May 2021

Molecular Weight Distributions and Size-dependent Composition of Dissolved Organic Matter in the Aquatic Continuum

Hui Lin
University of Wisconsin-Milwaukee

MOLECULAR WEIGHT DISTRIBUTIONS AND SIZE-DEPENDENT COMPOSITION OF DISSOLVED ORGANIC MATTER IN THE AQUATIC CONTINUUM

by

Hui Lin

A Dissertation Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

in Freshwater Sciences

at

The University of Wisconsin-Milwaukee

May 2021

ABSTRACT

MOLECULAR WEIGHT DISTRIBUTIONS AND SIZE-DEPENDENT COMPOSITION OF DISSOLVED ORGANIC MATTER IN THE AQUATIC CONTINUUM

by

Hui Lin

The University of Wisconsin-Milwaukee, 2021
Under the Supervision of Dr. Laodong Guo

Dissolved organic matter (DOM) in aquatic environments is one of the most important carbon reservoirs in global carbon cycling. The molecular weight of DOM is strongly related to a great number of biogeochemical reactions, influencing ecological functions and the fate of bioactive elements in aquatic ecosystems. A new technique coupling flow field-flow fractionation with fluorescence excitation-emission matrix (EEM) and parallel factor (PARAFAC) analysis was developed to elucidate the variations in DOM composition and optical properties with molecular weight in the individual samples and their changes along the aquatic continuum.

Based on the novel coupling technique, variations in DOM characteristics were investigated across a trophic gradient from the lower Fox River to open Green Bay. Humic-like components were predominantly partitioned in the lower molecular weight fractions while protein-like components were mostly in the higher molecular weight fractions. Furthermore, unique fluorescent components in the river and open bay waters were observed, which would not have been identified using classic EEM-PARAFAC approach.

Three fluorescent DOM components, including two humic-like and one protein-like components, were first reported for the lower Yukon River. Together with seasonal changes in chemical composition and optical properties between different DOM size-fractions, we provide a baseline dataset for future trend studies in the context of environmental change in Arctic river basins. A data fusion technique was applied to jointly analyze multiple datasets including fluorescence EEM and Fourier transform infrared spectra to decipher potential internal connections between fluorophores and

chemical bonds. Humic-like components were mostly associated with C-H, C=C, and C-O bonds, while protein-like components were correlated more with C-N and N-H bonds.

Based on year-long time-series sampling, seasonal variations of DOM size and composition were further investigated in two contrasting aquatic environments, i.e., the terrestrial-dominated Milwaukee River, with environmental changes and anthropogenic impacts from metropolitan Milwaukee, and the eutrophic Veterans Park Lagoon, where seasonal *Microcystis* blooms have been observed over the past decades. In addition to differences in DOM sources and composition, which were correlated to hydrological conditions and biomass, factors controlling DOM dynamics and seasonality were distinctive between the two aquatic ecosystems.

To

my parents Mr. Jianhua Lin and Mrs. Xueying Cai,

my wife Xiaojiao Wang

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First of all, I wish to express my deep gratitude to Dr. Laodong Guo. I started to ask him silly questions when I was in college in China. It has been almost ten years. His sensor of humor and enthusiasm for my projects have inspired me a lot.

I also thank Dr. Harvey Bootsma, who brought me to his soccer team, where I gained a lot of fun and exercise. Due to the pandemic COVID, I just played with him only in the first game in the winter session, 2020. In spite of only one game, I assisted him to have a goal, which is the only goal in our team in that game. After that, I scored at least one goal or assist every game.

I would like to thank other committee members, Dr. James T Waples, Dr. Shangping Xu, and Dr. Yin Wang. I really appreciate your kindness to my stupid answers in proposal and oral prelim presentations.

Special thanks to Dr. Urban Wünsch of Technical University of Denmark for nice talks in Ocean Science Meeting 2018, an inspiring virtual meeting, and many helpful and insightful comments on my published and publishing manuscript.

Finally, I really thank my parents Mr. Jianhua Lin and Mrs. Xueying Cai, my wife Xiaojiao Wang, and all my friends for their supports.

Chapter 1.General Introduction

1. 1. Background

Dissolved organic matter (DOM) in aquatic environments is an essential component in the global carbon biogeochemical cycles. The magnitude of carbon (670 PgC) in aquatic DOM or dissolved organic carbon (DOC) is comparable to that in the atmospheric carbon dioxide (CO_2) pool (747 PgC, Ogawa and Tanoue 2003; Søndergaar and Thomas 2004). The amount of CO_2 decomposed from ten percent of DOC is larger than annual global CO_2 emission from fossil fuels (36.4 PgC in 2019, Le Quéré et al. 2018; Ritchie and Roser 2020). Furthermore, DOM is of great importance in aquatic ecosystems due to not only its vast quantity but also its role in controlling microbial community composition, biogeochemical processes, and cycling of bioactive elements (Jiao et al. 2010; Romera-Castillo et al. 2018; Xu and Guo 2018). DOM, especially chromophoric DOM (CDOM), also manipulates the depth of photic zone in surface waters indirectly based on its photoreactive properties (Wetzel 2001; Søndergaar and Thomas 2004). Aquatic organisms are protected from harmful ultra-violet radiation in aqueous environments partly due to the existence of natural DOM (Wetzel 2001). Meanwhile, a fraction of DOM that absorbs blue light draws great attention from researchers in satellite remote sensing and ocean optics. Many researchers take advantage of this attribute to determine terrestrial inputs from continents into coastal seas (Binding et al. 2015; Zolfaghari et al. 2016). Therefore, the biogeochemical fate of aquatic DOM is an important issue that must be considered to understand the global carbon cycle in the context of climate change.

Molecular size is one of the most essential properties in DOM characterization. It is strongly related to various DOM biogeochemical reactions, such as degradation/aggregation processes (Sun et al. 2017; Santschi 2018; Xu et al. 2018b), complexation with trace metals (Tang et al. 2001; Chen et al. 2004; Lotfi-Kalahroodi et al. 2019), scavenging effects (Santschi 2018), and its reactivity and bioavailability (Amon and Benner 1996; Guo and Santschi 1997; Benner and Amon 2015).

A large fraction of natural DOM absorbs light and/or fluoresces when absorbing ultraviolet light. Therefore, UV-visible spectrophotometer and spectrofluorometer are the most widely used techniques for DOM characterization. In addition, optical characterization and measurements are fast and cost-effective.

Absorption coefficients (a_{CDOM}) of UV or visible light are widely used as indicators of DOM concentration (Fichot and Benner 2011). Spectrum-derived parameters, such as DOC-normalized Specific UV Absorbance at 254 nm ($SUVA_{254}$) and spectral slope between 275 nm and 295 nm ($S_{275-295}$), are used as proxy for structure-related DOM properties (Weishaar et al. 2003; Fichot and Benner 2012). With a spectrofluorometer, fluorescence emission spectrum (a range of λ_{em}) is measured over a range of excitation wavelengths (λ_{ex}) to generate three-dimensional fluorescence landscapes, so-called excitation-emission matrix (EEM). Each EEM represents a combination of fluorescence from an unknown number of underlying fluorophores with different excitation/emission wavelengths (Coble 2007; Coble et al. 2014). Over the past decade, it has become a common practice to decompose EEM database statistically using parallel factor analysis (PARAFAC, Murphy et al., 2014). PARAFAC is a machine learning algorithm that decomposes a batch of 3D fluorescence spectra by generating a statistic model that contains deconvoluted components (Malinowski 2002; Stedmon et al. 2003). However, PARAFAC is barely applicable for short-term observations and small region investigations such as streams, rivers, and groundwater due to its requirement for a large dataset. To make PARAFAC analysis more capable for limited samples (or even only one sample), I developed a method using flow field-flow fractionation (FlFFF) coupling with EEMs and PARAFAC analysis (FlFFF-EEM-PARAFAC) to conquer this problem. Details are provided in Chapter 2.

Variations in DOM molecular size and chemical composition in marine estuaries have been widely reported (e.g. Guo and Santschi 2007 and references therein). However, studies about the size and composition of DOM in lacustrine estuaries are limited, especially in negative estuaries (or inverse estuaries) where salt concentration is higher in the river-end than in the open water end. Using the FlFFF-EEM-PARAFAC method described in Chapter 2, DOM molecular size and optical properties were investigated along a river-estuary continuum from the lower Fox River to open Green Bay. Details are provided in Chapter 3.

As mentioned above, fluorescence spectroscopy is a sensitive and cost-effective determination tool, which has been widely used to characterize DOM in natural waters. However, since fluorescence spectroscopy is unable to provide molecular-level information, barriers to further interpret EEM spectra still exist.

Meanwhile, Fourier transform-infrared (FT-IR) spectroscopy is capable to detect chemical functional groups and chemical bonds, which makes it decent complementation to fluorescence spectroscopy (Abdulla et al. 2010a; Yang et al. 2015; Xu et al. 2018e). To jointly analyze FT-IR and EEM spectra, I developed a method to fine-tune a data fusion algorithm for both spectral datasets and applied this method to water samples from the Yukon River, an Arctic River in Alaska. Details are provided in Chapter 4.

Dynamic changes in source and fate of DOM lead to seasonal variations in DOM concentrations and properties, which can footprint regional hydrological and biogeochemical changes. Here, two different aquatic environments are selected to build up the linkage between DOM seasonality and hydrological patterns. One is the Milwaukee River, where the river basin is at the southeast quarter of Wisconsin. The downstream of the Milwaukee River drains through metropolitan downtown Milwaukee, where human activities had a great impact on the river and estuary. By comparing DOM indicators between upstream and downstream waters, anthropogenic impacts from metropolitan areas on the river could be unraveled. By comparing DOM properties collected during 2014-2015 (Teber 2016) and 2018-2019, longer-term variation trends and the influence of environmental change and human activity, such as dam removal, on water chemistry and DOM dynamics were evaluated. Details are provided in Chapter 5.

The other unique aquatic environment selected for this study is the Veterans Park Lagoon, a hypereutrophic lagoon, in the city of Milwaukee, where seasonal cyanobacteria *Microcystis* blooms have been observed over the past years (Wang et al., 2021). Using ultrafiltration membranes with different pore sizes, DOM was separated into several operationally defined size-fractions, including, e.g., the < 1 kDa, 1-10 kDa, and >10 kDa. Seasonal variations and biogeochemical behaviors of bulk- and size-fractionated DOM were investigated. Details are provided in Chapter 6.

1. 2. Scientific Questions

Below are the scientific questions addressed in each chapter of this dissertation.

- Chapter 2. When using the FlFFF-EEM-PARAFAC method, does the one-sample approach provide validated fluorescence components as same as those derived from large datasets? Is there any difference in the validation approach between one sample FlFFF-EEM-PARAFAC and

classic PARAFAC analysis? How to compare component loadings among different models from one-sample FlFFF-EEM-PARAFAC approach? What threshold settings for Tucker's Congruence Coefficient are appropriate to identify two components from different models as one fluorophore?

- Chapter 3. Is the fate of DOM in an inverse estuary different from that in a normal estuary? Does decreasing conductivity along an inverse estuary induce the change of DOM molecular size, stimulate the transformation of DOM, or swap the internal structure of DOM?
- Chapter 4. What are the differences in molecular composition and optical properties among different DOM size-fractions (i.e. HMW- and LMW- DOM) in the Yukon River? How to validate, evaluate, and interpret the results derived from the data fusion technique, a new method in terms of chemometrics? Is there any relation between components from PARAFAC and components from data fusion?
- Chapter 5. Does anthropogenic DOM influence the properties of DOM in a freshwater ecosystem? How severely the human activities impact the Milwaukee River? Is there any phenomenon, in light of DOM, that can be observed after the removal of the Estabrook dam?
- Chapter 6. How does DOM change with algal blooms in a hyper-eutrophic freshwater lagoon ecosystem? What are the differences in composition and optical properties between size-fractions (<1 kDa, 1-10 kDa, > 10 kDa) in the lagoon ecosystem?

1. 3. Objectives

The primary goals of this research are to develop novel methods to disentangle technical limitations on analyzing DOM fluorescent data, to characterize DOM chemistry from molecular and optical levels, and to further understand the aquatic DOM behaviors in response to the biogeochemical changing across different freshwater ecosystems.

The specific objectives of this research are to:

(1) combine FlFFF in tandem with EEM and PARAFAC analysis, evaluate the PARAFAC model from an individual water sample.

(2) investigate and characterize DOM in molecular level in the Fox River – Green Bay estuary using the newly developed FlFFF-EEM-PARAFAC coupling techniques.

(3) develop a novel method to fine-tune the data fusion technique (ACMTF algorithm) in order to fit FTIR spectra and fluorescence EEM spectra.

(4) evaluate DOM signatures in relation to potential anthropogenic impacts from the terrestrial DOM-dominated Milwaukee River basin.

(5) Interpret DOM biogeochemistry in response to seasonal cyanobacteria blooms in a hypereutrophic freshwater lagoon ecosystem.

1. 4. Organization of the Dissertation

Chapter 1 **General Introduction** (this chapter).

In this chapter, a description of study backgrounds, statements and scientific questions, and objectives of research projects are presented.

Chapter 2 **A novel method for dissolved organic matter characterization of an individual water sample using FlFFF-EEM-PARAFAC coupling techniques.**

In this chapter, a method was developed to characterize DOM size distribution and chemical composition using an FlFFF-EEM-PARAFAC coupling technique. Blank determinations, collection procedures, and carrier solutions were optimized for this method. Three freshwater samples were tested for repeatability and reliability of cross-validations. Errors between summarized EEM and bulk EEM were optimized to less than 5%.

Chapter 3 **Size and composition of dissolved organic matter in a negative estuary: a case study in the Fox River plume.**

In this chapter, applications of the FlFFF-EEM-PARAFAC method, presented in Chapter 2, were conducted to evaluate changes in DOM size-spectra and chemical compositions along the Fox River – Green Bay continuum.

Chapter 4 **Dynamic changes in size-fractionated dissolved organic matter composition in the seasonally ice-covered Yukon River as characterized using EEM-PARAFAC, FT-IR, and data fusion.**

In this chapter, a novel method to jointly analyze two spectral datasets (EEM and FT-IR) was developed. Applications of the method to interpret the seasonal variations of DOM in the lower Yukon River were presented. This manuscript has been submitted to a scientific journal, *Limnology, and Oceanography*.

Chapter 5 **Seasonal and multiyear variations of dissolved organic matter in the Milwaukee River, Wisconsin.**

This chapter presents seasonal variations in DOM concentration and optical properties based on two-year time-series observations (2017-2019) in the Milwaukee River, including source and composition of DOM, comparisons between upstream (pristine) and downstream (with human impacts), and comparisons with observations during 2014-2015 to evaluate the influence of environmental change, such as dam removal.

Chapter 6 **Dynamic cycling of organic matter under *Microcystis* blooms in a shallow eutrophic lagoon.**

This chapter presents seasonal variations in dissolved and particulate organic matter in the Veterans Park Lagoon, a hyper-eutrophic coastal lagoon, with cyanobacteria *Microcystis* blooms.

Chapter 2. A novel method for DOM characterization of an individual water sample using the FlFFF-EEM-PARAFAC coupling technique

This chapter is a slightly modified version of *Variations in Colloidal DOM Composition with Molecular Weight within Individual Water Samples as Characterized by Flow Field-Flow Fractionation and EEM-PARAFAC Analysis* published in the scientific journal *Environmental Science and Technology* in 2020 and has been reorganized with the permission of the copyright holder (ACS Publications). Below is a direct link to the published work: https://pubs.acs.org/doi/abs/10.1021/acs.est.9b07123

2. 1. Abstract

Fluorescence excitation emission matrices (EEM) and parallel factor (PARAFAC) analysis have been widely used in the characterization of dissolved organic matter (DOM) in the aquatic continuum. However, large sample sets are typically needed for establishing a meaningful EEM-PARAFAC model. Applications of the EEM-PARAFAC technique to individual samples require new approaches. Here, flow field-flow fractionation (FlFFF) combined with offline EEM measurements and PARAFAC analysis was used to elucidate the dynamic changes in DOM composition/optical properties with molecular weight within individual samples. FlFFF-derived size spectra of ultrafiltration-isolated colloidal DOM show that peak-C related fluorophores (Ex/Em= 350/450 nm) are present mostly in the 1−3 kDa size range, while peak-T associated fluorophores (Ex/Em = 275/340 nm) have a bimodal distribution with peaks in both the 1−3 kDa and the >100 kDa size fractions. The integrated EEM spectra from FlFFF size-fractionated subsamples closely resembled the bulk EEM spectra, attesting to the convincing comparability between bulk and FlFFF size-fractionated EEMs. The PARAFAC-derived DOM components are distinctive among individual samples with the predominant components being humic-like in river water, but protein-like in a highly eutrophic lagoon sample. This compelling new approach combining FlFFF and EEM-PARAFAC can be used to decipher the dynamic changes in size spectra and composition of individual DOM samples from sources to sinks or across the redox/hydrological/trophic interfaces.

2. 2. Introduction

Dissolved organic matter (DOM) is ubiquitous and heterogeneous in aquatic environments and plays an essential role in regulating water quality, chemical speciation of many trace elements, and the fate/transport of contaminants (Aiken et al. 2011; Philippe and Schaumann 2014; Kteeba et al. 2017; Kim and Kim 2017). Understanding the composition, reactivity, and cycling pathways of DOM is thus indispensable. Previous studies have shown the bulk DOM contains diverse components with different molecular sizes, ages, composition/functionalities, and chemical/biological reactivities in natural waters (Benner and Amon 2015; Xu and Guo 2017; Xu et al. 2018a). For example, different DOM components within the bulk DOM pool could be preferentially decomposed through different degradation pathways/mechanisms (Chen and Jaffé 2016; Hansen et al. 2016; Xu and Guo 2018), resulting in distinct DOM properties along the aquatic continuum (Abdulla et al. 2013; Xu et al. 2018b). Therefore, size and compositional characterization of individual DOM samples are needed to provide new insights into the lability, transformation, and environmental fate of natural organic matter in the aquatic continuum, especially at the groundwater-surface water, river-lake, and river-sea interfaces.

Due to the small volume requirement and user-friendly analytical procedure, optical properties measured by UV-visible spectroscopy techniques have been widely used in the characterization of DOM in natural waters (Mcknight et al. 2001; Weishaar et al. 2003; Coble 2007; Murphy et al. 2010; Osburn et al. 2012). Furthermore, applications of fluorescence excitation emission matrices (EEM) with parallel factor analysis (PARAFAC) have further enhanced our understanding of the sources, composition, and chemical properties of fluorescent DOM in aquatic environments, especially over the past decade (Stedmon et al. 2003; Stedmon and Bro 2008a). Nevertheless, the EEM-PARAFAC analysis requires a large number of samples to establish statistical models. For individual samples or when sample numbers are limited, new approaches are sorely needed to acquire PARAFAC-derived DOM components.

Natural DOM constantly endures degradation, modification, and transformation during its transport from sources to sinks along the aquatic continuum. As a result, each DOM sample should have its unique molecular size-spectrum and distinct DOM composition, especially along stations across river-lake and land-ocean interfaces (Giddings 1993; Abdulla et al. 2013; Zhou et al. 2016b; Xu et al. 2018b). Thus,

application of the EEM-PARAFAC to the characterization of individual DOM samples is urgently needed in order to thoroughly decipher dynamic changes in DOM molecular size and size-dependent chemical composition along a specific aquatic continuum.

Flow field-flow fractionation (FlFFF) is an analytical technique capable of simultaneous size-separation and chemical characterization of colloidal organic matter, macromolecules, and nanoparticles (Giddings 1993; Baalousha et al. 2011; Stolpe et al. 2013). In the past decades, the FlFFF has been frequently used in aquatic and environmental studies (Stolpe et al. 2010; Baalousha et al. 2011; Cuss and Guéguen 2013; Zhou et al. 2016b; Lin et al. 2016). Recently, Cuss and Guéguen combined the FlFFF with an online fluorescence detector, allowing discrete EEM characterization by halting elution-flow (Cuss and Guéguen 2013). In addition, Zhou and Guo reported the application of FlFFF and offline EEM measurements and showed variations in EEM spectra with selected DOM size-fractions in an individual sample (Zhou and Guo 2015). Similarly, Wünsch et al. and Murphy et al. demonstrated their one-sample PARAFAC approach, for the first time, using either high-performance size exclusion chromatography (HP-SEC) coupled with an online emission-spectra fluorescence detector through multiple-injections or photochemistry-EEM-PARAFAC approach (Wünsch et al. 2017a; Lee et al. 2018). Many previous studies used solid-phase extraction (SPE) and/or XAD resins to pre-concentrate DOM for further chemical characterization, which requires the acidification of natural waters to pH <2 and thus may fractionate DOM compositionally and structurally (Chen et al. 2016). In contrast, ultrafiltration-isolated DOM has a well-defined molecular weight range and is based on physical separation without chemical alteration to the DOM (Guo and Santschi, 2007). Despite recent progress in DOM characterization and one-sample PARAFAC approach (Wünsch et al. 2017a; Murphy et al. 2018a; Lee et al. 2018), combination of FlFFF with offline EEM-PARAFAC analysis has not been reported for single DOM samples, especially for colloidal DOM samples isolated using ultrafiltration.

In this study, a method combining FlFFF separation and characterization with offline EEM measurements was developed to elucidate variations in fluorescence EEM spectra with molecular weight in individual colloidal DOM samples preconcentrated using ultrafiltration with a 1 kDa membrane. Together with PARAFAC analysis on the EEM data from all sub-size-fractions, major fluorescent DOM components

were identified for each individual DOM sample collected from different aquatic environments. Applications of the FlFFF and EEM-PARAFAC to natural DOM samples provide new insights into both size-dependent chemical properties and PARAFAC-derived fluorescent DOM components within individual DOM samples.

2. 3. Materials and Methods

2. 3. 1 Sample processing

Three unique water samples were collected from different aquatic environments, including 1) the Milwaukee River, a terrigenous DOM-dominated river in Wisconsin, 2) open Green Bay, a mesotrophic sub-basin of Lake Michigan, and 3) Veterans Lagoon in the City of Milwaukee, a eutrophic lagoon with seasonal algal blooms over the past years. Detailed sampling locations and hydrographic parameters are given in Table 2-1.

Water samples were collected using acid cleaned Nalgene bottles. The collected water samples were immediately filtered through pre-combusted (450°C for 4 h) GF/F filters (Whatman, 0.7 μm). Aliquots of the <0.7 μm filtrate were sampled for the measurements of dissolved organic carbon (DOC), UV-visible absorption spectrum, and fluorescence EEM.

To optimize the signal-noise ratio for FlFFF fractional samples, an aliquot of the filtrate was pre-concentrated using a stirred cell ultrafiltration unit (Amicon 8200) with a 1 kDa membrane (regenerated cellulose, Millipore YM1, 63.5-mm diameter; actually cutoff >1.33 kDa based on a 80-85% rejection rate). Before ultrafiltration, the ultrafiltration disc membrane was pre-cleaned with 0.05 M NaOH solution followed by ultrapure water until the DOM and CDOM signatures were similar to those of ultrapure water (Xu and Guo 2017). To quantify the abundance of bulk colloidal DOM, time-series permeate samples, in addition to initial, integrated permeate and retentate samples, were collected for the measurements of CDOM using UV-visible and fluorescence spectrophotometers. The ultrafiltration permeation model was then applied to fit the time-series data to calculate the colloidal DOM abundance (Guo & Santschi, 1996), which ranged from 58% to 75% in these three samples (Figure 2-1).

Table 2-1. Sample descriptions, hydrographic parameters, and bulk DOM properties, including dissolved organic carbon (DOC), HMW-DOC (>1 kDa), LMW-DOC (<1 kDa), absorption coefficient at 254 nm (a_{254}), spectral slope between 275-295 nm ($S_{275\text{-}295}$), specific UV absorbance at 254 nm ($SUVA_{254}$), and fluorescence indices (HIX, humification index, and BIX, biological index) in water samples collected from the lower Milwaukee River, mesotrophic Green Bay/Lake Michigan, and eutrophic Veterans Lagoon/Milwaukee.

Parameters (Unit)	Milwaukee River	Green Bay	Veterans Lagoon
Longitude (°)	-87.910	-87.843	-87.894
Latitude (°)	43.035	44.718	43.046
Temperature (°C)	2.4	n/a	15.65
pH	8.28	8.05	8.9
Specific Conductivity (μS/cm)	882	321	971
DOC (μmol/L)	514	441	413
HMW-DOC (μmol/L)	298	290	309
LMW-DOC (μmol/L)	216	150	106
Colloidal DOC (%)	58	66	75
CDOM a_{254} (m^{-1})	49.9	27.7	23.2
$S_{275\text{-}295}$ (nm^{-1})	0.017	0.020	0.017
$SUVA_{254}$ (L mgC^{-1} m^{-1})	3.51	2.27	2.03
BIX	0.60	0.73	0.71
HIX	4.31	4.82	1.60

2. 3. 2 Measurements of DOC and CDOM

Concentrations of DOC were measured using a TOC-L analyzer (Shimazu, Japan). Ultrapure water, working standards, and community certified DOC samples (from University of Miami) were measured to monitor the blank level and instrument performance and to ensure data quality (Guo et al., 1995).

UV–visible absorption spectra were measured using an Agilent 8453 spectrophotometer (DeVilbiss et al. 2016a). CDOM absorption coefficients at 254 nm (a_{254}) were calculated using the following equation: a_{254} (m^{-1}) = $2.303 \times A(254nm)/l$, where a_{254} is the Naperian absorption coefficient at 254 nm, $A(254nm)$ is the absorbance at 254 nm, and l is the path length of cuvette (in m). Spectral slopes between 275-295 nm ($S_{275-295}$ in nm^{-1}) were calculated using the following equation: $a_\lambda = a_{\lambda_0} e^{-S_{275-295}(\lambda-\lambda_0)}$. Linear regression was performed to fit the absorption spectra to the log-transformed equation: $\ln a_\lambda = \ln a_{\lambda_0} - S_{275-295}(\lambda - \lambda_0)$(Helms et al. 2008). $SUVA_{254}$, an indicator of aromaticity, was calculated as: $SUVA_{254} = A(254nm)/DOC$, with a dimension of L m^{-1} mgC^{-1}.

2. 3. 3 Separation and characterization of DOM using FlFFF

The FlFFF system (AF2000, Postnova, Salt Lake City, UT), equipped with a 0.3 kDa polyether sulfone ultrafiltration membrane (with an actual cutoff of >1 kDa based on a 90% rejection rate), is coupled online with a UV absorbance detector and two fluorescence detectors (Zhou and Guo 2015). Carrier solution was prepared from 10 mmol L^{-1} NaCl (aq) and 5 mmol L^{-1} H_3BO_3 (aq) and pH of 8 (adjusted with NaOH) for optimal separation. Calibrations of DOM molecular size (in kDa or nm) against elution time have been described elsewhere (Stolpe et al. 2010; Zhou and Guo 2015).

Flow settings (including tip flow, focus flow, cross flow, etc.) followed those in previous studies. Briefly, the focus flow was 2.1 ml/min, while the tip flow was 0.1 ml/min. The resultant high ratio of focus to tip flow rate guaranteed an optimal pre-separation condition. The crossflow rate was set at 1.5 ml/min. After 15 min of focusing, cross flow and outlet flow remained constant at 1.5 ml/min and 0.7 ml/min, respectively. At the same time, the tip flow increased to 2.2 ml/min and the focus flow gradually declined to zero, shifting the FlFFF status from focusing to eluting. The eluting lasted 20 minutes to thoroughly separate colloidal DOM molecules in the channel. After eluting, cross flow linearly declined to zero while the tip flow linearly declined from 2.2 ml/min to 0.7 ml/min. In this 15-minute step, the flow field produced by crossflow gradually decreased to zero, which flushed out large molecules retaining above the membrane (Zhou and Guo 2015).

Online DOM characterization was conducted using a UV-absorbance detector (SPD-20A, Shimazu) and two fluorescence detectors (RF-20A, Shimadzu) with different excitation/emission (Ex/Em) combinations. Absorbance at 254 nm (UV_{254}) is commonly regarded as an indicator of CDOM abundance in natural waters (Weishaar et al. 2003; Spencer et al. 2012). Thus, UV-absorbance at 254 nm was selected to monitor changes in bulk CDOM along its colloidal size continuum within each sample. The fluorescence detectors were used to measure the UVC humic-like fluorophores, also known as Peak C (Ex/Em at 350/450 nm), and the tryptophan- or protein-like fluorophores, also known as Peak T (Ex/Em at 275/340 nm). All signatures are expressed in quinine sulfate equivalent unit (ppb-QSE).

The FlFFF-derived size fractionated sub-samples were collected for offline fluorescence EEM measurements using 5 ml pre-combusted glass vials. To optimize both size resolution and EEM analysis, size-fractioned subsamples were collected at either one or half minute time interval during elution depending on the distribution of major peaks and fractograms. To balance size spectrum resolution and subsamples' fluorescence signal intensity, multiple injections (3 for the one-minute subsamples and 7 for the half-minute subsamples) are needed depending on DOM abundances. In general, ≥40 subsamples were collected from each sample for offline EEM measurements. To eliminate the systematical EEM blanks from FlFFF, the carrier solution was run as a sample on the same FlFFF system and all size-fractionated carrier subsamples were collected for EEM measurements. Each corresponding background EEM signal was then subtracted from sample's EEM spectra.

2. 3. 4 Measurements of EEM spectra and PARAFAC analysis

Fluorescence EEM spectra of FlFFF-fractionated DOM subsamples were measured using a Fluoromax-4 (Horiba Jobin Yvon) spectrofluorometer (Zhou et al. 2016a). Excitation wavelength was set in the range from 250 nm to 450 nm with 5 nm intervals, and emission wavelengths varied from 220 nm to 600 nm with 5 nm intervals. Ultrapure water was referenced to eliminate background noise for the bulk samples. For the size-fractionated subsamples, the FlFFF size-fractionated carrier solution samples were used as their EEM blanks. To eliminate inner filter effects, samples were diluted to absorbance values at 254 nm < 0.02 before EEM measurements. Electric signals of emission and excitation spectra were corrected, normalized, and calibrated to QSE (Coble et al. 1998). First and second order of Raman and Rayleigh

scattering peaks were eliminated. Biological index (BIX) is the ratio of fluorescence intensity at 380 nm to 430 nm within excitation wavelength of 310 nm (Huguet et al. 2009). Humification index (HIX) is the ratio of fluorescence signals at the range of 435-480 nm to those at the range of 300-345 nm over excitation at 254 nm (Zsolnay et al. 1999).

PARAFAC analysis was performed on the EEM data acquired from all FlFFF-fractionated samples. The model was constrained to non-negative values and the results were validated using split half analysis.[18,19] All data handling and analysis described above were performed in MATLAB 2017a (MathWorks) using drEEM toolbox (Murphy et al., 2013). Detailed comparison and validation for the PARAFAC model using Tucker Congruence Coefficient (TCC) and PARAFAC model parameters are given in the following sections.

2. 3. 5 Comparison between components

Tucker Congruence Coefficient (TCC) was used to assess the similarity between fluorescence spectra from different models. Previously, it was adopted in the PARFAC; TCC was a criterion to validate components between half-split parts. It could be calculated using the following equation: $TCC = \frac{\sum XY}{\sqrt{\sum X^2 \sum Y^2}}$, where X and Y were the excitation spectral loadings or emission spectral loadings from two components in different halves or in different models. The TCC values ranged from 0 to 1. Two spectra could be recognized as in good similarity when the TCC > 0.95. The TCC between 0.90 to 0.95 indicates mediocre similarity between the two spectra. If the TCC is under 0.90, the compared spectra are regarded as no similarity/congruence. Here we used TCC to compare spectra from different halves and from different models. In PARAFAC analysis, TCC > 0.95 was set as the threshold of the split-half validation. When TCC was used to compare two spectra from different samples (or different aquatic environments), a slightly loose criterion was set (TCC > 0.90) in the excitation and emission comparison since they are from different PARAFAC models. The lower criterion has also been used in external comparison before.

2. 3. 6 PARAFAC model statistic details

Larger than forty offline subsamples were collected per bulk sample and analyzed for EEM spectra for PARAFAC analysis. Three different PARAFAC models were established, and half-split validated using

the TCC method. The criteria for the TCC_{em} and TCC_{ex} were set to 0.95. In the final validated models, 99.76%, 99.95%, 99.30% variance can be explained for samples from the Milwaukee River, Green Bay, and Veteran Lagoon, respectively. The relatively low explanation for the Green Bay sample is probably due to its low original fluorescence intensity. Core consistency could be regarded as a criterion to diagnose the PARAFAC model; low core consistency indicates high correlations between PARAFAC components. Here, the core consistencies for the river, bay and lagoon samples were 93.24%, 96.99%, 92.15%, respectively. Summaries of the squared errors for three models were presented in Figure 2-1. These model details were derived from PARAFAC model using drEEM (v.0.4.0) toolbox in Matlab (2017a, MathWorks).

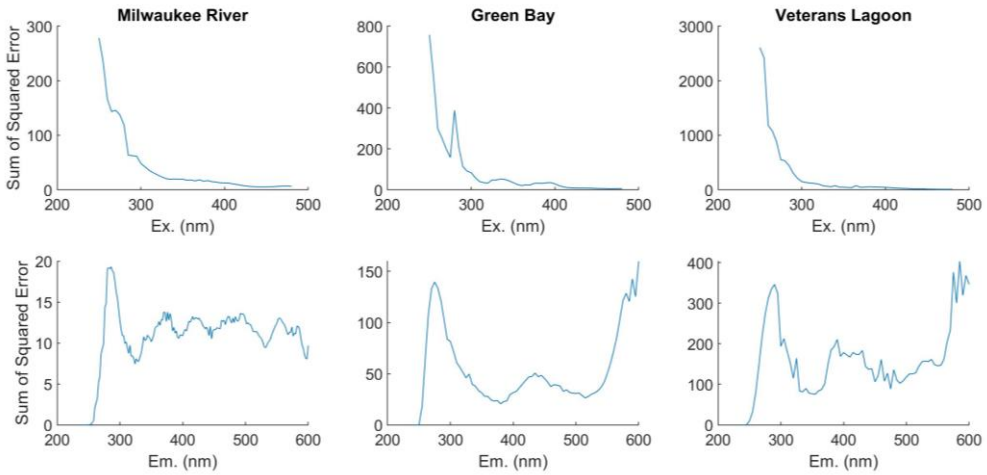

Figure 2-1. Sum of squared error from excitation (top) and emission (bottom) spectra for the three bulk samples.

2. 4. Results and discussion

2. 4. 1 Bulk DOM properties

Before further size fractionation and FlFFF analysis, water samples were characterized for the bulk properties including DOC concentration, absorption coefficient (a_{254} or CDOM), EEM spectra, and other optical properties, such as $SUVA_{254}$, spectral slope, and fluorescence indices, including HIX and BIX (Table 2-1and Figure 2-2).

For the Milwaukee River sample, bulk DOC concentration was 514 μM, and a_{254} value was 49.9 m^{-1}, both of which are within the ranges of major U.S. rivers (Spencer et al. 2012) and Wisconsin rivers (Teber

2016; DeVilbiss et al. 2016a). Values of $SUVA_{254}$ (3.51 L mgC^{-1} m^{-1}) and spectral slope ($S_{275-295}$ = 0.0170 nm^{-1}) were almost the same as those in the Fox River/Milwaukee River (3.01/3.60 L mgC^{-1} m^{-1}; 0.0197/0.0170 nm^{-1}, respectively) (Teber 2016; DeVilbiss et al. 2016a) although $SUVA_{254}$ was relatively higher than the lower Mississippi River and some other US rivers (Spencer et al. 2012; Cai et al. 2015), indicating higher aromaticity and prevailed vascular plant input in the water sample (Helms et al. 2008; Murphy et al. 2010). In terms of colloidal organic carbon abundance, 58% of the bulk DOC was measured in the >1 kDa colloidal fraction and the remaining 42% was partitioned in the <1 kDa low molecular weight (LMW) fraction (Table 1), similar to those found in other river waters (Guo and Santschi 2007).

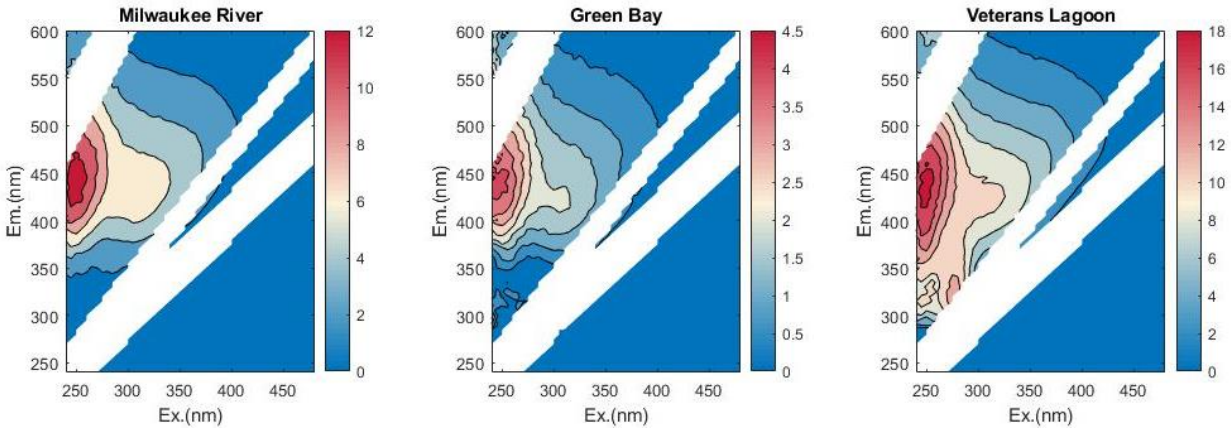

Figure 2-2. Excitation-emission matrix (EEM) spectra of the bulk water samples from the Milwaukee River (left), Green Bay (middle) and Veterans Lagoon (right).

For the mesoeutrophic Green Bay sample, both DOC concentration and a_{254} were in the ranges reported previously (Xu et al. 2018b), but significantly lower than those of the Milwaukee River sample (Table 2-1). Similar to other areas (Guo & Santschi, 1996), the $S_{275-295}$ value was slightly higher than those in river waters, indicating a relatively low in CDOM molecular weight. The colloidal DOC abundance in the Green Bay water sample was relatively high compared to that in Milwaukee River water, with 66% of DOC measured in the >1 kDa colloidal or HMW-DOM fraction (Table 2-1). For the Veterans Lagoon sample, the DOC concentration and a_{254} were relatively low compared to the other two samples. However, up to 75% of DOC was measured in the >1 kDa colloidal phase in this eutrophic water sample (Table 2-1), likely derived from algal exudates or degradation products.

The fluorescence EEM spectra of the three bulk water samples are depicted in Figure 2-2. The EEM contours of the Milwaukee River and Green Bay samples are similar, showing two obvious peaks, Peak A (ex/em = 260/400-460 nm) and Peak C (ex/em = 320-360/420-460 nm), which are related to humic-like fluorophores (Coble 2007; Murphy et al. 2010). The predominance of terrestrial humic-like DOM in the Milwaukee River sample is consistent with previous observations in the same river (Teber 2016). The EEM contour of the eutrophic lagoon sample was distinct from the other two samples (Figure 2-2). In addition to Peak A and Peak C in the sample, Peak T (ex/em = 275/340 nm) was another major peak, indicating a significant contribution of autochthonous DOM to the eutrophic lagoon sample.

Fluorescence indices (i.e. BIX and HIX) derived from EEM data have been used to explore potential DOM source that might be ignored using empirical peak-based method (Zsolnay et al. 1999; Huguet et al. 2009; Yang et al. 2012). A HIX <4 can be regarded as DOM source with scarce humic materials, while a BIX of 0.6-0.7 is related to DOM source with mediocre level of autochthonous materials (Huguet et al. 2009). As shown in Table 2-1, the value of HIX was 4.31 in the river water sample and 4.82 in the Green Bay water, which are similar to those previously observed in the Milwaukee River (Teber 2016), the Fox River, and Green Bay (DeVilbiss et al. 2016a), but higher than the HIX value measured for the Veterans Lagoon sample and those reported previously for the oligotrophic Great Lakes (Zhou et al. 2016a). The value of BIX was 0.60 for the river water sample, which is slightly lower than those measured for the Green Bay and Veterans Lagoon samples (Table 2-1) or the Great Lakes (Zhou et al. 2016a), but similar to those measured for the Milwaukee River (Teber 2016) and the Fox River (DeVilbiss et al. 2016a). Intermediate HIX and BIX values in the river and bay waters suggested a mixed DOM pool from both prevailed humic-like materials and autochthonous components derived from microbial transformation, while low HIX value in the eutrophic lagoon sample pointed to a predominant autochthonous DOM source.

2. 4. 2 Continuous size distribution of colloidal DOM characterized using FlFFF

The FlFFF technique is capable of simultaneous size separation and characterization depending on detectors. Fractograms derived from the FlFFF analysis are shown in Figure 2-3. Colloidal DOM in the river water had high UV_{254} absorbance and fluorescence intensities (both humic-like and protein-like) due

to its higher DOC concentration and $SUVA_{254}$ value compared to the other two samples. The UV_{254}-absorbance and both humic-like and protein-like fluorescent components (denoted by $Fluo_{350/450}$ and $Fluo_{275/345}$, respectively) have a major peak in the <10 kDa size range. In addition, there existed a secondary peak for protein-like fluorophores at the >100 kDa (Figure 2-3 and Figure 2-4). Although similar fractograms are observed among these three samples, their specific peak locations and intensities in UV_{254} absorbances and fluorescent components are clearly different (Figure 2-3).

When the continuous fractograms (Figure 2-3 with log-scaled molecular weight) were integrated and then divided into selected size-classes, e.g., the <1 kDa, 1-3 kDa, 3-10 kDa, 10-50 kDa, 50-100 kDa, and >100 kDa (Figure 2-4), the percentages of each size-class in the total integrated signal could be calculated. For example, UV_{254} absorbance was highest in the 3-10 kDa size-range for the river and bay water samples and in the 1-3 kDa size-range for the eutrophic lagoon sample (Figure 2-4). Similarly, Peak-C related fluorophores or humic-like substances (i.e. $Fluo_{350/450}$) were mostly partitioned in the 1-3 kDa size-range for all three samples followed by the 3-10 kDa and the <1 kDa size-fractions, with the >10 kDa size-fractions being the least (Figure 2-4). For Peak-T related fluorophores or protein-like substances (i.e. $Fluo_{275/340}$), however, the major molecular size-fraction was the >100 kDa or the 100 kDa-700 nm in all three samples, followed by the 1-3 kDa and 3-10 kDa or <1 kDa size-fractions (Figure 2-4). For both humic-like and protein-like substances, the intensities of different DOM size-fractions consistently decreased with increasing molecular weight within the size-range between 1 and 100 kDa, although there is a significant difference in the continuous molecular size distribution between humic-like and protein-like substances (Figure 2-4). The seemingly difference in the DOM size distribution between the log-scaled fractograms (Figure 2-3) and the integrated size-classes (Figure 2-4) are largely due to the amplification of signals of the lower molecular size-range (~1 kDa) and compression of those in the >10 kDa size-fraction (Figure 2-3). Similar DOM size distributions derived from FlFFF analysis have also been observed in other aquatic environments.

Overall, the humic-like substances (Ex/Em at 350/450 nm) are mostly partitioned in the 1-3 kDa size-range while the protein-like substances (Ex/Em at 275/340 nm) are largely present in the >100 kDa size-fraction with a bimodal distribution with a second peak at the 1-3 kDa size-fraction.

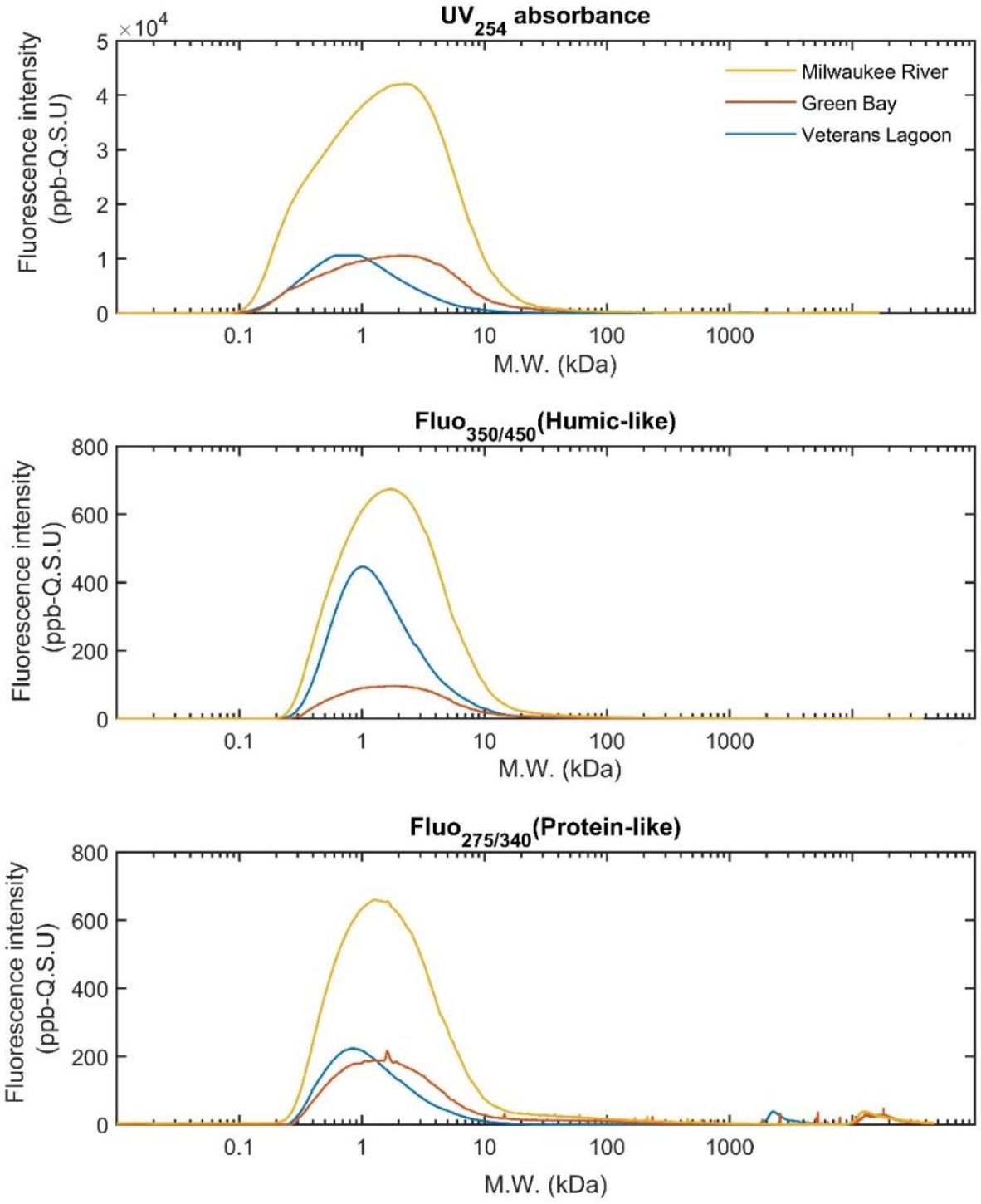

Figure 2-3. FlFFF fractograms of the three colloidal DOM samples, including UV absorbance at 254 nm (top), fluorescent humic-like at Ex/Em = 350/450 nm (middle) and fluorescent protein-like at Ex/Em = 275/340 nm (bottom). Notice that a log-scale was used here for x-axis or molecular weight (MW).

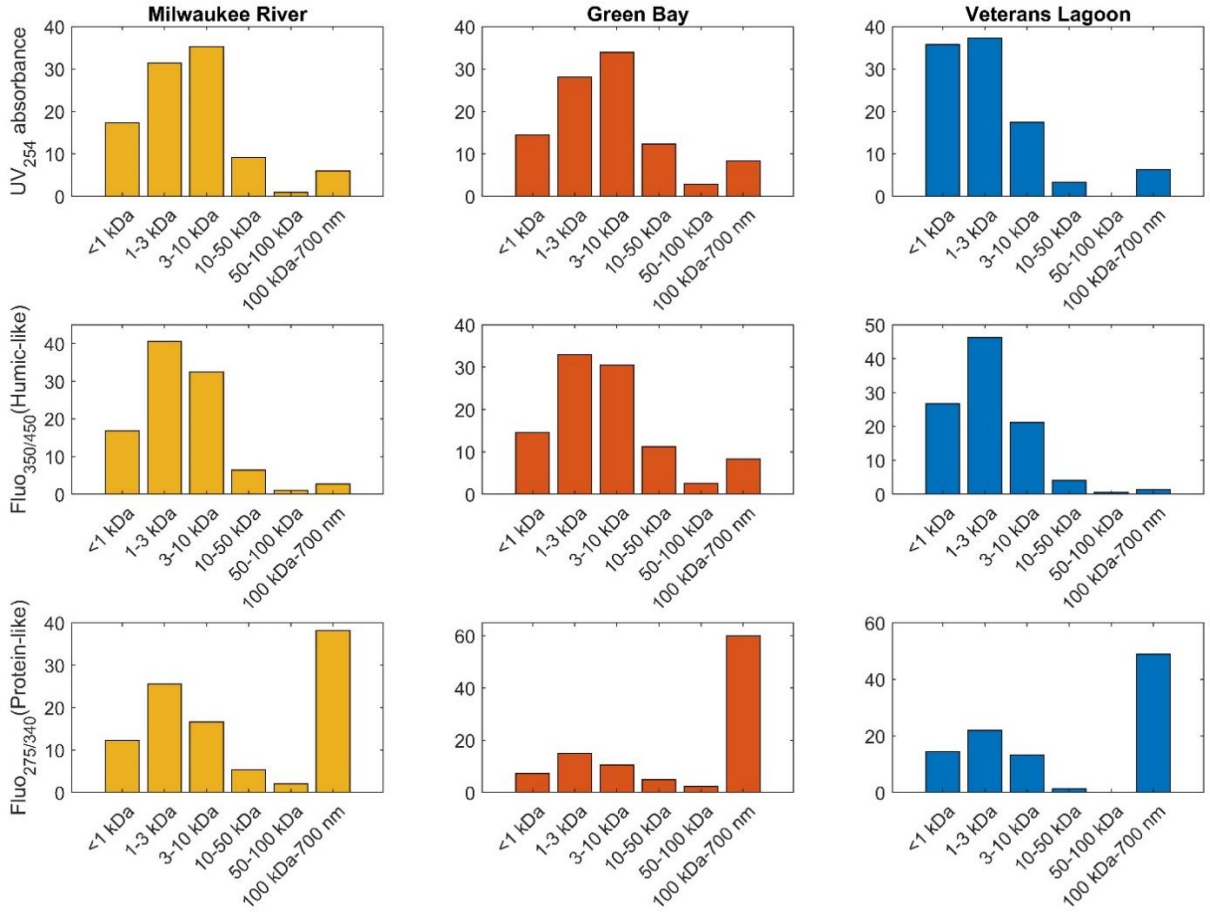

Figure 2-4. Size distributions of optical properties among the integrated size intervals in samples collected from the Milwaukee River (left), Green Bay (middle), and Veterans Lagoon (right), including UV absorbance at 254 nm (top), fluorescent humic-like at Ex/Em = 350/450 nm (middle), and fluorescent protein-like at Ex/Em = 275/340 nm (bottom). The seemingly difference in the distribution of larger size-classes between the continuous fractograms (log-scaled, Figure 2-3) and the integrated fractograms (linear scale, this figure) was largely due to the shrinking of signals in the size-fractions >10 kDa.

