Spring 1-1-2013

Characterization and Differentiation of Wastewater Effluent Organic Matter (EfOM) Versus Drinking Water Natural Organic Matter (NOM): Implications for Indirect Potable Reuse

Seong-Nam Nam
University of Colorado at Boulder, namsn76@gmail.com

CHARACTERIZATION AND DIFFERENTIATION OF WASTEWATER EFFLUENT ORGANIC MATTER (EfOM) VERSUS DRINKING WATER NATURAL ORGANIC MATTER (NOM): IMPLICATIONS FOR INDIRECT POTABLE REUSE

By

SEONG-NAM NAM

B.S., YONSEI UNIVERSITY, SOUTH KOREA, 1999

M.S., YONSEI UNIVERSITY, SOUTH KOREA, 2001

A Thesis submitted to the Faculty of Graduate School of the University of Colorado in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Civil, Environmental, & Architectural Engineering

2011

This Thesis Entitled:

Characterization and Differentiation of Wastewater Effluent Organic Matter (EfOM) versus Drinking Water Natural Organic Matter (NOM): Implications for Indirect Potable Reuse

Written by Seong-Nam Nam
Has been approved for the Department of Civil, Environmental & Architectural Engineering

_____
(Gary Amy, Committee Chair)

_____
(JoAnn Silverstein, Committee Member)

_____
(Diane McKnight, Committee Member)

_____
(Joerg Drewes, Committee Member)

_____
(R. Scott Summers, Committee Member)

Date: _____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

Nam, Seong-Nam (Ph.D., Civil, Environmental, & Architectural Engineering)

Seong-Nam Nam (Ph.D., Civil Engineering)

Characterization and Differentiation of Wastewater Effluent Organic Matter (EfOM) versus Drinking Water Natural Organic Matter (NOM): Implications for Indirect Potable Reuse

Dissertation directed by Professor Gary L. Amy

ABSTRACT

A major objective of this research was to investigate the characteristic differences/similarities between wastewater effluent organic matter (EfOM) and drinking water natural organic matter (NOM) derived from diverse water sources using multiple analytical and statistical techniques. Parallel factor analysis (PARAFAC), a three-dimensional statistical modeling technique, was applied to a dataset of excitation-emission matrix (EEM) of fluorescing organic matter components in order to identify organic matter fluorophores characteristic of EfOM versus NOM. Identified fluorophores were used to elucidate characteristic features of bulk organic matter as well as treatability of EfOM(-impacted waters) by conventional drinking water treatment processes. Results show that EfOM exhibits increased fractions of microbial-originated organic components of a hydrophilic nature. Size exclusion chromatographic (SEC) characteristics of EfOM and NOM reveal that EfOM is a mixture of various organic components which have different molecular weight distributions and light-absorbing properties. Incorporation of traditional NOM characterization results with multivariate statistical analyses shows that a hydrophilic fraction is a key discriminator of EfOM that differentiates it from NOM, and thus, the organic properties of any water source influenced by wastewater may be shifted to a more hydrophilic nature.

The evaluation of wastewater impacts on existing drinking water treatment plant in terms of treatability of an organic matter was also an aim of this study. Lab-scale simulations of the coagulation process suggest EfOM has potentially negative impacts on drinking water treatment efficiency due to changed organic properties of source waters. Consequently, adoption of a biodegradation process such as soil aquifer treatment (SAT) or river bank filtration (RBF) is proposed as a strategic management.

DEDICATION

To the deceased father

&

to my mother, Ms. Choon-Ok Kim

ACKNOWLEDGEMENTS

In his heart a man plans his course, but the LORD determines his steps. (Proverbs 16:9)

First of all, I thank God for being with me and helping me in everything necessary until now. I earnestly hope in my lifetime that I kneel down to acknowledge your way, and that I will be able to return the faithfulness that he has shown me. I would like to express my deepest thanks to my research advisor, Gary Amy, for providing me research opportunities and for his scientific insights, sage advice, encouragements and endless-like patience. Also, I thank my thesis committee-Diane McKnight, JoAnn Silverstein, Joerg Drewes, and Scott Summers for all the suggestions and comments that made me more grow as an engineer/scientist. I express my gratitude to Stuart Krasner (MWDSC) and all the AwwaRF #2948 research project crew (Paul Westerhoff, Baiyang Chen, Bruce E. Rittmann, Project manager-Alice Fulmer) for scientific discussion and many helps. I also acknowledge Korean government for providing me a Study Abroad scholarship in addition to an AWWA funding. To all the graduate students, staff and faculty that I have met or co-worked at CU and UNESCO-IHE (the Netherlands) over the past several years also go my thanks. Of them, I especially thank Rose Cory and Natalie Miladenov for their collaborations on fluorescence/PARAFAC. I also have a thankful heart to my two computers which had been always busy and too much tortured to run the model.

I am grateful with my church (UBF) members/friends for their prayers: Pastor David Kim & Sh. Sarah Kim (Korea), Msn. David & Ruth Kim (Tempe), Msn. Young & Curie Lee (Chicago), Sh. Maria Joo (Korea), and Denver UBF people. I thank Jeongsook and Mary Lee for their friendships, prayer supports and many laughing. I owe much thanks to all Yonhee UBF coworkers. Especially, I would like to say my special thanks to Seo-Hyun for supporting me with trust. Lastly, but not least, I would like to express my love and gratefulness to my mom and my family who have always believed in me and unfailingly supported me. Mom, without your tears and love for me, I would not have successfully finished this race. I was always thankful for being your daughter. Thank you

TABLES OF CONTENTS

LIST OF SYMBOLS AND ABBREVIATIONS

ANOVA: analysis of variance

AZ: Arizona State

BAP: biomass associated product

$BDOC_5$: biodegradable dissolved organic carbon at day 5

CA: Cluster Analysis

CA: California State

CO: Colorado State

COD: chemical oxidation demand

Da: Dalton(s), g/mol

DBPFPs: disinfection by-products formation potentials

DBPs: disinfection by-products

DIN: dissolved inorganic nitrogen

DOC: dissolved organic carbon

DOM: dissolved organic matter

DON: dissolved organic nitrogen

DW: drinking water

DWTP: drinking water treatment process

EDCs: Endocrine disrupting compounds

EEM: (fluorescence) excitation-emission matrix

EfOM: effluent organic matter

Em: fluorescence emission wavelength, nm

Ex: fluorescence excitation wavelength, nm

FI: fluorescence index, the ratio of fluorescence intensity of emission 450 to emission 500 nm at excitation 370 nm

HANFPs: haloacetonitrile formation potentials

HIX: humidification index

HMW: high molecular weight

HPI(A): hydrophilic (acid)

HPO(A): hydrophobic (acid)

HPSEC: high performance size exclusion chromatography

HS: humic substance

LMA: low molecular weight acid

LC-OND: Liquid chromatography organic nitrogen detection (or detector)

MI: Michigan State

MLR: Multiple Linear Regression

M_n: the number-average molecular weight

M_w: the weight-average molecular weight

MOC: mean oxidation number of carbon

MW: molecular weight

MWD: molecular weight distribution

NDMAFPs: *N*-nitrosodimethylamine formation potentials

NJ: New Jersey State

NOM: Natural Organic Matter

NRC: National Research Council

OH: Ohio State

OM: organic matter or organic material

PA: Pennsylvania State

PARAFAC: parallel factor analysis

PCA: principal component analysis

PEG: polyethylene glycol

PS: polysaccharide

PSS: polystyrene sulfonate

RBF: river bank filtration

RO: reverse osmosis

Rt: retention time (second or minute)

SAT: soil aquifer treatment

SMPs: soluble microbial products

SRFA: Suwannee River fulvic acid

SRHA: Suwannee River humic acid

SRNOM: Suwannee River natural organic matter

SRT: sludge retention time

SUVA: specific UV absorbance, usually at 254 nm, $(UVA_{254}/DOC) \times 100$, L/mg·m

TIN: total inorganic nitrogen

TPI(A): transphilic (acid)

TX: Texas State

UAP: utilization associated product

URI: UV ratio index, UVA_{210}/UVA_{254}

USA or US: United States of America or United States

U.S. EPA: United States Environmental Protection Agency

UVA_{254} or UV_{254}: ultraviolet absorbance at 254 nm, cm^{-1}

WRP: wastewater reclamation plant

WW: wastewater

WWTP: wastewater treatment plant

ε: molar absorptivity, L/mole·cm or $M^{-1}cm^{-1}$

ρ: polydispersity, M_w/M_n

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Problem statement

As the population growth and rapid urban development has been increased, the available potable water decreased, and consequently, the water demands (domestic, industrial, commercial, and agricultural purposes) increased. In order to handle increased water demand, the wastewater should be reused. As an approach to improve the water supplies, recycling wastewater for potable reuse have been assessed by the National Research Council (NRC). The NRC reported that planned, indirect potable reuse, which was defined as reclaimed water, was a viable application of reclaimed water (NRC, 1998). Water reuse in United States (US) is a large and growing practice. An estimated 1.7 billion gallons (6.4 million m^3) per day of wastewater (WW) is re-used, and reclaimed water use on a volume basis is growing at an estimated 15% per year according to the "*Guidelines for Water Reuse*" (EPA 2004, EPA/625/R-04/108).

Indirect potable reuse of wastewater occurs when the treated wastewater treatment plant (WWTP) effluent is discharged into rivers or lakes that are subsequently used as downstream drinking water sources. However, one of major concerns regarding the use of reclaimed water is the health risks associated with harmful disinfection by-products (DBPs) precursors, emerging contaminants or synthetic organic compounds (e.g., endocrine disrupting compounds (EDCs), pharmaceuticals and personal care products (PPCPs), hormones, etc.), and the contribution of

WW toward those compounds. In spite of many efforts of researchers to explicate the properties of EfOM matrix, still the problems of unknown organic species has been remained, and poor understanding on the EfOM may increase health risks.

This research attempts to elucidate the property of wastewater effluent organic matter found in 24 wastewater treatment plants (WWTPs) with different treatment processes, distinguishing it from pristine or drinking water natural organic matter.

Objectives of research

The researches have been performed with the following objectives;

1. To develop the EfOM-focused fluorophores' model using fluorescence spectroscopy combined with the parallel factor analysis (PARAFAC), in which contains fluorophore(s) that can be used as wastewater EfOM fingerprint (Chapter 3), and to compare the developed PARAFAC model to the existing model in order to validate the developed model in this study (Chapter 4)
2. To address the differentiation/similarity of WW EfOM from/to DW NOM by characterizing them using the multiple analytical techniques (Chapter 5)
3. To demonstrate the EfOM characteristics differentiating from NOM via incorporating the analytical parameters and multivariate statistical analysis (Chapter 6)
4. To investigate the EfOM impact on DWTP using conventional DWT process and to propose the alternative strategies for EfOM control in WW reuse (Chapter 7)

Research hypotheses

Based on research objectives, three hypotheses have been proposed:

Hypothesis #1:

WW EfOM has a distinct fluorescing signature that enables it to use as WW fingerprint to assess the WW impact.

Hypothesis #2:

Differentiation/similarity between EfOM versus NOM can be identified (specified) in terms of their physico-chemical properties via multiple analytical/statistical techniques.

Hypothesis #3:

EfOM has negative impacts (e.g., removal efficiency) on conventional drinking water treatment plants due to its different properties to NOM (e.g., more hydrophilic than NOM)

Organization of dissertation

This dissertation provides concepts and further understanding on drinking water NOM and wastewater EfOM as well as analytical approaches and an assessment of EfOM impact on drinking water treatment process in terms of removal efficiency. All chapters except chapters 1-2 have been written in a journal submission format.

Chapter 2 provides detail descriptions on the experimental methods, materials and analytical instrumentations including figure schematics of set-up.

Chapter 3 discusses the identification of wastewater EfOM fingerprint using fluorescence excitation-emission matrix (EEM) coupled with parallel factor analysis (PARAFAC). This

chapter explains a development of the EfOM-focused 5 components PARAFAC model using 423 EEMs from diverse sources of water (i.e., surface water, wastewater sources along with wastewater impacted sites), and describes how wastewater EfOM is different from NOM and can be used as wastewater fingerprint or as an assessing tool of wastewater impact in the watershed (e.g., South Platte River as a case study site) experiencing wastewater discharge. Application of 5 PARAFAC components to biodegradation is also introduced for understanding of EfOM biodegradability.

Chapter 4 describes additional validation of the 5 components identified in Chapter 3 through fitting to the NOM-centered existing model (i.e., 13 components model). The chapter discusses how fluorophores can be composites, and be resolved into separate or sub-fluorophores. The results in the chapter also provide additional information on both 5 components and 13 components through multivariate statistical analyses such as correlation analysis and principal component analysis.

Chapter 5 elucidates detail characteristics of EfOM using multiple spectroscopic and chromatographic analytical approaches. SEC-DOC-UVA-fluorescence detection techniques for several NOM and EfOM reference substances are employed to explore differentiation/similarity between NOM and EfOM (e.g., molecular weight distribution (MWD) and spectroscopic patterns). Fluorescence EEMs incorporated with EfOM and NOM isolates of hydrophobic (HPO), transphilic (TPI) and hydrophilic (HPI) materials explains the potent relationship of (physico-) chemical compositions/properties with spectroscopic trends. The chapter shows incorporations of spectroscopic and chromatographic techniques as a powerful tool for NOM and/or EfOM understanding.

Chapter 6 assesses the similarities or differentiations of NOM and EfOM by incorporating the analytical parameters into the multivariate statistical analyses such as principal component analysis (PCA), multivariate linear regression analysis (MLR), etc. This chapter shows that hydrophlilicity (HPI) as the key discriminating parameter of EfOM from NOM, and organic matter having hydrophilic or Transphilic properties as the most contributing to DOC of EfOM.

Chapter 7 evaluates treatability of EfOM by conventional drinking water treatment processes (e.g., coagulation) in terms of organic matter removal efficiency, and addresses possible negative impacts on DWTP, which may cause increased disinfection by-products (DBPs) or decreased bio-stability in the distribution system. This chapter introduces alternative strategies for EfOM control in water reuse for potable purpose (e.g., river bank filtration).

Chapter 8 includes conclusions of the dissertation based on previous research chapter, and provides the future research suggestion/directions.

Finally, appendices are added to supplement data that are mentioned as additional information in chapters.

CHAPTER 2

SAMPLING INFORMATION AND ANALYTICAL METHODS

Detail descriptions on studied sampling locations

WWTPs and DWTPs studied

Figure 2-1 shows process schematic diagrams for some generic examples of the types of WWTPs that were studied. Table 2-1 shows a summary of the unit processes that were used at the participating WWTPs in this study. Where possible, two or more WWTPs that treated wastewater from the same watershed or geographical region but used different treatment processes were sampled. In some cases, this was achieved at WWTPs with parallel treatment trains. Secondary treatment was typically accomplished through trickling filters or activated sludge. In addition, one WWTP used an aerated lagoon and another used an oxidation ditch. Some of the WWTPs had advanced biological treatment (nitrification with or without denitrification). In some cases (e.g., during the warmer months), WWTPs that were not designed for nitrification did produce well-nitrified EfOMs. In addition to biological treatment, some WWTPs used physical/chemical treatment processes (e.g., flocculation, lime softening, filtration, powdered and granular activated carbon (PAC, GAC)). Other advanced WWTP processes included membranes (membrane bioreactor (MBR), microfiltration (MF) and/or reverse osmosis (RO)) and/or soil aquifer treatment (SAT).

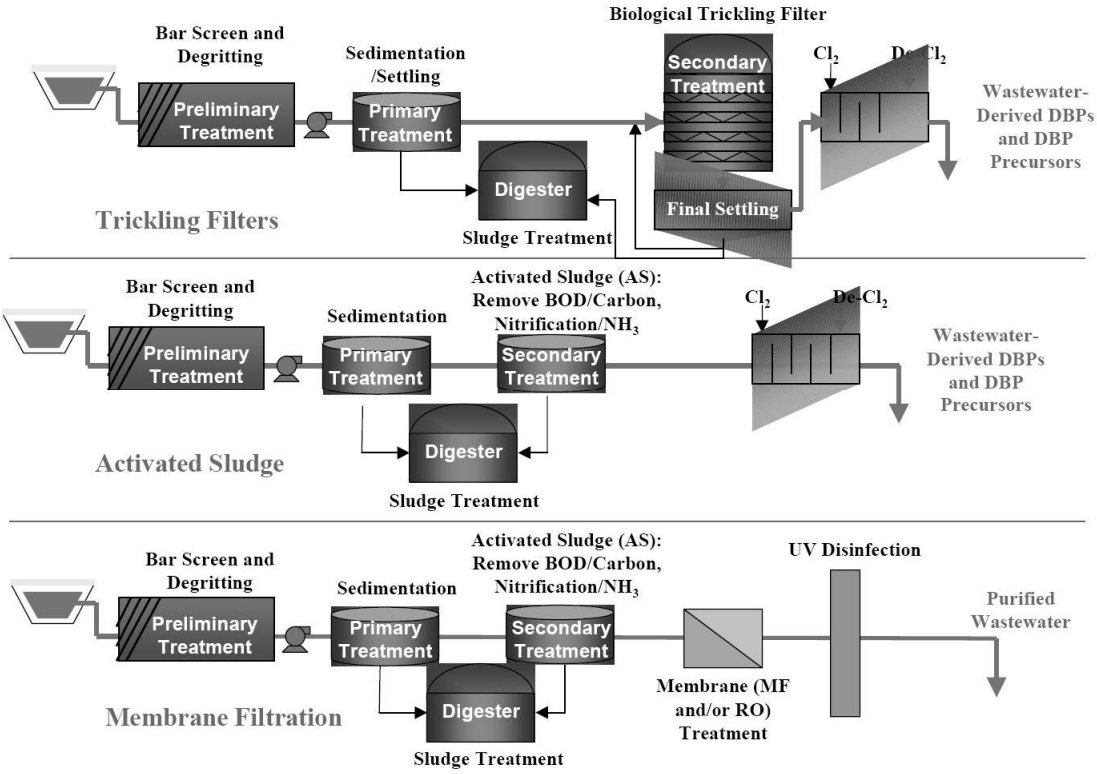


Figure 2-1 Examples of the types of WWTPs studied

Table 2-1 Unit processes at the participating WWTPs

Participating Utilities	Secondary Treatment	Advanced Biological Treatment	Physical/Chemical Treatment	Membranes	SAT	Disinfectant
West Region:						
West Basin Water Recycling Plant (CA):						
Title 22 train	X*		flocculation, filtration			chlorine
barrier trains	X*		lime softening, filtration	RO		chlorine
	X*			MF, RO		chlorine
Interim Water Factory 21/water recycling facility (Orange County Water District, CA):						
spring 2004	X*			MF		chlorine
fall 2004, summer 2005	X*			MF, RO		chlorine, UV + H_2O_2
Southwest Region:						
Nogales (Ariz.) International WWTP	aerated lagoon					chlorine
Roger Road facilities (Tucson, AZ):						
WWTP (Pima County)	trickling filters					chlorine
recharge facility (Tucson Water)	X*		filtration		X	chlorine
Anthem (AZ) WWTP		nitrification, denitrification		MBR		chlorine
Northwest Water Reclamation Plant (Mesa, AZ)	activated sludge	nitrification, denitrification	filtration		X	chlorine or UV
Mountain Region:						
Littleton/Englewood (CO) WWTP	trickling filters	Nitrification				chlorine
Denver Metro (CO) Central Treatment Plant:						
South Complex	activated sludge†					chlorine
North Complex	activated sludge	nitrification, denitrification				chlorine + ammonia
Lone Tree Creek WWTP (Arapahoe County Water and Wastewater Authority, CO)				MBR		UV

Table 2-1 Unit processes at the participating WWTPs (Continued)

Participating Utilities	Secondary Treatment	Advanced Biological Treatment	Physical/Chemical Treatment	Membranes	SAT	Disinfectant
South Central Region:						
El Paso Water Utilities (TX):						
Northwest WWTP	activated sludge	Nitrification	filtration			UV
Haskell R. Street WWTP	activated sludge	Nitrification				chlorine
Fred Hervey Water Reclamation Plant	activated sludge	nitrification, denitrification	PAC, lime softening, GAC			ozone, chlorine
Midwest Region:						
Northeast Ohio Regional Sewer District (OH):						
Westerly WWTP	trickling filters					chlorine[‡]
Easterly WWTP	activated sludge	nitrification, denitrification				chlorine[‡]
Southerly WWT Center	activated sludge	Nitrification	filtration			chlorine[‡]
Detroit (MI) WWTP	activated sludge[†]					chlorine
Northeast Region:						
Two Bridges Sewer Authority (NJ) WWTP	activated sludge	Nitrification	pressure filtration			chlorine
WWTP in Schuylkill River watershed (PA)	activated sludge	Nitrification				UV
Northeast Water Pollution Control Plant (Philadelphia, PA)	activated sludge					chlorine

[*]Influent to water recycling plant = secondary effluent of a conventional WWTP
[†]pure oxygen-fed activated sludge
[‡]During the winter, the treated wastewater was not chlorinated

Table 2-2 Information on the participating DWTPs

Participating Utilities	Impacted by WWTP	Conventional DWTP	Other Physical/ Chemical Treatment	Disinfectants
Val Vista WTP (Phoenix, AZ)	upstream of WWTPs*	X		chlorine
North Texas Municipal Water District DWTP #3	downstream of Wilson Creek Regional WWTP (Lake Lavon)	X		chlorine
Jonathan W. Rogers WTP (El Paso, TX)	downstream of Haskell R. Street WWTP	flocculation, sedimentation	lime softening, GAC filtration	2-stage ozone, chlorine
Garrett A. Morgan WTP (Cleveland, OH)	downstream of Westerly and Easterly WWTPs (Lake Eerie)	X		chlorine
Northeast DWTP (Detroit, MI)	upstream of Detroit WWTP (Detroit River)	X		chlorine
Southwest DWTP (Detroit, MI)	downstream of Detroit WWTP (Detroit River)	X		chlorine
Little Falls WTP (Passaic Valley Water Commission, NJ)	downstream of WWTPs on Passaic and Pompton Rivers (including Two Bridges Sewer Authority)	X		chlorine†
Queen Lane WTP (Philadelphia, PA)	downstream of WWTPs in Schuylkill River watershed (one of which was studied)	X		chlorine, chloramines
Baxter WTP (Philadelphia, PA)	impacted by Northeast Water Pollution Control Plant under some flow conditions (Delaware River)	X		chlorine, chloramines

*Treated drinking water from the Val Vista WTP is the carrier (baseline, background) water for the local WWTPs (e.g., 91st Avenue WWTP, Northwest Water Reclamation Plant)

†Part of Little Falls WTP treated with ozone during summer 2004 sample event

At the participating WWTPs, chlorine (gas or hypochlorite solution) or UV disinfection was utilized. At two WWTPs, ammonia was added with the chlorine (to well nitrified EfOMs) to form chloramines. At many of the plants that used chlorine, the plant effluent was dechlorinated prior to discharge. At one WWTP, UV plus hydrogen peroxide (H_2O_2) was used as an advanced oxidation process. At another WWTP, ozone was utilized.

Table 2-2 shows a summary of the participating DWTPs in this study. In addition to listing treatment processes and disinfectants used, information is provided on whether a DWTP may be impacted by an upstream WWTP. Note, the location of a WWTP upstream of a DWTP does not necessarily mean there is a significant impact of the WWTP on the DWTP. That is dependent on the dilution factor of the EfOM in the receiving water.

Rivers and groundwaters

In addition to WWTPs and DWTPs, selected rivers and groundwater impacted by wastewater discharge or recharge, respectively, were studied (summarized in Table 2-3).

Field survey

Samples were collected during a wet/cold season (winter or early spring 2004) and a dry/warm season (summer or early fall 2004), and once more in a second year (winter, spring or summer 2005). The two sampling events in year one were based on hydrology and treatment considerations. In general, in the summer, river flow is low, so some streams are more effluent-dominated; and there is more nitrification at the WWTP. In the winter there is more flow and less nitrification. These two seasons were selected to show the different impacts of hydrology and treatment. In year two, selected utilities were re-sampled in the season that provided

especially informative data for that system to ascertain temporal (year-to-year) variations or a third season was selected to better understand seasonal variability.

Table 2-3 Information on rivers and groundwater sites that were studied

River and Groundwater	Impacted by WWTP
monitoring well (Orange County, CA)	groundwater recharged with reclaimed wastewater
Santa Ana River (CA)	consists primarily of tertiary treated wastewater from upstream WWTP discharges
Santa Cruz River (AZ)	downstream of Nogales International WWTP
monitoring and extraction wells (Tucson, AZ)	groundwater recharged with reclaimed wastewater (Roger Road recharge facility)
monitoring wells (Mesa, AZ)	groundwater recharged with reclaimed wastewater (Northwest Water Reclamation Plant)
monitoring well (Scottsdale, AZ)	groundwater recharged with reclaimed wastewater (Scottsdale Water Campus)
South Platte River (CO)	Upstream and downstream of Littleton/Englewood and Denver Metro WWTPs
monitoring and extraction wells (El Paso, TX)	groundwater recharged with reclaimed wastewater (Fred Hervey Water Reclamation Plant)
Detroit River (MI).	Upstream of Detroit WWTP
Pompton River (NJ)	Upstream of Two Bridges Sewer Authority WWTP
Delaware River	Upstream of WWTP influences

For conventional WWTPs, the secondary effluent before and after chlorination were sampled. The sample before chlorination represented the background amount of DBPs in the treated wastewater. In addition, another aliquot of the sample before chlorination was chlorinated under FP conditions to determine the level of DBP precursors in the treated wastewater. For WWTPs with advanced wastewater treatment processes, samples were collected before and after each major unit process to evaluate their ability to remove DBP precursors.

For DWTPs impacted by WWTP discharges or groundwaters that were recharged with reclaimed wastewater, the primary sampling location was the DWTP influent or well, respectively. Selected rivers were sampled upstream of WWTPs. For DWTPs with advanced treatment processes (e.g., ozone), samples were collected after such unit processes to determine the impact of those processes (e.g., DBPFP) of the water.

Analyses performed by the University of Colorado

Sample handling and filtration

The ice chest containing 4-L sample bottles was shipped to the utilities. Collected samples were overnight-shipped to the University of Colorado, laboratory. Immediately or at least in a day, samples were filtered through pre-rinsed Millipore nylon or cellulose acetate 0.45 μm filter(s). Filtered water samples were used for measuring the water qualities parameters, except BDOC. If not analyzed immediately, water samples were kept at below 4 ^{o}C until being analyzed.

Ultraviolet and visible light absorbance (UVA) and scanning

UV spectrophotometer Shimadzu UV-VIS Model UV-160/CL-750 was used for UV absorbance measurement of 0.45 μm pre-filtered sample. Samples were placed into a 1-cm quartz cell and measured at selected wavelengths.

Dissolved organic carbon (DOC)

The Sievers Total Organic Carbon Analyzer (Model 800) is used for DOC measurement. The analyzer is based on the oxidation of organic compounds to form CO_2 using UV radiation and a chemical oxidizing agent (ammonium persulfate). Carbon dioxide is measured using a membrane permeation/conductivity detector. The Model 800 TOC can monitor water samples ranging form high-purity water containing < 0.05 ppb carbon to water samples containing up to 50 mg/L of TOC.

The measurements of TOC and DOC are based on calibration with potassium hydrogen phthalate (KHP) standards. Using a certified grade KHP, calibration is based on a standard curve

between the ranges zero to 10 mg/L as C. Samples for DOC are filtered through 0.45 μm filter and acidified to pH ≤ 2 within 48 hours of collection, and stored at ≤ 4 °C [1].

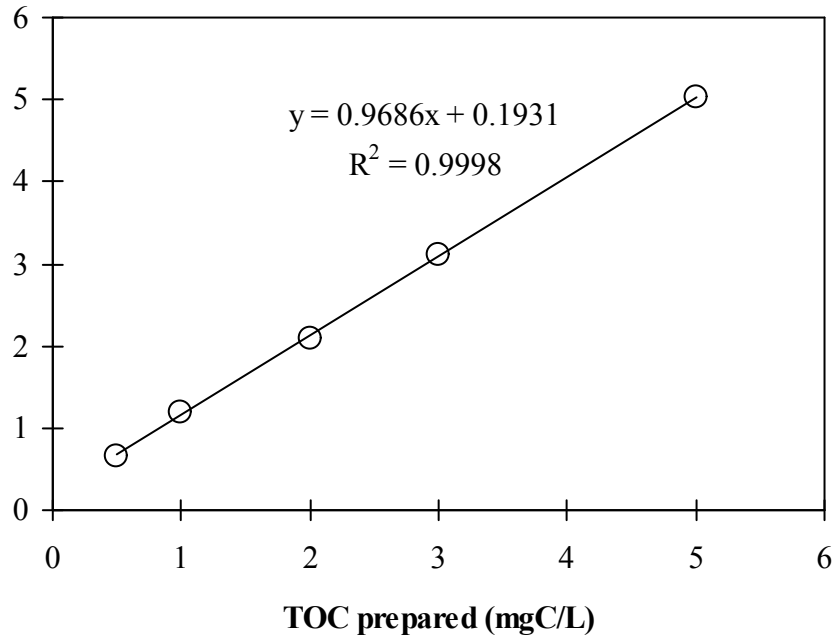


Figure 2-2 Standard calibration of TOC analyzers: a linear relation between prepared DOC versus measured DOC

Molecular weight distribution (High Performance Size Exclusion Chromatography (HPSEC) with DOC and UVA detection)

Molecular weight (MW) distributions are determined by a HP-SEC (High Performance Size Exclusion Chromatography) method. A high performance liquid chromatograph (HPLC, LC600 Shimadzu) is used with UVA (SPD-10Avp Shimadzu) and on-line DOC detectors (modified Sievers Turbo Total Organic Carbon Analyzer) (Figure 2-4). The on-line DOC detector is capable of providing molecular weight distribution of non-aromatic carbon compounds as well as aromatic carbon compounds. UVA and DOC data were acquired at every 2 seconds or 6 seconds by the modified Labview software (National Instrument Inc.).

The system used a TSK HW-50S column (ID 2 cm × Length 25 cm, 35 μm Toyopearl HW

resin) and the flow rate was 1 mL/min. The mobile phase was prepared with Milli-Q water buffered with phosphate (0.0024 M NaH_2PO_4 + 0.0016 M Na_2HPO_4, pH 6.8) and 0.025 M Na_2SO_4, producing 0.1 M of an ionic strength. All samples were filtered through 0.45 μm polycarbonate membrane filter, and conductivity of sample was adjusted to the same level as that of the mobile phase before injection. The sample injection volume was 2 mL. A detailed description of HPSEC instrumentation can be found in Her *et al* [2].

The column separates compounds on the basis of hydrodynamic molecular size. Molecules that are larger than the average pore size of the column packing material pass through more readily, and thus are eluded first (i.e., a shorter retention time corresponds with a larger MW). Smaller molecules permeate through the pores of the column packing and thus correspond with longer retention times. The molecular weight was calibrated with the polyethylene glycol (PEG).

To check the response sensitivity of the area with respect to changes of DOC concentration, potassium hydrogen phthalate (KHP) was used and the very good linear relation was showed (Figure 2-5).

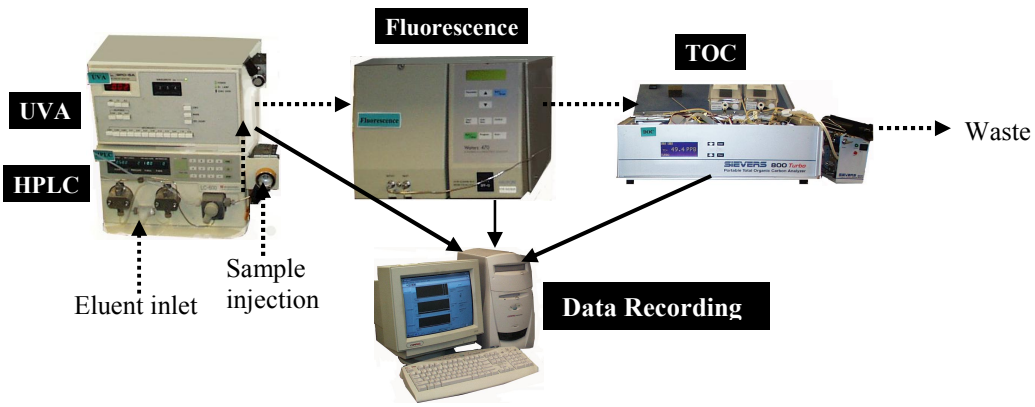


Figure 2-3 HPSEC-DOC/UVA/Fluorescence detection system (adapted from ref [3]), But in this study, SEC/fluorescence detection was not used quite often

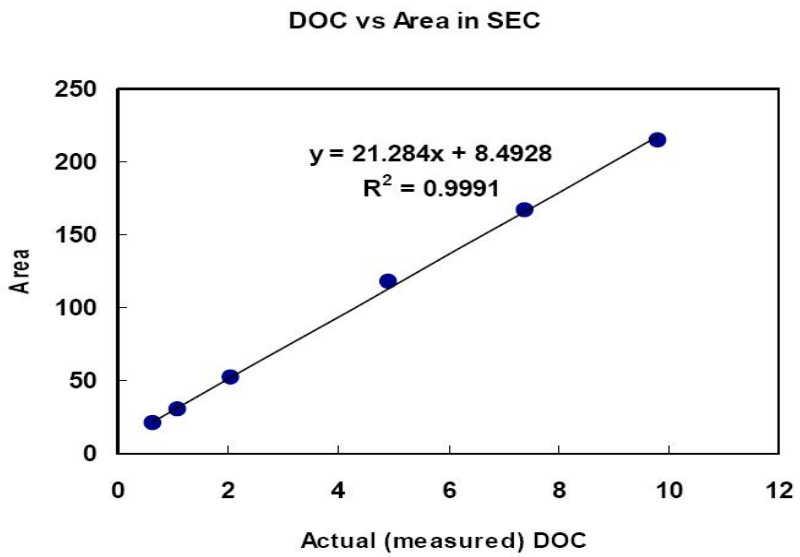


Figure 2-4 A Linear Relation between DOC and the integrated area in HP-SEC

Non-ionic resin chromatography

NOM and EfOM isolation was performed by the use of non-ionic macroporous Amberite XAD-8 (acrylic ester, specific surface area = 140 m^2/g) and XAD-4 (styrene-divinylbenzene, specific surface area = 750 m^2/g) resins as the method described in Aiken *et al* [4]. Two glass chromatography columns with Teflon fittings housed the two resins. 0.45 μm filtered samples were acidified to pH=2 prior to application to the columns. 2 N phosphoric acid (H_3PO_4) was used to acidify samples since a presence of an excessive chloride has an impact on DOC results by reacting with persulfate during persulfate-UV light oxidation and subsequent detection in Sievers TOC 800 analyzer [5].

