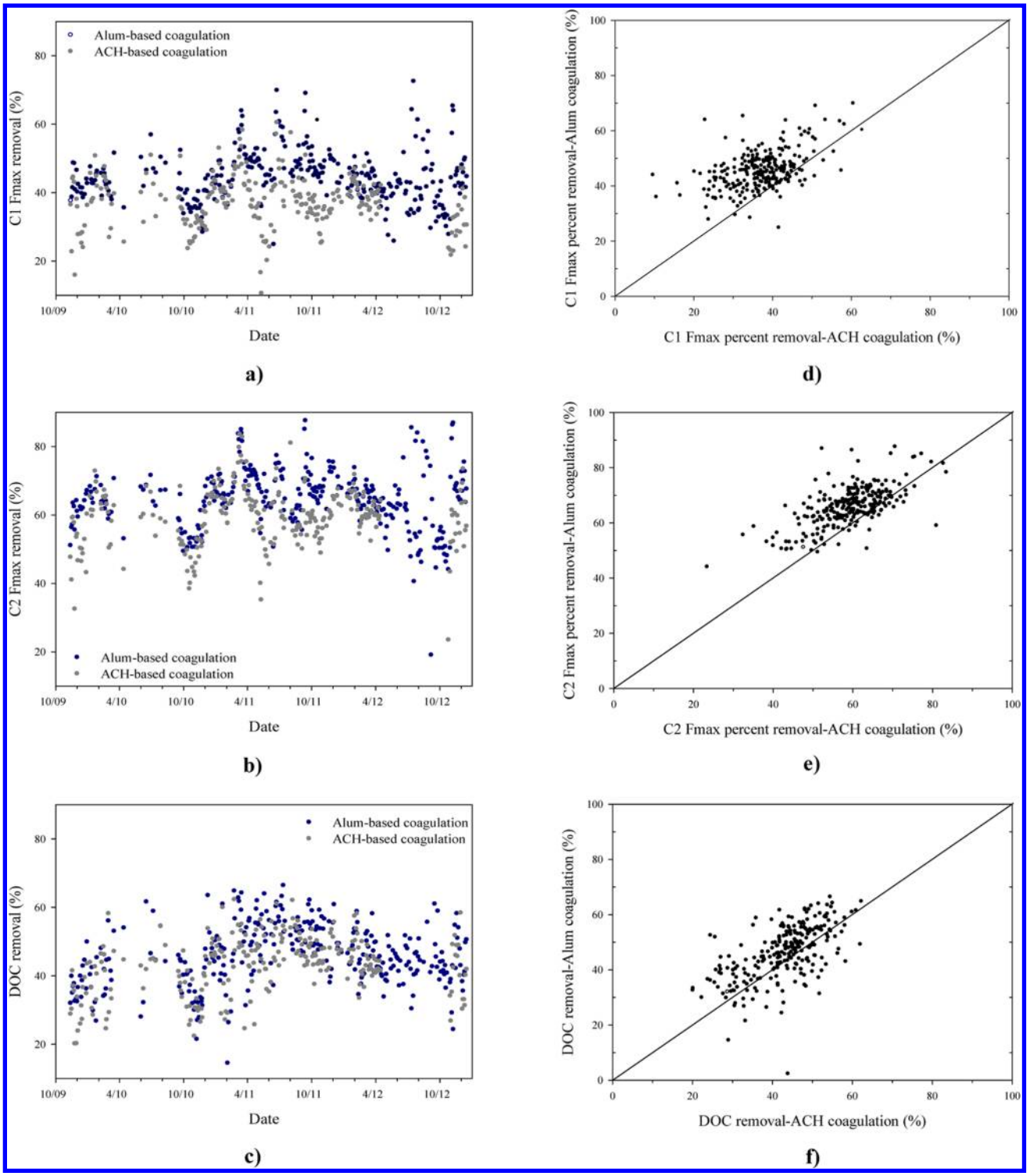


Figure 3. Fluorescence and DOC removal after alum and ACH-based coagulation during the sampling period (November 2009 and December 2012), (a) Fmax removal levels of PARAFAC component 1(C1), (b) F_{max} removal levels of PARAFAC component 2 (C2), (c) DOC removal levels. (d), (e), and (f) comparison of alum and ACH in C1, C2, and DOC removal, respectively. Line in (d), (e), and (f) indicates equal removal.

evaluation of alternative parameters such as OFI should also be introduced in this analysis.

Evaluation of removal percentages of PARAFAC components using these approaches leads to equivalent results, as F_{max} and OFI are extensions of the PARAFAC sample loadings (Figure 2). The highest average removal after coagulation was observed for C2 for both coagulants (average removal levels of 64 and 60% for alum and ACH treatment, respectively), confirming the results of the DF models. These results concur with observations from previous studies[17,22] and reflect the findings of a recent critical review in which this component was identified as the most affected by coagulant application when different recurrent PARAFAC structures were investigated.[19]

Seasonal Analysis of Coagulation Dynamics. SI Table S5 and figures S4 and S5 present the seasonal variation of the contribution of the PARAFAC components to the total F_{max}

dx.doi.org/10.1021/es4049384 | Environ. Sci. Technol. 2014, 48, 1582−1591

Table 1. Association between PARAFAC Components in Raw and Treated Water Samples Collected from Akron Drinking Water Treatment System

data set	C_x	C_y	slope K/K	intercept log (C/C)	number of samples	R^2 (eq 2)	r (correlation analysis based on F_{max})[a]
raw water	C1	C2	1.002	−0.155	364	0.60	0.80
	C1	C3	0.663	−0.332		0.31	0.48
	C2	C3	0.211	−0.422		0.06	0.22
treated water- alum	C1	C2	1.110	−0.156	355	0.87	0.93
	C1	C3	0.893	−0.270		0.44	0.65
	C2	C3	0.580	−0.310		0.37	0.56
treated water-ACH	C1	C2	1.184	−0.161	287	0.91	0.95
	C1	C3	0.849	−0.267		0.58	0.77

[a]Pearson correlation coefficient. Significant correlation at α=0.01.

before and after coagulation by alum and ACH. C1, C2, and C3 varied between 41 and 47%, 24 and 30%, and 24 and 34% in the treated-alum water, respectively, while these components ranged between 42 and 45%, 24 and 32%, and 23 and 33% in the treated-ACH water, respectively. Significant quarterly variation in the fluorescence signal of each component in the treated waters (SI Figure S4) was observed during the three-year period. A moderate association between precipitation levels and raw water fluorescence signal (total F_{max}, r = 0.70 and C2 F_{max}, r = 0.77) was observed, and important increases in the fluorescence signal of C1, C2, and C3 in the raw water were noticed during the 2010 and 2011 wet season (April–June). This increase was also observed in the alum and ACH-treated water. Despite variations in the fluorescence signal of the PARAFAC components, the contribution of the PARAFAC groups to the total F_{max} after coagulation remained mostly constant during the period of monitoring (variation between quarters not exceeding ∼5% for both alum and ACH).

As presented in SI Figure S6, preferential removal of C2 was observed during the three-year study for both coagulants, indicating that this tendency prevailed under a wide range of coagulant doses and raw water DOM quality. The removal trend C2 > C1 > C3 was consistent during the three-year monitoring for alum and ACH. Average quarterly reduction of F_{max} for C2 ranged between 58 and 71%, and 57 and 69% for alum and ACH-based coagulation, respectively. Removal of C1 and C3 after alum treatment varied between 38 and 52%, and 22 and 44%, respectively, while reductions between 35 and 44%, and 26 and 37% were observed for these components in the treated-ACH data set, respectively. The levels of removal of the protein-like component exhibited a high variability during the entire period of analysis for both coagulation treatments. C1 and C2 presented a more consistent behavior, with slightly higher variability in the removal levels of C1 after alum and ACH application.