2. 4. 3 Molecular size-dependent EEM spectra

As shown in Figure 2-3, the FlFFF technique coupled with online detectors provides a continuous size distribution of bulk DOM. Offline measurements on the FlFFF size-fractionated subsamples using 3D spectrofluorometer, on the other hand, offer EEM spectra of different DOM molecular size-fractions in the same sample, revealing continuous changes in fluorescence EEM characterization from small molecular size to large molecular size (>100 kDa) fractions within individual samples (Figure 2-5).

Examples showing the application of FlFFF combined with offline EEM measurements and changes in EEM spectra with DOM molecular weight in the river water sample are given in Figure 4. Indeed, the results demonstrated clearly a change in fluorescence EEM characteristics with increasing DOM

molecular size within the specific individual water sample (Figure 2-5). For example, the fluorophores characterized with Peak A and Peak C were dominantly present in the molecular size-fractions between 3-100 kDa, then faded away gradually with increasing size.

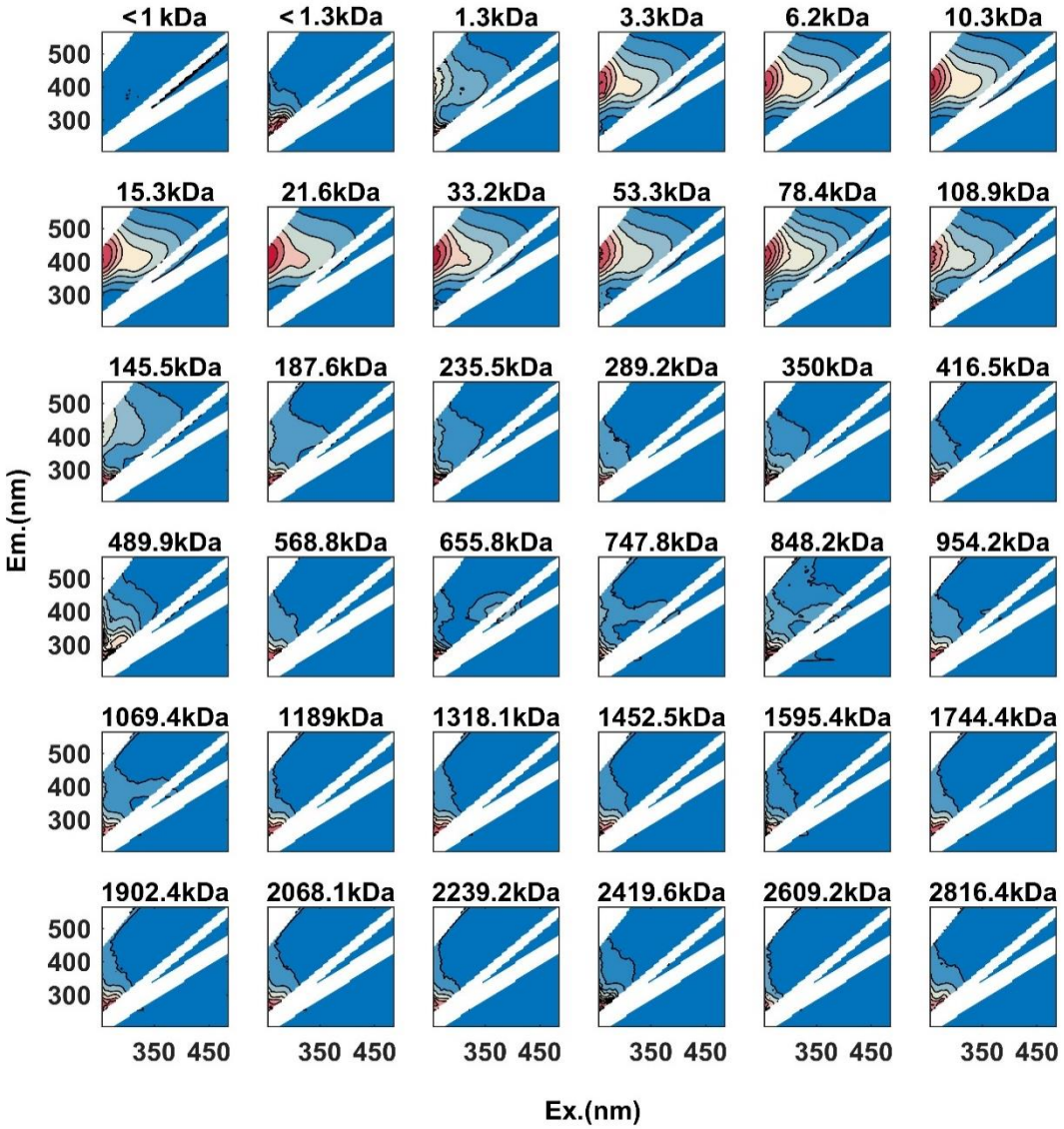

Figure 2-5. Examples of EEM spectra in FlFFF-generated size-fractions along the size continuum, showing dynamic changes in the fluorescence EEM characteristics with molecular weight within the individual DOM sample from the Milwaukee River.

Peak M (Ex/Em = 290-310/370-410 nm) appeared between ~650-850 kDa although somewhat weak. Part of the peak was masked by the first order Rayleigh and Raman scatter peaks of water, making it difficult to be identified. Its intensity was relatively low compared with Peak B (Ex/Em = 275/305 nm).

Fluorophores associated with Peak B (Ex/Em=275/305 nm) appeared mostly in the > 78 kDa. Intriguingly, the humic-like substances, including Peaks A and C (Zhou and Guo 2015), were largely partitioned in the lower MW size-fractions (<100 kDa), while the protein-like substances, such as Peaks B and T were present in the higher MW DOM fractions (>78 kDa). These unique size distributions between humic-like and protein-like DOM, revealed by the FlFFF-EEM technique, are consistent with their characteristics in molecular size, surface-reactivity (such as dispersiveness, aggregation, self-assembly), and environmental behavior (von Wandruszka 2000; Perminova et al. 2003). For example, humic substances have been shown to be detergent-type highly dispersive organic matter (von Wandruszka 2000), while proteins can be formed from aggregation of amino acids and peptides and contain mostly HMW materials (Perminova et al. 2003). Thus, the application of FlFFF-EEM and the resultant molecular size-dependent EEM spectra in individual bulk DOM samples as exemplified in Figure 2-5 allow the identification of not only the variation in the relative importance of different fluorescent components with molecular weight, but also the less visible minor fluorophores in the bulk EEM spectra.

2. 4. 4 PARAFAC-derived DOM components for individual samples

PARAFAC analysis was performed for each water sample based on the EEM data of FlFFF-derived size-fractionated samples, resulting in three different models that were validated by half-split analysis. Figure 2-1 shows the sum of squared error from excitation and emission spectra for each sample. As shown in Figure 2-6, four different fluorescent DOM components were identified for the river water and the eutrophic lagoon water, but only three fluorescent components were identified for the mesotrophic bay water, demonstrating different major fluorescent DOM components among the three water samples.

Within the four components in the river water sample, three (C1, C2 and C3 or C_{445}, C_{515}, and C_{405}) are humic-like materials with their Ex/Em maxima at 255/445 nm, 250/515 nm, and 250/405 nm, respectively. They are related to Peak A and Peak C (Coble et al. 1998; Coble 2007). These three terrestrially derived DOM components have been widely observed in aquatic environments (Stedmon and Markager 2005; Guo et al. 2014; Zhou et al. 2016a; DeVilbiss et al. 2016a). The fourth component (C4 or C_{305}) is identified as protein-like materials with the Ex/Em centered at 250/305 nm (Table 2-2), which is

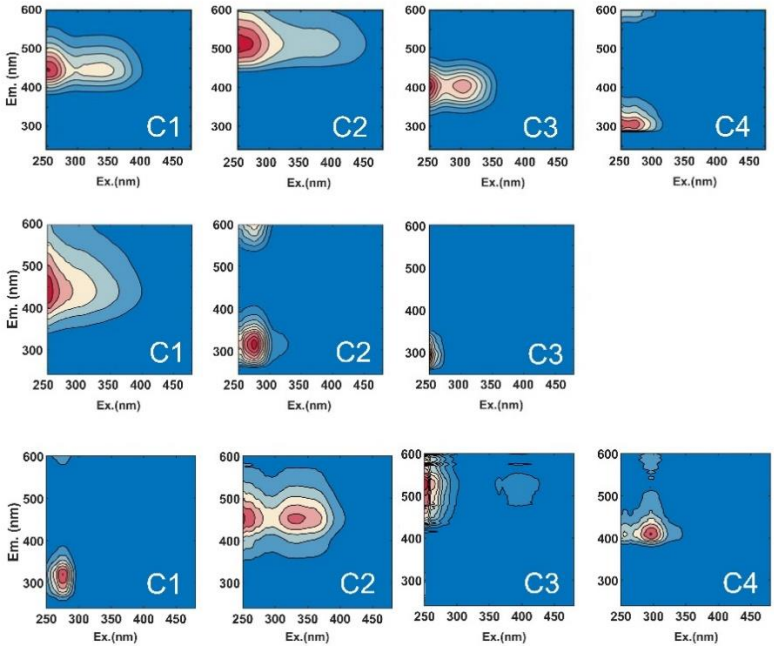

Figure 2-6. Contour plots showing distinctive PARAFAC-derived major fluorescent DOM components for three different water samples from the Milwaukee River (top), Green Bay (middle), and Veterans Lagoon (bottom) using FlFFF and EEM-PARAFAC analysis (see Table 2-2 for the characteristic Ex/Em wavelengths of each DOM component).

empirically regarded as the combination of Peak B and Peak T from autochthonous sources (Coble 2007). However, it has the least contribution (15.9%) compared to those humic-like components, which contributed up to 84.1% of the total fluorescence intensity (Table 2-2). Within the three fluorescent components in the mesotrophic bay water, the predominant component is the UVC humic-like materials mainly from terrestrial input (Coble 1996). The other two are tryptophan-like and tyrosine-like components, which are typical protein-like materials in natural waters (Figure 2-6 and Table 2-2). For the eutrophic lagoon water sample, protein-like component (Ex/Em at 275/320 nm), contributing 28.2 % of the total fluorescence, is the dominant component related to tyrosine-like (Ex/Em = 275/305 nm) and tryptophan-like (Ex/Em = 275/340 nm). The other three fluorescent components are humic-like components (Figure 2-6 and Table 2-2).

Quantitative comparisons in the PARAFAC-derived components between the three samples using the TCC method show that only the C_{445} fluorescent component was found common in all three samples (with TCCex/em values >0.90) although all the PARAFAC-derived fluorescent components have been reported previously (Coble, 2007; Murphy et al., 2018; Wünsch et al., 2017, 2019). This again manifests

the uniqueness in major DOM components among samples collected from different aquatic environments. When the Ex/Em loadings from these components were uploaded for online comparisons using the OpenFluor database (Murphy et al., 2014) and the thresholds for the Ex/Em minimum similarity scores were set at 0.95, the loadings of C1, C2, C3 and C4 (or C_{445}, C_{515}, C_{405} and C_{305}) from the Milwaukee River sample match components in 40, 34, 37, and 6 cases/datasets, respectively. In addition, comparing components C_{445}, C_{515}, C_{405} and C_{305} identified for the Milwaukee River with those (F_{400}, F_{420}, F_{450} and F_{520}) reported in Murphy et al., three out of the four components (i.e. C_{445}, C_{515} and C_{405}) are highly congruence with their F_{450}, F_{520} and F_{400}, respectively ($TCC_{ex/em}$ >0.95). These three humic-like components (C_{445}, C_{515} and C_{405}) are also highly similar with those derived from HPSEC-EEM-PARAFAC method for the Rio Negro sample ($TCC_{ex/em}$ >0.95, Wünsch et al., 2017). Fewer cases reported the same protein-like component (C4 or C_{305} here), probably due to the low signal/noise ratio in the UV-A area and sample's unique DOM regime under different hydrological and/or trophic conditions.

To further validate the new approach, the "mass balance" of fluorophores between the bulk and sub-fractional samples was evaluated using comparisons between the integrated EEM spectra from all subsamples and the EEM spectra from the bulk water sample. As shown in Figure 2-6 for the Milwaukee River sample, the integrated EEM and the bulk EEM spectra resemble each other in both fluorescence intensity and spectral characteristics. Fluorescence intensities of major peaks (Peaks A, C, and M) recovered from all subsamples comprised up to 98.4%, 99.5% and 99.8%, respectively, of their corresponding intensities in the bulk sample (Table S4). Differences between the integrated EEM and the original EEM spectra are reflected in the residuals given in Figure 2-7. In the major peak areas (i.e. Peak C and Peak A; Em > 400 nm), the residuals are negligible since their fluorescence signals are high and the spectral corrections are reliable. Outside the major peaks in the EEM spectra, some residuals are visible due to their low intensities, but they are all within ±10% and mostly observed in the UV-A area (Em <400 nm, Figure 2-6). Negative residuals in Em at 320-400 nm area are probably related to the loss of DOM through sorption onto the FlFFF system (e.g. tubing and membrane). On the other hand, positive residuals below the 320 nm likely result from the low signal/noise ratio in the specific wavelength region. Overall, the spectra of the integrated EEM from size-fractionated sub-samples and the bulk EEM spectra from the

Table 2-2. PARFAC-derived major fluorescent DOM components for each sample and their relative contributions to the total fluorescence within individual water samples.

Components	Ex. (nm)	Em. (nm)	Description	Contribution (%)
Sample from the Milwaukee River				
C1	255	445	Humic-like	34.9
C2	250	515	Humic-like	23.9
C3	250	405	Humic-like	25.2
C4	250	305	Protein-like (Tryptophan-like)	15.9
Sample from Green Bay				
C1	250	445	Humic-like	44.1
C2	275	315	Protein-like (Tryptophan-like)	31.9
C3	250	290	Protein-like;	24.0
Sample from Veterans Lagoon				
C1	275	320	Protein-like (Tryptophan-like)	28.2
C2	250/335	455	Humic-like	28.4
C3	250/390	530	Humic-like	19.0
C4	295	410	Humic-like	24.4

original sample closely resemble each other and the residuals are in general negligible, attesting to the applicability and validity of the approach combining FlFFF and EEM-PARAFAC.

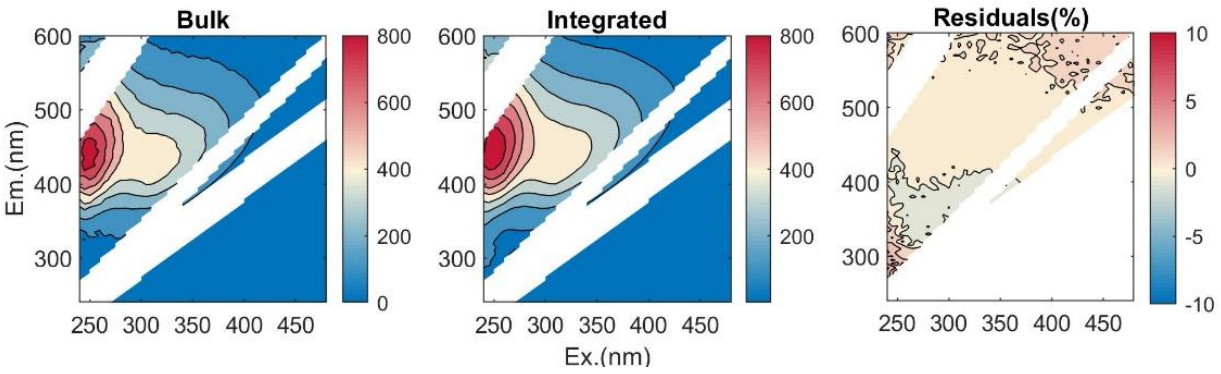

Figure 2-7. EEM spectra of the bulk water sample (left panel) and the integrated subsamples (middle panel) for the Milwaukee River sample, and mass-balance or EEM residuals (right panel) between the bulk and the integrated-subsamples.

2. 4. 5 Changes in DOM properties with molecular weight in individual samples

As shown in Figure 2-7 for the riverwater sample, the three humic-like components (C_{445}, C_{515} and C_{405}) all presented mostly in the <20 kDa size-range, with a minor peak in the 25-45 kDa and negligible in the >45 kDa. For the protein-like component (C_{305}), its relative abundance increased in general with molecular weight, with peaks mostly between ~30 kDa and >100 kDa. High protein-like to humic-like ratio (i.e. $C_{305}/(C_{445}+C_{515}+C_{405})$) was found in the <0.3 kDa and >20 kDa size-ranges, indicating protein-like DOM was present mainly in either lower MW <0.3 kDa or higher MW >20 kDa, especially the >100 kDa size fraction. In contrast, humic-like components occurred preferentially in the 0.3-20 kDa size-range.

Similarly, fluorescence indices also demonstrated a dynamic variation with molecular weight within the sample (Figure 2-8c&d), showing high humified DOM components had a MW of ~6 kDa and elevated BIX occurred in the <0.3 kDa and >100 kDa size-ranges. This bimodal distribution of BIX resembled those of protein-like components and the $C_{305}/(C_{445}+C_{515}+C_{405})$ ratio (Figure 2-8b). Similar results are shown for the other samples in Figure 2-9 and Figure 2-10.

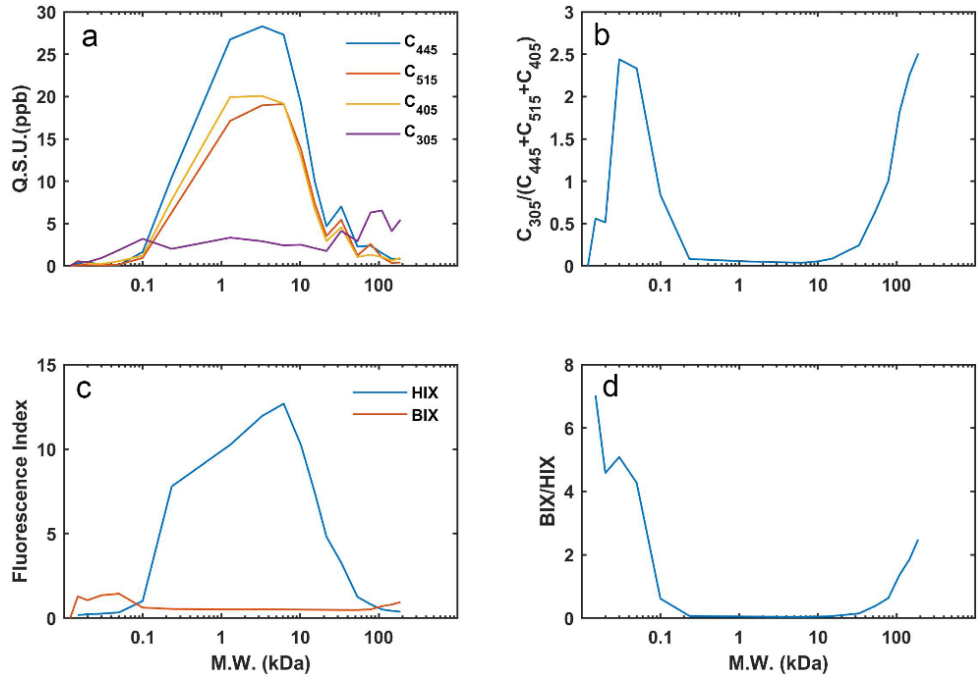

Figure 2-8. Variations in the four PARAFAC-derived fluorescent components (a), fluorescence intensity ratio of C_{305} to (C_{445}+C_{515}+C_{405}), an indicator for the relative contribution of protein-like substances (b), fluorescence index (HIX and BIX) (c), and the BIX/HIX ratio (d) with molecular weight (M.W.) within the Milwaukee River water sample (see also Figure S4 for changes in the distribution of PARAFAC-derived fluorescent components with molecular weight).

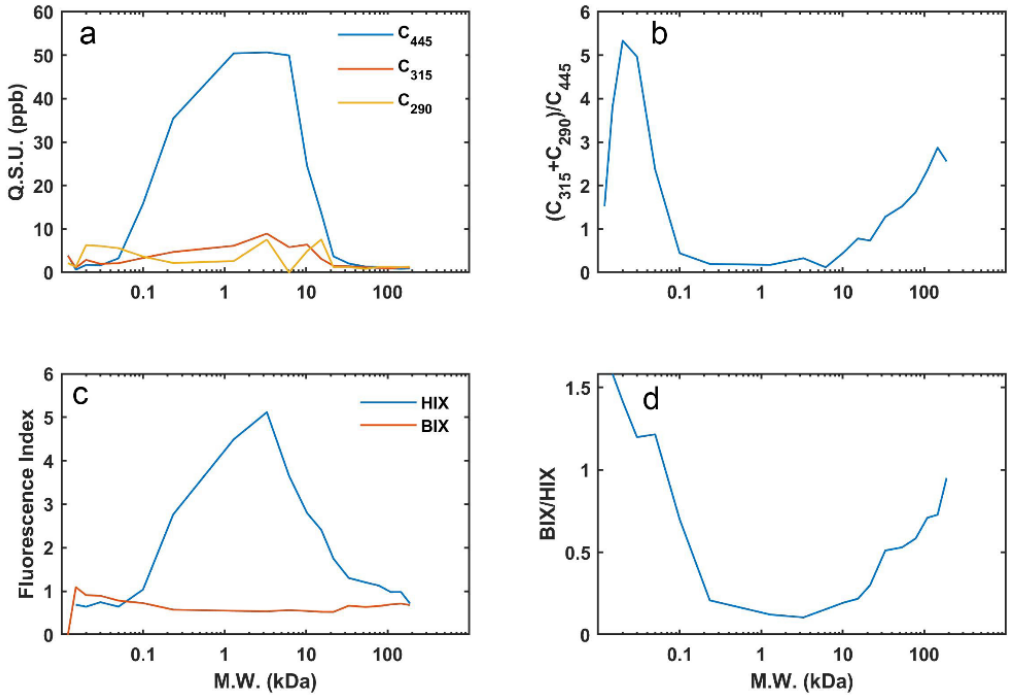

Figure 2-9. Changes in the three PARAFAC-derived fluorescent components (a), fluorescence intensity ratio of (C_{315}+C_{290}) to C_{445} (b), fluorescence index (HIX and BIX) (c), and the BIX/HIX ratio (d) with increasing molecular weight within the Green Bay water sample.

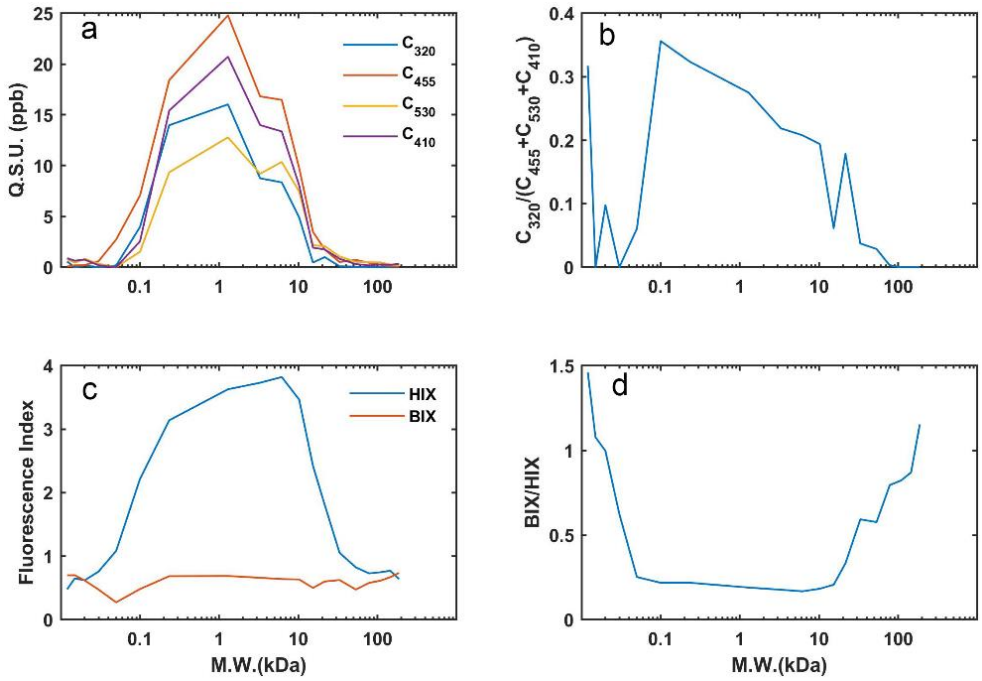

Figure 2-10. Changes in the four PARAFAC-derived fluorescent components (a), fluorescence intensity ratio of C_{320} to (C_{455}+C_{530}+C_{410}) (b), fluorescence index (HIX and BIX) (c), and the BIX/HIX ratio (d) with increasing molecular weight within the Veterans Lagoon water sample.

These detailed features showing dynamic changes in the size-dependent distribution of optical properties and PARAFAC-derived DOM components within an individual sample could not have been revealed by the EEM spectra of bulk samples. Therefore, the combination of FlFFF and EEM-PARAFAC provides a compelling approach to the characterization of DOM composition and molecular size for individual samples. When applied to samples in the aquatic continuum, they will provide new insights into the source, environmental fate, mixing behavior, and cycling pathways of DOM, especially at the groundwater-surface water, river-lake and land-ocean interfaces.

Chapter 3. Size and composition of dissolved organic matter in a negative estuary of the Fox River -Green Bay plume

3. 1. Abstract

The size and chemical composition of dissolved organic matter (DOM) were first investigated in the Fox River-Green Bay estuary, a negative estuary where salt concentration decreases from river to open bay waters. Water samples were collected along a transect from the lower Fox River to Green Bay in the summer of 2018 and individual samples were characterized using a newly developed analytical approach coupling the Flow Field-Flow Fractionation separation with excitation emission matrix – parallel factor analysis modeling to elucidate variations in DOM size spectra and molecular composition along the aquatic continuum. Concentrations of dissolved organic carbon, chromophoric-DOM (in terms of absorption coefficient), specific UV absorbance (or aromaticity), and humification index all decreased consistently from river to estuary, while spectral slope and biological index increased from the river to open bay. Using the novel one-sample EEM-PARAFAC approach, four fluorescent components, including two humic-like (C_{475} and C_{410}) and two protein-like components (C_{320} and C_{290}), were identified in the study area. While humic-like fluorescent components (C_{475} and C_{410}) decreased from the river to open bay water, one of the protein-like components (C_{290}) increased along the river-bay transect, demonstrating a dynamic change in DOM source and composition along the river-bay transect. Furthermore, the percentage of the major humic-like component (C_{475}) remained consistently similar, while the second humic-like component (C_{410}) decreased along the transect and protein-like components (C_{320} and C_{290}) consistently increased from the river to open bay waters, showing a continuous and dynamic variation in DOM composition and size spectra across the river-bay interface along the trophic gradient, which can only be revealed by the new one-sample EEM-PARAFAC approach.

3. 2. Introduction

Dissolved organic matter (DOM) in aquatic environments plays an important role in the global carbon cycle because of its huge amount and dynamic reactivity (Laird and Scavia 1990; Hansell and Carlson 2002). It has been recognized that DOM strongly participates in a series of critical environmental and

ecological processes particularly in biogeochemical cycling (Coble 2007; Coble et al. 2014). Previous studies showed that several biogeochemical processes were significantly affected by DOM properties. For example, the size of DOM influences the coagulation, flocculation, and precipitation of colloidal DOM and its complexation abilities with trace metals (Guo et al. 2000; Santschi 2018). Meanwhile, DOM size and composition are the most important DOM properties reflecting the reactivity and age of DOM (Santschi et al. 1995; Loh et al. 2004; Benner and Amon 2015).

The size distribution and composition of DOM have been widely investigated in marine estuaries (Guo and Santschi 2007; Baalousha et al. 2011). However, only a limited number of studies have reported the DOM size-spectra and composition in the lacustrine estuaries, much less for negative estuaries. Negative estuaries are the estuaries where ionic strength or specific conductivity decreases from river to coastal open water (Xu and Guo 2018; Yang et al. 2021). In marine estuaries, the abundance of colloidal DOM decreases with salinity from the river to open ocean while the C-14 age and bioactive components, such as carbohydrates and proteins, increase along the river-ocean transect (Guo et al. 2009; Benner and Amon 2015). Flocculation can remove aggregated colloids, leading to shifts in DOM size distribution (Stolpe et al. 2014; Zhou et al. 2016a). Meanwhile, based on the double layer model of colloids (Laboratories 1939), an increase in ionic strength of ambient solution would compress the size of colloids and *vice versa*. Thus, it was expected that a decrease in ionic strength would change the thickness of the double electric layer of colloidal DOM. Therefore, unlike the DOM in marine estuaries, the size of DOM is expected to increase or remain consistent from river to coastal water in a negative estuary.

Currently, optical measurements (e.g., UV-visible absorbance spectroscopy, infrared spectroscopy, and fluorescence spectroscopy) have been massively applied on characterization of natural DOM and tracing DOM sources in aquatic environments. Particularly, fluorescence spectroscopy has been widely used since it provides information on DOM optical properties and fluorescent components for examining sources and biogeochemical cycling of DOM in the aquatic continuum. Coupling with parallel factor analysis (PARAFAC), a statistical tool to deconvolute excitation-emission matrix (EEM) spectra into specific fluorescent components, EEM delivers fingerprints information about DOM with various sources (Stedmon et al. 2003). It provides a powerful tool to characterize DOM components both quantitatively

and qualitatively. However, since PARAFAC requires a large number of samples, the performance is far away from satisfaction if with limited samples. To address this issue, a novel method, using FlFFF combined with EEM-PARAFAC to analyze individual water samples (one-sample FlFFF-EEM-PARAFAC), was developed to characterize the size distribution and fluorescent DOM components within one individual sample (Lin and Guo 2020).

In this study, the Fox River estuary, the largest freshwater estuary in the Great Lakes region, was selected as the study area to evaluate variations in DOM size distribution, optical properties, and chemical composition in this negative estuary from the lower Fox River to open waters in the southern Green Bay. Surface water samples were collected using a Niskin bottle, filtered with 0.7 μm GF/F filters, and further ultrafiltered using stirred cell ultrafiltration. The size and composition of DOM were determined using the one-sample FlFFF-EEM-PARAFAC method. Our objectives were to (1) apply the one-sample FlFFF-EEM-PARAFAC technique and compare results with those from the traditional EEM-PARAFAC method to examine the applicability of our new method; (2) evaluate changes in DOM size-spectra and composition during estuarine mixing along the lower Fox River-Green Bay transect; (3) investigate DOM sources, reactivity, and environmental fate during estuarine mixing as well as potential influencing factors in the negative estuary.

3. 3. Methods

3. 3. 1 Study area and sample collection

Green Bay is an important arm of Lake Michigan, one of the largest freshwater reservoirs on Earth. Its watershed area is about 40,000 km^2, almost one-third of the watershed of Lake Michigan. The Fox River is one of the eleven rivers and streams draining into Green Bay. Also, it is a major source of terrestrial DOM and nutrient input to the Green Bay Basin. Specific conductivity (μS/cm), an indicator of ionic strength in freshwater ecosystems, is usually higher in river water than in Green Bay due to the widely distributed limestone in the drainage river basin. In addition to being a typical negative estuary (Xu et al., 2018), the lower Fox River – Green Bay ecosystem has experienced nutrient pollution, increasing algal blooms and seasonal eutrophication and hypoxia over the past decades (Bartlett et al., 2018; Klump et al.,

2018; Robertson and Saad, 2019), making it an ideal estuary to study the environmental fate and mixing behavior of DOM and variations in the characterization of DOM size and chemical composition along the Fox River plume.

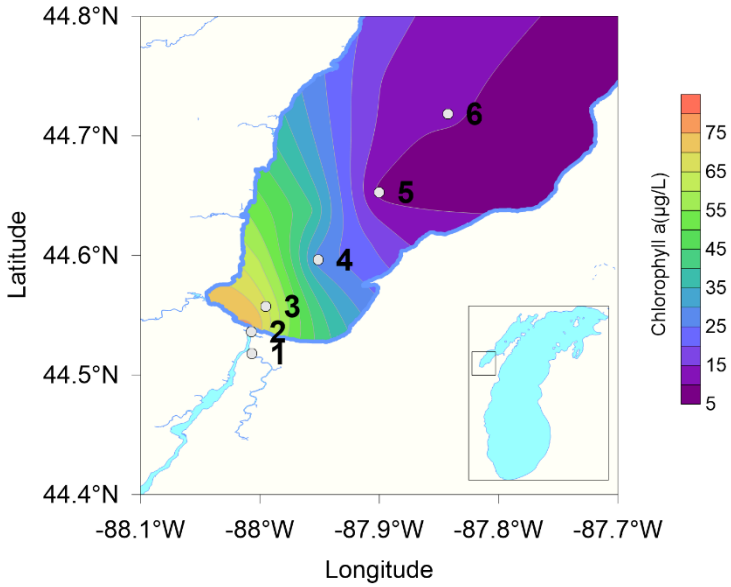

Figure 3-1. Map of the sampling area in the lower Fox River - southern Green Bay. Chl *a* concentration at the surface layer is shown as background (Data from Dr. S. Bartlett, New Water/GBMSD).

In the summer of 2018, a six-station transection from the lower Fox River to the southern part of Green Bay was investigated (Figure 3-1). The *in situ* measured specific conductivity decreased from 463.4 µS/cm to 321.8 µS/cm (Table 3-1). Station 1 (Figure 3-1) is on the downstream of the East River, a tributary draining to the Fox River main channel at the estuarine mouth. Other stations are located on the Fox River and its plume in Green Bay.

In this study, approximately 2 liters of surface water were collected using a Niskin bottle on July 10th, 2018. Water samples were transferred immediately to pre-rinsed HDPE bottles and stored in a cooler filled with ice bags. Samples were shipped back to the lab and filtered on the same day using pre-combusted (450°C with 4 h) GF/F membranes (0.7µm pore-size, Whatman). Filtrates were transferred into Nalgene HDPE bottles for optical measurements and pre-combusted glass vials for dissolved organic carbon (DOC) concentration measurements. All samples were stored in a 4 °C refrigerator overnight before stirred-cell ultrafiltration preconcentration and analytical measurements. Temperature (°C),

Chlorophyll *a* (Chl *a*) concentration (μg/L), and specific conductivity (μS/cm) were both obtained using a multi-sensor probe (YSI Sonde).

Table 3-1. Sampling locations, temperature (Temp.), specific conductivity (Sp. Cond.), pH, suspended particulate matter (SPM), and Chlorophyll a (Chl *a*) concentration for surface waters in the Fox River plume.

Station	Latitude (°N)	Longitude (°W)	Temp. (°C)	Sp. Cond. (μS/cm)	pH	SPM (mg/L)	Chl *a* (μg/L)
1	44.52	88.01	26.4	463.4	8.7	28.7	55.18
2	44.54	88.01	26.5	381.8	9.1	33.8	76.90
3	44.56	88.00	25.5	375.8	8.9	26.3	65.86
4	44.60	87.95	24.6	343.4	9.0	12.2	27.23
5	44.65	87.90	24.3	331.8	8.9	5.7	8.54
6	44.72	87.84	23.4	321.8	8.8	4.0	10.68

3. 3. 2 Ultrafiltration

To preconcentrate or isolate the > 1 kDa DOM (commonly known as colloidal DOM) for further FlFFF analysis, a stirred cell unit (Amicon 8200) with 1 kDa membrane (regenerated cellulose, YM1, Millipore) was used to process the <0.7 μm filtrate samples. Detailed ultrafiltration procedures have been described previously (Lin and Guo 2016). Briefly, the membrane was rinsed with 0.05 M NaOH and deionized (DI) water followed by membrane conditioning using about 20 ml water sample before processing the water samples. For each sample, 180 ml of the <0.7 μm filtrate was ultra-filtered to a final retentate volume of 18 ml (concentration factor = 10). During the ultrafiltration, high pressure (60 psi) zero-grade nitrogen gas (Airgas), providing air pressure, was applied on the stirred cell unit to make the < 1kDa DOM penetrate through the ultrafiltration membrane. The retentate samples were stored in a 4°C fridge until FlFFF analysis within 4 days.

3. 3. 3 Measurements of DOC and optical properties

DOC concentrations were measured by Shimadzu TOC-L analyzer. Details of DOC measurements were described in previous studies (Zhou et al. 2013). To remove inorganic carbon in water, samples were acidified to pH 2 before measurements. Calibration curves were made before sample measurements. Ultrapure water samples were measured every five samples to monitor the background signals. Standards with known concentrations were measured every ten samples for quality assurance.

UV-visible spectrophotometer (Agilent 8453) provides absorbance spectra from ultraviolet (i.e. 200 nm) to visible light (i.e. 900 nm) with a 10-mm quartz cuvette. Ultrapure water was used as a reference blank, and the refractive index effect was corrected (Weishaar et al. 2003; DeVilbiss et al. 2016b). The absorption coefficient at 254 nm (a_{254}, in m^{-1}), commonly recognized as CDOM concentration indicator due to its linear correlation with DOC concentration in terrigenous DOM dominant areas (Fichot and Benner, 2011), was calculated using the following equation: $a_{254} = 2.303 \times A_{254}/l$, where A_{254} is the absorbance at 254 nm, and l is the light path length of the cuvette (in meter). Spectral slope ($s_{275\text{-}295}$) was calculated using non-linear regression to fit the absorption spectra ranging from 275 nm to 295 nm (Fichot and Benner, 2012) based on this equation: $a_\lambda = a_{280} e^{-s_{275\text{-}295}(\lambda-280)}$, where λ is the wavelength, a_{280} is the absorption coefficient at 280 nm, a_λ is absorption coefficient at wavelength λ and $s_{275\text{-}295}$ is the spectral slope between 275 nm and 295 nm, which shows inversely correlated with DOM molecular weights (Fichot and Benner, 2012). $SUVA_{254}$ (in L mg^{-1} m^{-1}), an indicator of aromaticity, was calculated using the following equation: $SUVA_{254} = A_{254}/DOC$ (Weishaar et al. 2003). Absorption ratio between λ = 250 nm and λ = 365 nm (known as E2:E3 ratio) was used to track relative molecular size change, while the ratio of absorption at λ = 465 nm to λ = 665 nm (known as E4:E6) could be used as a CDOM aromaticity indicator (Helms et al. 2008).

EEM provides fluorescent fingerprint patterns of DOM. EEM measurements were carried out using a spectrofluorometer (Horiba Fluoromax-4, Japan) in 3D scan mode with a 10 mm quartz cuvette. Excitation wavelength (Ex) ranged from 220 nm to 480 nm with 5 nm increment, while emission wavelength (Em) ranged from 260 nm to 600 nm with 2 nm increment. DI water was scanned for each measurement batch as a blank reference. To eliminate the inner filter effect, water samples were diluted

using ultrapure water until the A_{254} is no larger than 0.02. First-order and second-order of Rayleigh scattering peaks and Raman scattering peaks were removed, and every EEM was double-checked with visual inspection (Zhou et al. 2016a). A series of quinine sulfate solutions were determined to calibrate fluorescence intensities, and quinine sulfate equivalent (QSE) was used as a dimension of the fluorescence.

Fluorescence indices derived from EEM data have been used to explore potential DOM sources that might be ignored using the empirical peak-based method (Huguet et al., 2009; Yang et al., 2012; Zsolnay et al., 1999). Biological index (BIX) and humification index (HIX) are two dimensionless indicators showing autochthonous source and allochthonous source, respectively. BIX was calculated using the following equation: $BIX = \left({I_{em=380}}/{I_{em=430}}\right)_{ex=310}$, where the I_{em} is the fluorescence intensity at a specific Em. HIX was calculated from an emission spectrum where excitation wavelength is 254 nm: $HIX = \left({I_{em=435-480}}/{I_{em=300-345}}\right)_{ex=254}$, where the $I_{em=435-480}$ is integrated fluorescence intensities from a wavelength band from 435 nm to 480 nm, and $I_{em=300-345}$ is integrated fluorescence intensities from 300 nm to 345 nm (Huguet et al. 2009).

3. 3. 4 One-sample FlFFF-EEM-PARAFAC DOM characterization

Colloidal DOM in water samples were size-fractionated and characterized simultaneously using an FlFFF system (Postnova, Salt Lake City, UT). Procedures for injections and separation were thoroughly described in previous studies (Zhou et al. 2016b; Lin and Guo 2020). Briefly, ultrafiltration-concentrated water samples were injected into the system using a rinsed plastic syringe. The injected sample was carried to the FlFFF separation channel with carrier solution, which is made of 10 mM NaCl, 5 mM H_3BO_3, and pH = 8 (adjusted using NaOH). The size spectrum was calibrated using a series of molecular-weight-known protein standards (e.g., Bovine Serum Albumin, ovalbumin, Vitamin B_{12}, etc.). Size distribution spectra were obtained using online optical detectors including a two-channel UV-vis spectroscopy (SPD-20A, Shimadzu, Japan) and two fluorescence detectors (RF-20A, Shimadzu, Japan). The online UV-Visible detector measured the absorbance at λ = 254 nm. Two fluorometers targeted at

Ex/Em = 350/450 nm ($Fluor_{350/450}$) and Ex/Em = 275/340 nm ($Fluor_{275/340}$), which empirically indicated the UVC humic-like fluorophore (Peak C) and protein-like fluorophore (Peak T), respectively.

Details of the one-sample FlFFF-EEM-PARAFAC method are thoroughly described in Chapter 2 (Lin and Guo 2020). Briefly, size-fractionated sub-samples were collected manually at the end of FlFFF separation every minute. Injections were repeated triply to collect adequate subsamples for EEM measurements. The subsample fractions were collected into acid-cleaned 5 ml centrifugal tubes and measured offline. The EEM measurement procedures for the subsample fractions are the same as the EEM measurements for the bulk samples. PARAFAC was performed to mathematically deconvolute the EEM database according to the following equation: $x_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} + e_{ijk}$. The x_{ijk} is the fluorescence signal intensity of sample *i* measured at Em = *j* and Ex = *k*. On the right side of the equation is the sum of factorized components and the model residuals (e_{ijk}). a_{if} is model scores, proportionated to the fluorescence intensity, of component *f* in the sample *i*. b_{jf}, is model-estimated excitation loadings of component *f* at $\lambda_{Em} = j$, while c_{kf} is model-estimated emission loadings of component *f* at $\lambda_{Ex} = k$. Non-negativity constraints were applied to guarantee all positive values in scores (a_{if}) and spectra loadings (b_{jf} and c_{kf}). As PARAFAC is an unsupervised algorithm, a half-split validation is a reliable approach to evaluate the model results (Stedmon and Bro 2008a). The number of components was evaluated and eventually determined by the sum of squared errors. Ten random initialized models were developed and compared using a convergence criterion setting to 10^{-7} to optimize the model.

Tucker's congruence coefficient (TCC) was used in the half-split validation. In this study, it was also used to statistically compare and identify components. It is an index to evaluate the similarity between factors calculated from the following equation: $\mathrm{TCC} = \frac{\sum XY}{\sqrt{\sum X^2 \sum Y^2}}$, where *X* and *Y* represent the excitation loadings and emission loadings from two models. Generally, a congruence coefficient of 0.90 is interpreted as a high degree of factor similarity, while a coefficient of 0.95 or higher indicates that the factors are virtually identical (Kathleen R. Murphy et al., 2014; Stedmon and Bro, 2008). To enhance the reliability of our model, the threshold of TCC in the half-split validation was set at 0.95, but it was set at

0.90 in the inter-model comparison. The drEEM toolbox (version 0.5.0) was used to perform the PARAFAC analysis on the Matlab (R2018b) platform (Murphy et al., 2013).

3. 4. Results

3. 4. 1 Bulk DOM characteristics

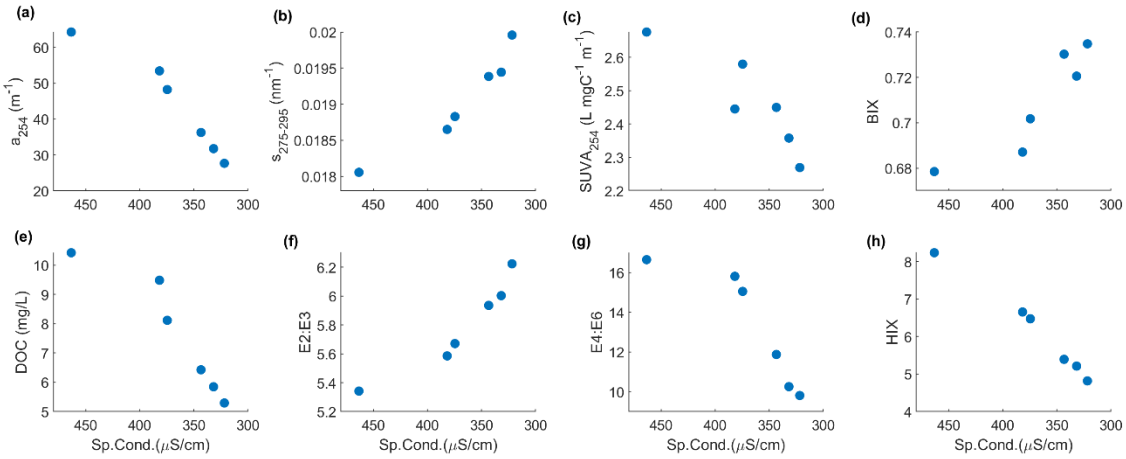

Figure 3-2. Variations in (a) UV absorption coefficient at 254 nm (a_{254}), (b) spectral slope between 275-295 nm ($s_{275-295}$), (c) specific UV absorbance at 254 nm ($SUVA_{254}$), (d) biological index (BIX), (e) dissolved organic carbon (DOC), (f) E2:E3 ratio, (g) E4:E6 ratio, and (h) humification index (HIX) with decreasing specific conductivity (Sp. Cond.) in the Fox River plume from river to open bay.

In the Fox River - Green Bay transect, DOC concentrations ranged from 441 μmol/L to 869 μmol/L, and a_{254} ranged from 27.7 m^{-1} to 64.2 m^{-1}, which are consistent with observations reported previously (DeVilbiss et al. 2016b; Xu et al. 2018b). Along the trophic gradient with declining specific conductivity, both DOC and CDOM a_{254} decreased consistently from the lower Fox River to open Green Bay, showing a conservative mixing behavior (Figure 3-2) excluding the data point from the tributary of the East River draining into the lower Fox River (Figure 1). The aromaticity indicators ($SUVA_{254}$ and E4:E6) and humification index (HIX) also decreased from the river to the bay, demonstrating the strong influence of terrestrial DOM, while the molecular weight indicators (E2:E3 and $s_{275-295}$) showed an opposite pattern (Figure 3-2). The quantitative and qualitative optical properties demonstrated that, although DOM behaved conservatively during estuarine mixing between river and bay waters, the composition of DOM evidently changed along the river-bay transect. In addition, the Chl *a* concentration in river water was exceedingly higher than those in the open Green Bay (Figure 3-1). BIX showed an asynchronous pattern; it increased gradually from river water to bay water.

3. 4. 2 Changes in DOM molecular size distribution along the river plume

Colloidal DOC (>1 kDa) was 66.0% to 76.7% of bulk DOC with an average of 71.3±4.1%, while colloidal CDOM a_{254} was 65.2% to 78.0% of bulk CDOM a_{254} with an average of 72.7±5.0% (Figure 3-3). Both colloidal DOC and CDOM a_{254} were the predominant size-fraction in the bulk DOM pool and remained fairly constant along the river-bay transect in this estuary. However, the percentage of colloidal DOC is significantly larger than those reported by Xu et al. (2018a). The major reason for the large difference is the gap between authentic membrane cutoffs and the cutoffs estimated by manufacturers. As calibrated by Xu and Guo (2017), the authentic membrane cutoff of a manufacture-rated 1 kDa membrane in a centrifugal ultrafiltration unit is 2.5 kDa. It would significantly underestimate the colloidal fraction (> 1 kDa) of DOM in natural waters since the humic-like DOM is mostly within the 1-3 kDa size fraction (Lin and Guo 2020). Moreover, different sampling seasons might be another reason for the difference. The waters from this investigation contained more autochthonous DOM from algae and phytoplankton due to summer algal blooms, while in May 2016, more soil-leached humic substances contributed to the < 1kDa and 1-3 kDa DOM fractions.

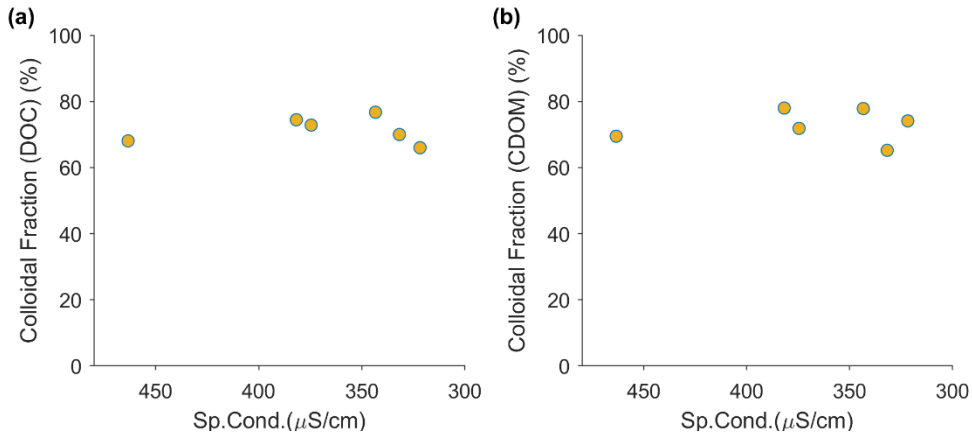

Figure 3-3. Variations in colloidal abundance (in percent) of (a) DOC and (b) CDOM a_{254} with specific conductivity in the Fox River plume and Green Bay

Size spectra from the FlFFF system showed that bulk UV_{254} (Figure 3-4) and $Fluor_{350/450}$ (humic-like, Figure 3-5) exhibited a single-narrow-peak fractogram, while $Fluor_{275/340}$ (protein-like, Figure 3-6) was a single peak with a shoulder extending toward a larger molecular size. Normalized size fractogram of UV_{254}, $Fluor_{350/450}$, and $Fluor_{275/340}$ were highly similar between samples (TCC >0.99 for UV_{254} and $Fluor_{350/450}$, TCC>0.90 for $Fluor_{275/340}$ (Table 3-2), although the major peak intensities were coherent with

the declining DOC concentration and CDOM a_{254} along the transect. The integrations of the fine-divided size intervals showed that $Fluor_{275/340}$ (protein-like) was dominated in 100 kDa -0.7μm fraction (Figure 3-6), while UV_{254} (bulk DOM) and $Fluor_{350/450}$ (humic-like) were majorly in the relatively small size ranges i.e. 1-3 kDa and 3-10 kDa (Figure 3-4 and Figure 3-5).