The organic matter found within the samples was separated into three defined groups: 1) hydrophobic (HPO) compounds, which adsorb onto XAD-8 resin, 2) transphilic (TPI) compounds, which adsorb onto XAD-4 resin, and 3) hydrophilic (HPI) compounds, which pass through both resins. The quantity of organic matter in each fraction was determined by DOC

measurements of pre- and post-column sample. XAD resin was cleaned through the use of a Soxhlet reactor. The resin was first rinsed with methanol and then with acetonitrile for a minimum time of 48 hours each. The clean resin was transferred to a clean flask, and rinsed and stored in Milli-Q™.

Fluorescence excitation-emission matrix measurement

Based on the measured DOC level, samples were diluted to ~ 1.0 mg/L of DOC with 0.01M KCl solution which pH was pre-adjusted to 2.8 using HCl. This dilution procedure was performed to correct the inner-filter effect, and to minimize possible metal complexation of DOC [6]. All samples were performed at room temperature. Fluorescence excitation-emission matrix (EEM) was recorded on a FluoroMax-2 or -3 spectrofluorometer (Horiba Jobin Yvon Inc., USA). A 150-W ozone-free xenon arc-lamp was used as a light source for excitation. The slit widths were set to 5 nm for both excitation and emission and the scan speed were set to 120 nm/s. All EEMs were obtained by measuring the emission spectra in the range of 300-500 nm at 2 nm intervals, with an excitation range of 240-450 nm at 10 nm intervals. EEMs of each sample were subtracted with an EEM of 0.01M KCl (pH 2.8 adjusted with HCl) solution (set as a blank EEM) to remove Raman scatter peaks. Correction steps were applied to each blank-subtracted EEM using emission and excitation correction factors provided by the manufacturer. Intensities were normalized to the area under water Raman peak of excitation at 350 nm.

Fluorescence Index (FI)

Fluorescence index is a ratio of fluorescence intensity of emission wavelength 450 nm and of emission wavelength 500 nm at excitation 370 nm [7]

Humidification Index (HIX)

Humidification index, which is a measure of the extent of humidification, was introduced by Zsolnay *et al*. [8, 9], and later, Ohno T. [10] proposed slightly modified method. Based on the method of Zsolnay *et al*., HIX is calculated from the fluorescence data as follows (equation 2-1)

$$HIX = \frac{\sum_{\lambda=435}^{480} I_\lambda}{\sum_{\lambda=300}^{345} I_\lambda} \quad \text{or} \quad HIX = \frac{\int_{435}^{480} I(\lambda)\, d\lambda}{\int_{300}^{345} I(\lambda)\, d\lambda} \tag{2-1}$$

Where I(λ) represents the fluorescence intensity of emission spectral curve with excitation at 254 nm. Therefore, an HIX value is determined by the sum of the fluorescence intensity in the emission wavelengths 300 nm → 345 nm region divided by the sum of the fluorescence intensity in emission wavelengths 435 nm → 480 nm.

Biodegradable Dissolved Organic Carbon (BDOC) measurement

5-day BDOC is defined as initial sample DOC minus the sample DOC after five days within the BDOC reactors (i.e., $BDOC_5=DOC_0-DOC_5$) with biologically active sand (BAS). BDOC was measured in a modified method of the Simplified BDOC method developed by Allgeier *et al.* [11]. 100 ml of clean sand was placed into 1-L acid-washed bottle. Prior to this, sand was with tap water rinsing, combustion at over 550 ^oC, rinsing with Milli-Q water, followed by drying at room temperature for 3 days. Biologically activated sand (BAS) was kept active inside of 1L amber serum bottles. Algal growth was limited by keeping the reactors in a dark environment. When the reactors were not in use, the water in the reactors was replaced every seven to ten days to keep the active biofilm.

The BAS consists of garnet drinking water filtration media inoculated with a mixed bacteria culture obtained from unfiltered secondary wastewater effluent. All of reactors are shaken at 100

rpm to ensure substrate and gas transfer to the biomass. The setup for the aerobic tests varies from that of the anoxic (anaerobic) tests. For the aerobic tests, the dissolved oxygen levels are maintained near saturation by atmospheric gas transfer. This ensures that oxygen is the dominant electron acceptor.

For the anoxic tests, the dissolved oxygen is purged out by the nitrogen gas (Air Gas Ultra High Purity). Once the dissolved oxygen (DO) concentration drops below 0.5 mg/L, under this setup, nitrate, instead of oxygen, is the dominant electron donor. BDOC is calculated by the difference of DOC between the initial sample and the biodegraded sample for 5 days.

$$BDOC_5\ (mg/L) = DOC_0\text{-}DOC_5 \tag{2-2}$$

where DOC_0 and DOC_5 mean DOC at day 0 and day 5, respectively.

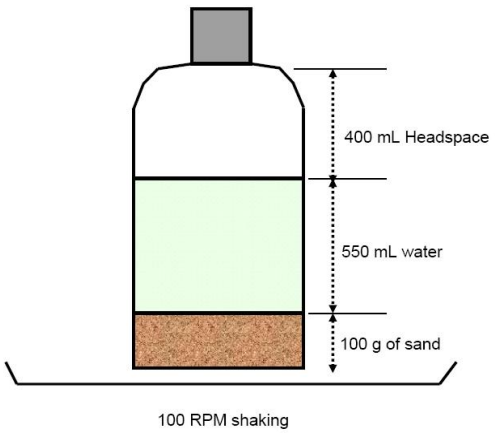


Figure 2-5 Batch reactor set-up for BDOC test

Ratio of Phenol to Aromatic (Phenol/Arom)

Phenol to Arom signifies the ratio of phenolic over highly conjugated aromatic compounds and denotes the intensity ratio of the fluorescence intensity at $\lambda_{ex}/\lambda_{em}$ =340/440 over $\lambda_{ex}/\lambda_{em}$=440/515 nm [12].

Analyses performed at other research group

Dissolved Organic Nitrogen (DON) by Arizona State University

A low-level DON method was used, which employed dialysis pretreatment to remove dissolved inorganic nitrogen (DIN). DON concentration was calculated by the difference of total dissolved nitrogen (TDN) and dissolved inorganic nitrogen (DIN):

$$DON = TDN - NO_3^- - NO_2^- - NH_4^+ \quad (2\text{-}3)$$

$$DIN = NO_3^- - NO_2^- - NH_4^+ \quad (2\text{-}4)$$

Since the standard DON measurement method has limitations in waters with elevated DIN/DON ratios, a recently developed dialysis-based pretreatment was also used to quantify DON [13]. The sample solution was placed in a clean cellulose ester dialysis tube (Spectra/Pore, Spectrum Laboratories Inc., CA) and was dialyzed for 24 hours in a covered dialysis system. The dialysis membrane has a nominal MW cut-off of 100 Da. The membrane is rigid and symmetric with a nominal molecular weight cutoff of 100 Da. All membranes were cleaned prior to use by rinsing thoroughly with deionized water over several days until DOC was less than 0.01 mg/L in the supernatant solution, indicating minimal leaching from the membrane. After comparing DON values using dialysis pretreatment DON values without pretreatment, (1) DON without dialysis pretreatment was reported for samples with DIN/TDN ratios $\leq$ 0.6 mgN/mgN, and (2) DON with dialysis pretreatment was reported for samples with DIN/TDN ratios > 0.6 mgN/mgN. A specific explanation for the choice of cut-off is given elsewhere [13].

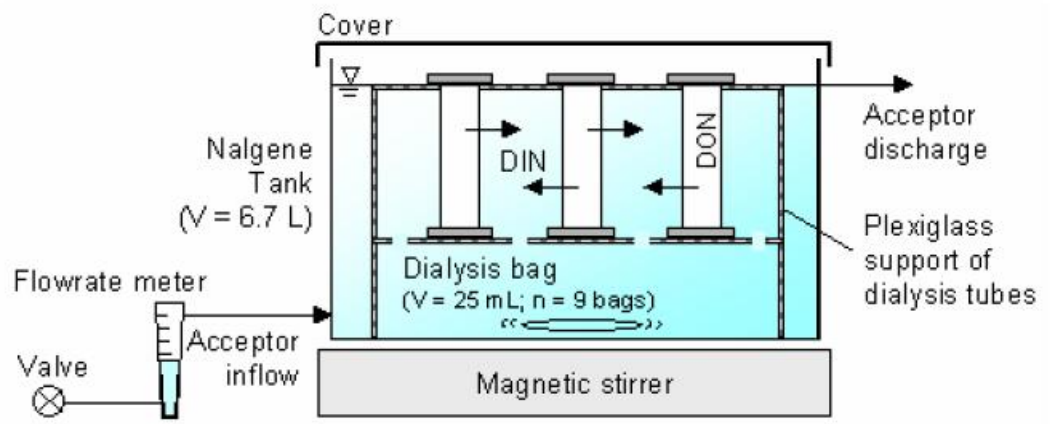


Figure 2-6 schematic diagram of the dialysis system (9 dialysis tubes; 3 rows of 3 tubes). (adapted from ref. [14])

A Shimadzu TOC-V$_{CSH}$ analyzer (high temperature combustion at 720 °C; nondispersive infrared detection) with a TNM-1 TN unit (chemiluminescence detection) (Shimadzu Corp., Japan) was used to measure DOC and TDN simultaneously. MDLs of TDN and DIN measurements were calculated for several model compounds [13]; MDLs of the TDN method ranged from 0.007 to 0.019 mgN/L depending on the chemicals added. Nitrate (MDL = 0.005 mgN/L) and nitrite (MDL = 0.005 mgN/L) were measured using a Dionex DX-120 Ion Chromatography system (Dionex Corp., CA). Ammonia (MDL = 0.005 mgN/L) was measured by the automated phenate method (SM 4500-NH_3 G) [15] using a TRAACS 800 autoanalyzer (Bran-Luebbe, Germany).

Other water quality parameters

Br^- was measured by ion chromatography [15]. The minimum reporting level (MRL) is 0.02 mg/L. Iodine was measured by inductively coupled plasma (ICP)/MS (MDL: 0.006 mg/L) [16].

NH_3 was measured either using an NH_3-selective electrode [15] or by the salicylate method

(Hach® Method 8155), which is a colorimetric method.

As part of the DON measurements, NO_3^- and NO_2^- (SM4500-NO3F) and NH_3 (SM4500-NH3G) were measured on a high-speed continuous-flow wet chemistry analyzer (TRAACs 800 Autoanalyzer, Bran-Luebbe) [15].

CHAPTER 3

WASTEWATER FINGERPRINTING USING EXCITATION-EMISSION MATRIX (EEM) FLUORECENCE SPECTROSCOPY COMBINED WITH PARALLEL FACTOR ANALYSIS (PARAFAC)

Introduction

Fluorescence spectroscopy is one of the analytical tools most frequently used for dissolved organic matter characterization, and in environmental sample: its increasing application to wastewater organic matter as well as drinking water natural organic matter and sea water organic matter has been successfully performed by numbers of researchers [6, 17-20]. Overall, through many research studies on fluorescing organic matter in EEMs, categorization according to humic-like peaks (e.g., humic acids and fulvic acids) and protein-like peaks (e.g., amino acids-like, tyrosine-like, and tryptophan-like peaks) has been generalized. At the same time, statistical approaches such as parallel factor analysis (PARAFAC) have made it possible to identify underlying fluorophores consisting of a single EEM, and to look at the environmental processes in more detailed [21-29]. The fact that fluorescence coupled with PARAFAC resolves bulk excitation-emission spectra into (known or unknown) fluorophores also has a strong attraction when being used as a tool in fingerprinting of wastewater effluent organic matter (EfOM).

Wastewater reuse is becoming increasingly emphasized as a water scarcity solution, and it will be continue to be an important means for indirect or direct potable sources due to increasing population density, accompanying water demand, and global water pollution. However, it should

be noted that wastewater EfOM has different characteristics compared to NOM, that is, mainly the former consists of more hydrophilic substances whereas the latter tends to be more hydrophobic. Since most drinking water treatment processes have been focused on removing NOM to control disinfection by-products (DBPs) formation during or after treatment processes, this may cause unexpected problems, such as a poor removal of hydrophilic EfOM, and consequently, microbial regrowth in the distribution system, or increases of DBP formation (e.g., NDMA or HAN for DWTPs receiving treated wastewater with a high content of DON). Increases of DBP formation potentials (DBPFPs) by chlorination of proteineous substances or wastewater impacted water have been reported by several studies [30-32]. In cases of negative impacts, appropriate monitoring and fingerprinting of EfOM will be crucially important not only in controlling water quality, but also in cost efficiency for treatment and modification of operation scenarios. Recently, Holbrook *et al.* (2006) have shown three (3) PARAFAC components identified by using 55 EEMs from surface waters and wastewater reclamation facilities [22]. Although 3 components were validated using the split-half analysis, more than 4 components could not be verified according to authors' statement (possibly due to the relatively few number of EEMs). In contrast, some researchers have included relatively small numbers of wastewater EEMs in their PARAFAC modeling [23, 33]. However, fluorophores associated with most PARAFAC results may be constrained to a specific region or specific type of WWTP, so that fluorophores may not be generalized as representative wastewater-derived fluorophores unless those are impacted sites or WWTPs having various different types of treatment processes. Moreover, the results tend to be of little interest to probe (physico-) chemical properties of identified components (fluorophores) and their treatability in environmental processes, although to understand characteristics and/or treatability of fluorophores (regardless of microbial-derived

or terrestrially-derived) is important to be able to assess the extent of anthropogenic impact, and to provide an appropriate control strategy in wastewater reuse.

The objective of this study is to identify fluorophores indicating wastewater impacts through the use of fluorescence EEM coupled with PARAFAC, which fluorophores have distinct spectral signatures differentiated from NOM, depending on type of water and sampling location, and to apply the results to the environmental processes and/or watersheds dealing with EfOM to better understand the impact of EfOM properties on treatment process. To author's knowledge, this study is the first of its kind to determine wastewater-indicative fluorophores using fluorescence and to apply the PARAFAC analysis to a large dataset of 423 EEMs, in which there are a variety of different source waters such as surface waters (lake and river), drinking water treatment plants (DWTPs), wastewater reclamation plants (WRPs), and wastewater treatment plants (WWTPs). Moreover, the dataset of this study can also be a representative EEM dataset of wastewater from different treatment processes widely being applied, which can address the treatability of each fluorophore according to different treatment aspects.

Brief background on PARAFAC

Parallel factor analysis is a decomposition method for multivariate data. In fluorescence excitation-emission spectroscopy, each sample is measured by exciting the sample at several wavelengths and measuring the emitted light at several wavelengths, which result in a measurement called an excitation-emission matrix (EEM). The fluorescence intensity reflecting the number of photons emitted is plotted versus the excitation and emission wavelengths and the resulting landscape (or contour map) is a function of the amounts and types of fluorophores (light-emitting or fluorescing compounds) in the sample. Since the theory and application of

PARAFAC is described well in several references [34-36], an overview of the method will be discussed here. Theoretically, the systematic chemical part of the EEM can be described in a simplified way as in Equation (3-1). For a measurement of one sample, i.e., for one EEM with R fluorophores measured at emission wavelength j and excitation wavelength k the intensity x_{jk} is

$$x_{jk} = \sum_{r=1}^{R} a_r b_r c_r + e_{ijk} \tag{3-1}$$

Therefore, the raw fluorescence excitation/emission spectra (i.e., the data array, $\underline{X}$) of the i numbers of the measured EEMs, j numbers of excitation wavelengths, and k numbers of emission wavelengths, R numbers of the factors is

$$x_{ijk} = \sum_{r=1}^{R} a_{ir} b_{jr} c_{kr} + e_{ijk} \tag{3-2}$$

a_r (first mode) is the object score (magnitude of the fluorophore), b_r (second mode) and c_r (third mode) are the excitation loading and the emission loading, respectively. e_{ijk} is the residual ($\underline{E}$) and contains the variation not explained by the PARAFAC model [35]. The principle behind the PARAFAC decomposition is to minimize the sum of squares of e_{ijk}. A graphical illustration of the decomposition of the data array $\underline{X}$ is displayed in Figure 3-1. The model of fluorescence data is based on the assumption that the total absorbance is small and that there is no energy transfer between analytes, which ensures no inner filter effect.

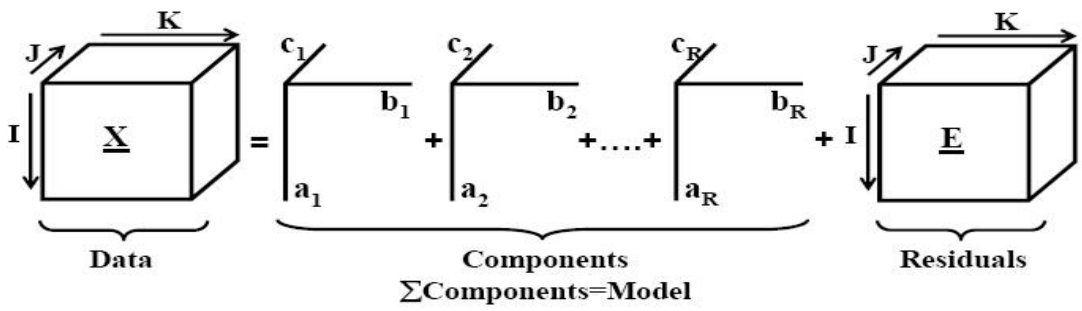


Figure 3-1 The PARAFAC model with R components (adapted from reference [36]). X represents the fluorescence EEM, and E means the residual EEM

Experimental Methods

Water sample collection and preparation

The water samples were collected from multiple locations of 6 river watersheds, ~25 WWTPs (including wastewater reclamation plants) and DWTPs, downstream WWTPs, and two groundwater recharge sites in eight states (Figure 3-2) of the USA during a 1.5-year period (Feb. 2004-Aug. 2005). Sampling location and time were selected with consideration of organic matter in pristine or minimally/heavily wastewater-impacted water. The selected WWTPs utilize conventional secondary treatment with partial or full nitrification or nitrification/denitrification, less expensive treatment (e.g., trickling filters, aerated lagoons), and advanced wastewater treatment (e.g., membranes). The DWTPs use conventional drinking water treatment processes (coagulation/flocculation) as well as advanced treatment processes such as GAC, membranes, advanced oxidation processes (AOPs) using ozone or UV/H_2O_2. In addition to WWTPs and DWTPs, river watersheds studied included pristine sites, WWTPs discharging EfOM into the watersheds, and downstream sites or DWTPs impacted by upstream WWTPs. Groundwater sampling sites were impacted by wastewater recharge. (More detailed descriptions on the water and wastewater treatment plants utilized samples is available in Chapter 2 Methods and Materials)

Figure 3-2 Locations of studied utilities (DWTPs, WRPs, and WWTPs), and surface water (lake and river watersheds) in the eight states of USA

Samples were taken without headspace in Nalgene screw-cap bottles and immediately transferred to the laboratory in an ice chest. Samples were filtered through a 0.45 μm pore sized glass filter which was pre-rinsed with Milli-Q water and sample water in series. Dissolved organic carbon (DOC) concentrations of samples were measured using Sievers 800 total organic analyzer (Ionics Instruments, USA). Based on the measured DOC level, samples were diluted to ~ 1.0 mg/L of DOC with 0.01M KCl solution, and pH was pre-adjusted to 2.8±0.1 using HCl. This dilution procedure was performed to correct the inner-filter effect, and to minimize possible metal complexation of DOC [6]. Some literatures studies recommend that at the wavelength of excitation the absorbance of the (potentially) fluorescing species below 0.02 is appropriate because the molar absorptivity at that nominal wavelength is linear and the self-shadowing effect does not occur [37, 38]. In our samples, ~1.0 mg/L of DOC had an absorbance much below the criteria at the selected excitation wavelengths.

Fluorescence Measurements

Fluorescence excitation-emission matrix (EEM) was recorded on a FluoroMax-2 or -3

spectrofluorometer (Horiba Jobin Yvon Inc., USA). A 150-W ozone-free xenon arc-lamp was used as a light source for excitation. The slit widths were set to 5 nm for both excitation and emission and the scan speed were set to 120 nm/s. Prior to measurement, the lamp maximum intensity at 467 nm, and the water Raman area under the maximum peak 350 nm were monitored to verify the instrument stability condition. All EEMs were obtained by measuring the emission spectra over the range of 300-500 nm at 2 nm intervals, with an excitation range of 240-450 nm at 10 nm intervals. EEMs of each sample were subtracted with an EEM of 0.01M KCl (pH 2.8 adjusted with HCl) solution (set as a blank EEM) to remove Raman scatter peaks. Correction steps were applied to each blank-subtracted EEM using emission and excitation correction factors provided by the manufacturer. Intensities were normalized to the area under water Raman peak of excitation at 350 nm. All measurements were performed at room temperature (22±2°C).

PARAFAC modeling procedures

EEMs were combined into a dataset consisting of 423 samples × 201 emission wavelengths × 43 excitation wavelengths and analyzed by the PARAFAC method. Of the entire EEMs, ~75% of EEMs were representing wastewater EfOM, in which samples were from WWTPs or wastewater-*impacted* or *–dominant* locations, and ~25% of EEMs were representing drinking water NOM samples including standard reference NOM isolates (i.e., NOM, humic and fulvic acid from Suwannee River, IHSS). Some of EEMs from biodegradation and coagulation experiments of NOM/EfOM in this study were also included into the dataset in order to accelerate the PARAFAC modeling because as the number of EEMs increases, the explained variation of the fluorophores increases [39]. Rayleigh scattering effects were removed from the dataset by deleting emission measurements made at wavelengths ≤ excitation wavelength + 20

nm [25]. A triangle of zeros also was inserted into the EEMs in the region of missing data, which is the region where excitation wavelength is larger than emission wavelength. Non-negativity constraints and split-half analysis were applied to validate the identified components. The split-half analysis involves dividing the whole dataset into two random equal-sized groups (in this case, one dataset was made with every odd sample and the other with every even sample), and making an independent PARAFAC model based on both half datasets [25, 27]. The PARAFAC analysis was run in MATLAB® version 7.0 (Mathworks Inc, USA) using the N-way 2.1 version toolbox, available at http://www.model.kvl.dk/source [40]. Through a preliminary analysis, EEMs with high leverage, which are considered as strong outliers, were removed from the dataset in order to enhance the goodness-of-fit and those EEMs identified as outliers were mostly from the mixed liquor tank in wastewater treatment plants. Their DOC and fraction of hydrophilic organics were extremely high compared to other samples. In addition, emission wavelengths below 300 nm were excluded from the dataset due to a deteriorating signal to noise ratio in this region. The above-described data processing prior to PARAFAC application dataset gives 34.8259% of the missing values in the entire dataset. Finding the number of components was performed by a spectral comparison between calibration and validation datasets, and by core consistency diagnostics. The detailed introduction and application of PARAFAC are described in [35, 36].

Statistical analysis

Pearson correlation test ($p < 0.05$) of individual PARAFAC components with 18 commonly used NOM characterization analytical parameters was performed using SPSS® version 13.0. The 18 NOM parameters used were as follows; UV absorbance at 254 and 280 nm (UVA_{254} and

UVA_{280}), molar absorptivity at 280 nm (ε280), DOC, specific UV absorbance (SUVA) at 254 nm, hydrophobic/transphilic/hydrophilic organics (HPO, TPI and HPI DOC) by XAD-8/-4 fractionation method [4, 41], amount of polysaccharide, humic substances and low molecular weight acids determined by high performance size-exclusion chromatography with DOC detector (HPSEC-DOC) and by converting each integrated peak area to corresponding DOC [2, 42]. Additionally, through fluorescence data, fluorescence index (FI) [43], and humidification index (HIX) [9] were obtained.

Results and Discussion

Model validation

Five components were established in this study. The explained variation and Sum-of-Squares of residuals with the components showed 99.08% and 29.0026, respectively. For model validation, the PARAFAC algorithm was applied step-wise to two individual dataset made by the split-half method, for which the two datasets are usually called the calibration and validation data arrays. The appropriate number of components was determined by comparing the excitation and emission spectra of the components between two datasets. Also, a core consistency of 83% confirms that 5 components are an appropriate number of components. Comparison of the measured, modelled and residual EEMs revealed that the five-component model explains the majority of the variation of EEMs, and remaining signals of the instrumental noise level had no distinct shape. Examples of the measured, modelled and residual EEMs are shown in Figure 3-3.

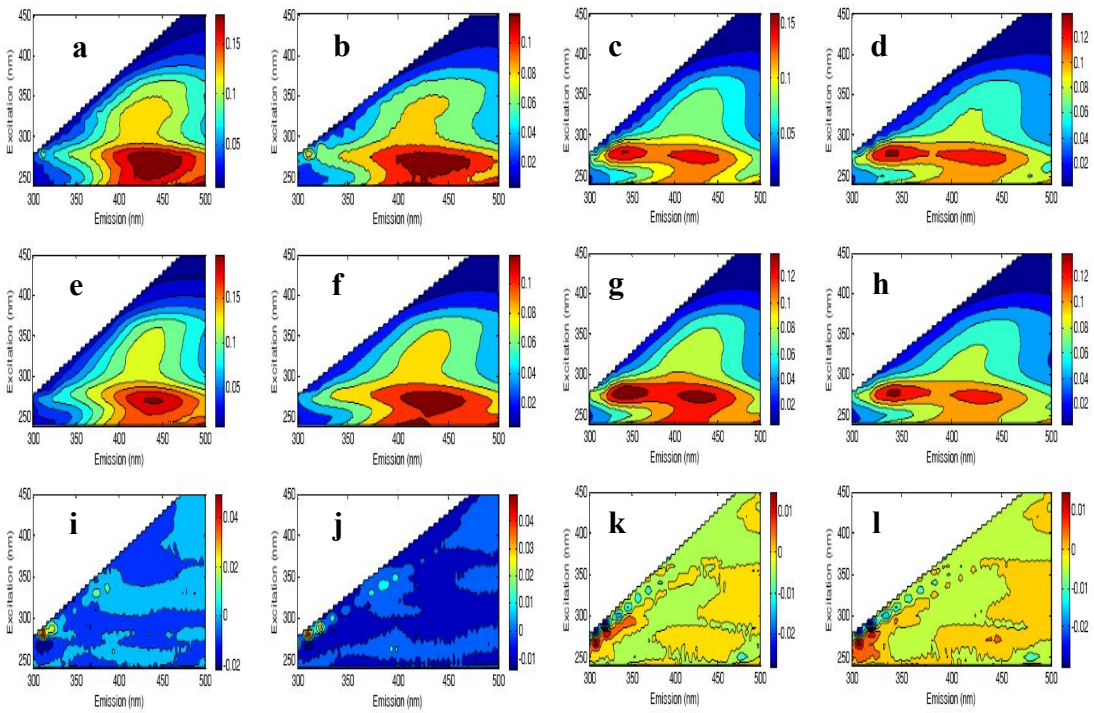


Figure 3-3 Examples of measured (a-d), modelled (e-h), and residual (i-l) EEMs for four different water samples taken from the State of Pennsylvania; a and b samples were from upstream of WWTP as a NOM, c and d samples represent WW effluent as an EfOM

Characteristics of components identified

The five components identified are summarized in Figure 3-4 and Table 3-1. In Figure 3-4, hollow (pink) triangle and diamond (green) symbols display the excitation and emission spectra of references comparable to each component. Spectral loadings and contour features of these components showed many similarities to results in previous studies [4, 10, 29], and also to the EEMs of fluorophores (shown in Chapter 5).

Component 1, which resembles the EEM of SRFA (shown in chapter 5), showed two maxima: the first maxima peak at an excitation/emission of 270 nm/455 nm, and the second at 360/455 nm. Thus, component 1 is likely to represent a group of fulvic-like fluorophores and probably hydrophobic organics. However, its wide-ranged excitation and emission spectra indicate that it may be a complex mixture of NOM, rather than a single fluorophore.

Component 2 and 4 exhibited their maxima in 295 nm/343 nm and 275 nm/326 nm, respectively. Their EEMs showed similar spectral features to protein-like fluorophores in spite that in component 4, a shoulder peak below excitation of 240 nm was observed. More specifically, component 2 resembled a tryptophan peak though a shifted excitation maxima, while component 4 was very similar to an albumin peak (see the contour introduced in Chapter 5) in both its shape and position. Component 3, which showed spectral similarities with component 3 reported in Cory *et al.* [9], exhibited its maxima at 325 nm/401 nm of excitation/emission wavelengths. Compared to the contour/peak position of the hydrophilic fraction of organic matter, it showed similarities in shape and maxima location and this may mean that component 3 is close to humic-like organic matter with a more hydrophilic nature than component 1.

Component 5 absorbed only at a shorter wavelength (a high energy) below 240 nm and its emission spectra were broad with a maximum at 397 nm [30]. Until quite recently, this component was not often reported in the literatures, but recently, components similar to this component were shown in Cory *et al.* [9] and Murphy *et al.* [44]. Component 6 shown in Cory *et al.* is a virtual co-plot with component 5 in Figure 3-3.

In general, peaks at excitation wavelengths of 250-280 nm and shorter emission wavelengths below 380 nm are related to soluble microbial product-like (SMP-like) materials [31]. Also, it is known that fluorescence at Ex=270-290 nm, Em=330-350 nm are attributed to the indole moiety seen in tryptophan (see structure of indole moiety in Appendix), which is a common structure in nature and is abundant in alkaloids [19]. The 'indole alkaloids' is a major class of legal/illegal drugs. Peaks at longer excitation wavelengths over 280~270 nm and emission wavelengths at more than 380 nm are related to humic acid-like organics [38].