SI Figures S6 and S7 indicate that in quarterly terms, average removal of C1, C2, and C3 was consistently higher after coagulation with alum, with major differences between treatments observed during 2011. As presented in SI Figure S6, the concentration of coagulant used in each treatment was comparable, with an average ratio of 1.1 for alum and ACH during the monitoring period, indicating a weak association between coagulant dose and levels of fluorescence removal in this specific DWTP.

DOC Removal vs PARAFAC-based Analysis for Coagulation Assessment. Average levels of DOC removal for the coagulation based on alum and ACH were 45.7 ± 9.1% and 43.1 ± 8.5%, respectively, during the three-year period. Although alum exhibited a slightly higher DOC removal during most of the period of monitoring, ACH showed to be more efficient at some intervals (Figure 3). In general, differences between both coagulants in terms of DOC removal were often comparable as reflected by the average percentages above.

Figure 3 shows the seasonal variability of the levels of reduction of C1, C2, and DOC observed after alum and ACH application. Figure 3c indicates a high overlapping when both coagulants are examined in terms of DOC removal. Although Figure 3f indicates a tendency of alum to attain higher reductions in DOC, the degree of coincidence between both coagulants suggests that comparison and subsequent selection of coagulants based only on DOC removal may obscure the inherent differences between alum and ACH on the DOM pool. This indicates that additional parameters would be required in order to conduct a more comprehensive evaluation.

Figure 3d and e show that unlike a DOC-based comparison, an evaluation based on PARAFAC components is more specific about the effect of alum and ACH in the DOM pool. Figure 3 shows that higher removals of C1 and C2 were observed when alum was used as coagulant. This figure also indicates that similar levels of reduction in DOC can be related to significantly different impacts on the DOM quality of the raw water. This observation is consistent with preliminary results (not shown) that indicate significant differences in the reactivity of DOM toward trihalomethane, haloacetic acid, and total organic halide formation after alum and ACH-based treatment; and emphasizes the necessity of more specific measurements of the effect of the coagulation process on the DOM character (e.g., reactivity). Highly specific information on changes introduced by coagulation in the molecular composition of DOM in drinking water treatment trains have been obtained by high resolution techniques such as Fourier transform ion cyclotron resonance mass spectrometry.[44−46] However, these techniques are laborious and the instrumental and personnel requirements might limit their practical application. Despite the lower resolution of PARAFAC-based analysis, results suggest that evaluation of coagulation based on this technique might constitute an appropriate alternative for the formulation of particular operation strategies in DWTPs in order to (i) remove specific components that are relevant to their treatment objectives (e.g., a moiety associated with high DBP formation[26,28]), (ii) define more specific coagulation goals, and (iii) make informed decisions about potential modifications in the operational conditions in the treatment train. Although analysis of EEMs by PARAFAC, unlike other more

sophisticated techniques,[44−46] provides an estimate of the composition of the fluorescent DOM that allows quantitative determination of the effect of coagulation, it should be recognized that lack of information on the number of fluorescent isomers and the quantum yield of each PARAFAC component limits the potential determination of the mass-based concentration of each fluorophore entity in the DOM.[17,46]

Association Between PARAFAC Components Before and After Coagulation. The degree of correlation between PARAFAC components before and after treatment was examined (Table 1). Correlation analysis using F_{max} levels in the treated water showed the trend C1−C2 > C1−C3 > C2−C3 for the correlation levels. A significant increase in the linear association between components was observed after coagulation. The coefficients of correlation between C1 and C2 varied from 0.80 to 0.93 and 0.95 after application of alum and ACH, respectively. This behavior was also observed for the association between humic-like and protein-like components. Coefficients of correlation between C1−C3 and C2−C3 varied from 0.48 to 0.65 and 0.77, and from 0.22 to 0.56 and 0.70 in the raw and treated-alum and treated-ACH water, respectively. This increase in the linear association between PARAFAC components suggests the existence of a common controlling treatment-related factor that determines the concentration of the fluorophore groups after coagulation. Based on a previous approach presented by Stedmon and Markager,[37] the concentration of each PARAFAC component can be represented by an exponential function including a parameter D, that corresponds to a factor controlling its concentration. According to this approach, the concentration of a generic component x (C_x), can be represented by eq 1.[37]

$$C_x = C_{Xo}e^{KxD} \quad (1)$$

Where C_{xo} indicates the concentration of the component x when $D = 0$ and K_x is the rate of change of C_x with respect to D.

When the relationship between two components (e.g., C_x and C_y) is examined assuming that D is a common factor for these, the function presented in eq 2 is obtained.[37]

$$\log C_y = \log\left(\frac{C_{yo}}{C_{xo}}\right) + \frac{K_y}{K_x}(\log C_x) \quad (2)$$

Using F_{max} as a surrogate of the concentration of the PARAFAC components, eq 2 indicates that a linear relationship between the logarithms of F_{max} of C_x and C_y should be obtained if a common factor is affecting the occurrence of these components. The slope in eq 2 (K_y/K_x) corresponds to the ratio between the rate of removal of the components due to the factor D, and the intercept indicates the logarithm of the ratio of the concentration of the components x and y in the water before treatment ($D = 0$). The analysis presented by Stedmon and Markager was focused on multiple natural systems where D was associated to parameters such as the temperature. Here, D is assumed to be a treatment-related factor (i.e., coagulant type and dose) whose effect is to remove the fluorophore structures leading to a decrease in their concentration.

The results of the linear regression (eq 2) between each pair of PARAFAC components in the treated waters indicate a trend C1−C2 > C1−C3 > C2−C3 for the levels of association between the moieties (coefficients of determination in Table 1), suggesting that the treatment-related factor was particularly affecting the concentration of humic-like components.

Considering that K_x and K_y represent the rate of change of C_x and C_y associated with the application of coagulant, these can be considered as coagulant-affinity factors. The ratio of K_x and K_y in the case of the PARAFAC components with humic nature (x = C1 and y = C2) showed similar values for both coagulants (Table 1). According to these results, $K_{C2} = 1.11\ K_{C1}$ and $K_{C2} = 1.18\ K_{C1}$ for alum and ACH, respectively. This indicates that the coagulant-affinity of C2 for alum and ACH (i.e., the rate of change of C2 associated with the treatment factor) is 11 and 18% higher than the degree of affinity exhibited by C1 toward these coagulants respectively. This result concurs with the preferential removal of C2 observed in the seasonal analysis (SI Table S5).

When the association between humic and protein-like PARAFAC components is examined, lower coefficients of determination than those obtained for C1 and C2 are observed (Table 1), suggesting that despite the existence of a common treatment-related factor, coagulation has a substantially different impact on the concentration of the fluorophores with protein and humic character. This observation is consistent with previous reports on preferential removal of humic-like structures by coagulation.[22,27,47,48]

Environmental Relevance. This study constitutes the first approach to evaluate a three-year, long-term, multicoagulant treatment (i.e., engineered system) configuration using PARAFAC. It demonstrated that the use of a fluorescence-PARAFAC analysis can evaluate specific coagulation goals according to the treatment objectives and account for the dynamic character of DOM in water sources. Results also suggest that a minor impact on the structure of the PARAFAC components in the raw water and physical removal with no further transformations might be expected in the coagulation process irrespective of the type of coagulant being applied. Two different aluminum-based coagulants impact the same fraction of fluorescent DOM in the water and lead to highly similar contribution of the PARAFAC components to the total F_{max} and OFI, suggesting that differences in treatment by each coagulant are limited to variations in the removal levels of each fluorescence entity. This result further validates the application of PARAFAC including water samples before and after coagulation, and suggests that application of this procedure on data sets likely including chemical transformation of the DOM should first evaluate aspects such as the degree of modification that a treatment should introduce to a fluorophore before PARAFAC resolution leads to its identification as a different component. This paper questions the exclusive use of F_{max} to study the variation in the contribution of the PARAFAC component after coagulation treatment, and offers further evidence related to the higher impact of coagulation on the levels of a humic-like component (C2) extensively reported in freshwater sources irrespective of the coagulant being applied, the DOM quality and quantity, and the specific operational conditions. Furthermore, the analyses conducted demonstrate for the first time, that a fluorescence-PARAFAC approach can enhance the traditional TOC/DOC- based criterion used in DWTPs for evaluation of a particular coagulant, providing better insight on its effect on the different fractions of DOM.

dx.doi.org/10.1021/es4049384 | Environ. Sci. Technol. 2014, 48, 1582−1591

■ ASSOCIATED CONTENT

Supporting Information

The results presented in this paper are supported by five tables, seven figures, and a calculation example of the PARAFAC component contribution to F_{max} and OFI. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: (713) 348-3036; fax: (713) 348-5268; e-mail: nps1@rice.edu.