Table 3-2.Tucker congruence coefficients (TCC) showing the similarity between size distribution of all samples.

	UV_{254}	$Fluor_{350/450}$	$Fluor_{275/340}$
Station 1 vs Station 2	0.99847	0.99701	0.96014
Station 1 vs Station 3	0.99669	0.99936	0.97845
Station 1 vs Station 4	0.99362	0.99843	0.97649
Station 1 vs Station 5	0.99513	0.99855	0.97822
Station 1 vs Station 6	0.99600	0.99886	0.97387
Station 2 vs Station 3	0.99882	0.99535	0.95015
Station 2 vs Station 4	0.99252	0.99250	0.94318
Station 2 vs Station 5	0.99849	0.99219	0.94551
Station 2 vs Station 6	0.99851	0.99569	0.94271
Station 3 vs Station 4	0.99524	0.99900	0.97132
Station 3 vs Station 5	0.99878	0.99908	0.96819
Station 3 vs Station 6	0.99941	0.99906	0.96256
Station 4 vs Station 5	0.99220	0.99929	0.98922
Station 4 vs Station 6	0.99511	0.99889	0.98557
Station 5 vs Station 6	0.99934	0.99817	0.98678

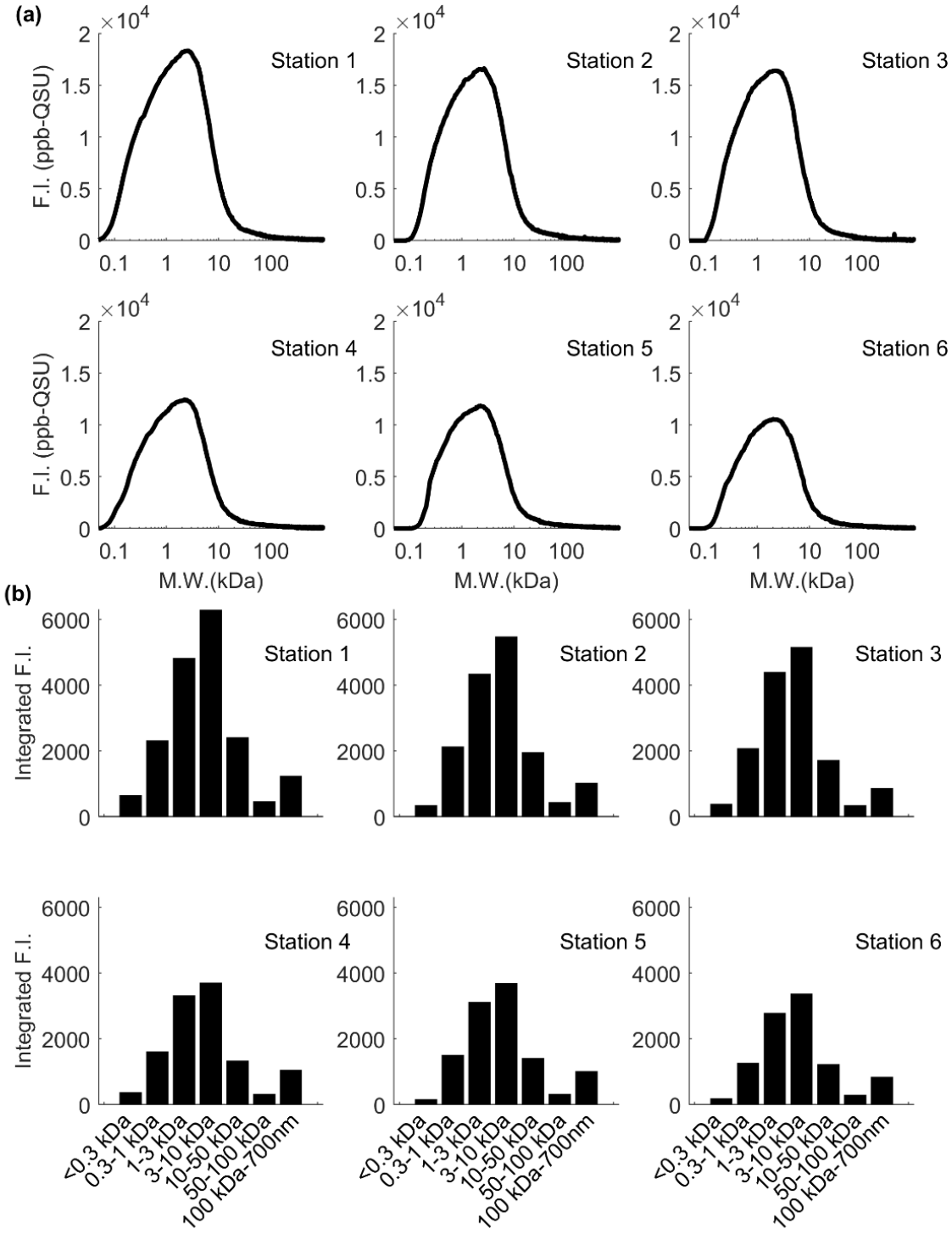

Figure 3-4. (a) Size spectra of DOM in different samples based on UV_{254} absorbance and (b) the distributions of DOM (UV_{254}) among seven continuous size intervals.

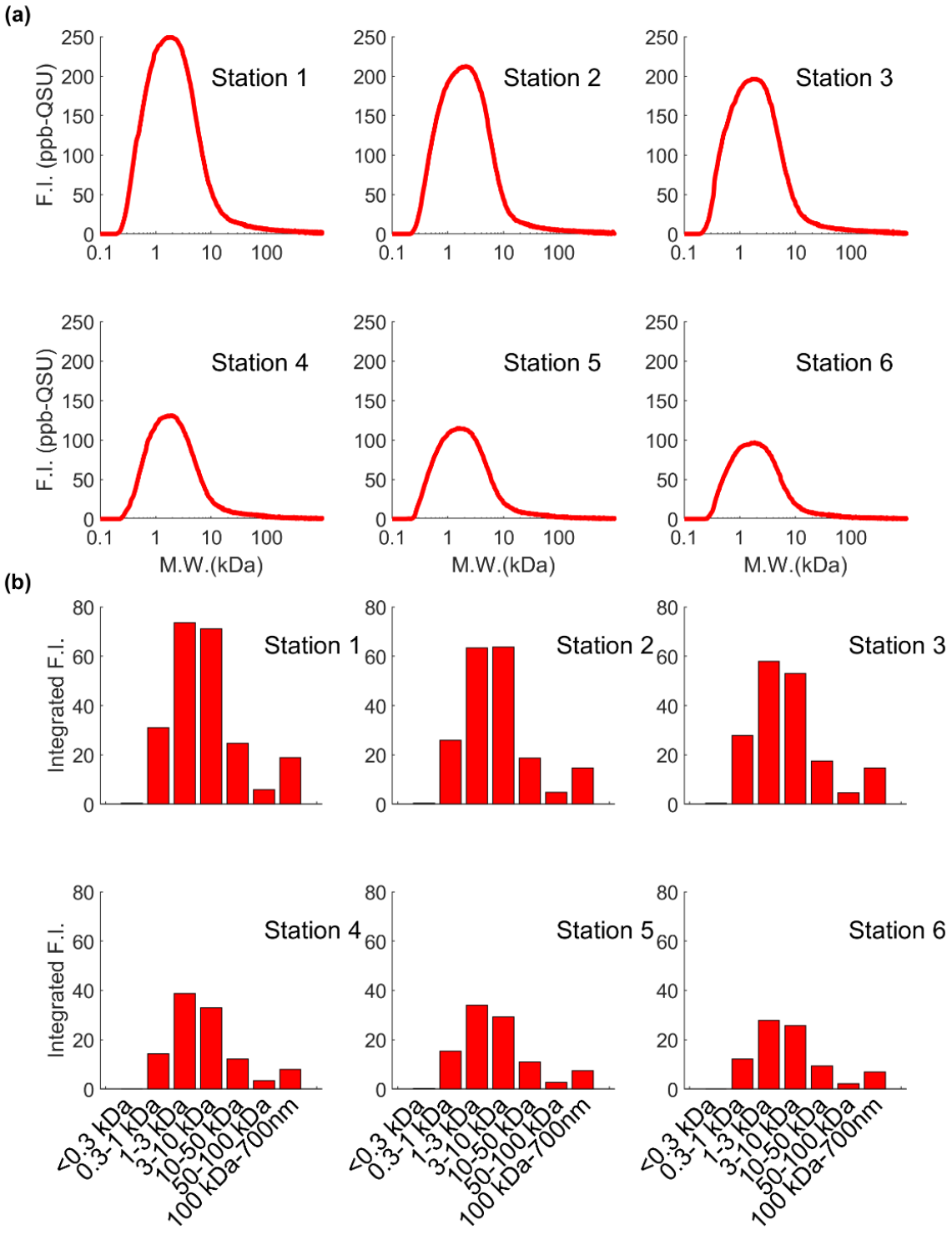

Figure 3-5. (a) Size spectra of DOM in different samples based on humic-like (Ex/Em = 350/450 nm) fluorescence intensity (F.I.) and (b) the size distributions of humic-like FDOM among seven continuous size intervals.

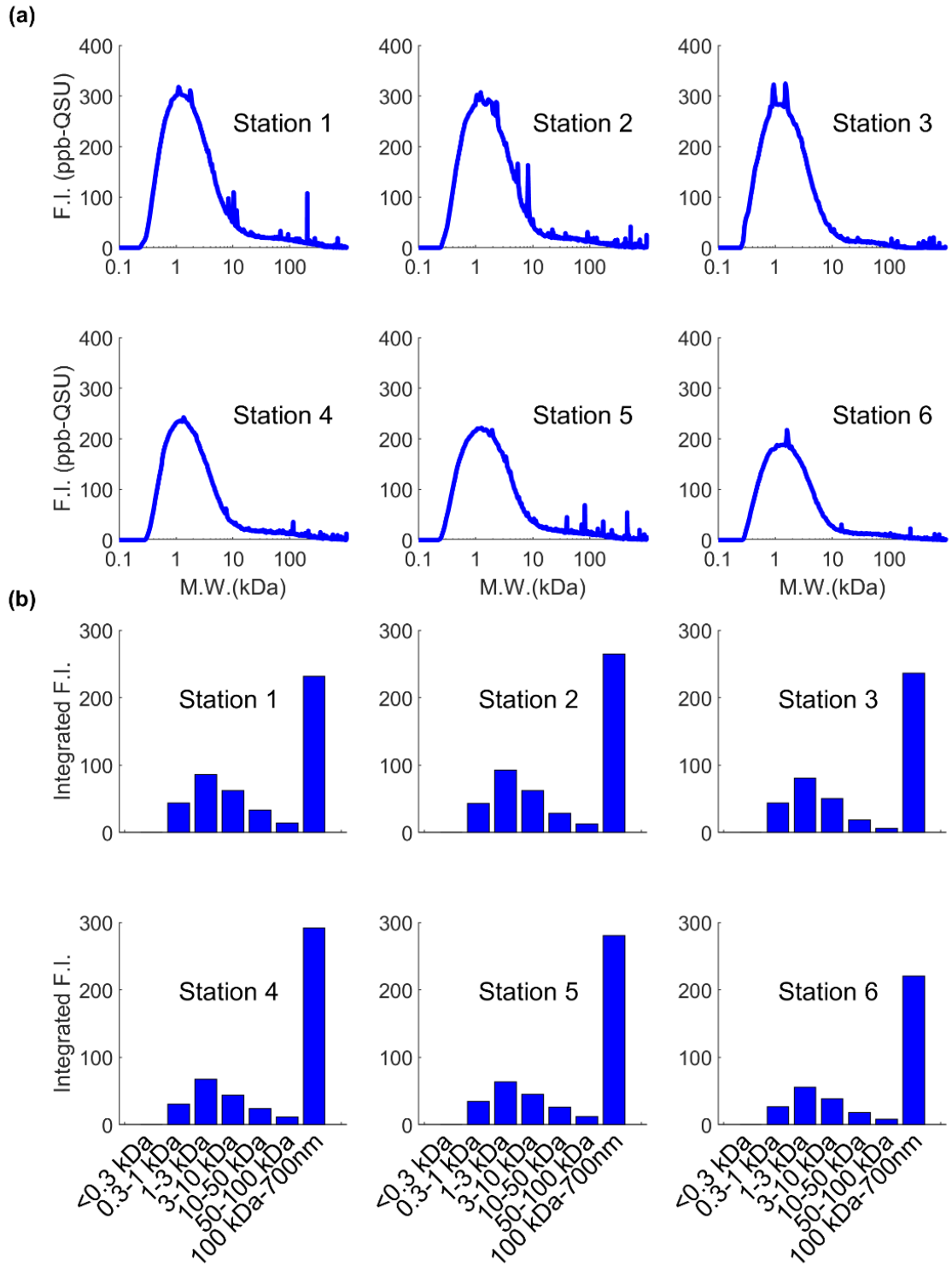

Figure 3-6. (a) Size spectra of DOM in different samples based on protein-like (Ex/Em = 350/450 nm) fluorescence intensity (F.I.) and (b) the size distributions of protein-like FDOM among seven continuous size intervals.

To compare with previous studies, the fine-divided size fractions were integrated into four major size fractions: 0.3-1 kDa, 1-3 kDa, 3-10 kDa, and >10 kDa. Obviously, conservative mixing behaviors were shown in terms of UV_{254} and $Fluor_{350/450}$; all integrated UV_{254} and $Fluor_{350/450}$ size fractions were linearly correlated with the specific conductivity (n=6, r > 0.882, p<0.02, Figure 3-7(a) and (b)). The UV_{254} in Station 6 was 54% of the UV_{254} in Station 1, while the humic-like fluorescence ($Fluor_{350/450}$) in Station 6 was 37% of $Fluor_{350/450}$ in Station 1. As for the $Fluor_{275/340}$, the largest fraction (>10 kDa) was the most abundant fraction among these size fractions (> 60%, Figure 3-7 (c)). No significant conservative dilution was clearly observed along the estuary.

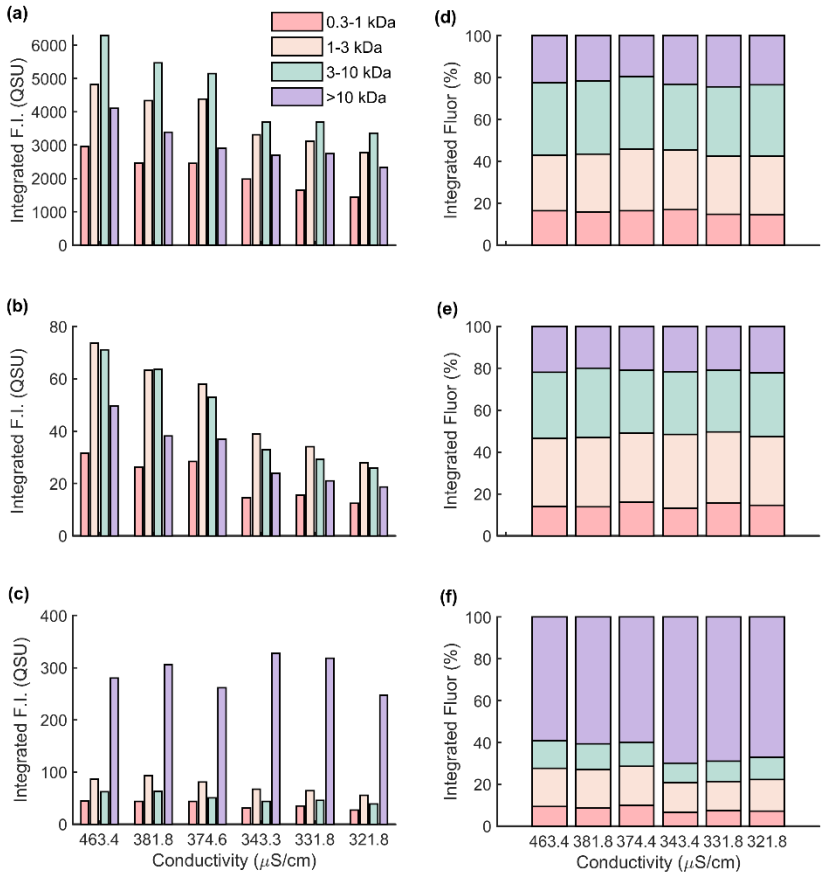

Figure 3-7. Variations in colloidal DOM abundances of (a) UV_{254}, (b) humic-like $Fluor_{350/450}$ (c) protein-like $Fluor_{275/340}$ and their relative percentages of (d) UV_{254}, (e) humic-like $Fluor_{350/450}$ (f) protein-like $Fluor_{275/340}$ in the bulk colloids (0.3 kDa-700 nm) with specific conductivity along the Fox River plume. The four different colloidal size fractions, including the 0.3-1 kDa, 1-3 kDa, 3-10 kDa, and >10 kDa, were integrated from the colloidal size spectra analyzed by FlFFF.

How much autochthonous protein-like DOM was added along the estuary? To address this question, we assumed that the part of protein-like DOM from terrestrial input diluted linearly along the estuary like

UV_{254}. To eliminate the impacts from this part of DOM, the linear diluted terrestrial $Fluor_{275/340}$ was removed from the bulk $Fluor_{275/340}$. The remaining $Fluor_{275/340}$ could be regarded as net autochthonous addition provided by algal and microbial activities. From Station 2 to Station 6, the percentage of the autochthonous $Fluor_{275/340}$ addition are 19.4%, 11.2%, 35.3%, 37.0%, and 30.0% of the bulk $Fluor_{275/340}$, which indicated that the autochthonous protein-like DOM contribution in offshore areas (> 30%, Station 4, 5 and 6) is larger than those in onshore areas (< 20%, Station 1,2 and 3).

In light of the relative fluorescent intensity of each size fraction, all three DOM indicators remained stable in the estuary (Figure 3-7(d), (e), and (f)). For all size fractions of UV_{254} and $Fluor_{350/450}$, coefficients of variations (CV) were < 10%, while for the $Fluor_{275/340}$, CVs were < 20%. It implied that for all size fractions of DOM, the size distribution remained constant through the estuary, which implied that the colloidal DOM size suffered no substantial alteration within this negative estuary.

3. 4. 3 Fluorescent EEM and PARAFAC-derived components

A similar conservative mixing behavior was also observed for the fluorescence maximum (Peak A) intensity of EEM spectra ($n=6$, $r=0.965$, $p<0.05$, Figure 3-8). Although the fluorescent DOM concentration decreased with decreasing specific conductivity, the EEM contour patterns of fluorescent DOM were highly similar between stations. Obviously, Peak A (Ex/Em = 260/400-460 nm) and Peak C (Ex/Em = 320-360/400-460 nm) are dominant on these EEM patterns, which are ubiquitous humic-like fluorophores among the freshwater ecosystems (Coble, 2007; Xu, Lin, et al., 2018; Xu and Guo, 2017; Zhou et al., 2013; Zhou, Guo, et al., 2016; Zhou, Stolpe, et al., 2016). Based on the increasing BIX along the estuary (Figure 3-2), autochthonous fluorophores were expected to emerge on EEM at the open water end, but none was observed because the fluorescence from autochthonous DOM (e.g., Peak T and Peak B) is somewhat overwhelmed by the prevailing terrestrial fluorophores. Thus, the one sample FlFFF-EEM-PARAFAC method is in need to decompose the EEM spectra and unravel the underlying components.

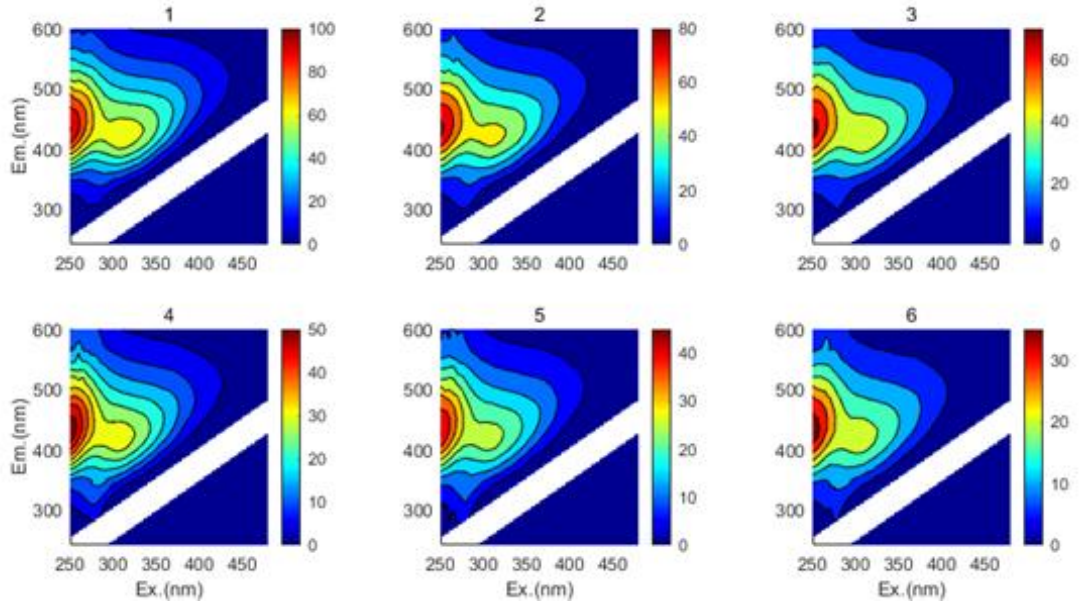

Figure 3-8. Bulk EEM spectra (<0.7 µm filtrate) from Station 1 to Station 6.

Using the one sample FlFFF-EEM-PARAFAC technique, four-component models were validated for water samples with strong riverine influence, but for sample from Station 6 in open Green Bay, only a three-component model was validated (Figure 3-9). To compare and identify these fluorescent components, inter-model comparisons and fluorescence components identifications were performed using the TCC method (Table S2). These components were named after their characteristic emission peak from their emission loadings, e.g. C_{475} for the first component in Station 1 (Stedmon and Bro, 2008). C_{475}, C_{320}, C_{410}, and C_{290} were then identified from those components (Figure 3-10). C_{475} and C_{410} are humic-like fluorescent components that were ubiquitous in natural freshwaters (Coble et al. 1998; Stedmon and Markager 2005; Murphy et al. 2018b; Wünsch et al. 2019b). According to the conventional peak pick-up method, C_{475} was a combination of Peak A and Peak C which are UVC and UVA humic-like fluorophores. Its Em loadings from all six samples were highly similar, while its Ex loading maximum varied from 435 nm to 475 nm (Figure 3-10b). The most deviated Ex loading spectra were of Station 6 model, which contained only three components. C_{410} was a combination of Peak M and Peak A. It has been reported that C_{410} was derived from microbial activities, and it correlates to DOM exported from agricultural catchments, possibly resulting from widely-used organic fertilizer on fields (Stedmon and Markager, 2005). Additionally, it is worth noting that C_{410} was not found in the water sample collected in Station 6. These two humic-like components were highly similar to F_{450} and F_{420}, which were recently

reported as ubiquitous components in aquatic environments (Murphy et al. 2018b). The other two fluorescent components are both UVB protein-like autochthonous components. C_{320} has similar fluorescence characteristics to the tryptophan-like fluorophore (Peak T) detected in several aquatic ecosystems, and C_{290} resembles the tyrosine-like peak (Peak B) in the Em < 300 nm area (Coble, 1996; Stedmon et al., 2003; Stedmon and Markager, 2005).

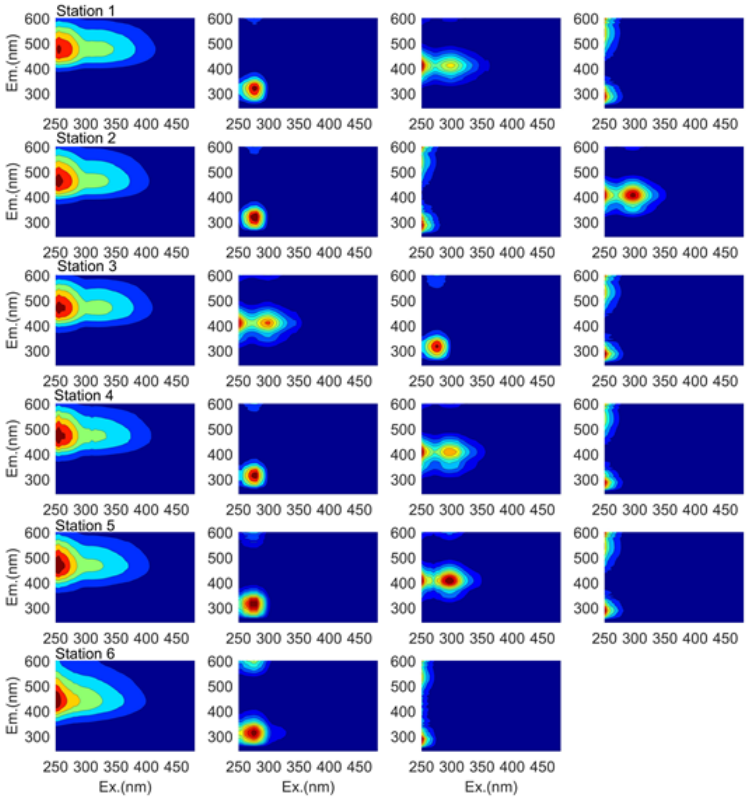

Figure 3-9. PARAFAC-derived fluorescent DOM components from individual water samples at different stations along the river-bay transect (from top to the bottom: Station 1 to Station 6).

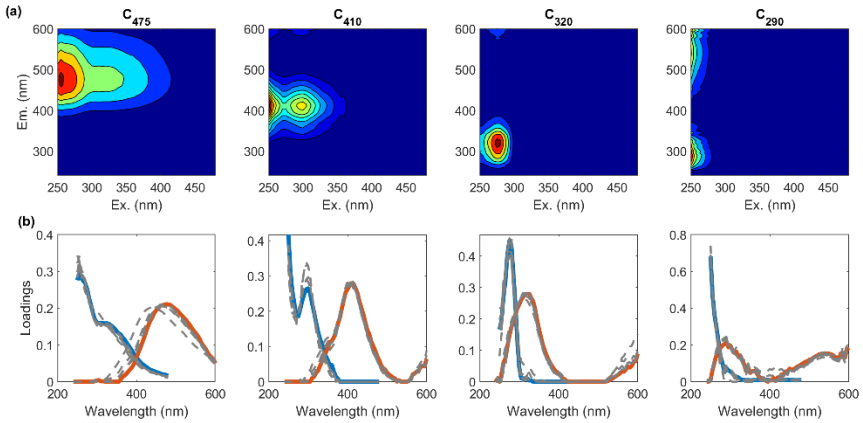

Figure 3-10. (a) The EEM contours and (b) the Ex and Em loadings for the four DOM components, namely C_{475}, C_{320}, C_{410}, C_{290}, identified from PARAFAC models. Ex loadings and Em loadings from the first sample were shown in blue line and red line. Other Ex/Em loadings were plotted with dash grey lines.

To better interpret the mixing behavior of each component in the negative estuary, fluorescence intensities of subsamples in each station were summarized. F_{max}, converted from PARAFAC model scores, represents the maximum intensity of each component. F_{max} of each component was integrated from subsamples in each station (a). The humic-like components (C_{475} and C_{410}) were overwhelming in all 6 stations, while the protein-like components (C_{320} and C_{290}) fluoresced relatively less. Significant negative linear correlations were found between summarized F_{max} ($\sum F_{max}$) and specific conductivity in C_{475}, C_{410}, and C_{320} ($n=6$, $p<0.05$, Table S3), though C_{320} is a protein-like component. The behaviors of the two humic-like components (C_{475}, C_{410}) were consistent with the decreasing HIX (Figure 3-2), despite that the C_{410} was absent in Station 6. The C_{290} increased minorly along the decreasing conductivity gradient although no significant linear correlation was found between them (Figure 3-11, Table S3).

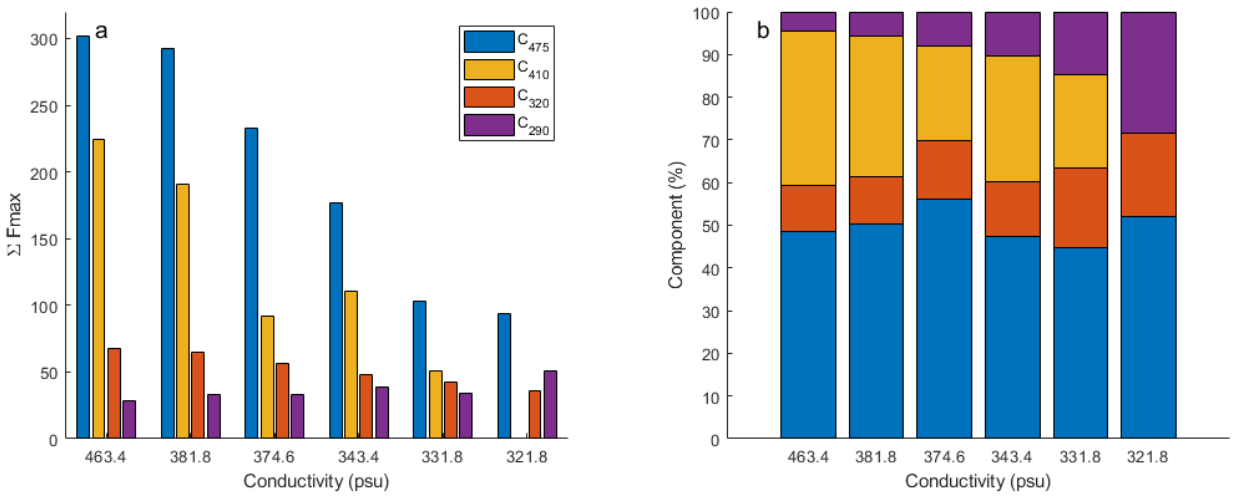

Figure 3-11. Variations in (a) F_{max} intensities and (b) percentage of each component with specific conductivity along the sampling transect from the river (Station 1) to open Green Bay (Station 6) as characterized by the FlFFF-EEM-PARAFAC technique.

Figure 3-11b shows the relative intensity of these fluorescent components from Station 1 to Station 6. The component percentage was calculated using the following equation:

$$\text{Component percentage (\%)} = \frac{\sum F_{max}}{\sum_{i=1}^{4}(\sum F_{max})_i}$$

C_{475}, varying between 44.8% and 55.2%, stayed relatively constant in the estuary. C_{410} gradually decreased from 36.0 % to 21.9 % except the station 6. Humic-like fluorescent components (C_{475} and C_{410}) were totally decreasing from 84.1 % to 52.0 %. Although they declined through the estuary, they were

still dominant in this region. The percentage of C_{320} and C_{290} were all increasing from the river to the open lake water.

3. 5. Discussion

3. 5. 1 Variations in PARAFAC-derived EEM components along the river-bay transect

This is the first time the one sample FlFFF-EEM-PARAFAC technique was applied to evaluate variations in DOM composition in the aquatic continuum. To demonstrate the advantages and novelty of the newly developed method, results were compared between the one sample FlFFF-EEM-PARAFAC coupling technique based on size-fractionated subsamples from an individual sample (Lin and Guo, 2020) and the traditional EEM-PARAFAC method using all samples from all stations in the study area (DeVilbiss et al., 2016). To address this question, all fractional subsample EEMs from all these six stations were assembled into one dataset and applied a conventional PARAFAC analysis. The final half-split validated model had four components, which were found no difference from those components identified from individual samples (TCC >0.95). Thus, Component 1 to Component 4 in this model were also named C_{475}, C_{320}, C_{290}, and C_{410}. However, the F_{max}s in each station were different from those derived from the one-sample FlFFF-EEM-PARAFAC coupling technique.

Unlike the conservative mixing behaviors of both humic-like components (C_{475} and C_{410}) from the one sample FlFFF-EEM-PARAFAC, $\sum F_{max}$ of these components deviated from the conservative mixing line (Figure 3-12a). No significant linear correlation between specific conductivity and the $\sum F_{max}$ of C_{475} (n=6, r = 0.72, p=0.11) and C_{410} (n=6, r=0.50, p=0.31) was observed. In terms of the relative fluorescence intensity, the relative humic-like C_{475} decreased along the transect, and two protein-like components (C_{320} and C_{290}) relatively increased with decreasing specific conductivity (Figure 3-12b). The relative intensity of C_{410}, recognized as a microbial derived fluorescent component, remained unchanged along the transect.

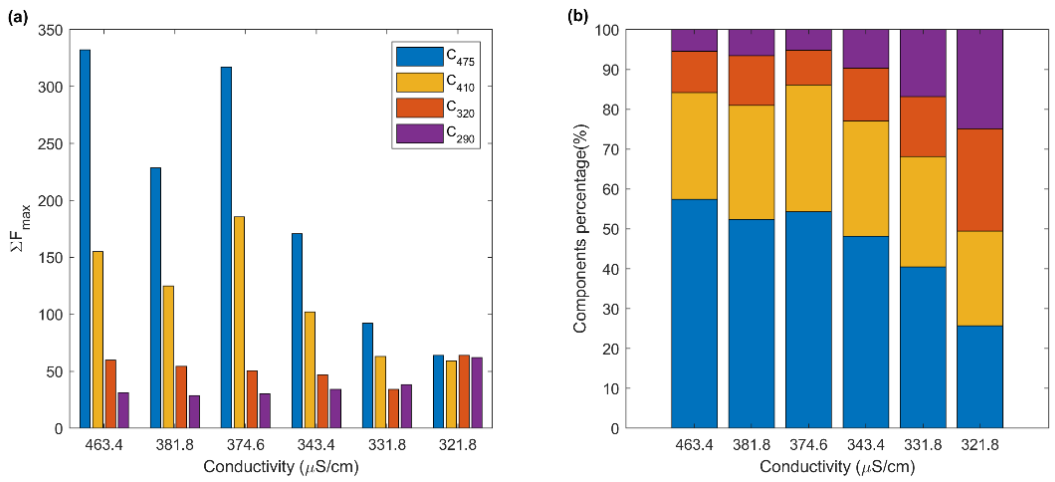

Figure 3-12. Variations in (a) F_{max} summary and (b) the percentage of each DOM component along the sampling transect from the Fox River to open Green Bay (Station 6).

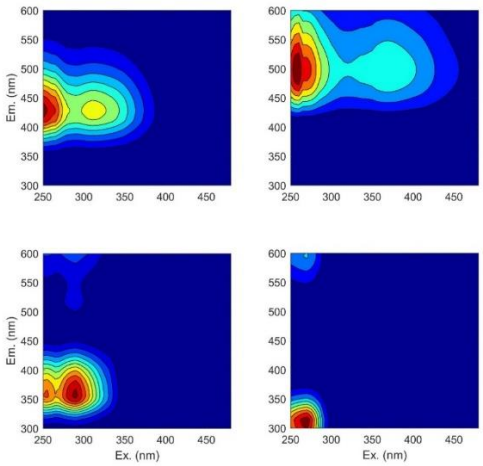

Figure 3-13. Component contours from EEM data collected during 2014 Green Bay sampling.

DeVilbiss et al. showed four components in a previous survey covering almost the entire Green Bay but less stations in the Fox River plume (DeVilbiss et al. 2016b). A four-component PARAFAC model was obtained with two humic-like and two protein-like components (Figure 3-13). Peak Maxima of excitation and emission spectra loadings were highly similar with those published components (Table S3). Here we also name the components using the emission wavelength (i.e. C_{430}, C_{500}, C_{360}, C_{320}). Compared the four components with our one sample components, we can see the difference. C_{320} are highly similar with the C_{320} (TCC > 0.95). C_{430} are similar with the C_{410} (TCC >0.90). Except these two components, no similarity was found.

Since the Fox Rive plume is a natural ecological continuum, everything herein should change gradually and consistently. Thus, we believed that results from one sample FlFFF-EEM-PARAFAC are more reliable than all subsample PARAFAC regarding to $\sum F_{max}$ or relative fluorescence intensities.

To better compare and interpret these fluorescent components, we separated them into three categories based on their biological lability: refractory, semi-labile and aquagenic. Although it has been reported that Peak A and Peak C have high photo-chemical reactivities (especially Peak C, impacted by predominant UVA in solar irradiation), C_{475}, containing majorly Peak A and Peak C, has been proven as a biological refractory component (Murphy et al. 2018a). C_{290} and C_{320}, specifically tyrosine-like and tryptophan-like component, are sorted into the aquagenic categories. C_{410}, as described above, is categorized as a semi-labile component. Figure 3-14a shows the relative importance of the three categories. Aquagenic DOM increased along the estuary, while the semi-labile DOM disappear gradually, and the refractory DOM remained consistently. If components from all subsample PARAFAC are sorted into these three categories, the relative importance of them are shown in Figure 3-14b. The refractory DOM decreased along the river plume while the semi-labile DOM were unvaried, which makes no sense. These results, on the other hand, further shows the advantages of the one sample FlFFF-EEM-PARAFAC.

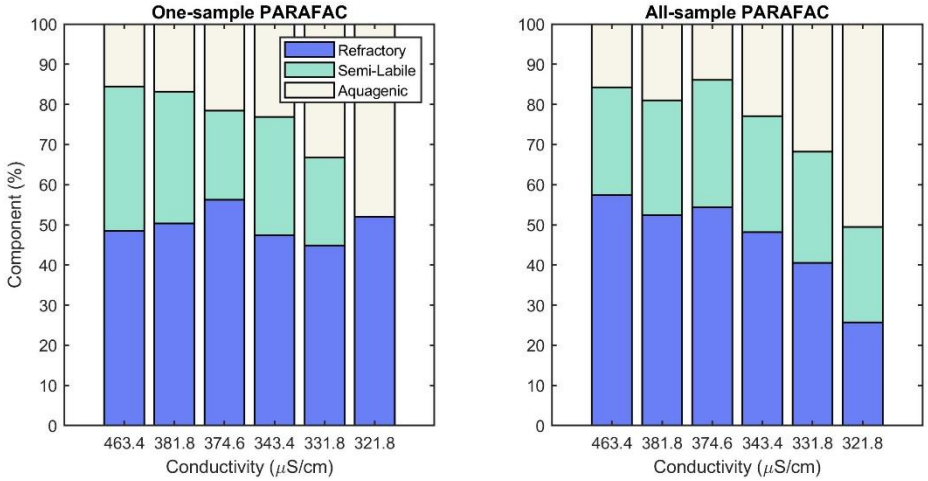

Figure 3-14. Comparisons in major DOM components between (left panel) individual sample PARAFAC approach and (right panel) all sample PARAFAC results. For discussion purposes, both C_{320} and C_{290} were combined to be "aquagenic" component. C_{475} was regarded as "refractory" component while C_{410} was recognized as a "semi-labile" component.

3. 5. 2 Size distributions and composition of DOM in the Fox River estuary

In comparison with a normal coastal estuary, our observations showed that molecular size of DOM in the negative estuary is consistent. The reasons for homogeneity of DOM size in a negative estuary, at least at Fox River – Green Bay estuary, are discussed below.

First, based on double layer theory, diminish of ionic strength ambient environment would thicken the electric double layer, and *vice versa* (Markus et al. 2016; Xu et al. 2018d). The decrease of DOM molecular size had been observed in marine estuaries by several field studies (Wells 2004; Lin et al. 2016). However, the gradient of ion density in the negative estuary is relatively insignificant in this small spatial scale compared with salinity boost in a normal coastal estuary; the specific conductivity here declines from 463.4 μS/cm to 321.8 μS/cm, which is nearly 100 times less than the specific conductivity rise in normal estuaries (e.g. from 500 μS/cm to 50,000 μS/cm). Second, negligible flocculation, aggregation was observed. Despite the UV_{254} and $Fluor_{350/450}$ (humic-like) of different size fractions decreased linearly, the relative each size fraction intensity remained constant (Figure 3-7). Meanwhile, the suspended particulate matter (SPM) was conservative mixed through the estuary, which implies that no addition or removal was found in this transportation (Table 1). Last but not the least, although the size distribution of DOM was partially impacted by pH (Romera-Castillo et al. 2014), the pH value, ranged from 8.7 to 9.1 in this negative estuary, was relatively constant, resulting in little impact on DOM molecular size (Table 1).

As for the DOM composition in the negative estuary, one sample FlFFF-EEM-PARAFAC method unveils the compositional structure change of fluorescent DOM. Although the fluorescent DOM is only a fraction of DOM, it potentially reflects the change of bulk DOM pool. Previous studies have proved that ion strength have in aquatic environments slight impact on the optical properties of CDOM (Del Castillo et al. 2000; Guo et al. 2007b). Thus, this DOM composition variations were mostly due to the various sources and sinks of natural DOM.

3. 5. 3 Evidence for humic self-assembly supramolecular hypothesis

Humic substances are regarded as a mixture of multiple molecules. Soil researchers proposed that the structure of natural humic substances is supramolecular associations of relatively small heterogenous molecules held by weak dispersive forces such as van der Waals force and hydrogen bonds (Thurman 2012; Verdugo 2012). Supporting evidence to this hypothesis from several different aspects has been reported. For example, Romera-Castillo et al. found that size spectra from high performance size exclusion chromatography (HPSEC) were overlapped significantly between different fluorescence components (Romera-Castillo et al. 2014). It indicates the inclusion of fluorophores in different supramolecular assemblies. However, the size fractionated EEM spectra was manipulated by fitting into a PARAFAC model originally from natural water samples collected in the Everglades National Park freshwater ecosystem. It would potentially miss some characteristic fluorescent signals from size fractionated EEM, and it would have additional artifacts during fitting processes since the PARAFAC model is not the optimal one for size fractionated EEM spectra. Another piece of evidence to support the hypothesis was using a coupling technique (HPSEC-PARAFAC2) to decompose the size spectra for natural marine DOM (Wünsch et al. 2017b). However, the online fluorescent DOM spectra were not provided. Herein, in this study, the one sample FlFFF-EEM-PARAFAC results contribute an additional and robust piece of evidence for the self-assembly supramolecular hypothesis. Figure 3-15 shows size spectra of two humic-like components (C_{475} and C_{410}) in all six stations. C_{475} and C_{410} were accumulated and highly overlapped at 1-10 kDa size interval despite of their different sources. They represent two groups of small heterogeneous humic-like moieties conjugating to form supramolecular assemblies around 1-10 kDa. Thus, consistent with the overlaps found in previous studies, our results provide a reliable support for the supramolecular assemblies hypothesis (Her et al. 2003; Boehme and Wells 2006; Romera-Castillo et al. 2014; Wünsch et al. 2017b).

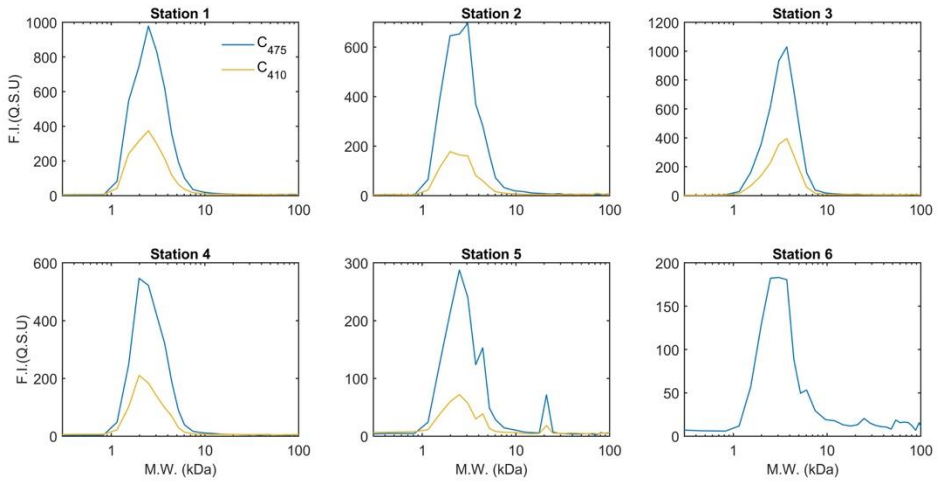

Figure 3-15. Size distributions of two humic-like fluorescence components (C_{475} and C_{410}) in the Fox River plume.

Chapter 4. Dynamic changes in size-fractionated DOM composition in the seasonally ice-covered Yukon River as characterized using EEM-PARAFAC, FT-IR and data fusion

This chapter is majorly from a manuscript *Dynamic changes in size-fractionated dissolved organic matter composition in a seasonally ice-covered Arctic river* submitted to a scientific journal *Limnology & Oceanography*. It was submitted on October 29th, 2020 and accepted on May 12th, 2021.

4. 1. Abstract

Arctic rivers are sensitive to climate and environmental change, but the biogeochemical response remains poorly understood. Monthly size-fractionated dissolved organic matter (DOM) samples from the lower Yukon River were characterized using UV-visible, fluorescence, and Fourier transform-infrared (FT-IR) spectroscopy techniques. The EEM-PARAFAC analysis revealed three major fluorescent DOM components, including two humic-like components (C_{480} and C_{400}) and one protein-like component (C_{310}), with their relative importance following the order of $C_{480} \geqslant C_{400} > C_{310}$ in the high-molecular-weight DOM (1 kDa-0.4 μm) and $C_{400} > C_{480} > C_{310}$ in the low-molecular-weight DOM pool (<1 kDa). Transformation in DOM and change in sources were manifested in major fluorescent components, and optical properties, including biological index (BIX), humification index (HIX), spectral slope ($S_{275\text{-}295}$) and specific UV absorbance at 254 nm ($SUVA_{254}$). These changes occurred within different DOM size-fractions and among ice-covered, spring freshet, and open seasons. Joint analysis of EEM and FT-IR spectra using a data fusion technique showed that humic-like DOM is mostly associated with C-H, C=C, and C-O bonds, while protein-like DOM is correlated more with C-N and N-H related structures. DOM aromaticity and the ratios of HIX to BIX and protein-like to humic-like components may be used as a compelling proxy to measure change in source waters and to infer permafrost dynamics. Our results provide insight into the seasonal variation in DOM composition for different size-fractions in the lower Yukon River, and a baseline dataset against which future changes can be understood in the context of arctic basin biogeochemical cycling.

4. 2. Introduction

Arctic terrestrial and aquatic ecosystems have experienced rapid climate and environmental change over the last several decades, resulting in evident impacts on hydrological cycles, permafrost dynamics, biogeochemical processes, and ecosystems as a whole. Of particular concern is the biogeochemical response of Arctic ecosystems to amplified warming, increased discharge, and accelerated permafrost degradation (Mcguire et al. 2009; Abbott et al. 2016; Turetsky et al. 2020). Indeed, changes in sources, export fluxes, and chemical composition of dissolved organic matter (DOM) and other chemical species in Arctic rivers have attracted considerable attention (Frey and Smith 2005; Finlay et al. 2006; Guo et al. 2007a; Raymond et al. 2007; Striegl et al. 2007; Spencer et al. 2009; O'Donnell et al. 2010; Pokrovsky et al. 2011; Cory et al. 2014; Vonk et al. 2015). While recent studies have advanced our understanding of the sources, composition, and fluxes of DOM in Arctic rivers, little is known about the seasonal variations in DOM dynamics and cycling pathways especially when there is an ice cover (Shogren et al. 2020; O'Donnell et al. 2021).

The Yukon River, one of the largest northern rivers, has a prolonged seasonal ice cover and drainage basins that contain large areas of continuous and discontinuous permafrost (Brabets et al. 2000). Although there has been relatively little anthropogenic disturbance within the drainage basin itself, the position of this river just below the Arctic Circle places it in a state of high vulnerability to amplified global warming with significant impacts have already been observed (Striegl et al. 2005; O'Donnell et al. 2021). Most of the previous studies on the dynamics of DOM in the Yukon River basin have relied on data collected during open seasons between spring freshet and early fall (Guo and Macdonald 2006; Zou et al. 2006; Spencer et al. 2008; Striegl et al. 2007), partly due to the extremely harsh sampling conditions during winters and partly due to cost. Although a few studies have accumulated a sparse set of winter data for organic carbon (Spencer et al. 2008; Stedmon et al. 2011; Guo et al. 2012; Aiken et al. 2014), detailed seasonal variations in DOM compositions/functionalities, optical properties, and PARAFAC-derived fluorescent DOM components in the Yukon River, especially during cold seasons under the ice, remain poorly observed.

In addition to bulk dissolved organic carbon (DOC), measurements of chromophoric DOM (CDOM), including UV-absorbance and fluorescence DOM (FDOM) using excitation-emission matrices (EEM), have been reported for the Yukon River (Belzile and Guo 2006; Guéguen et al. 2006; Spencer et al. 2008, 2009; Stedmon et al. 2011; Aiken et al. 2014). Over the last two decades, EEM coupled with parallel factor (PARAFAC) analysis have been widely used to characterize natural DOM in both freshwater and saltwater ecosystems (Murphy et al. 2013; Stedmon and Bro 2008; Yang et al. 2019). To the best of our knowledge, there is not a single study thus far reporting the characteristics and fluorescent components of the DOM in the lower Yukon River using these recently developed techniques. Likewise, reports characterizing CDOM and FDOM in size-fractionated DOM, such as the high-molecular-weight (HMW-) and low-molecular-weight (LMW-) DOM fractions are also scarce for the Yukon River.

In the upper Yukon River, elemental compositions (C and N) and fatty acid contents have been shown to differ significantly between the <1 kDa LMW-DOM and the >1 kDa HMW-DOM during ice open season from May to September (Zou et al. 2006). In the lower Yukon River, Belzile and Guo (2006) observed contrasting DOM absorbance and fluorescence signatures between the LMW- and colloidal DOM size-fractions, but these authors did not further investigate the differences using EEM-PARAFAC analysis. Specific differences in functional groups, fluorescent-DOM components, and other optical properties between the LMW- and HMW-DOM fractions, as well as the temporal variation from growing season to winter ice-covered periods, remain to be discovered.

Fluorescence spectroscopy, a highly sensitive and cost-effective method, has been widely used in the optical characterization of DOM in aquatic environments (Coble 2007; Murphy et al. 2014; Osburn et al. 2019; Zhou et al. 2013). Based on fluorescence EEM, aquatic DOM consists predominantly of two broad groups of chemical species: protein-like and humic-like. However, there are limitations in using only EEM data given that they do not provide detailed molecular level information. Similarly, Fourier transform-infrared (FT-IR) spectroscopy is another cost-effective analytical technique for DOM characterization (Jung et al. 2002; Yu 2005; Abdulla et al. 2010b; Yang et al. 2015). It provides molecular level information, such as functional groups, which can be highly complementary to fluorescence spectroscopy.

Data fusion, the process of joint analysis of multiple datasets to generate more useful information than what a single dataset can provide (Acar et al. 2015), has been applied to the characterization of DOM in aquatic environments. For example, Wünsch et al. (2018) used advanced coupled matrix tensor factorization (ACMTF) to integrate two datasets of fluorescence EEM, and Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) to improve the understanding of DOM biogeochemistry. Similarly, joint analysis of the FT-IR spectra and fluorescence EEM datasets should provide more comprehensive information about DOM dynamics.