Table 3-1 Descriptions of the five components identified. (Second maxima are presented in brackets)

Component	Excitation Maximum	Emission maximum	Component type	References	Origin description
1	270(360)	455(455)	Fulvic acid-like	[25]	allochthonous
2	295	343	UVB protein-like	[25]	autochthonous
3	325	401	UVA humic, terrestrial humic	[25]	allochthonous
4	275	326	Tryptophan-like, protein-like or phenol-like A (or γ) peak	[37, 45, 46] or [46, 47]	autochthonous
5	< 240	397	Humic-like	[24, 26, 44]	allochthonous

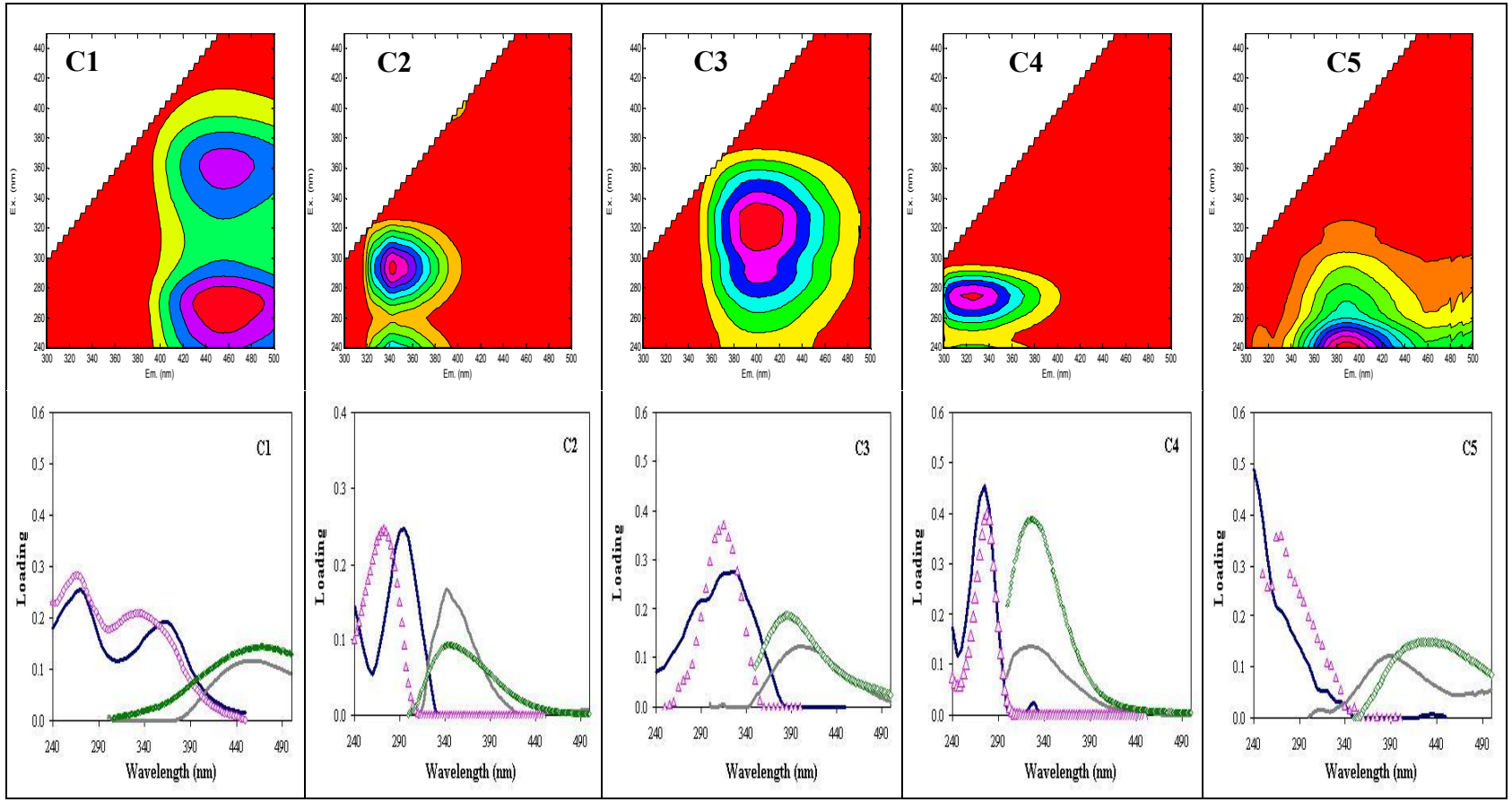


Figure 3-4 Contour plots (top row) and graphs (bottom row) of excitation and emission loadings of five components (dark solid line for excitation loadings, and gray light line for emission loadings); pink hollow triangle and green diamond are comparable excitation emission spectra (for C1: SRFA, C2: tryptophan, C3: CM3 in Cory *et al*.[24], C4: albumin and C5: CM6 in Cory *et al*.[24])

Variability and properties of components in reference NOMs/EfOMs

Proportions of each fluorophore in reference NOMs and EfOMs (isolated EfOM and SMPs) were investigated by including their EEMs into a validation dataset for PARAFAC (Figure 3-5, see also Figure 5-9 for their EEM contours in Chapter 5). In NOM, component 1 is presented as the most abundant fluorophore (ca 50% or more), followed by component 5. For component 2, there were present 5.5%, 6.0% and 0.0%, and for component 4 there were present 0.0%, 2.7% and 12.4% in SRHA SRFA and SRNOM, respectively. 6.0% and 2.7% of component 2 and 4, respectively. The sum of the total fractions of these two components did not exceed 12.4% in NOM samples, meaning that these components are (quite) minor constituents in NOMs.

In addition to investigating the variability of components in reference materials, comparing the peak locations of components to those of organic matter isolates may provide information on properties of components. Figure 3-6 shows peak maxima locations of 5 components for various NOM and EfOM standards and isolates, showing that NOMs from IHSS are located in longest emission wavelengths, and overall Ex/Em wavelengths of HPI OMs showed shorter (blue-shifting) wavelengths than TPI OMs. Protein-like substances appeared at their maxima at shorter Ex/Em wavelengths. Supposing that IHSS NOMs are hydrophobic, component 1 is likely to be a mostly hydrophobic fluorophore. Component 3 is close to be hydrophilic. Component 5 is not overlapped with any of the three fractions (HPO, TPI, and HPI), however, when considering the broad emission spectra ranging to an emission wavelength of 500 nm, it is thought to be hydrophobic.

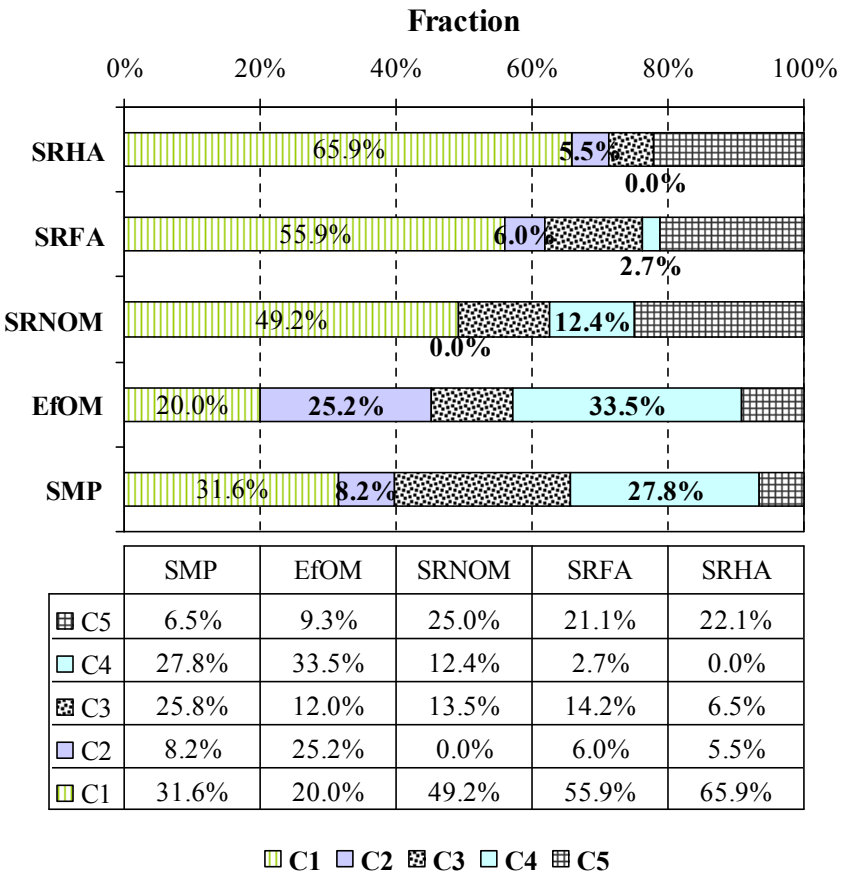
According to the results on molecular weight distributions (MWDs) of these reference IHSS NOMs (Figure 5-5 in Chapter 5), it was observed that the weight-average molecular weight (M_w) and number-average molecular weight (M_n) of SRHA was 2,795 and 1,719 daltons (Da),

respectively, and M_w and M_n values of SRFA and SRNOM were less than SRHA and decreased in turn (refer to Table 5-4 in Chapter 5). Their polydispersity values (ρ) did not vary noticeably (~1.6), showing unimodal-like distributions.

Theoretically, M_w and M_n is a function of f_i, the probability of a compound having molecular weight M_i in a MWD [48], and except an extremely polydispersed MWDs, the higher f_i of a compound i having M_i, the closer M_i is to M_w, and the higher fraction is in overall M_n. Therefore, in linking PARAFAC components to MW concepts, the fact of a dominance of component 1 (along with component 5) indicates that this components is the most responsible component for the M_w and M_n of reference NOMs, a component is attributable to humic substance with higher molecular weight. In addition, several researchers using fulvic acids with various same sources such as soils, marines, etc. have reported a spectral red-shifting (to longer wavelengths) in fluorescence emission as the molecular weights increase [49, 50], as well as red-shifting of the absorption (excitation) bands of aromatic compounds as the number of rings increases [51]. The facts that the maxima locations in EEMs were blue-shifted and that M_w decreased as the fraction of component 1 in NOM samples decreased support the notion that component 1 is a moiety highly related to NOMs, and this fraction dominates over the overall natures of NOMs. To the author's knowledge, while component 5 is also considered as a common fluorophore of NOMs and its MW would be smaller than component 1 according to its maxima positions, still one must be cautious to assign its molecular weight simply based on the Ex/Em maxima locations. Of course, since the NOM is heterogeneous, it is difficult to state the molecular weight of component 1 in specific numbers, and still there remains an unknown possibility to be resolved into another sub- or independent fluorophores. Nevertheless, based on these observations, this argument is reasonable.

On the other hand, in contrast, component 4 was plentiful in EfOMs, followed by component 2, and the sum of component 2 and component 4 ranged over 35-58.9%. Component 1 and 3 were also abundant fluorophores in EfOMs, especially in SMP samples. Moreover, in linking to the MWDs of EfOM and SMP (MWDs shown as Figure 5-5 in Chapter 5), both MWDs were broadly distributed from low molecular weights (LMWs) to high molecular weights (HMWs), resulting in high polydispersities. Therefore, fractional distributions of each component not showing extremely significant differences may be linked to high polydispersity of their SEC, and furthermore, the abundance of component 1 in EfOM and SMP supports the notion that organic matter from DWTP is one of constituents of EfOM along with organic matter from WWTP biological treatment. (However, the fractional contribution or presence of NOM in EfOM is dependent upon internal/external influences such as sources, physicochemical reactions, geology, etc.)

In comparison of maximum peak locations of component 2 and component 4, both were shown to be close to amino-acid or protein-like fluorophores. Shorter Ex/Em wavelengths of component 2 peak maxima compared to component 4 imply that chemical bonds in component 2 require higher energy in absorption and fluorescence (emission) than component 4 whose maxima were very close to that of Albumin. Component 2 had similar fluorescence wavelengths with tryptophan and the 1st peak of EfOM.



	SMP	EfOM	SRNOM	SRFA	SRHA
C5	6.5%	9.3%	25.0%	21.1%	22.1%
C4	27.8%	33.5%	12.4%	2.7%	0.0%
C3	25.8%	12.0%	13.5%	14.2%	6.5%
C2	8.2%	25.2%	0.0%	6.0%	5.5%
C1	31.6%	20.0%	49.2%	55.9%	65.9%

Figure 3-5 Fractional compositions of fluorophores in different organic sources (NOM and EfOM)

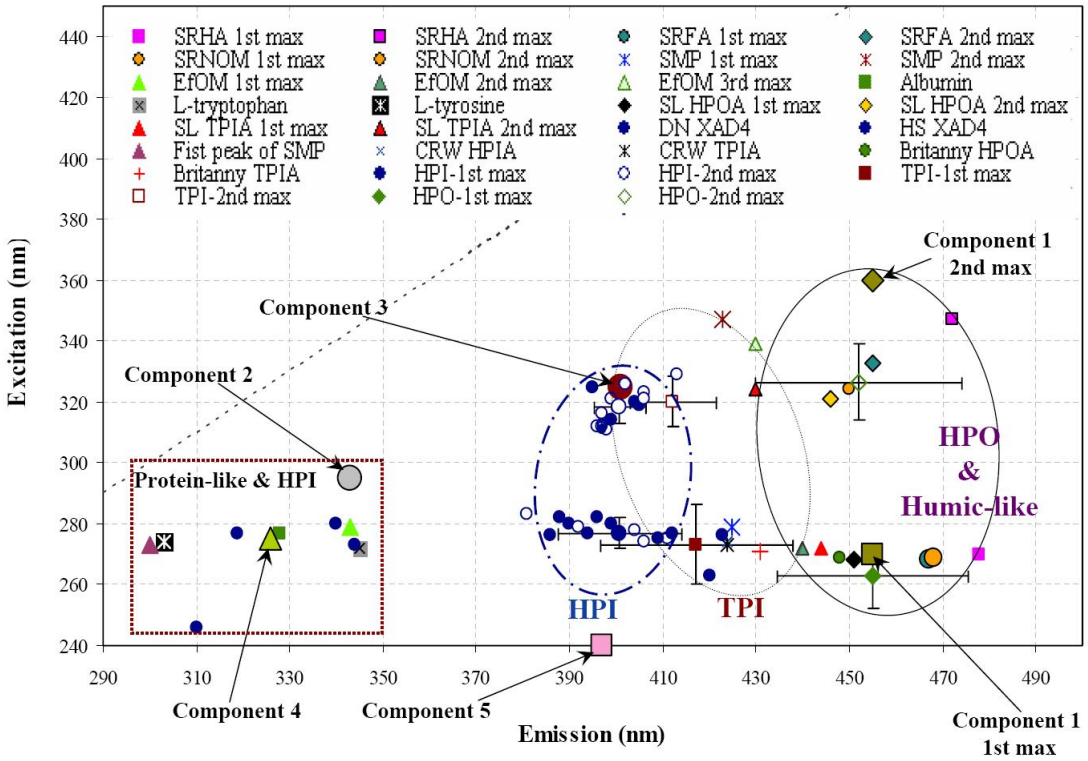


Figure 3-6 Peak maxima locations for 5 components and comparison with various NOM and EfOM references (standards and isolates); 1st (◆) and 2nd (◇) maxima of HPO and Humic-like OM, 1st (■) and 2nd (□) maxima of TPI OM 1st maxima, and 1st (●) and 2nd (○) maxima of HPI. Error bars mean standard deviations

Statistical analysis

Correlation analysis and hierarchical cluster analysis (Ward's methods, squared Euclidean distances, z-score data transformation) were used to statistically assess the relationship of components with NOM characterization parameters. Table 3-2 summarizes the Pearson correlation (r) results between fractions of 5 components and NOM characterization parameters. Component 1 showed relatively strong negative correlation with parameters that have increased levels in wastewaters (HIX, component 2 and 4, bromide concentration, fraction of $BDOC_5$, and FI). On the other hand, component 2 had strong relations with FI, $BDOC_5$, and bromide, and negative correlation with component 1 and component 5. Bromide is widely used in making

brominated flame retardants, methyl bromide used as pesticide to fumigate soil, dyes, agrichemicals, pharmaceuticals, etc[52, 53]. Also it is one of major constituents in cationic detergents such as hexadecyltrimethyl ammonium bromide (CTAB) and trimethylammonium bromide (TTAB), hence, an increase of bromide concentration is likely linked to an increase of wastewater impact.

Component 3 did not show strong positive or negative correlations with any parameters. For component 4, correlations showed similar trends to the ones shown in for component 2 with additional positive correlation with PS and LMA, which implies that these two components are explainable by an impact of wastewater input. Finally, component 5 also exhibited negative correlation with parameters increasing in wastewater samples. Strongly inverse correlations of both component 1 and component 5 with $BDOC_5$ mean that these components are not readily biodegradable, thus constituting fluorophoric moieties of hydrophobic humic substances.

Figure 3-7 shows a parametric clustering between 5 components identified by the PARAFAC analysis and 15 NOM characterization (organic) variables. As shown, variables were grouped into two groups: group I and group II which differentiate parameters according to closeness to their polarity (group I-hydrophobic, and group II-hydrophilic natures). Component 1 and component 5 were clustered in group I, whereas component 2, component 3 and component 4 belonged to group II. For correlation analysis, component 3 was weakly correlated with parameters indicating the aromatic extent ($0.3< r < 0.5$). However, in cluster analysis, it appeared to belong to group II.

Group I can to be divided into group I-1 explaining organic (nutrient) loading (i.e., extent of nutrients contamination) – e.g., UV_{254}, COD, DOC, and DON, and parameters (group I-2) that explain the degree of aromaticity (usually linear to hydrophobicity).

Table 3-2 Summary of Pearson correlations between PARAFAC components and NOM characterization parameters

PARAFAC Component	Positive correlation				Negative correlation			
	Water parameters	R	*p*	n	Water parameters	R	*p*	n
%C1					HIX	-0.786	0.000	115
					%C2	-0.534	0.000	115
					%C4	-0.736	0.000	115
					Br^-	-0.555	0.000	36
					$BDOC_5$	-0.513	0.001	41
					FI	-0.503	0.000	115
%C2	FI	0.652	0.000	115	%C5	-0.787	0.000	115
	$BDOC_5$	0.538	0.000	41	%C1	-0.534	0.000	115
	Br^-	0.536	0.001	36				
%C3*	UV254	0.382	0.000	115	%C4	-0.459	0.000	118
	SUVA	0.365	0.000	115	$BDOC_5$	-0.414	0.007	41
					HIX	-0.309	0.001	118
%C4	HIX	0.776	0.000	115	%C1	-0.736	0.000	115
	$BDOC_5$	0.671	0.000	41	%C5	-0.566	0.000	115
	PS	0.633	0.000	115				
	FI	0.596	0.000	115				
	LMA	0.525	0.000	115				
%C5					%C2	-0.787	0.000	115
					FI	-0.687	0.000	115
					$BDOC_5$	-0.609	0.000	41
					%C4	-0.584	0.000	118
					PS	-0.555	0.000	115
					LMA	-0.536	0.000	115
					DON	-0.525	0.000	98

* displayed for R > |0.3|

DON-mg/L as N, $BDOC_5$-mg/L, SUVA-L/mg·m, Br^--mg/L

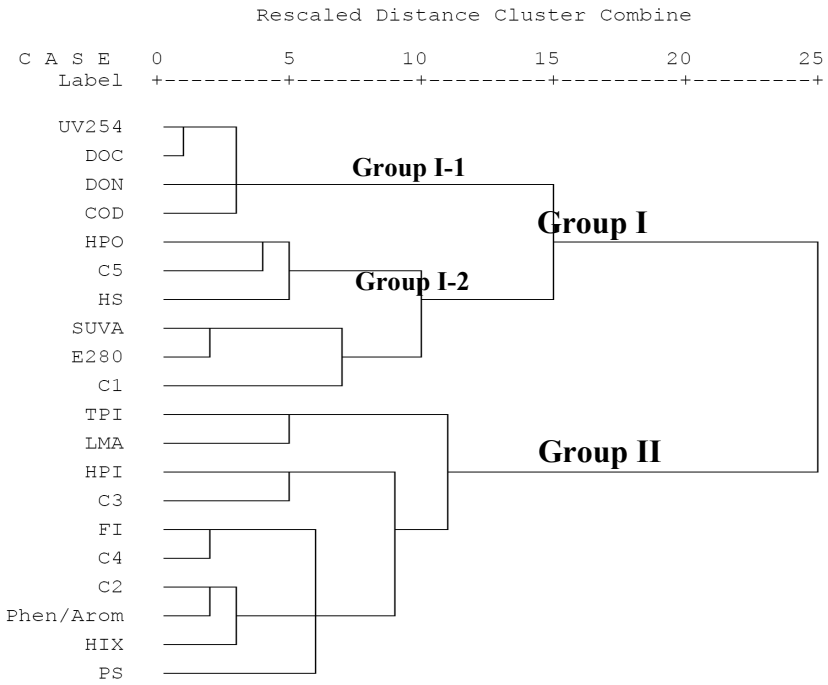


Figure 3-7 Dendrogram showing parametric clustering of NOM characterizing variables (measured for 120 sampling sites, 5 sampling sites were removed due to strong outliers)

A Case Study of PARAFAC application to Effluent-dominant stream: South Platte River Watershed, Colorado

This section of the chapter introduces a case study showing the applicability of PARAFAC components as a wastewater impact indicator or fingerprint, and the case is the South Platte River (SPR) watershed (Colorado, USA), which is known as an EfOM-*dominant* stream, with detailed site descriptions and sampling history. Wasteater impact evaluations have been conducted using 3 approaches; 1) flows, 2) concentrations of primidone, a pharmaceutical, serving as conservative/non-adsorbing tracer, and 3) two microbial-derived PARAFAC components (i.e., C2 and C4) were used to estimate the extent of wastewater discharges on the SPR.

Descriptions of sampling sites

The South Platte River is the major watershed in the Denver (CO) Metropolitan region. Overall, the watershed is a combination of pristine mountain areas, highly urbanized areas, and intense agricultural activities. Its headwaters originate in the Rocky Mountains to the southwest of the Denver Metropolitan area. The South Platte River flows from the south to the north through the cities of Littleton, Englewood, Denver, and Thornton (CO). There are several drinking water and wastewater treatment plants (WWTPs) located on the South Platte River in the Denver Metropolitan area and further downstream. Effluent organic matter (EfOM) from WWTPs, storm water runoff, and irrigation return flows have a major impact on the quality of the river water. The South Platte River is considered to be an effluent-*dominated* stream and wastewater effluent discharges will be further discussed later.

Sampling collection

Samples were collected during three seasons: winter (February 2004), summer (September 2004), and spring (April 2005). The three sampling events were based on hydrology and treatment considerations. In the winter and summer, river flow is low, so the river is more effluent-*dominated* (as compared to the spring season). In addition, there may be more nitrification at the WWTPs in warmer months. These three seasons will show the different impacts of hydrology and treatment.

Figure 4-8 describes a map of sampling sites. Site names on the figure followed the direction of the river flows (from site 1 to site 11), and detailed explanations of sampling sites are as followings.

- Site 1: Chatfield Reservoir (upstream of the Centennial WWTP)
- Site 3: upstream of the Allen DWTP (downstream of the Centennial WWTP)
- Site 5U: upstream of the Littleton/Englewood WWTP
- Site 5: effluents from the Littleton/Englewood WWTP (Site 5W1: after nitrification before chlorination, Site 5W2: after chlorination after denitrification)
- Site 6: at 14th Street (upstream of Farmers and Gardeners Ditch) (downstream of the Littleton/Englewood WWTP)
- Site 7: at the Burlington Canal (drinking water source for the City of Thornton). This site is downstream of Site 6 and Cherry Creek. Although there is a small WWTP (Glendale WWTP) on Cherry Creek, it has been assumed that Cherry Creek will primarily bring in fresh water.

  Water can be diverted at this point.
  - During the February 2004 sampling event, no water was diverted into the Ditch. Hence, all of it continued to flow down the river (and through the Denver Metro area).
  - During the other 2 sample events, water was diverted into the Ditch. So, in those cases, the "total river flow" was equal to what was diverted through the Burlington Ditch and what passed by Site 11.
  - In September 2004, the flow at Site 7 and the flow diverted to the Burlington Ditch were the same, but were different in the other 2 sampling events.
- Sites 8 and 9: effluents from the Denver Metro WWTP (the North and South outfalls)
- Site 11: downstream of the Denver Metro WWTP (although Thornton was downstream of this site, they took water from upstream of the Denver Metro WWTP); this site represents a potential worst-case scenario as a drinking water source for Thornton

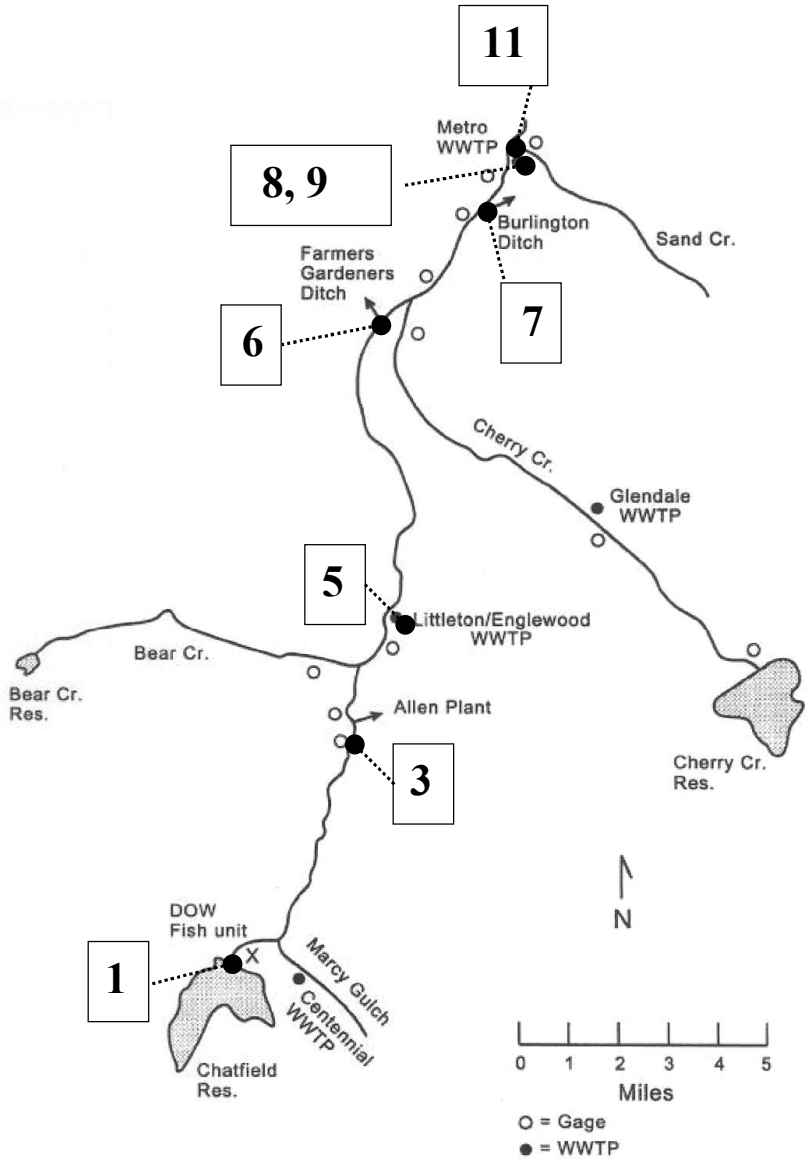


Figure 3-8 Description of the South Platte River (CO) sampling locations

Specific detail descriptions of WWTPs in South Platte River watershed, CO

The Littleton/Englewood WWTP's secondary treatment process included trickling filters and solids contactors, which were followed by nitrifying trickling filters for ammonia removal, and chlorination/dechlorination. The Denver Metro WWTP had two parallel treatment processes. The

South Complex used activated sludge with high-purity oxygen, and chlorination/dechlorination. The North Complex used a single-sludge nitrification/denitrification process (the "modified Ludzack-Ettinger" process), with anoxic cells for the removal of nitrate and an oxic zone where oxidation of organic material and nitrification occurred. The North Complex practiced chlorination/dechlorination, where a small portion of the South secondary effluent water was transferred to the North effluent flow to provide ammonia so that disinfection was accomplished by chloramines.

Results and discussion

Wastewater impact by approach 1) flows, and by 2) primidone concentrations

Discharges from WWTPs provide a large part of the river flow most of the year. Wastewater impact on the river was assessed based on the river/wastewater flow data balancing, and primidone data (a pharmaceutical compound present in wastewater, but not present in pristine samples).

Table 4-3 shows the stream flows on the days of sampling at selected sites, as well as the flows of the WWTPs on the South Platte River. In addition, there were significant diversions taken from the river between the Littleton/Englewood and the Denver Metro WWTPs, specifically the Burlington Ditch (the ditch was off-line during the winter sample event). Each of the WWTPs operated at similar flows during each sample event (Centennial WWTP = 4.1–5.5 mgd, Littleton/Englewood WWTP = 20–22 mgd, and Denver Metro WWTP = 120–130 mgd). Initial flows on the South Platte River (at site 1) were low (15–34 mgd) during the winter and summer sample events and moderate (122 mgd) during the spring sampling. By site 11, flows on the South Platte River were up to 149–249 mgd. The total output—which was

the sum of the South Platte River at site 11 plus the flow diverted to the Burlington Ditch—was 202–342 mgd. The collective flows from the Centennial, Littleton/Englewood, and Denver Metro WWTPs on the days of sampling were 145–157 mgd, which represented 62–77% and 43% of the total flow during the low- and moderate-flow events, respectively. These flow data indicate that the South Platte River was an effluent-*dominated* stream (i.e., >50% WW) during the winter and summer sampling events and effluent-*impacted* (i.e., 10–50% WW) during the spring sampling.

Primidone was present—on the days of sampling (Table 4-4)—at 96–266 ng/L in the effluent of both WWTPs in Colorado that were evaluated. This level is comparable to that of other U.S. WWTPs tested to date in this study. Primidone was not detected at site 1, which is upstream of the WWTPs. During the low-river-flow events (winter and summer), it was detected at sites 3 and 5U—which are downstream of the Centennial WWTP (and site 5U is downstream of Bear Creek)—at 57 and 10–21 ng/L, respectively. It was not detected at site 5U during the moderate-river-flow event (spring). The concentrations of primidone at the sites downstream of the Littleton/Englewood WWTP (sites 6 and 7)—where site 7 is also downstream of Cherry Creek—were 149 and 22–68 ng/L, respectively. The low end of primidone occurrence at site 7 (i.e., 22 ng/L) was during the spring sample event. By site 11, which is downstream of the Denver Metro WWTP and Sand Creek, primidone was up to 40–140 ng/L (40 ng/L during the spring sampling). It is unknown if there are any human health concerns over exposure to this low of a level of a pharmaceutical in a drinking water supply. Moreover, some drinking water treatment processes are capable of removing a variety of pharmaceutically active compounds.

Based on the primidone data, 55–59% of the river flow at site 11 was due to treated wastewater during the winter and summer sampling events (based on the average WWTP-

effluent primidone value during each sampling), whereas only 37% was due to treated wastewater in the spring sample event. Thus, these analyses confirm that the river was effluent-dominated during the winter and summer sample events and highly impacted during the spring sampling (the summer event was the most impacted by WW, whereas the spring event was the least impacted).

In addition to site 11, at site 7 it was found that 26-33% and 14-50% of the river was wastewater-impacted based on flow and primidone data (the bottom rows in Table 3-3 and 3-4), respectively. These values were much less (~30%) than those at site 11, and suggests that the wastewater impact on the South Platte River was greatly increased by Denver Metro WWTP effluent than the 2 previous WWTPs (Centennial WWTP and Littleton/Englewood WWTP) (Table 3-5). Both flow-based and/or primidone-based approaches prove seemingly that the SPR received wastewater discharges in significant portion of its total flows, especially by the Denver Metro WWTP.

In general, the calculated percent impact for each sample event did not differ significantly when based on flow or primidone data (the relative differences between the two methods were 12% to 27%).

Table 3-3 Changes of stream and WWTP flow on the South Platte River watersheds (three sampling events during Feb., 2004 to April, 2005)

Stream and WWTP Flow on South Platte River Location	2/18/2004		9/15/2004		4/06/2004	
	Flow (cfs)	Flow (mgd)	Flow (cfs)	Flow (mgd)	Flow (cfs)	Flow (mgd)
Site 1: Upstream of WWTPs (South Platte River below Chatfield Reservoir	24	15	52	34	188	122
Centennial WWTP		**4.1**		**4.7**		**5.5**
Littleton/Englewood WWTP		**20.9**		**21.8**		**20**
Site 7: downstream of Littleton/Englewood WWTP (South Platte River at Burlington Canal head gate)	125	81	82	53	284	184
Burlington Ditch flow	off	0	82	53	284	126
Denver Metro WWTP		**128**		**130**		**120**
Site 11	385	249	231	149	334	216
WWTPs: total flow for sample event		**153**		**157**		**145**
Total river flow=Burlington Ditch flow + South Platte River at Site 11		249		202		342
% of river flow at Site 11 due to treated wastewater (from all WWTPs) based on flow data		**61%**		**78%**		**42%**
% of river flow at Site 7 due to treated wastewater (from 2 WWTPs) based on flow data*		**31%***		**50%***		**14%***
% of river flow at Site 3 due to treated wastewater (from Centennial WWTP) based on flow data*		**21%***		**12%***		**4.5%***

Source: Reproduced from ref [54]. Copyright 2008 Awwa Research Foundation, *: Additionally added by the author of the dissertation.

Table 3-4 Presence of primidone (a wastewater tracer) at the South Platte River watersheds (three sampling events during Feb., 2004 to April, 2005)

Location	Primidone (ng/L)		
	2/18/2004	9/15/2004	4/06/2005
Site 1: Chatfield Reservoir	< 10	< 8	<8
Site 3: Upstream of Allen DWTP	57	–	–
Site 5U: Upstream of Englewood WWTP	21	10	<8
Site 5: Littleton/Englewood WWTP	172	266	96
Site 6: Upstream of Farmers Ditch	149	–	–
Site 7: Burlington Canal	48	68	22
Site 8: Denver Metro WWTP/North	161	208	118
Site 11: Downstream of Denver Metro WWTP	91	140	40
WWTPs: average value for sample event	167	237	107
% of river flow at Site 11 due to treated wastewater (from all WWTPs) based on primidone data	**55%**	**59%**	**37%**
% of river flow at Site 7 due to treated wastewater (from 2 WWTPs) based on primidone data**	**28%****	**26%****	**23%****

Source: Reproduced from ref [54]. Copyright 2008 Awwa Research Foundation, **: Additionally added by the author of the dissertation.