Present Address

†Department of Civil and Envnironmental Engineering, Rice University, 6100 Main St, Houston, Texas 77005, United States.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Matilainen, A.; Gjessing, E. T.; Lahtinen, T.; Hed, L.; Bhatnagar, A.; Sillanpää, M. An overview of the methods used in the characterisation of natural organic matter (NOM) in relation to drinking water treatment. *Chemosphere* **2011**, *83* (11), 1431−1442.

(2) Xu, W.; Gao, B.; Wang, Y.; Yue, Q.; Ren, H. Effect of second coagulant addition on coagulation efficiency, floc properties and residual Al for humic acid treatment by Al13 polymer and polyaluminum chloride (PACl). *J. Hazard. Mater* **2012**, *215−216* (0), 129−137.

(3) Fabris, R.; Chow, C. W. K.; Drikas, M. Comparison of coagulant type on natural organic matter removal using equimolar concentrations. *Aqua* **2012**, *61* (4), 210−219.

(4) Pelekani, C.; Newcombe, G.; Snoeyink, V. L.; Hepplewhite, C.; Assemi, S.; Beckett, R. Characterization of natural organic matter using high performance size exclusion chromatography. *Environ. Sci. Technol.* **1999**, *33* (16), 2807−2813.

(5) Hua, G.; Reckhow, D. A. Characterization of disinfection byproduct precursors based on hydrophobicity and molecular size. *Environ. Sci. Technol.* **2007**, *41* (9), 3309−3315.

(6) Xu, Z.; Jiao, R.; Liu, H.; Wang, D.; Chow, C.; Drikas, M. Hybrid treatment process of using MIEX and high performance composite coagulant for DOM and bromide removal. *J. Environ. Eng.* **2012**, *139* (1), 79−85.

(7) Krasner, S. W.; Amy, G. Jar-test evaluations of enhanced coagulation. *J. Am. Water Works Assoc.* **1995**, *87* (10), 93−107.

(8) Chen, C.; Zhang, X.-j.; Zhu, L.-x.; Liu, J.; He, W.-j.; Han, H.-D. Disinfection by-products and their precursors in a water treatment plant in North China: Seasonal changes and fraction analysis. *Sci. Total Environ.* **2008**, *397* (1−3), 140−147.

(9) Rebhun, M.; Lurie, M. Control of organic matter by coagulation and floc separation. *Water Sci. Technol.* **1993**, *27* (11), 1−20.

(10) Gregory, J.; Duan, J. Hydrolyzing metal salts as coagulants. *Pure Appl. Chem.* **2001**, *73* (12), 2017−2026.

(11) Bell-Ajy, K.; Abbaszadegan, M.; Ibrahim, E.; Verges, D.; LeChevallier, M. Conventional and optimized coagulation for NOM removal. *J. Am. Water Works Assoc.* **2000**, *92* (10), 44−58.

(12) Matilainen, A.; Vepsäläinen, M.; Sillanpää, M. Natural organic matter removal by coagulation during drinking water treatment: A review. *Adv. Colloid Interface Sci.* **2010**, *159* (2), 189−197.

(13) Dempsey, B. A.; Sheu, H.; Ahmed, T. M. T.; Mentink, J. Polyaluminum chloride an alum coagulation of clay-fulvic acid suspensions. *J. Am. Water Works Assoc.* **1985**, *77* (3), 74−80.

(14) Dempsey, B. A.; Ganho, R. M.; O'Melia, C. R. The coagulation of humic substances by means of aluminum salts. *J. Am. Water Works Assoc.* **1984**, *76* (4), 141−150.

(15) Van Benschoten, J. E.; Edzwald, J. K. Chemical aspects of coagulation using aluminum salts—I. Hydrolytic reactions of alum and polyaluminum chloride. *Water Res.* **1990**, *24* (12), 1519−1526.

(16) Cheng, W. P.; Chi, F. H.; Yu, R. F. Evaluating the efficiency of coagulation in the removal of dissolved organic carbon from reservoir water using fluorescence and ultraviolet photometry. *Environ. Monit. Assess.* **2004**, *98* (1), 421−431.

(17) Baghoth, S. A.; Sharma, S. K.; Guitard, M.; Heim, V.; Croué, J.-P. Removal of NOM-constituents as characterized by LC-OCD and F-EEM during drinking water treatment. *Aqua* **2011**, *60* (7), 412−424.

(18) Ho, L.; Hainthaler, M.; Newcombe, G. Using UV spectroscopy and molecular weight determinations to investigate the effect of various water treatment processes on NOM removal: Australian case study. *J. Environ. Eng.* **2013**, *139* (1), 117−126.

(19) Ishii, S. K. L.; Boyer, T. H. Behavior of reoccurring PARAFAC components in fluorescent dissolved organic matter in natural and engineered systems: A critical review. *Environ. Sci. Technol.* **2012**, *46* (4), 2006−2017.

(20) Ohno, T.; Bro, R. Dissolved organic matter characterization using multiway spectral decomposition of fluorescence landscapes. *Soil Sci. Soc. Am. J.* **2006**, *70* (6), 2028−2037.

(21) Stedmon, C. A.; Markager, S.; Bro, R. Tracing dissolved organic matter in aquatic environments using a new approach to fluorescence spectroscopy. *Mar. Chem.* **2003**, *82* (3−4), 239−254.

(22) Baghoth, S. A.; Sharma, S. K.; Amy, G. L. Tracking natural organic matter (NOM) in a drinking water treatment plant using fluorescence excitation−emission matrices and PARAFAC. *Water Res.* **2011**, *45* (2), 797−809.

(23) Luciani, X.; Mounier, S.; Redon, R.; Bois, A. A simple correction method of inner filter effects affecting FEEM and its application to the PARAFAC decomposition. *Chemometr. Intell. Lab.* **2009**, *96* (2), 227−238.

(24) Bieroza, M.; Baker, A.; Bridgeman, J. Classification and calibration of organic matter fluorescence data with multiway analysis methods and artificial neural networks: An operational tool for improved drinking water treatment. *Environmetrics* **2011**, *22* (3), 256−270.

(25) Bieroza, M.; Baker, A.; Bridgeman, J. New data mining and calibration approaches to the assessment of water treatment efficiency. *Adv. Eng. Softw.* **2012**, *44* (1), 126−135.

(26) Johnstone, D. W.; Sanchez, N. P.; Miller, C. M. Parallel factor analysis of excitation−emission matrices to assess drinking water disinfection byproduct formation during a peak formation period. *Environ. Eng. Sci.* **2009**, *26* (10), 1551−1559.

(27) Sanchez, N. P.; Skeriotis, A. T.; Miller, C. M. Assessment of dissolved organic matter fluorescence PARAFAC components before and after coagulation−Filtration in a full scale water treatment plant. *Water Res.* **2013**, *47* (4), 1679−1690.

(28) Pifer, A. D.; Fairey, J. L. Improving on SUVA254 using fluorescence-PARAFAC analysis and asymmetric flow-field flow fractionation for assessing disinfection byproduct formation and control. *Water Res.* **2012**, *46* (9), 2927−2936.