In this study, monthly water samples from the lower Yukon River were collected to size-fractionate the bulk DOM into the LMW- (<1 kDa) and HMW-DOM (1 kDa-0.4 μm) size-fractions using cross-flow ultrafiltration. The size-fractionated DOM samples were characterized using UV-visible absorption, fluorescence, and FT-IR spectroscopy techniques, as well as EEM-PARAFAC modeling and data fusion. Our major goals were to (1) provide the first baseline dataset on EEM-PARAFAC derived fluorescent DOM components in the lower Yukon River, (2) investigate biogeochemical characteristics of size-fractionated DOM and determine seasonal variations in DOM abundance, size distributions, optical properties, and chemical composition, and (3) apply a state-of-the-art data fusion technique (ACMTF) to elucidate relationships between fluorescent components decomposed from EEM and organic functional groups revealed by FT-IR spectra. Our results show that selected DOM optical properties have a promising potential to provide proxies for change in hydrological conditions and permafrost dynamics consequent to warming, as reflected by the composition of DOM fractions carried by the Yukon River.

4. 3. Methods and materials

4. 3. 1 Study area

The Yukon River is fed by three major classes of tributaries (Figure 4-1), which are commonly categorized as blackwater, clearwater, and meltwater rivers (Brabets et al. 2000; Striegl et al. 2007). Blackwater rivers, such as the Porcupine River, exhibit highly colored river water with high DOC concentrations and relatively low dissolved inorganic carbon (DIC) loadings because they drain through permafrost-dominated wetlands and peatlands. Clearwater rivers are supplied by headwater lakes with

low DOC and suspended sediment concentrations. Glacier-fed rivers, such as the Tanana River, are dominated by meltwater with low DOC concentrations, but high DIC and sediment loadings. Our sampling station, which is located in the lower Yukon River (Figure 4-1), is supplied by a mix of water from all major tributaries of the Yukon River and therefore integrates the varied water chemistry, vegetation cover, DOM and other biogeochemical characteristics of the drainage basin (Brabets et al. 2000).

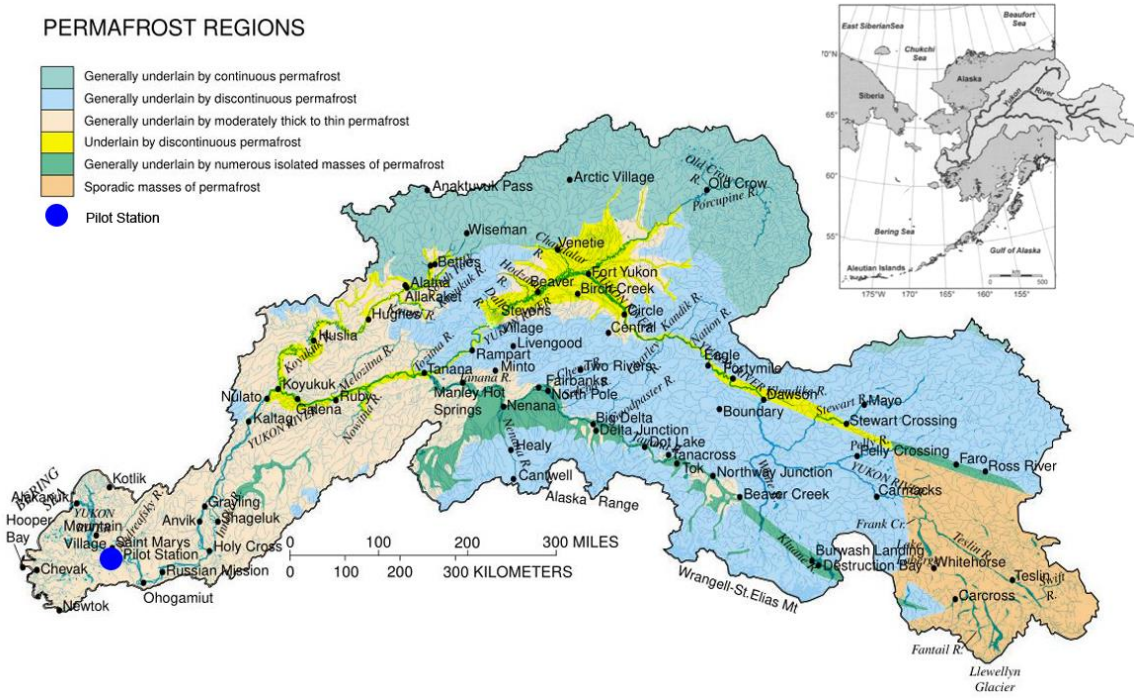

Figure 4-1. Map of the Yukon River basin showing the distribution of permafrost, classified from continuous to sporadic, and the sampling location at Pilot Station, Alaska, (blue dot) in the lower Yukon River (modified from Brabets et al. 2000; Guo and Macdonald 2006).

4. 3. 2 Sampling procedures

Nine water samples from the lower Yukon River were collected monthly or bi-monthly for DOM size-fractionation using ultrafiltration between September 2004 and September 2005 at Pilot Station, Alaska (65°56′10″ N, 162°53′0″ W, Figure 4-1). Other sampling details and measurements of water quality parameters have been described elsewhere (Guo et al. 2012) and the instantaneous discharge and selected hydrographic parameters during the sampling time period are given in Figure 4-2. Briefly, using a peristaltic pump with acid-cleaned polypropylene tubing, water samples were pumped from the river through a 0.45 μm polycarbonate filter cartridge (Osmonics) into acid-cleaned 20 L HDPE containers. A

manual ice auger was used to drill through the ice before sampling when the river was ice-covered between November and May. Nutrients (N, P, and Si), bulk DOC, and other hydrographic data for these water samples have been presented in Guo et al. (2012).

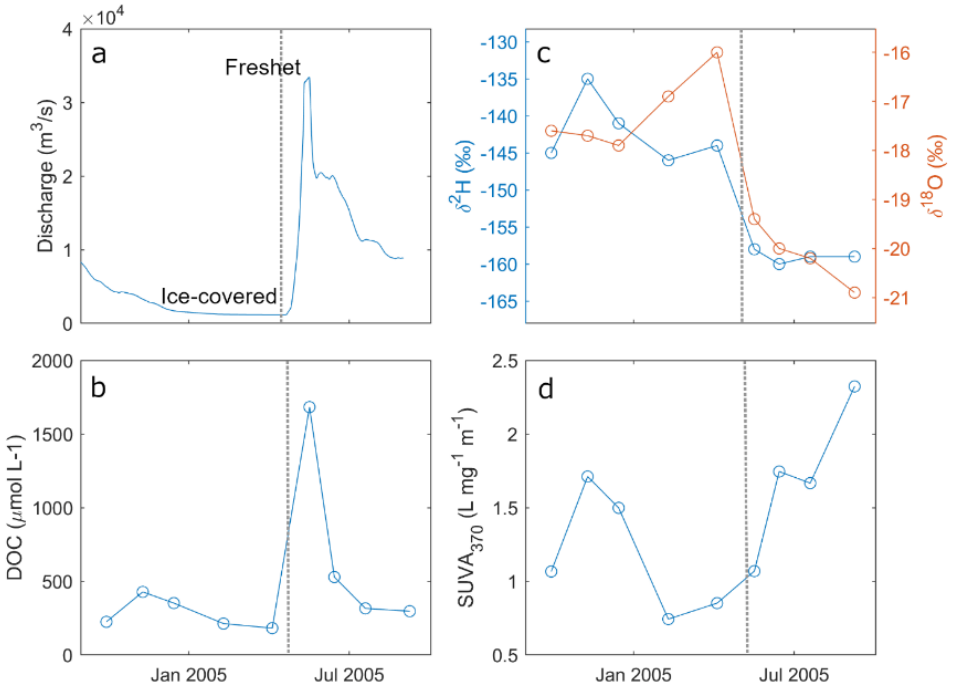

Figure 4-2. Temporal variations in (a) daily discharge, (b) DOC concentration, (c) water isotopes including δ^2H and $\delta^{18}O$, and (d) specific UV absorbance at 370 nm ($SUVA_{370}$) in the lower Yukon River during the 2004 and 2005 sampling period (discharge data are from USGS; Others are calculated and/or replotted from data in Guo et al. 2012). The gray dashed line divides the periods into the ice-covered and ice-free season.

The <0.45 μm filtrate was further ultrafiltered using a spiral-wound 1 kDa cartridge (regenerated cellulose, Amicon, S10Y1) to size-fractionate the bulk DOM into HMW- (1kDa-0.45 μm) and LMW- (<1 kDa) DOM size-fractions. A total of 40 L of pre-filtered water sample was used for ultrafiltration to isolate sufficient amounts of HMW-DOM for other chemical and isotopic characterization (e.g. Guo and Macdonald 2006; O'Connor et al. 2020). The first 20 L was ultrafiltered to quantify the abundance of HMW-DOC using the ultrafiltration permeation model and time series DOC concentrations of permeate samples (Belzile and Guo 2006; Guo and Santschi 2007). The final retentate volume was reduced to 500 mL, corresponding to a concentration factor of 80. The cleaning and calibration of the ultrafiltration system are described in previous studies (Guo et al., 2000; Guo and Santschi, 2007). Right after ultrafiltration, the retentate or HMW-DOM samples were further diafiltrated to remove residual salts and LMW-DOM. The HMW-DOM and aliquots of the LMW-DOM and bulk DOM samples, including 9

HMW-, 9 LMW-, and 9 bulk DOM samples (a total of 27), were freeze-dried immediately after sample processing each month for further optical/fluorescence and chemical/molecular characterization.

4. 3. 3 Measurements of DOC, UV-vis absorption, and fluorescence EEM

Concentrations of DOC were measured on a TOC-Analyzer (Shimadzu TOC-L) using the high-temperature combustion method (Cai et al. 2008). Since one of our major objectives was to integrate both EEM and FT-IR datasets using advanced coupled matrix tensor factorization (ACMTF) modeling, freeze-dried DOM samples were consistently used for both optical characterization and FT-IR analyses. For the measurements of UV-vis absorbance and fluorescence EEM spectra, predetermined amounts of the freeze-dried DOM samples, including HMW-, LMW-, and bulk-DOM, were re-dissolved in boric acid-borate buffer solution with a pH=8 similar to the pH range (7.2-8.4) of Yukon River waters (Guo et al. 2012) and shaken for 2 h (on a VWR Mini Shaker) followed by ultrasonication for 2 h (using a VWR Symphony Ultrasonic cleaner). Absorption spectra were obtained using a UV-visible spectrophotometer (Agilent 8453) with a 1-cm quartz cuvette. Specific UV Absorbance at 254 nm ($SUVA_{254}$), which has been used as an indicator of DOM aromaticity (Weishaar et al. 2003), was calculated as: $SUVA_{254} = A_{254}/[DOC]$, where A_{254} is the UV-absorbance at 254 nm (m^{-1}), and [DOC] is the DOC concentration (mg-C L^{-1}). The spectral slope between 275 nm and 295 nm ($S_{275\text{-}295}$), which is a proxy for apparent DOM molecular weight, was calculated using a non-linear exponential model based on $a_\lambda = a_0 e^{S(\lambda-\lambda_0)}$ (Helms et al. 2008).

Fluorescence EEM were measured using a spectrofluorometer (Fluoromax-4, Horiba). Samples in a 1 cm path-length quartz cuvette were scanned at excitation wavelengths ranging from 250 nm to 480 nm with 5 nm intervals, and emission wavelengths ranging from 240 nm to 600 nm with 2 nm intervals. Humification index (HIX) is the ratio of fluorescence signals over the emission range of 435-480 nm to those over the range of 300-345 nm with excitation at 254 nm (Zsolnay et al. 1999). Biological index (BIX), an indicator of autochthonous fluorescent DOM (Huguet et al. 2009; Zhou et al. 2016b), is the ratio of fluorescence intensity over the emission range of 380 nm to 430 nm within the excitation wavelength of 310 nm

4. 3. 4 Fourier transform infrared spectroscopy

The freeze-dried DOM samples were further characterized using FT-IR spectroscopy (IRTracer-100, Shimadzu). To increase the reproducibility, FT-IR absorbance spectra were obtained by collecting 30 scans at each wavenumber to reduce errors and subtracted from the air background scan. These FT-IR spectra ranging from 4000 to 400 cm^{-1} had a resolution of 1 cm^{-1}. The spectra in the 4000-1800 cm^{-1} region contain one broad peak from O-H stretching, while the 900-400 cm^{-1} region contains mostly noise (Figure 4-3). Therefore, only the spectral region between 1800-900 cm^{-1} was selected for data fusion modeling. This region covers major bands from natural organic matter, such as amide, carboxylic, ester and carbohydrate functional groups (Abdulla et al. 2010b, 2013) although O-H and N-H stretching at the >3000 cm^{-1} has been observed in treated wastewater samples (Yang et al. 2015).

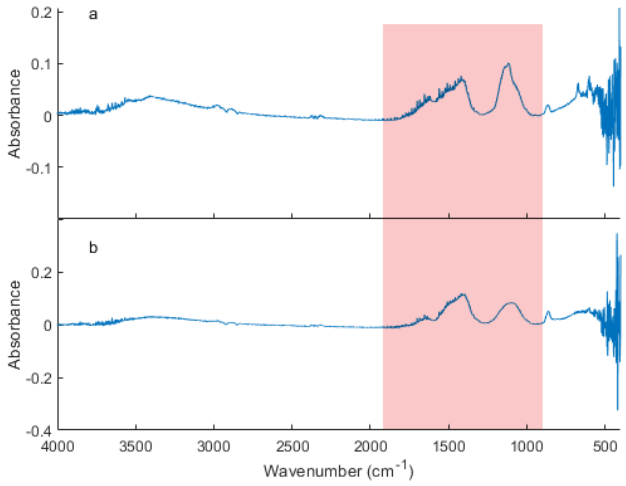

Figure 4-3. Examples of FT-IR spectra of (a) HMW- and (b) LMW- DOM in the Yukon River. The red band shows the selected region for data fusion analysis.

4. 3. 5 Parallel factor (PARAFAC) analysis

PARAFAC analysis was conducted on all EEM data using the drEEM toolbox (v.0.5.0) in MathWorks® Matlab (R2020a) software platform (Murphy et al., 2013). Rayleigh and Raman scattering peaks were first eliminated for conformity to the linear assumptions required for PARFAC models. Fluorescence intensities were corrected for the inner-filter effect using the UV-vis absorbance spectra and calibrated into quinine sulfate equivalent (QSE). A total of 27 samples, including 9 HMW-DOM, 9 LMW-DOM and 9 bulk DOM samples, were assembled and normalized to create validated models. The models were

constrained to non-negative values. Split-half validations were performed based on a randomized split of the dataset and the comparisons of Tucker Congruence Coefficient (TCC) between split models.

4. 3. 6 Advanced coupled matrix tensor factorization (ACMTF) model

To explore the potential connections between EEM and FT-IR spectra, ACMTF analysis was conducted to analyze these spectroscopic data jointly. Before ACMTF analysis, EEM spectra were corrected, calibrated, normalized, and assembled as described above. Data of FT-IR were scaled to the maximum absorption of each FT-IR spectra. The dimension of the assembled EEM tensor and FT-IR spectra matrix are 27×181×47 (sample×emission×excitation) and 27×935 (sample×wavenumber), respectively.

ACMTF analysis was performed on both EEM and FT-IR datasets using the CMTF toolbox (Acar et al. 2011, 2014), Tensor toolbox (Bader and Kolda, 2006, 2007), and SNOPT toolbox (Gill et al. 2005) in Matlab. ACMTF analysis is based on three assumptions, similar to PARAFAC models, including variability, additivity, and linearity (Murphy et al., 2013). For the variability assumption, all components have distinct excitation/emission/FT-IR spectral loadings. Second, fluorescence spectroscopy and FT-IR spectroscopy are both based on the Beer-Lambert Law, which satisfies the additivity assumption. Lastly, fluorescence EEM have tri-linearity in terms of component concentrations, excitation spectra, and emission spectra. The FT-IR absorption is linearly related to FT-IR spectral loadings and its invariant spectral wavenumber.

ACMTF decomposed the EEM and FT-IR spectra into a set of trilinear fluorescence components and bilinear molecular formula components. Briefly, the ACMTF model can be expressed as:

$$x_{ijk} = \sum_{f=1}^{F} \mu_f a_{if} b_{jf} c_{kf} + \varepsilon_{ijk} \quad \ldots\ldots \quad (1)$$

$$y_{im} = \sum_{f=1}^{F} \sigma_f a_{if} v_{mf} + \eta_{im} \quad \ldots\ldots \quad (2)$$

where i=1…I; j = 1…J; k=1…K; f=1…F; m=1…M. The x_{ijk} is the fluorescence signal intensity of sample i measured at $\lambda_{Em} = j$ and $\lambda_{Ex} = k$. f is the underlying component (factor in statistics) with a range from 1 to F (the total number of factors). μ_f and σ_f are penalty terms, which are the differences between ACMTF and PARAFAC models. These terms determine whether one specific component, f, is shared between the

EEM and FT-IR spectra (Wünsch et al., 2018). For example, in a case where the μ_f is 0.9, but the σ_f is 0.03, component *f* might be unshared between these datasets. a_{if}, b_{jf}, and c_{kf} are model-derived parameters representing the fluorescence maximum (F_{max}, which could be used to approximate fluorescence concentration) of f^{th} component in sample *i*, excitation loadings of the f^{th} component in $\lambda_{Em} = j$, and emission loadings of the f^{th} component in $\lambda_{Ex} = k$, respectively. The final terms, ε_{ijk} and η_{im}, are residuals representing noise and variation uninvolved in the model. y_{im} is the FT-IR absorption of sample *i* measured at wavenumber *m*. The v_{mf} is the FT-IR spectral loading of the f^{th} component at wavenumber *m*. Overall, ACMTF models two datasets using equations (1) and (2), fitting the equation by minimizing the sum of squares of the residuals (ε_{ijk} and η_{im}). All models were constrained to non-negative values. Model comparisons and validations are quantified using Tucker Congruence Coefficient (TCC) similar to those in Murphy et al. (2018).

Statistical data analyses, including t-test and F-test for comparisons of variances and correlations, were conducted with a Matlab statistics toolbox (MathWorks, R2020a). A significant level of 0.05 was selected for F-test and 0.01 for t-test.

4. 4. Results

4. 4. 1 Environmental parameters in the lower Yukon River

In the lower Yukon River, the discharge during the spring freshet of 2005 (up to 33428 m^3/s, Figure 4-2) at the downstream USGS hydrological station at Pilot Station is almost double the maximum discharge in spring 2004 (at 18406 m^3/s), showing high variability in river discharge between hydrological years. Water isotopic composition ($\delta^{18}O$ and $\delta^{2}H$) exhibited similar patterns, with higher values during the ice-covered season due to a larger relative contribution of isotopically heavier groundwater input but lower values during snowmelt and river open seasons (Figure 4-2c). Intense precipitation during 2005 (including snow and rainwater) resulted in lower stable isotope values compared with summer 2004. In 2004, [DOC] increased slightly between September and November (Figure 4-2b), likely due to the exclusion of DOC from ice during freeze-up, followed by a steady decrease from 428 µM to 182 µM during the ice-covered season, reaching the lowest concentration in April, 2005 (Guo et al. 2012). During

the spring freshet (May 2005), [DOC] increased abruptly to as high as 1682 μM, dropped rapidly after freshet, and then decreased monotonically between June and September (Figure 4-2b). In contrast to [DOC], $SUVA_{370}$ showed elevated values during both ice-covered (November 2004) and ice-free (September 2005) periods with the lowest values observed between February and May 2005 (Figure 4-2d).

4. 4. 2 Variations in $SUVA_{254}$ and $S_{275\text{-}295}$ between HMW- and LMW-DOM

For the >1 kDa HMW-DOM pool, the $SUVA_{254}$ value, an indicator of DOM aromaticity (Weishaar et al. 2003), varied from 3.30 ± 0.60 L mg-C^{-1} m^{-1} in the ice-covered season (with a water temperature at or below 0 °C) to 3.95 ± 0.36 L mg-C^{-1} m^{-1} in the ice-free season (post-May 2005, Figure 4-4a). The abrupt increase in $SUVA_{254}$ values during the freshet period is concurrent with the peak [DOC] during spring freshet with heavy loads of terrestrial DOM. After the spring freshet, $SUVA_{254}$ declined from June to September, similar to the trend of [DOC], but to a relatively smaller extent. On average, the $SUVA_{254}$ value of the HMW-DOM is higher during the summer growing season than the ice-covered season (p <0.05), indicating overall higher aromaticity in the HMW-DOM pool during summer. In the ice-covered season, similar to changes in [DOC] (Figure 4-4b), the initial river-ice formation and thus the exclusion of DOM from ice (Belzile et al. 2002) also resulted in a slight increase of the HMW-DOM's $SUVA_{254}$ value in November, followed by a general decrease between November and April (Figure 4-4a).

For the LMW-DOM pool (Figure 4-4a), $SUVA_{254}$ values are relatively less variable compared to those of the HMW-DOM (p<0.025). However, they showed a general decrease during the ice-covered season from 2.57 L mg-C^{-1} m^{-1} in September 2004 to 1.86 ± 0.03 L mg-C^{-1} m^{-1} during December-April, followed by an increase during the spring freshet with a $SUVA_{254}$ value of 2.12 L mg-C^{-1} m^{-1} in May 2005. After snowmelt (May 2005), $SUVA_{254}$ values changed little, ranging from 2.12 to 2.19 L mg-C^{-1} m^{-1} with an average of 2.15 ± 0.05 L mg-C^{-1} m^{-1} in the lower Yukon River.

Overall, $SUVA_{254}$ values of the LMW-DOM samples are systematically lower than those of the HMW-DOM (2.06 ± 0.23 vs. 3.31 ± 0.49 L mg-C^{-1} m^{-1}, p < 0.01), indicating that the HMW-DOM pool is of higher aromaticity than the LMW-DOM pool. $SUVA_{254}$ values observed here are consistent with previous

measurements for the bulk DOM pool at Pilot Station (Striegl et al. 2007; Spencer et al. 2008; Stedmon et al. 2011; Aiken et al. 2014). Compared to $SUVA_{254}$, the values of $SUVA_{370}$ showed greater temporal variability in the HMW-DOM pool (Figure 4-5), although both the HMW- and LMW-DOM pools had similar trends in their variability and were highly correlated with one another (r^2= 0.979, p<0.01).

Spectral slope ($S_{275\text{-}295}$), a proxy for DOM apparent molecular weight (Helms et al. 2008), showed a general temporal change opposite to the pattern of $SUVA_{254}$ (Figure 4-4b). For both HMW- and LMW-DOM, the $S_{275\text{-}295}$ increased gradually when $SUVA_{254}$ declined during the ice-covered season, suggesting a general decrease in the apparent DOM molecular weight with decreasing aromaticity. In particular, the $SUVA_{254}$ of the HMW-DOM sharply increased during spring freshet, accompanied by a simultaneous decline in the $S_{275\text{-}295}$ values of the HMW-DOM (Figure 4-4b). After freshet, the decrease in $SUVA_{254}$ of the HMW-DOM pool was accompanied by an increase in its $S_{275\text{-}295}$ values, indicating a decrease in the DOM molecular weight. These relationships suggest that terrestrial HMW-DOM during spring freshet in the lower Yukon River was of both higher aromaticity and higher molecular weight, while DOM under the ice was of lower $SUVA_{254}$ with lower molecular weight or higher spectral slope values.

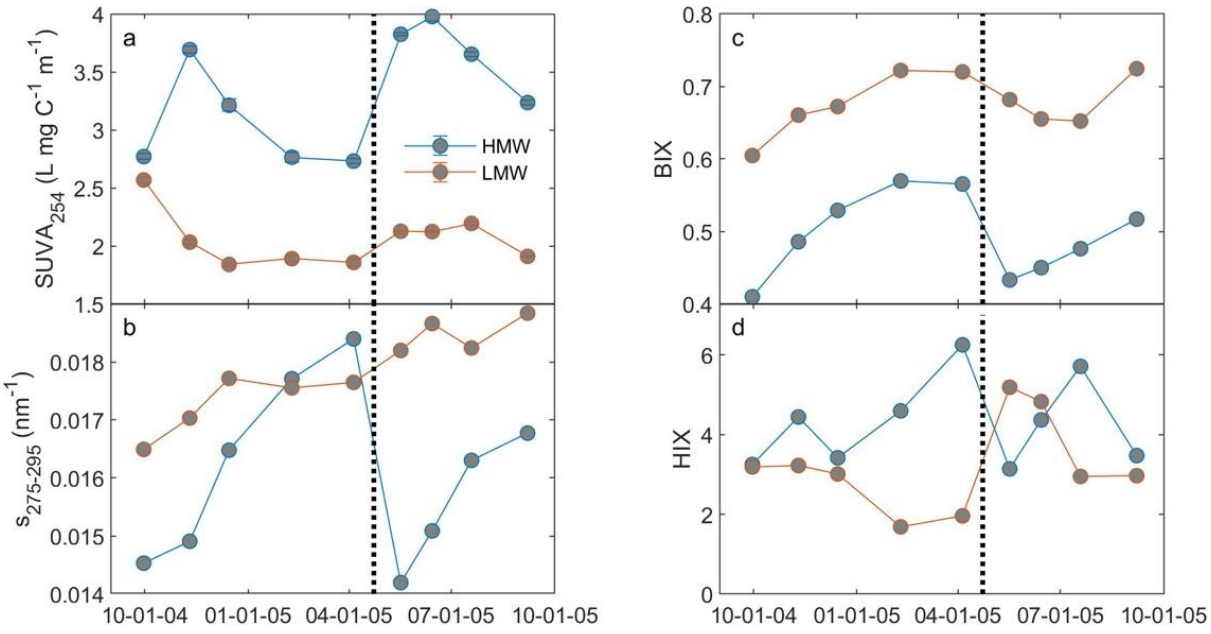

Figure 4-4. Temporal variations in (a) $SUVA_{254}$, (b) $S_{275\text{-}295}$, (c) BIX, and (d) HIX between HMW-DOM and LMW-DOM pools in the lower Yukon River with dashed lines denoting the beginning of spring freshet.

Compared with the HMW-DOM, values of $S_{275\text{-}295}$ in the LMW-DOM pool changed little (p<0.05) but showed a general increase over the entire sampling period from September 2004 to September 2005

(Figure 4-4b). $SUVA_{254}$ in both HMW- and LMW-DOM decreased to the lowest values during the ice-covered season, whereas the corresponding $S_{275\text{-}295}$ values increased, suggesting that the average DOM molecular weight decreased in both HMW- and LMW-DOM pools during winter under the ice. After the spring freshet, the $S_{275\text{-}295}$ values in the LMW-DOM pool remained relatively high but stable.

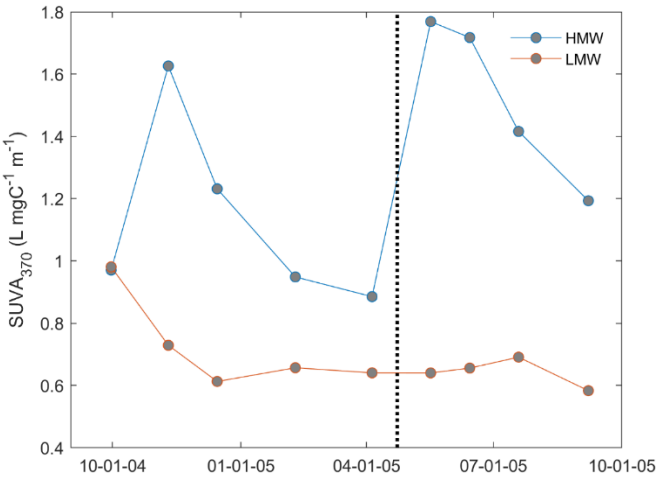

Figure 4-5. Temporal variations in $SUVA_{370}$ of the HMW-DOM and LMW-DOM in the lower Yukon River.

4. 4. 3 Seasonal variations in fluorescence indices between HMW- and LMW-DOM

Values of BIX in the HMW-DOM pool increased consistently during the ice-covered season from September 2004 up until May 2005 (Figure 4-4c), indicating an increase in autochthonous DOM sources under ice. However, during the ice melt season, the BIX dropped abruptly from 0.57 in April under the ice to 0.43 in May (spring freshet). Values of BIX in the LMW-DOM pool also had a variation trend similar to the HMW-DOM, increasing from September (0.60) to early April (0.72) during the ice-covered season, declining slightly during the spring freshet with a slight decrease from May (0.68) to July (0.65), and an elevated value in September 2005 (Figure 4-4c). During the spring freshet, values of BIX in the HMW-DOM pool dropped to their lowest value (0.42), while the LMW-DOM pool always exhibited a higher BIX value than the HMW-DOM pool ($p<0.01$).

Similar to BIX, the values of HIX in the HMW-DOM pool also increased generally from September 2004 to early May 2005 during the ice-covered season (Figure 4-4d). The highest HIX value of the HMW-DOM was found in April under the ice, followed in mid-May by the lowest HIX during the spring freshet. For the LMW-DOM pool, the HIX values during the winter season were, on average, lower than those

during the open season (p<0.05, Figure 4-4d). HIX reached its maximum during spring freshet and then decreased monotonically from May to September 2005. Overall, HIX values in the HMW-DOM pool were slightly higher than those in the LMW-DOM (4.30 ± 1.11 vs. 3.22 ± 1.15, $p < 0.10$), consistent with the difference in $SUVA_{254}$ between HMW- and LMW-DOM pools in the lower Yukon River.

4. 4. 4 Major functionalities of DOM derived from FT-IR spectra

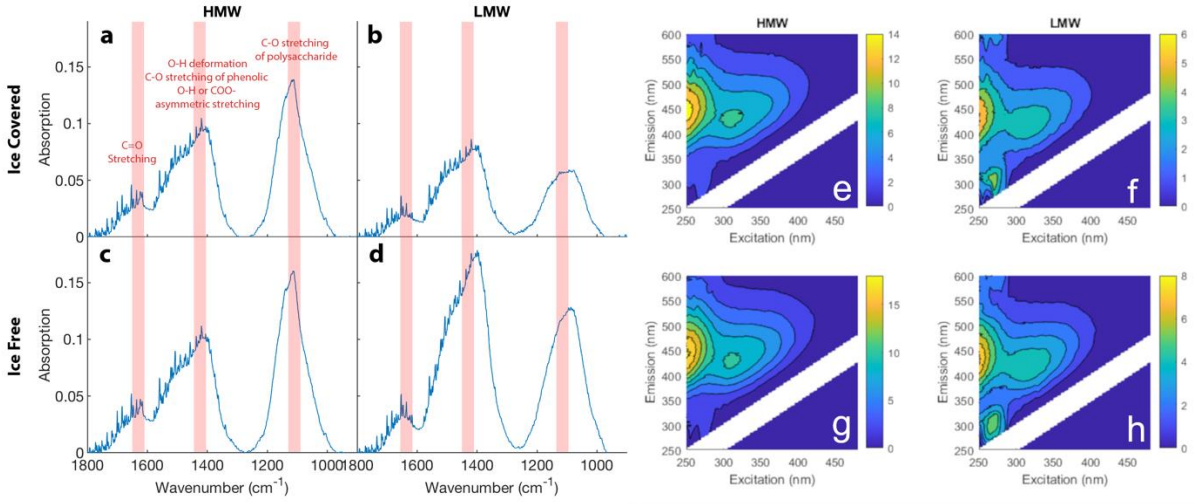

Figure 4-6. Examples of (a-d) FT-IR spectra and (e-f) EEM contours for HMW- and LMW-DOM samples during the ice-covered season (top panels) and ice-free season (bottom panels).

FT-IR spectroscopy was applied to identify major functional groups in size-fractionated DOM and their seasonal variations in the lower Yukon River. Samples from February and May were used to represent the ice-covered and ice-free seasons, respectively, and examples of FT-IR spectra of both HMW- and LMW-DOM samples, along with examples of their EEM spectra, are shown in Figure 4-6. There are three main peaks with locations at 1620-1640 cm^{-1}, 1400-1410 cm^{-1}, and 1040-1120 cm^{-1}, respectively. These peaks correspond to functional groups of C=O stretching of amide, quinone, or ketones at 1620-1640 cm^{-1}, O-H deformation, C-O stretching of phenolic OH^- or COO^- asymmetric stretching at 1400-1410 cm^{-1}, and C-O stretching of polysaccharide or polysaccharide-like substances at 1040-1120 cm^{-1} (Xu et al. 2018c). In general, major functional groups are broadly similar, especially in the HMW-DOM between the winter and spring-freshet samples. For the LMW-DOM, the absorbance for each major functional group was lower in the winter sample compared to the spring-freshet sample (Figure 4-6). In addition, specific peak

locations in wavenumbers between 1040-1120 cm^{-1} were slightly different between the HMW-DOM and the LMW-DOM pools, with a lower wavenumber in the LMW-DOM samples (Figure 4-6).

4. 4. 5 Major fluorescent DOM components derived from EEM-PARAFAC analysis

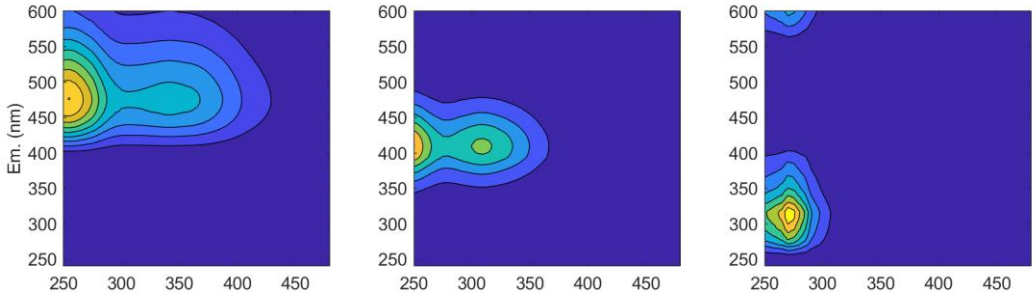

Figure 4-7. Fluorescence EEM contours of PARAFAC-derived major fluorescent DOM components including, from left to right, components C_{480}, C_{400}, and C_{310} in the lower Yukon River (see Table 4-1 for their specific Ex/Em maxima).

Table 4-1. Characteristics of the PARAFAC-derived major fluorescent DOM components.

Component	C_{480}	C_{400}	C_{310}
Ex/Em (nm)	250/480	260,310/400	275/310
Fluorophore type	Humic-like	Humic-like	Protein-like
Major peaks	A and C	A and M	T and B

Examples of EEM spectra of size-fractionated HMW- and LMW-DOM samples are shown in Figure 4-6. Based on all EEM data, PARAFAC analysis revealed three major fluorescent DOM components (Figure 4-7), including two humic-like components (C_{480} and C_{400}) and one protein-like component (C_{310}), whose characteristics are given in Table 4-1. Component C_{480} (Ex/Em = 250/480 nm) is one of the most widely distributed humic-like fluorescent components in aquatic environments (Coble 2007; Lin and Guo 2020; Stedmon et al. 2003; Wünsch et al. 2019; Yamashita and Jaffe 2008; Zhou et al. 2013). Its EEM contour covers Peak A and Peak C areas defined in Coble (2007), which are UVC and UVA humic fluorophores. Component C_{400} (Ex/Em = 260/400 nm) is the second most abundant fluorescent DOM component, with characteristics of Peak M and part of Peak A (Wünsch et al. 2019a). This C_{400} component was previously recognized as a marine-derived humic-like fluorophore (Coble et al. 1998; Coble 2007), but has recently been found widely distributed in freshwater environments (DeVilbiss et al., 2016; Zhou et al., 2016).

Component C_{310} (Ex/Em = 275/310 nm) is a protein-like component covering tryptophan-like Peak T (Ex/Em = 275/340 nm) and tyrosine-like Peak B (Ex/Em = 275/305 nm). Component C_{310} was found to be strongly related to microbial activities and autochthonous DOM production (Coble 2007; Zhao et al. 2017).

Online comparisons to the OpenFluor dataset (Murphy et al. 2014) show that these three fluorescent DOM components are similar to PARAFAC-derived components identified in previous studies ($TCC_{ex,\ em}$ > 0.95). For example, component C_{480} has 64 pairs of match-ups, including components found in fresh waters to salt waters in coastal seas (e.g. Osburn et al. 2016; Gonçalves-Araujo et al. 2016). Within the 64 model pairs, $TCC_{ex,em}$ for 5 pairs are larger than 0.99, 28 pairs between 0.97 and 0.99, and 31 pairs between 0.95 and 0.97. The five highly similar pairs include samples from drinking water (Shutova et al. 2014), wetlands (Yamashita et al. 2010), and Japan Sea (Tanaka et al. 2014). Match-ups for components C_{400} and C_{310} are also found in 58 and 27 previous studies, respectively, from both natural waters (Yamashita et al. 2010; Tanaka et al. 2014; Dainard et al. 2015; Kulkarni et al. 2018) and wastewaters/drinking waters (Shutova et al. 2014). However, both components did not seem to exhibit a high matchup (i.e. $TCC_{ex,em}$ >0.99) with any previous reported components in the OpenFluor database. Apparently, the higher the threshold of $TCC_{ex,em}$, the stronger the support for the identification of fluorescent components between different models (Wünsch et al. 2019a). In fact, samples in the OpenFluor database are from different aquatic environments, and data were obtained under different instrumental specifications (such as, sensitivity, signal/noise ratio, etc.) and handling procedures. Thus, the probability is likely low for $TCC_{ex,em}$ to reach the highest values during model comparisons between different studies; nevertheless, a stringent $TCC_{ex,em}$ criterion is highly recommended for PAFARAC model validations.

To compare these DOM components on an intensive-property basis, fluorescence intensities (quinine sulfate equivalent, QSE) of each fluorescent component were normalized to DOC concentration (Figure 4-8). The DOC-normalized C_{480} (C_{480}/DOC) in the HMW-DOM (Figure 4-8a) increased in general from September 2004 to April 2005, decreased during the spring freshet, and then increased again from June to September 2005. In contrast, the C_{480}/DOC in the LMW-DOM (Figure 4-8a) decreased during the winter

season between September 2004 and April/May 2005, but increased during the spring freshet, and decreased again in September 2005. The C_{400}/DOC (Figure 4-8b) was less variable in the HMW-DOM over the entire sampling period but was elevated during spring freshet and then decreased from May to September 2005 in the LMW-DOM pool (p<0.01). The seasonal variation in the C_{310}/DOC (Figure 4-8c) differs distinctly between the HMW- and LMW-DOM pools. For example, this DOC-normalized protein-like component, C_{310}/DOC, in the HMW-DOM pool generally decreased over the whole sampling time period except for a slightly higher value during ice opening and an elevated value in September 2005, while values in the LMW-DOM increased consistently below the ice, attaining its lowest values during spring freshet (May and June), and increasing thereafter in July and September.

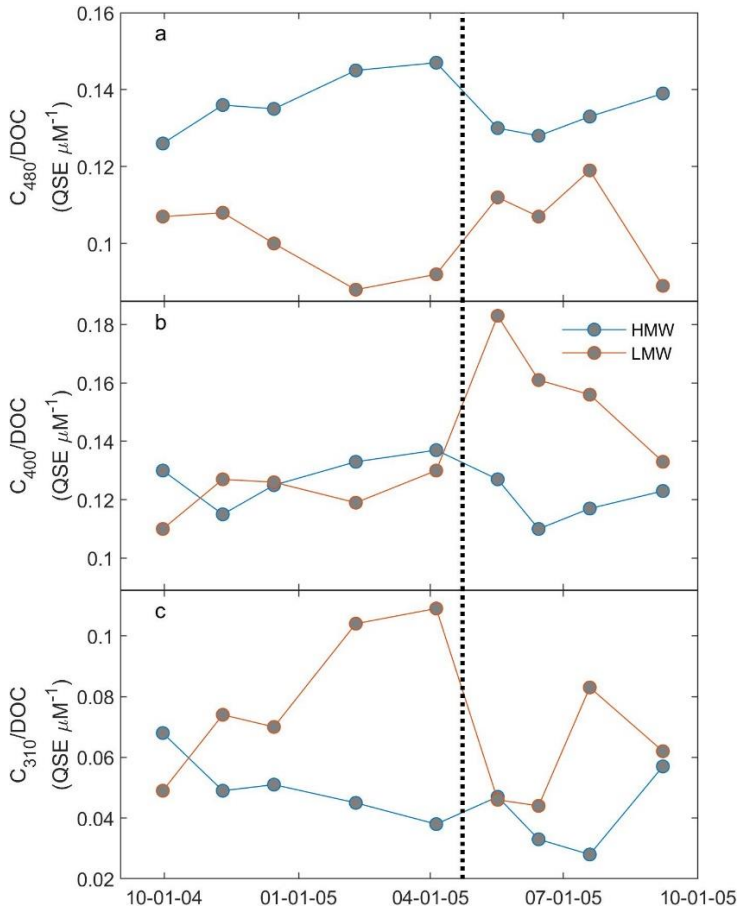

Figure 4-8. Temporal variations in DOC-normalized major fluorescent components, including C_{480}/DOC (a), C_{400}/DOC (b), and C_{310}/DOC (c) in both HMW- and LMW-DOM pools (dashed lines denote the beginning of spring freshet).

To compare the relative contribution of each component to the total fluorescence, percentages were calculated as $C\% = C_i / \Sigma C_i \times 100$, where C_i is the F_{max} of the i^{th} component derived from EEM-PARFAC

analysis. Both C_{480}% and C_{400}% are approximately equally dominant DOM components, comprising 44.4±2.7 % and 40.6±1.5 % in the HMW-DOM pool and 33.0±3.8 % and 44.3±5.4 % in the LMW-DOM pool, respectively, while C_{310} contributes a smaller component, making up 15.0±3.4 % of the fluorescence in the HMW-DOM and 22.8±7.0 % in the LMW-DOM (Figure 4-9). The percentage of each component varied less in the HMW-DOM pool than in the more dynamic LMW-DOM pool. For example, in the LMW-DOM pool, C_{480}% decreased consistently, whereas C_{310}% increased from September to April, while C_{400}% changed little. These compositional changes resulted in a large increase in the C_{310}% in the LMW-DOM during winter (Figure 4-9). During the ice-free season, the C_{480}% in the LMW-DOM pool changed little, and the C_{400}% became higher or predominated, while the C_{310}% was lower than it was in the winter season. The main implication of these data in terms of monitoring is that dynamic changes occur between ice-covered and ice-free seasons in the DOM components for both HMW- and LMW-DOM pools, and these two pools respond differently.

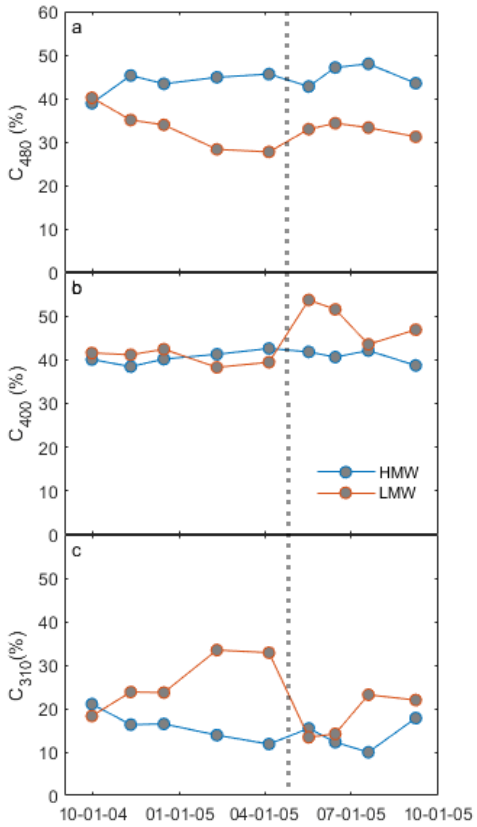

Figure 4-9. Temporal variations in the percentages (%) of C_{480}, C_{400}, and C_{310} in the HMW- and LMW- DOM pools in the lower Yukon River.

4. 4. 6 Integrating EEM and FT-IR datasets using data fusion

The ACMTF modeling on both EEM and FT-IR datasets identified three model-derived components, C_{480}, C_{400} and C_{310}, named after their characteristic emission wavelengths (Table 4-2). Their scores, excitation loadings, and emission loadings are highly congruent with those generated from the EEM-PARAFAC model (Figure 4-10). $TCC_{ex,em}$ values for C_{480}, C_{400} and C_{310} are 0.9993, 0.9998, and 0.9999, respectively, indicating that both the ACMTF model and the EEM-PARAFAC model highly resemble each other (Figure 4-10). Online comparisons between our model-derived components and those previously identified in the OpenFluor database also show identical results ($TCC_{ex, em}$ >0.95), further validating the ACMTF and PARAFAC models.

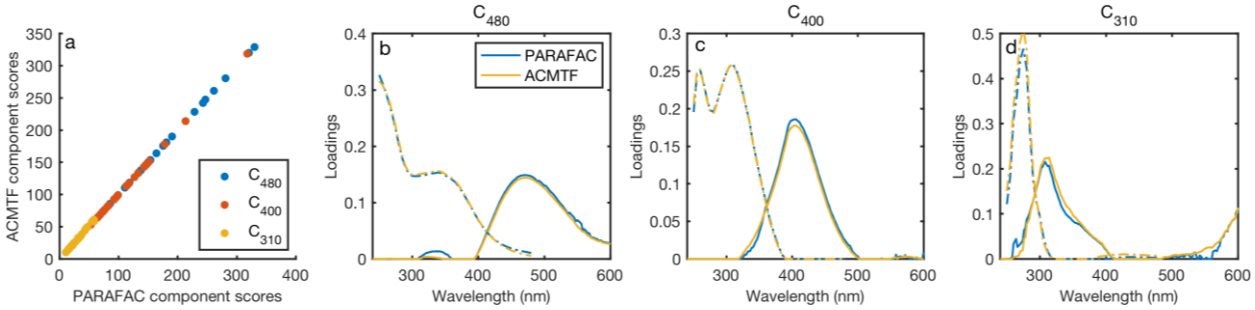

Figure 4-10. Comparisons in (a) component scores and (b-d) excitation loadings (dashed lines) and emission loadings (solid lines) for C_{480}, C_{400}, and C_{310} between the ACMTF and PARAFAC models.

Table 4-2. Characteristics and properties of three ACMTF model-derived components.

Component	C_{480}	C_{400}	C_{310}
Ex/Em (nm)	250/480	260,310/400	275/310
μ_f/σ_f*	20.35±1.03/9.36±0.65	21.28±1.08/9.15±0.67	21.63±0.96/9.20±0.68
Peaks in EEM	A and C	A and M	T and B
Peaks in FT-IR (cm^{-1})	1060, 1230, 1400, 1470	1085, 1190, 1400, 1650	1114, 1417, 1621
Major Bonds	C-H bending, C-O stretching	C-H bending, C-O stretching, C=C stretching	C-N stretching, C-H bending, N-H bending

* μ_f and σ_f are factors representing the sharedness of components for EEM and FTIR spectra, respectively.

Similar to those derived from EEM-PARAFAC analysis, the ACMTF model-derived C_{480} and C_{400} are humic-like components, while the C_{310} is a protein-like component (Table 4-2). In addition, both C_{480} and C_{400} fluorescent components are associated with the functional groups that are composed of C-O bonds, C-H bonds, and C=C double bond (Figure 4-11). The protein-like C_{310} component matches two unique nitrogen-related bonds, including C-N bending and N-H bending. These results imply that humic-like DOM components represent mostly carbon-related bonds, such as C-O, C-H, and C=C, while the protein-like fluorescent component comprises more nitrogen-related molecular structures, such as C-N and N-H (Figure 4-11). Similar results have been reported by Maie et al., (2007), in which a protein-like fluorophore (Peak T) was positively correlated with nitrogen concentration, and Wünsch et al. (2018), who found that C_{310} (similar component with the emission maximum at 300 nm) contained the highest N/C ratio among all six DOM components.

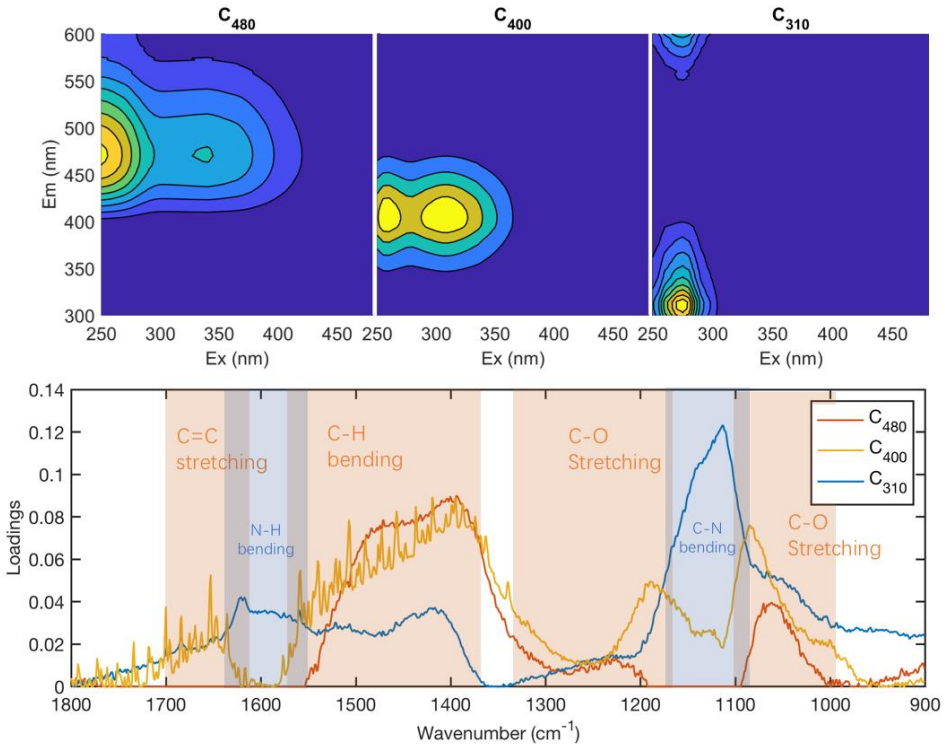

Figure 4-11. Integrated EEM fluorescent components and FT-IR spectra using the data fusion technique, showing major functional groups associated with each fluorescent component (C_{480}, C_{400}, and C_{310}).

4. 5. Discussion

4. 5. 1 Details on ACMTF model validations

The validation of the ACMTF model includes explained variance, the randomness of residuals, chemical coherence of component loadings, and reproducibility. Three models, including a two-component, a three-component, and a four-component model, were tested. The two-component model explained 95.66% and 96.67% of total variances for EEM and FT-IR spectra, respectively, which was considered to be a low explanation of total variance. The three-component model (98.69% EEM, 96.76% FTIR) and four-component model (99.21% EEM, 96.82% FT-IR) have very close explained percentages of total variances, which also implies that increasing a model's component number from three to four does not significantly improve the fit for the FT-IR spectra.