Wastewater impact by approach 3) two PARAFAC components

It is believed that flow- or primidone-based approaches were successfully comparable to each other. Nevertheless, the author questioned that "wastewater impact" measured by these two methods could reflect changes of organic matters properties in a wastewater impacted stream. In other words, "is the % WW impact equal to the % EfOM impact?" Because amounts or fractions of organic matter discharged to a stream can be changed depending on degrees or types of wastewater treatment processes, and even if a WWTP discharges small volumes of effluent, if water contains high amounts of EfOM per unit volume, the "actual" influence upon the stream may more significant than the measured (or vice versa), resulting in a misjudgement in wastewater management or a in selection of treatment processes. Thus, the volume-based method may be better for assessing wastewater impact (% wastewater impact). Also, primidone may be

inappropriate when evaluating wastewater impact in terms of % EfOM influence in spite of advantages as a wastewater tracer because it is solely anthropogenic and does not exist in pristine water. Although there are increasing concerns about pharmaceuticals in drinking water and wastewater treatment processes, still major interest is placed on efficient removals of organic matter targeting most parts of the DOC, and trace synthetic compounds (pharmaceuticals, endocrine disrupting compounds, etc) are considered to be negligibly responsible for bulk DOC concentration. Thus, wastewater impact evaluated by primidone data may not explain correctly the impact by wastewater-derived organic matter out of total organic carbon (i.e., EfOM impact) in corresponding locations. Furthermore, recent studies [55-58] showing sorptional removal possibilities of these pharmaceuticals or EDCs by colloidal and/or dissolved organic carbon make this chemical appropriate as a suitable wastewater indicator in terms of % EfOM impact. Therefore, in using two microbially-derived PARAFAC components, two perspectives were taken into consideration: first, if the PARAFAC components are successful in assessment of wastewater impact, and second, if the components can evaluate the % EfOM impact on downstream, i.e., how much EfOM actually contributed to changing of the overall organic matter properties in wastewater impacted downstream samples.

To begin with, variations of five PARAFAC components along the river flows were observed, and average fractions of each component for three sampling events are presented in Figure 3-9. Component 1 and component 5 showed high presences in the pristine water (site 1, the Chatfield reservoir) and relatively in less impacted water by wastewater, while low fraction in wastewater. Fraction of component 1 showed gradual decreases along the river, possibly due to relative high contribution by other fluorophore, or degradations or transformations into other organics by biodegradation and/or photolysis. Component 2 were present very low in pristine

water (site 1), but its fraction suddenly increased in site 3, where were impacted by upstream Allan WWTP discharge even though wastewater flow was minimal (<10%) compared to river flow. In WW and WW-impacted waters, component 2 showed always higher. Component 3 did not vary compared to other components. Component 4 showed mostly similar variations with component 2, showing increases in WWTPs and WW-impacted water. Of two different treatment processes in Denver Metro WWTPs, the South complex including the activated sludge process with the high purity oxygen and shorter sludge retention time (~3 days) exhibited higher fraction of both component 2 and component 4 than those of the north complex using nitrification/denitrification process (SRT: ~5 days). Considering they receive same water, the fractional differences of these two components present in final effluents may imply connectivity with differences in treatment processes. Component 5 showed more clearly opposite trend to component 2 or component 4, decreasing in WW and increasing in upstream water or less WW-impacted water.

On the whole, it was observed that the patterns of component 2 and component 4 were mostly similar to that of primidone and this supports that these components are applicable as indicator(s) for wastewater impact.

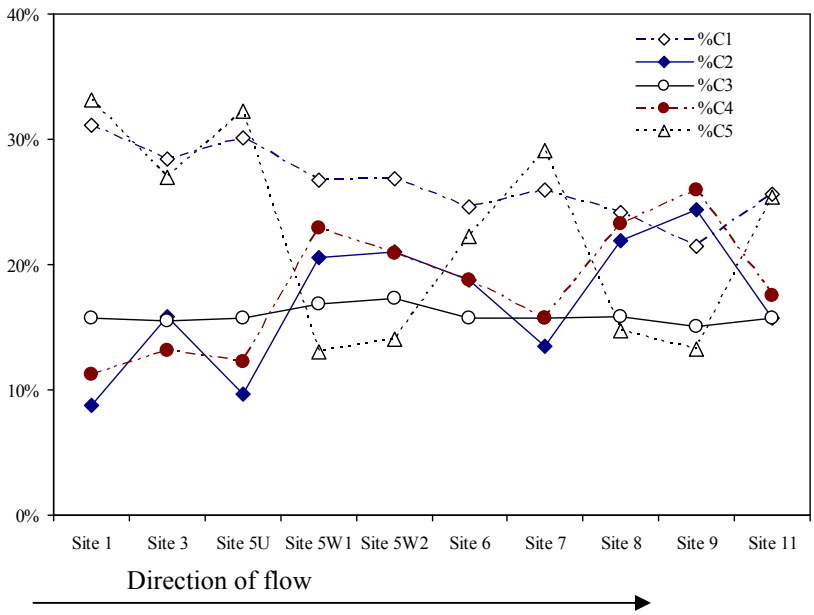


Figure 3-9 Variations of five components at each sampling sites in South Platte watershed (mean during three sampling events-Feb. and Sep. in 2004, and April in 2005); Site 5W1 and 5W2 denote respectively wastewater effluents Littleton/Englewood WWTP, which are after nitrification before chlorination (W1), and after chlorination (W2)

Table 3-5 shows the summary of wastewater-impact percentages at site 7 and site 11, in which amounts of component 2 and component 4 were calculated based on fractions of components and bulk DOC of samples. Wastewater impact by PARAFAC component showed similar trends by the flow and primidone approaches, i.e., <50% impact at site 7, >50% impact at site 11, indicating that not only is there a significant influence to SPR by Denver Metro WWTP discharges but selected PARAFAC components also would be successful indicators for wastewater impact evaluation. Figure 3-10 presents relations among wastewater impacts drawn by different approaches. Comparisons to flow-based results showed clear linearity to both primidone-based and PARAFAC components-based methods. Relatively great deviations between flow-based and other approaches appeared in the Sep, 2004 Site 7 results (summer sampling event). By the flow-based method, SPR at Site 7 in Sep. 2004 was ~50% of river water impacted to wastewater, but by PARAFAC components and by primidone-based approaches,

~39% (an average from C2, C4 or C2+C4) and ~26% of the flow were wastewater-impacted, respectively (also see Table 3-5). Especially, the result by the primidone-based method showed almost a 2- fold discrepancy with the flow-based method. In the regression equations, the y-axis intercepts mean that overall the flow-based method predicted less wastewater impact than when performed by the primidone-based or the PARAFAC components-based methods, showing ~10% less (i.e., intercept 0.1046) compared to the primidone-based method, and 16-24% less (intercepts 0.1649~0.2421) than the PARAFAC components-based method.

Figure 3-11 shows linear relations between by the primidone-based and by PARAFAC components-based methods. Compared to the flow-based results, generally, this showed better matches (particularly component 2 showed an r^2 of more than 0.99), and slopes closed to one.

The comparison of PARAFAC components-based results with the other two results confirmed that the PARAFAC components are successful indicators for addressing wastewater impact.

Table 3-5 Summary of wastewater impact by flow, primidone, organic carbon and PARAFAC components, and contribution of Denver Metro WWTP effluent to SPR

		%WW					%EfOM	%SMP
		Flow	Primidone	C2	C4	C2+C4		
Site 3	Feb, 2004	22%					57%	22%
	Sep, 2004	12%						
	April, 2005	4%						
Site 7	Feb, 2004	31%	28%	40%	45%	42%	60%	25%
	Sep, 2004	50%	26%	36%	43%	39%	73%	27%
	April, 2005	14%	23%	29%	37%	34%	20%	9%
	Average	32%	26%	35%	42%	38%	51%	21%
Site 11	Feb, 2004	61%	55%	70%	58%	64%	72%	33%
	Sep, 2004	78%	59%	74%	89%	81%	98%	38%
	April, 2005	42%	37%	50%	66%	59%	59%	30%
	Average	60%	50%	65%	71%	68%	76%	34%
		% increased WW contribution to SPR Site 11 by Denver Metro WWTP/North						
		Flow	Primidone	C2	C4	C2+C4	%EfOM	%SMP
	Feb, 2004	30%	27%	30%	13%	22%	12%	8%
	Sep, 2004	28%	33%	38%	46%	42%	25%	11%
	April, 2005	28%	14%	21%	29%	25%	38%	21%
	Average	29%	25%	30%	29%	30%	25%	13%

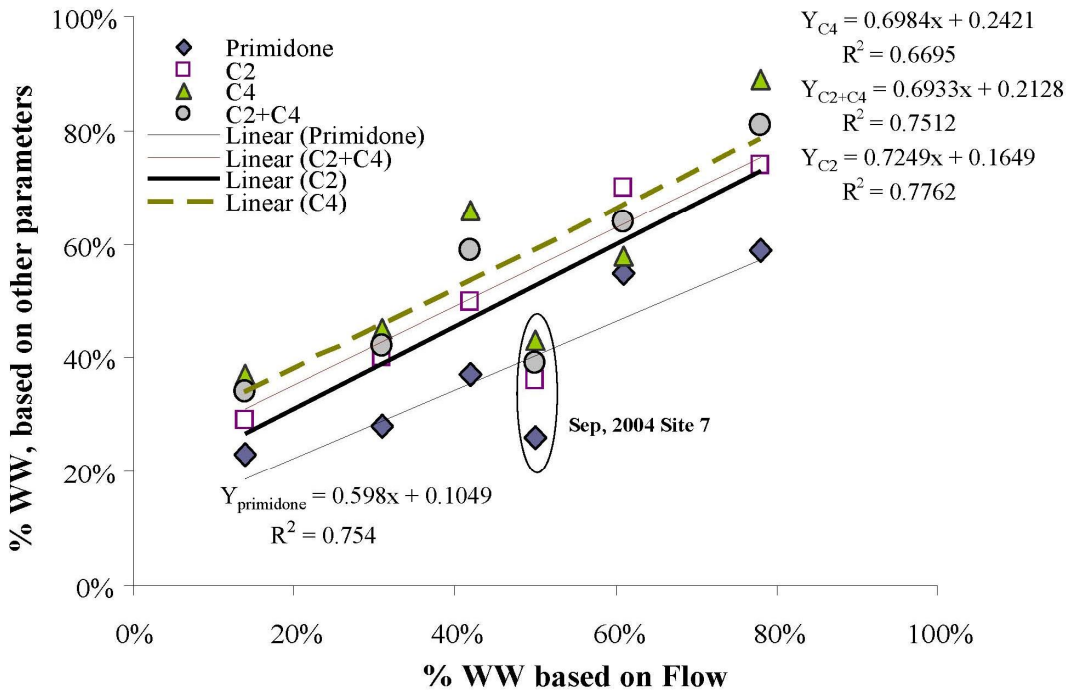


Figure 3-10 Wastewater impact (%) at Site 7 and Site 11 in South Platte River based on different wastewater indicators (flow, primidone and two PARAFAC components)

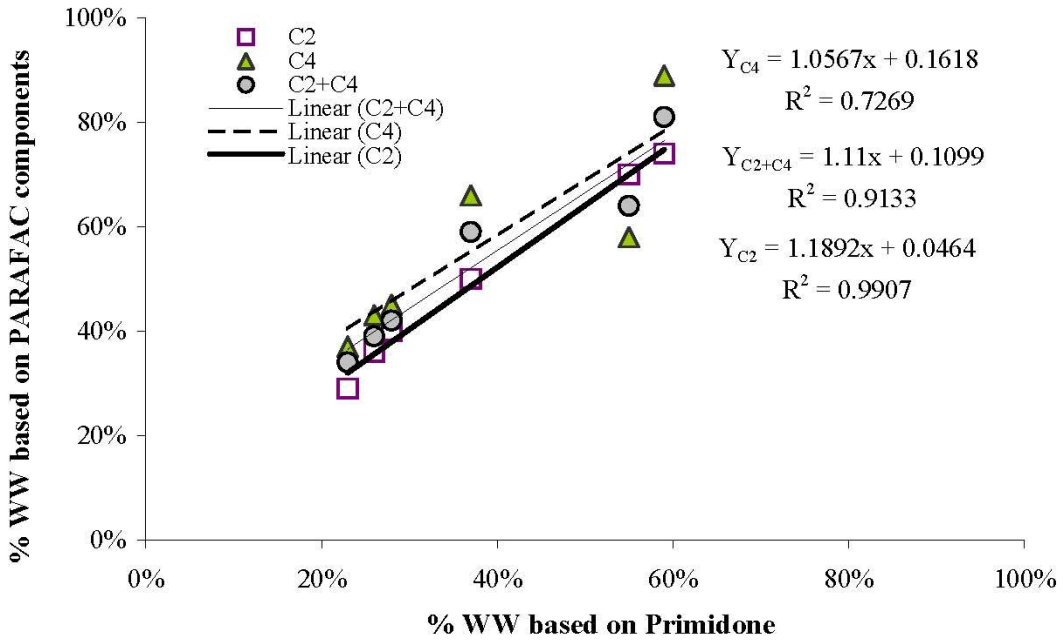


Figure 3-11 Wastewater impact (%) at Site 7 and Site 11 in South Platte River based on different wastewater indicators (flow, primidone and two PARAFAC components)

Meanwhile, an attempt has been attempted to determine balances of th mass of total organic carbon (kg) and the two microbially-derived PARAFAC components along the river flows to see how the river would be influenced by wastewater effluent organic matter, in terms of %EfOM and %SMP impact

By the widely-accepted definition of EfOM [59], the sum of persistent NOM from the DWTP, SMPs from the WWTP, and the amount of EfOM was counted as total mass of organic carbon (kg) discharged from the WWTPs, whose mass could be estimated using flow and DOC data of each WWTP. The total masses of organic carbon or PARAFAC components at each site are explained by the followings:

- X_1: amount (kg) of organic carbon or PARAFAC component present in Site 1 sample
- X_2: amount (kg) of organic carbon or PARAFAC component present in Centennial WWTP effluent
- X_3: an amount (kg) of organic carbon or PARAFAC component present in Site 3 sample (upstream of Allan DWTP, downstream of Centennial WWTP)
- $Unknown_1$: a possible amount (kg) of organic carbon or PARAFAC component present in water from Bear

Creek

- X_5: an amount (kg) of organic carbon or PARAFAC component present in Littleton/Englewood WWTP effluent sample
- X_6: an amount (kg) of organic carbon or PARAFAC component present in downstream of Littleton/Englewood WWTP
- $Unknown_2$: a possible amount (kg) of organic carbon or PARAFAC component present in water from Cherry Creek and Glendale WWTP
- X_7: an amount (kg) of organic carbon or PARAFAC component present in Site 7 sample
- X_B: an amount (kg) of organic carbon or PARAFAC component present in sample diverted to Burlington Ditch
- X_{11}: an amount (kg) of organic carbon or PARAFAC component present in Site 11 sample

The sites are related to as follows, and effects by unknown flows ($Unknown_1$ and $Unknown_2$) are assumed negligible:

- $X_3=X_1+X_2$ (here X_2 is unknown because the sample was not taken. Using unknown X_3 and X_1, however, can be found for Feb, 2004 (winter) sampling event.
- $X_6=X_3+X_5+Unknown_1=(X_1+X_2)+X_5+Unknown_1$
- $X_7=X_3+X_5+X_6=(X_1+X_2)+X_5+Unknown_1+Unknown_2$
- $X_{11}=X_7+X_8=(X_3+X_5+X_6)+X_8=\{(X_1+X_2)+X_5+X_6\}+X_8+X_B+Unknown_1+Unknown_2$

Hence, a portion out of the total mass of organic carbon at the downstream sites of SPR could be addressed according to %EfOM impact. Likewise, assuming that two microbially-derived PARAFAC components (component 2 and 4) were major constituents of SMPs, their contributions to SPR were assessed as %SMP impact to the stream. Equation 3-3 and 3-4 explains how %EfOM and %SMP were calculated, where *i* could be 3, 7 and 11, meaning sites 3, 7 and 11.

$$\%\text{EfOM} = \frac{mass\ of\ organic\ carbon\ from\ WWTP\ effluent\ to\ site_i\ (kg)}{mass\ of\ organic\ carbon\ present\ at\ site_i\ (kg)} \times 100 \qquad (3\text{-}3)$$

$$\%\text{SMP} = \frac{mass\ of\ C2+C4\ from\ WWTP\ effluent\ to\ site_i\ (kg)}{mass\ of\ organic\ carbon\ present\ at\ site_i\ (kg)} \times 100 \qquad (3\text{-}4)$$

Table 3-6 shows results of wastewater impact in terms of %EfOM impact and %SMP impact at downstream sites of WWTPs for three sampling events. By %EfOM impact, 20-98% of total organic carbon appeared to be wastewater-originated, and 9-38% of the organic carbon was related to SMPs. Both had linear relationships with the flow-based method, showing the regression equations of %EfOM = 0.9421 × %Flow-based + 0.227, r^2=0.82; %EfOM = 0.4119 × %Flow-based + 0.0812, r^2=0.87, respectively, and depeding on sites/sampling seasons 6-36% of higher influences than when using flow data. This implies that wastewater impact by only balancing flow does not necessarily explain the impact of EfOM because the flow-balancing does not count for variations in organic matter properties and it is strongly affected by the upstream flow. Especially for the case of spring sampling (April, 2005), the increase of upstream flow rate by a snow melting event at site 1 led to possibly underestimations of wastewater impact than actual impact on the stream. Between %EfOM and %SMP impact, it was also shown a linear relationship (%EfOM = 0.3733 × %SMP + 0.031, r^2=0.90).

Overall, in agreement with the flow-based and the primidone-based methods, it was observed that wastewater impact upon SPR was greatly increased by Denver Metro WWTP discharge at site 11 compared to site 7 (and site 3 for Feb, 2004 sampling), By the balancing-approaches of organic carbon and PARAFAC components, it was confirmed that SPR is an EfOM-*dominant* stream receiving more than 50% of total organic carbon from WWTPs as well as great amounts (up to 38% in maximum) of SMPs.

Table 3-6 Summary of wastewater-derived organic carbon impact to South Platte River

February, 2004 (Winter sampling event)														
		Flow	DOC	Total organic	C2			C4			C2+C4			
Locations		(mgd)	(mg/L)	carbon (kg)	% of DOC	mg/L	kg/day	% of	mg/L	kg/day	% of	mg/L	kg/day	
Site 1: Upstream of WWTPs (South Platte River below Chatfield Reservoir	X_1	15	3.1	174	8.9%	0.274	16	8.0%	0.24	14	16.9%	0.52	29	
Centennial WWTP	X_2	**4.1**	-	236	-	-	50		-	40		-	90	
Site 3 (upstream of Allan DWTP)=Site 1 + Centennial WWTP	X_3	**19.1**	5.7	410	15.9%	0.901	65	13.1%	0.75	54	29.0%	1.65	119	
Littleton/Englewood WWTP	X_5	**20.9**	10.9	860	21.3%	2.319	183	21.4%	2.32	184	42.7%	4.64	367	
Site 6: downstream of Littleton/Englewood WWTP, upstream of Farmers and Gardeners Ditch	X_6	**40.0**	6.66	1008	18.8%	1.250	189	18.7%	1.25	189	37.5%	2.50	378	
Site 7: downstream of Littleton/Englewood WWTP (SPR at Burlington Canal head gate)	X_7	81	6.0	1827	15.5%	0.925	284	17.6%	1.05	321	33.1%	1.97	604	
Burlington Ditch flow	X_B	0	-	0	-	-	-	-	-	-	v	-		
Denver Metro WWTP	X_8	**128**	11.7	5676	21.7%	2.531	1230	25.0%	2.92	1418	46.7%	5.45	2649	
Site 11	X_{11}	249	10.0	9381	17.1%	1.705	1606	15.1%	1.51	1419	32.2%	3.21	3026	
Total organic carbon (kg)=Burington Ditch + South Platte River at Site 11				9381			1606			1419			3026	
WWTPs: total flow for sample event		**153**												
Total river flow=Burington Ditch flow + South Platte River at Site 11		249												
		Flow-based		% EfOM impact			% OCwwtp C2 to Total OC			% OCwwtp C4 to Total OC			% SMP	
% of river flow at Site 11 due to treated wastewater (from all WWTPs)		**62%**		**72%**			16%			18%			**33%**	
% of river flow at Site 7 due to treated wastewater (from 2 WWTPs)		**31%**		**60%**			13%			12%			**25%**	
% of river flow at Site 3 due to treated wastewater (from Centennial WWTP)		**22%**		**57%**			12%			10%			**22%**	
September, 2004 (Summer sampling event)														
		Flow	DOC	Total organic	C2			C4			C2+C4			
Locations		(mgd)	(mg/L)	carbon (kg)	% of DOC	mg/L	kg/day	% of	mg/L	kg/day	% of	mg/L	kg/day	
Site 1: Upstream of WWTPs (South Platte River below Chatfield Reservoir	X_1	34	3.7	470	8.5%	0.311	40	8.2%	0.300	39	16.7%	0.610	79	
Centennial WWTP	X_2	**4.7**	-	-	-	-	-	-	-	-		-	-	
Site 3 (upstream of Allan DWTP)=Site 1 + Centennial WWTP	X_3	**38.7**	-	-	-	-	-	-	-	-		-	-	
Littleton/Englewood WWTP	X_5	**21.8**	9.2	755	20.7%	1.898	157	16.7%	1.531	126	37.5%	3.429	283	
Site 6: downstream of Littleton/Englewood WWTP, upstream of Farmers and Gardeners Ditch	X_6	**-**	-	-	-	-	-	-	-	-		-	-	
Site 7: downstream of Littleton/Englewood WWTP (SPR at Burlington Canal head gate)	X_7	53	5.2	1035	13.3%	0.686	138	12.7%	0.653	131	25.9%	1.339	269	
Burlington Ditch flow	X_B	53	5.2	1035	13.3%	0.686	138	12.7%	0.653	131	25.9%	1.339	269	
Denver Metro WWTP	X_8	**130**	10.2	5019	21.2%	2.166	1066	17.7%	1.809	890	39.0%	3.975	1956	
Site 11	X_{11}	149	8.6	4839	17.6%	1.507	850	17.4%	1.493	842	35.0%	3.000	1692	
Total organic carbon (kg)=Burington Ditch + South Platte River at Site 11				5875			987			973			1960	
WWTPs: total flow for sample event		**157**												
Total river flow=Burington Ditch flow + South Platte River at Site 11		202												
		Flow-based		% EfOM impact			% OCwwtp C2 to Total OC			% OCwwtp C4 to Total OC			% SMP	
% of river flow at Site 11 due to treated wastewater (from all WWTPs)		**77%**		**98%**			21%			17%			**38%**	
% of river flow at Site 7 due to treated wastewater (from 2 WWTPs)		**50%**		**73%**			15%			12%			**27%**	
% of river flow at Site 3 due to treated wastewater (from Centennial WWTP)		**12%**		-			-			-			-	
April, 2005 (Spring sampling event)														
		Flow	DOC	Total organic	C2			C4			C2+C4			
Locations		(mgd)	(mg/L)	carbon (kg)	% of DOC	mg/L	kg/day	% of	mg/L	kg/day	% of	mg/L	kg/day	
Site 1: Upstream of WWTPs (South Platte River below Chatfield Reservoir	X_1	122	3.05	1409	8.9%	0.270	125	17.6%	0.538	248	26.5%	0.808	373	
Centennial WWTP	X_2	**5.5**	-	-	-	-	-	-	-	-	-	-	-	
Site 3 (upstream of Allan DWTP)=Site 1 + Centennial WWTP	X_3	**127.5**	-	-	-	-	-	-	-	-	-	-	-	
Littleton/Englewood WWTP	X_5	**20**	8.80	666	20.8%	1.831	139	24.4%	2.151	163	45.3%	3.983	302	
Site 6: downstream of Littleton/Englewood WWTP, upstream of Farmers and Gardeners Ditch	X_6	**-**	-	-	-	-	-	-	-	-	-	-	-	
Site 7: downstream of Littleton/Englewood WWTP (SPR at Burlington Canal head gate)	X_7	184	4.67	3253	11.5%	0.539	376	17.1%	0.800	557	28.7%	1.340	933	
Burlington Ditch flow	X_B	126	4.67	2227	11.5%	0.539	257	17.1%	0.800	382	28.7%	1.340	639	
Denver Metro WWTP	X_8	**120**	9.84	4470	25.3%	2.493	1133	27.1%	2.666	1211	52.4%	5.160	2344	
Site 11	X_{11}	216	7.99	6533	12.6%	1.010	826	19.9%	1.592	1302	32.6%	2.602	2128	
Total organic carbon (kg)=Burington Ditch + South Platte River at Site 11				8760			1083			1684			2767	
WWTPs: total flow for sample event		**146**												
Total river flow=Burington Ditch flow + South Platte River at Site 11		342												
		Flow-based		% EfOM impact			% OCwwtp C2 to Total OC			% OCwwtp C4 to Total OC			% SMP	
% of river flow at Site 11 due to treated wastewater (from all WWTPs)		**43%**		**59%**			15%			16%			**30%**	
% of river flow at Site 7 due to treated wastewater (from 2 WWTPs)		**14%**		**20%**			4%			5%			**9%**	
% of river flow at Site 3 due to treated wastewater (from Centennial WWTP)		**4%**		-			-			-			-	

Application of PARAFAC to EfOM biodegradation: contribution of each component to BDOC

In order to understand the biodegradability of EfOM in terms of the contribution of each component to BDOC, PARAFAC analysis was applied to BDOC samples. Figure 3-12 depicts the aerobic 5-day kinetic (daily) BDOC test of an EfOM-*dominant* stream in the Southeast of USA. The $BDOC_5$ was 7.1 mg/L which corresponds to 47.3% of initial DOC. For the first 2 days, biodegradation was somewhat fast showing 4.88 mg/L of BDOC (68.6% of total $BDOC_5$), this is generally known as easily (fast) biodegradable DOC, and then later slow biodegradation was observed.

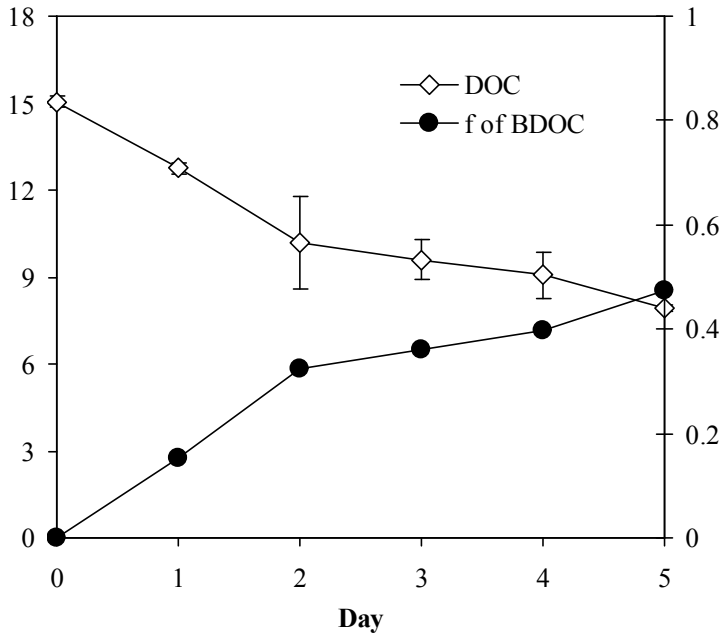


Figure 3-12 5-day BDOC kinetic tests for the EfOM-*dominant* stream in the Southeast of USA (taken in June, 2005). Bars represent the standard deviation (n=3)

Figure 3-13 shows that the fractional changes (a) and variations of concentration (b) of each component during the 5-day kinetic BDOC test for the same sample shown in Figure 3-10. Sub-graphs (c) and (d) represent the contributions of each component to $BDOC_t$ in percentages and

amounts, respectively. As shown in (a), in the sample, there was a predominance of component 4, followed by component 1. Component 2 and 3 showed similar amounts and component 5 represented less than 10% (below 1 mg/L). During biodegradation, decreases in amounts of component 4 were greater than other components, as sub-graphs (c) and (d) show that this component greatly contributed to BDOC at each day. Component 1 also contributed somewhat to BDOC, and its degradation was high at the early stage of $BDOC_t$ with a later reduced contribution to BDOC (sub-graph-(c)). Considering its higher amounts of the initial concentration compared to component 2, component 1 is a less biodegradable substance than component 2, and its (relative) fraction showed an increase as also depicted in sub-graph (a) (see the slope of each component). Figure 3-14 displays the fractional contribution of each component to the $BDOC_5$. Overall, in this EfOM-*dominant* stream, more than 60% of $BDOC_5$ was attributed to component 4 and component 2, and 23% of the $BDOC_5$ by component 1, implying that the protein-like materials contribute most to the biodegradability of EfOM. Of humic-like components 1, 3 and 5, as previously discussed, component 1 and component 5 have clear strong relevance with hydrophobic and aromatic properties with carbon double bonds difficult to degrade by microorganisms. As many studies have proved that a advanced oxidation processes such as an ozonation or any hydroxyl radical generating reaction, can enhance the biodegradability of refractory organic substances (in this case, mostly component 1 and component 5), with the biodegradability enhancing mechanism mainly being breaking the double bonds and transforming molecules to more easily microbially-accessible and smaller molecules [60, 61]; additional pre-treatment may help to further understand the PARAFAC components as well as reaction mechanisms including their fates and intermediates.

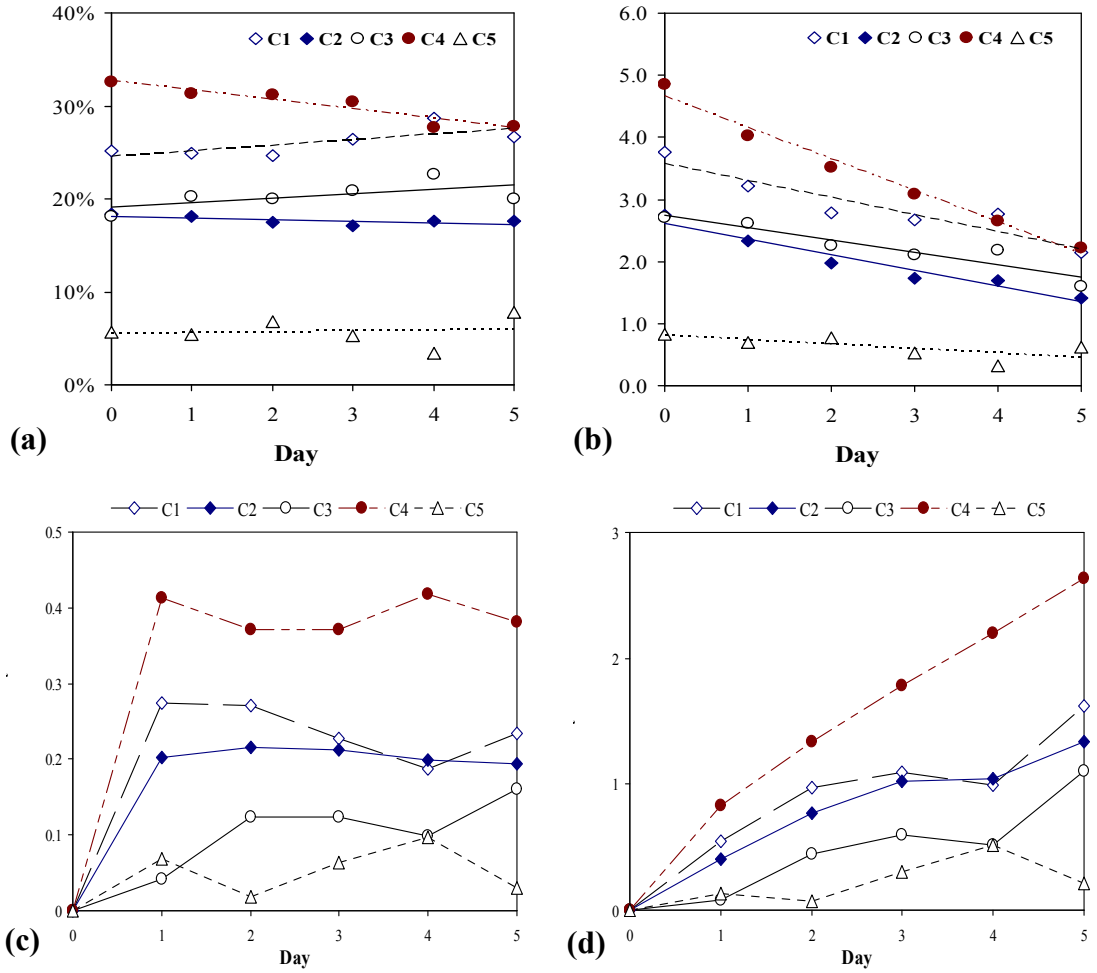


Figure 3-13 Fractional and concentration changes of each component during 5-day aerobic BDOC kinetic test for the EfOM-*dominant* stream in the Southeast of USA (taken in June, 2005)

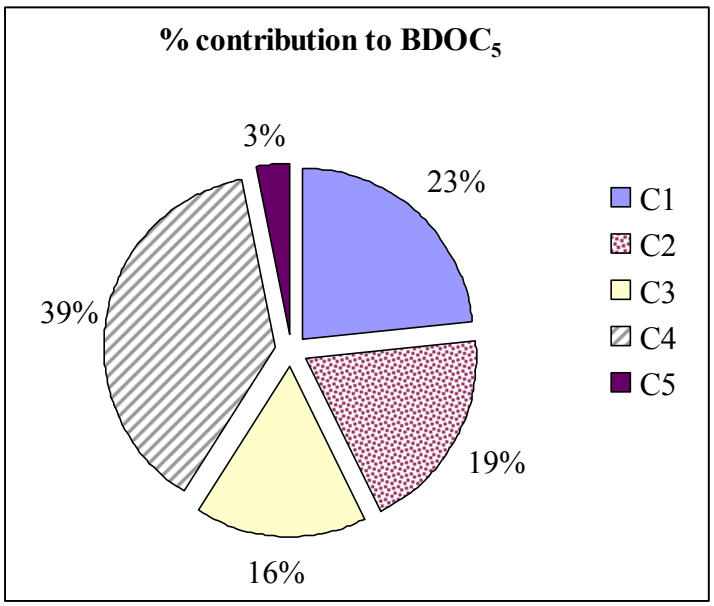


Figure 3-14 Contributions to aerobic $BDOC_5$ of each component for the EfOM-*dominant* stream in the Southeast of USA (taken in June, 2005)

Concluding remarks, implications, and further study of the chapter

This study showed the development of a distinct PARAFAC model that identified 5 components based on a wastewater EfOM-focused dataset, but still successfully elucidating surface water and drinking water NOMs (~25% of whole EEMs). As wastewater fingerprints, component 2 and component 4 turned out to be appropriate along with one NOM-similar component (component 5). Furthermore, in addition to comparison to literatures' results, investigation of components in reference NOM and EfOM isolates in terms of fraction and MWs as well as statistical analyses for fractions of components and other external NOM characterization parameters from the whole dataset have been explainable based on physico-chemical properties of components, and this approach was unique compared to other groups. A case study of PARAFAC components applied to the EfOM-*dominant* South Platte River watershed along with primidone data variations also confirmed a successful applicability of PARAFAC components to a real watershed system for assessing wastewater impacts.