(29) Markechová, D.; Tomková, M.; Sádecká, J. Fluorescence excitation-emission matrix spectroscopy and parallel factor analysis in drinking water treatment: A review. *Pol. J. Environ. Stud.* **2013**, *22* (5), 1289−1295.

(30) Senesi, N. Molecular and quantitative aspects of the chemistry of fulvic acid and its interactions with metal ions and organic chemicals: Part II. The fluorescence spectroscopy approach. *Anal. Chim. Acta* **1990**, *232* (0), 77−106.

(31) Thygesen, L. G.; Rinnan, Å.; Barsberg, S.; Møller, J. K. S. Stabilizing the PARAFAC decomposition of fluorescence spectra by insertion of zeros outside the data area. *Chemom. Intell. Lab. Syst.* **2004**, *71* (2), 97−106.

(32) Yamashita, Y.; Tanoue, E. Chemical characterization of protein-like fluorophores in DOM in relation to aromatic amino acids. *Mar. Chem.* **2003**, *82* (3−4), 255−271.

(33) Andersson, C. A.; Bro, R. The N-way Toolbox for MATLAB. *Chemom. Intell. Lab. Syst.* **2000**, *52* (1), 1−4.

(34) Stedmon, C. A.; Bro, R. Characterizing dissolved organic matter fluorescence with parallel factor analysis: A tutorial. *Limnol. Oceanogr. Methods* **2008**, *6*, 572−579.
(35) Bro, R. Multi-Way Analysis in the Food Industry. Models, Algorithms, and Applications. PhD Thesis, Royal Veterinary and Agricultural University, Denmark, 1998.
(36) Bro, R.; Kiers, H. A. L. A new efficient method for determining the number of components in PARAFAC models. *J. Chemom.* **2003**, *17* (5), 274−286.
(37) Stedmon, C. A.; Markager, S. Resolving the Variability in dissolved organic matter fluorescence in a temperate estuary and its catchment using PARAFAC analysis. *Limnol. Oceanogr.* **2005**, *50* (2), 686−697.
(38) Harshman, R. A.; DeSarbo, W. S., An application of PARAFAC to a small sample problem, demonstrating preprocessing, orthogonality constraints, and split-half diagnostic techniques. In *Research Methods for Multi-Mode Data Analysis*; Law, H. G., Snyder, C. W., Hattie, J. A., McDonald, R. P., Eds.; Praeger: New York, 1984; pp 602−642.
(39) Kowalczuk, P.; Durako, M. J.; Young, H.; Kahn, A. E.; Cooper, W. J.; Gonsior, M. Characterization of dissolved organic matter fluorescence in the South Atlantic Bight with use of PARAFAC model: Interannual variability. *Mar. Chem.* **2009**, *113* (3−4), 182−196.
(40) Yamashita, Y.; Jaffé, R. Characterizing the Interactions between trace metals and dissolved organic matter using excitation−emission matrix and parallel factor analysis. *Environ. Sci. Technol.* **2008**, *42* (19), 7374−7379.
(41) Burdick, D. S.; Tu, X. M. The wavelength component vectorgram: A tool for resolving two-component fluorescent mixtures. *J. Chemom.* **1989**, *3* (2), 431−441.
(42) Granderson, C. W.; Pifer, A. D.; Fairey, J. An improved chloroform surrogate for chlorine dioxide and alum-treated waters. *J. Am. Water Works Assoc.* **2013**, *105* (3), E103−E114.
(43) Beggs, K. M. H.; Summers, R. S.; McKnight, D. M. Characterizing chlorine oxidation of dissolved organic matter and disinfection by-product formation with fluorescence spectroscopy and parallel factor analysis. *J. Geophys. Res.* **2009**, *114* (G4), G04001.
(44) Lavonen, E. E.; Gonsior, M.; Tranvik, L. J.; Schmitt-Kopplin, P.; Köhler, S. J. Selective chlorination of natural organic matter: Identification of previously unknown disinfection byproducts. *Environ. Sci. Technol.* **2013**, *47* (5), 2264−2271.
(45) Zhang, H.; Zhang, Y.; Shi, Q.; Hu, J.; Chu, M.; Yu, J.; Yang, M. Study on transformation of natural organic matter in source water during chlorination and its chlorinated products using ultrahigh resolution mass spectrometry. *Environ. Sci. Technol.* **2012**, *46* (8), 4396−4402.
(46) Zhang, H.; Zhang, Y.; Shi, Q.; Ren, S.; Yu, J.; Ji, F.; Luo, W.; Yang, M. Characterization of low molecular weight dissolved natural organic matter along the treatment trait of a waterworks using Fourier transform ion cyclotron resonance mass spectrometry. *Water Res.* **2012**, *46* (16), 5197−5204.
(47) Beggs, K. M. H. Characterizing temporal and spatial variability of watershed dissolved organic matter and disinfection byproduct formation with fluorescence spectroscopy; University of Colorado: Boulder, 2010.
(48) Beggs, K. M. H.; Summers, R. S. Character and chlorine Reactivity of dissolved organic matter from a mountain pine beetle impacted watershed. *Environ. Sci. Technol.* **2011**, *45* (13), 5717−5724.

Supporting Information

A PARAFAC-Based Long-Term Assessment of DOM in a Multi-Coagulant Drinking Water Treatment Scheme

Authors

Nancy P. Sanchez[a*¥], Andrew T. Skeriotis[a] and Christopher M. Miller[a]

Authors' affiliations

[a] Department of Civil Engineering, The University of Akron, Akron, Ohio 44325

[*] Corresponding author:

Phone: (713) 348-3036

Fax: (713) 348-5264

Email: nps1@rice.edu

This supplementary material includes 5 tables, 7 figures and an example of calculation.

[¥]Present Address: Department of Civil and Environmental Engineering, Rice University, 6100 Main St, Houston, TX 77005.

Supporting Information- Table of Contents

Table S1. Water quality parameters of raw and treated water samples collected from Akron treatment facility from November 2009 to December 2012.

Parameter[a]	Raw water	Treated-alum	Treated-ACH
DOC (mg/L)	5.20±1.06	2.83±0.66	2.96±0.73
UV_{254} (cm^{-1})	0.12±0.04	0.04±0.01	0.04±0.01
pH	7.64±0.33	6.99±0.30	7.50±0.29

[a]Average values and associated standard deviation for the period of monitoring.

Table S2. Number of samples included in PARAFAC models fit on Akron raw and treated water data sets. Samples collected from November 2009 to December 2012[a].

Model	Number of samples-initial	Samples removed/Outliers	Number of samples-final
Raw water	384	20	364
Treated water-alum	384	29	355
Treated water-ACH	314	27	287
Raw and treated waters-alum and ACH (combined data set)	1082	190	892
Differential data set-alum	384	44	340
Differential data set-ACH	314	44	270

[a] Only samples removed from the raw and treated water data sets correspond to outliers. Samples removed from the combined data set differ from the sum between raw and treated water outliers, as these include the outliers identified in the group of raw and treated samples and their respective treated/raw water sample. 77 samples were identified as outliers (20 in the raw, 29 in the treated water-alum and 27 in the treated water-ACH data sets).

Table S3. Explained variance and core consistency levels of PARAFAC models including from one to six components fit on the Akron raw and coagulant-treated water data sets.

	Raw water		Treated-alum		Treated-ACH	
Number of components	Explained variance (%)	Core consistency (%)	Explained variance (%)	Core consistency (%)	Explained variance (%)	Core Consistency (%)
1	98.0	100	97.1	100	97.3	100
2	99.2	99.0	98.6	98.2	98.7	97.5
3	99.6	73.2	99.4	86.6	99.5	78.2
4	99.7	38.3	99.5	14.3	99.6	32.1
5	99.8	37.9	99.6	13.8	99.7	4.7
6	99.8	3.89	99.6	3.5	99.8	0.6

Table S4. Explained variance and core consistency levels of PARAFAC models fit on the differential excitation-emission matrices of the Akron raw and treated samples-alum (DF-alum model) and raw and treated samples-ACH (DF-ACH model).