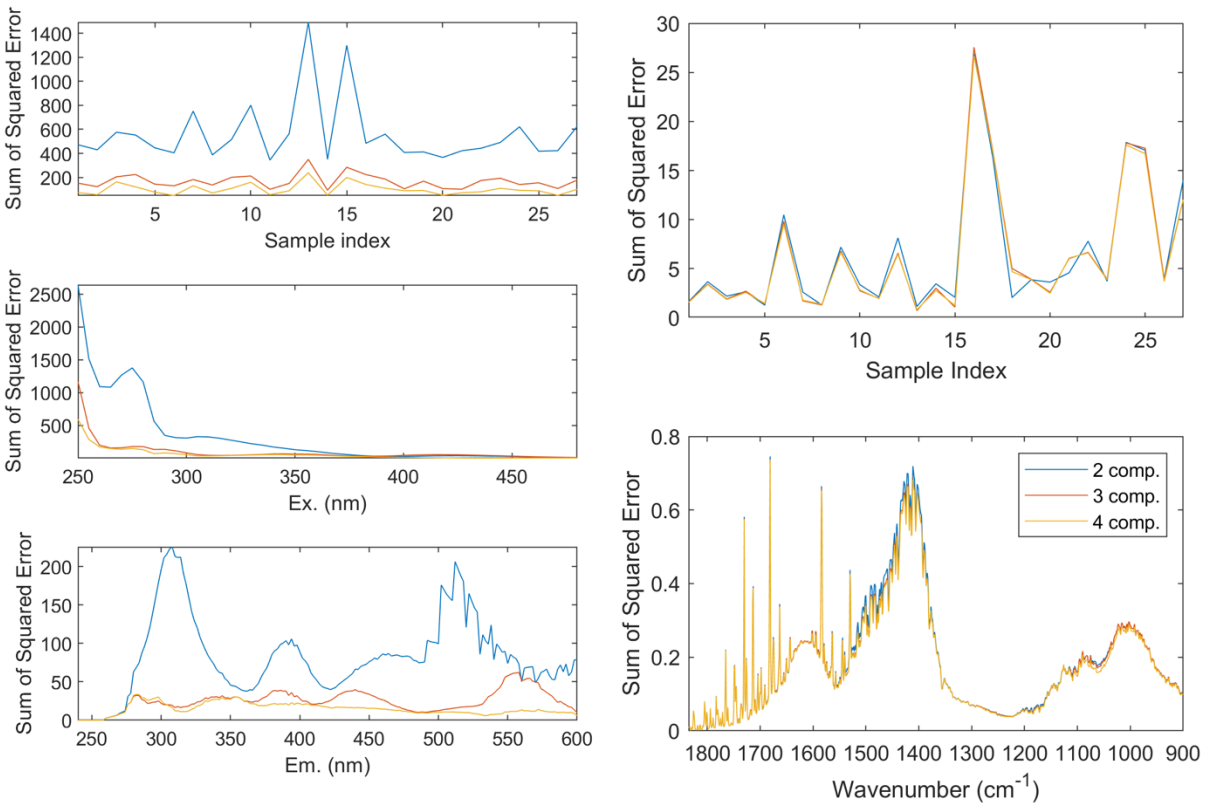

Figure 4-12. Sum of squared errors of the two-component, three-component, and four-component ACMTF models for both EEMs (left column) and FT-IR (right column).

As shown in Figure 4-12, the three-component and the four-component models for EEM have similar sum of squared error (SSE) values and distributions, and both of them have much smaller SSE compared

to the two-component model. For the FT-IR spectra, however, increasing the component number in the ACMTF models did not yield a smaller SSE. As for the distributions of residuals between all three models (Figure 4-13), the EEM contour from the two-component model poorly resembled the original EEM contour, and most of the residuals are larger than ±10% of the original signals. Residuals from both the three-component and four-component models are smaller and are within ±10% of the original signals, with major losses in the UVA areas (Ex 300-400 nm, Em>400 nm) and major gains in the UVC areas (Ex 250-300 nm), having residuals mostly around Ex = 250 nm. For the FT-IR spectra (Figure 4-14), the residuals of all three models are within ±5%. The negative residuals are mostly in the regions between 1700 cm^{-1} - 1650 cm^{-1} and 1000 - 900 cm^{-1}, representing double-bond stretching (majorly C=C and C=O) and bending (C=C), respectively. On the other hand, the positive residuals are mostly within the 1650-1000 cm^{-1} area, especially the 1500-1400 cm^{-1} and 1100-1000 cm^{-1}, representing bonds with C-H bending and C-O stretching, respectively. Overall, the model seemed to miss small signals of double bonds and added some single-bond signals, but the model results are all within ±5%, which is similar to errors associated with most analytical methods.

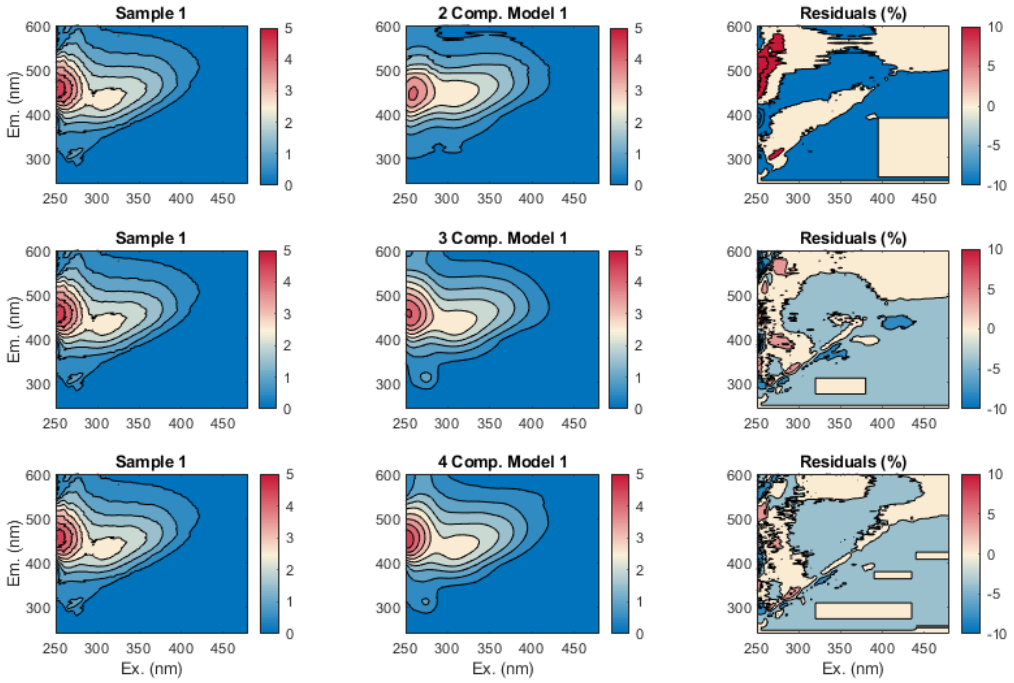

Figure 4-13. Comparisons in EEM contours between the original sample (left column) and model-derived (middle column) based on the two-component (top panel), three-component (middle panel), and four-component (bottom panel) models as well as their EEM residuals (in %, right column) between the original sample and model-derived spectra (using a HMW-DOM sample collected in September 2004 as an example).

Compared to the three-component model, the four-component model seems to perform better in residuals, errors, and the percentage of explained variances. However, the most important reason for adopting the three-component model is that some derived FT-IR spectra in the four-component model exhibited several unexpected sharp peaks. Thus, a three-component ACMTF model was adopted as more chemically realistic even though the explained variance was slightly lower compared to the EEM-PARAFAC model where >99% of the variance can be explained.

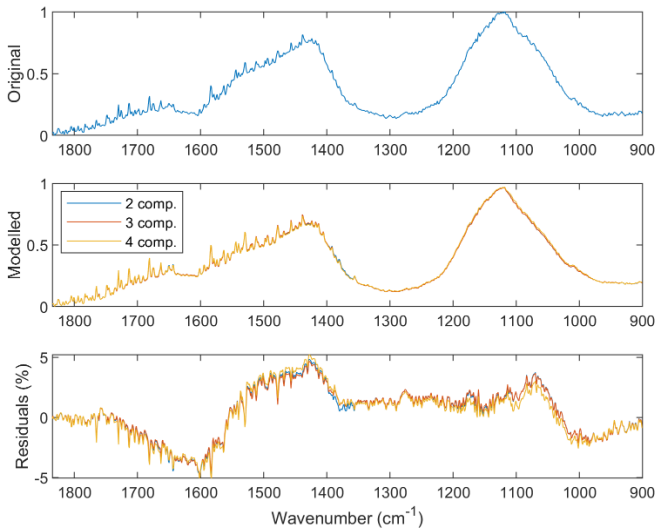

Figure 4-14. Comparisons in the FTIR spectra between the original sample (top) and model-derived spectra (middle) based on the two-component, three component, and four-component models as well as the residuals (in %, bottom) between original sample and model-derived spectra (using a HMW-DOM sample collected in September 2004 as an example).

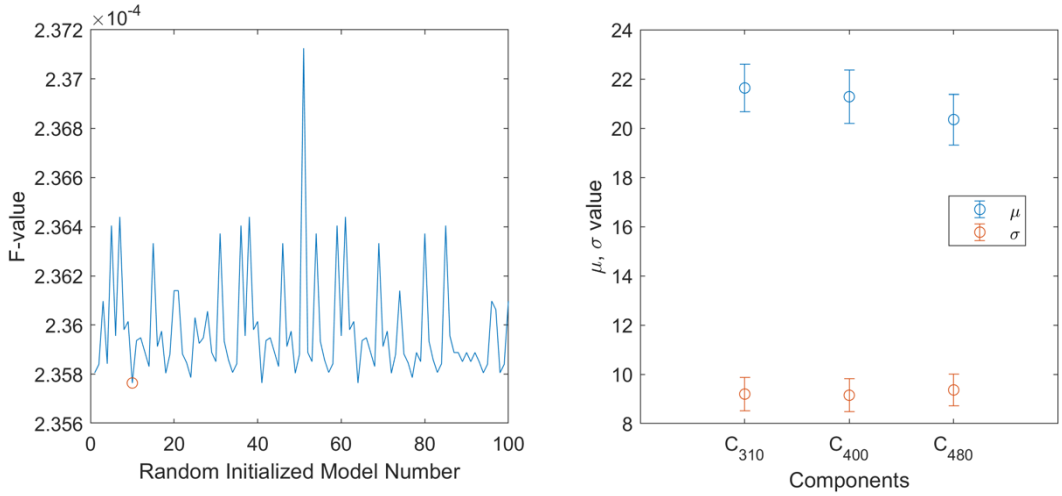

Figure 4-15. F-value of 100 random initialized models (left panel) and the factor weights (μ and σ) of validated models (right panel).

Uniqueness is validated using random initialization validation, which was run 100 times to obtain an optimized model least square error (Figure 4-15). Due to limitation of computational resources, the criterion for the convergence is set to 10^{-3}. The average λ and σ values for the three model-derived components with standard deviations (Table 4-2 and Figure 4-15) indicate that all components are shared.

The penalty terms μ_f and σ_f were used to gauge whether a specific component is shared between the EEM and FT-IR datasets in the ACMTF model. Theoretically, values of both μ_f and σ_f should be either 0 or 1, representing unshared or shared components, respectively. However, when the model is applied on field data instead of a randomly-distributed experimental dataset, the μ_f and σ_f are influenced by several distinct factors or parameters. For example, fluorescence quantum yield and ionization efficiency have been identified as potential factors controlling the sharedness of model-derived components between EEM and FT-ICR-MS datasets (Wünsch et al. 2018), while centering across the subject mode could have significant effects on these terms of a model integrating multiple instrument datasets (Acar et al. 2019). Thus, the μ_f and σ_f in this case could be regarded as weights or contribution of one component to the total variances to EEM and FT-IR spectra. As shown in Table 4-2, weights of all components for EEM (μ_f) are significantly higher ($p<0.01$) than those for FT-IR spectra (σ_f), indicating that FT-IR spectra are potentially unshared (Figure 4-15). However, considering that the percentage of total explained variances for FT-IR spectra is larger than 96%, lower contributions from FT-IR spectra (σ_f) might result from a different instrumental response since fluorescence is generated from excited electrons, while FT-IR spectra are based on rotational and/or vibration energy of chemical bonds. Similarly, Wünsch et al. (2018) found that C_{310} had significantly different component-weights for EEM and FT-ICR-MS and was regarded as a shared component. There is likely a small possibility of unsharedness between EEM and FT-IR spectra, which should be further evaluated using larger datasets in future studies. In addition, the more chemically meaningful results also render a strong support for the sharedness of C_{310} between EEM and FTIR spectra. Overall, although values of the penalty terms, μ_f and σ_f, are significantly different, all three components are regarded as shared components.

4. 5. 2 Dynamic changes in composition, source and sink of DOM

Arctic terrestrial ecosystems have undergone the effects of amplified warming, which have been manifested in the Arctic by changes including increasing river discharge, accelerated permafrost degradation, and changes in sources and fluxes of fluvial organic carbon (Mcguire et al. 2009; Schuur et al. 2015). Despite the importance of such changes, the detailed biogeochemical responses of Arctic rivers to climate and environmental change remains poorly characterized (Pokrovsky et al. 2011; Shogren et al. 2020; O'Donnell et al. 2021). Arctic rivers generally have a strong seasonality in discharge and DOM fluxes compared to other world rivers (Finlay et al. 2006; Guo and Macdonald 2006; Shogren et al. 2020). Time-series sampling that includes all seasons is thus indispensable to a better understanding of river DOM biogeochemistry and a more robust way to design long-term monitoring. However, sampling and characterization of DOM during ice-covered seasons remain scarce, due partly to the difficulty of sampling during a prolonged ice-covered season and harsh weather conditions during winter, and partly to the remoteness of most locations. The results of the year-long time-series study, presented here, clearly elucidate a dynamic change in DOM aromaticity, composition, fluorescent components, and optical properties during the transition from low productivity, ice-covered conditions, to spring freshet and higher productivity seasons. Given the resolution implied by nine samples collected monthly/bimonthly during the 2004-2005 sampling year, the data set presented here will likely not capture events of short duration, especially during spring freshet. Future studies with more frequent sampling, producing higher temporal resolution during periods of seasonal change are needed.

Other than changes in water and DOM sources, we speculate that physicochemical and microbial processes under the ice might significantly change the molecular structure of DOM and transform aromatic components to less aromatic DOM (Lee et al. 2018; Murphy et al. 2018), resulting in both a aromaticity decrease and an increase in $S_{275\text{-}295}$ or decrease in DOM molecular weight (Helms et al. 2008; Liu et al. 2020). Bacterial transformation and production would also induce higher and increased BIX values during the ice-covered period (Birdwell and Engel 2010; Frenette et al. 2008; Huguet et al. 2009). The concurrently high HIX and BIX values are seemingly contradictory, but the higher HIX values in the HMW-DOM pool under the ice may signal a shift in DOM sources from mostly surface soils during the

open seasons to increasing groundwater and deeper soil during winter (Striegl et al. 2007; O'Donnell et al. 2010) compounded by the preservation and accumulation of humic-like DOM components under the ice due to their relative resistance to biological degradation (Chen and Jaffé 2016; Xu and Guo 2018). In addition, the assembly of less aromatic LMW-DOM into HMW-DOM components (Xu and Guo 2018) may be another factor contributing to the $SUVA_{254}$ decrease in the HMW-DOM, as supported by the changes in spectral slope or apparent DOM molecular weight (Figure 4-4b).

During the spring freshet, BIX in the HMW-DOM pool dropped to its lowest value (0.42), while $SUVA_{254}$ and HIX increased abruptly, owing to the increased contribution from terrestrial DOM input resulting in a decrease in the proportion of autochthonous DOM. Interestingly, the LMW-DOM pool always exhibited a higher BIX value than the HMW-DOM pool ($p<0.01$), indicating the former LMW-DOM pool contained a relatively higher proportion of autochthonous components in this terrestrial DOM-dominated river, probably due to the preferential decomposition of autochthonous HMW-DOM (Benner and Amon 2015) and the transformation and production of DOM from HMW- to LMW-DOM during decomposition either under the ice or from a deeper active layer. After spring freshet, the decrease in $SUVA_{254}$ may reflect increased photodegradation due to decrease in turbidity and prolonged sun irradiation and the change in DOM sources/composition due to a deepened active layer, permitting leaching of older soils (Guo and Macdonald 2006; Zou et al. 2006; Aiken et al. 2014).

Compared to other aquatic ecosystems, DOM from the Yukon River is generally characterized as having intermediate to high HIX (between 1 and 7) and low BIX values (<0.6 - 0.8), consistent with a terrestrial DOM-dominated river system. For example, DOM from the Milwaukee River and Green Bay, both in Wisconsin and somewhat hypereutrophic, had medium HIX (<8) and BIX (0.55-0.8) values (Teber 2016; DeVilbiss et al. 2016a), whereas DOM from the Laurentian Great Lakes with predominantly autochthonous DOM (Zhou et al. 2016a) and Veterans Park Lagoon, a highly eutrophic ecosystem, both were characterized by lower HIX (<4), especially the open Great Lakes samples, but high BIX (>0.8) values (Figure 4-16). There is generally an inverse relationship between HIX and BIX, with a non-linear transition along a trophic gradient from high HIX-low BIX in terrestrial DOM-dominated river systems to medium BIX (with a wide range in HIX) in hypereutrophic ecosystems, to low HIX-high BIX values in

eutrophic, mostly autochthonous ecosystems (Figure 4-16). Thus, a tandem plot of both HIX and BIX indices and/or their ratios may be used as a compelling proxy to illustrate relative changes in DOM sources and trophic states as responses to a changing climate.

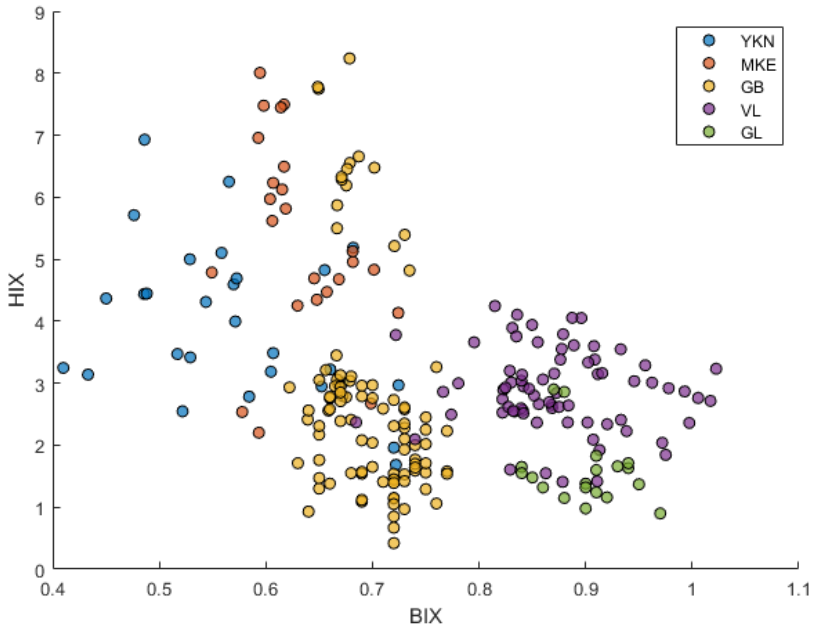

Figure 4-16. A scatter plot between HIX and BIX combining data from the Yukon River (YKN), the Milwaukee River (MKE, data from Teber 2016), Green Bay (GB, data from DeVilbiss et al. 2016), the Veterans Park Lagoon (VL), and the Laurentian Great Lakes (GL, data from Zhou et al. 2016).

4. 5. 3 Heterogeneity of DOM in between HMW- and LMW-DOM pools

Natural DOM has been shown to be highly heterogeneous in chemical properties, composition, and reactivities (Santschi et al. 1995; Benner and Amon 2015; Xu and Guo 2017). In the upper Yukon River, the HMW-DOM was found exclusively contemporary while the LMW-DOM had an average ^{14}C age of 976±212 y-BP during river open seasons (Guo and Macdonald 2006). In addition, fatty acids and biomarker composition were also distinctly different between the HMW- and LMW-DOM pools (Zou et al. 2006). In the lower Yukon River, the HMW-DOM is mostly associated with aromatic DOM compounds with much higher $SUVA_{254}$ values compared to its LMW-DOM counterpart (Figure 4-4a), deriving from either high vascular plant-derived DOM or leaching of surface organic-rich soil layers (Guo et al. 2007; Striegl et al. 2007). Similarly, higher $SUVA_{254}$ values and higher contents of hydrophobic organic acid were found for the HMW-DOM pool in the lower Mississippi River (Cai et al. 2015). During

the cold season, the decrease of $SUVA_{254}$ in the HMW-DOM (November 2004 – April 2005) suggests a change in DOM sources: a decrease in surface soil leached DOM and increase in groundwater DOM, consistent with increased ^{14}C ages between November and April (Aiken et al. 2014; Guo and Lin 2017).

In the lower Yukon River, the HMW-DOM is dominated by terrigenous organic components. In contrast, the LMW-DOM is mostly derived from microbial activity and degraded organic matter and has a higher BIX and protein-like component (C_{310}) during the ice-covered season (Figure 4-4c and Figure 4-8c). River water below the ice is generally much warmer than the air temperature during winter, and not so different from water temperature in summer, which promotes microbial activity during early spring. Psychrophilic bacteria can thrive in such conditions and contribute to the production of LMW-DOM through DOM transformation from either HMW-DOM or particulate organic matter (Deming 2002). Indeed, the LMW-DOM in Yukon River water contained abundant protein-like DOM components during the late winter season, which may have a priming effect on the degradation of terrigenous DOM although these components are lower during spring freshet (Figure 4-8c). Later in the summer season, both LMW- and HMW-DOM had a higher abundance of the protein-like C_{310} component (Figure 4-8c) due to *in-situ* production and transformation processes. Our results show that DOM characteristics of the LMW-DOM and HMW-DOM pools are distinct from one another, attesting to the need to characterize not just the bulk DOM, but also the DOM associated with different molecular weight- or size-fractions.

Previous studies have shown that DOM during the spring freshet is mostly terrigenous with high bio-lability. For example, Spencer et al. (2008) reported a maximum concentration of DOM in the Yukon River during spring freshet with high aromaticity and biological lability. Cooper et al. (2005) and Holmes et al. (2008) also reported higher DOM biological reactivity during spring freshet in Alaskan Arctic rivers, with a loss of labile DOM as high as 20-40% in samples incubated for 90 days. This rate of loss is much higher than the actual loss in DOM in the Yukon River during summer. High biological lability of DOM during spring freshet is likely associated with its contemporary or young ^{14}C age in Arctic rivers (Guo and Macdonald 2006; Raymond et al. 2007). However, old DOC released from ancient permafrost in Arctic river basins can also be labile (Gao et al. 2018; Vonk et al. 2013). Together, these observations imply that natural DOM is highly heterogeneous and contains a variety of components with diverse

functional groups distributed across a wide range of molecular size-fractions. Furthermore, DOM of varied composition and molecular size would likely have varying degradation potentials or varying photochemical and biological reactivities (Benner and Amon 2015; Chen and Jaffé 2016; Xu and Guo 2018). Thus, understanding changes in DOM sources, composition, optical properties, and molecular size distributions is essential to constrain the response of DOM biogeochemistry to seasonal processing and to understand what drives long-term environmental changes in arctic river basins.

4. 5. 4 Implications

The data presented here show that dynamic changes in DOM composition and spectral properties occurs throughout the year (ice covered, freshet, open season) in the lower Yukon River, and such changes differ among HMW- and LMW-DOM size fractions. Physicochemical and microbial processes, changes in sources of river water (especially groundwater vs. surface runoff), climate factors (e.g., temperature, precipitation, hydrology), and permafrost dynamics likely all play a role in regulating river DOM quantity and composition, and the response of river biogeochemistry to climate and environmental change. The results presented here provide a baseline dataset against which future changes in the context of a warming climate may be assessed. We strongly suggest that multiyear observations incorporating additional organic biomarkers and organic molecular structural data are needed in future studies to make further progress in establishing specific linkages between the forcing by climate change, and the biogeochemical response occurring in Arctic river basins.

Chapter 5.Seasonal variations of DOM in the Milwaukee River, Wisconsin

5. 1. Abstract

Although the Milwaukee River is a small river in the southeast quarter of Wisconsin, it attracts attentions and interests of local Wisconsin residents. In this study, I will address two questions: 1) Can changes in DOM quantity and quality in the Milwaukee River be related to the human activities and environmental change in the watershed including the removal of Estabrook dam? 2) Can changes in the size and composition of DOM be used as proxies for agricultural runoffs and inputs from Milwaukee metropolitan area? River water samples from upstream and downstream of the Milwaukee River were collected monthly from Oct. 2017 to Oct. 2019. Multiple indicators, including dissolved organic carbon (DOC) concentration, chromophoric dissolved organic matter (CDOM) absorption coefficient (a_{254}), spectral slope ($S_{275\text{-}295}$), specific UV absorbance ($SUVA_{254}$), biological index (BIX) and humification index (HIX), were used to evaluate spatial and temporal variations of dissolved organic matter (DOM) in the Milwaukee River. Seasonal variations in DOM concentration and compositions were observed in the two-year observation period; in wet seasons, large river flow brought a great amount of terrestrial humic-like DOM with high molecular weight and high aromaticity, while in dry seasons, autochthonous protein-like DOM was released by microbial communities under ice. Comparisons were made between the upstream and downstream of the river, as well as between 2014-2015 and 2018-2019. Analysis of the multiple indicators indicates that 1) DOM in the Milwaukee River is majorly terrigenous; 2) Anthropogenic impacts from metropolitan area on the riverine ecosystem can be observed from several indicators; 3) Based on the comparisons of BIX and HIX between 2014-2015 and 2017-2019, the Estabrook dam removal at the beginning of 2018 has minor impacts on the fluvial ecosystem.

5. 2. Introduction

The Milwaukee River drains through metropolitan City of Milwaukee into Lake Michigan, which is the source of drinking water for more than 900,000 people (Rath et al. 2013; Milwaukee Riverkeepers 2015; Teber 2016). The water quality and the health of the riverine ecosystem are the big concerns of local

citizens, although the river discharge is not at high level. The river has been currently monitored by several organizations. For example, the discharge and gage height of the Milwaukee River has been monitored by the United States Geological Survey (USGS) and the Milwaukee Metropolitan Sewerage District (MMSD) since 1986. The Milwaukee Riverkeeper have investigated the Milwaukee River basin since 2012, and they obtained fundamental water quality data including water temperature, dissolved oxygen, pH, turbidity and phosphorus concentration. Despite the routine investigation, little is known about source and compositions of DOM in the river and how the river impacts on Lake Michigan.

Metropolitan areas bring anthropogenic impacts on aquatic environments including the Milwaukee River. Downstream of the Milwaukee River is strongly influenced by human activities. To evaluate the influence of human activities from metropolitan area, I compared DOM optical properties of water samples collected in the upstream of the Milwaukee River with those from downstream river water.

In the beginning of 2018, the 80-year-old Estabrook Dam, originally located in the Estabrook County Park (43.0993° N, 87.9078° W), was totally removed. Before the demolition of the dam, the dam gates kept open since 2008 due to the safety and maintenance violations. Here I propose a hypothesis that since all gates of the Estabrook Dam are kept open, the existence of the dam would not increase residence time of the river water, and the removal of the dam would not have huge impacts on the riverine ecosystem. In order to test the hypothesis, I compared DOM parameters before and after the dam removal. Six parameters from this study, including DOC, CDOM a_{254}, $SUVA_{254}$, $S_{275\text{-}295}$, BIX, and HIX, are compared with corresponding data collected between 2014 and 2015 (Teber 2016).

5. 3. Methods and Materials

5. 3. 1 Study sites and water sample collection

Water samples from the Milwaukee River were collected monthly or bimonthly from two stations located at Upstream and Downstream of the river (Figure 5-1). Despite that they are named as Upstream and Downstream, they are all located at the southern quarter of the Milwaukee River basin (Figure 5-1), where is the most densely populated area in the state, holding 90% of the basin's population. The Upstream sampling point is at the Kletzsch Park, Glendale, WI (43.1405°N, 87.9227°W), where less

human activities and anthropogenic impacts on the river are expected. The Downstream sampling location is at the very lower Milwaukee River (43.0325°N, 87.9096°W), challenged by metropolitan pollution and habitat modification. In the freezing winter, ice auger was used when the river was frozen (Jan 18, 2018, Jan 31, 2019, and Mar 8, 2019), and water samples were collected under the river ice. In the freezing winter, ice auger was used when the river was frozen (specifically, Jan 18, 2018, Jan 31, 2019, and Mar 8, 2019), and water samples were collected under the river ice. Temperature and specific conductivity were measured *in situ* using a YSI probe. Chlorophyll (Chl *a*), pH, dissolved organic carbon (DOC), UV-visible absorption spectra, fluorescence EEM, and size fractogram of DOM were measured immediately after water samples shipped back to lab.

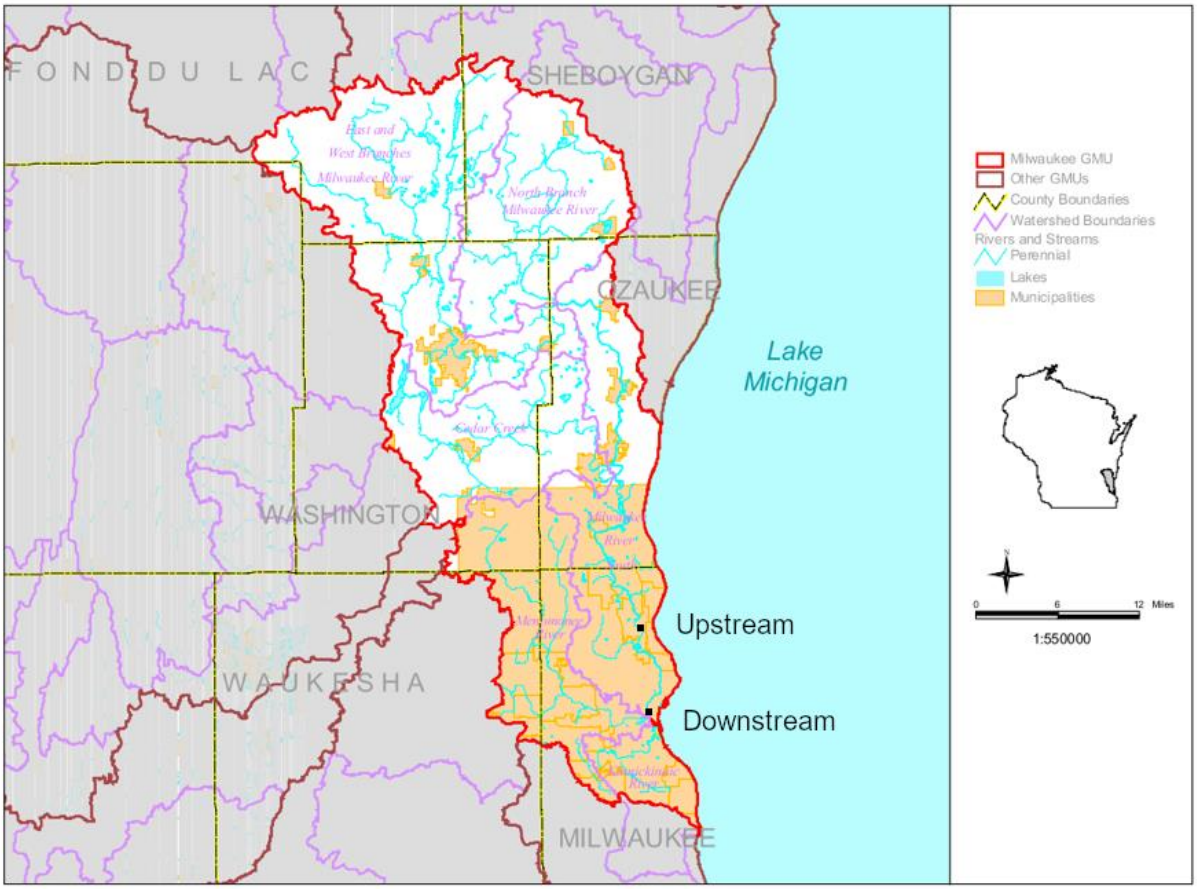

Figure 5-1. Map of the Milwaukee River Basin and our sampling stations (black dots). Modified from Mccallum et al., (2001).

Chl *a* concentration was analyzed using a spectrophotometric method (Strickland and Parsons 1972). Briefly, water samples (500 ml) were immediately filtered using GF/F filters (Whatman, 47 mm), which then stored frozenly at -20 ˚C. The filters were cut into small pieces in 10 ml 90% acetone solution

followed by 4 ˚C extraction on a shaker for 12 hrs. After centrifugation (4000 r/min, 10 min), absorbance at 750, 664, 647, and 630 nm were measured using a spectrophotometer (Agilent, US). Chl *a* concentration was calculated based on these four absorbance values. The measurements of DOC concentration, UV-visible absorption spectra, and fluorescence EEM were describe in previous chapters.

5. 4. Results

5. 4. 1 Hydrology and water chemistry in the Milwaukee River

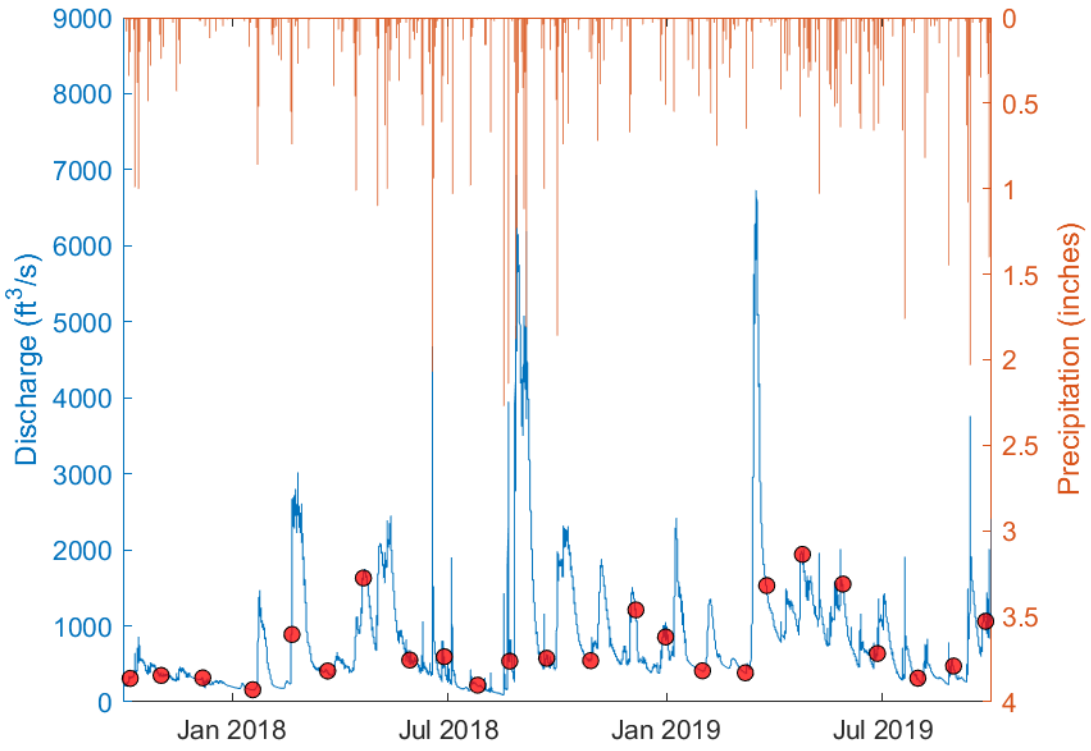

Figure 5-2. The daily discharge of the Milwaukee River (downloaded from USGS website. http://waterdata.usgs.gov/nwis) during our sampling and the daily precipitation in the Downtown Milwaukee Metropolitan area (downloaded from National Centers for Environmental Information in National Oceanic and Atmospheric Administration, https://www.ncdc.noaa.gov/cdo-web/). The red dots denote our sampling date.

The daily discharge of the Milwaukee River from Oct 2017 to Oct 2019 is shown in Figure 5-2. Based on temperature and precipitation in the Milwaukee, there are two seasons in one year: a wet season and a dry season. Genearlly, wet season in the Milwaukee lasts 6.4 months, from March 29th to October 11th (Spark 2017; Reichle et al. 2017). The discharge of the Milwaukee River is significantly higher in wet seasons than in dry seasons (p>0.05). The discharge is strongly influenced by the local precipitation in the river basin; large discharge spikes are shown subsequently after local precipitation events (Figure 5-2).

However, the daily averaged river discharge is not significantly linearly correlated with the daily precipitation in the Milwaukee area (R^2 = 0.02, p >0.05). This indicates that some discharge spikes are associated with precipitation events in wet season (e.g. the large spike in the Sept 2019), but the others are related to other factors, such as the spring freshet due to air temperature rising. Monthly sample collection might miss some small events in terms of discharge, but important events can still be observed in this sampling resolution.

Water temperature in the Milwaukee River varied from 0.0 °C during winters to 24.9 °C in summers. Obviously, the water temperature in the downtown area (average: 11.45 °C) is higher than that in the upper stream (average: 10.81 °C, Figure 5-3 left panel), especially in summers (average 23.43 °C to 22.1°C). Specific conductivity ranged from 490 to 1504 µS cm^{-1}. The maximum of the specific conductivity exhibited in winter due to the road salt distribution and the ice formation exclusion effects (Figure 5-3 right panel). The two lowest specific conductivity values shown in Figure 5-3 are corresponding to the heavy precipitation events before sampling dates (2/20/2018 and 8/22/2018).

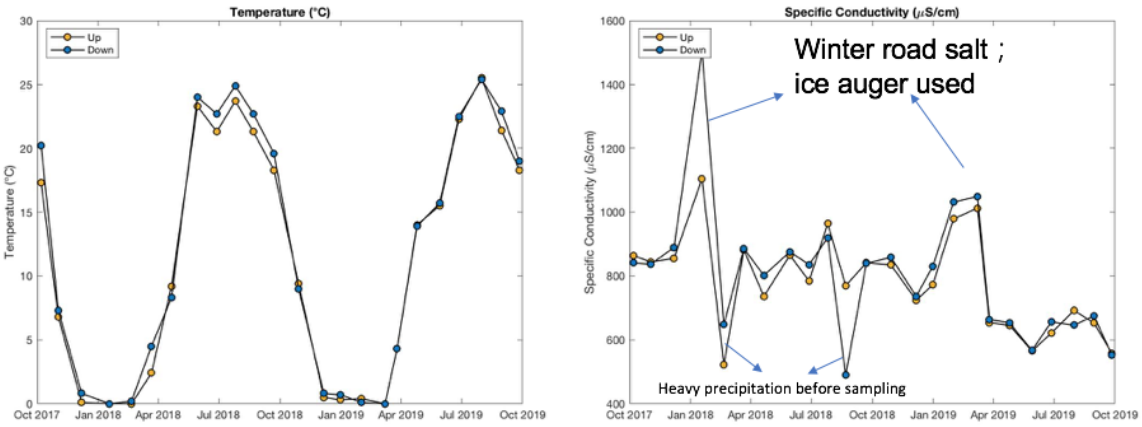

Figure 5-3. Seasonal variations in water temperature (°C, left panel) and specific conductivity (µS/cm, right panel) at upper stream station (yellow dots) and downstream station (blue dots) in the lower Milwaukee River.

The range of pH was 7.75 to 8.67, with an average of 8.29 ± 0.24 (Figure 5-4 left panel). Based on the Milwaukee Riverkeeper annual reports, pH in the river from 2012-2018 was vibrating between 6 and 9, which was defined as *good* for swimmers (Milwaukee Riverkeepers 2011, 2012, 2013, 2015, 2018). In light of the Chl *a* concentration, there were five peaks showing in the time series curve. These peaks are all associated with the large discharge events (Figure 5-4).

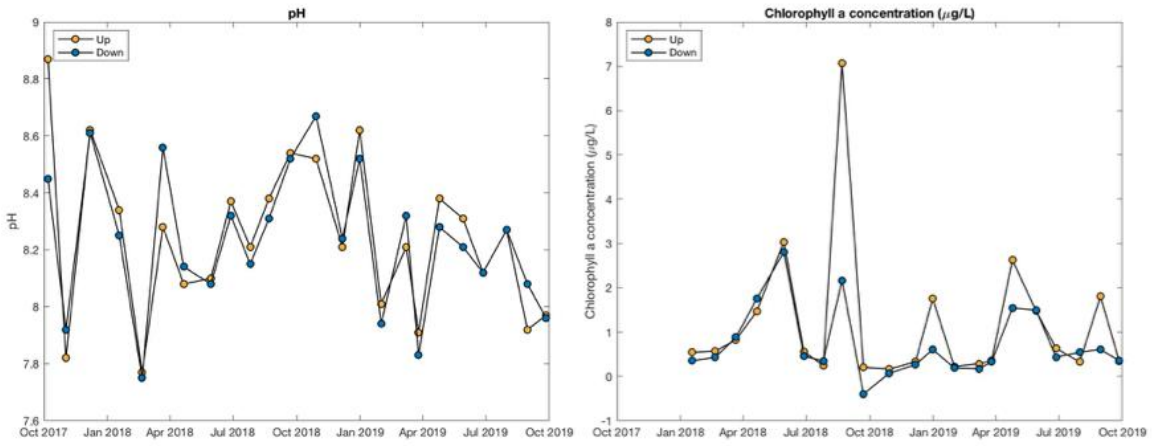

Figure 5-4. Seasonal variations in pH (left panel) and Chlorophyll a (in μg/L, right panel) at upper stream station (yellow dots) and downstream station (blue dots) in the Milwaukee River.

5. 4. 2 DOM concentrations and optical properties

Concentration and properties of the DOM in the Milwaukee River were largely influenced by the hydrological and physio-chemical properties. Seasonal variations of CDOM a_{254}, DOC, $S_{275\text{-}295}$ and $SUVA_{254}$ are shown in Figure 5-5. CDOM a_{254} ranged from 28.07 to 121.6 m^{-1}, which is consistent with CDOM a_{254} observed in 2014-2015 (24.7~145 m^{-1}, Teber, 2016). CDOM a_{254} exhibited the highly similar trend as DOC concentration. High correlation between DOC and CDOM a_{254} (including both upstream and downstream samples) indicated that DOC and CDOM pools are largely overlapped in the Milwaukee river (Figure 5-6, $R^2 = 0.869$). Additionally, this also implied the significance of terrestrial input to this river basin. DOC concentration ranged from 334.7 to 1132 μmol/L, which is consistent with the determinations in 2014-2015 (318~1189 μmol/L, Teber, 2016). Seasonal variations of DOC concentration are highly related to the discharge events. Due to the allochthonous DOM source input, when the river discharge was high, the DOC was correspondingly higher. DOC concentration in the upstream station (647.15 ± 200.42 μmol/L) were slightly higher than DOC in the downstream station (616.78 ± 182.03 μmol/L, $p<0.05$).

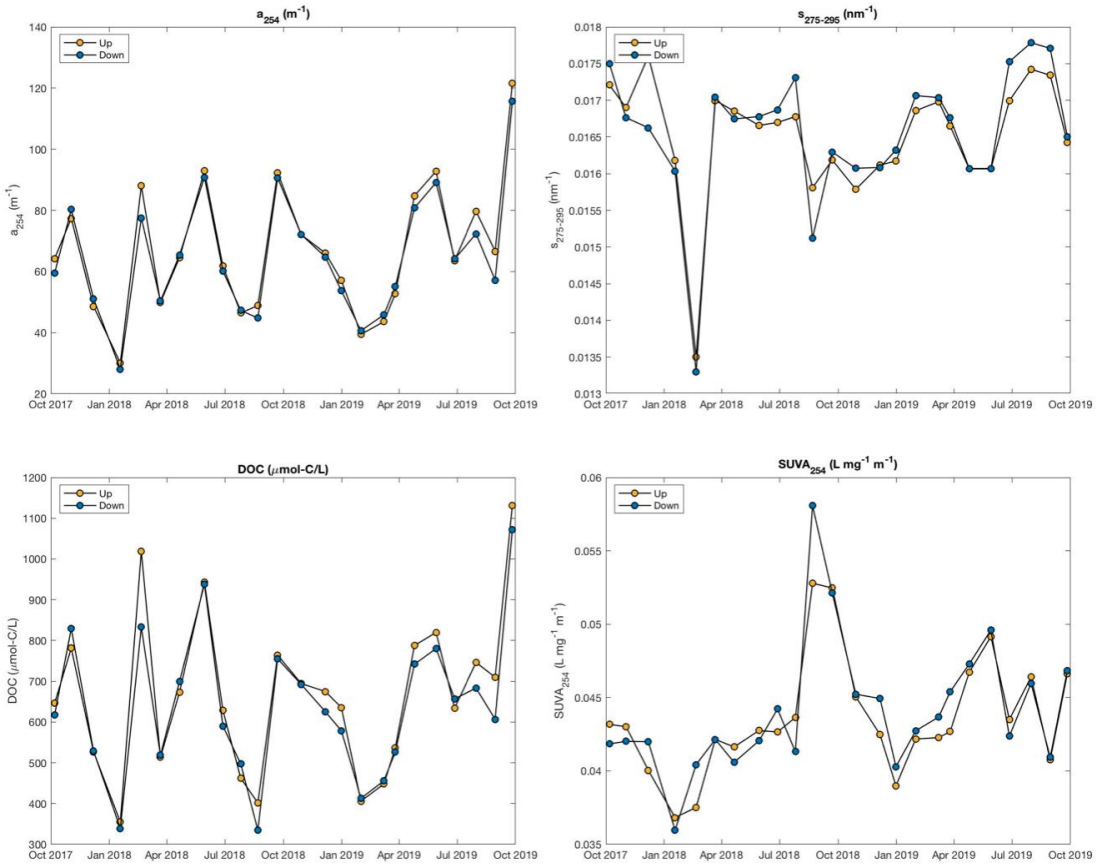

Figure 5-5. Seasonal variations in a_{254} (m^{-1}), $S_{275\text{-}295}$ (cm^{-1}), DOC (μmol/L) and $SUVA_{254}$(L mg^{-1} m^{-1}) at upper stream station (blue line and dots) and downstream station (orange line and dots) in the lower Milwaukee River.

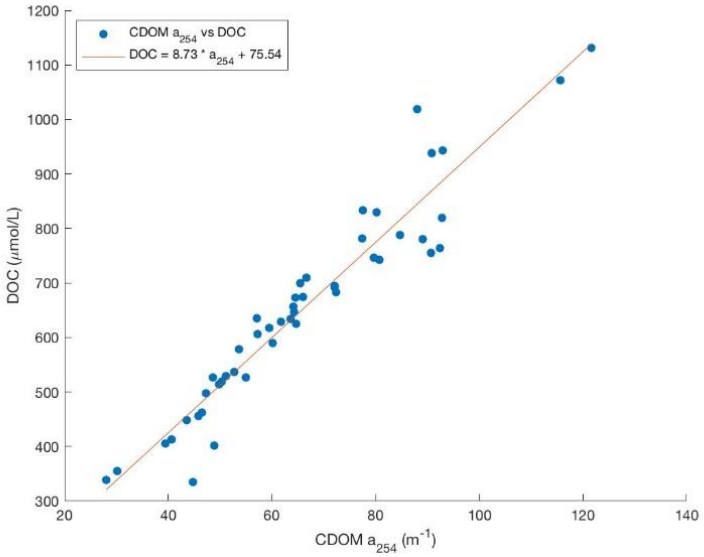

Figure 5-6. Correlation between DOC and CDOM (a_{254}) in the lower Milwaukee River waters.

$S_{275\text{-}295}$ ranged from 0.0133 to 0.0173 with an average of 0.162 ± 0.001 nm^{-1} (Figure 5-5). The high molecular weight DOM input, indicated by the low values in the February and August, was corresponded

to the high discharge event, which implied that the high molecular weight humic substances from soil were leached into river water. $SUVA_{254}$ indicated the aromaticity of DOM in the river water. It ranged from 3.00 to 4.84 L μmol^{-1} m^{-1} with an average of 3.63 ± 0.44 L μmol^{-1} m^{-1} (Figure 5-5). The $SUVA_{254}$ peak in the late summer illustrate the source of the high flux terrestrial input.

5. 4. 3 Properties of fluorescent DOM

BIX values in the river varied from 0.50 to 1.41 with an average of 0.72±0.26, while HIX values ranged from 1.45 to 19.21 with an average of 8.97±4.63 (Figure 5-7). Further discussion of the trending of these two fluorescence indices are presented in the Discussion section.

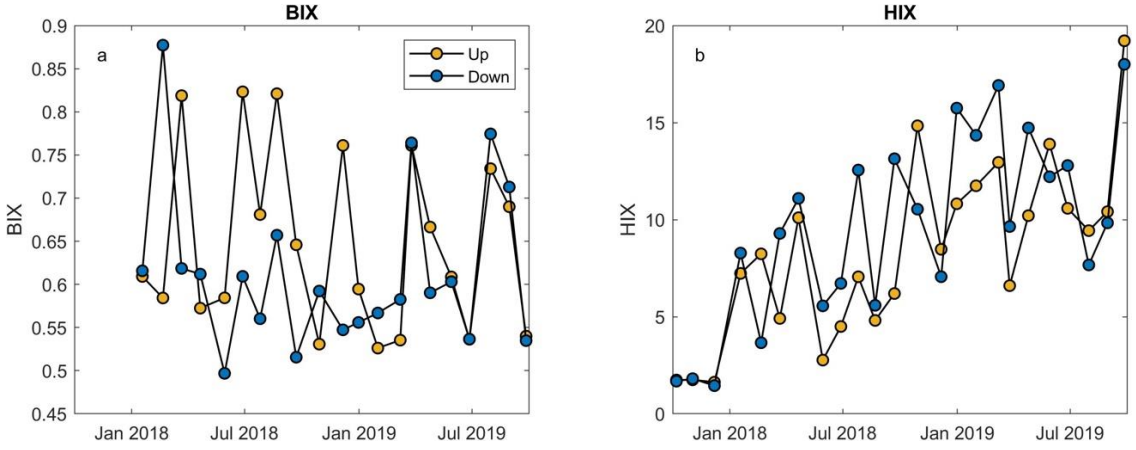

Figure 5-7. Seasonal variations in (a) BIX and (b) HIX at the upper stream station (blue line and dots) and downstream station (yellow line and dots) in the lower Milwaukee River in 2018-2019.

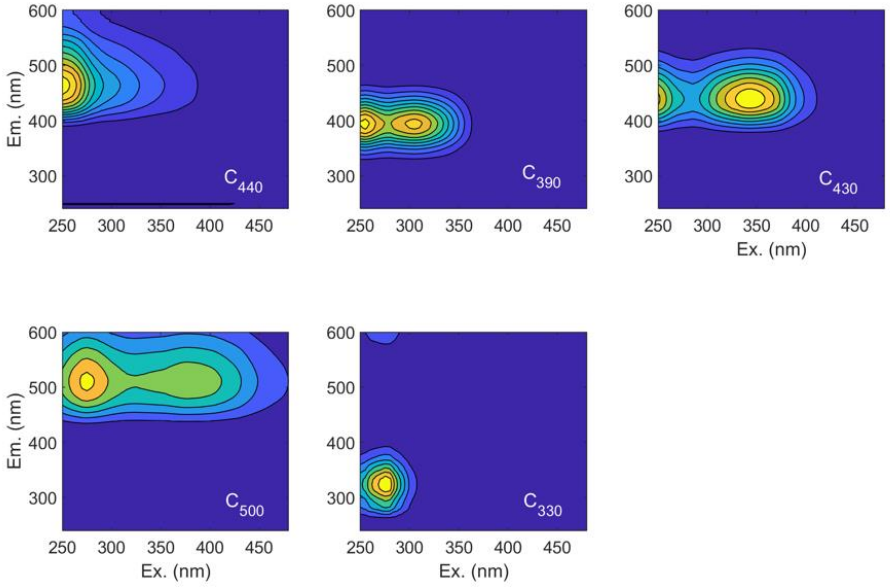

Figure 5-8. Excitation-emission contours of the five fluorescent components derived from PARAFAC model.