PARAFAC application to a $BDOC_5$ test showed that biodegradation was more preferentially targeting protein-like components (components 2 and 4) rather than humic-like components. This also suggests that the 5-components PARAFAC model can be practically used for the purpose of reaction mechanisms or investigation of NOM fates/transformation pathways in any fluorophore-involved process. Regarding aspects of further study, there are few things to remark. In PARAFAC, there are two most important conditions that enable it to enhance resolving fluorophores in addition to reliable instrument, correct inter-filter correction, etc., and those are the number of EEMs, and the presence of reproducible spectral patterns of fluorophores. Increased numbers of EEMs not only can facilitate the modeling, but can also provide better variability of different fluorophores. Stedmon *et al.* (2003) identified 5 components in marine and fresh waters using 90 EEMs [25], and the same researchers (2005) have presented an extended 8 components for the same sampling region by increasing the number of EEMs to n=1,276 [39]. Recently, Cory *et al.* (2006) presented a 13- components PARAFAC model using 379 EEMs, in which most of samples were from well-controlled water systems (such as lakes in the Antarctica and sub-alpine lakes), which exhibited less variations of contained fluorophores, and were isolated from impact by other contamination sources other than bulk organic matter, so that the DOM contained may be more homogeneous terrestrial or microbial DOM end members. Moreover, fluorophores may experience a similar or routine reaction pathway, resulting in generation of repeatable/reproducible patterns of intermediate or by-product fluorophores. On the other hand, fluorophores in a water system that is more vulnerable to environmental controlling factors such as wastewater may be associated with various (unpredictable) reactions and physicochemical parameters, consequently, less possible to exhibit the same spectral patterns as a single component. Therefore, these previous research studies evoke a possibility that for the

extended sampling periods, accumulating the number of EEMs at the same sampling locations depending on types of WWTP may identify more components which may be sub-fluorophores of some of the original 5 components, or independent new components that could not be precisely resolved and overlapped under some of the 5 components. Meanwhile, since collection of a large dataset takes much time and energy, another approach would be to try to fit EEMs used for developing the 5 -components PARAFAC model (or any EEMs in general cases) to a built-in PARAFAC model (e.g., the model shown in Cory *et al.*).

CHAPTER 4

COMPARISONS OF THE EfOM-CENTERED PARAFAC COMPONENTS TO THE EXISTING PARAFAC MODEL

Introduction

Over the past decade, organic matter characterization using fluorescence excitation-emission matrix (EEM) spectroscopy has been increasingly applied in many research areas including environmental research [17, 20, 62, 63]. At the same time, multivariate statistical analyses such as parallel factor analysis (PARAFAC) incorporated with EEMs has been shown as a powerful tool to understand dissolved organic matter derived from different origins/locations [23, 24, 44].

Theoretically, the spectra of each component resolved by PARAFAC are the pure fluorophores, which means the numbers of components are matched to the numbers of fluorophores. In reality, however, the complex nature of fluorophores, especially dissolved organic matter, and their diverse origins, and the daily-based resolution of measurements may lead to a single fluorophore assigned to a class of compounds. Such possibilities have already been discussed by several other researchers [23, 39]. Correct resolution of composites of fluorescing organic matter into subordinate or separated individual fluorophores enables one to examine physicochemical phenomena in further depth although incorrect resolution may lead to confusion. In the previous chapter, a five (5) component PARAFAC model was established with 423 EEMs. Although 5 components were considered to be appropriate, there is the possibility of

some of the individual components being a mixture (composite) of compounds' fluorophores. Accordingly, the work presented in this chapter is based on the hypothesis that the 5 components or some of components identified in the N&A model (i.e., 5-components PARAFAC model in previous chapter, hereafter will be called as N&A model) may be a complex of separate fluorophores which can be resolved separately in fitting to a 13-component PAFAFAC model developed by Cory and McKnight [24].

The purpose of this study was to compare the 5 components identified in this study with the 13 component PARAFAC model [24] (hereafter called the C&M model) in order to see whether the components are composites or not and, if so, to address in more detail which components are involved with processes such as coagulation or biodegradation, and other drinking water treatment processes.

While there have been similar types of fitting studies, most of the fitting was constrained to elucidate the unknown concentration of known fluorophores or for the purpose of validating the reproducibility of the built model [21, 28, 29]. Therefore, in practical aspects, this work has meaning in revalidating the 5 N&A model components, and in providing further validation of the 13 C&M model components as well.

Materials and Methods

Fitting procedures

Since matrix dimensions or data structures of EEMs should be consistent to those of the C&M model, pre-dataset processing was necessary to be able to fit to the model. In fitting to the C&M model, there are two important processing steps in preparation of a dataset. One was to have both datasets to have the same matrices dimensions, and the other was to make the measured

excitation and emission ranges matched to each other because the MATLAB program run by a written-code does not automatically recognize the starting/ending ranges of data. A single EEM in the 13 component C&M model, which was established with 379 EEMs, consists of Ex: 250-400 nm by 5 nm increments, and Em: 350-550 nm by 2 nm increments, which results in a 101 × 43 (number of Emission × number of Excitation) matrix, whereas the EEM used in the 5 components N&A model originally conformed to a 201 × 43 matrix measured in the ranges of Ex: 240-450 by 5 nm and Em: 300-500 nm by 1 nm increments. Hence, the EEMs of the 5 component model should be cut in the 240-250 nm and 400-450 nm ranges of Ex, and in the 300-350 nm ranges of Em, which the data removed corresponding to 31.3% of the total data, also, it was a limitation that most of the protein-like peaks shown in Em: 300-350 nm could not be fit to the C&M model since the model does not have data in this region. Also in the C&M model, data within Ex: 250-400 nm and Em: 500-550 nm, explaining 28.6% of the total information, could not be utilized in the fitting. Thus, ca 70% of the EEMs in the N&A model were fit to ca 70% of the C&M model. The region in the fitting from both datasets is described by the shaded-area in Figure 4-1. Detailed explanations including sample conditions, measurement and data handlings associated with the C&M model can be found in Cory et al. [24]. Table 4-1 summarizes the experimental conditions applied to generate the EEM dataset.

Table 4-1 Summary of experimental conditions between two PARAFAC dataset

	Nam & Amy 5 components	Cory & McKnight 13 components
Excitation ranges, increments (nm)	240-450 nm, 5 nm	250-400 nm, 5 nm
Emission ranges, increments (nm)	300-500* nm, 2 nm	350-550 nm, 2 nm
Dilution criteria	$DOC_{sample} \leq 1.0$ mg/L	$UVA_{254\ nm} \approx 0.02$ cm^{-1}
Dilution solution	0.01M KCl	Milli-Q® water
Sample pH	2.7-2.8 (below 3)	≈ near neutral pH (with HCl or NaOH)

*: measured over 290-530 nm, and prepared for data with 300-500 nm

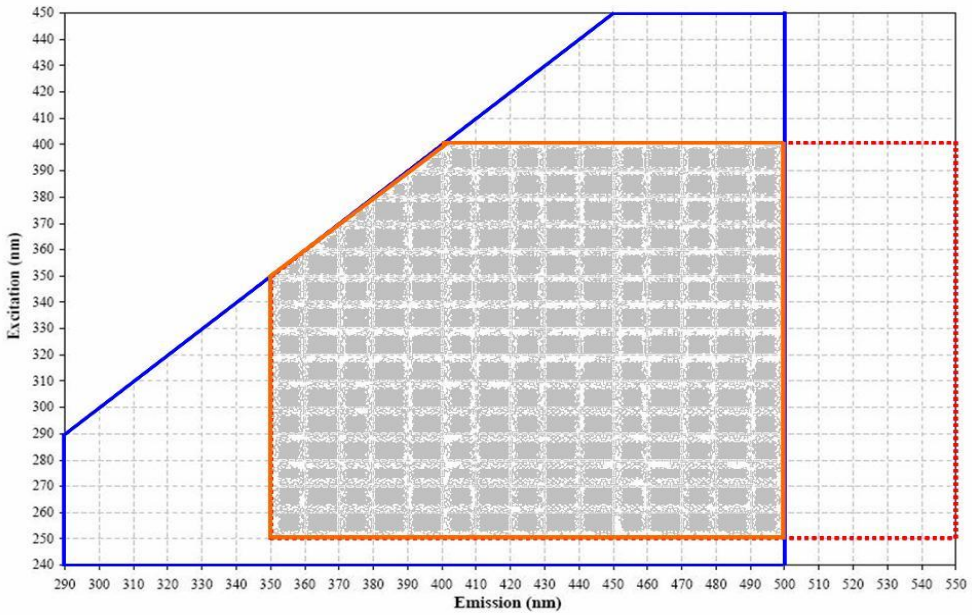
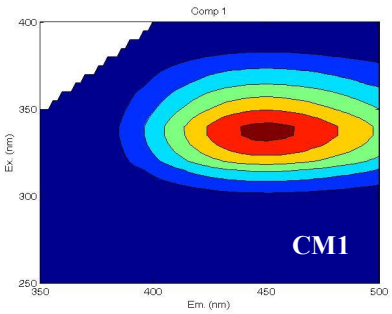
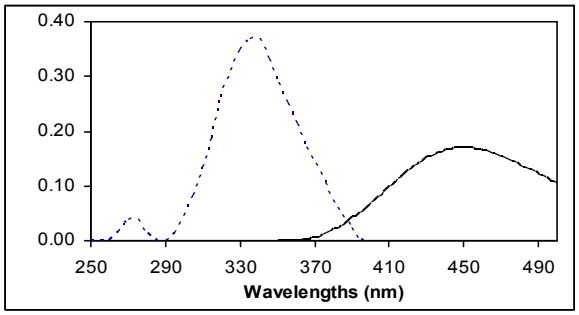
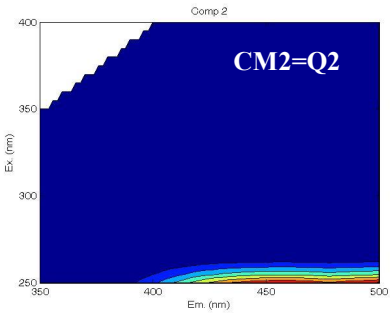
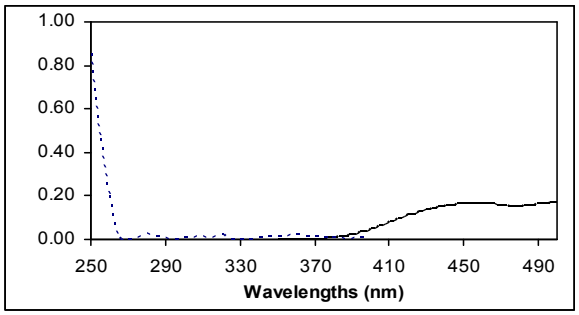


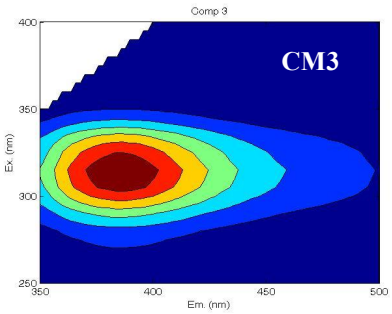
Figure 4-1 Comparison of excitation and emission wavelengths ranges of original EEMs between Cory and McKnight 13 components PARAFAC model [24] and this study; dotted-line for 13 components model (Ex: 350-550 nm and Em: 350-400 nm), thick bold line for this study (Ex: 240-450 nm, Em: 290-500 nm), shaded-region for fitting to Cory and McKnight 13 components model (Ex: 350-500 nm, Ex: 250-400 nm)

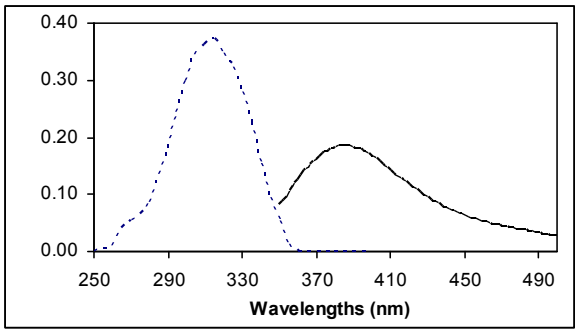

Comp 1
CM1
Ex. (nm)
Em. (nm)

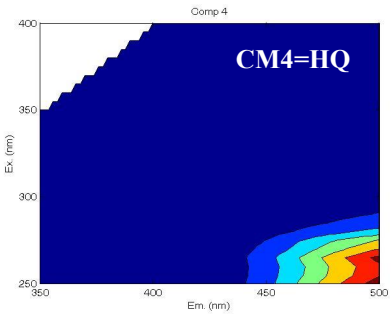

Wavelengths (nm)

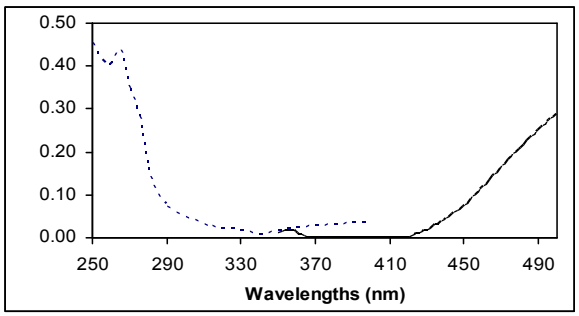

Comp 2
CM2=Q2
Ex. (nm)
Em. (nm)

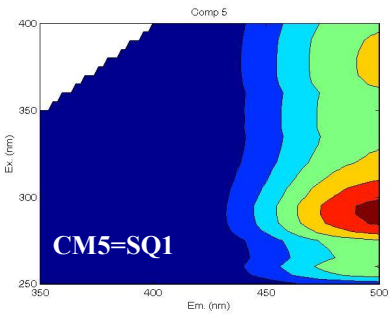

Wavelengths (nm)

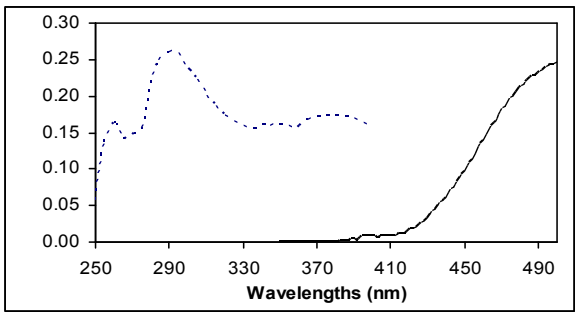

Comp 3
CM3
Ex. (nm)
Em. (nm)

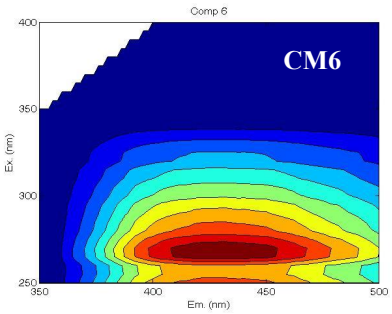

Wavelengths (nm)

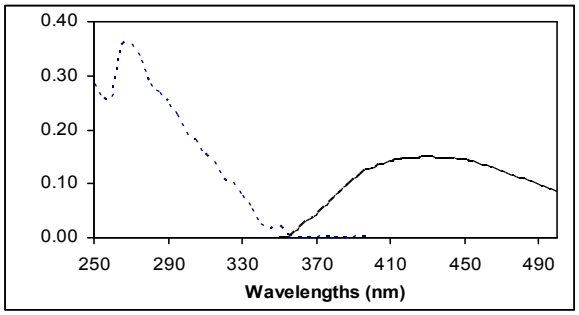

Comp 4
CM4=HQ
Ex. (nm)
Em. (nm)

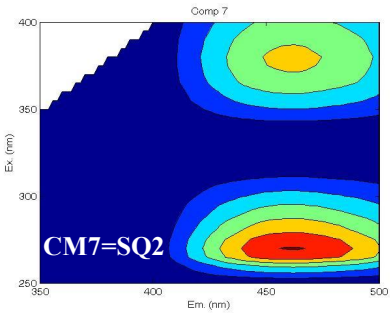

Wavelengths (nm)

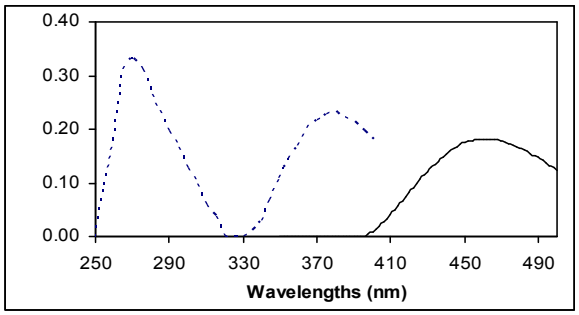

Comp 5
CM5=SQ1
Ex. (nm)
Em. (nm)

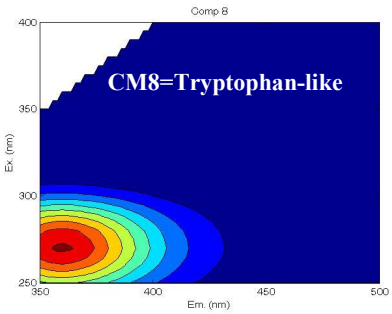

Wavelengths (nm)

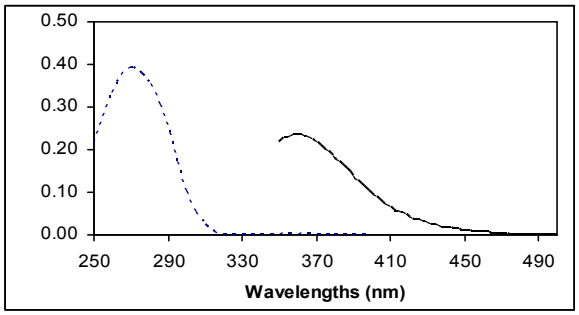

Comp 6
CM6
Ex. (nm)
Em. (nm)

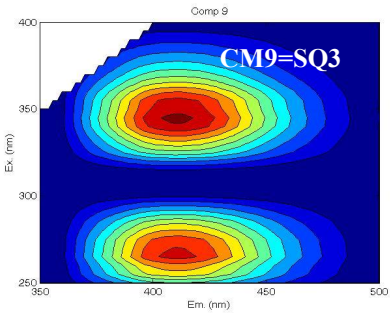

Wavelengths (nm)

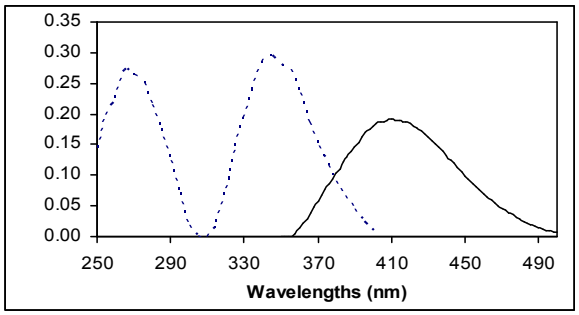

Comp 7
CM7=SQ2
Ex. (nm)
Em. (nm)

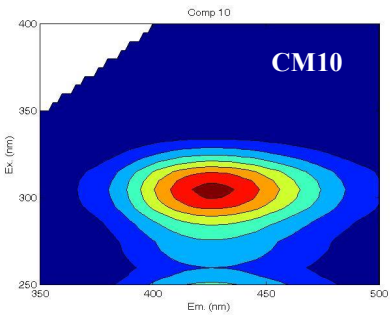

Wavelengths (nm)

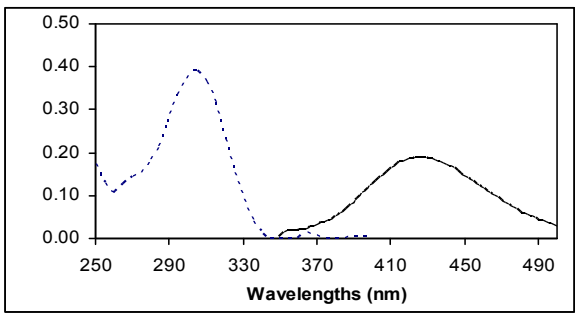

Comp 8
CM8=Tryptophan-like
Ex. (nm)
Em. (nm)


Wavelengths (nm)


Comp 9
CM9=SQ3
Ex. (nm)
Em. (nm)


Wavelengths (nm)


Comp 10
CM10
Ex. (nm)
Em. (nm)


Wavelengths (nm)

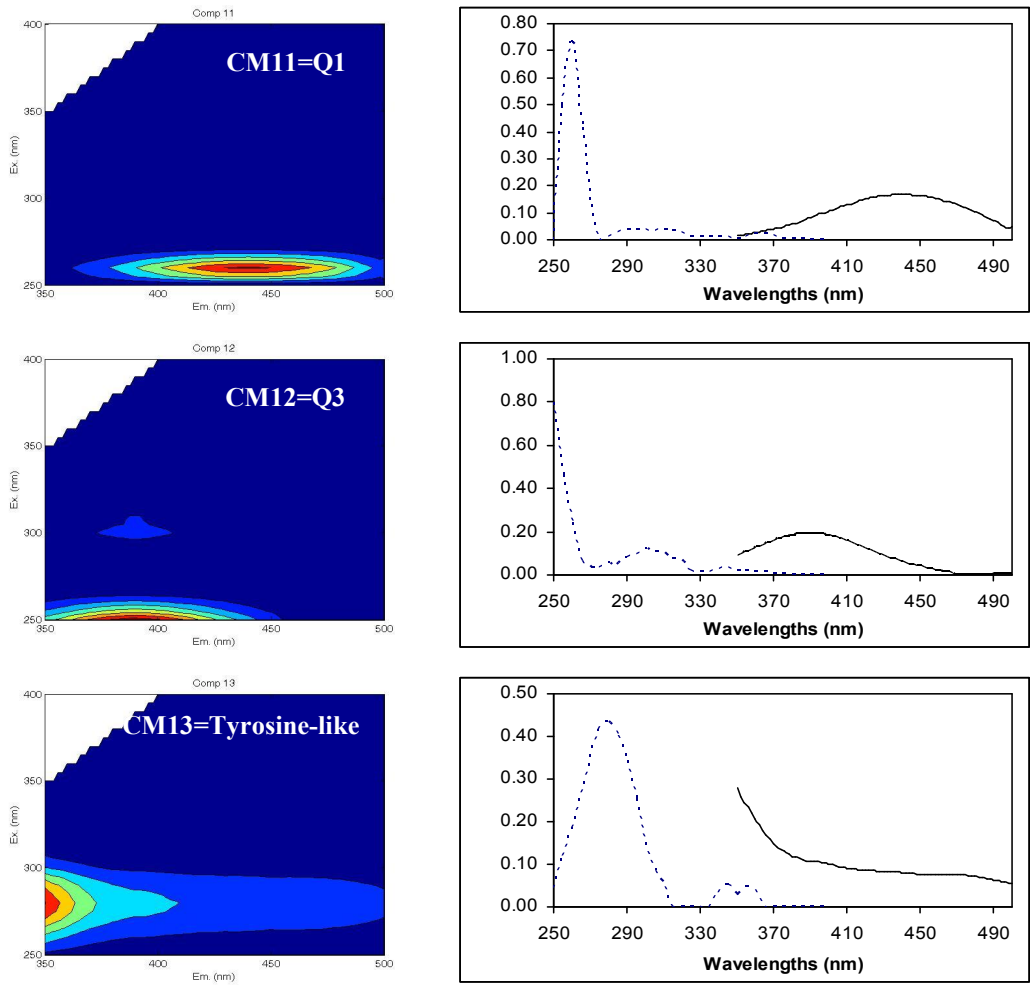


Figure 4-2 13 components contours (left column) and excitation and emission loadings of corresponding components (right column); also available in Ref. [24] and [64]

Figure 4-2 shows the contours and excitation/emission spectra of 13 components identified by the C&M model. For more easy recognition, components from the C&M model will be referred to as "CM1, CM2, ... , CM13" whereas components identified in this N&A study will be referred to as "C1, C2, ..., C5". According to the classification by Cory and McKnight, CM11, CM2, CM12, CM5, CM7, CM9 and CM4 were originally designated as Q1, Q2, Q3, SQ1, SQ2, SQ3 and HQ, respectively and these 7 components are quinone-like components. In addition, CM8 and CM13 were identified as tryptophan-like and tyrosine-like components, and CM1, CM3, CM9 and CM10 were designated as unknown compounds [64].

Results and discussion

Fitting new EEMs (i.e., here EEMs of the N&A model) to verified components of the established PARAFAC model (i.e. here the C&M model), known as second-order calibration, is distinctly different from building a new PARAFAC model. It is comparable to fitting of some parameters to a linear (or nonlinear) regression model although there are differences in that the former is a tri-linear fitting into x, y and z axes (in the case of EEM, excitation wavelength, emission wavelength and concentration (i.e., intensity)) while the latter is a bi-linear fitting (x, y axes regardless of whether linear or non-linear, or numbers of x variables). In fitting EEMs, assuming that score values and concentrations of each component are linear (i.e., trilinearity), a PARAFAC model will try to separate fluorophores of fitted EEMs uniquely into a unit sample, unit excitation spectra, and unit emission spectra. If the PARAFAC model can fully explain all fluorophores in the model, the residual EEMs simply have low residuals compared with the measured spectra (EEMs), and residuals are randomly distributed showing an irregular noise peak. Rather, if the PARAFAC model explains a part of the samples, a high residual and large leverage are shown, meaning that some of the fluorophores in the new EEM cannot be predicted well either by the presence of unknown analytes that are not present in the model, or by other analytical factors influencing the fluorescence spectra of analytes [36, 65].

General remarks or trends in fitting

Overall, it was found that a majority of the EEMs showed clear residual EEMs. The intensity of the residual was often high up to 10% of original fluorescence intensity (loading values) at pertinent positions. In most of the residual EEMs, similar patterns with discernable peaks were repeated, where residual EEMs are from measured sample EEMs (5-component PAFAFAC

model) minus model EEMs by fit to the C&M model. Mostly, four distinct positive residual peaks (peaks A, B, C, D in Figure 5-3) were observed: peak A at Ex: 270-290 nm and Em: 360-420 nm, peak B at Ex: 350-370 nm and Em: 400-450 nm, peak C at Ex: 270-290 and Em: after 460 nm, and peak D at Ex: 350-370 nm and Em: 440-490 nm. In terms of trends, peaks A and B become intense and clear in wastewater effluent or samples with high hydrophilicity, while peaks C and D are associated with upstream river samples. Peaks B and D appeared in most of the EEMs with variations in extent of wastewater and river water.

Peak A occurs in regions of tryptophane- and tyrosine-like components of the C&M model. Peak B overlaps with CM1 and the second maximum peak of SQ2, and peak C belongs to the first maximum peak of SQ2. Interestingly, Suwannee River humic/fulvic acid (SRHA, SRFA) and SRNOM samples also showed distinct residual peaks C and D, and these residual levels/patterns were very similar. Fitting of isolated effluent organic matter in 5 different concentrations also showed similar residuals of peaks A, B and C, indicating the presence of a consistent error between model and (similar types of) samples.

With two datasets, several factors such as pH, compositions of sample (or sample type), analytes concentrations, ionic strength, etc., can be considered as reasons for discrepancies between sample EEMs and model EEMs regardless of a negligible residual or significance level. Of these factors, two probable reasons, pH and samples types, have initially been considered based on major differences in producing datasets by the two different research groups. Mobed *et al*. [66] have reported that 0-1M of KCl did not affect the fluorescence of HS. Hence, although the ionic strength of our samples (0.01M KCl) was higher than those of samples for the C&M model, a possible effect by ionic strength was considered negligible. Thus, different sample pH can be one of the main reasons. As aforementioned, fluorophores in the C&M model were

measured at near neutral pH either by adding HCl and NaOH, or by no control, whereas those in this study were measured under acidic conditions (~2.8) since a large portion of samples included in the dataset were from WWTP or wastewater related samplings sites. Lowering pH to an acidic level with dilution was necessary to exclude or minimize any interference effect by metal-binding or complexation (e.g., quenching effect) between organic matter and inorganics.

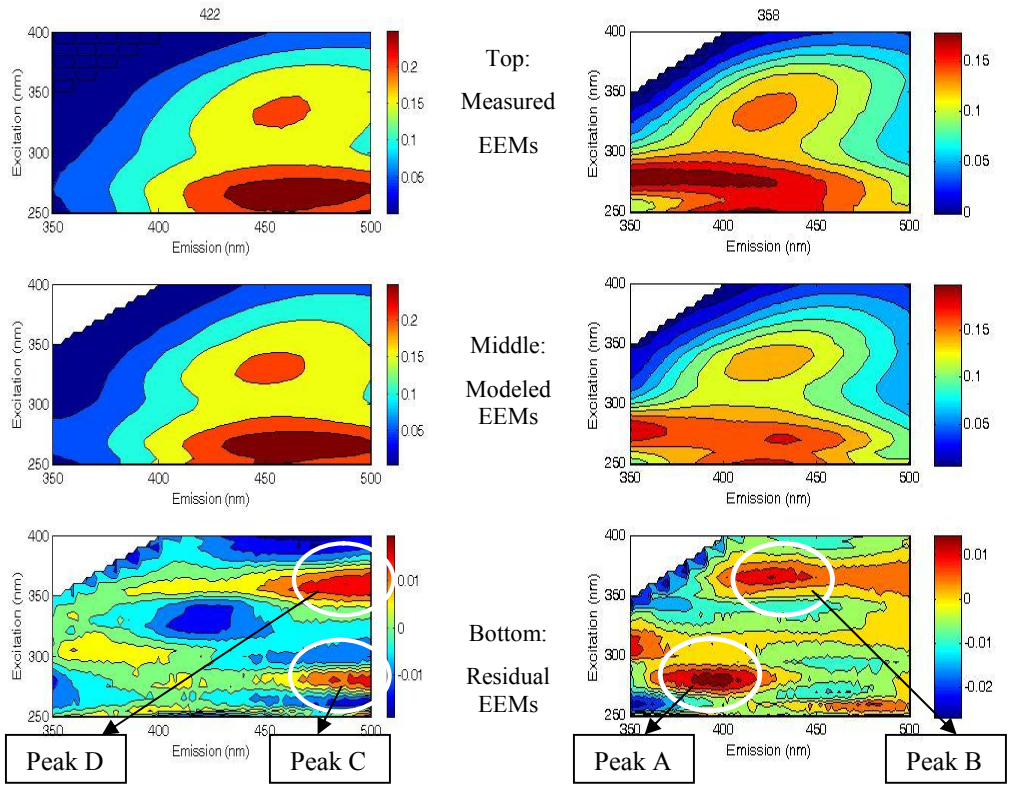


Figure 4-3 Examples of measured (top), modeled (middle), and residual (bottom) EEMs ; left-upstream river sample, right-wastewater with high residual

pH involves protonation or dissociation of acidic/basic functional groups, and the functional group undergoing protonation or dissociation may be directly connected to an aromatic structure, thus, whether it gains or loses electronic charge when going from the ground to the excited state, the electronic charge can cause changes in the spectral shifting, the spectral shaping as well as the fluorescence intensity. It is known that the protonation of electronic-withdrawing groups (e.g.,

carboxyl and carbonyl) causes shifts to longer wavelengths while the protonation of electronic groups such as amino groups shifts spectra to shorter wavelengths [37, 38]. Ionization of the phenolic hydroxyl group is also a function of pH. Ewald *et al*. [61] reported the spectral shifting of emission spectra of a fraction of FA of microbial origin with red-shifting at higher pH, and blue-shifting at low pH. Senesi N. [50] has presented the effect of pH upon the fluorescence intensity and the changes of spectrum shaping for soil FA. Reduction of a quinone-moiety to semiquinone, followed by reduction to the hydroquinone is also affected by pH.

Sample condition for the N&A dataset might have provided a favorable environment for a reduced quinone (hydroquinone) compared to an oxidized one, resulting in slightly different patterns of spectra compared to ones in the C&M model for the same fluorophores.

The composition of samples in relation to fluorophore origin has also been considered as a possible reason. Although the C&M model contains fluorophores of a microbial-end member, chemical properties or major constituents of microbial-derived organic matter between the C&M model and the N&A model may differ. The location of the residual peak A is in a valley between protein-like and humic peaks, this valley is known as the region of a SMP-like peak (Ex: 280 nm, Em: 350 nm) [62, 67]. According to previous researchers (see Ref. [68, 69]) and the N&A model survey [32, 70], and considering that 78% of the EEMs used in the N&A model contain more than 30% of microbial-related fluorophores (i.e., C2 and C4) regardless of whether these are composites of some sub-fluorophores, or not, it may be said that the N&A model is more wastewater effluent organic matter-centered/inclined, whereas the C&M model is relevant to dissolved organic matter in surface waters such as lakes or reservoirs.