Number of components	Model			
	DF-alum		DF-ACH	
	Explained variance (%)	Core consistency (%)	Explained variance (%)	Core consistency (%)
1	97.0	100	96.9	100
2	98.4	99.5	98.4	99.5
3	99.1	82.5	99.1	87.3
4	99.2	72.8	99.2	35.11
5	99.3	60.9	99.4	55.8
6	99.4	48.2	99.5	23.5

Table S5. Quarterly analysis of seasonal variation of removal of PARAFAC components after coagulation and their distribution in Akron raw and treated water during the sampling period. Composition and removal levels are calculated based on fluorescence maximum intensity (F_{max})[a].

Year	Period		Average percent contribution to F_{max} (%)									Average Fmax removal (%)					
			Raw water			Treated water-alum			Treated water-ACH[b]			Alum-based treatment			ACH-based treatment[b]		
			C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3
2010	Jan-March	Average	39.7	38.6	21.7	45.7	27.6	26.7	44.5	26.9	28.6	43.5	64.8	40.6	40.5	62.7	36.4
		Std. dev.	*3.0*	*3.7*	*6.4*	*4.7*	*3.3*	*7.7*	*5.3*	*3.6*	*8.6*	*3.3*	*3.7*	*17.5*	*5.7*	*5.4*	*17.3*
	Apr-June	Average	39.4	37.1	23.5	41.1	24.4	34.5	43.5	26.5	30.0	46.2	65.7	22.8	38.7	59.7	36.8
		Std. dev.	*2.5*	*4.3*	*6.1*	*3.5*	*1.5*	*4.7*	*2.4*	*1.9*	*3.9*	*8.1*	*7.3*	*11.1*	*10.3*	*9.6*	*4.7*
	July-Sept	Average	40.5	31.7	27.8	45.0	24.7	30.3	43.3	23.9	32.8	45.7	61.6	43.7	39.0	56.7	32.6
		Std. dev.	*2.8*	*2.3*	*4.4*	*6.1*	*2.6*	*8.6*	*4.9*	*2.8*	*7.3*	*4.8*	*4.9*	*11.4*	*5.1*	*5.3*	*12.1*
	Oct-Dec	Average	38.1	35.4	26.5	41.7	25.5	32.8	41.7	26.1	32.2	38.3	58.4	31.8	34.8	57.9	26.1
		Std. dev.	*2.9*	*5.1*	*5.8*	*5.2*	*2.7*	*7.4*	*3.7*	*2.3*	*5.4*	*4.4*	*7.4*	*12.8*	*6.0*	*9.1*	*13.9*
2011	Jan-March	Average	37.1	41.3	21.6	45.3	27.7	27.0	44.1	26.9	29.0	48.1	70.8	42.5	44.3	68.8	30.4
		Std. dev.	*3.5*	*3.0*	*6.2*	*6.4*	*4.4*	*10.1*	*7.8*	*4.2*	*11.3*	*7.6*	*7.3*	*16.7*	*6.2*	*6.9*	*17.1*
	Apr-June	Average	38.2	42.4	19.4	43.8	30.1	26.1	43.1	30.8	26.1	48.2	67.9	38.5	34.8	57.9	26.1
		Std. dev.	*2.5*	*1.3*	*2.7*	*3.2*	*2.1*	*4.3*	*2.7*	*2.0*	*3.5*	*8.1*	*6.1*	*15.2*	*10.6*	*8.8*	*15.5*
	July-Sept	Average	39.4	37.9	22.7	43.2	27.2	29.6	43.2	27.9	28.9	52.0	67.7	42.6	40.4	59.4	33.5
		Std. dev.	*2.0*	*3.3*	*4.5*	*2.9*	*2.8*	*4.2*	*2.9*	*2.0*	*3.9*	*6.1*	*9.9*	*13.3*	*5.6*	*6.3*	*15.1*
	Oct-Dec	Average	39.3	43.5	17.2	44.9	29.0	26.1	45.5	31.7	22.8	46.9	68.7	32.5	34.6	58.6	25.7
		Std. dev.	*2.8*	*2.9*	*5.4*	*4.1*	*3.0*	*6.5*	*4.6*	*3.8*	*8.1*	*3.2*	*3.9*	*12.6*	*4.2*	*4.6*	*19.8*
2012	Jan-March	Average	38.8	40.0	21.2	46.8	29.7	23.5	44.6	28.7	26.7	43.9	65.7	32.9	39.7	62.3	27.8
		Std. dev.	*2.0*	*1.7*	*3.3*	*6.8*	*5.3*	*12.0*	*6.1*	*4.5*	*10.3*	*3.4*	*3.3*	*12.8*	*3.9*	*4.3*	*15.4*
	Apr-June	Average	40.0	36.5	23.5	44.5	25.5	30.0	44.1	27.8	28.1	40.0	62.2	32.8	39.7	62.3	27.8
		Std. dev.	*0.9*	*2.4*	*2.2*	*2.3*	*2.6*	*3.4*	*1.5*	*0.7*	*1.2*	*4.9*	*4.7*	*9.6*	*3.0*	*2.5*	*10.4*
	July-Sept	Average	40.8	33.1	26.1	45.8	24.2	30.0	NA	NA	NA	43.2	60.2	42.4	NA	NA	NA
		Std. dev.	*3.1*	*2.5*	*3.0*	*4.9*	*6.0*	*4.5*	*NA*	*NA*	*NA*	*10.5*	*16.7*	*14.4*	*NA*	*NA*	*NA*
	Oct-Dec	Average	38.0	38.9	23.1	42.6	26.7	30.7	40.4	28.1	31.5	42.4	63.0	34.2	38.0	60.1	25.9
		Std. dev.	*2.3*	*3.5*	*2.7*	*3.2*	*3.0*	*5.1*	*3.6*	*4.0*	*6.4*	*9.0*	*11.8*	*13.2*	*7.5*	*10.6*	*10.0*

[a]Results are based on PARAFAC model fit on the combined data set of Akron raw and treated water samples (n= 892 samples).
[b]ACH-based treatment was temporarily interrupted from May to September 2012.