Table 5-1. Peak excitation wavelength (Ex.) and emission wavelength (Em.) of components from PARAFAC in the lower Milwaukee River 2018-2019. Numbers in the parentheses in the Ex.(nm) column are the wavelength of secondary peak in the component. The Peak Name were from empirical peak-selection method (Coble 2007).

Component Name	Ex. (nm)	Em. (nm)	Peak Name	Properties
C_{500}	265	506	D	Humic-like
C_{440}	250(335)	442	A and C	Humic-like
C_{430}	250	428	A	UVC humic-like
C_{390}	250(310)	392	M	Microbial-derived; Humic-like
C_{330}	280	332	T	Protein-like

Five fluorescent DOM components were deconvoluted from the assembled EEM dataset using PARAFAC analysis (Figure 5-8). Details of these PARFAC-derived components were displayed in Table 5-1. $C_{500,}$ C_{440}, and C_{430} were typically recognized as allochthonous terrigenous fluorescent DOM. C_{500} was a terrestrial fluvic acid fluorophore group commonly observed in a wide range of freshwater environments (Stedmon and Markager 2005; Wünsch et al. 2019a). It contained a UVA major peak and a visible secondary peak named Peak D. The major peak of C_{440} covers the Peak A which is a UVA humic-like fluorophore and the secondary peak covers Peak C, a UVC humic-like fluorophore (Coble 2007). This component was dominating the DOM exported from the natural catchments during the warmer months of the year. It was also exported from agricultural catchments (Stedmon and Markager 2005). C_{430} had high similarity as the major peak of C_{440}, which covers Peak A as well. Surprisingly C_{440} and C_{430} were intermediately linear correlated ($R^2 = 0.43$). The different behaviors could probably be due to the existence of secondary peak in C_{440}. C_{390} was the most intriguing component, whose source were complicated. It was first regarded as a visible marine humic-like fluorescent component, which was nearly recognized as an endmember of marine end in an estuary (Coble et al. 1998). However, some studies observed it in freshwater field (Stedmon et al. 2003). Furthermore, it was related to the microbial activity or anthropogenic activities because its existence might due to the spreading of animal waste on fields as fertilizer (Stedmon and Markager 2005). C_{330} was an autochthonous component named Peak T previously. It was the only protein-like fluorescent component observed in this study.

C_{500} and C_{440} showed high values in high-discharge periods (May and September) and low values in winters (Figure 5-9). These two components are significantly linear correlated (R^2=0.26, n=48, p<0.05). C_{430} and C_{390}, which are highly correlated (R^2=0.32, n=48, p<0.05), showed high values in the summer and fall but not in the snowmelt season, which might be due to the different sources and DOM composition. Instead, the C_{330} exhibited the maximum value in the February (Figure 5-9) probably due to the release of microbial communities in the static under-ice circumstances.

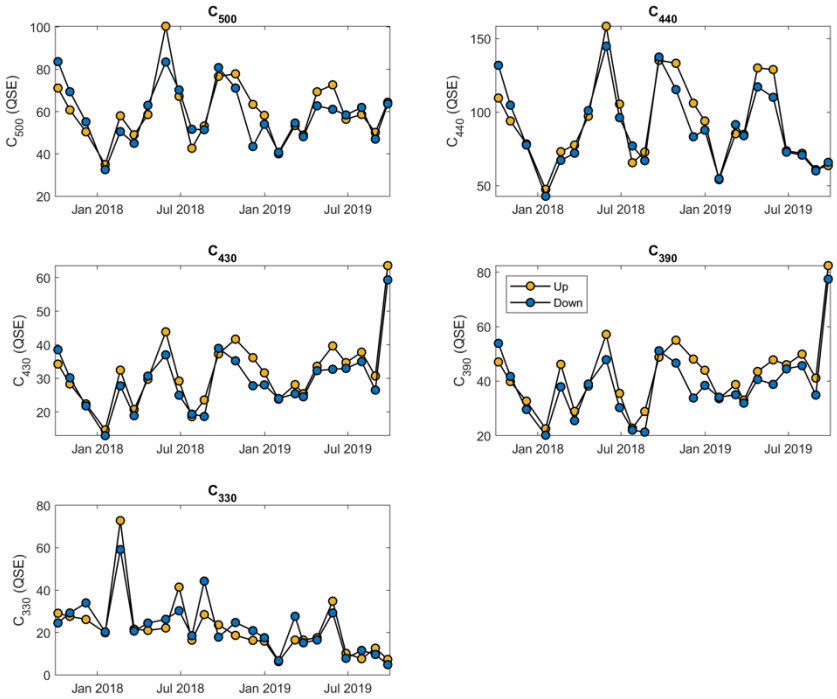

Figure 5-9. Seasonal variations of PARAFAC-decomposed components at upper stream station (yellow dots) and downstream station (blue dots) in the lower Milwaukee River.

5. 4. 4 Partitioning of carbon and nitrogen in particulate and dissolved phase

Concentration of DOC ranged from 334.7 to 1131.8 μmol L^{-1}, while the POC ranged from 18.8 to 171.2 μmol L^{-1} (Figure 5-10 left panel). The percentage of POC in bulk organic carbon pool is less than a quarter in most of the time (4.67 to 27.64 %, Figure 5-11). The highest POC percentage was observed in the summer of 2018 (Aug 22, 2018), when DOC (TDN as well) concentration reached to the minima and the Milwaukee River contained highest concentration of POC (PN as well, Figure 5-10). TDN ranged from 0.22 to 1.89 mg L^{-1}, and PN ranged from 0.021 to 0.39 mg L^{-1} (Figure 5-10 right panel). TDN was high in winters and relatively low in summers, which is contrary to the trending of PN concentration. POC and PN are significantly linear correlated (R^2=0.906, p<0.01). The C/N ratio (mole ratio) varied

from 6.0 to 12.3 with an average of 8.4±1.5, which is slightly higher than the Redfield ratio (106:16). It can be explained by the greater input of terrigenous chemical molecules, containing more carbon than nitrogen in the river. The bulk DOM is further size-fractionated into HMW-DOM (> 1 kDa) and LMW-DOM (< 1 kDa, Figure 5-12). In terms of both DOC concentration and CDOM a_{254}, HMW-DOM are greater than LMW-DOM (Figure 5-12 left panel). HMW-DOC is 48.0~85.8 % of bulk DOC (Figure 5-12b), and HMW-CDOM a_{254} is 61.7~87.9 % of bulk CDOM a_{254} (Figure 5-12d). Thus, the seasonal variation trendings of HMW-DOC and HMW-CDOM a_{254} are highly similar with bulk DOC and bulk CDOM a_{254}, respectively. DOC concentration and CDOM a_{254} for HMW-DOM are highly linear correlated ($R^2 = 0.93$, $p<0.05$), but not for LMW-DOM ($R^2 = 0.07$, $p>0.1$), probably resulting from the major contributor of HMW-DOM: the natural terrestrial DOM, whose size distribution is mostly between 1 kDa and 3 kDa (Lin and Guo 2020; Asmala et al. 2021). Due to the highly aromatic chemical properties of terrestrial DOM, the terrestrial-dominant area shows high correlation between DOC and CDOM a_{254} (Fichot and Benner 2014).

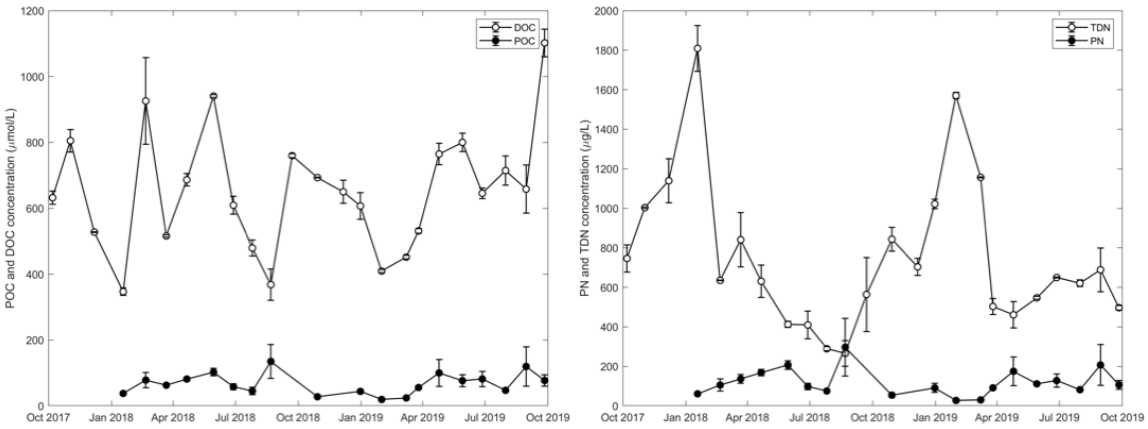

Figure 5-10. Seasonal variation of averaged DOC and POC concentration (left panel) and averaged TDN and PN concentration (right panel) in the Milwaukee River.

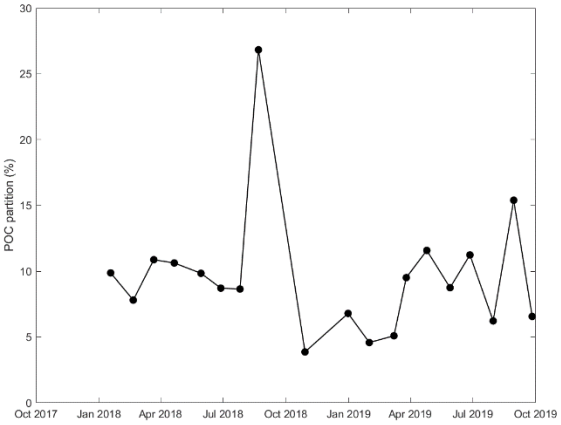

Figure 5-11. Percentage of POC in the organic carbon pool in the Milwaukee River.

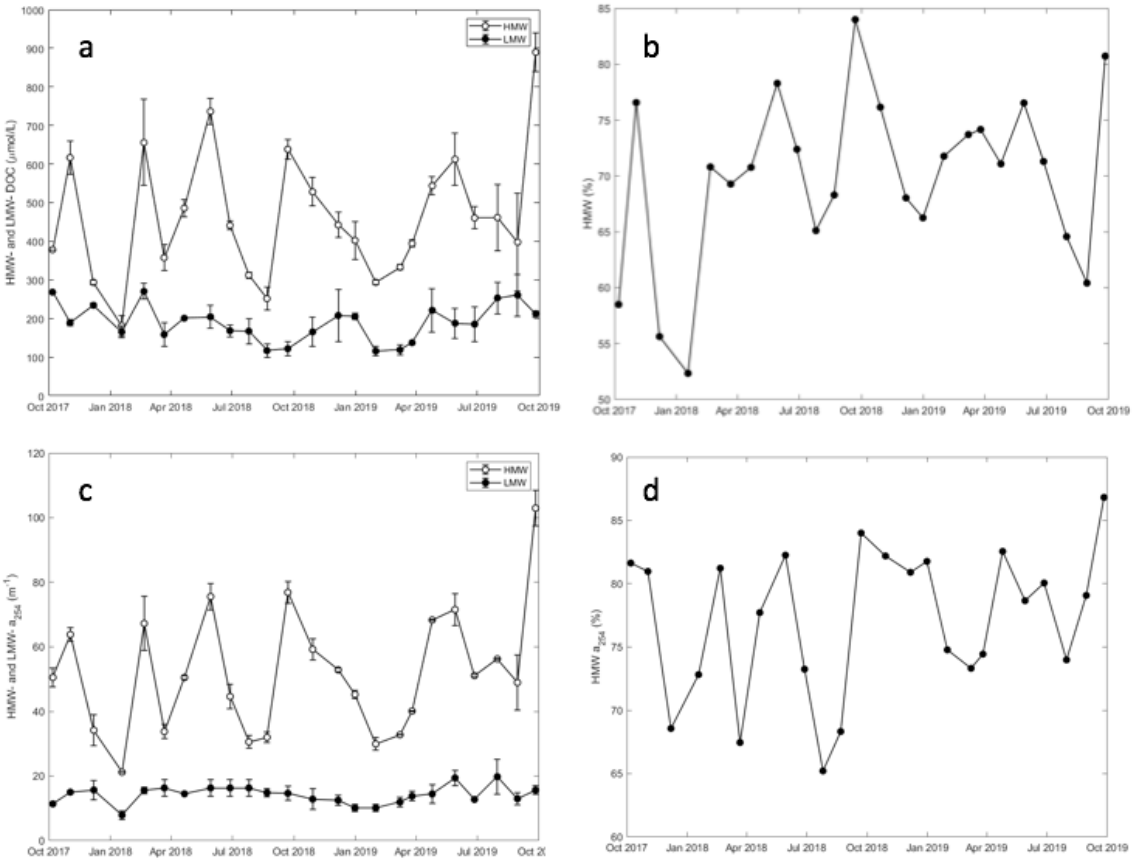

Figure 5-12. Seasonal variation of (a) averaged HMW-DOC and LMW-DOC concentration, (b) percentage of HMW-DOC, (c) averaged HMW a_{254} and LMW a_{254}, and (d) percentage of HWM a_{254} in the Milwaukee River.

5. 5. Discussion

5. 5. 1 Seasonal variations and relations of DOM indicators in the Milwaukee River

The Milwaukee River is predominated by terrigenous DOM, though Chl *a* spikes (Figure 5-4, right panel) were observed sometimes in summers. River discharge is an essential key point in this river. In wet seasons, high river discharge, induced by heavy precipitation, leads to significant DOM flux (high DOC and CDOM a_{254} in Figure 5-5) with characteristic features of terrigenous DOM: high molecular weight (indicated by $S_{275\text{-}295}$ in Figure 5-5), high aromaticity (indicated by $SUVA_{254}$ in Figure 5-5), and emitting humic-like fluorescence (Figure 5-9). Conversely, in dry seasons, especially ice-covered season, the under-ice DOM was of low molecular weight, low aromaticity and high autochthonous fluorophores (Figure 5-5 and Figure 5-9). The dynamic DOM composition and properties controlled by river flow were commonly observed in riverine ecosystems, such as Changjiang River (Yangtze River, Bao et al., 2015; Lv et al., 2019), Mississippi River (Del Castillo and Miller 2008; Guo et al. 2009), and Yukon River (Spencer et al. 2008).

To further understand the potential internal relationship between these optical indicators and measurements, the Pearson's correlation coefficients are calculated and shown in Figure 5-13. A couple of interesting pairs are found in the figure. First, Chl *a* concentration is highly correlated with POC and PN concentrations (r=0.76 and 0.88), and POC and PN are also positively correlated (r=0.95). It's quite common to observe these correlations since they are all measured from 0.7 μm GF/F filters, which collect particulate organic matter (POM) in water samples. Second, as mentioned above, CDOM a_{254} are correlated with DOC concentration (r=0.96), DOC retentate concentration (r=0.96) and CDOM a_{254} of retentates (r=0.99). Nevertheless, CDOM a_{254} are highly correlated with fluorescent component C_{430} (r=0.87) and C_{500} (r=0.90), which are both terrestrial humic-like components. This is additional evidence to prove that this river is terrestrial-input dominant watershed.

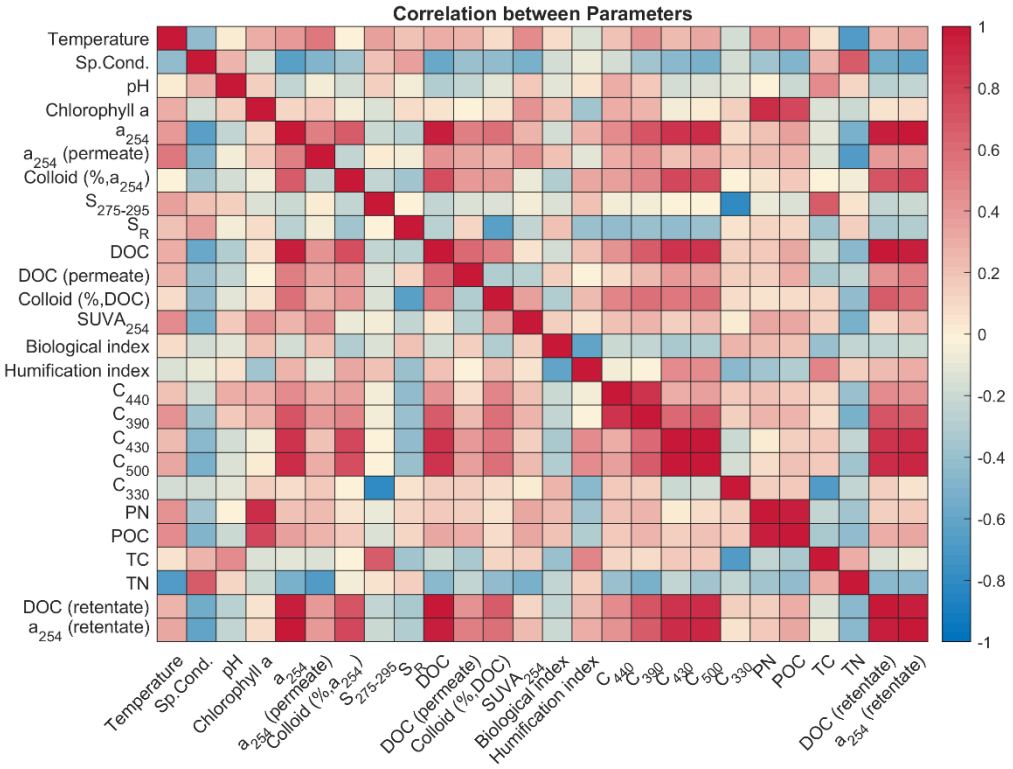

Figure 5-13. Linear correlations between measurements. The color of blocks indicates the Pearson's correlation coefficient.

5. 5. 2 Potential anthropogenic impacts on the river

As shown in Figure 5-14 and Table 5-2, several hydro-chemical parameters are significantly different between upstream and downstream samples, including water temperature, CDOM a_{254}, DOC, POC, PN, and fluorescent component C_{430} and C_{500}. Water temperature in the upstream is obviously lower than those in the downstream (Figure 5-3). The differences are more obvious during the summer because in the winter, the whole river is ice-covered, and water temperature under ice are close to 0 °C. The differences of water temperature between upstream and downstream is part of urban heat island effect due to the human activities in metropolitan area (Solecki et al. 2005). The anthropogenic impacts could also be observed in winters. Due to the road salt distribution, specific conductivity is relatively higher in winters. (Figure 5-3).

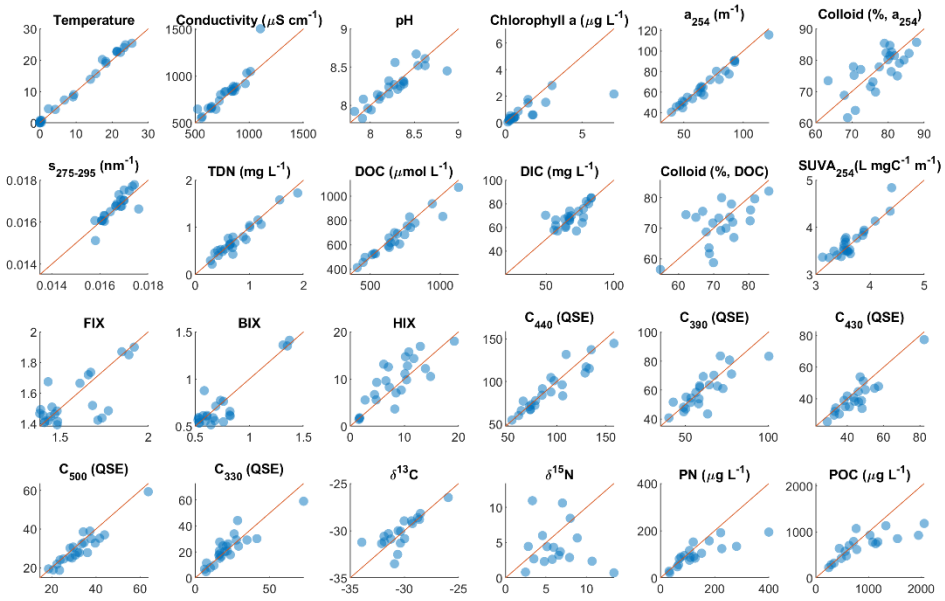

Figure 5-14. Comparison of different parameters between upstream (in x-axis) and downstream samples (in y axis) in the Milwaukee River.

DOM concentration, indicated by DOC and CDOM a_{254}, were significantly higher in the upstream than in the downstream (Table 5-2, Figure 5-5). Meanwhile, POM is of similar trend as DOM; POC and PN are higher in the upstream than in the downstream (Table 5-2). These differences might be due to the mixing of river water and Lake Michigan water in the downstream of the Milwaukee River.

C_{430} and C_{390} were also significantly higher in the upstream than in the downstream (Figure 5-9, Table 5-2). However, based on the experiences from previous studies and observations, the decrease of C_{430} and C_{390} in the downstream can hardly be resulted from human activities. The most possible reasons for these differences might be the different with the DOM and POM mentioned above. These two components were reported as microbial-derived components. The decrease of these components might be due to the microbial community changes in the Milwaukee River.

It is worth noted that C_{330} is the most likely to be one of the anthropogenic indicators among these five components, since this component could be frequently observed in influent and effluent of sewer water plants, which is most probably related to human activities. However, no significant difference of C_{330} was observed between upstream and downstream. This implies that the Milwaukee River received minor sewer water from the metropolitan areas. Other anthropogenic DOM (like pharmatheutical organic

compounds or micro-plastic) is negligible (at least using these optical determinations) compared with the natural soil-released DOM.

Table 5-2. Student t-test results for the comparison between upstream and downstream samples in the Milwaukee River. H0 is a null hypothesis which assumes that there is no effective difference for parameters between upstream and downstream.

Parameters	Reject H_0	*p* value	Parameters	Reject H_0	*p* value
Temperature	1	0.0033	FIX	0	0.31
Conductivity (μS cm^{-1})	0	0.31	BIX	0	0.27
pH	0	0.6	HIX	0	0.05
Chlorophyll *a* (μg L^{-1})	0	0.077	C_{500} (QSE)	0	0.18
a_{254} (m^{-1})	1	0.015	C_{440} (QSE)	0	0.39
Colloid (%, a_{254})	0	0.69	C_{430} (QSE)	1	0.0006
$S_{275-295}$ (nm^{-1})	0	0.5	C_{390} (QSE)	1	0.001
TDN (mg L^{-1})	0	0.21	C_{330} (QSE)	0	0.72
DOC (mg L^{-1})	1	0.018	$\delta^{13}C$ (‰)	0	0.6
DIC (mg L^{-1})	0	0.82	$\delta^{15}N$ (‰)	0	0.14
Colloid (%, DOC)	0	0.24	PN (μg L^{-1})	1	0.019
$SUVA_{254}$ (L mgC^{-1} m^{-1})	0	0.14	POC (μg L^{-1})	1	0.028

5. 5. 3 Influence of Estabrook dam removal on DOM in the Milwaukee River

Comparisons of CDOM properties and fluorescence indices were made between measurements from 2018-2019 and measurements from 2014-2015, and the results are shown in Figure 5-15. DOC and CDOM a_{254} varied within a similar range, and the variation of them are very similar. Statistical tests showed that there is no difference between two studies (t-tests, p>0.1). In terms of $SUVA_{254}$ and $S_{275-295}$, the range of $SUVA_{254}$ and $S_{275-295}$ are smaller in 2017-2020 than in 2014-2015, but there is still no

difference between two periods (t-test, p>0.1). Speaking of the fluorescence indices, the values of BIX in 2017-2020 have same range with those measured in 2014-2015. HIX is increasing from 2017 to 2020. Compared with the previous observations, the HIX values in 2017-2020 are significantly higher than those measured in 2014-2015 (t-test, p<0.005).

Why did HIX keep increasing since 2018? Is that the subsequent effects from the removal of Estabrook Dam? The dam removal might have an impact on the hydrological settings of the Milwaukee River, and thus resulted in the increasement of HIX. Previous studies reported that protein-like and microbial derived humic-like DOM was relatively more enriched for the impoundment-affected waters than those upstream pristine waters, where terrestrial humic-like DOM was predominant (Hur et al. 2014; Wang et al. 2019). The increase of water residence time resulted in water eutrophication (Hur et al. 2007; Park et al. 2009), with more autochthonous DOM inputs, leading to a BIX increasement and HIX decline. These pieces of evidence explain the HIX increasing in 2017-2020. However, if the dam removal has positive effects for HIX, we should have observed that HIX had a surge when the dam was totally removed and kept at the same level, but not a gradual increasement from 2017 to 2020.

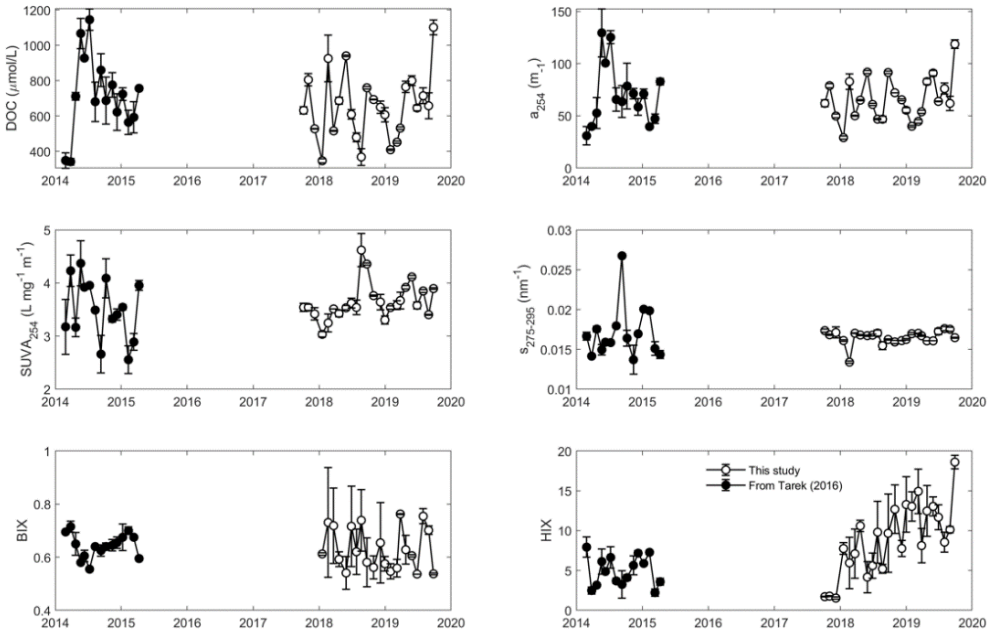

Figure 5-15. Comparison between measurements from this study and the measurements from Tarek (2016), including DOC, CDOM a_{254}, $SUVA_{254}$, $S_{275-295}$, BIX, and HIX.

Another alternative explanation might be the boost of river discharge. Due to the increase of precipitation (Figure 5-16), the annual river discharge in 2018-2019 is significantly higher than the discharge in 2013-2014 (Figure 5-16). Large precipitation would increase the river discharge, and flush more soil-derived DOM into river, which lead to the gradual HIX climbing (Figure 5-15).

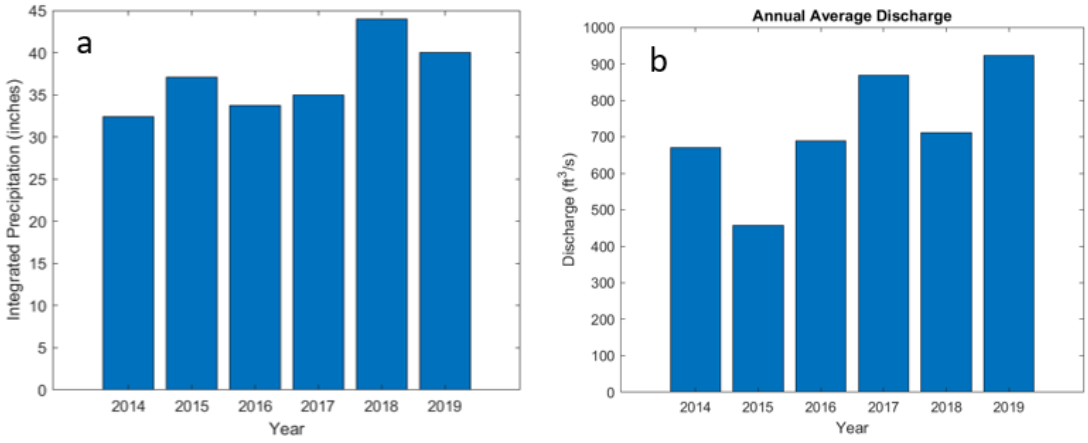

Figure 5-16. (a) Annual integrated precipitation. (b) Annual average discharge of the Milwaukee River.

5. 5. 4 Use $S_{275\text{-}295}$ as a MW indicator

The spectral slope $S_{275\text{-}295}$ has been recognized as indicators for lignin (a terrestrial-derived biomarker (Opsahl and Benner 1997), average molecular weight (MW) of DOM (Helms et al. 2008), and terrestrial DOM fraction (Fichot and Benner 2012) in aquatic environments. Here the reliability of $S_{275\text{-}295}$ as a MW indicator was evaluated for the Milwaukee River. Although the average MW of DOM is absent, the colloidal fraction (%) is capable to indicate the MW of DOM, since the higher MW, the larger colloidal fraction (DOM larger than 1 kDa). As shown in Figure 5-17 (left panel), linear relationship between colloidal fraction (%) and $S_{275\text{-}295}$ are ambiguous, which probably due to different DOM sources between wet season and dry season. Once the samples that water temperature lower than 15 °C are excluded, a clear linear relationship between $S_{275\text{-}295}$ and colloidal fraction (%) is shown (Figure 5-17 right panel, R^2=0.54, n = 22, p<0.01), which implies that $S_{275\text{-}295}$ can be used as a MW indicator where terrestrial DOM was predominant in the aquatic environments, just like where the CDOM a_{254} can be used as a DOC indicator.

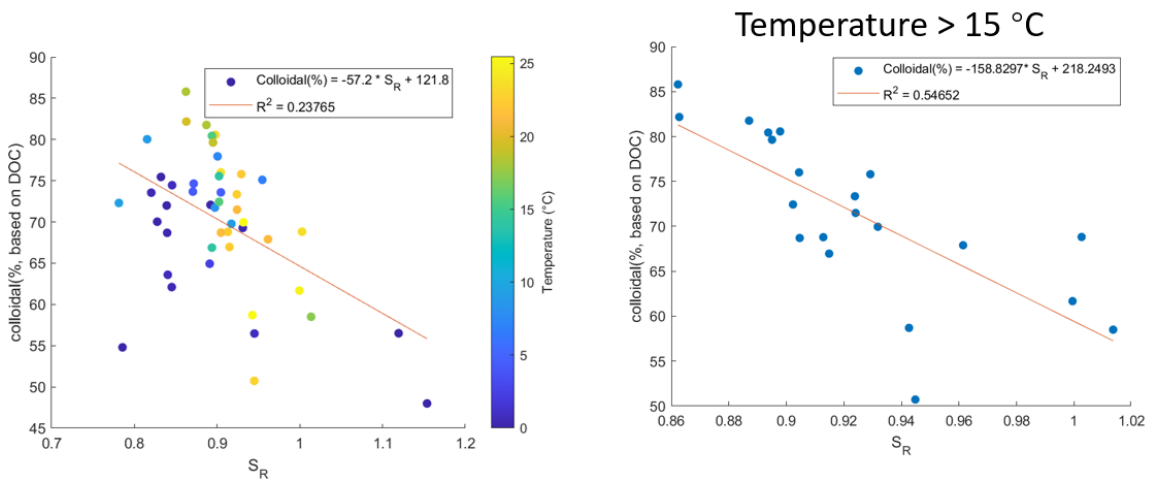

Figure 5-17. Relationship between S_R and colloidal fraction (%, based on DOC concentration, left panel). The color of dots indicates the water temperature at sampling points. Right panel shows the relationship between S_R and colloidal fraction percent when water temperature is higher than 15 ˚C.

Chapter 6. Dynamic cycling of organic matter under *Microcystis* blooms in a shallow eutrophic lagoon

6. 1. Abstract

To better understand the influence of cyanobacteria blooms on the cycling of dissolved organic matter (DOM) and particulate organic matter (POM), I conducted a time-series field study in the hyper-eutrophic Veterans Park Lagoon with seasonal cyanobacterial blooms, located in the city of Milwaukee, Wisconsin, where *Microcystis*, a genus of cyanobacteria, is predominant during summer. Weekly or bi-weekly samples were collected to measure optical characteristics of DOM and stable isotope composition (δ^{13}C-POC and δ^{15}N-PN) of POM from March 2019 to January 2020. Dissolved organic carbon concentration (DOC), highly correlated with chromophoric DOM absorption coefficient (a_{254}), increased gradually in the summer, and increased rapidly after the collapse of cyanobacteria blooms. The DOM spike after bloom collapse was of high molecular size and low aromaticity, indicated by spectral slope ($S_{275\text{-}295}$) and Specific UV absorbance ($SUVA_{254}$). Using parallel factor analysis, fluorescent DOM were decomposed into four components, including three terrestrial humic-like components and one autochthonous protein-like component. Three humic-like components were strongly related to the temperature and precipitation, while the protein-like component showed a spike corresponding to the boosted DOC and a_{254}. The δ^{13}C-POC ranged from -28.18 to -18.23 ‰, while δ^{15}N-PN ranged from -10.62 to 5.63 ‰, both of which are consistent with an autochthonous origin from cyanobacteria *Microcystis* blooms. The δ^{13}C enrichment in POC when bloom occurred was attributed to reduced isotope fractionation due to carbon (CO_2) limitation. Using stirred-cell ultrafiltration, three size fractions of DOM, i.e. < 1 kDa, 1-10 kDa, and > 10 kDa were collected. Small fractions (<1 kDa and 1-10 kDa) are single-peak seasonal variations with the maxima in July and May, while large fraction (> 10 kDa) is a bimodal distribution with two maxima in May and November. In terms of seasonal variations of four fluorescent DOM components, three humic-like components are largely distributed among < 1kDa and 1-10 kDa size fractions from March to September, and protein-like components are mostly in > 10 kDa fraction. But from November to January, both humic-like and protein-like fluorescent components distributed evenly in the three size fractions.

6. 2. Introduction

Shallow lagoons and bays are important features of marine or lacustrine coastal areas (Gobler and Sañudo-Wilhelmy 2003; Minor et al. 2006) because of high primary production and high organic carbon concentration, especially in summers. The fate of the production and the organic pool in coastal waters (e.g. comsumption by microbial communities and burial in sediments), especially when harmful algal blooms occurr, is of great interest in terms of its impact on global carbon budget (Hedges et al. 1997). Currently, our understanding of the source and sink of the organic reservior during algal bloom progression has been partially hampered by the poor characterization of their size and chemical composition.

Veterans Park Lagoon (VL), where cyanobacteria *Microcystis* blooms were observed in the summertime of recent decades, is a small-area lagoon in the coast area of Lake Michigan (Juneau Park investigation report, 1980s). *Microcystis*, a genus of cyanobacteria, is capable to form summer algal blooms in eutrophic freshwater ecosystems, such as Lake Taihu in China (De Kluijver et al. 2012), and Lake Biwa in Japan (Kumagai et al. 2000; Ozawa et al. 2003). *Microcystis* releases neurotoxins, such as microcystin, which cause water quality deterioration and have attracted increasing attention. Meanwhile, the unpleasant odor after collapse of blooms makes the cyanobacteria one of the most major focus targets among the harmful bloom-forming algae. There were many studies focusing on the ecological role, physical characteristics, and health risks of the toxins (Shen et al. 2011; De Kluijver et al. 2012; Patriarca et al. 2020). The inter-relationship between DOM dynamics and cyanobacterial blooms remains poorly quantified. Therefore, the focus of this study is on variations in carbon speciation and DOM quantity and quality and their changes during a whole bloom cycle in the VL.

In this study, we investigated the particulate organic matter and dissolved organic matter in the VL. The objective of this study is to understand the seasonal variation of POM and DOM along a progression of *Microcystis* bloom using stable isotopes and optical measurements. Using ultrafiltration with alternative membranes, DOM can be separated into several size fractions, e.g. < 1 kDa, 1-10 kDa, and > 10 kDa. Biogeochemical behaviors of DOM are investigated with these different size DOM.

6. 3. Methods

6. 3. 1 Study site and sample collection

The VL is a small (14.7 acres), shallow (averagely 1.5 m), well-mixed enclosed lagoon in Veterans County Park (Figure 6-1), close to downtown metropolitan Milwaukee, Wisconsin (Juneau Park investigation report, 1980s). Absence of surface runoff and point-source anthropogenic inputs make this lagoon a simple ecosystem, where 68.2% of water input is from surface precipitation. The hydraulic time is nearly one-year (0.9 yr), which indicates the lagoon as an ideal natural mesocosm for documenting the cycling of DOM during phytoplankton blooms (Juneau Park investigation report, 1980s).

We selected two sampling points (43.049837 °N, 87.890379 °W; 43.044197 °N, 87.894670 °W, Figure 6-1) in the lagoon. Surface water samples were collected weekly or biweekly between March 11th, 2019 and Jan 10th, 2020, using a lab-made water sampler and one-liter pre-rinsed HDPE bottles. Water samples were immediately filtered using pre-combusted 0.7 μm GF/F filters (Whatman) after shipping back to lab. The GF/F filters were stored frozen before further measurements.

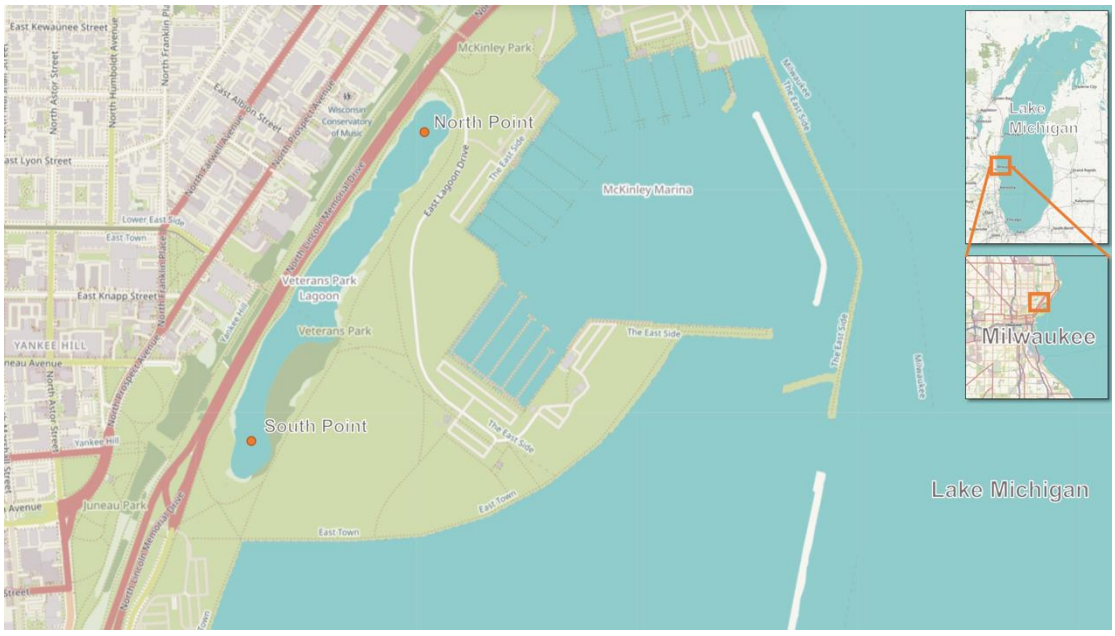

Figure 6-1. Map of the lagoon in the Veterans Park in Milwaukee, Wisconsin, and two sampling point (north and south). The map was derived from Leaflet (1.6.0) 'Thunderforest' map provider.

6. 3. 2 Ultrafiltration for DOM size fractionation

Due to the shortage of time and efforts for conducting ultrafiltration, sampling resolution for ultrafiltration cannot be as high as those bulk samples. Samples were selected every other month as representatives to determine the size-dependent DOM properties. Although small events might be missed due to these sampling timing, seasonal variations in this lagoon can still be recorded. An aliquot of the filtrate (180 ml) was size-fractionated into three size-fractions, including < 1 kDa, 1-10 kDa, and 10 kDa - 0.7 μm, using a 180-ml stirred cell ultrafiltration system (Amicon 8200) with 1 and 10 kDa membrane (regenerated cellulose, Millipore YM1, 63.5-mm diameter). Details of the procedures are described in Xu and Guo (2017).

6. 3. 3 Sample analysis

Dissolved organic carbon concentration (DOC) and dissolved inorganic carbon concentration (DIC) were determined using a high-temperature catalytic oxidation method with a Shimadzu TOC-L Total Organic Carbon (TOC) analyzer (Guo et al. 1994). Two aliquots of one water sample were determined; one was acidified to pH 2 to remove DIC in the water, the other was not acidified to have total dissolved carbon (TDC = DOC + DIC). DIC was determined by substraction DOC from TDC. Total dissolved nitrogen concentrations (TDN) were determined simultaneously with the DOC using a TN module coupling on the TOC analyzer. The GF/F filter samples for particulate organic carbon concentration (POC) and particulate nitrogen concentration (PN) were acid-fumigated for 12 h and dried at 60 ℃ for 24 h before analysis on an elemental analyzer (ECS 4010, Costech instruments). Stable isotopes $\delta^{13}C$ and $\delta^{15}N$ were measured using an isotope ratio mass spectrometry (Delta V Plus) on tandem with the elemental analyzer. Chlorophyll *a* concentration (Chl *a*) was analyzed using standard spectrophotometric method. (Strickland and Parsons 1972). Chromophoric dissolved organic matter (CDOM) was measured using a spectrophotometer (Agilent 8453). CDOM absorption coefficient at 254 nm (a_{254}) is calculated using the following equation: $a_{254} = A_{254}/l$, where A_{254} is the absorbance at 254 nm and l is the cuvette length. Spectral slope between 275 nm and 295 nm ($S_{275\text{-}295}$) is determined using a non-linear fit for the equation: $S_{275\text{-}295} = a_{280}\, e^{S(\lambda - 280)}$. $SUVA_{254}$ is calculated using the following equation: $SUVA_{254} = A_{254}/DOC$. Fluorescence excitation emission matrix (EEM) and parallel factor analysis (PARAFAC) has been

described in previous studies. Briefly, fluorescence EEMs were determined using a Fluoromax-4 spectrofluorometer (Horiba) within excitation 250-480 nm and emission 240-600 nm. Humification index (HIX), a fluorescent index related to the degree of humification, is a ratio of fluorescence signals at the emission range of 435-480 nm to those at the range of 300-345 nm with excitation at 254 nm (Zsolnay et al. 1999). Biological index (BIX) is a ratio of fluorescence intensity at the emission 380 nm to 430 nm within the excitation wavelength of 310 nm (Huguet et al., 2009; Zhou et al., 2016). ΔEEM can be calculated using difference between two size fractionated EEMs. For example, the ΔEEM for 1-10 kDa is the difference between EEM <1 kDa and EEM < 10 kDa. PARAFAC was performed on all EEMs data, including bulk water samples, size-fractionated samples, and ΔEEM, using drEEM toolbox (v.0.5.0) and Matlab (R2020a) software. Components were compared with other literatures online using OpenFluor (Murphy et al. 2014). Tucker Congruence Coefficient (TCC) was used as comparison criteria, which were set to 0.95 for excitation and emission spectra.

6. 4. Results and discussion

6. 4. 1 Variations in temperature, precipitation, and Chl *a*

Based on the available water temperature data from *in situ* YSI sensors, a complete seasonal variation in water temperature was constructed using the correlation between water temperature and air temperature. According to the historical records, daily- and weekly- average air temperature is strongly related to the *in situ* measured water temperature (Figure 6-2a). Thus, we predict those unmeasured water temperatures using daily or weekly average air temperature with a least square linear regression algorithm. Both daily and weekly average air temperatures are significantly correlated with the water temperature, but the daily temperature predict better (higher R^2 and lower percentage of mean average error, Figure 6-3). Therefore, here we reported the predicted *in situ* water temperature using daily average air temperature.

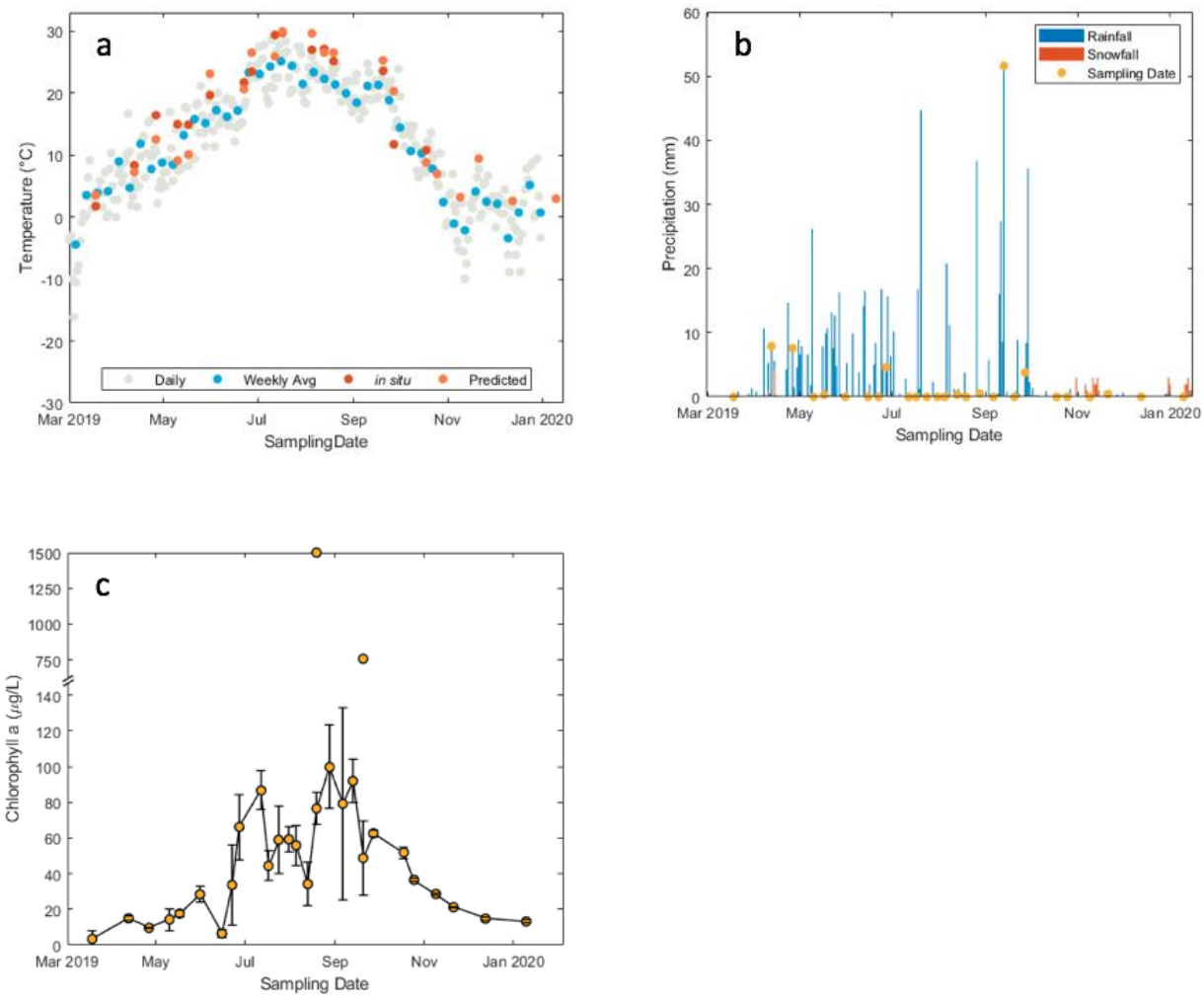

Figure 6-2. (a) Daily air temperature (gray dots), weekly averaged air temperature (blue dots), *in situ* measured water temperature (red dots) and predicted water temperature (orange dots). (b) Precipitation, including rainfall (blue bars) and snowfall (red bars) (c) Seasonal variation of Chlorophyll *a* concentration.

In term of small lake ecosystem, water temperature is largely influenced by ambient air temperature (Figure 6-2a). Precipitation severely impact lacustrine ecosystems, especially small and shallow lagoons (Pace and Cole 2002; Andersson et al. 2005; Sobek et al. 2007). Based on temperature and precipitation in the Milwaukee, two seasons in one year are defined: a wet season and a dry season. Generally, wet season in the Milwaukee lasts 6.4 months, basically from March 29th to October 11th (Spark 2017; Reichle et al. 2017). The observation is majorly in the wet season (Figure 6-2b). Small snowfall events are observed between November 2019 and January 2020. The last sampling date is right before the major snowfall event in the beginning of 2020. For Chl *a* concentration (Figure 6-2c), there are two outliers (Aug. 19 South Surface and Sept. 20 South Surface) corresponding to cyanobacteria blooms (scums), which can be seen on live photos taken *in situ* (provided by Wilson). The cyanobacteria genus was

identified as *Microcystis* by microscopy (conducted by Wilson). The outliers were both surface samples at the south sampling point, resulted from the southbound wind in summer accumulated surface algae to the south part of lagoon. Other than that, Chl *a* was higher in the wet season than in the dry season due to stimulated primary production driven by elevated temperature and flushed nutrients in summer. There were several peaks in the wet season, including July 12–24 and Aug 28–Sep 13, associated with algal blooms dominated by diatoms/green algae.

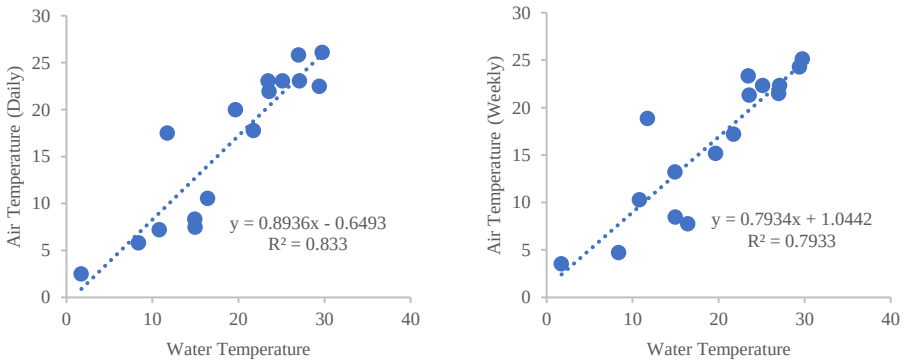

Figure 6-3. Comparison between different linear regression models. Left panel presents the linear regression between water temperature and daily air temperature. Right panel presents the linear regression between water temperature and weekly average air temperature.