Figure 4-4 shows a combination of all of 13 components (left), and 5 components (right), composing the C&M model and the N&A model, respectively identified by PARAFAC. These

figures indicate that the EEMs used in the two models are mostly like these contours, further supporting the above statement about relevance/applicability of the two models.

Therefore, in brief conclusion on overall fitting results, it was found that the shifting effect of emission spectra by pH change may have significance and should not be excluded in fitting EEMs to the built PARAFAC model. Moreover, this implies that the pH of samples should be consistent if it is proved that PARAFAC can resolve precisely the same fluorophore to different fluorophores at different pHs levels. Otherwise, a new PARAFAC model for each pH should be made, which is a cumbersome task.

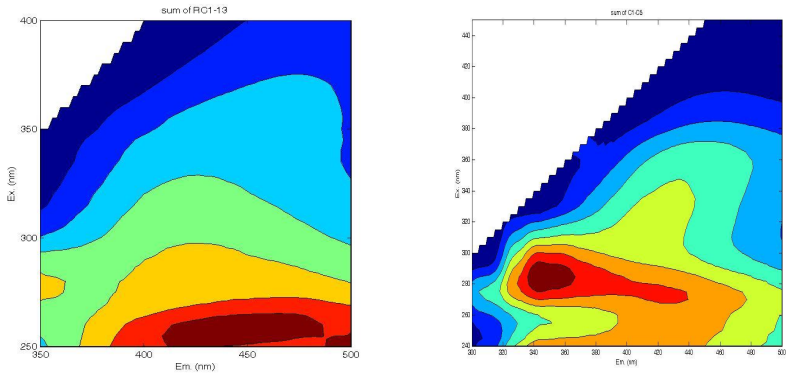


Figure 4-4 Merged contours of the C&M PARAFAC 13 components and the N&A PARAFAC 5 components (right), respectively

Therefore, in brief conclusion on overall fitting results, it was found that the shifting effect of emission spectra by pH change may have significance and should not be excluded in fitting EEMs to the built PARAFAC model. Moreover, this implies that the pH of samples should be consistent if it is proved that PARAFAC can resolve precisely the same fluorophore to different fluorophores at different pHs levels. Otherwise, a new PARAFAC model for each pH should be made, which is a cumbersome task.

The second probable reason is that the 5 component PARAFAC model of the N&A study can be stand-alone as an independent model that is wastewater effluent organic matter (EfOM)-

focused, while still one fourth of EEMs derived from non-wastewater samples (reservoir, drinking water treatment plant, etc.) successfully describe natural organic matter (NOM).

Components comparisons between the two models

Components between the two models were compared by observing the spectra and contour features as well as the variations of loadings in samples. Based on the spectra locations, an attempt has been made to combine several C&M components that may be sub-components resolved from components in the N&A model, and to compare each merged C&M component' EEM to components identified the N&A model. Merged EEMs of the C&M components are shown in Figure 4-5.

Combinations among CM1, SQ1, SQ2 and HQ, either by combining two or three CM components, could generate an imitation EEM that resembles C1, and C2 was similar to CM13 (tyrosine-like), indicating a high possibility that C2 is a single fluorophore. C3 was described by combinations of CM3 and SQ3. Merging of CM13 (tyrosine-like) and CM8 (tryptophan-like) resembled that of C4. C5 showed a similarity to Q3 and HQ combinations. Although this approach of combination is a basic step to compare components in the models, it shows the possibility of resolving composites of fluorophores into sub- or separate fluorophores.

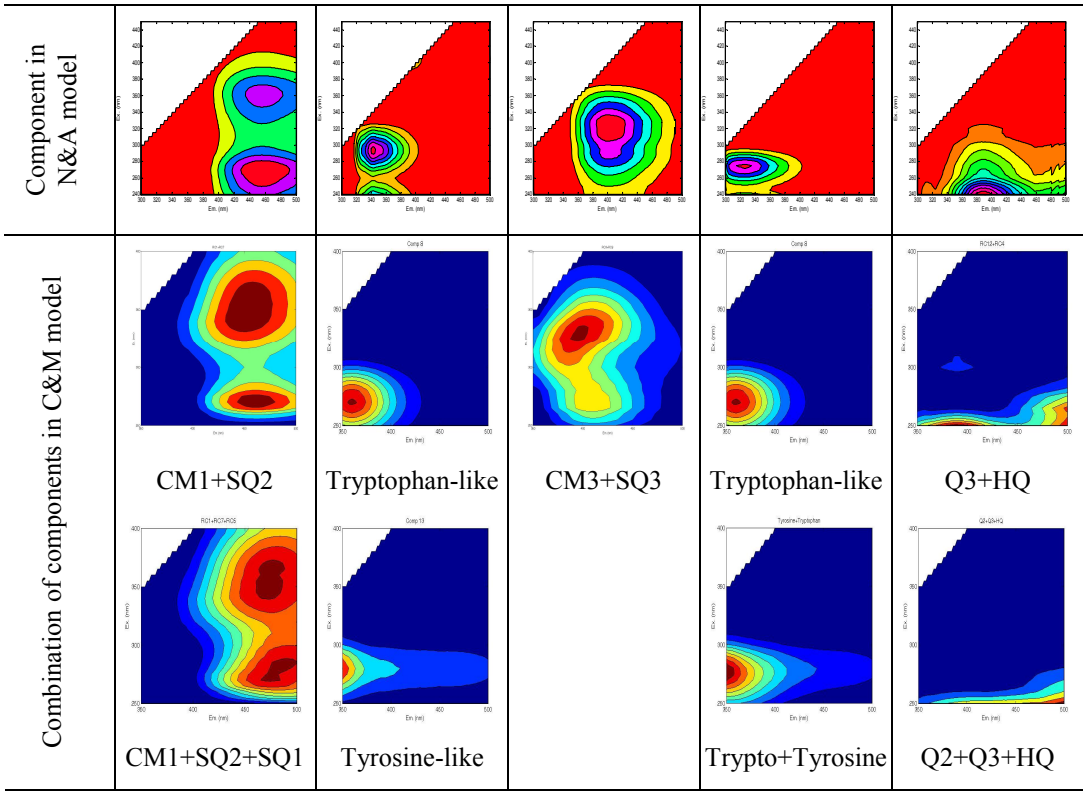


Figure 4-5 Examples of combinations of components in the C&M model; top row from left to right: components C1-C5 in the N&A model (plotted over Ex: 240-450nm, Em: 300-500 nm), middle and bottom row: combined components of the C&M model (plotted over Ex: 250-400 nm, Em: 350-500 nm); it should be noted that ranges of plotted wavelengths are different when comparing

Correlation analysis

Correlations among components in the two models can give some insight into properties of each component. Table 4-2 includes Pearson correlations among components. Investigation of significant positive and negative correlations was made based on correlation coefficients (r) greater than 0.5 in absolute values. C1 showed moderately positive relations to HQ, SQ1 and SQ2, while exhibiting a strong negative relation with a tryptophan-like component. C2 was inversely correlated to HQ, CM6, Q1, and showed a strong linear relation with a tyrosine-like fluorophore. C3 was proportional to SQ2 and SQ3. C4 exhibited a positive relation to a tyrosine-like component, with inverse correlations to SQ1, CM6 and Q1. Lastly, C5 showed somewhat strong linearity to CM6, Q1 and Q3, however, it was moderately inversely related to CM1, Q2, SQ2, CM10 and a tyrosine-like component. Overall, there were strong relations to tyrosine- or tryptophan-like components for C2 and C4, supporting the results (presented in Chapter 3) that these are microbially-derived fluorophores. On the other hand, C1, C3 and C5 appear to be related to quinone-like components identified in the C&M model. Summarizing the spectral position of components in the N&A model, C1 has a longer emission wavelength than C3, with a longer wavelength for C3 than C5. If C1, C3 and C5 change within the quinone-like family by redox-state changes, monitoring anoxic biodegradation may a good approach to prove this premise, as it will be discussed in a later section.

Table 4-2 Pearson correlation matrix among components in two PARAFAC models

	C1	C2	C3	C4	C5	CM1	Q2	CM3	HQ	SQ1	CM2	SQ2	Tryp	SQ3	CM10	Q1	Q3	Tyro
C1	1																	
C2	-0.491	1																
C3	0.197	0.210	1															
C4	**-0.606**	0.221	-0.459	1														
C5	0.184	**-0.653**	-0.194	**-0.584**	1													
CM1	0.321	0.332	0.234	0.112	**-0.626**	1												
Q2	0.050	0.395	0.241	0.216	**-0.584**	0.384	1											
CM3	-0.428	0.500	0.334	0.258	-0.397	-0.196	0.334	1										
HQ	**0.600**	**-0.709**	-0.034	-0.402	0.460	-0.193	-0.448	**-0.519**	1									
SQ1	**0.577**	-0.414	0.124	**-0.505**	0.326	-0.096	-0.382	-0.351	**0.667**	1								
CM6	0.379	**-0.785**	-0.139	**-0.576**	**0.864**	-0.362	**-0.630**	**-0.621**	**0.601**	0.472	1							
SQ2	**0.580**	0.121	**0.506**	-0.135	**-0.513**	0.482	**0.647**	0.151	0.076	0.015	-0.365	1						
Tryp	**-0.814**	0.235	-0.313	0.482	0.040	-0.357	-0.251	0.335	-0.407	-0.275	-0.236	**-0.621**	1					
SQ3	0.135	0.039	**0.603**	-0.003	-0.312	0.020	0.384	0.480	-0.041	-0.003	-0.411	**0.578**	-0.032	1				
CM10	-0.123	0.324	0.358	0.452	**-0.682**	0.404	0.427	0.313	-0.307	-0.227	**-0.662**	0.312	0.101	0.383	1			
Q1	0.435	**-0.643**	-0.166	**-0.523**	**0.688**	-0.361	-0.276	**-0.515**	**0.615**	0.257	**0.743**	-0.092	-0.445	-0.309	**-0.631**	1		
Q3	-0.300	-0.376	-0.303	-0.237	**0.782**	**-0.730**	-0.438	-0.050	0.139	0.017	0.479	**-0.594**	0.320	-0.134	**-0.590**	0.418	1	
Tyro	-0.435	**0.804**	-0.239	**0.533**	**-0.674**	0.374	0.256	0.222	**-0.646**	**-0.507**	**-0.624**	-0.051	0.138	-0.312	0.296	**-0.537**	-0.477	1

a. C1 to C5: components from the N&A model, the reminders: components from the C&M model corresponding to CM1 to CM13 in order.
b. Shaded region: correlation among components.
c. Values in bold: Strong correlations with coefficients (r) greater than an absolute value, 0.500.

Principal component analysis (PCA)

Principal components analysis (PCA) was used to portray the grouping of components and relations between components of the C&M model and components in the N&A model. Four principal components (PCs) could be extracted with criteria of eigenvalues greater than 1 [71]. Since the scales of all components were in same range (0-1), neither rotation nor transformation of data was applied. There were no missing values. Table 4-3 summarizes the loading of each PC. The fourth PC was discounted due to relatively low weight and explained variation compared to other three PCs. Figure 4-6 shows a loading plot of PC1 and PC2. The first two PCs account for 62.7% of the total variations in the EEMs dataset. PC1 explained 39.2% of the total variability. C2 and C4 were clearly separated from other components by being related to tyrosine-like and tryptophan-like fluorophores and CM3, and the opposite positioning of C5 means that C5 has a significantly different property compared to C2 and C4. Albeit a unknown component, it was discussed that the spectral characteristics of CM3 as well as SQ2, SQ3, HQ, CM6, tryprophan-like fluorophores (CM8), and a combination of Q2 and Q3 were well matched with those of microbial components in the original C&M study [24]. SQ2, SQ3 and Q2 were somewhat strongly related to PC1, although HQ and CM6 exhibited strong negative loading on PC1. Thus, this indicates that PC1 is a strong microbial-related factor having high biodegradability. Also, the fluorescence index was linear to PC1 (result shown in Chapter 6).

On the other hand, PC2 explains 23.5% of the total variability and includes C1, C3, CM1, SQ1, SQ2, SQ3, a tryptophan-like fluorophore and Q3. C1 and C3 showed high PC2 loadings along with SQ2 and CM1, opposite to a tryptophan-like fluorophore and Q3. In an extended statistical analysis performed with other water quality parameters (detailed graphs not shown in this chapter, but in Chapter 6), C1 and C3 showed strong inverse correlations to bromide

concentration in wastewater and positive relation to the fraction of humic substances from SEC-DOC chromatographic analysis. Extended PCA (shown in Chapter 6) showed that they are in same direction to SUVA and molar absorptivity at 280 nm, two parameters that are commonly used for indicating the degree of aromaticity of organic matter [72]. It was reported SQ2 has a linear correlation to % aromatic carbon [24]. Hence, PC2 is likely to explain the degree of redox state of quinone-like fluorophores having aromatic properties. PC3 (11.7% of the variance) is positively contributed by C3, CM3 and SQ3, and negatively by CM1, a tyrosine-like fluorophore.

PCA results show, overall, that fluorophores can be further separated mainly by source orientation, information corresponding to PC1, and by the extent of chemical transformation/reaction experienced, which is distinguished by PC2 to some degree, after all influencing fluorescing patterns and/or physicochemical characters of fluorophores.

Table 4-3 Component loading matrix of principal component analysis for 18 components identified from PARAFAC modeling (5 components from the N&A model, 13 components from the C&M model)

Variables Names		Principal Component 1	2	3	4
C1	C1	-0.423	**0.880**	-0.123	-0.021
C2	C2	**0.809**	-0.161	-0.142	-0.191
C3	C3	0.139	**0.677**	**0.504**	-0.001
C4	C4	**0.633**	-0.457	-0.172	0.280
C5	C5	**-0.904**	-0.276	0.182	-0.126
CM1	CM1	0.462	**0.526**	**-0.514**	0.097
Q2	CM2	**0.623**	0.392	0.063	-0.422
CM3	CM3	**0.588**	-0.132	**0.636**	-0.152
HQ	CM4	**-0.738**	0.347	-0.066	0.343
SQ1	CM5	**-0.562**	0.362	-0.002	0.478
CM6	CM6	**-0.942**	-0.034	-0.122	-0.005
SQ2	CM7	0.328	**0.845**	0.075	-0.182
Tryptophan-like	CM8	0.235	**-0.785**	0.308	0.318
SQ3	CM9	0.320	0.442	**0.730**	0.107
CM10	CM10	**0.695**	0.247	0.109	0.430
Q1	CM11	**-0.814**	0.104	-0.135	-0.363
Q3	CM12	**-0.583**	**-0.571**	0.388	-0.237
Tyrosine-like	CM13	**0.727**	-0.256	**-0.502**	-0.142
% Variance explained		39.2	23.5	11.7	6.8
%Cumulative variance		39.2	62.7	74.4	81.2
Eigenvalue		7.047	4.232	2.110	1.226

Extraction Method: Principal Component Analysis.

a. 4 components extracted.

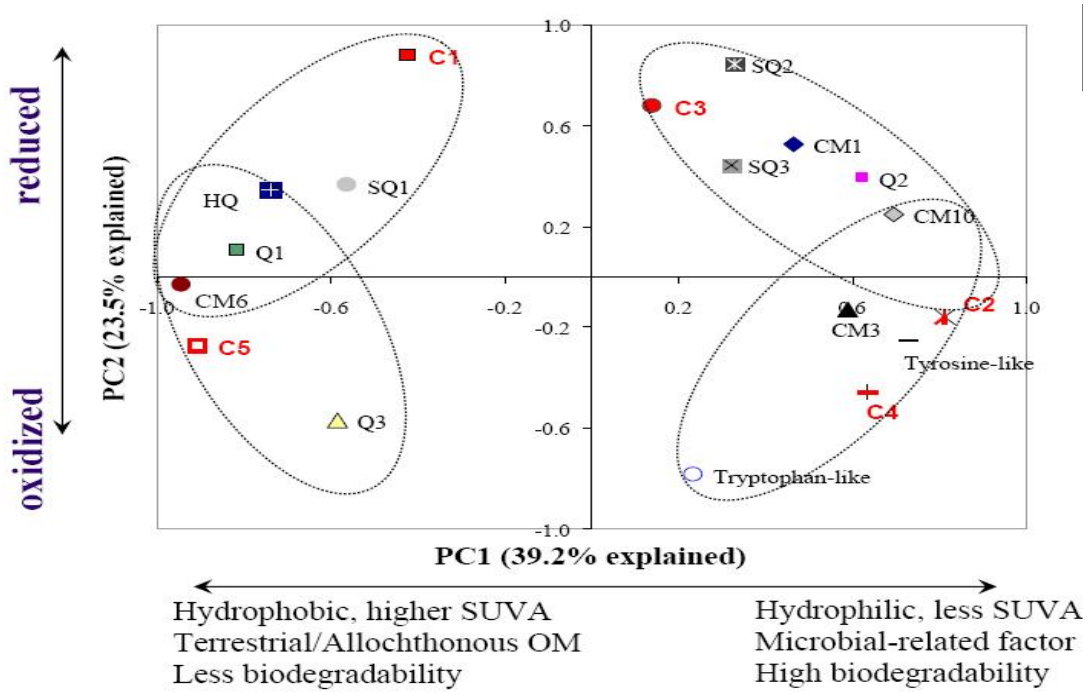


Figure 4-6 PCA loading plot of the two first PCs of 18 PARAFAC components; 13 components of the C&M model and 5 components of the N&A model; C1-C5 represents components from the N&A model, Q1-Q3 for oxidized quinone-like, SQ1-SQ3 for semiquinone-like, and HQ for hydroquinone (reduced quinone-like), Tyro and Tryp for tyrosine-like and tryptophan-like, respectively

Biodegradation

As previously mentioned, C1, C3 and C5 are likely to have relationships with quinone-like components although it is not clear to identify which ones are specifically related to which quinone-like fluorophores of the C&M model. In order to examine the possibility of following the quinone-like fate in an aquatic environment, were investigated fractional changes of C1, C3, and C5 components were investigated during a 5-day kinetic anoxic biodegradation experiment and compared to an aerobic test (Figures 4-7 and 4-8) for a wastewater (EfOM)-*dominant* river sample (this sample was the same water used for results shown in Figures 4-9 to 4-11. In Figure 4-7, (a) displays fractional changes of each component for anoxic biodegradation and (b) represents the regression slopes of each component shown in anoxic (a) and in aerobic (Figure 3-10-(a) in Chapter 3) tests. For this selected water, fractions of aerobically and anoxically biodegradable organic carbon over 5-days were 47% (n=3) and 21% (maximum of n=4), respectively and the corresponding amounts of BDOC were 7.1 and 3.3 mg/L. As seen in Figure 4-7-(a), fractions of C2 and C4 decreased during the course of biodegradation time, showing faster degradation of C4 rather than C2. Of the remaining components, fractions of C1 and C3 increased gradually while the fraction of C5 was reduced. Although an increase of a component' fraction does not always mean "no" biodegradation or "formation" of such components, an increase of C1 and C3 components shown in figure 4-7-(b), and this may indicate the fate of quinone-like fluorophores under reducing environments like anoxic conditions, the possible pathway of their fates may be C5 → C3 → C1 while degradation of each of these components along with labile components like C2 and C4 are still happening: This route is possible since C5 seems most close to oxidized quinone fluorophores (i.e., Q1-Q3 components), while C1 is one of components similar to semiquinone or HQ (by correlation) and has the longest emission

wavelength of the three components. Also, a small shoulder peak at 440 nm of excitation, and at near 490-500 nm of emission may be considered as a peak of reduced humics, and this peak that was not completely separated from C5 might have appeared as another component with C5.

On the other hand, compared to anoxic biodegradation, aerobic biodegradation showed different patterns in fractional changes, especially in C5 (Figure 4-7-(c)). Under aerobic conditions, the fraction of C5 was almost unchanged, and the fractional increases of C1 and C3 were slow compared to those under anoxic conditions. Obviously, this means that anoxic biodegradation of organic matter follows different mechanisms, compared to aerobic biodegradation. Regarding the fate of organic matter, a more comprehensive understanding of these components may be required in treatment processes for wastewater reuse such as soil aquifer treatment and river bank filtration.

Changes of the fluorescence intensity have also been observed by subtracting the EEM of day 5 from the EEM of day 0. After 5-days of biodegradation, the protein-like peak showed a significant decrease for both aerobic and anoxic tests. However, the peak corresponding to excitation wavelength 340-480 nm, and emission wavelength 420-470 nm increased, resulting in negative values for the differential EEM (Figure 5-8). This peak overlaps CM1 and the second maxima peak of C1. Besides the case of this sample, similar results have been experienced for several different samples that have gone through anoxic biodegradation conditions.

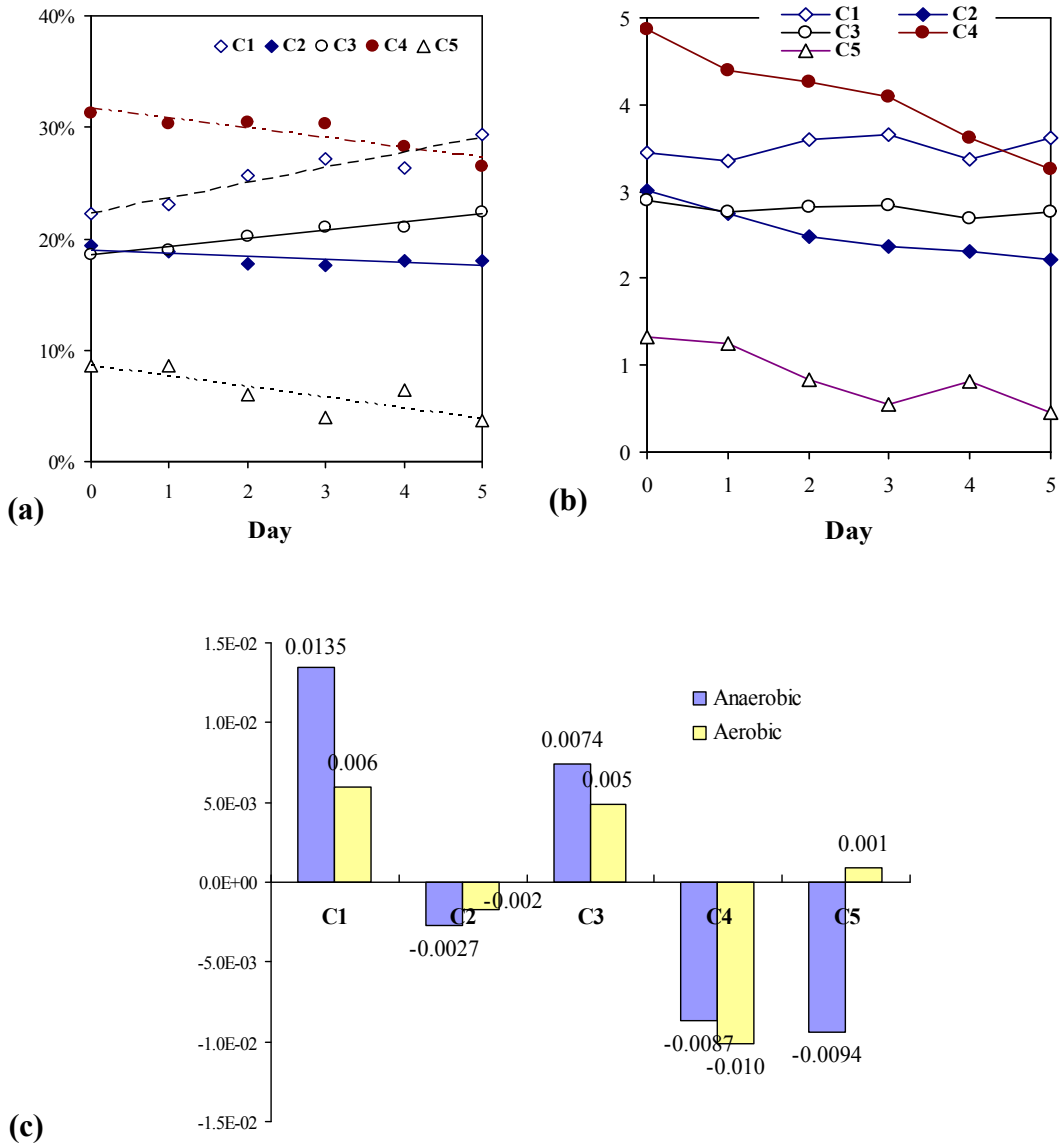


Figure 4-7 Fractional changes (a) and DOC (b) of each component during 5-day anoxic BDOC kinetic test and comparisons of regression slopes (c) of fractional changes for an EfOM-*dominant* stream in the Southeast of USA (taken in June, 2005)

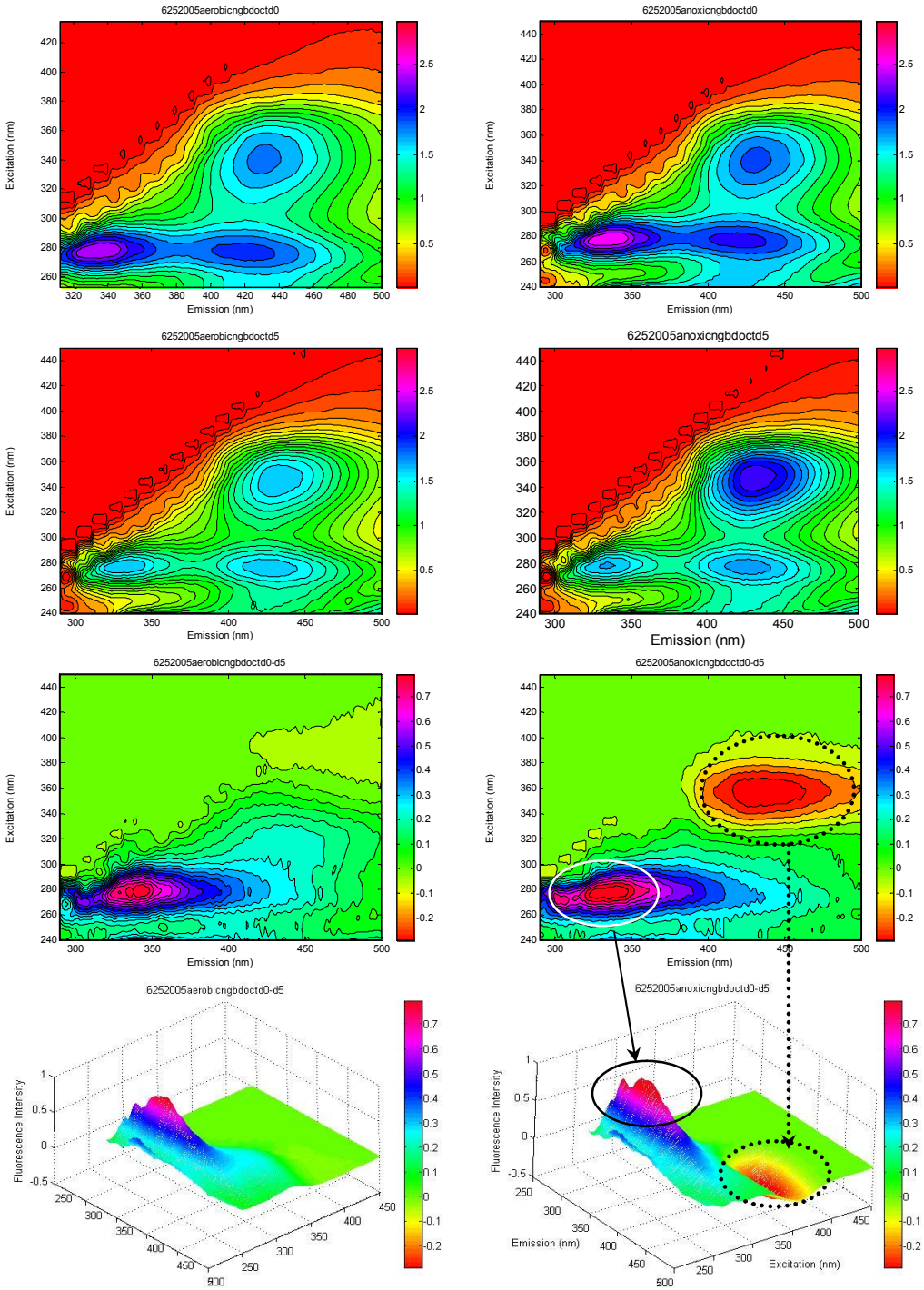


Figure 4-8 Variations of EEM contour and landscape plots by 5-day aerobic (left) and anoxic (right) biodegradation; top row-day 0, second row-day 5, third row-differential EEMs by subtracting day 5 from day 0, bottom row-3-D plotting of the differential EEMs. EEMs were drawn according to same z-axis elevation; fixed to 0~3 ranges for day 0 and 5, -0.3~0.8 ranges for differential EEMs. Source water was from the Santa Cruz River (AZ) whose watershed was a wastewater-impacted stream. For anoxic biodegradation, oxygen was purged with pure nitrogen (>99.99%), confirmed by close to zero concentration

Components in the South Platte River watershed

Observation of components variation in the South Platte River watershed is meaningful in understanding how components vary according to source water changes (e.g., wastewater input). Detailed descriptions of the sampling sites were contained in Chapter 3.

Since a loading value of a component is linear to concentration or amount of the component under the tri-linearity assumption, it is possible to judge that a higher loading for component one than component two means greater abundance of component one compared to component two in respective samples. Components in the same sample are under the same extent of influences by residuals. Of course, the extents of residuals may be varied for samples, resulting in making a direct comparison of components among different samples impossible, in terms of each fraction of components. That conclusion is only valid when there are negligible errors and the sum of loadings of all components, or normalized fraction of components' loading divided by sum of loadings, is responsible for explaining the complete variation or explained variation for the component. However, it is reasonable for the same component to compare a trend of its abundance based on the absolute comparison of loading scale itself among sampling sites along a river.

As Figure 4-9 depicts, of the 13 components, upstream sites (Site 1, SP/S1CR) or minimally wastewater-impacted sampling sites (Site 3, SP/S3UA, and Site 5W-1 and 5W-2, SP/S5AU), Q1, Q3, SQ1, HQ and CM6 showed increased loading, whereas CM1, CM3, CM10, Q2, SQ2, SQ3, and tryptophan- and tyrosine-like components increased in wastewater effluent or sampling sites that heavily wastewater impacted (SP/S6FG, SP/S7BC and SP/S11D). Of those, variation of CM6 was definitely in contrast between wastewater effluent and upstream or wastewater-impacted water. It was present to a large extent in non-wastewater effluent, indicating this

component can be used as a possible indicator component for tracking wastewater input.

For wastewater samples, Q3, Q2 and tyrosine-like components were present in large abundance while SQ1 was found mostly in the least abundance as a fraction. Since wastewater is a mixture of NOM, SMPs and traces of synthetic organics [73] and usually wastewater has high level of organic carbon, a fractional increase of certain components does not necessarily mean that those components can be used as an indicator of wastewater input. As organic loadings increase, constitutional fluorophores of NOM also increases along with those of EfOM. Fractional trends should be considered with this understanding.

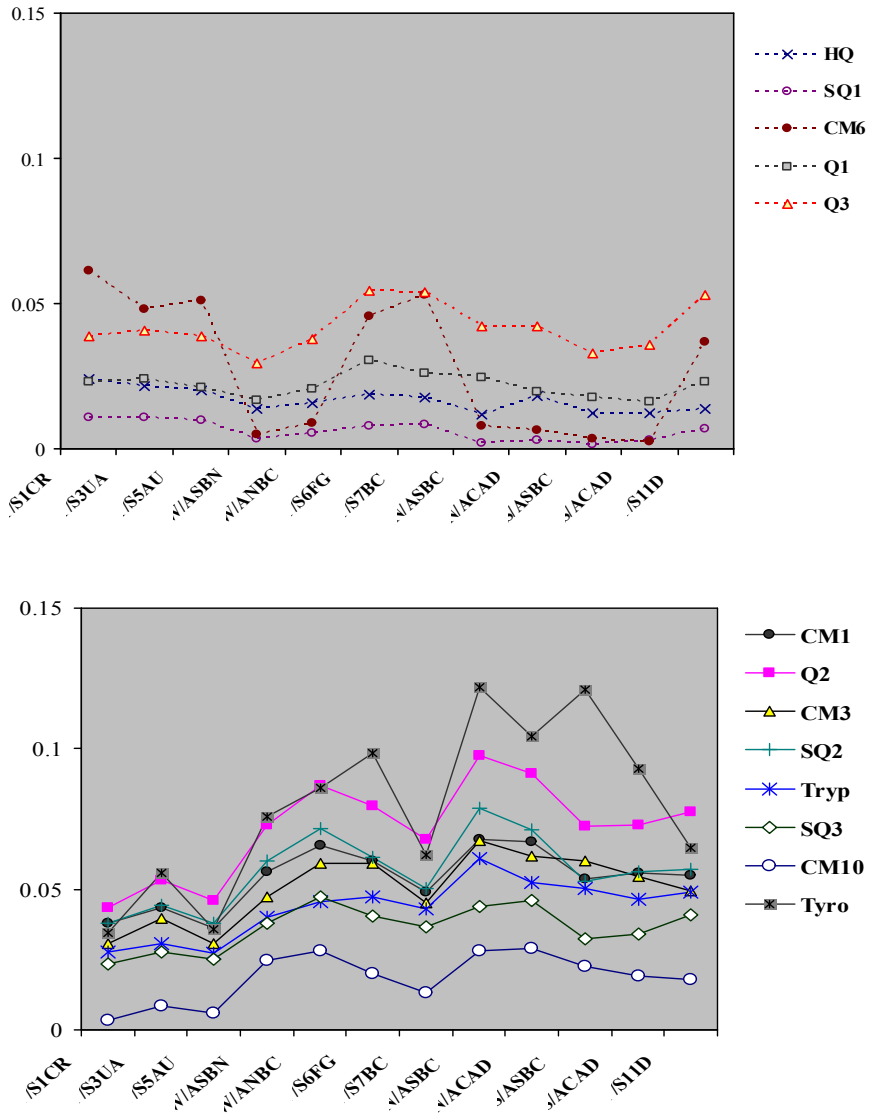


Figure 4-9 Loading values of 13 components in the SPR watershed; the upper and the lower figures show components decreasing and increasing in wastewater or heavily-impacted sampling sites, respectively. Loadings shown are means of three samplings (Feb, Sep. in 2004, April in 2005). The River flow corresponds to left to right in figures

Assessment of properties of components

Assessment of components' properties was performed by EEMs of EfOM and NOM or corresponding isolates. References and ~30 isolates extracted from WWTPs, rivers and lakes, and ~20 hydrophilic NOM isolates' samples have been compared with the C&M model. The hydrophilic fraction has been defined as the fraction that passes though XAD-8 resin [4] but, in

this work, it is defined as the fraction after sequentially passing through XAD-8/-4 resins. Results shown here are representative since it is too extensive to cover all results.