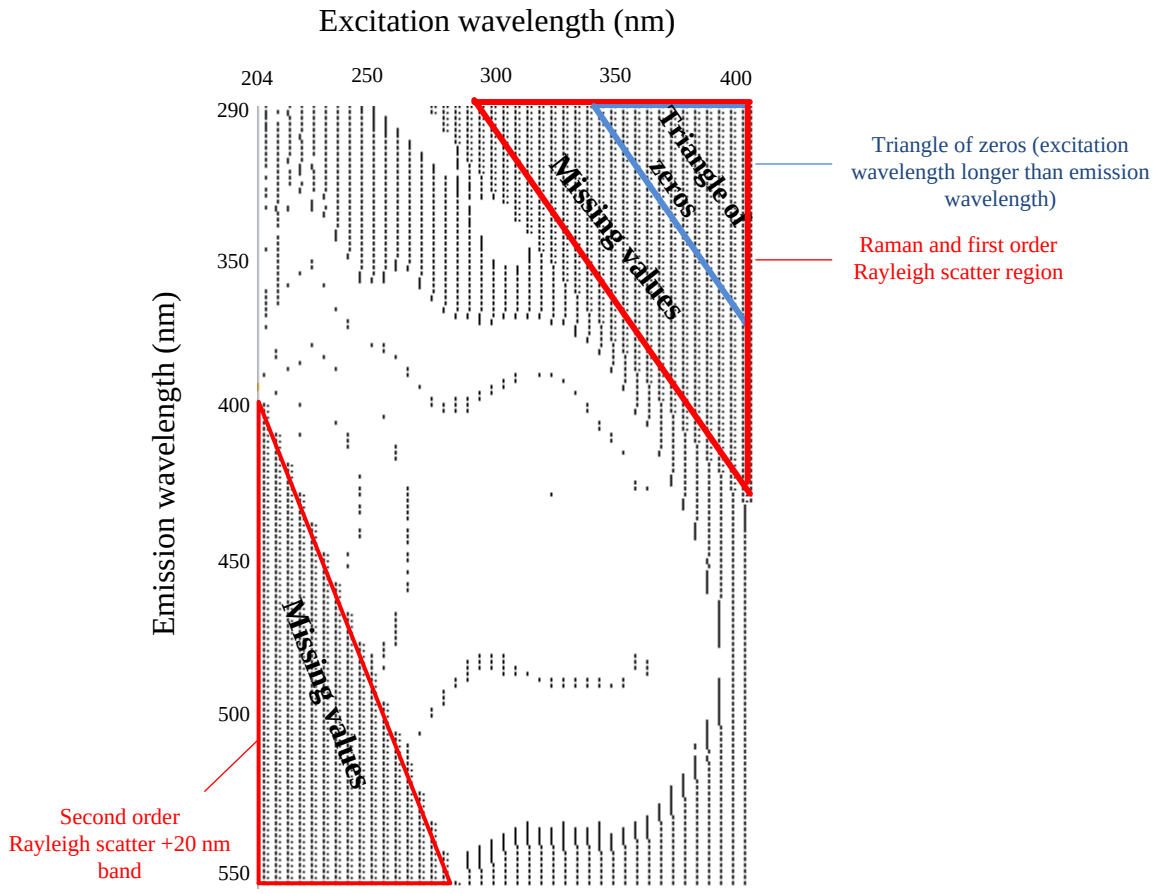


Figure S1. Location of missing values and zeros inserted in the EEMs during the pre-processing step prior to PARAFAC application.

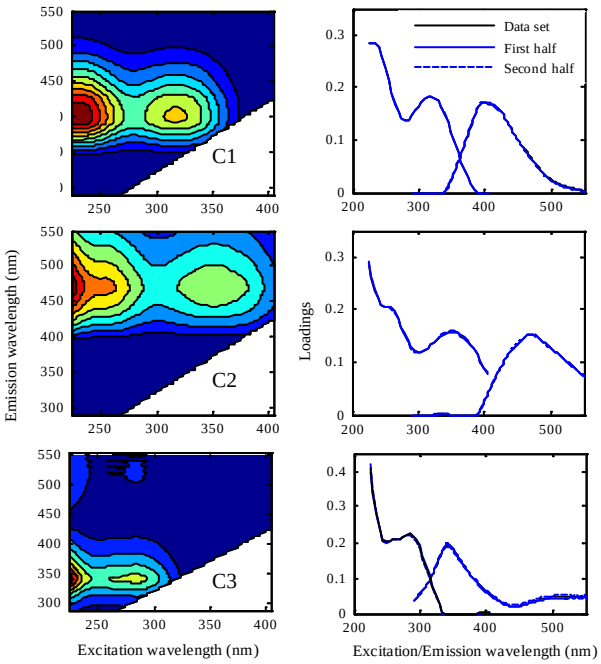


Figure S2. PARAFAC components and results of split-half validation for Akron raw water. Excitation and emission loadings at the left and right side respectively.

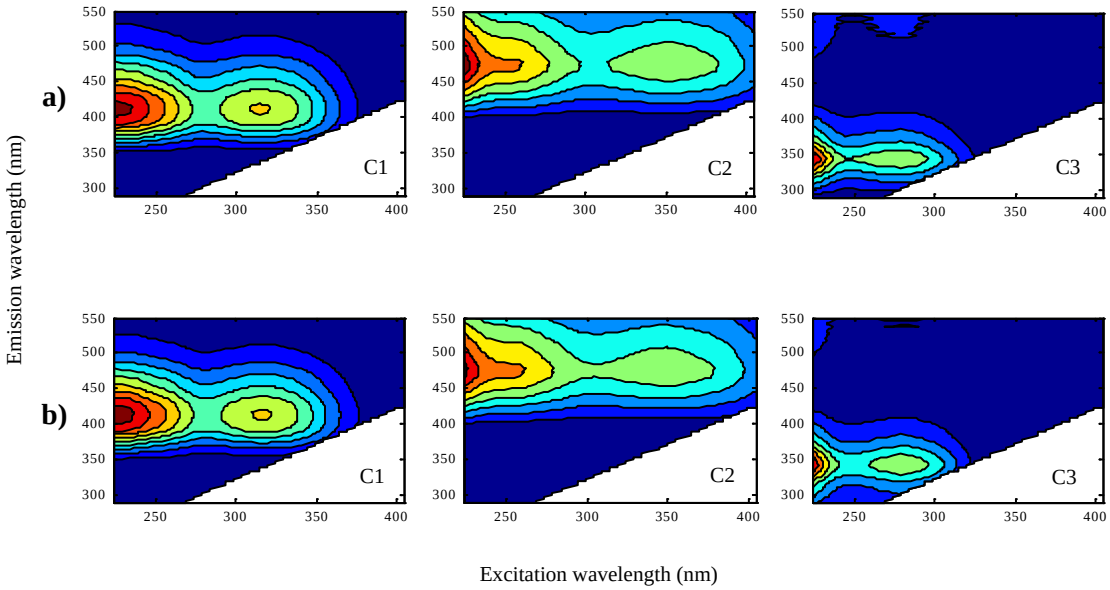


Figure S3. PARAFAC components present in the differential excitation-emission matrices of the Akron water samples. (a) alum-based coagulation (DF-alum model), (b) ACH-based coagulation (DF-ACH model).