6. 4. 2 Seasonal variation of organic matter concentration

The VL had a strong seasonality of CDOM and DOC concentrations, varying nearly 4-fold for both of them. CDOM a_{254} ranged from 9.6 to 44.0 m^{-1}, and DOC ranged from 171.9 to 714.5 μM. CDOM a_{254} and DOC both increased monotonically between March and May 2019 (Figure 6-4a and b). They remained at high level stably from June to November except the large spike surging in the deep November. After the spike, CDOM-a_{254} and DOC decreased dramatically to winter baseline level. Except the spike in the November, the DOC seasonal trending was highly consistent with temperature; two little peaks on July 17 and September 6 were corresponding to the highest temperature in the downtown Milwaukee area. DOC concentration and CDOM a_{254} in the lagoon are not significantly linear correlated (Figure 6-5). The scatter points in the summer are more converged than those in the cold seasons, which was probably due to the dominant terrigenous DOM input in the summertime. TDN concentrations were generally invariant (average: 29.7±6.5 μM), except the spike in the November which was 59.6 ± 5.4 μM.

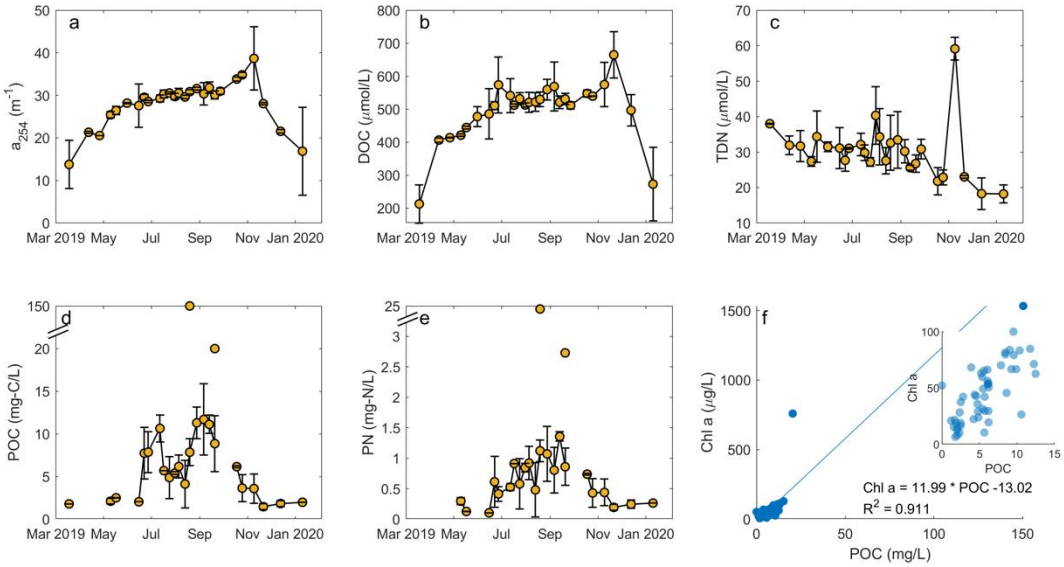

Figure 6-4. Seasonal variation of (a) chromophoric dissolved organic matter (CDOM) a_{254} (b) dissolve organic carbon (DOC) concentration (c) total dissolved nitrogen (TDN) (d) particulate organic carbon (POC) concentration (e) particulate nitrogen concentration, and (f) linear correlation between POC and Chl *a* concentration.

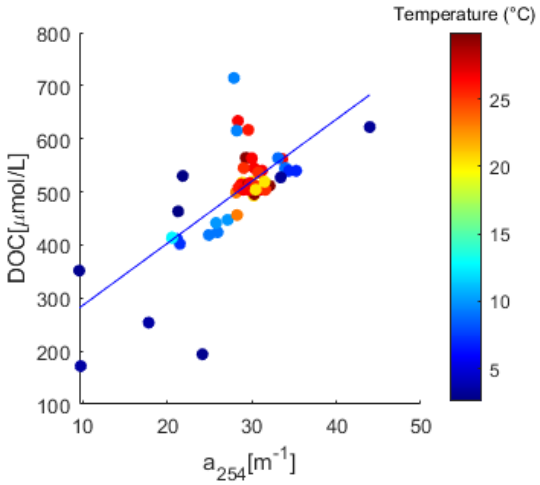

Figure 6-5. Relation between CDOM a_{254} and DOC concentration. The color of dots indicate water temperature of that sample, implying the sampling season.

The particulate organic carbon (POC) ranged from 1.2 to 150.3 mg-C/L, and particulate nitrogen (PN) ranged from 0.1 to 22.6 mg-N/L. Two POC and PN outliers are corresponding to the Chl *a* outliers, and also corresponding to the algal blooms in the lagoon. Besides, there were two peaks over the one-year POC and one peak over the PN seasonal variations (Figure 6-4d and 2e). The first POC peak was shown at the beginning of July 2019, and the second POC peak was shown in September 2019, which is concurrent with the PN peak. They were both indicated the algal bloom in the lagoon this year, since POC

was significantly correlated with Chl *a* concentration (n = 20, R^2 = 0.91, p<0.05, Figure 6-4f). The second POC and PN concentration peaks were relatively higher than the first one (nearly 2-fold for POC and 4-fold for PN), which indicated that second algae bloom was much more intense than the first one. The difference between POC and PN peaks implied the different species of the algal blooms.

6. 4. 3 Seasonal variations of DOM optical properties

$SUVA_{254}$, an indicator of DOM aromaticity, ranged from 1.41 to 2.56 L m^{-1} mg^{-1}. It was stably increasing, except the two low values in the April and November 2019 (Figure 6-6a). The first low value, indicating relatively weak aromaticity, implied ice melted water running into the ecosystem. The second low value indicated that the high DOC and CDOM a_{254} in the late fall was associated with the degraded fresh DOM or detritus from phytoplankton. $S_{275\text{-}295}$, a DOM molecular size indicator, ranged from 17.2 to 29.4 nm^{-1}. $S_{275\text{-}295}$ was relatively low in the wet season, and high in the dry season (Figure 6-6b). The $SUVA_{254}$ and $S_{275\text{-}295}$ indicated that 1) in the spring (Mar-May 2019) low molecular weight and low aromatic DOM is dominated; 2) the outbreak of DOM in the November was of high molecular size with low aromaticity. BIX ranged from 17.2 to 29.4. It was higher in the wet season (avg 0.85±0.03) and low in the dry season (avg 0.79± 0.04, Figure 6-6c). HIX seasonal variation, ranged from 0.69 to 0.94, showed opposite trend to the BIX (Figure 6-6d). HIX is relatively stable except the low value in the November.

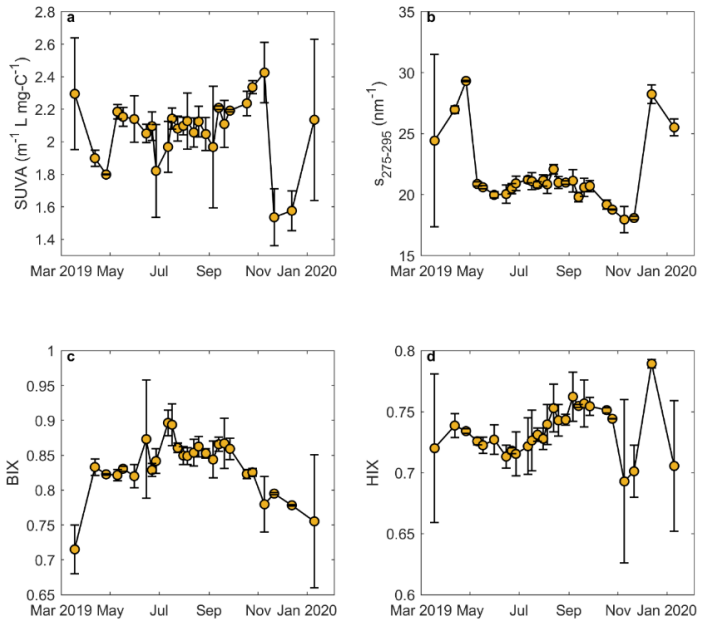

Figure 6-6. Seasonal variation of (a) spectral slope ($s_{275\text{-}295}$), (b) $SUVA_{254}$, (c) BIX, and (d) HIX.

6. 4. 4 PARAFAC-derived fluorescent DOM components

Three humic-like components and one protein-like component were identified and validated from the EEM dataset using PARAFAC (Figure 6-7a, Table 6-1). Fluorescence intensities of C1~C3 varied seasonally which were strongly congruent with temperature (Figure 6-7b); they were relatively higher in the wet season than in dry season due to the high terrestrial input in wet season. C1, C2 and C3 were highly correlated with each other (R^2 values > 0.70, p values < 0.01). Other than three humic-like components, the autochthonous C4 had a spike in the November 2019. This spike also existed in the CDOM a_{254}, DOC concentration, TDN concentration, $SUVA_{254}$, $S_{275\text{-}295}$, and HIX. In Figure 6-8, the ratio of C4 to C1 (Protein-like to Humic-like) is showed, which largely eliminates impacts from terrigenous DOM. The ratio ranged from 0.35 to 0.74, except for the November spike, which is 0.92. It decreased slowly from March to September, then increased rapidly. After the November spike, it decreased again rapidly.

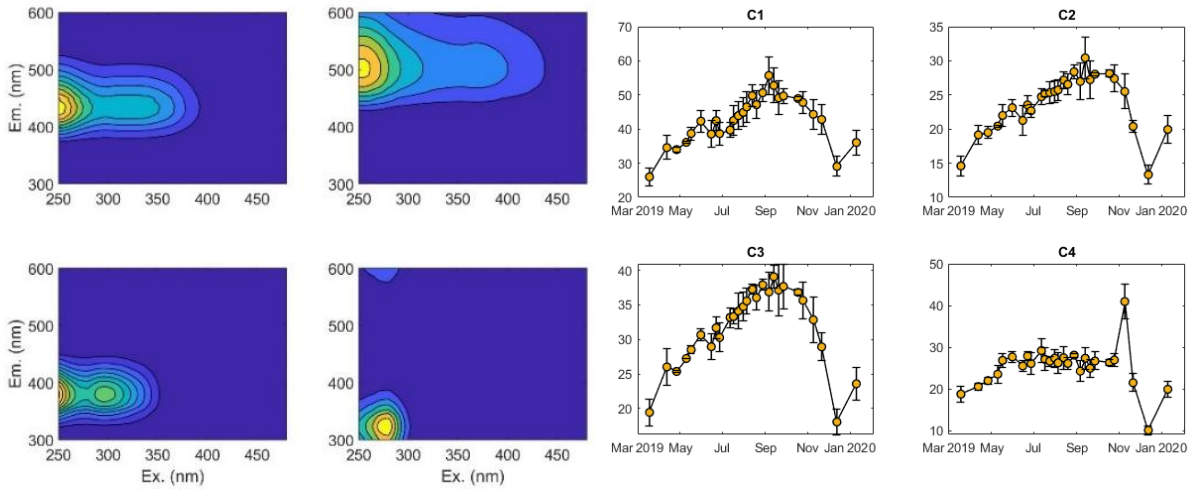

Figure 6-7. Left panel: Contours for the four PARAFAC-derived components. Right panel: Seasonal variations in the four fluorescent components derived from a PARAFAC model.

Using the OpenFluor online comparisons, we found that the four PARAFAC-derived fluorescent components are highly similar (TCC $_{Ex,Em}$ >0.95) as components observed in freshwater ecosystems worldwide (Murphy et al. 2014). For instance, in the Zambezi River, west Greenland lakes and Lake Lillsjön in Sweden, humic-like fluorescent components similar as C1 are found (Lambert et al. 2016; Wünsch et al. 2017a; Osburn et al. 2017). As for the protein-like C4, there are highly similar components observed in Antarctic freshwaters and streams in Mediterranean areas (Harjung et al. 2018; Kida et al.

2019). In summary, there are 30, 47, 9, and 23 online published components highly similar as C1, C2, C3 and C4, respectively, which implies that the fluorescent DOM in the lagoon is ubiquitous.

Table 6-1. Parameters of all PARAFAC-derived DOM components.

Component	Ex. (nm)	Em. (nm)	Properties	Peaks from Coble (2007)
C1	250/310	425	Humic-like	Peak B and Peak T
C2	255	490	Humic-like	Peak A and Peak C
C3	<250/290	375	Humic-like	Peak M
C4	280	320	Protein-like	Peak B and Peak T

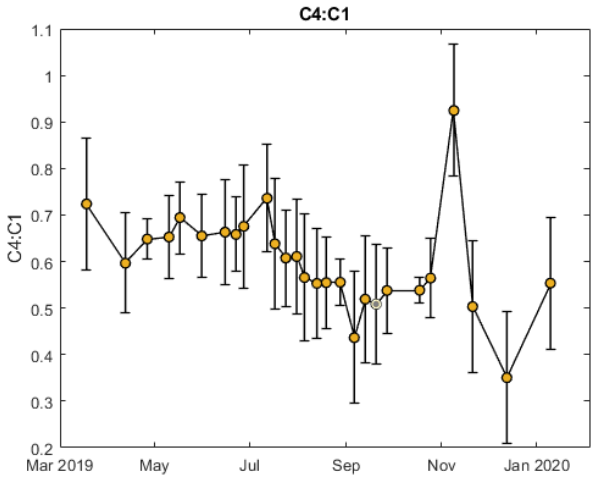

Figure 6-8. Seasonal variations in the F_{max} ratio of Component 4 to Component 1.

6. 4. 5 Variations in stable isotope compositions of particulate organic matter

Stable carbon isotope ($\delta^{13}C$) values in POC ($\delta^{13}C$-POC) ranged from -28.18 to -18.23 ‰ (Figure 6-8a). It vibrated in a range between -27.82 and -23.04 ‰ in the spring (Mar 18-May 31, 2019), gradually became heavier in the summer and the early fall (from -23.84±0.40 to -18.61±0.5 ‰, June 15-Sep 20, 2019) when blooms and POC were highest, then decreased sharply in the late fall and winter (to -28.18±0.2‰, Jan 10, 2020) under low biomass but high DOM. Noted that only samples from September 20 contained $\delta^{13}C$-POC heavier than -20 ‰. Previous studies reported that during *Microcystis* blooms in freshwater ecosystems, $\delta^{13}C$-POC in the surface water are significantly heavier than other months (Table 2, Kumagai et al., 2000; Maki et al., 2010; Ozawa et al., 2003). This isotopic enrichment could be explained by the

limitation of carbon caused by the high-density algal bloom (De Kluijver et al. 2012). However, there was a few studies reporting δ^{13}C-POC during a *Microcystis* bloom heavier than -20 ‰ (Xu et al. 2007). The duplicated δ^{13}C-POC measurements were double-checked to increase the confidence to these two data points. This abnormally heavy δ^{13}C-POC signal can be explained by the extremely high density of the *Microcystis* bloom. The Chl *a* concentration in the lagoon was much higher than (approximately 10-fold) any other previously reported (Table 2). The carbon (carbonate) limitation due to the high-density of algal bloom would enrich the δ^{13}C-POC and decline isotope fractionation in photosynthesis in *Microcystis* community, which is discussed in next session 1. 1. 1.

δ^{15}N values in PN ranged from -3.06 to 5.63 ‰, vibrating between 0 ‰ to 4 ‰ (Figure 6-8b). It was probably due to the exist of nitrogen-fixation cyanobacteria cell in this ecosystem (Gu et al. 2006). POC:PN mole ratios (C/N ratio) in the surface exhibited a two-stage variation: they were mostly larger than 20.0 in the spring and summer (May 17 to Sep 6, 2019), then the ratios decreased to 7.8±2.3 from Sep 13 2019 to Jan 10 2020 (Figure 6-8c). The high stage of C/N ratio in the summer probably due to the terrestrial inputs, such as leaf litter (up to 20) and soil organic matter (up to 25, Michener & Lajtha, 2008). It is worth noted that the C/N ratio of the two bloom events samples (South Surface sample on Aug 19 and Sep 20, 2019) were 6.65 and 7.45, which were very closed to Redfield ratio of 6.6. The C/N ratio in the bottom ranged from 7.27 to 12.33. It kept increasing from July to September. The C/N ratio in the bottom were comparatively closed to Redfield ratio than those in the surface.

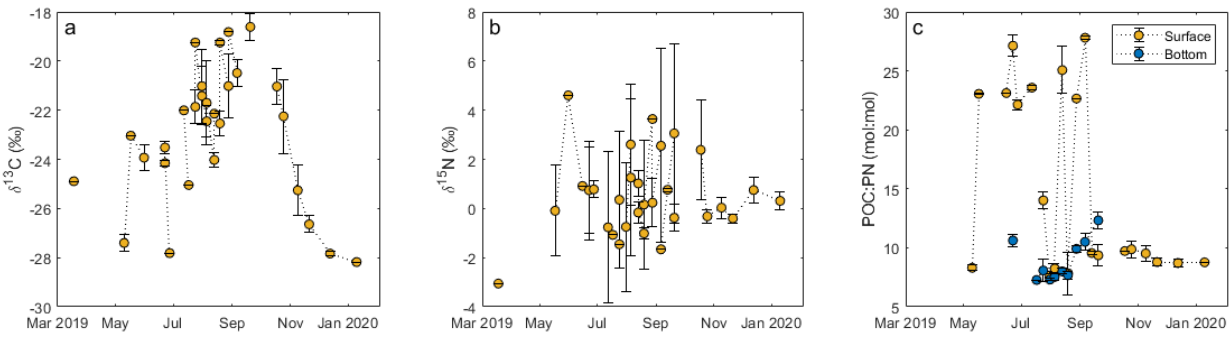

Figure 6-9. Seasonal variation of stable isotope (a) δ^{13}C, (b) δ^{15}N of particulate organic matter, and (c) POC:PN molar ratio.

Table 6-2. Lakes influenced by cyanobacteria *Microcystis*, and comparisons in $\delta^{13}C$ -POC.

Lakes	$\delta^{13}C$ -POC Range (mean±SD)	Chl *a* (μg L^{-1})	Reference
Veterans Lagoon, WI, US.	-28.18~-18.23 ‰ (-22.9 ± 2.7 ‰)	0.3~1532;	This study
Lake Biwa, Japan	-29.6 ~ -20.4 ‰ (-22.7 ± 0.7 ‰)	3.1~9.6	(Kumagai et al. 2000; Ozawa et al. 2003; Maki et al. 2010)
Lake Taihu, China	Surface bloom (scum): -20.9 ± 0.4 ‰	-*	(De Kluijver et al. 2012)
Lake Kirkkojarvi, Finland	*Microcystis botrys*: -29.2~-18.5 ‰ (-23.0 ± 2.5 ‰) *Microcystis* spp.: -32.1~-20.7 ‰ (-24.3 ± 4.1 ‰)	128 ± 21	(VUORIO et al. 2006)
Venice Cut, San Joaquin River, CA	-29 ~ -20 ‰	0~4.0	(Lehman et al. 2014)
Lake Chaohu, China	-17.6~-18.2 ‰ (-17.9±0.4 ‰)	-	(Xu et al. 2007)
Lake Xingyun, China	-19.6~-20.9 ‰ (-20.4±0.6 ‰)	-	(Xu et al. 2007)

*: data not available.

6. 4. 6 Bloom progression from the perspective of carbon

As Figure 6-10 shown, POC, as mentioned above, had two spikes: one at the beginning of July, the other at the middle of September. They are both corresponded to the *Microcystis*-induced Chl *a* spikes with heavier δ^{13}C-POC values and relatively less DIC (Figure 6-10). The DIC was decreasing gradually from May to July (from ~30 mgC L^{-1} to ~ 25 mgC L^{-1}), probably due the rising temperature (from ~10 ˚C to 25 ˚C) and growing phytoplankton communities. It kept consistent in July and August when temperature is around highest, which indicated the lowest CO_2 and carbonate dissolution. It decreased again to less than 20 mgC L^{-1} when *Microcystis* bloom occurred in September, which implied that high-density cyanobacteria community resulted in removal of DIC, and further less δ^{13}C-POC isotope fractionation. This relationship between carbon species were commonly observed in coastal bays and lakes (Gu et al. 2006; Maki et al. 2010).

DOC spike had a nearly two-month lag after the POC peaks in the summer, which indicates it took nearly 60 days for DOM to responds to the algal bloom in the summer. The DOC released by the cyanobacteria *Microcystis* was of high lability (Patriarca et al. 2020), indicating that it was consumed quickly by heterotrophic bacteria in this ecosystem, herein there were no obvious DOC spikes appearing with the POC spikes. DOC kept increasing slowly in the whole wet season until the late fall. The spike of DOC implied a boost source and a shrink sink to DOM pool. The boost in DOM source was probably due to DOM releasing from phytoplankton induced by zooplankton grazing. Previous studies revealed that DOM is one of the essential elements in the energy flow: phytoplankton → DOM → zooplankton (De Kluijver et al. 2012). Summer algal bloom would result in the increasing DOM source. On the other hand, the decreased sink can be explained by reduce of *Microcystis*-associated heterotrophic bacteria with collapse of *Microcystis* blooms (Shen et al. 2011). Natural heterotrophic bacterial communities around the cyanobacteria provide a quick pathway to consume and transfer DOM molecules. When the duration and intensity of solar radiation was unable to sustain growth of the cyanobacteria, the mutually beneficial colonies formed by *Microcystis* and its associated bacterial cells would collapse, leading to a shrink sink of DOM in the late fall.

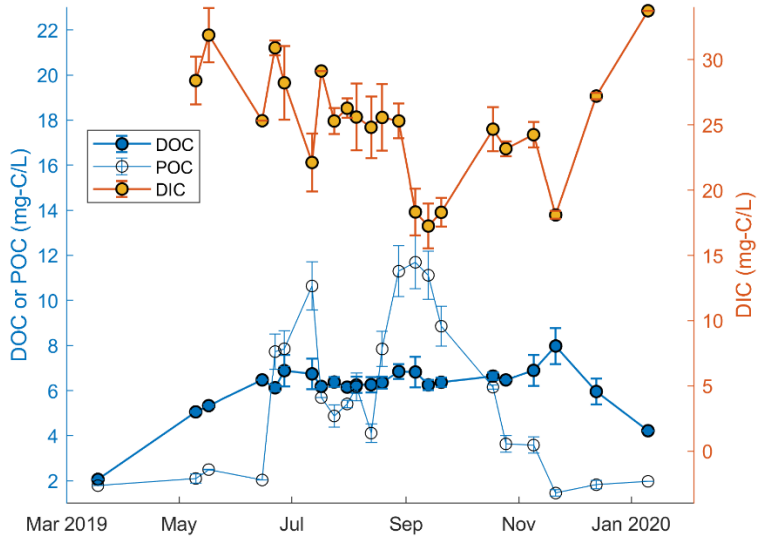

Figure 6-10. Seasonal variations in POC, DOC and DIC in the lagoon in the Veterans Park, Milwaukee, WI. Two extremely high POC outliers were excluded in this chart.

6. 4. 7 EEM spectra of size-fractionated DOM

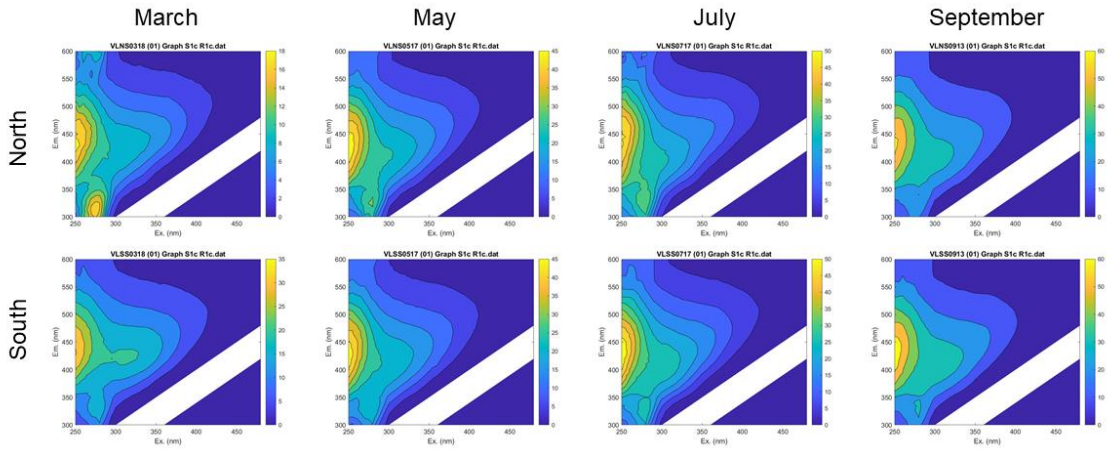

Figure 6-11. EEM contours of selected samples (bulk water, no size fractionated) in March, May, July and September 2019.

Surface water samples, collected on March 18, May 10, July 12, and September 6, 2019 at the north point, were selected as examples to characterize the size-fractionated DOM. Each sample was ultrafiltered into three size fractions (< 1kDa, 1-10 kDa, 10 kDa-0.7μm). In terms of the CDOM a_{254}, intermediate size fractions (1-10 kDa) have largest contributions to bulk CDOM a_{254}, followed by < 1 kDa, and 10 kDa-0.7μm size fractions. Two small fractions (< 1 kDa and 1-10 kDa) are single-peak

distribution (although their maxima are in different month), while the > 10 kDa fraction is a bimodal distribution with two maxima in May and November. In light of the fluorescence signals, EEM contours of bulk samples are shown in the Figure 6-11, and the averaged EEM contours for these three size fractions and bulk sample are presented in Figure 6-12. While there are several tiny differences in the bulk water EEM contours (Figure 6-11), including the total fluorescence intensity, Peak T (ex/em = 275/340 nm) intensity, shape of Peak A (ex/em = 250/400-500 nm), etc., the bulk sample EEMs contours are generally highly similar. However, the size fractional EEMs clearly show a clear distinction (Figure 6-12 bottom panel), which reflects the chemical composition difference between them. This implies that the fluorophores in DOM is highly size-dependent. Humic-like fluorophores are majorly presented in < 1 kDa and 1-10 kDa size fractions, while protein-like fluorophore is mostly observed in 10 kDa–0.7 μm fraction. Seasonal variations of four PARAFAC-derived fluorescent components are size-dependent as well (Figure 6-13). All components in < 1 kDa fraction are increasing from March to July 2019, then decrease from July 2019 to January 2020. For 1-10 kDa size fraction, C1 and C2 are more abundant than C3 and C4. The 10 kDa-0.7μm size fraction is intriguing. From March to September, protein-like C4 contributes much more than other humic-like components, while in November and January humic-like components are dominant in the > 10 kDa fraction.

It is unsurprising to find size-dependent properties for the fluorescent DOM (Figure 6-13) because of the heterogeneity of natural DOM. Recently, Lin and Guo (2020) report a humic-to-protein trend on freshwater DOM using a continuous EEM measurement, which is consistent with this observation. Lin et al. (2016) also report the connection between terrestrial sources and smaller size fraction of DOM in coastal Arctic region. In addition, Zhou and Guo (2015) presented EEMs of three integrated size fractions of colloidal organic matter from soil leachate (0.5-4 nm, 4-8 nm and > 55 nm). They demonstrated that humic-like fluorophores (peak A and peak C) are dominated in the smallest size fraction (0.5-4 nm), and protein-like fluorophores (peak T and peak B) are the major fluorophores in the largest size fraction (> 55 nm). Humic-like components mingled with protein-like components in the intermediate size fraction (4-8 nm). Here we assume that >55 nm size fraction can be regarded as

the >8 nm size fraction, since the fluorescence signals for both humic-like and protein-like fluorophores ($Fluor_{275/340}$ and $Fluor_{350/450}$) are close to baseline in the size interval between 8 nm and 55 nm (Figure 9 in Zhou & Guo, 2015). Based on the known relationship between nanometer and molecular weight (Lin et al. 2016), these three size intervals in nanometer scale (0.5-4 nm, 4-8 nm, >8 nm) can be transferred to kilo-Dalton scale (0.06-16.66 kDa, 16.66-108.5 kDa, >108.5 kDa, respectively). Since the nominal membrane pore-size is 0.3 kDa, the smallest size fraction could be corrected to 0.3–16.66 kDa (Zhou & Guo, 2015). It shows a molecular size spectrum (from >1 kDa to > 100 kDa) associated with switching from humic-like to protein-like fluorophores, which is consistent with our observations (Figure 6-13).

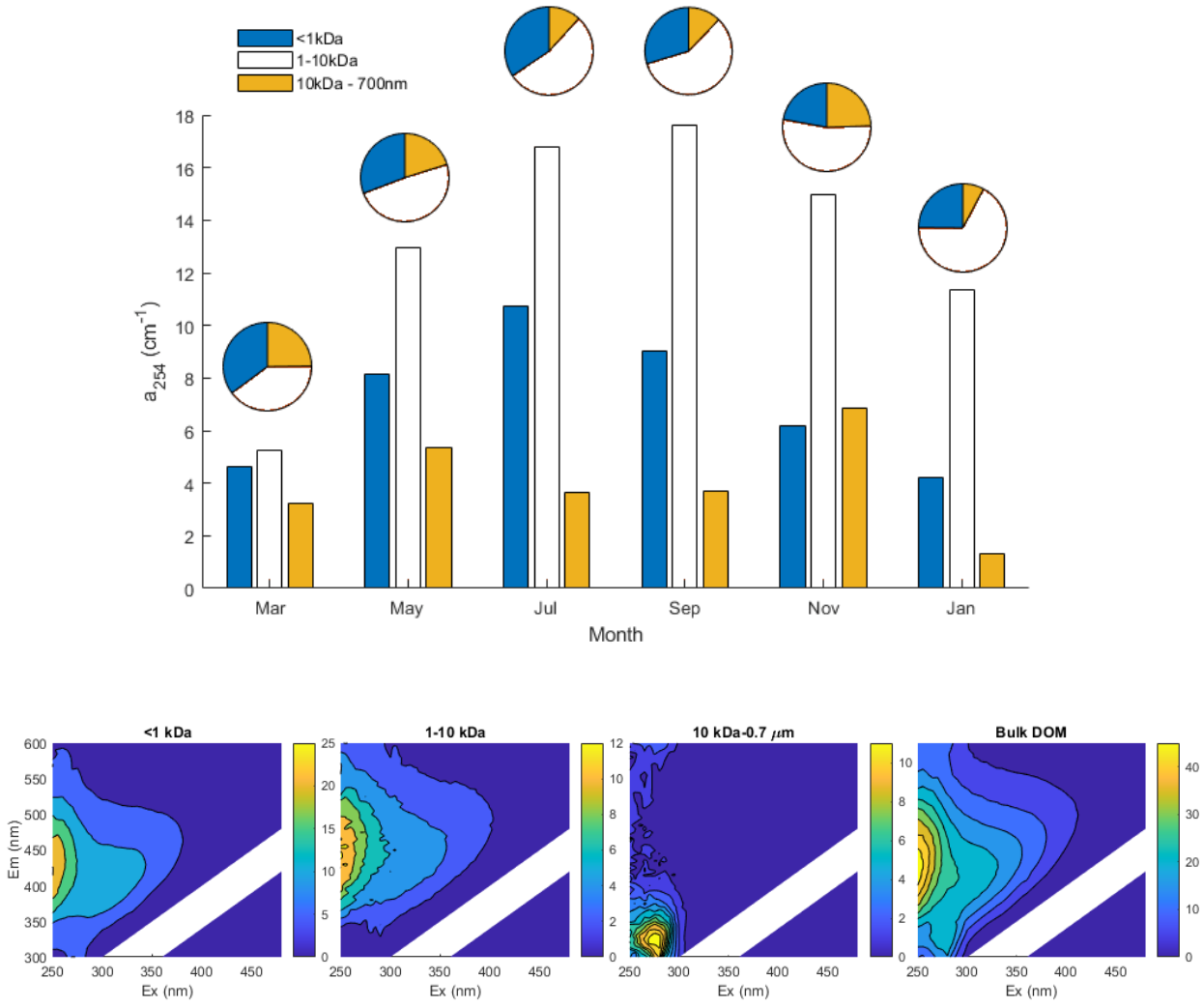

Figure 6-12. Top panel: Three size fractions (< 1 kDa, 1-10 kDa, >10 kDa) of CDOM a_{254} in the Veterans Lagoon

on March 18, May 17, July 24 and September 20, 2019, representatives of March, May, July and September. Pie charts on the top are CDOM a_{254} from water samples collected from northern part of lagoon while those on the bottom are from southern part of lagoon. Bottom panel: EEM contours for bulk water samples and different size fractions including <1 kDa, 1-10 kDa and 10 kDa – 0.7µm.

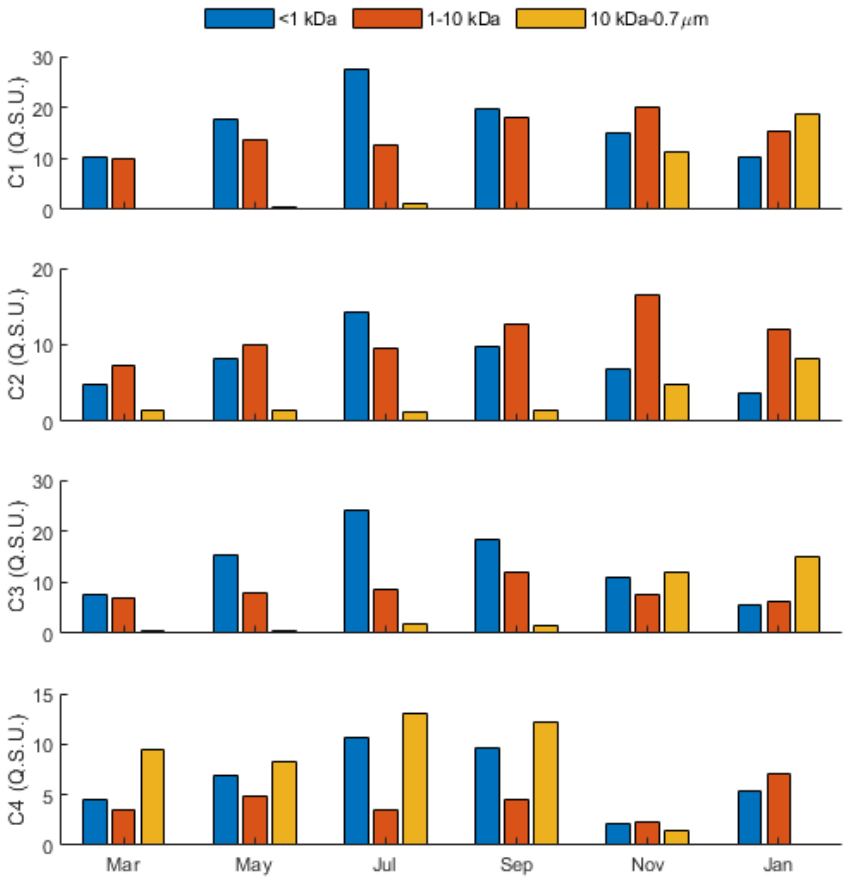

Figure 6-13. Seasonal variation of four fluorescent DOM components in three different DOM size-fractions (<1kDa, 1-10 kDa, 10 kDa – 0.7µm).

Chapter 7. Summary and Conclusions

In this dissertation, two new coupling techniques were developed to study the molecular weight distribution and size-dependent chemical composition of natural DOM in different aquatic environments.

The one-sample PARAFAC approach, coupling FlFFF size-fractionation with EEM and PARAFAC, can convincingly illustrate the variations in DOM composition and optical properties with molecular weight within individual sample and dynamic changes in DOM composition and molecular size across geochemical and hydrological interfaces. Using this method, we demonstrated that fluorescence composition of DOM is highly correlated with molecular size. The PARAFAC-derived DOM components are distinctive between samples from different aquatic environments or along a trophic gradient. This new method provides an opportunity to inspect PARAFAC-derived fluorescent components with DOM molecular size using single water sample. It benefits those researchers focusing on dynamic interface with a steep redox/ionic-strength/hydrological/trophic gradient.

The lower Fox River-Green Bay ecosystem is the largest freshwater estuary with a steep trophic gradient, where the one-sample PARAFAC method was performed for the first time. Surface water samples along a river-bay transect were analyzed individually. There existed four fluorescent PARAFAC-derived components in this estuary, including two humic-like components (C_{475} and C_{410}), which are ubiquitous in freshwater ecosystems, and two protein-like components (C_{320} and C_{290}), which were absent on the bulk EEM contours. The one-sample PARAFAC approach clearly reveals different DOM composition between stations along the river-bay transect, with the humic-like components declined, while one of the protein-like components (C_{290}) increased. Additionally, in terms of percentage of fluorescent component in the total fluorescence, the refractory humic-like component (C_{475}) remained relatively constant and the semi-labile component (C_{410}) decreased along the transect, while the protein-like components (C_{320} and C_{290}) increased. The variations in fluorescent DOM composition and size spectra along the trophic gradient could not have been revealed by the classic EEM-PARAFAC method, demonstrating the power of the new one-sample PARAFAC approach.

The data fusion technique was applied, for the first time, to integrate different DOM datasets for size-fractionated DOM samples and to provide new insights into DOM biogeochemistry in the lower Yukon River, one of the largest Arctic rivers, which are sensitive to climate change. Three EEM-PARAFAC components in the Yukon River are firstly reported; two are humic-like (C_{400} and C_{480}) and one are protein-like (C_{310}). Heterogeneity of chemical properties and composition, including three fluorescent components, fluorescent indices, and CDOM properties, between the HMW- and LMW-DOM fractions is revealed. Application of data fusion technique to jointly analyze EEM and FT-IR spectra elucidates internal connections between chemical functional groups and fluorescent components; humic-like components are majorly associated with C-H, C=C, and C-O bonds, while protein-like component is corresponding to C-N and N-H bonds. These results clearly demonstrate the seasonal variations in size-fractionated DOM composition in the Yukon River, and offer a baseline dataset to understand the Arctic biogeochemical cycling.

The Milwaukee River is one of the most important Wisconsin freshwater ecosystems. Water samples from the river were collected at the upstream and downstream monthly between 2017 and 2019. Several indicators, including DOC, CDOM a_{254}, $S_{275\text{-}295}$, $SUVA_{254}$, BIX and HIX were used to evaluate tempo-spatial variations in the riverine DOM. Seasonal variations in DOM quality and quantity in the Milwaukee River are highly related to the river discharge; large river flow brought terrestrial humic-like DOM with high molecular weight and high aromaticity in wet seasons, while autochthonous DOM was observed under ice in dry seasons. Comparisons of multiple parameters between upstream and downstream indicated signs of human impacts on river waters. Comparisons in longer-term BIX and HIX data show that the removal of Estabrook Dam (occurred at Jan 2018) had little impact on the DOM composition in the Milwaukee River. The gradual increase in HIX largely resulted from increasing precipitation and surface runoffs.

Veterans Lagoon is an ideal natural mesocosm to examine seasonal variations in DOM with progression of cyanobacterial blooms. Based on measurements of weekly or bi-weekly water samples between March 2019 and January 2020, we observed that at the end of the wet season, warm and calm surface water promoted cyanobacterial, *Microcystis,* blooms. POM and DOM reflected the bloom

progression; δ^{13}C-POC was more positive when bloom occurred, resulting from carbon (CO_2) limitation within the ecosystem, while DOC, CDOM a_{254}, and a protein-like fluorescent component showed a spike at the end of wet seasons, which might be due to the collapse of the bloom. The size fractionated EEMs showed heterogeneity and seasonality of DOM in the Veterans Lagoon; the <1 kDa and 1-10 kDa size fractions are single-peak distribution while the > 10 kDa size fraction is a bimodal distribution with two peaks in May and November, respectively. These results of size-fractionated DOM under progressions of cyanobacterial blooms provide fundamental observations for the harmful algal bloom and improve our understanding of the relationship between DOM changes and cyanobacterial blooms.

References

Abbott, B. W., J. B. Jones, E. A. G. Schuur, and others. 2016. Biomass offsets little or none of permafrost carbon release from soils, streams, and wildfire: An expert assessment. Environ. Res. Lett. **11**. doi:10.1088/1748-9326/11/3/034014

Abdulla, H. A. N., E. C. Minor, R. F. Dias, and P. G. Hatcher. 2010a. Changes in the compound classes of dissolved organic matter along an estuarine transect: A study using FTIR and 13C NMR. Geochim. Cosmochim. Acta **74**: 3815–3838. doi:10.1016/j.gca.2010.04.006

Abdulla, H. A. N., E. C. Minor, R. F. Dias, and P. G. Hatcher. 2013. Transformations of the chemical compositions of high molecular weight DOM along a salinity transect: Using two dimensional correlation spectroscopy and principal component analysis approaches. Geochim. Cosmochim. Acta **118**: 231–246. doi:10.1016/j.gca.2013.03.036

Abdulla, H. A. N., E. C. Minor, and P. G. Hatcher. 2010b. Using Two-Dimensional Correlations of 13 C NMR and FTIR To Investigate Changes in the Chemical Composition of Dissolved Organic Matter along an Estuarine Transect. Environ. Sci. Technol. **44**: 8044–8049. doi:10.1021/es100898x

Acar, E., R. Bro, and A. K. Smilde. 2015. Data Fusion in Metabolomics Using Coupled Matrix and Tensor Factorizations. Proc. IEEE **103**: 1602–1620. doi:10.1109/JPROC.2015.2438719

Acar, E., T. G. Kolda, and D. M. Dunlavy. 2011. All-at-once Optimization for Coupled Matrix and Tensor Factorizations.doi:https://arxiv.org/abs/1105.3422

Acar, E., M. Nilsson, and M. Saunders. 2014. A flexible modeling framework for coupled matrix and tensor factorizations. 22nd Eur. Signal Process. Conf. **IEEE**: 111–115.

Acar, E., C. Schenker, Y. Levin-Schwartz, V. D. Calhoun, and T. Adali. 2019. Unraveling diagnostic biomarkers of schizophrenia through structure-revealing fusion of multi-modal neuroimaging data. Front. Neurosci. **13**. doi:10.3389/fnins.2019.00416

Aiken, G. R., H. Hsu-Kim, and J. N. Ryan. 2011. Influence of Dissolved Organic Matter on the Environmental Fate of Metals, Nanoparticles, and Colloids. Environ. Sci. Technol. **45**: 3196–3201. doi:10.1021/es103992s

Aiken, G. R., R. G. M. Spencer, R. G. Striegl, P. F. Schuster, and P. A. Raymond. 2014. Influences of

glacier melt and permafrost thaw on the age of dissolved organic carbon in the Yukon River basin. Global Biogeochem. Cycles **28**: 525–537. doi:10.1002/2013GB004764

Amon, R. M. W., and R. Benner. 1996. Bacterial utilization of different size classes of dissolved organic matter. Limnol. Oceanogr. **41**: 41–51. doi:10.4319/lo.1996.41.1.0041

Andersson, T., Å. Nilsson, and M. Jansson. 2005. Coloured substances in Swedish lakes and rivers — temporal variation and regulating factors, p. 243–253. *In* Humic Substances in the Aquatic and Terrestrial Environment. Springer-Verlag.

Asmala, E., P. Massicotte, and J. Carstensen. 2021. Identification of dissolved organic matter size components in freshwater and marine environments. Limnol. Oceanogr. 1–13. doi:10.1002/lno.11692

Baalousha, M., B. Stolpe, and J. R. Lead. 2011. Flow field-flow fractionation for the analysis and characterization of natural colloids and manufactured nanoparticles in environmental systems: a critical review. J. Chromatogr. A **1218**: 4078–4103. doi:10.1016/j.chroma.2011.04.063

Bader, B. W., and T. G. Kolda. 2006. Algorithm 862: MATLAB tensor classes for fast algorithm prototyping. ACM Trans. Math. Softw. **32**: 635–653. doi:10.1145/1186785.1186794

Bader, B. W., and T. G. Kolda. 2007. Efficient MATLAB computations with sparse and factored tensors. SIAM J. Sci. Comput. **30**: 205–231. doi:10.1137/060676489

Bao, H., Y. Wu, and J. Zhang. 2015. Spatial and temporal variation of dissolved organic matter in the Changjiang: Fluvial transport and flux estimation. J. Geophys. Res. G Biogeosciences **120**: 1870–1886. doi:10.1002/2015JG002948

Belzile, C., J. A. E. Gibson, and W. F. Vincent. 2002. Colored dissolved organic matter and dissolved organic carbon exclusion from lake ice: Implications for irradiance transmission and carbon cycling. Limnol. Oceanogr. **47**: 1283–1293. doi:10.4319/lo.2002.47.5.1283

Belzile, C., and L. Guo. 2006. Optical properties of low molecular weight and colloidal organic matter: Application of the ultrafiltration permeation model to DOM absorption and fluorescence. Mar. Chem. **98**: 183–196. doi:10.1016/j.marchem.2005.08.009

Benner, R., and R. M. W. Amon. 2015. The size-reactivity continuum of major bioelements in the ocean. Ann. Rev. Mar. Sci. **7**: 185–205. doi:10.1146/annurev-marine-010213-135126

Binding, C. E., T. A. Greenberg, S. B. Watson, S. Rastin, and J. Gould. 2015. Long term water clarity changes in North America's Great Lakes from multi-sensor satellite observations. Limnol. Oceanogr. **60**: 1976–1995. doi:10.1002/lno.10146

Birdwell, J. E., and A. S. Engel. 2010. Characterization of dissolved organic matter in cave and spring waters using UV-Vis absorbance and fluorescence spectroscopy. Org. Geochem. **41**: 270–280. doi:10.1016/j.orggeochem.2009.11.002

Boehme, J., and M. Wells. 2006. Fluorescence variability of marine and terrestrial colloids: Examining size fractions of chromophoric dissolved organic matter in the Damariscotta River estuary. Mar. Chem. **101**: 95–103. doi:10.1016/j.marchem.2006.02.001

Brabets, T. P., B. Wang, and R. H. Meade. 2000. Environmental and Hydrologic Overview of the Yukon River Basin, Alaska and Canada. U.S. Geol. Surv. 106.

Cai, Y., L. Guo, and T. A. Douglas. 2008. Temporal variations in organic carbon species and fluxes from the Chena River, Alaska. Limnol. Oceanogr. **53**: 1408–1419. doi:10.4319/lo.2008.53.4.1408

Cai, Y., L. Guo, X. Wang, and G. Aiken. 2015. Abundance, stable isotopic composition, and export fluxes of DOC, POC, and DIC from the Lower Mississippi River during 2006-2008. J. Geophys. Res. Biogeosciences **120**: 2273–2288. doi:10.1002/2015JG003139

Del Castillo, C. E., F. Gilbes, P. G. Coble, and F. E. Müller-Karger. 2000. On the dispersal of riverine colored dissolved organic matter over the West Florida Shelf. Limnol. Oceanogr. **45**: 1425–1432. doi:10.4319/lo.2000.45.6.1425

Del Castillo, C. E., and R. L. Miller. 2008. On the use of ocean color remote sensing to measure the transport of dissolved organic carbon by the Mississippi River Plume. Remote Sens. Environ. **112**: 836–844. doi:10.1016/j.rse.2007.06.015

Chen, M., and R. Jaffé. 2016. Quantitative assessment of photo- and bio-reactivity of chromophoric and fluorescent dissolved organic matter from biomass and soil leachates and from surface waters in a subtropical wetland. Biogeochemistry **129**: 273–289. doi:10.1007/s10533-016-0231-7

Chen, M., S. Kim, J.-E. Park, H.-J. Jung, and J. Hur. 2016. Structural and compositional changes of dissolved organic matter upon solid-phase extraction tracked by multiple analytical tools. Anal. Bioanal. Chem. **408**: 6249–6258. doi:10.1007/s00216-016-9728-0

Chen, M., W.-X. Wang, and L. Guo. 2004. Phase partitioning and solubility of iron in natural seawater controlled by dissolved organic matter. Global Biogeochem. Cycles **18**. doi:10.1029/2003GB002160

Coble, P. G. 1996. Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy. Mar. Chem. **51**: 325–346. doi:10.1016/0304-4203(95)00062-3

Coble, P. G. 2007. Marine optical biogeochemistry: the chemistry of ocean color. Chem. Rev. **107**: 402–418.

Coble, P. G., C. E. Del Castillo, and B. Avril. 1998. Distribution and optical properties of CDOM in the Arabian Sea during the 1995 Southwest Monsoon. Deep. Res. Part II Top. Stud. Oceanogr. **45**: 2195–2223. doi:10.1016/S0967-0645(98)00068-X

Coble, P. G., J. Lead, A. Baker, D. Reynolds, and R. G. M. Spencer. 2014. Aquatic Organic Matter Fluorescence, 1st ed. P.G. Coble, J. Lead, A. Baker, D.M. Reynolds, and R.G.M. Spencer [eds.]. Cambridge University Press.

Cooper, L. W., R. Benner, J. W. McClelland, and others. 2005. Linkages among runoff, dissolved organic carbon, and the stable oxygen isotope composition of seawater and other water mass indicators in the Arctic Ocean. J. Geophys. Res. **110**: 1–14. doi:10.1029/2005JG000031

Cory, R. M., C. P. Ward, B. C. Crump, and G. W. Kling. 2014. Sunlight controls water column processing of carbon in arctic fresh waters. Science (80-.). **345**: 925–928. doi:10.1126/science.1253119

Cuss, C. W., and C. Guéguen. 2013. Distinguishing dissolved organic matter at its origin: Size and optical properties of leaf-litter leachates. Chemosphere **92**: 1483–1489. doi:10.1016/j.chemosphere.2013.03.062

Dainard, P. G., C. Guéguen, N. McDonald, and W. J. Williams. 2015. Photobleaching of fluorescent dissolved organic matter in Beaufort Sea and North Atlantic Subtropical Gyre. Mar. Chem. **177**: 630–637. doi:10.1016/j.marchem.2015.10.004

Deming, J. W. 2002. Psychrophiles and polar regions. Curr. Opin. Microbiol. **5**: 301–309. doi:10.1016/S1369-5274(02)00329-6

DeVilbiss, S. E., Z. Zhou, J. V. Klump, and L. Guo. 2016a. Spatiotemporal variations in the abundance and composition of bulk and chromophoric dissolved organic matter in seasonally hypoxia-influenced Green Bay, Lake Michigan, USA. Sci. Total Environ. **565**: 742–757. doi:10.1016/j.scitotenv.2016.05.015

DeVilbiss, S. E., Z. Zhou, J. V. Klump, and L. Guo. 2016b. Spatiotemporal variations in the

abundance and composition of bulk and chromophoric dissolved organic matter in seasonally hypoxia-influenced Green Bay, Lake Michigan, USA. Sci. Total Environ. **565**: 742–757. doi:10.1016/j.scitotenv.2016.05.015

Fichot, C. G., and R. Benner. 2011. A novel method to estimate DOC concentrations from CDOM absorption coefficients in coastal waters. Geophys. Res. Lett. **38**. doi:10.1029/2010GL046152

Fichot, C. G., and R. Benner. 2012. The spectral slope coefficient of chromophoric dissolved organic matter (S275-295) as a tracer of terrigenous dissolved organic carbon in river-influenced ocean margins. Limnol. Oceanogr. **57**: 1453–1466. doi:10.4319/lo.2012.57.5.1453

Fichot, C. G., and R. Benner. 2014. The fate of terrigenous dissolved organic carbon in a river-influenced ocean margin. Global Biogeochem. Cycles **28**: 300–318. doi:10.1002/2013GB004670

Finlay, J., J. Neff, S. Zimov, A. Davydova, and S. Davydov. 2006. Snowmelt dominance of dissolved organic carbon in high-latitute watersheds: Implications for characterization and flux of river DOC. Geophys. Res. Lett. **33**: 2–6. doi:10.1029/2006GL025754

Frenette, J. J., P. Thibeault, J. F. Lapierre, and P. B. Hamilton. 2008. Presence of algae in freshwater ice cover of fluvial Lac Saint-Pierre (St. Lawrence River, Canada). J. Phycol. **44**: 284–291. doi:10.1111/j.1529-8817.2008.00481.x

Frey, K. E., and L. C. Smith. 2005. Amplified carbon release from vast West Siberian peatlands by 2100. Geophys. Res. Lett. **32**: 1–4. doi:10.1029/2004GL022025

Gao, L., Z. Zhou, A. V. Reyes, and L. Guo. 2018. Yields and Characterization of Dissolved Organic Matter From Different Aged Soils in Northern Alaska. J. Geophys. Res. Biogeosciences **123**: 2035–2052. doi:10.1029/2018JG004408

Giddings, J. C. 1993. Field-flow fractionation : analysis of macromolecular, colloidal, and particulate materials. Science (80-.). **260**: 1456–1465.