Figure 4-10 shows fractional components present in several different concentrations of the EfOM (used in this study as a reference EfOM which had gone through electrodialysis and RO by CSM) resolved by the C&M model, and reveals that the tyrosine-like component is a dominant constituent of EfOM, and CM1, Q2, CM3 and SQ2 also constitute a large part of EfOM (no. 1-3 and 7 in figure). On the other hand, HQ, SQ1 and CM6 (no. 4-6 in figure) were minor members. Unexpectedly, the fraction of tryptophan-like (no.8) was present within 7~9%. Over a range at seven different concentrations, general trends were reliable and results based on 1 mg/L showed little variations compared to others.

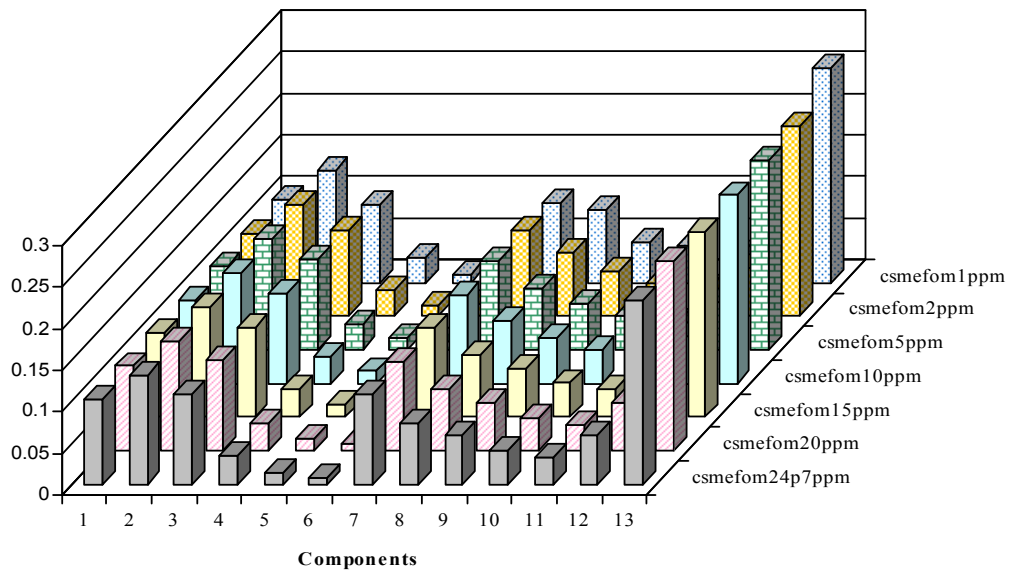


Figure 4-10. Fraction of components of the C&M model in reference EfOM at seven different concentrations; initial DOC was 1, 2, 5, 10, 15, 20 and 24.7 mg/L, and diluted samples were measured and later dilution factors were applied

Figure 4-11 shows pie-graphed fractional components in diverse types of organic matters. As shown, CM1, Q2, CM6, HQ, SQ1 and SQ2 were common in reference NOMs (SRHA, SRFA and SRNOM) and these trends were quite uniform among samples. Of these, CM6 was the most

abundant component followed by CM1. SQ3, tryptophan-like, and CM10 were present at a negligible level. In comparison to EfOM and SMPs, the tyrosine-like component was present less, and the fraction of SQ1 and CM6 were greatly increased. Also, fractions of SQ3 and CM10 were opposite to that of EfOM and SMP trends.

In the wastewater hydrophilic fraction, CM3 was dominantly present while it was minimal in wastewater HPOA. In terms of HPOA between wastewater and lake water, fractions of each component were very different. However, the distribution of each fraction in a lake HPOA sample were very similar to reference NOM, meaning that constituents of wastewater HPOA and lake/river HPOA are different although their natures based on XAD-8 isolation are categorized as HPOA. Commonly, Q2 was a ubiquitous fluorophore in all samples, showing an increased fraction in wastewater HPOA. The tyrosine-like fluorophore was not observed in wastewater HPOA. In general, it is thought that HQ, CM6, CM1, Q1-Q3 and SQ1 have hydrophobic natures, while CM3, SQ2, tryptophan-like, SQ3 and tyrosine-like components are closely related to hydrophilic properties. CM0 is also thought to be hydrophilic. Components showing hydrophilic natures were likely to have shorter emission wavelengths compared to components showing transphilic and/or hydrophobic properties (also shown in Chapter 5).

Linking to components in the N&A model, obviously tryptophan- and tyrosine-like components were similar to C2 and C4 which are major fluorophores reflecting EfOM and its hydrophilic nature. C1 which was dominant in reference NOM appears to have a relationship with CM1 and SQ1. C3 may be linked to CM3 and CM10 since this component was abundant in wastewater hydrophilic fraction, compared to a low fraction in reference NOMs. However, there is a discrepancy with PCA results and the linkage of C3 to wastewater/hydrophilicity may be due to an increased amount of organic carbon, simultaneously resulting in an increased fraction of C3.

Lastly, C5 is similar to Q2 and Q3 in aspects of hydrophobic character and similar ranges of wavelengths.

Components in the each group (hydrophobic/hydrophilic) were mostly consistent in fractional trends in the SPR watershed, i.e., wastewater or wastewater impacted samples showed increased fractions of components assigned to hydrophilic group, and vice versa.

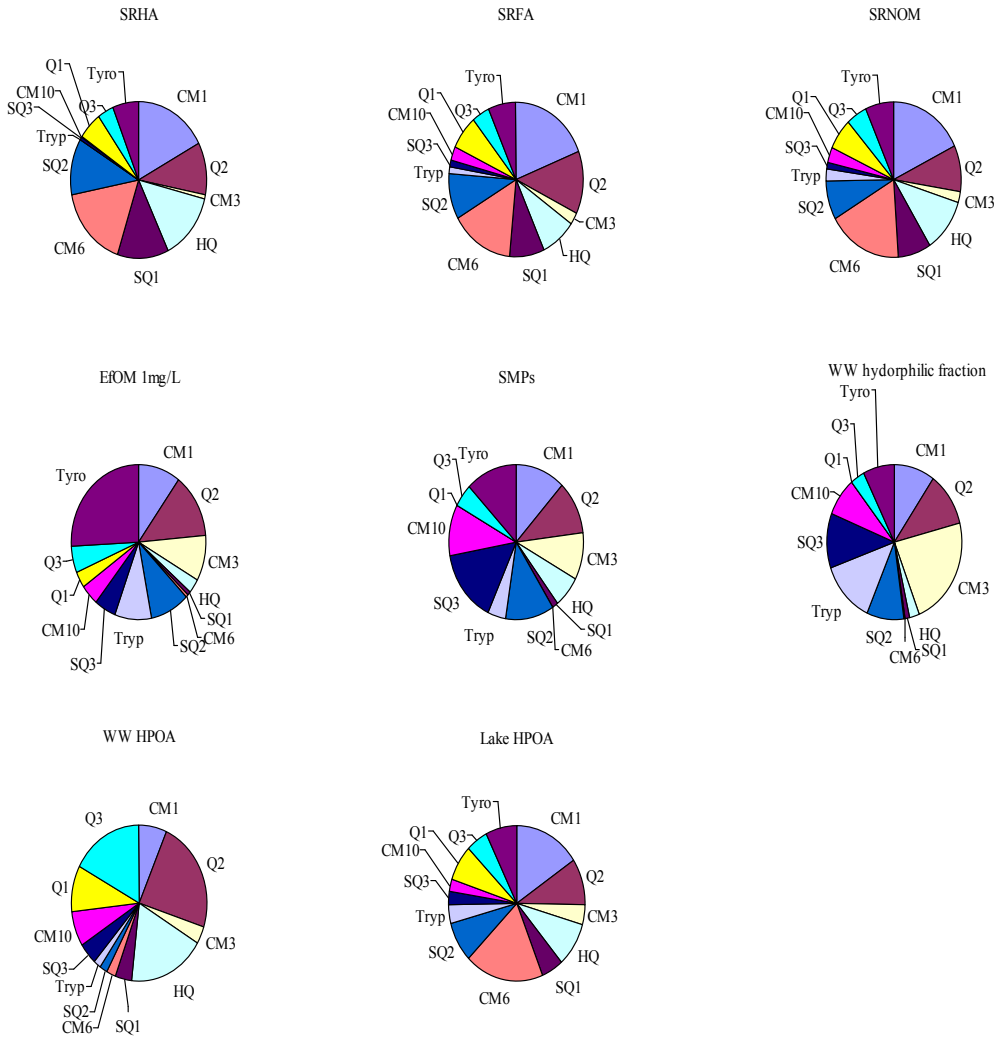


Figure 4-11 Fraction of components in reference NOM and EfOM; upper row-SRHA, SRFA, SRNOM, middle row-EfOM (1mg/L DOC), SMPs, WW hydrophilic fraction bottom row-wastewater HPOA, and lake HPOA

Conclusions and implications

This study was initiated with the hypothesis that some of the fluorophores identified through the 5 components PARAFAC model may be composites and might be resolved into more pure or sub-fluorophores. For this, EEMs which were used in building the 5 components PAFARAC model were fitted to the C&M 13 components PARAFAC model. Spectra and contour comparisons of individual components showed indication of mixtures of several fluorophores. Overall fitting resulted in relatively high residuals with similar patterns, meaning that the C&M 13 component PARAFAC model was not able to explain all the information in EEMs used for the 5 components model, and there may still be unresolved fluorophores which are more related based on wastewater-originated spectral properties. In spite of unexplained information (assuming at least more than 10%), results by fitting to 13 components were valuable because trends were consistent based on assumption of similar level of errors. However, it was thought that the N&A PARAFAC model can be more appropriately applied to wastewater effluent organic matter including SMPs whereas the C&M model would be more adaptable to organic matter present in surface water.

In anoxic biodegradation tests, it was found that C1, C3 and C5 are possibly related to quinone-like components. As a monitoring tool of wastewater input, C2 and C4 in the N&A model are successful, and these components are obviously comparable to the tyrosine-like component of the C&M model. CM6 can also be used as a supporting component in terms of minimal or no wastewater impact to a stream. Based on reference NOM, EfOM and isolates results, C1 and C5 are hydrophobic fluorophores, while C2, C3 and C4 reflect hydrophilic natures.

Overall, through fitting EEMs to the built PARAFAC model, it was supported that the 5

components identified in the N&A model are appropriate and well-explainable in terms of dissolved organic matter regardless of origin. However, as mentioned in an earlier part of this study, it should be understood that there are still some portions of errors in the fitting results.

For further research, relations of each component with respect to molecular weight may be necessary to provide a better understanding in environmental processes.

CHAPTER 5[1]

CHARACTERIZATION OF WASTEWATER EFFLUENT ORGANIC MATTER (EfOM) AND DIFFERENTIATION FROM DRINKING WATER NATURAL ORGANIC MATTER (NOM) USING SIZE EXCLUSION CHROMATOGRAPHY AND FLUORESCENCE EXCITATION EMISSION MATRIX

Introduction

Dissolved natural organic matter (DNOM) is a ubiquitous organic material in all types of waters, and the properties of NOM vary according to source/origin and various environmental factors. Proper characterizations of NOM in physical and chemical aspects has been a central activity since the understanding of NOM plays important roles in the fate, transport and the reactivity of NOM in all physicochemical and biological water treatment processes [74-80]. Due to the complexity and heterogeneity of NOM in both properties and structures, investigations on its characterization have increased with multiple analytical methods [12, 72, 81, 82].

On the other hand, limitation of available water sources has led to wastewater reuse as an alternative strategy. Since wastewater reclamation/reuse has become more common, it should be noted that wastewater effluent organic matter (EfOM) issues are linked to NOM in water treatment processes and potential health concerns. Through previous studies, it has been shown

[1] Partial data discussed in this chapter were published in:

Seong-Nam Nam, Stuart W. Krasner, Gary L. Amy, (2007). Differentiating Effluent Organic Matter (EfOM) from Natural Organic Matter (NOM): Impact of EfOM on Drinking Water Sources, *Advanced Environmental Monitoring*, Chapter 20, p259-270, Springer-Verlag GmbH, Germany

Seong-Nam Nam and Gary L. Amy, Differentiation of wastewater effluent organic matter (EfOM) from natural organic matter (NOM) using multiple analytical techniques, *Water Science and Technology*, 57(7), p109-115, 2008

that EfOM is not only more complex, but also more diverse in composition than NOM, due to combinations of organic matters derived from various sources. The composition of bulk EfOM is mainly a combination of bulk natural organic matter (NOM) and soluble microbial products (SMPs) [59, 83, 84]. A significant portion of the EfOM originates from the corresponding drinking water, and is mostly refractory humic substances which are not easily removed by wastewater treatment process, while SMPs are derived from biological wastewater treatment.

Previous studies on NOM and EfOM have shown that organic matter from the wastewater-derived samples exhibits lower C and O contents and a significantly higher amount of H, N, and S [85]. Debroux (1998) found that EfOM has higher solubility, lower SUVA, and higher elemental concentrations of nitrogen and sulfur [86]. From studies using analytical techniques including XAD-8/-4 resin fractionation, Drewes and Croué (2002) suggest that the moeties of EfOM are probably of different origin than those of NOM, for example, they found that NOM usually represents well degraded organic matter, while EfOM contains more easily degradable, non-aromatic carbon [87]. There have been increasing numbers of researchers who investigated both NOM and EfOM as references standards or isolates using size-exclusion chromatography (SEC) with multiple detectors and fluorescence excitation-emission matrix (EEM). Therefore, the objective of this research was to investigate the characteristics of EfOM differentiating from NOM in terms of its molecular weight distribution (size exclusion chromatography coupled with DOC/UVA detection) and fluorescence spectra (excitation emission matrix and fluorescence index). The results and discussion presented in this chapter will provide further understanding of NOM and EfOM in terms of their similarities and differences, as well as a synthesis with previously reported studies.

Experimental Methods

Reference NOMs/EfOMs sources: Suwannee River humic acid (SRHA), fulvic acid (SRFA) and natural organic matter (SRNOM) purchased from the IHSS were used as reference standards of NOM. Laboratory-produced soluble microbial products (SMP), and electrodialysis/reverse osmosis (ED/RO) isolated effluent organic matter (EfOM) were used as representative references of EfOMs. SMPs were generated in the Arizona State University Laboratory by the following procedures. To generate SMP, synthetic wastewater was prepared with glucose and inorganic salts in a bench-scale sequencing batch reactor (SBR) to simulate an activated sludge process. The synthetic wastewater contained a simple organic (C) substrate and inorganics with a recipe of glucose-10 mgC/L, NH_4Cl-38.2 mg/L, $CaCl_2$-10.1 mg/L, $MgSO_4$-24 mg/L, KH_2PO_4-20.3 mg/L and Na_2HPO_4-42.8 mg/L in deionized water. To start up the system, activated sludge was obtained from a municipal wastewater treatment plant and used to inoculate the SBR. Glucose was used as an organic carbon/energy source for the bacteria. An initial concentration of dissolved organic carbon (DOC) in the reactor was approximately 100 mgC/L, providing a carbon to nitrogen ratio (C:N) of 8:1. Phosphate buffer with a concentration of 10^{-3} M was used to maintain a pH of 7.5± 0.5 and also to provide phosphate for bacterial growth.

The EfOM sample was generated at the Colorado School of Mines and provided to University of Colorado. Initially, WW was collected from the secondary effluent of the Boulder wastewater treatment plant, was filtered through a microfilter (0.04 μm) and concentrated by an ultra low pressure RO (FILMTEC XLE-2540 membrane) followed by electrodialysis to remove mono/divalent cations and anions.

***Analyses and instrumentations*:** UV absorbance (UVA) and dissolved organic carbon

(DOC) were measured using a UV-Vis spectrophotometer (Shimadzu LC, Japan) and an Ionics Sievers 800 TOC analyzer, respectively. The HP-SEC measurements were performed with a HPLC (Shimadzu LC600) coupled with a UV-Vis detector (Shimadzu SPD-10Avp) and an on-line DOC detector (modified Ionics Sievers Turbo 800 TOC analyzer). A detailed description of the HPSEC instrumentation can be found in Her *et al.*[2]. The system used a TSK HW-50S column (ID 2 cm × Length 25 cm, 35 μm Toyopearl HW resin) and the flow rate was 1 mL/min. The mobile phase was prepared with Milli-Q water buffered with phosphate (0.0024 M NaH_2PO_4 + 0.0016 M Na_2HPO_4, pH 6.8) and 0.025 M Na_2SO_4, producing 0.1 M of an ionic strength. All samples were filtered through a 0.45 μm polycarbonate membrane filter, and conductivity of the sample was adjusted to the same level as that of the mobile phase before injection. The sample injection volume was 2 mL. UV ratio index (URI) value, which was introduced by Her *et al.* [88, 89], was calculated by absorbance ratio at 210 and 254 nm.

***Molecular weight calculation*:** Chin *et al.*[72] introduced polystyrene sulfonate (PSS) as representative calibrants for humic substances in the SEC analysis with UV detection only. However, due to both the electrostatic and hydrophobic interactions between the PSS and the gel in Toyopearl column [90], in this study polyethylene glycol (PEG) ranging from 200 to 10,000 g/mol (daltons) was used as standards to determine the apparent molecular weight. A semi-log linear calibration curve (Figure 5-1) was used to calculate the MWs of samples. The M_n, M_w and ρ were determined using the following equations (1)-(3).

$$M_n = \frac{\sum_{i=1}^{n} h_i}{\sum_{i=1}^{n} \frac{h_i}{M_i}} \tag{5-1}$$

$$M_w = \frac{\sum_{i=1}^{n} h_i M_i}{\sum_{i=1}^{n} h_i} \tag{5-2}$$

$$\rho = \frac{M_w}{M_n} \tag{5-3}$$

Where M_n is the number average molecular weight, M_w is weight average molecular weight, ρ is polydispersity, h_i is the response of the sample SEC curve at retention time (Rt) or at volume i, and M_i is the molecular weight at eluted volume i, respectively. A pure substance will have $M_w = M_n$ (i.e., $\rho = 1$). However, for a mixture of molecules, M_w will be greater than M_n (i.e., $\rho > 1$).

In the case of a somewhat unsmooth chromatogram, the chromatogram was smoothed by the Savitzky-Golay method [91] for M_w and M_n calculation using the GRAM/AI® program (Thermo Inc., USA).

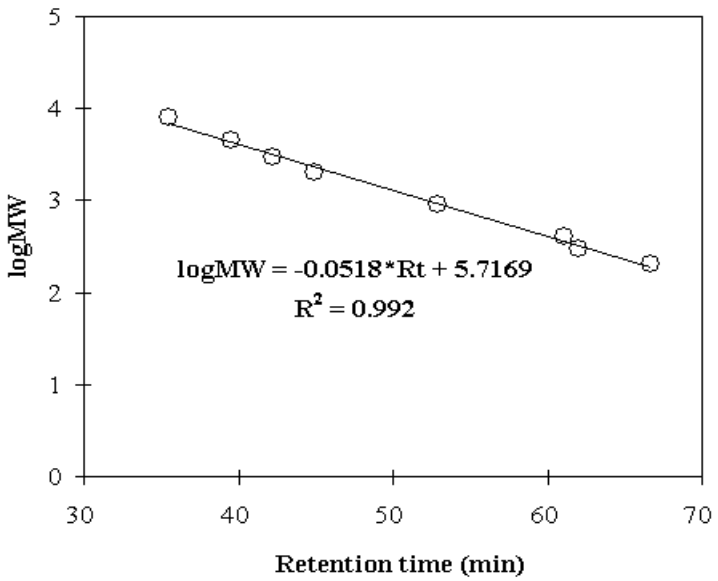


Figure 5-1 Calibration curve obtained by a semi-log linear regression of standards using PEGs 0.2k-10k Daltons: The calibration curve was made using the peak-position method

***Fluorescence excitation-emission matrix measurement*:** Fluorescence excitation-emission matrix (EEM) was performed with a FluoroMax-2 or -3 spectrofluorometer (Horiba Jobin Yvon Inc., USA). A 150-W ozone-free xenon arc-lamp was used as a light source for excitation. The slit widths were set to 5 nm for both excitation and emission. All EEMs were obtained by measuring the emission spectra over the range of 290-530 nm at 2 nm intervals, with an excitation range of 240-450 nm at 10 nm intervals. The EEMs of each sample was adjusted by subtracting an EEM of 0.01M KCl (pH 2.8 adjusted with HCl) solution (set as a blank EEM) to remove Raman scatter peaks. Correction steps were carried through each blank-subtracted EEM using emission and excitation correction factors provided by the manufacturer, and then the EEM was went through the interpolation of excitation and emission by 1 nm interval at both excitation and emission. Intensities were normalized to the area under the water Raman peak of excitation at 350 nm and the Raman peak area was also used for checking the stability of the instrument by measuring it on the measurement date. All data processing including the aforementioned correction procedures were performed with MATLAB® version 7.0 using a custom made code (Appendix). In order to check reproducibility of the normalized fluorescence intensity, some samples were measured in triplicate, showing a coefficient of variation of 2±1%.

Results and discussion

Methodology establishment

***Conversion of UV and DOC electrical signal to UV absorbance*:** The HPSEC detecting system in this work consists of a variable wavelength UV detector, a variable wavelength fluorescence detector and a DOC detector in series. Prior to electrical signal interpretation from the HPSEC-DOC/UV detectors including the calculation of SEC-SUVA, it was necessary to

investigate the relationship between the electrical response as milli-volts (mV) and the corresponding amount of UV absorbance (UVA) in absorbance unit (A.U.) or DOC in mgC/L. First, the UV detector is able to detect absorbance at two different wavelengths at the same time. However, between the two signal outputs and data acquisition channels, there exists different relation (i.e., sensitivity) which was intentionally designed not only for cross-use depending on concentrations (i.e., less sensitive channel for high absorbance and vice versa), but also to provide detection at two wavelengths simultaneously. The relationship of each electrical signal to absorbance is presented in Figure 5-2 and equation 4 and 5. Since the optical absorbance is sensitive to the amount of light-absorbing moieties with σ, π, or n electrons per molecule or per unit volume of organic matter and especially, UV absorbance at 254 nm is sensitive to unsaturated bonds. UVA_{254} serves as a measure of the aromatic carbon content of NOM by normalizing to DOC concentration which is defined as specific UV absorbance (SUVA), whatsoever chromophores with unsaturated bonds can be expected to respond within the detecting ranges. It was assumed that A.U. is the same as cm^{-1} for a path length of the flow cell in the detector of 1 cm. Yet, it should be taken into consideration that this relationship may change as the UV lamp approaches its life span because the *intrinsic* A.U. of chromophores at a certain Rt (retention time) will be diminished to the *experimental* A.U. Therefore new calibration is required intermittently; however, for the time period of this study, the relation was virtually consistent.

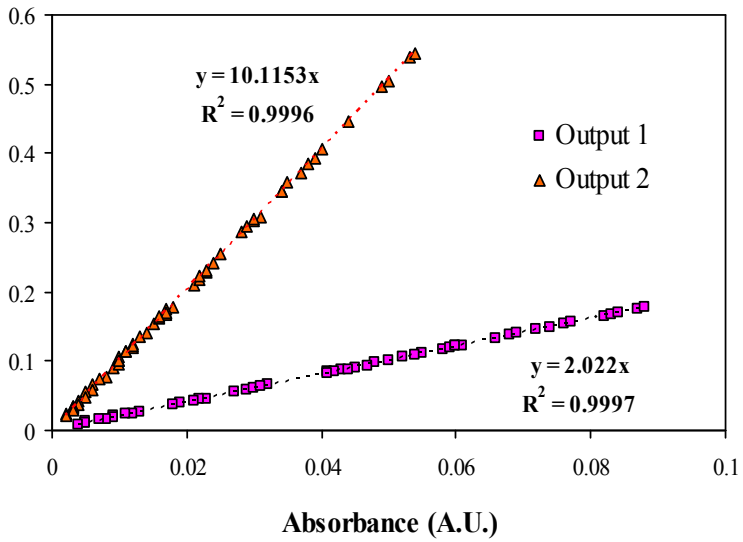


Figure 5-2 The relation curve of $UV_{mV\_output1}$ and $UV_{mV\_output2}$ in electrical response (mV) versus $UV_{abs\_output1}$ and $UV_{abs\_output2}$ (A.U.); Wavelengths selection to output channels may vary depending on researcher's choice. In this study, in case of SEC-URI, UVA_{210} and UVA_{254} were measured using output 1 and output 2, respectively

Second, in order to be able to investigate the SEC-SUVA (equation 4) at the eluted volume *i*, it is required to know the amount of DOC at the eluted volume *i*, which is possible by simply determining the relation of electrical responses based on the amount of DOC, and by converting the relation directly to the signal at the eluted volume *i*. As shown in Figure 5-3, it showed a good linear relationship ($mV_{DOC} = 0.0596 \times mg/L_{DOC}$, $R^2 = 0.9932$) using potassium hydrogen phthalate (KHP) which is a recommended TOC analyzer standard reagent [15]. The relationship is applicable to any (natural or synthetic) organic compounds that contain organic carbons regardless of physicochemical properties because detection does not respond to chemical properties but to the amount of organic carbon. Note that if the concentration of samples is too high to be completely mineralized under the given oxidation conditions of the TOC detector, the signal will yield an underestimation of actual DOC. Therefore, it is crucial to adjust DOC level of samples to appropriate ranges which are dependent on the system capability. Additionally, the oxidation efficiency also is a function of chemical structures, i.e., highly refractory compounds

such as polycyclic aromatic hydrocarbons (PAHs) may have much lower limits in terms of a linear relationship. In this work, it was confirmed that up to ~7 mg/L of DOC, SRHA showed the same linear relation to that of KHP. SRHA was selected because it is a known heterogeneous mixture and a common reference NOM. The DOC of all samples was adjusted to below 5 mg/L. Based on UV absorbance and DOC values, the SEC-SUVA was calculated at the elution volume *i* (equation 6).

$$\text{A.U.}_{\text{wavelength 1, } i} = 2.022 \times \text{mV}_{\text{output 1, } i} \tag{5-4}$$

$$\text{A.U.}_{\text{wavelength 2, } i} = 10.115 \times \text{mV}_{\text{output 2, } i} \tag{5-5}$$

$$SEC - SUVA_{254,i} = \frac{A.U._{254,i} \times 100}{Converted\ DOC_i} \tag{5-6}$$

Where $A.U._{254,i}$ is the absorbance at the retention time or at eluted volume *i*, and *converted* DOC_i denotes that the amount of DOC calculated from the electrical response at the retention time or eluted volume *i*.

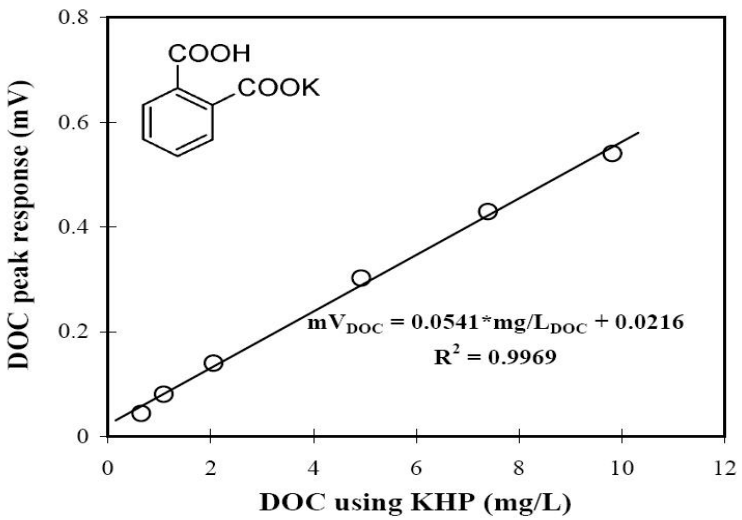


Figure 5-3 The relation curve of DOC concentrations (mg/L) versus its corresponding electrical responses (mV) using KHP. (A chemical structure of KHP is shown inside figure)

Her *et al.* [88, 89] suggested the URI as the UV absorbance ratio at 210 and 254 nm (UVA_{210}/UVA_{254}). Figure 5-4 presents the comparisons of SEC-URI and SEC-SUVA of KHP calculated using the electrical signal and absorbance converted from the electrical signal. As shown, the average SEC-URI and SEC-SUVA of KHP using the electrical signal were 0.1 (dimensionless) and 11.2 mg/L·m, respectively, while these values were 0.5 and 1.1 mg/L·m when using absorbance (A.U.). SEC-URI by electrical signal was found to yield one fifth of the values by absorbance. SEC-SUVA by electrical signal was ten times greater than values calculated using absorbance. Nevertheless, in the work by Her *et al.* [88, 89], URI values of several reference DOM samples calculated using the electrical signals, not the absorbance, were demonstrated, for example, the lowest for humic acids (1.59 for a humic acid), medium for fulvic acids (1.88 for a fulvic acid), highest for proteins (13.50 for bovine serum albumin). Therefore, although the overall URI patterns (i.e., higher values for proteins or organic matter with high contents of organic nitrogen, lower for aromatic organic matter, etc.) may be a trend, URI values presented as specific numbers are not universal and cannot be used as a reference numbers because the electrical signal responding to the same concentration of light-absorbing moieties will vary depending on the system set-up. Conversely, absorbance is universally reproducible regardless of the system configuration. Therefore, correction of URI values suggested in Her *et al.* [88] for modified URI values using reference NOM and EfOM samples will be applied to understanding of the SEC characterization of NOMs with the same reasoning, modified SEC-SUVA using absorbance is proposed.

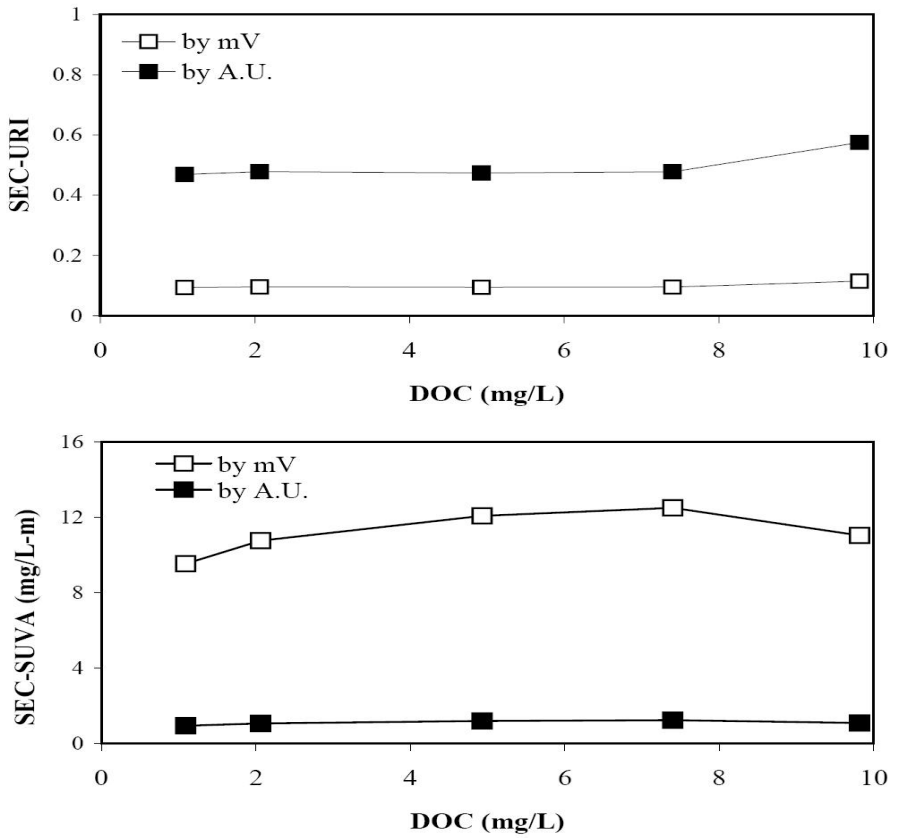


Figure 5-4 Calculation of SEC-URI and SEC-SUVA of KHP (□: by the electrical signal (mV), ■: by A.U. converted from mV using equation 4 or 5 (mV) of UVA_{254}, i.e., SUVA=UVA_{254} (as mV) ×100/DOC (as mV))

Characterization of NOM and EfOM

***SUVA and Fractions*:** SEC-SUVA is useful in understanding the degree of aromaticity of organic substances at a certain Rt and provides an improved assessment of compounds eluted at Rt. Table 5-1 shows the levels of DOC, UVA and SUVA for the organic matters components used in this study. For the EfOM samples, the polarity of organic matter was investigated by XAD-8/-4 fractionation. SRHA and SRNOM showed high SUVA values, compared to isolated EfOM and SMP samples while EfOM samples exhibited a relatively low SUVA and a significant hydrophilic fraction. In general, waters with a high SUVA are enriched in hydrophobic NOM. Therefore, SRHA and SRNOM consist of hydrophobic organics and have a higher degree of

aromaticity compared to EfOM samples. Also, for isolated EfOM, hydrophobic fraction reflected a greater SUVA compared to the SMP sample.