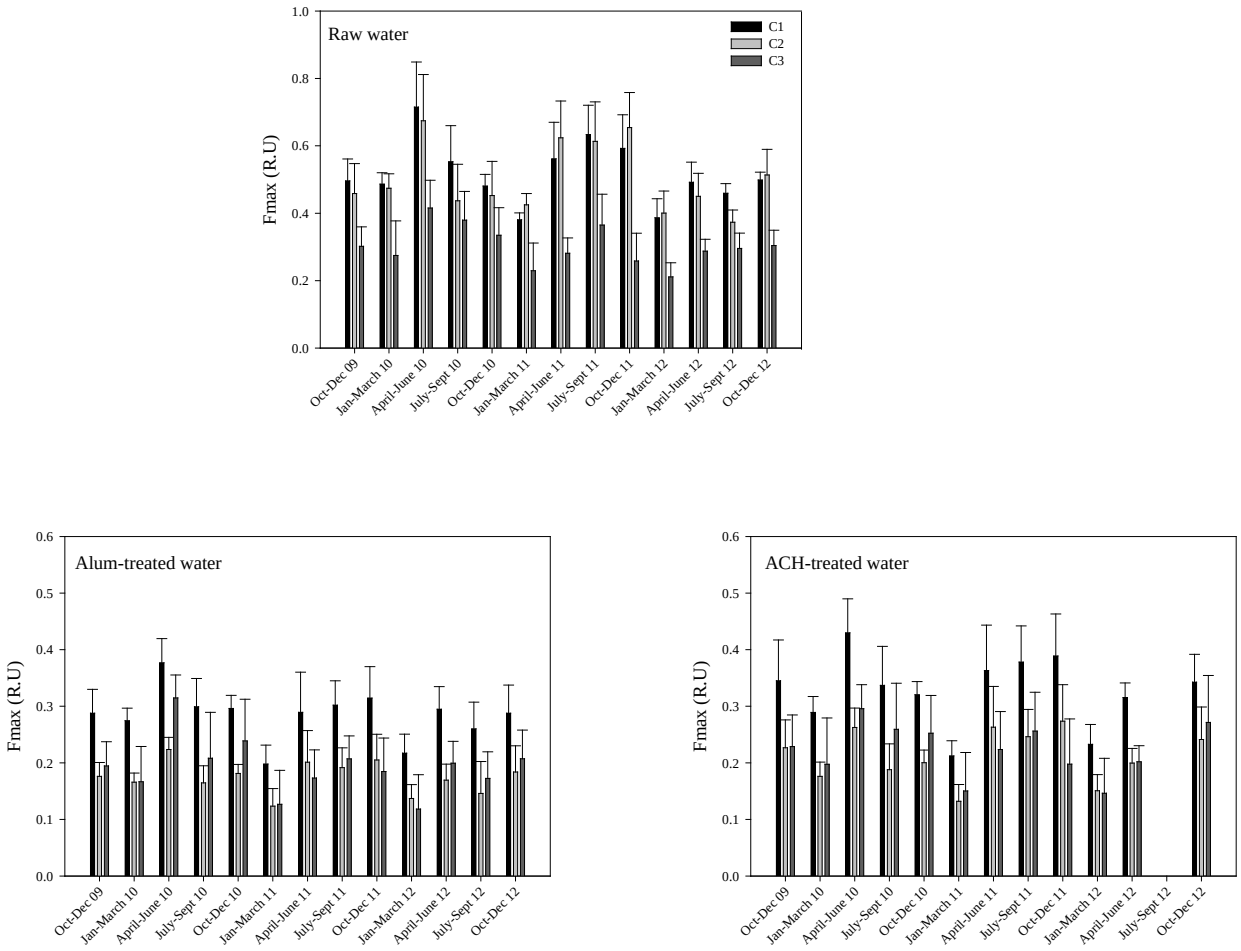


Figure S4. Seasonal variation of fluorescence signal of raw and alum and ACH-treated water DOM PARAFAC components during the monitoring period comprised between November 2009 and December 2012.

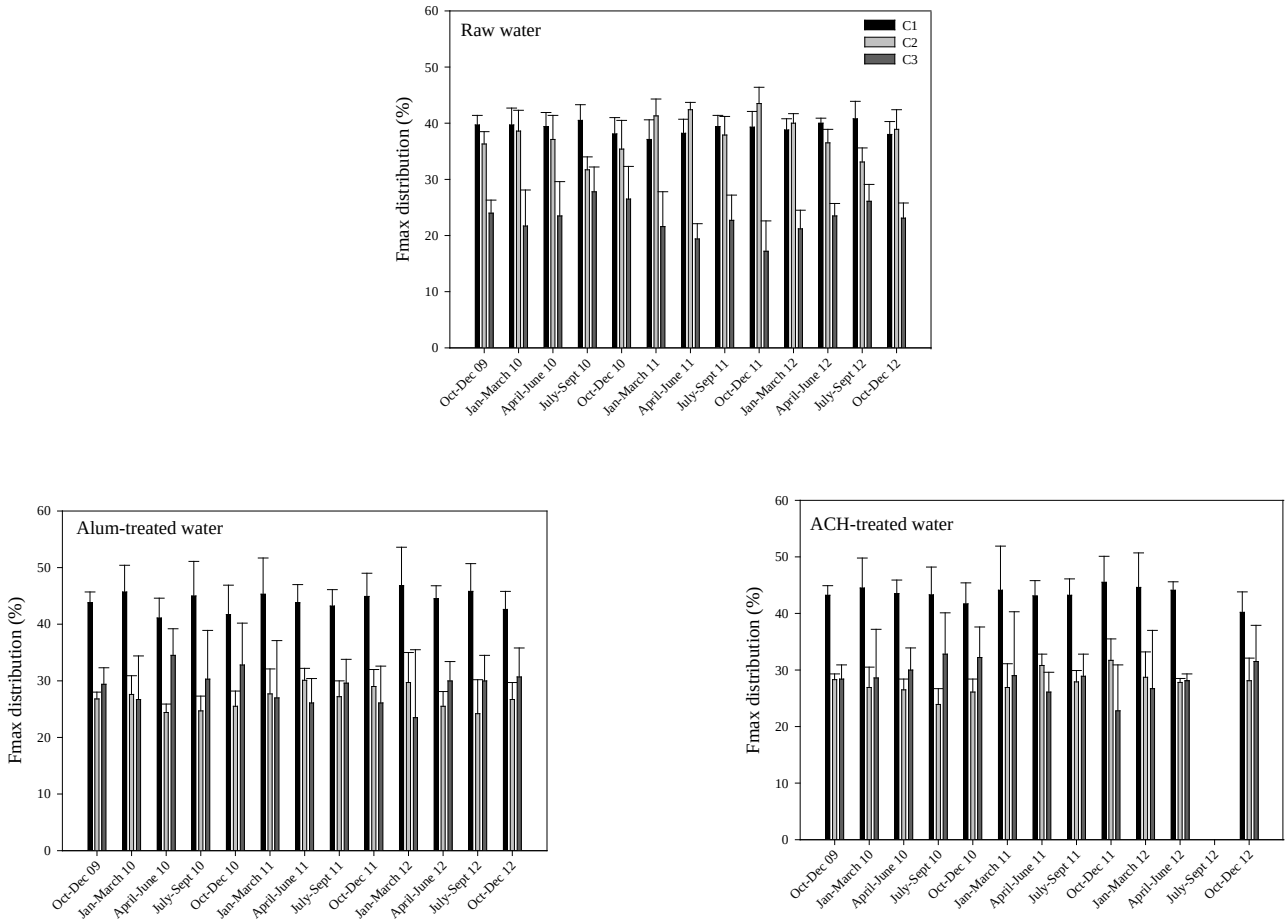


Figure S5. Seasonal variation of PARAFAC component fluorescence maximum intensity-based distribution in the raw and alum and ACH-treated water samples during the monitoring period comprised between November 2009 and December 2012.