Gill, P. E., W. Murray, and M. A. Saunders. 2005. SNOPT: An SQP algorithm for large-scale constrained optimization. SIAM Rev. **47**: 99–131. doi:10.1137/S0036144504446096

Gobler, C. J., and S. A. Sañudo-Wilhelmy. 2003. Cycling of colloidal organic carbon and nitrogen during an estuarine phytoplankton bloom. Limnol. Oceanogr. **48**: 2314–2320. doi:10.4319/lo.2003.48.6.2314

Gonçalves-Araujo, R., M. A. Granskog, A. Bracher, K. Azetsu-Scott, P. A. Dodd, and C. A. Stedmon. 2016. Using fluorescent dissolved organic matter to trace and distinguish the origin of Arctic surface waters. Sci. Rep. **6**: 1–12. doi:10.1038/srep33978

Gu, B., A. D. Chapman, and C. L. Schelske. 2006. Factors controlling seasonal variations in stable isotope composition of particulate organic matter in a soft water eutrophic lake. Limnol. Oceanogr. **51**: 2837–2848. doi:10.4319/lo.2006.51.6.2837

Guéguen, C., L. Guo, D. Wang, N. Tanaka, and C. C. Hung. 2006. Chemical characteristics and origin of dissolved organic matter in the Yukon River. Biogeochemistry **77**: 139–155. doi:10.1007/s10533-005-0806-1

Guo, L., Y. Cai, C. Belzile, and R. W. Macdonald. 2012. Sources and export fluxes of inorganic and organic carbon and nutrient species from the seasonally ice-covered Yukon River. Biogeochemistry **107**: 187–206. doi:10.1007/s10533-010-9545-z

Guo, L., C. H. Coleman Jr, and P. H. Santschi. 1994. The distribution of colloidal and dissolved organic carbon in the Gulf of Mexico. Mar. Chem. **45**: 105–119.

Guo, L., and H. Lin. 2017. Characterization of dissolved organic matter (DOM) from the lower Yukon River. *AGU Fall Meeting*.

Guo, L., and R. W. Macdonald. 2006. Source and transport of terrigenous organic matter in the upper Yukon River: Evidence from isotope (δ13C, Δ 14C, and δ15N) composition of dissolved, colloidal, and particulate phases. Global Biogeochem. Cycles **20**: 1–12. doi:10.1029/2005GB002593

Guo, L., C. L. Ping, and R. W. Macdonald. 2007a. Mobilization pathways of organic carbon from permafrost to arctic rivers in a changing climate. Geophys. Res. Lett. **34**: 1–5. doi:10.1029/2007GL030689

Guo, L., and P. H. Santschi. 1996. A critical evaluation of the cross-flow ultrafiltration technique for sampling colloidal organic carbon in seawater. Mar. Chem. **55**: 113–127. doi:10.1016/S0304-4203(96)00051-5

Guo, L., and P. H. Santschi. 1997. Composition and cycling of colloids in marine environments. Rev. Geophys. **35**: 17. doi:10.1029/96RG03195

Guo, L., and P. H. Santschi. 2007. Ultrafiltration and its Applications to Sampling and Characterisation of Aquatic Colloids K.J. Wilkinson and J.R. Lead [eds.]. Environ. Colloids Part. 159–221. doi:10.1002/9780470024539.ch4

Guo, L., P. H. Santschi, and K. Warnken. 1995. Dynamics of dissolved organic carbon (DOC) in oceanic environments. Limnol. Oceanogr. **40**: 1392–1403.

Guo, L., L.-S. Wen, D. Tang, and P. H. Santschi. 2000. Re-examination of cross-flow ultrafiltration for sampling aquatic colloids: evidence from molecular probes. Mar. Chem. **69**: 75–90. doi:10.1016/S0304-4203(99)00097-3

Guo, L., D. M. White, C. Xu, and P. H. Santschi. 2009. Chemical and isotopic composition of high-molecular-weight dissolved organic matter from the Mississippi River plume. Mar. Chem. **114**: 63–71. doi:10.1016/j.marchem.2009.04.002

Guo, W., C. A. Stedmon, Y. Han, F. Wu, X. Yu, and M. Hu. 2007b. The conservative and non-conservative behavior of chromophoric dissolved organic matter in Chinese estuarine waters. Mar. Chem. **107**: 357–366. doi:10.1016/j.marchem.2007.03.006

Guo, W., L. Yang, W. Zhai, W. Chen, C. L. Osburn, X. Huang, and Y. Li. 2014. Runoff-mediated seasonal oscillation in the dynamics of dissolved organic matter in different branches of a large bifurcated estuary-The Changjiang Estuary. J. Geophys. Res. Biogeosciences **119**: 776–793. doi:10.1002/2013JG002540

Hansell, D. A., and C. A. Carlson. 2002. Biogeochemistry of marine dissolved organic matter, 1st ed. D.A. Hansell and C.A. Carlson [eds.]. Academic press.

Hansen, A. M., T. E. C. Kraus, B. A. Pellerin, J. A. Fleck, B. D. Downing, and B. A. Bergamaschi. 2016. Optical properties of dissolved organic matter (DOM): Effects of biological and photolytic degradation. Limnol. Oceanogr. **61**: 1015–1032. doi:10.1002/lno.10270

Harjung, A., F. Sabater, and A. Butturini. 2018. Hydrological connectivity drives dissolved organic matter processing in an intermittent stream. Limnologica **68**: 71–81. doi:10.1016/j.limno.2017.02.007

Hedges, J., R. Keill, and R. Benner. 1997. What happens to terrestrial organic matter in the ocean? Org. Geochem. **27**: 195–212.

Helms, J. R., A. Stubbins, J. D. Ritchie, E. C. Minor, D. J. Kieber, and K. Mopper. 2008. Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. Limnol. Oceanogr. **53**: 955–969. doi:10.4319/lo.2008.53.3.0955

Her, N., G. Amy, D. McKnight, J. Sohn, and Y. Yoon. 2003. Characterization of DOM as a function

of MW by fluorescence EEM and HPLC-SEC using UVA, DOC, and fluorescence detection. Water Res. **37**: 4295–4303. doi:10.1016/S0043-1354(03)00317-8

Holmes, R. M., J. W. McClelland, P. A. Raymond, B. B. Frazer, B. J. Peterson, and M. Stieglitz. 2008. Lability of DOC transported by Alaskan rivers to the Arctic Ocean. Geophys. Res. Lett. **35**: L03402. doi:10.1029/2007GL032837

Huguet, A., L. Vacher, S. Relexans, S. Saubusse, J. M. Froidefond, and E. Parlanti. 2009. Properties of fluorescent dissolved organic matter in the Gironde Estuary. Org. Geochem. **40**: 706–719. doi:10.1016/j.orggeochem.2009.03.002

Hur, J., N. C. Jung, and J. K. Shin. 2007. Spectroscopic distribution of dissolved organic matter in a dam reservoir impacted by turbid storm runoff. Environ. Monit. Assess. **133**: 53–67. doi:10.1007/s10661-006-9559-0

Hur, J., B. M. Lee, S. Lee, and J. K. Shin. 2014. Characterization of chromophoric dissolved organic matter and trihalomethane formation potential in a recently constructed reservoir and the surrounding areas - Impoundment effects. J. Hydrol. **515**: 71–80. doi:10.1016/j.jhydrol.2014.04.035

Jiao, N., G. J. Herndl, D. A. Hansell, and others. 2010. Microbial production of recalcitrant dissolved organic matter: Long-term carbon storage in the global ocean. Nat. Rev. Microbiol. **8**: 593–599. doi:10.1038/nrmicro2386

Jung, Y. M., H. S. Shin, S. Bin Kim, and I. Noda. 2002. New Approach to Generalized Two-Dimensional Correlation Spectroscopy. 1: Combination of Principal Component Analysis and Two-Dimensional Correlation Spectroscopy. Appl. Spectrosc. **56**: 1562–1567. doi:10.1366/000370202321116020

Kida, M., T. Kojima, Y. Tanabe, K. Hayashi, S. Kudoh, N. Maie, and N. Fujitake. 2019. Origin, distributions, and environmental significance of ubiquitous humic-like fluorophores in Antarctic lakes and streams. Water Res. **163**: 114901. doi:10.1016/j.watres.2019.114901

Kim, J., and G. Kim. 2017. Inputs of humic fluorescent dissolved organic matter via submarine groundwater discharge to coastal waters off a volcanic island (Jeju, Korea). Sci. Rep. **7**: 7921. doi:10.1038/s41598-017-08518-5

De Kluijver, A., J. Yu, M. Houtekamer, J. J. Middelburg, and Z. Liu. 2012. Cyanobacteria as a carbon source for zooplankton in eutrophic Lake Taihu, China, measured by 13C labeling and fatty acid biomarkers. Limnol. Oceanogr. **57**: 1245–1254. doi:10.4319/lo.2012.57.4.1245

Kteeba, S. M., H. I. El-Adawi, O. A. El-Rayis, A. E. El-Ghobashy, J. L. Schuld, K. R. Svoboda, and L. Guo. 2017. Zinc oxide nanoparticle toxicity in embryonic zebrafish: Mitigation with different natural organic matter. Environ. Pollut. **230**: 1125–1140. doi:10.1016/j.envpol.2017.07.042

Kulkarni, H. V., N. Mladenov, S. Datta, and D. Chatterjee. 2018. Influence of monsoonal recharge on arsenic and dissolved organic matter in the Holocene and Pleistocene aquifers of the Bengal Basin. Sci. Total Environ. **637**–**638**: 588–599. doi:10.1016/j.scitotenv.2018.05.009

Kumagai, M., S. I. Nakano, C. Jiao, K. Hayakawa, S. Tsujimura, T. Nakajima, J. J. Frenetie, and A. Quesada. 2000. Effect of cyanobacterial blooms on thermal stratification. Limnology **1**: 191–195. doi:10.1007/s102010070006

Laboratories, C. S. 1939. Problems of stability in hydrophobic colloidal solutions II. On the interaction of two colloidal metallic particles: m athem atical theory. Proc. R. Soc. London. Ser. A. Math. Phys. Sci. **170**: 165–182. doi:10.1098/rspa.1939.0024

Laird, G. a., and D. Scavia. 1990. Distribution of labile dissolved organic carbon in Lake Michigan. Limnol. Oceanogr. **35**: 443–447. doi:10.4319/lo.1990.35.2.0443

Lambert, T., C. R. Teodoru, F. C. Nyoni, S. Bouillon, F. Darchambeau, P. Massicotte, and A. V. Borges. 2016. Along-stream transport and transformation of dissolved organic matter in a large tropical river. Biogeosciences **13**: 2727–2741. doi:10.5194/bg-13-2727-2016

Lee, M.-H., C. L. Osburn, K.-H. Shin, and J. Hur. 2018. New insight into the applicability of spectroscopic indices for dissolved organic matter (DOM) source discrimination in aquatic systems affected by biogeochemical processes. Water Res. **147**: 164–176. doi:10.1016/j.watres.2018.09.048

Lehman, P. W., C. Kendall, M. A. Guerin, M. B. Young, S. R. Silva, G. L. Boyer, and S. J. Teh. 2014. Characterization of the Microcystis Bloom and Its Nitrogen Supply in San Francisco Estuary Using Stable Isotopes. Estuaries and Coasts **38**: 165–178. doi:10.1007/s12237-014-9811-8

Lin, H., M. Chen, J. Zeng, Q. Li, R. Jia, X. Sun, M. Zheng, and Y. Qiu. 2016. Size characteristics of chromophoric dissolved organic matter in the Chukchi Sea. J. Geophys. Res. Ocean. **121**: 6403–6417. doi:10.1002/2016JC011771

Lin, H., and L. Guo. 2020. Variations in Colloidal DOM Composition with Molecular Weight within Individual Water Samples as Characterized by Flow Field-Flow Fractionation and EEM-PARAFAC Analysis. Environ. Sci. Technol. **54**: 1657–1667. doi:10.1021/acs.est.9b07123

Lin, P., and L. Guo. 2016. Do invasive quagga mussels alter CO2 dynamics in the Laurentian Great Lakes? Sci. Rep. **6**: 39078. doi:10.1038/srep39078

Loh, A. N., J. E. Bauer, and E. R. M. Druffel. 2004. Variable ageing and storage of dissolved organic components in the open ocean. Nature **430**: 877–881. doi:10.1038/nature02780

Lotfi-Kalahroodi, E., A.-C. Pierson-Wickmann, H. Guénet, and others. 2019. Iron isotope fractionation in iron-organic matter associations: Experimental evidence using filtration and ultrafiltration. Geochim. Cosmochim. Acta **250**: 98–116. doi:10.1016/j.gca.2019.01.036

Lv, S., Q. Yu, F. Wang, Y. Wang, W. Yan, and Y. Li. 2019. A Synthetic Model to Quantify Dissolved Organic Carbon Transport in the Changjiang River System: Model Structure and Spatiotemporal Patterns. J. Adv. Model. Earth Syst. 1–18. doi:10.1029/2019ms001648

Maie, N., N. M. Scully, O. Pisani, and R. Jaffé. 2007. Composition of a protein-like fluorophore of dissolved organic matter in coastal wetland and estuarine ecosystems. Water Res. **41**: 563–570. doi:10.1016/j.watres.2006.11.006

Maki, K., C. Kim, C. Yoshimizu, I. Tayasu, T. Miyajima, and T. Nagata. 2010. Autochthonous origin of semi-labile dissolved organic carbon in a large monomictic lake (Lake Biwa): Carbon stable isotopic evidence. Limnology **11**: 143–153. doi:10.1007/s10201-009-0299-z

Malinowski, E. R. 2002. Factor analysis in chemistry, Wiley.

Markus, A. A., J. R. Parsons, E. W. M. Roex, P. de Voogt, and R. W. P. M. Laane. 2016. Modelling the Release, Transport and Fate of Engineered Nanoparticles in the Aquatic Environment – A Review, p. 53–87. *In* Reviews of Environmental Contamination and Toxicology.

Mccallum, S., T. A. Solberg, J. E. Tiefenthaler, G. M. O 'brien, H. F. Behnke, H. D. Poulson, C. L. Stepp, and S. D. Willett. 2001. The State of the Milwaukee River Basin.

Mcguire, A. D., L. G. Anderson, T. R. Christensen, and others. 2009. Sensitivity of the carbon cycle in the Arctic to climate change. Ecol. Monogr. **79**: 523–555. doi:10.1890/08-2025.1

Mcknight, D. M., E. W. Boyer, P. K. Westerhoff, P. T. Doran, T. Kulbe, and D. T. Andersen. 2001. Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. Limnol. Oceanogr. **46**: 38–48. doi:10.4319/lo.2001.46.1.0038

Michener, R., and K. Lajtha. 2008. Stable Isotopes in Ecology and Environmental Science, Second Edi. R. Michener and K. Lajtha [eds.].

Milwaukee Riverkeepers. 2011. Milwaukee River Basin Report Card 2010.

Milwaukee Riverkeepers. 2012. Milwaukee River Basin Report Card 2012.

Milwaukee Riverkeepers. 2013. Milwaukee River Basin Report Card 2013.

Milwaukee Riverkeepers. 2015. Milwaukee River Basin Report Card 2015.

Milwaukee Riverkeepers. 2018. Milwaukee River Basin Report Card 2018.

Minor, E. C., J. P. Simjouw, and M. R. Mulholland. 2006. Seasonal variations in dissolved organic carbon concentrations and characteristics in a shallow coastal bay. Mar. Chem. **101**: 166–179. doi:10.1016/j.marchem.2006.02.005

Murphy, K. R., K. D. Butler, R. G. M. Spencer, C. A. Stedmon, J. R. Boehme, and G. R. Aiken. 2010. Measurement of Dissolved Organic Matter Fluorescence in Aquatic Environments: An Interlaboratory Comparison. Environ. Sci. Technol. **44**: 9405–9412. doi:10.1021/es102362t

Murphy, K. R., C. A. Stedmon, D. Graeber, and R. Bro. 2013a. Fluorescence spectroscopy and multi-way techniques. PARAFAC. Anal. Methods **5**: 6557–6566. doi:10.1039/c3ay41160e

Murphy, K. R., C. A. Stedmon, D. Graeber, and R. Bro. 2013b. Fluorescence spectroscopy and multi-way techniques. PARAFAC. Anal. Methods **5**: 6557. doi:10.1039/c3ay41160e

Murphy, K. R., C. A. Stedmon, P. Wenig, and R. Bro. 2014. OpenFluor– an online spectral library of auto-fluorescence by organic compounds in the environment. Anal. Methods **6**: 658–661. doi:10.1039/C3AY41935E

Murphy, K. R., S. A. Timko, M. Gonsior, L. C. Powers, U. J. Wünsch, and C. A. Stedmon. 2018a. Photochemistry Illuminates Ubiquitous Organic Matter Fluorescence Spectra. Environ. Sci. Technol. **52**: 1–12. doi:10.1021/acs.est.8b02648

Murphy, K. R., S. A. Timko, M. Gonsior, L. C. Powers, U. J. Wünsch, and C. A. Stedmon. 2018b. Photochemistry Illuminates Ubiquitous Organic Matter Fluorescence Spectra. Environ. Sci. Technol. **52**: 11243–11250. doi:10.1021/acs.est.8b02648

O'Connor, J. A., K. Lu, L. Guo, B. E. Rosenheim, and Z. Liu. 2020. Composition and lability of riverine dissolved organic matter: Insights from thermal slicing ramped pyrolysis GC–MS, amino acid, and stable isotope analyses. Org. Geochem. **149**: 104100. doi:10.1016/j.orggeochem.2020.104100

O'Donnell, J. A., G. R. Aiken, E. S. Kane, and J. B. Jones. 2010. Source water controls on the character and origin of dissolved organic matter in streams of the Yukon River basin, Alaska. J. Geophys. Res. Biogeosciences **115**: 1–12. doi:10.1029/2009JG001153

O'Donnell, J., T. Douglas, A. Barker, and L. Guo. 2021. Changing Biogeochemical Cycles of Organic Carbon, Nitrogen, Phosphorus, and Trace Elements in Arctic Rivers, p. 315–348. *In* Y. D. and K. D. [eds.], Arctic Hydrology, Permafrost and Ecosystems. Springer International Publishing.

Ogawa, H., and E. Tanoue. 2003. Dissolved Organic Matter in Oceanic Waters. J. Oceanogr. **59**: 129–147. doi:10.1023/A:1025528919771

Opsahl, S., and R. Benner. 1997. Distribution and cycling of terrigenous dissolved organic matter in the ocean. Nature **386**: 480–482. doi:10.1038/386480a0

Osburn, C. L., N. J. Anderson, C. A. Stedmon, M. E. Giles, E. J. Whiteford, T. J. McGenity, A. J. Dumbrell, and G. J. C. Underwood. 2017. Shifts in the Source and Composition of Dissolved Organic Matter in Southwest Greenland Lakes Along a Regional Hydro-climatic Gradient. J. Geophys. Res. Biogeosciences **122**: 3431–3445. doi:10.1002/2017JG003999

Osburn, C. L., T. J. Boyd, M. T. Montgomery, T. S. Bianchi, R. B. Coffin, and H. W. Paerl. 2016.

Optical proxies for terrestrial dissolved organic matter in estuaries and coastal waters. Front. Mar. Sci. **2**. doi:10.3389/fmars.2015.00127

Osburn, C. L., L. T. Handsel, M. P. Mikan, H. W. Paerl, and M. T. Montgomery. 2012. Fluorescence Tracking of Dissolved and Particulate Organic Matter Quality in a River-Dominated Estuary. Environ. Sci. Technol. **46**: 8628–8636. doi:10.1021/es3007723

Osburn, C. L., J. D. Kinsey, T. S. Bianchi, and M. R. Shields. 2019. Formation of planktonic chromophoric dissolved organic matter in the ocean. Mar. Chem. **209**: 1–13. doi:10.1016/j.marchem.2018.11.010

Ozawa, K., A. Yokoyama, K. Ishikawa, M. Kumagai, M. F. Watanabe, and H. D. Park. 2003. Accumulation and depuration of microcystin produced by the cyanobacterium Microcystis in a freshwater snail. Limnology **4**: 131–138. doi:10.1007/s10201-003-0106-1

Pace, M. L., and J. J. Cole. 2002. Synchronous variation of dissolved organic carbon and color in lakes. Limnol. Oceanogr. **47**: 333–342. doi:10.4319/lo.2002.47.2.0333

Park, H. K., M. S. Byeon, Y. N. Shin, and D. Il Jung. 2009. Sources and spatial and temporal characteristics of organic carbon in two large reservoirs with contrasting hydrologic characteristics. Water Resour. Res. **45**: 1–12. doi:10.1029/2009WR008043

Patriarca, C., V. T. Sedano-Núñez, S. L. Garcia, J. Bergquist, S. Bertilsson, P. J. R. Sjöberg, L. J. Tranvik, and J. A. Hawkes. 2020. Character and environmental lability of cyanobacteria-derived dissolved organic matter. Limnol. Oceanogr. 1–14. doi:10.1002/lno.11619

Perminova, I. V., F. H. Frimmel, A. V Kudryavtsev, N. A. Kulikova, G. Abbt-Braun, S. Hesse, and V. S. Petrosyan. 2003. Molecular Weight Characteristics of Humic Substances from Different Environments As Determined by Size Exclusion Chromatography and Their Statistical Evaluation. Environ. Sci. Technol. **37**: 2477–2485. doi:10.1021/es0258069

Philippe, A., and G. E. Schaumann. 2014. Interactions of Dissolved Organic Matter with Natural and Engineered Inorganic Colloids: A Review. Environ. Sci. Technol. **48**: 8946–8962. doi:10.1021/es502342r

Pokrovsky, O. S., L. S. Shirokova, S. N. Kirpotin, S. Audry, J. Viers, and B. Dupré. 2011. Effect of permafrost thawing on organic carbon and trace element colloidal speciation in the thermokarst lakes of western Siberia. Biogeosciences **8**: 565–583. doi:10.5194/bg-8-565-2011

Le Quéré, C., R. M. Andrew, P. Friedlingstein, and others. 2018. Global Carbon Budget 2017. Earth Syst. Sci. Data **10**: 405–448. doi:10.5194/essd-10-405-2018

Rath, J., C. Nenn, and B. Lukasz. 2013. 2014 Milwaukee River Basin Report Card.

Raymond, P. A., J. W. McClelland, R. M. Holmes, and others. 2007. Flux and age of dissolved organic carbon exported to the Arctic Ocean: A carbon isotopic study of the five largest arctic rivers. Global Biogeochem. Cycles **21**: 1–9. doi:10.1029/2007GB002934

Reichle, R. H., Q. Liu, R. D. Koster, C. S. Draper, S. P. P. Mahanama, and G. S. Partyka. 2017. Land Surface Precipitation in MERRA-2. J. Clim. **30**: 1643–1664. doi:10.1175/JCLI-D-16-0570.1

Ritchie, H., and M. Roser. 2020. CO2 emissions. Our World Data 1–31.

Romera-Castillo, C., M. Chen, Y. Yamashita, and R. Jaffé. 2014. Fluorescence characteristics of size-fractionated dissolved organic matter: Implications for a molecular assembly based structure? Water Res. **55**: 40–51. doi:10.1016/j.watres.2014.02.017

Romera-Castillo, C., M. Pinto, T. M. Langer, X. A. Álvarez-Salgado, and G. J. Herndl. 2018. Dissolved organic carbon leaching from plastics stimulates microbial activity in the ocean. Nat. Commun. **9**. doi:10.1038/s41467-018-03798-5

Santschi, P. H. 2018. Marine colloids, agents of the self-cleansing capacity of aquatic systems: Historical perspective and new discoveries. Mar. Chem. **207**: 124–135. doi:10.1016/j.marchem.2018.11.003

Santschi, P. H., L. Guo, M. Baskaran, S. Trumbore, J. Southon, T. S. Bianchi, B. Honeyman, and L. Cifuentes. 1995. Isotopic evidence for the contemporary origin of high-molecular weight organic matter in oceanic environments. Geochim. Cosmochim. Acta **59**: 625–631. doi:10.1016/0016-7037(94)00378-Y

Schuur, E. A. G., A. D. McGuire, C. Schädel, and others. 2015. Climate change and the permafrost carbon feedback. Nature **520**: 171–179. doi:10.1038/nature14338

Shen, H., Y. Niu, P. Xie, M. Tao, and X. Yang. 2011. Morphological and physiological changes in Microcystis aeruginosa as a result of interactions with heterotrophic bacteria. Freshw. Biol. **56**: 1065–1080. doi:10.1111/j.1365-2427.2010.02551.x

Shogren, A. J., J. P. Zarnetske, B. W. Abbott, F. Iannucci, and W. B. Bowden. 2020. We cannot shrug off the shoulder seasons: addressing knowledge and data gaps in an Arctic headwater. Environ. Res. Lett. **15**: 104027. doi:10.1088/1748-9326/ab9d3c

Shutova, Y., A. Baker, J. Bridgeman, and R. K. Henderson. 2014. Spectroscopic characterisation of dissolved organic matter changes in drinking water treatment: From PARAFAC analysis to online monitoring wavelengths. Water Res. **54**: 159–169. doi:10.1016/j.watres.2014.01.053

Sobek, S., L. J. Tranvik, Y. T. Prairie, P. Kortelainen, and J. J. Cole. 2007. Patterns and regulation of dissolved organic carbon: An analysis of 7,500 widely distributed lakes. Limnol. Oceanogr. **52**: 1208–1219. doi:10.4319/lo.2007.52.3.1208

Solecki, W. D., C. Rosenzweig, L. Parshall, G. Pope, M. Clark, J. Cox, and M. Wiencke. 2005. Mitigation of the heat island effect in urban New Jersey. Environ. Hazards **6**: 39–49. doi:10.1016/j.hazards.2004.12.002

Søndergaar, M., and D. Thomas. 2004. Dissolved Organic Matter (DOM) in Aquatic Ecosystems: A study of European Catchments and Coastal Waters, M. Søndergaar and D.N. Thomas [eds.]. The Domaine project.

Spark, W. 2017. Average Weather in Milwaukee Wisconsin, United States.

Spencer, R. G. M., G. R. Aiken, K. D. Butler, M. M. Dornblaser, R. G. Striegl, and P. J. Hernes. 2009. Utilizing chromophoric dissolved organic matter measurements to derive export and reactivity of dissolved organic carbon exported to the Arctic Ocean: A case study of the Yukon River, Alaska. Geophys. Res. Lett. **36**: 1–6. doi:10.1029/2008GL036831

Spencer, R. G. M., G. R. Aiken, K. P. Wickland, R. G. Striegl, and P. J. Hernes. 2008. Seasonal and spatial variability in dissolved organic matter quantity and composition from the Yukon River basin, Alaska. Global Biogeochem. Cycles **22**: 1–13. doi:10.1029/2008GB003231

Spencer, R. G. M., K. D. Butler, and G. R. Aiken. 2012. Dissolved organic carbon and chromophoric dissolved organic matter properties of rivers in the USA. J. Geophys. Res. Biogeosciences **117**. doi:10.1029/2011JG001928

Stedmon, C. A., R. M. W. Amon, A. J. Rinehart, and S. A. Walker. 2011. The supply and characteristics of colored dissolved organic matter (CDOM) in the Arctic Ocean: Pan Arctic trends and differences. Mar. Chem. **124**: 108–118. doi:10.1016/j.marchem.2010.12.007

Stedmon, C. A., and R. Bro. 2008a. Characterizing dissolved organic matter fluorescence with parallel factor analysis : a tutorial. Limnol. Oceanogr. Methods 572–579.

Stedmon, C. A., and R. Bro. 2008b. Characterizing dissolved organic matter fluorescence with parallel factor analysis: a tutorial. Limnol. Oceanogr. Methods **6**: 572–579.

doi:10.4319/lom.2008.6.572

Stedmon, C. A., and S. Markager. 2005. Resolving the variability in dissolved organic matter fluorescence in a temperate estuary and its catchment using PARAFAC analysis. Limnol. Oceanogr. **50**: 686–697.

Stedmon, C. A., S. Markager, and R. Bro. 2003. Tracing dissolved organic matter in aquatic environments using a new approach to fluorescence spectroscopy. Mar. Chem. **82**: 239–254.

Stolpe, B., L. Guo, and A. M. Shiller. 2013. Binding and transport of rare earth elements by organic and iron-rich nanocolloids in Alaskan rivers, as revealed by field-flow fractionation and ICP-MS. Geochim. Cosmochim. Acta **106**: 446–462. doi:10.1016/j.gca.2012.12.033

Stolpe, B., L. Guo, A. M. Shiller, and M. Hassellöv. 2010. Size and composition of colloidal organic matter and trace elements in the Mississippi River, Pearl River and the northern Gulf of Mexico, as characterized by flow field-flow fractionation. Mar. Chem. **118**: 119–128. doi:10.1016/j.marchem.2009.11.007

Stolpe, B., Z. Zhou, L. Guo, and A. M. Shiller. 2014. Colloidal size distribution of humic- and protein-like fluorescent organic matter in the northern Gulf of Mexico. Mar. Chem. **164**: 25–37. doi:10.1016/j.marchem.2014.05.007

Strickland, J. D. H., and T. R. Parsons. 1972. A practical handbook of seawater analysis,.

Striegl, R. G., G. R. Aiken, M. M. Dornblaser, P. A. Raymond, and K. P. Wickland. 2005. A decrease in discharge-normalized DOC export by the Yukon River during summer through autumn. Geophys. Res. Lett. **32**: 1–4. doi:10.1029/2005GL024413

Striegl, R. G., M. M. Dornblaser, G. R. Aiken, K. P. Wickland, and P. A. Raymond. 2007. Carbon export and cycling by the Yukon, Tanana, and Porcupine rivers, Alaska, 2001-2005. Water Resour. Res. **43**: 2001–2005. doi:10.1029/2006WR005201

Sun, L., C. Xu, S. Zhang, and others. 2017. Light-induced aggregation of microbial exopolymeric substances. Chemosphere **181**: 675–681. doi:10.1016/j.chemosphere.2017.04.099

Tanaka, K., K. Kuma, K. Hamasaki, and Y. Yamashita. 2014. Accumulation of humic-like fluorescent dissolved organic matter in the Japan Sea. Sci. Rep. **4**: 1–7. doi:10.1038/srep05292

Tang, D., K. W. Warnken, and P. H. Santschi. 2001. Organic complexation of copper in surface waters of Galveston Bay. Limnol. Oceanogr. **46**: 321–330. doi:10.4319/lo.2001.46.2.0321

Teber, T. 2016. Seasonal and spatial variations in chemical composition and fluxes of dissolved organic matter and nutrients in the lower Milwaukee River. University of Wisconsin, Milwaukee.

Thurman, E. M. 2012. Organic geochemistry of natural waters, Springer Science & Business Media.

Turetsky, M. R., B. W. Abbott, M. C. Jones, and others. 2020. Carbon release through abrupt permafrost thaw. Nat. Geosci. **13**: 138–143. doi:10.1038/s41561-019-0526-0

Verdugo, P. 2012. Marine microgels. Ann. Rev. Mar. Sci. **4**: 375–400. doi:10.1146/annurev-marine-120709-142759

Vonk, J. E., P. J. Mann, S. Davydov, and others. 2013. High biolability of ancient permafrost carbon upon thaw. Geophys. Res. Lett. **40**: 2689–2693. doi:10.1002/grl.50348

Vonk, J. E., S. E. Tank, W. B. Bowden, and others. 2015. Reviews and syntheses: Effects of permafrost thaw on Arctic aquatic ecosystems. Biogeosciences **12**: 7129–7167. doi:10.5194/bg-12-7129-2015

VUORIO, K., M. MEILI, and J. SARVALA. 2006. Taxon-specific variation in the stable isotopic

signatures (delta13C and delta15N) of lake phytoplankton. Freshw. Biol. **51**: 807–822. doi:10.1111/j.1365-2427.2006.01529.x

von Wandruszka, R. 2000. Humic acids: Their detergent qualities and potential uses in pollution remediation. Geochem. Trans. **1**: 10. doi:10.1039/b001869o

Wang, K., Y. Pang, C. He, and others. 2019. Optical and molecular signatures of dissolved organic matter in Xiangxi Bay and mainstream of Three Gorges Reservoir, China: Spatial variations and environmental implications. Sci. Total Environ. **657**: 1274–1284. doi:10.1016/j.scitotenv.2018.12.117

Weishaar, J. L., G. R. Aiken, B. A. Bergamaschi, M. S. Fram, R. Fujii, and K. Mopper. 2003. Evaluation of Specific Ultraviolet Absorbance as an Indicator of the Chemical Composition and Reactivity of Dissolved Organic Carbon. Environ. Sci. Technol. **37**: 4702–4708. doi:10.1021/es030360x

Wells, M. L. 2004. The colloidal size spectrum of CDOM in the coastal region of the Mississippi Plume using flow field-flow fractionation. Mar. Chem. **89**: 89–102. doi:10.1016/j.marchem.2004.02.009

Wetzel, R. G. 2001. Limnology: Lake and River Ecosystems, 3rd,.

Wisconsin Department of Natural Resources, W. 1984. Chapter 7 Juneau Park Lagoon.

Wünsch, U. J., E. Acar, B. P. Koch, K. R. Murphy, P. Schmitt-Kopplin, and C. A. Stedmon. 2018. The Molecular Fingerprint of Fluorescent Natural Organic Matter Offers Insight into Biogeochemical Sources and Diagenetic State. Anal. Chem. acs.analchem.8b02863. doi:10.1021/acs.analchem.8b02863

Wünsch, U. J., R. Bro, C. A. Stedmon, P. Wenig, K. R. Murphy, R. Bro, and P. Wenig. 2019a. Emerging patterns in the global distribution of dissolved organic matter fluorescence. Anal. Methods **11**: 888–893. doi:10.1039/C8AY02422G

Wünsch, U. J., K. R. Murphy, and C. A. Stedmon. 2017a. The one-sample PARAFAC approach reveals molecular size distributions of fluorescent components in dissolved organic matter. Environ. Sci. Technol. **51**: acs.est.7b03260. doi:10.1021/acs.est.7b03260

Wünsch, U. J., C. A. Stedmon, K. R. Murphy, R. Bro, and P. Wenig. 2019b. Emerging patterns in the global distribution of dissolved organic matter fluorescence. Anal. Methods 888–893. doi:10.1039/c8ay02422g

Wünsch, U. J., C. A. Stedmon, L. J. Tranvik, and F. Guillmette. 2017b. Unraveling the size-dependent optical properties of dissolved organic matter. Limnol Ocean. doi:10.1002/lno.10651

Xu, H., D.-X. Guan, L. Zou, H. Lin, and L. Guo. 2018a. Contrasting effects of photochemical and microbial degradation on Cu(II) binding with fluorescent DOM from different origins. Environ. Pollut. **239**: 205–214. doi:10.1016/j.envpol.2018.03.108

Xu, H., and L. Guo. 2017. Molecular size-dependent abundance and composition of dissolved organic matter in river, lake and sea waters. Water Res. **117**: 115–126. doi:10.1016/j.watres.2017.04.006

Xu, H., and L. Guo. 2018. Intriguing changes in molecular size and composition of dissolved organic matter induced by microbial degradation and self-assembly. Water Res. **135**: 187–194. doi:10.1016/j.watres.2018.02.016

Xu, H., E. M. Houghton, C. J. Houghton, and L. Guo. 2018b. Variations in size and composition of colloidal organic matter in a negative freshwater estuary. Sci. Total Environ. **615**: 931–941. doi:10.1016/j.scitotenv.2017.10.019

Xu, H., H. Lin, H. Jiang, and L. Guo. 2018c. Dynamic molecular size transformation of aquatic colloidal organic matter as a function of pH and cations. Water Res. **144**: 543–552.

doi:10.1016/j.watres.2018.07.075

Xu, H., M. Xu, Y. Li, X. Liu, L. Guo, and H. Jiang. 2018d. Characterization, origin and aggregation behavior of colloids in eutrophic shallow lake. Water Res. **142**: 176–186. doi:10.1016/j.watres.2018.05.059

Xu, H., M. Yan, W. Li, H. Jiang, and L. Guo. 2018e. Dissolved organic matter binding with Pb (II) as characterized by differential spectra and 2D UV e FTIR heterospectral correlation analysis. Water Res. **144**: 435–443. doi:10.1016/j.watres.2018.07.062

Xu, J., M. Zhang, and P. Xie. 2007. Stable carbon isotope variations in surface bloom scum and subsurface seston among shallow eutrophic lakes. Harmful Algae **6**: 679–685. doi:10.1016/j.hal.2007.02.002

Yamashita, Y., and R. Jaffe. 2008. Assessing the dynamics of dissolved organic matter (DOM) in coastal environments by excitation emission matrix fluorescence and parallel factor analysis (EEM-PARAFAC). Limnol. Oceanogr. **53**: 1900–1908.

Yamashita, Y., L. J. Scinto, N. Maie, and R. Jaffé. 2010. Dissolved Organic Matter Characteristics Across a Subtropical Wetland's Landscape: Application of Optical Properties in the Assessment of Environmental Dynamics. Ecosystems **13**: 1006–1019. doi:10.1007/s10021-010-9370-1

Yang, B., H. Lin, S. L. Bartlett, E. M. Houghton, D. M. Robertson, and L. Guo. 2021. Partitioning and transformation of organic and inorganic phosphorus among dissolved, colloidal and particulate phases in a hypereutrophic freshwater estuary. Water Res. **196**: 117025. doi:10.1016/j.watres.2021.117025

Yang, L., W. Chen, W. E. Zhuang, and others. 2019. Characterization and bioavailability of rainwater dissolved organic matter at the southeast coast of China using absorption spectroscopy and fluorescence EEM-PARAFAC. Estuar. Coast. Shelf Sci. **217**: 45–55. doi:10.1016/j.ecss.2018.11.002

Yang, L., D. H. Han, B. M. Lee, and J. Hur. 2015. Characterizing treated wastewaters of different industries using clustered fluorescence EEM-PARAFAC and FT-IR spectroscopy: Implications for downstream impact and source identification. Chemosphere **127**: 222–228. doi:10.1016/j.chemosphere.2015.02.028

Yang, L., H. Hong, W. Guo, C.-T. A. Chen, P.-I. Pan, and C.-C. Feng. 2012. Absorption and fluorescence of dissolved organic matter in submarine hydrothermal vents off NE Taiwan. Mar. Chem. **128**–**129**: 64–71. doi:http://dx.doi.org/10.1016/j.marchem.2011.10.003

Yu, P. 2005. Applications of Hierarchical Cluster Analysis (CLA) and Principal Component Analysis (PCA) in Feed Structure and Feed Molecular Chemistry Research, Using Synchrotron-Based Fourier Transform Infrared (FTIR) Microspectroscopy. J. Agric. Food Chem. **53**: 7115–7127. doi:10.1021/jf050959b

Zhao, Z., M. Gonsior, J. Luek, and others. 2017. Picocyanobacteria and deep-ocean fluorescent dissolved organic matter share similar optical properties. Nat. Commun. **8**: 1–10. doi:10.1038/ncomms15284

Zhou, Z., and L. Guo. 2015. A critical evaluation of an asymmetrical flow field-flow fractionation system for colloidal size characterization of natural organic matter. J. Chromatogr. A **1399**: 53–64. doi:10.1016/j.chroma.2015.04.035

Zhou, Z., L. Guo, and E. C. Minor. 2016a. Characterization of bulk and chromophoric dissolved organic matter in the Laurentian Great Lakes during summer 2013. J. Great Lakes Res. **42**: 789–801. doi:10.1016/j.jglr.2016.04.006

Zhou, Z., L. Guo, A. M. Shiller, S. E. Lohrenz, V. L. Asper, and C. L. Osburn. 2013. Characterization of oil components from the Deepwater Horizon oil spill in the Gulf of Mexico using

fluorescence EEM and PARAFAC techniques. Mar. Chem. **148**: 10–21. doi:10.1016/j.marchem.2012.10.003

Zhou, Z., B. Stolpe, L. Guo, and A. M. Shiller. 2016b. Colloidal size spectra, composition and estuarine mixing behavior of DOM in river and estuarine waters of the northern Gulf of Mexico. Geochim. Cosmochim. Acta **181**: 1–17. doi:10.1016/j.gca.2016.02.032

Zolfaghari, K., C. R. Duguay, and H. K. Pour. 2016. Satellite-Derived Light Extinction Coefficient and its Impact on Thermal Structure Simulations in a 1-D Lake Model. Earth Syst. Sci. Discuss. Earth Syst. Sci 377–391. doi:10.5194/hess-2016-82

Zou, L., M. Y. Sun, and L. Guo. 2006. Temporal variations of organic carbon inputs into the upper Yukon River: Evidence from fatty acids and their stable carbon isotopic compositions in dissolved, colloidal and particulate phases. Org. Geochem. **37**: 944–956. doi:10.1016/j.orggeochem.2006.04.002

Zsolnay, A., E. Baigar, M. Jimenez, B. Steinweg, and F. Saccomandi. 1999. Differentiating with fluorescence spectroscopy the sources of dissolved organic matter in soils subjected to drying. Chemosphere **38**: 45–50. doi:10.1016/S0045-6535(98)00166-0

Curriculum Vitae

Hui Lin

EDUCATION

Ph.D in Aquatic Biogeochemistry May 2021
School of Freshwater Science
University of Wisconsin-Milwaukee, USA
Advisor: Dr. Laodong Guo

M.S. in Marine Chemistry June 2016
Xiamen University, Xiamen, China
Advisor: Dr. Chen Min

B.A. in Marine Chemistry May 2013
Xiamen University, Xiamen, China
Advisor: Dr. Yihua Cai

PEER-REVIEWED PUBLICATIONS

1. **Lin, H**., Bartlett, S.L. and Guo, L. (2021). Variations in DOM molecular size and PARAFAC-derived components during estuarine mixing along the lower Fox River-Green Bay ecosystem. (to be submitted).

2. **Lin, H.**, Guo, L., Miller, T.R., Newton, R.J., (2021) Dynamic cycling of organic matter in a shallow eutrophic lagoon with seasonal *Microcystis* blooms. (to be submitted).

3. **Lin, H.**, Matsui, K., Newton, R.J. and Guo, L. (2021). Changes in composition and molecular size spectra of dissolved organic matter between influent and effluent from a metropolitan wastewater treatment plant. *Environmental Science and Technology Letters*. (submitted)

4. **Lin, H.**, Xu, H., Cai, Y. Belzile, C. Macdonald, R.W. and Guo, L. (2021). Dynamic changes in size-fractionated DOM composition from a seasonally ice-covered Arctic River. *Limnology and Oceanography*. (Accepted).

5. Wang, C.Y., Yang, Y., Yang, B., **Lin, H**., Miller, T.R., Newton, R.J., and Guo, L. (2021). Causal relationship between alkaline phosphatase activities and phosphorus dynamics in a eutrophic coastal lagoon in Lake Michigan. *Science of Total Environments*. doi: 10.1016/j.scitotenv.2021.147681. (IF = 6.51)

6. Yang, B., **Lin, H.**, Bartlett, S.L., Houghton, E.M., Robertson, D.M., and Guo, L. (2021). Partitioning and transformation of organic and inorganic phosphorus among dissolved, colloidal and particulate phases in a hypereutrophic freshwater estuary. *Water Research*. doi:10.1016/j.watres.2021.117025 (IF = 7.91)

7. **Lin, H.**, and Guo, L. (2020). Variations in Colloidal DOM Composition with Molecular Weight within Individual Water Samples as Characterized by Flow Field-Flow Fractionation and EEM-PARAFAC Analysis. *Environmental Science & Technology*, 54(3):1657-1667. (IF = 7.27)

8. Zeng, J., Chen, M., Guo, L., **Lin, H.**, Mu, X., and Fan, L. (2019). Role of organic components in regulating denitrification in the coastal water of Daya Bay, southern China. *Environmental Science: Processes & Impacts*, 21(5), 831–844. (IF = 3.24)

9. Xu, H., Lin, H., Jiang, H., and Guo, L. (2018). Dynamic molecular size transformation of aquatic colloidal organic matter as a function of pH and cations. *Water Research*, 144, 543–552. (IF = 7.91)

10. Xu, H., Guan, D.X., Zou, L., **Lin, H.**, and Guo, L. (2018). Contrasting effects of photochemical and

microbial degradation on Cu(II) binding with fluorescent DOM from different origins. *Environmental Pollution*, 239, 205–214. (IF = 5.71)

11. Li, Q., Chen, M., Jia, R., Zeng, J., **Lin, H.**, Zheng, M., and Qiu, Y. (2017). Transit time of river water in the Bering and Chukchi Seas estimated from $\delta^{18}O$ and radium isotopes. *Progress in Oceanography*, 159, 115–129. (IF = 4.06)

12. **Lin, H.**, Chen, M., Zeng, J., Li, Q., Jia, R., and Sun, X. (2016). Size characteristics of chromophoric dissolved organic matter in the Chukchi Sea. *Journal of Geophysical Research: Oceans*, 121(8), 6403–6417. (IF = 3.56)

13. **Lin, H.**, Cai, Y., Sun, X., Chen, G., Huang, B., Cheng, H., and Chen, M. (2016). Sources and mixing behavior of chromophoric dissolved organic matter in the Taiwan Strait. *Marine Chemistry*, 187, 43–56. (IF = 2.93)

PRESENTATIONS/CONFERENCE ABSTRACTS

1. **Lin, H.** and Guo, L. 2021. Dynamic changes in size-fractionated DOM composition in the seasonally ice-covered Yukon River as characterized using EEM-PARAFAC, FT-IR and data fusion. Abstract submitted to 2021 ASLO Aquatic Sciences Meeting, Jun 22-27, 2021. Virtual meeting.
2. Guo, L., **Lin, H**., Miller, T.R., and Newton, R.J. 2021. Response of dissolved organic matter dynamics to algal blooms in a coastal eutrophic lagoon in Lake Michigan. Abstract submitted to 2021 ASLO Aquatic Sciences Meeting, Jun 22-27, 2021. Virtual meeting.
3. **Lin, H.** and Guo, L. 2020. Variations in size and fluorescent components of dissolved organic matter in a negative estuary. 2020 Ocean Science Meeting, February 16-21, 2020, San Diego, CA. Abstract#: CT44B-1005 (Poster presentation).
4. **Lin, H.** and Guo, L. 2019. Data fusion applications: Interpreting DOM properties from coupling Fourier Transform infrared (FT-IR) spectra and fluorescence excitation emission matrices (EEMs), Independent study public presentation, December 11, 2019, School of Freshwater Sciences, University of Wisconsin-Milwaukee, WI. (Oral presentation).
5. **Lin, H.** and Guo, L. 2019. Molecular size distribution and size-dependent composition of a single DOM sample as characterized using FlFFF-EEM-PARAFAC coupling techniques. The fourth Xiamen Symposium on Marine Environmental Sciences, January 6-9 (8), 2019. Xiamen, China (Oral presentation).
6. **Lin, H.,** Joung, D.J., Kessler, J., and Guo, L. 2018. Composition and size-distribution of dissolved organic matter in the two largest Great Lakes: Lakes Superior and Michigan. 2018 Ocean Science Meeting, February 11-16, 2018, Portland, OR, Abstract#: CT24B-1316 (Poster presentation).
7. **Lin, H**. and Guo L. 2017. Optical and size characterization of dissolved organic matter from the lower Yukon River. 2017 AGU Fall Meeting, December 11-15, 2017, New Orleans, LA. Abstract# 266361 (Oral presentation).
8. **Lin, H.** and Chen, M. 2017. Size-fractionated dissolved organic nitrogen profile in the Chukchi Sea and Canada Basin. The third Xiamen Symposium on Marine Environmental Sciences, Student Session, January 9-11, 2017. Xiamen, China (Oral presentation).
9. **Lin, H.,** Chen, M., Zeng, J. 2016 Size characteristics of chromophoric dissolved organic matter in the Chukchi Sea. The Arctic Science Summit Week (ASSW), Student Session. March 8-11, 2016, Fairbanks, AK. (Poster presentation).

AD-HOC MANUSCRIPT REVIEWERS

-Scientific Reports.

-Limnology and Oceanography Methods

OUTREACH ACTIVITIES

- Volunteer for Open Doors Milwaukee at School of Freshwater Sciences, University of Wisconsin-Milwaukee - September 28, 2019.
- Volunteer for Lake Sturgeon Bowl Competition at School of Freshwater Sciences, University of Wisconsin-Milwaukee - 2018 and 2019.

OTHER SKILLS

- **Laboratory Instrument**:
  - Asymmetric Flow Field-Flow Fractionation (AF2000, Postnova)
  - Spectrophotometer (Agilent 8453 and Shimadzu UV2450)
  - Spectrofluorometer (Fluoromax-4, Horiba and Varian)
  - Liquid Scintillation Counter (TriCarb 2900TR, PerkinElmer)
  - TOC V_{CPH} Analyzer (Shimadzu)
  - FT-IR (IRTracer-100, Shimadzu)
  - Knowledge of EA-IRMS (Carlo Erba NC2500-Finnigan MAT Delta^{plus} XP).
- **Computer Programs/Softwares**:
  - Matlab
  - R software
  - ODV
  - Origin
  - Surfer
  - Grapher
  - Sigmaplot
  - MS Office
  - Photoshop
- **Computer Languages**:
  - Microsoft Visual Basic (VB), HTML, Python.
  -

HONORS AND AWARDS

- Distinguished Dissertator Fellowship (University of Wisconsin-Milwaukee) 2020-2021
- Graduate Student Excellence Fellowship (University of Wisconsin-Milwaukee) 2019-2020
- Distinguished Graduate Student Fellowship (University of Wisconsin-Milwaukee), 2017-2018
- Fellowship from China Scholarship Council, 2016-2020
- First State University Scholarship & TA fellowship (Xiamen University), 2013-2016
- First State University Scholarship (Xiamen University), 2010-2013
- The honor of "3A Students" (Xiamen University), 2010-2013