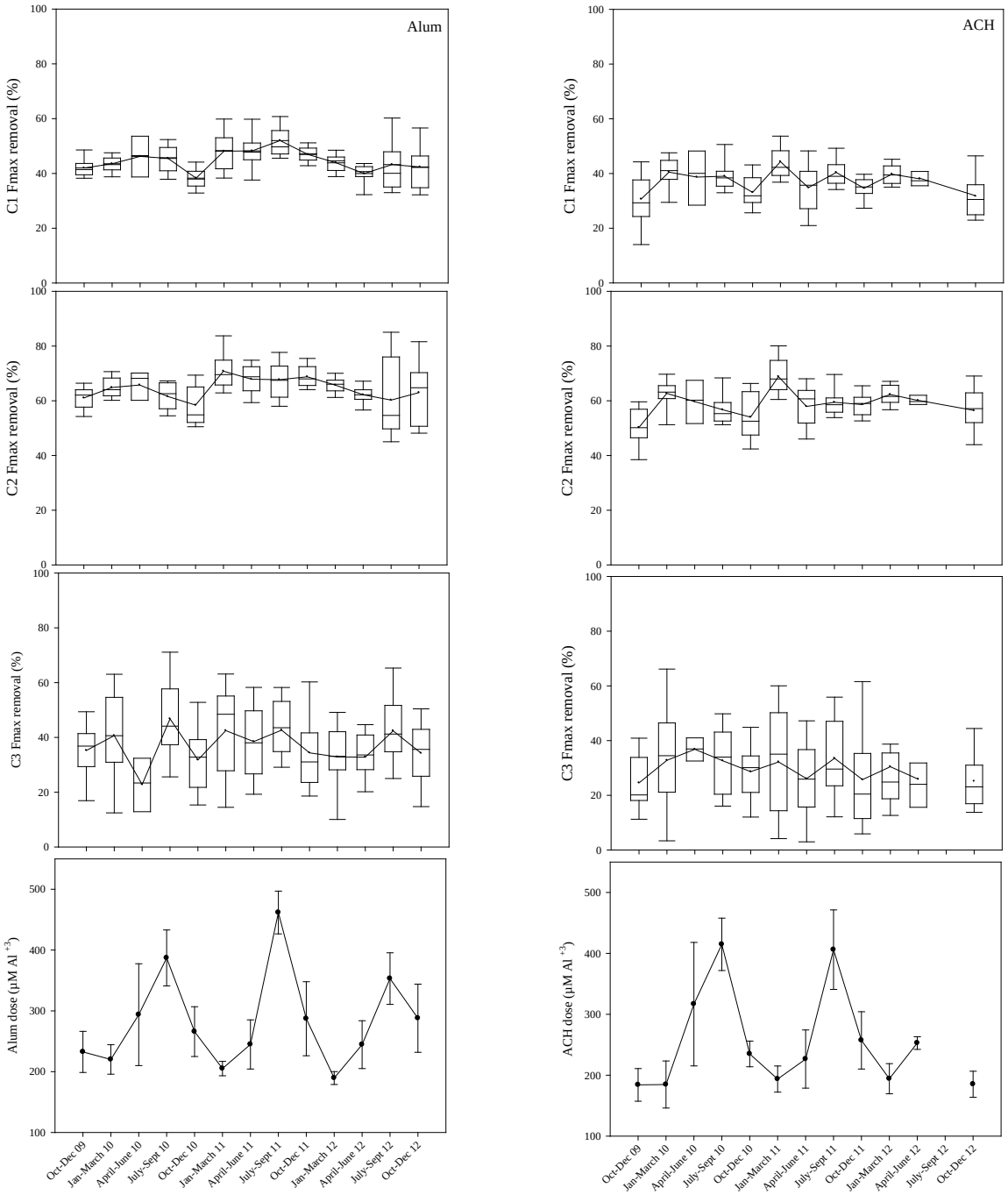


Figure S6. Seasonal variation of PARAFAC component removal by alum and ACH-based coagulation, and coagulant dose during the monitoring period comprised between November 2009 and December 2012. Bottom whisker, bottom box line, top box line and top whisker indicate the 10th, 25th, 75th and 90th percentile of the percentage of fluorescence maximum intensity (F_{max}) removal respectively. Line inside the boxes and continuous solid line represent the median and mean of the percentage of F_{max} removal respectively

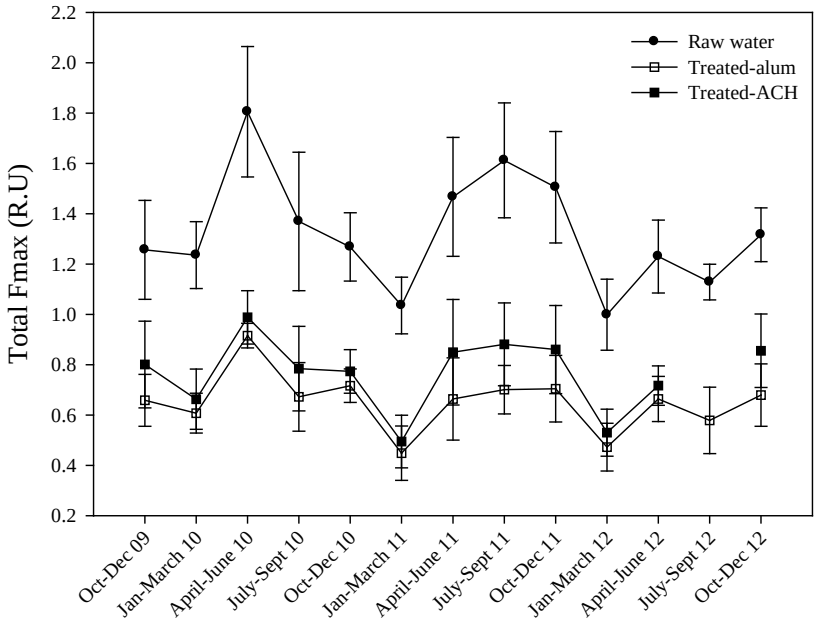


Figure S7. Seasonal variation of total fluorescence maximum intensity (F_{max}) for raw and alum and ACH-treated water samples during the monitoring period comprised between November 2009 and December 2012.

Example of Calculation

Calculation of maximum fluorescence intensity (F_{max}) and overall fluorescence intensity (OFI)-based distribution of PARAFAC components (calculation example for a raw water sample collected on 12/3/2012 at Akron drinking water treatment facility)

After PARAFAC application to EEMs expressed in Raman Units, excitation, emission and sample mode loading vectors are obtained for the components identified in the data set. If X, Y and Z represent the matrices containing the loading vectors in the sample, emission and excitation modes respectively, I and J represent the number of excitation and emission intervals set for analysis of fluorescence respectively, and F components are identified in a data set containing S samples, the size of X, Y and Z corresponds to:

X: S x F

Y: I x F

Z: J x F

In this way, the matrix representing the spectral signal of the component f in the sample s is:

$$A_{f,s} = X(s,f) * (Y(:,f) * Z(:,f)^T)$$

Where

$$Y(:,f) * Z(:,f)^T$$

Is a matrix corresponding to the outer product (OP) of the emission and excitation loading vectors of component f. Therefore, $A_{f,s}$ can be expressed as:

$$A_{f,s} = X(s,f) * OP_f$$

Fluorescence intensity values in $A_{f,s}$ are expressed in Raman Units (R.U).

F_{max} and F_{max}-based distribution calculation

The maximum fluorescence intensity of the component f in the sample s is:

$$Fmax_{f,s} = max\ (A_{f,s})$$

Which is equivalent to:

$$Fmax_{f,s} = X(s,f) * max\ (OP_f)$$

Where $Fmax_{f,s}$ is expressed in R.U

The F_{max}-based contribution of the PARAFAC component f in the sample s can be calculated as:

$$\%Fmax_{f,s} = \frac{Fmax_{f,s}}{\sum_{f=1}^{F} Fmax_{f,s}} * 100$$

OFI and OFI-based distribution calculation

The OFI of the component f in the sample s correspond to the total summation of the fluorescence intensity values in the matrix $A_{s,f}$:

$$OFI_{f,s} = \sum A_{f,s}$$

Having:

$$OP_f = Y(:, f) * Z(:, f)^T$$

Where OP_f is a matrix, OFI of the component f in the sample s can be calculated as:

$$OFI_{f,s} = X(s, f) * \sum OP_f$$

Where $OFI_{f,s}$ is expressed in R.U

The OFI-based contribution of the PARAFAC component f in the sample s can be calculated as:

$$\%OFI_{f,s} = \frac{OFI_{f,s}}{\sum_{f=1}^{F} OFI_{f,s}} * 100$$

For the specific raw water sample (s) collected on 12/3/2012, the F_{max}, OFI and Fmax and OFI-based distributions of PARAFAC components can be calculated as:

Initial information (after PARAFAC application)*

PARAFAC component (f)	**(X(s,f))**	**Max OP_f**	**ΣOP_f,**
1	9.958	0.0519	39.45
2	11.688	0.0459	48.86
3	3.355	0.0870	35.73

*F= 3 components

According to the equations above,

F_{max}-based distribution

Calculating F_{max} and F_{max}-based distribution for C1, C2 and C3:

PARAFAC component (f)	$Fmax_{f,s}$ (R.U)	F_{max}-based contribution (%)
1	0.517	39.1
2	0.536	40.5
3	0.271	20.4
Total	**1.324**	**100**

OFI-based distribution

Calculating OFI and OFI-based distribution for C1, C2 and C3:

PARAFAC component (f)	$OFI_{f,s}$ (R.U)	OFI-based contribution (%)
1	392.84	36.2
2	571.07	52.7
3	119.90	11.1
Total	**1083.81**	**100**