Table 5-1 Specific UV absorbance of NOMs (SRHA and SRNOM) and EfOMs (isolated EfOM and SMP) samples

DOM	DOC (mg/L)	UVA_{254} (cm^{-1})	$SUVA_{254}$ (L/mg·m)	HPO (%)	TPI (%)	HPI (%)
SRHA	2.24	0.085	3.810	93.5*	5.2*	1.4*
SRNOM	1.19	0.043	3.646	-	-	-
Isolated EfOM	6.69	0.113	1.693	41.1	21.4	37.5
SMP	2.35	0.027	1.170	17.5	28.7	53.8

* from [88]

***Molecular weight distribution of samples by HPSEC*:** The HPSEC-DOC chromatograms of representative NOM and EfOM samples are presented in Figure 5-5. Based on the retention time of the SMP, DOC chromatograms were divided into 3 fractions (Zone-1: high MW, 30 min; Zone-2: moderate MW, 30-70 min; Zone-3: low MW, after 70 min). HPSEC-DOC chromatograms for SRHA and SRNOM showed close to unimodal and symmetric features although some tailing phenomena at high retention time (low MW) were displayed due to possible interaction with the stationary phase in the column. The MW distribution of SRHA started from a shorter Rt than that of SRNOM, meaning that SRHA contains more higher MW substances compared to that in SRNOM. Maximum peaks of each were shown at a Rt of ~42 min and ~45 min, corresponding to MWs of 3,030 and 2,400 Da, respectively.

On the other hand, HPSEC-DOC chromatograms of EfOM samples displayed distinct differences compared to NOM samples. Both isolated EfOM and SMP samples showed clear/high peaks in Zone-1 (~1000 Da, Rt ~30 min) with quite long Rt intervals to Zone-2, showing that the MWs of SRHA and SRNOM are in the boundary between Zone-1 and Zone-2. Isolated EfOM showed the presence of a significant amount of moderate-sized materials (Zone-

2) at ~50 to ~70 min (a few hundreds to a few thousands Da) as well as an intense peak of high MW substances (Zone-1). A long tailing was shown for the low molecular weight substances (the Zone-3). By simple comparison of the MW distribution between EfOM and NOM samples, it was observed that EfOM samples consist of a diverse-range of MW substances over very low MWs through high MWs, and the humic substances contained in EfOM samples (Zone-2, typically known as humic substances peak) represent lower MW materials than humic materials in SRHA or SRNOM although the diversity of their MWs is more wide than corresponding constituents in NOM samples.

Yet, there are a few things to consider in terms of chromatographic interactions. Fractionation by size-exclusion is assumed to occur in the stationary phase (the gel in the column) which is inert to the solute. In practice, however, it is not possible to synthesize a hydrophilic gel which would be perfectly inert in aqueous medium and a variety of hydrophilic gels whose spatial structure is based on ether bridging (for example, the Toyopearl resin used in this study) are weakly negatively-charged due to partial hydrolysis of the ether sites in aqueous medium [90, 92]. Therefore, organic substances with negatively charged (more polar, more hydrophilic) functional groups may experience electrostatic repulsive interactions which result in the over-exclusion of these substances to a shorter Rt than expected.

Since substances in Zone-1 are known to be mainly from polysaccharides, proteins, or high MW deprotonated acids associated with carboxyl group at the experimental pH, the shifting effect to large MW should also be considered although their MWs are still high. In addition, the cross-linked structure of the gels creates hydrophobic sites on its surface [92]. Hydrophobic interactions with the gels can cause a retardation of the analyte, especially for the hydrophobic organic matter. The greater the hydrophobicity of the analyte is, the stronger the effect will be, so

that the MW distribution will be shifted to a smaller MW by eluting later. Thus, the long tailing phenomenon observed in Zone-3 of the isolated EfOM, along with the high amount of smaller organics compared to SRHA/SRNOM may be due to a hydrophobic interactions effect of humic substances.

For the SMP chromatogram, distinct MW distributions were exhibited in the three regions (Zone-1 ~ 3). Significant peaks appeared at Zone-1 and Zone-3, and a rather broad and low peak at Zone-2. The peak in Zone-3 was not observed for isolated EfOM, however, the wide range in MW distributions were similar to isolated EfOM. Individual integration of each fraction were defined percentages of each peak as 18.5, 26.6 and 54.8% of total DOC, meaning a dominance of low MW compounds among a mixture of different MW compounds. Additionally, the M_w and M_n of each peak were 9,170 and 8,358 Da (ρ=1.1) for Zone-1, 752 and 496 Da (ρ=1.5) for Zone-2, and 60 and 56 Da (ρ=1.1) for Zone-3, respectively while gross M_w and M_n calculated from the entire chromatogram were 598 and 90 Da (ρ=6.6), respectivly. A big difference of ρ between each peak and the entire chromatogram means that although bulk SMPs consist of various MWs compounds, major constituents of SMPs can be divided into several MWs segments and these major constituents are likely an assemblage of compounds with similar MWs and/or properties. In terms of M_w and M_n calculations, Her *et al.* [93] proposed that separate values for each fraction should be presented independently in case of multiple MW fractions. In agreement to his suggestion, the results presented herein indicate that the gross M_w and M_n values for SMPs without consideration of fractional importance may result in misleading information on actual MW values. Therefore, gross values of M_w and M_n may help to understand the heterogeneity of DOM in samples whereas M_w and M_n of individual fractions may provide better elucidation of the composition of DOM, or aid in better interpretation of behaviors of specific fractions in

environmental processes.

Jarusutthirak *et al.* [94] presented the transformation of organic compounds during biological processes in which glucose was initially biodegraded and transformed to intermediate MW substances (>300 Da) assumed to be utilization associated products (UAP) related to biomass growth and substrate utilization; afterward, large MW substances (>10,000 Da) were formed gradually as the sludge retention time (SRT) increased to 2, 5, and 10 days. The large MW substances were attributed to biomass associated products (BAP) caused by cell lysis during the endogenous phase. Thus, it is may be presumed that peaks in Zone-2 and Zone-1 are UAPs and BAPs of SMPs, respectively.

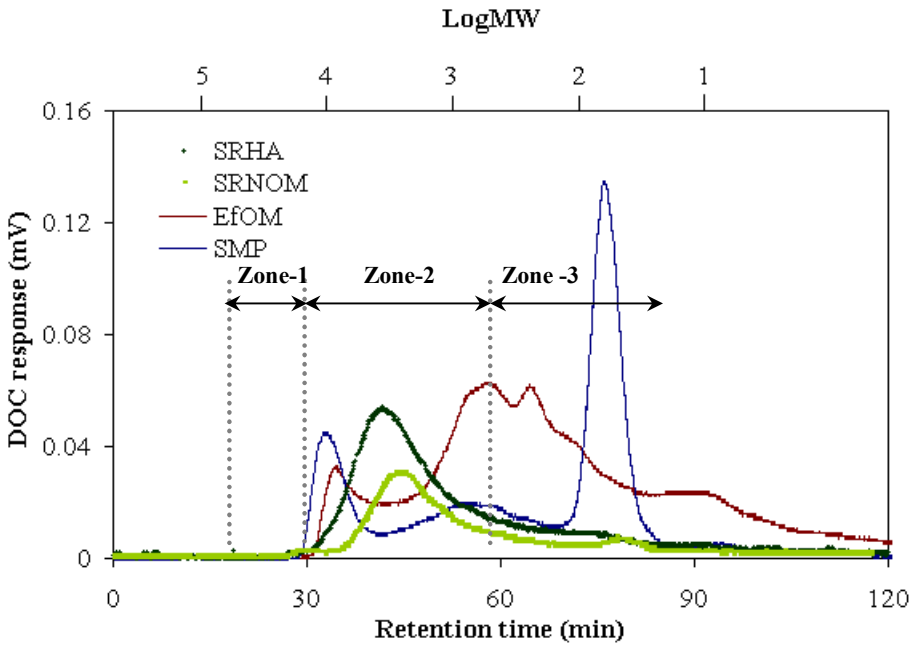


Figure 5-5 HPSEC-DOC-UVA chromatogram of SRHA, (aquatic) SRNOM, EfOM and SMP; TSK HW-50S column (2×25 cm), Na_2SO_4 eluent with phosphate buffer (pH: 6.8, ionic strength: 0.1 M), flow rate 1 mL/min

***SEC-UVA*:** Figure 5-6 shows the SEC-UVA profiles of NOM and EfOM samples measured at multiple wavelengths of 210, 230, 254, 280, 300 and 350 nm. Wavelengths used were selected

based on typical UV/Vis ranges that are able to detect chromophores since in general a wavelength of greater than 400 nm is not commonly used due to a low molar absorptivity for most of DOM constituents. For NOM samples (SEC-UVA for SRFA omitted due to similarity to that of SRHA), organic matter showed a maximum UV absorption at 210 nm followed by 230 nm which means that the maximum molar absorptivity of NOM is at 210 nm. SEC-UVA at 254 and at 280 nm did not show big differences. UVA patterns at the maximum peak locations (shown in Figure 5-6-(a)) with respect to wavelengths were similar to a typical UV spectrum of organic matter which is featureless, and (exponentially) decreasing as wavelength increases. For SRNOM, absorption after 280 nm significantly decreased compared to SRHA. This is thought to be because SRHA contains molecular moieties that absorb low energy light more abundantly than SRNOM, and the greater molar absorptivity (ε, L/mol-cm) of SRHA than SRNOM at corresponding peak regions confirmed this (results not shown, but data included in appendix). In addition, large differences between 254 nm and 280 nm, which are commonly used as an indicator of aromaticity of NOM and which do not exhibit significant interferences by inorganic species, may indicate that UVA_{254} is a more sensitive or better aromaticity indicator than UVA_{280}.

On the other hand, for EfOM samples (Figure 6-6-(c)), isolated EfOM did not show absorption at shorter Rt values corresponding to ~10,000 Da of MWs (Zone-1 region). Considering the high DOC peak in this region, the most probable substances would be macromolecules such as polysaccharides that neither respond to UV nor possess conjugated bonds. In the Zone-2 region, UVA at 230 nm showed the greatest absorption, whereas UVA at 210 became a maximum at low MW ranges. Biopolymers in NOM have different portions of functional groups (ex. –COOH, $-NH_2$, etc) and aromatic rings, resulting in different absorptivities, depending on wavelengths. For an aromatic ring compound (i.e., benzene), a wide

absorption appears at 180-210nm by the ethylenic band (the E band), at 250-295 nm for the benzenoid band (the B band), and at 275-330 nm for the K band. Benzene displays several absorption bands in the UV region and shifting of the absorption band and/or the ε varies, depending on the degree of conjugation. The light absorption of organic matter varies with substituents on the benzene ring [95]. The main NOM functional groups associated with non-aromatic groups display absorption maxima only at shorter wavelengths. In the case of the un-conjugated form, absorption maxima wavelengths are 206 nm for carboxylic acids and esters, and 210 nm for amides in water (by n → π* transitions). When carboxylic acids or esters are conjugated, peak maxima absorption occurs by π → π* transitions at longer wavelengths than 210 nm. The peak maxima absorption shifts to longer wavelengths with the substituents from the base absorption of 187 nm in water. The absorption band increment is dependent upon the substituent type and position of the unsaturated carbonyl chromophore [89, 96]. Therefore, in somewhat the same sense of structural or spectroscopic understanding, it may be said that the HS fraction of EfOM samples differs from the HS fraction of NOM, showing lower MWs than NOM, possibly due to biodegradation/transformation of NOM to humic substances in EfOM and consequently shifting to HS with lower MWs. For the low MW region, high absorbance at 210 nm, followed by 230 nm, and subsequent absorption decrease with increasing wavelengths, indicating that LMW compounds in Zone-3 are attributable to analogues of humics or smaller molecules from microbial sources along with nitrate species (MW=62 Da).

For SMP compared to isolated EfOM (Figure 6-6-(d)), large MW compounds in Zone-1 showed a high absorption peak only at 210 nm, indicating that there are large macromolecules composed of polypeptide bonds such as proteins. High absorption by 210 and 230 nm in Zone-3 is thought to be due to similar reasons. Further investigation of LC-OND (measured by DOC-

LABOR, Dr. Stefan Huber) confirmed that SMP contained significant amounts of low MW organic nitrogenous substances (Figure 5-7). For an early retention time (approximately matching to the Zone-1 region in the SEC-DOC chromatogram in Figure 5-5), DON was shown to increase, whereas the humic substances peak showed a relatively low DON peak.

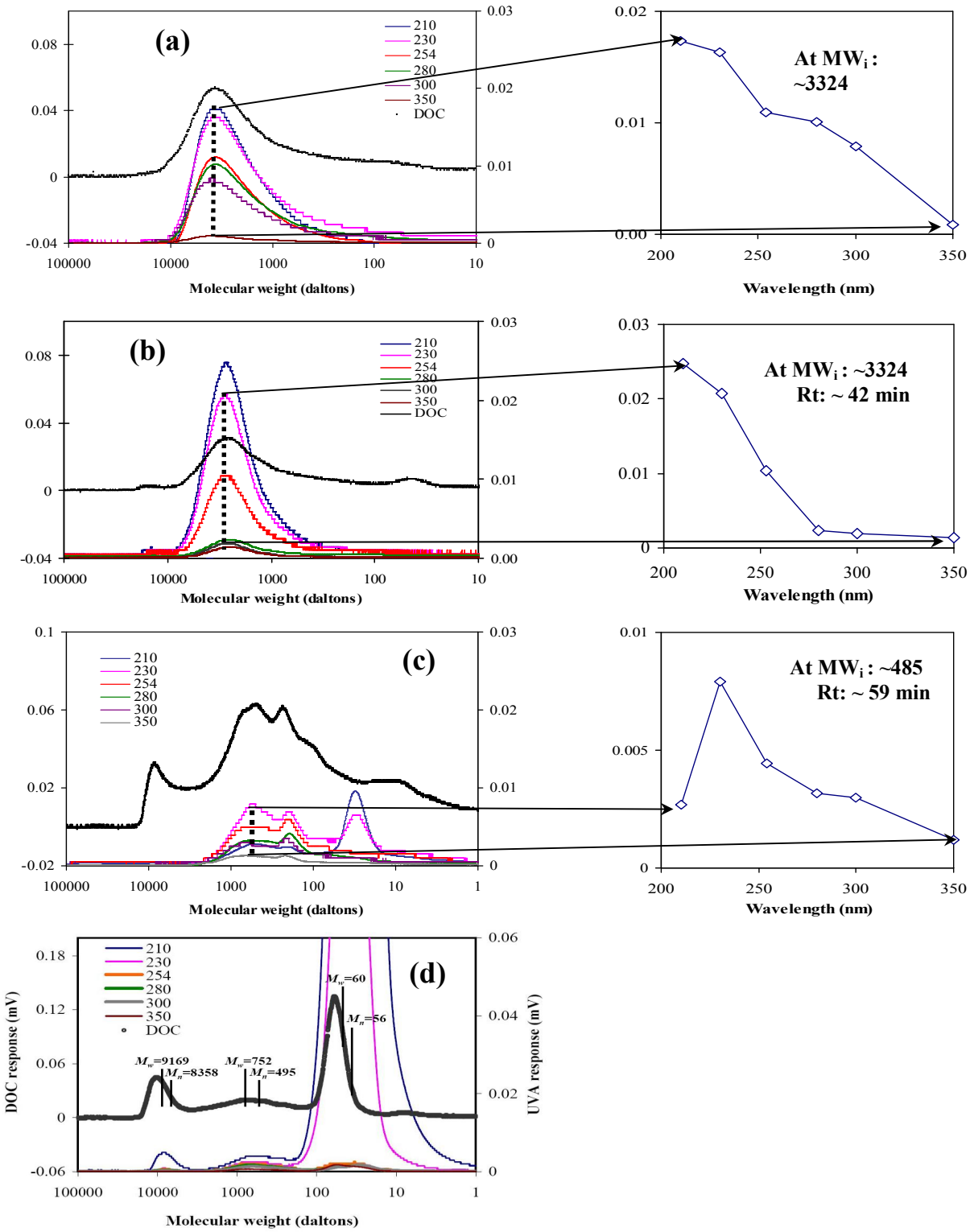


Figure 5-6 HPSEC-UVA chromatograms and UVA maxima of SRHA (a), (aquatic) SRNOM (b), EfOM (c) and SMP (d) at various wavelengths. The wavelengths were 210, 230, 254, 280, 300 and 350 nm; TSK HW-50S column (2×25 cm), Na_2SO_4 eluent with phosphate buffer (pH: 6.8, ionic strength: 0.1 M), flow rate 1 mL/min

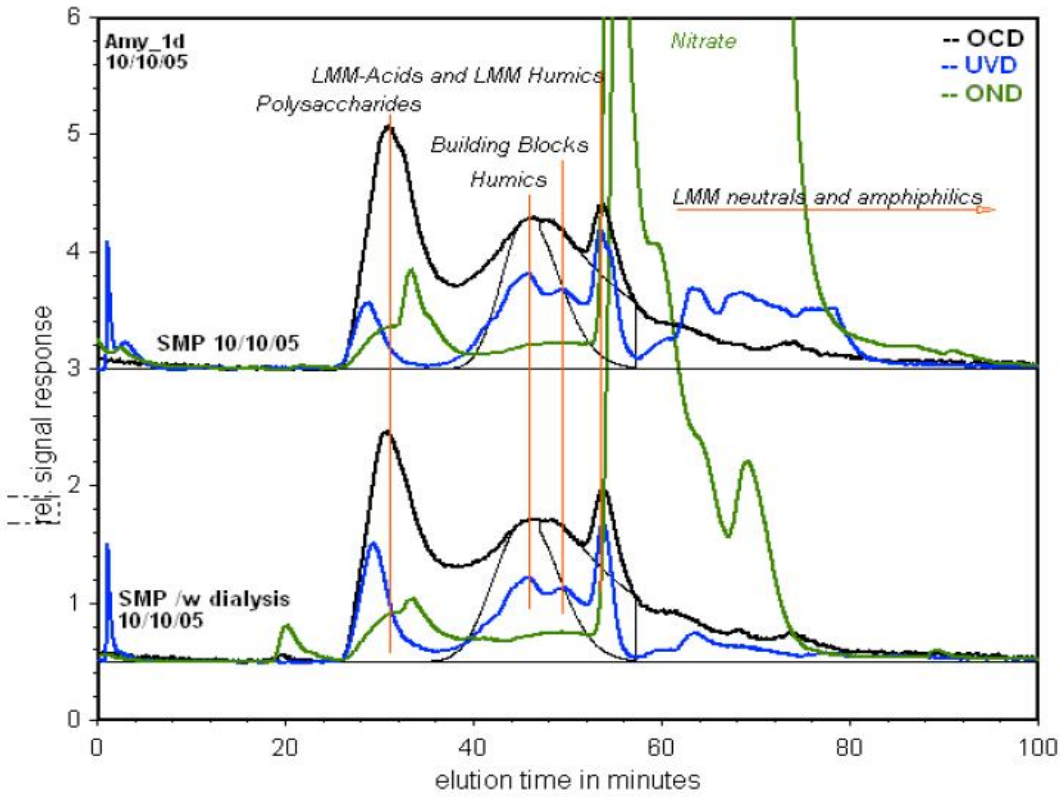


Figure 5-7 LC-OCD/UVD/OND chromatograms of SMP derived from the 100 Da dialysis (measured in DOC-LABOR, Dr. Stefan Huber, in October, 2005): This SMP sample was from the same SBR as the one shown in chapter 3, but for a different sampling date and thus may be slightly different chromatogram from the one in chapter 3. Also, the detailed eluent compositions and buffer may cause a different elution chromatograph as well as retention time

Table 5-2 compares SEC-URI using two methods by the electrical signal (mV) ratio and the absorbance ratio. For URI chromatograms of EfOM and SMP, it was difficult to choose certain specific points or trends, so it was not included in the table, but presented with SEC-SUVA profiles in Figure 5-8. As seen, URI variations by absorbance were similar among NOMs but for proteins, big differences were observed between the two methods, displaying approximately 5 folds greater values using absorbance than the electrical signals.

Table 5-2 Comparisons of SEC-URI values by electrical signal (mV) and absorbance (cm^{-1}) around at maximum peak positions

	SRHA	SRFA	SRNOM	Tryptophan	Bovine serum albumin
By mV	1.8 (1.59)*	1.7~1.8 (1.88)*	2.3	2.8-3.0	9.5 (13.5)*
By Abs.	1.57	2.7-2.8	2.6-2.7	14.8	47.5

*values in brackets introduced in Her et al [88].

Table 5-3 Molar absorptivity (ε) values of various compound types at specified wavelengths (summarized from references [15]** and [97])

Name	Chromophores	Wavelength	Molar extinction, ε
Acetylide	-C=C	175-180	6,000
Aldehyde	-CHO	210	1,500
Amine	$-NH_2$	195	2,800
Azo	-N=N-	280-400	3-25
Bromide	-Br	208	300
Disulphide	-S-S-	194	5,500
Ester	-COOR	205	50
Ether	-O-	185	1,000
Ketone	>C=O	195	1,000
Carboxyl	-COOH	200-210 230¶	50-70 10,000¶
Nitrate	$-ONO_2$	270 210	12 7,490**
Nitrite	-ONO	220-230	1,000-2,000
Nitro	-N=O	210	Strong
Phenyl	-OH	210.5¶ 270¶	6,030¶ 1,480¶

¶substituent of benzene

Figure 5-8 depicts SEC-SUVA and SEC-URI of EfOM and SMP samples. As expected, the trend shows lower SUVA and higher URI in Zone-1 and Zone-3 regions, and increased SUVA and decreased URI for Zone-2 (humic peak). These results confirm that high MW substances (Zone-1) and low MW substances (Zone-3) of EfOM are not aromatic and presumably closely related to a hydrophilic and labile nature. Comparison of the URI chromatogram of SMP to the DON chromatogram (LC-OND included in Appendix) showed similar trends although there

were differences in retention time and features of the entire chromatogram due to analytical differences and different sampling dates. Through the overall URI results, it can be concluded that URI can be used a surrogate detector for DON detection in SEC when a DON detector is not available. However, it should be noted that URI has inferences by the presence of inorganic species, especially nitrate (ε=7,490 $M^{-1}cm^{-1}$ at 210 nm) or nitrite (1000-2000 $M^{-1}cm^{-1}$ at 220-230 nm), on organic absorption (e.g., aldehyde (-CHO): 1,500 $M^{-1}cm^{-1}$ at 210 nm, carboxyl (-COOH): 50-70 $M^{-1}cm^{-1}$ at 200-210 nm) (refer to Table 5-3); thus, the URI chromatogram may be useful in recognizing the high MW peak.

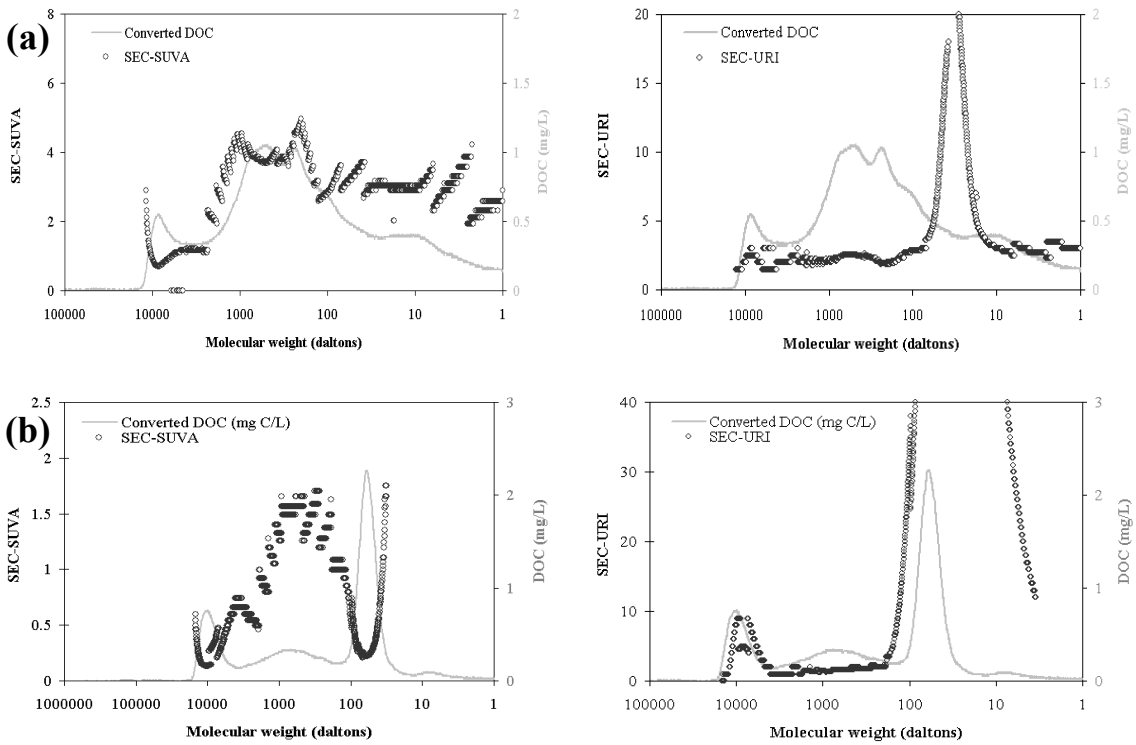


Figure 5-8 HPSEC-SUVA-URI profiles of EfOM and SMP. The left column is SEC-SUVA, and the right column is SEC-URI; TSK HW-50S column (2×25 cm), Na_2SO_4 eluent with phosphate buffer (pH: 6.8, ionic strength: 0.1 M), flow rate 1 mL/min. "Converted DOC" in the graphs means "DOC" chromatogram converted from its electrical signal (mV) to mgC/L

***M_w of NOM and EfOM*:** Based on the SEC-DOC and UVA_{280} chromatogram with the criterion of 1% MW cutoff at the maximum response, M_w, M_n and ρ were calculated and summarized in Table 5-4. M_w and M_n values exhibited ~400 daltons higher in SRHA than in SRNOM, but the ρ values were similar. UVA_{280} detection showed a tendency to decrease M_w, and M_n as well as resulting in smaller ρ numbers. It is important to note that DOC detects any substances with organic carbon, but UV detects only light absorbing substances (chromophores). Therefore, substances such as high MW polysaccharides, which can be recognized by DOC detection, are undetectable to UV. For EfOM samples, ρ assumed greater values, almost 5-fold over NOM, while M_w and M_n were much less than NOM, although there was a high peak of large MW substances. However, this is reasonable because their molecular weight distribution was relatively wide spread from high MW to very small MW as well as a tailing effect for low MW region, which was not fit in applying the 1% MW cutoff criterion. Moreover, the distribution was far from a normal distribution. Less M_w and M_n for EfOM imply that a high amount of humic substances in Zone-2 forced the apparent overall (average) molecular weight of EfOM to smaller values than those of NOM, although EfOM contains a significant amount of high molecular weight substances. High ρ means that EfOM is more heterogeneous and is a complex mixture of organics in composition.

Table 5-4 Molecular weight parameters of NOM (SRHA, SRFA and SRNOM) and EfOM (isolated EfOM and SMP) samples (*PSS as MW calibration standards)

			In this study			Reported values			
DOM	Type of organic origin	Detection	M_w	M_n	ρ	M_w	M_n	P	Ref
SRHA	Allochthonous	DOC	2795	1719	1.63	3305	1934	1.71	[93]
		280	3070	1996	1.54	3703	1807	2.05	[98]*
SRFA	Allochthonous	DOC	2412	1552	1.55	2114	1385	1.53	[93]
		280	2760	1765	1.56	2310	1360	1.70	[72]*
						1950	1112	1.75	[99]
SRNOM	Allochthonous	DOC	2385	1452	1.64				
		280	2065	1386	1.49				
EfOM	Autochthonous	DOC	1339	132	10.14				
		280	409	126	3.25				
SMP	Autochthonous	DOC	598	90	6.63				
		280	505	75	6.77				

***Fluorescence Excitation-Emission Matrix (EEM)*:** The EEMs of reference NOMs (a-d: SRHA, SRFA, SRNOM and Lake Fryxell in order), EfOMs (e-g: isolated EfOM and SMP), and single protein fluorophores (h-j: tyrosine, tryptophan, and albumin) are shown in Figure 5-9. Excitation and emission wavelength positions for different organic matter isolates/materials are summarized in Table 5-5.

EEMs of NOMs exhibit similar contours with two distinct excitation and emission maxima, in excitation/emission ranges of 268-270 nm/467-478 nm. (humic-like peak, 1st maxima). But shifting of both the 1st maxima and 2nd maxima to shorter wavelengths (known as blue-shifting[¶]) among NOM samples was observed (i.e., blue shifting from SRHA to SRFA, and to SRNOM). Several factors such as molecular weight (structure), elemental compositions, functional groups, pH, etc., are involved in spectral shifting. An increase in the extent of the π-electron system (i.e., a degree of π-π* conjugation in the aromatic) leads to a shift to longer excitation/emission

[¶] a blue shifting: a shift to shorter excitation/emission wavelength (corresponding to higher energy).

wavelengths[§], and an increase of conjugation is generally accompanied by increased molecular weight and fluorescence quantum yield. As demonstrated earlier, SEC showed SRHA had the highest MW followed by SRFA and SRNOM in addition to many similarities for chromatographic and UV spectroscopic patterns. According to previous researchers, it has also been reported that increases of MW for the "same or similar" types of organic matter caused shifting of emission spectra to longer wavelength [49, 50]. Hayase and Tsubota (1985) stated that excitation/emission maxima in sedimentary fulvic acid were blue-shifted relative to humic acid [100]. Thus, of the various factors, molecular weight differences by degree of conjugation among NOM samples would be mainly responsible for the spectral shifting.

Single proteineous compounds: albumin, tyrosine and tryptophan; exhibited a single excitation and emission maxima in high energy regions of excitation 270-280 nm/emission 300-350 nm (protein-like regions). EEMs absorbing high energy light (UV-C, less than 290 nm) were obvious discrepancies compared to NOM samples. On the other hand, the EEM of EfOM showed three distinct maxima: the most intense peak in a protein-like region, the second and third maxima resembles those of NOM, but their wavelengths were located at blue-shifted positions. The EEM of SMP was also a blue-shifted EEM compared to NOM.

The fluorescence index (FI) of samples has been used to distinguish the source of DOM between terrestrially-derived and microbial-derived origin [43]. The increased FI values in EfOMs or proteineous DOMs compared to NOM samples were also consistently observed. Hydrophilic DOM showed a higher FI than hydrophobic DOM and this is likely because fluorescence intensity at Ex/Em-370/450 nm (i.e., the numerator of FI) is close to the maxima of hydrophilic DOM and the intense protein-like peak location of EfOMs. In order to further

[§] a red shifting: a shift to longer excitation/emission wavelength (corresponding to lower energy).

understand SMP, Zone-1, 2 and 3 samples of SMP was collected from SEC and measured by EEM. As shown as Figure 5-9-e, the EEM revealed very similar contour patterns compared to single protein-like EEMs, especially, the tyrosine-like peak (Figure 5-9-h). This means that protein-like fluorophores with large molecular weight are also one of the SMP members along with polysaccharides that are not detectable to UV or fluorescence. The EEM from the fractions 2 and 3 of SEC-DOC showed mostly like blue-shifted EEM of NOM, but the shape between humic-like and protein-like peaks was more broadly spread (data not shown).

The maxima of EEMs of EfOMs and SMP were compared with several isolated organic fractions such as HPO, TPI, HPI listed in Table 5-5, and WW HPI samples (DN XAD4 and HS XAD4 in Table 5-5) passed through XAD-4 resin. Samples named "DN XAD4" and 'HS XAD4" were thought to be composed of "hydrophilic" substances (i.e., HPI) after passing through the XAD-4 resin. As a result, their EEMs showed similar trends to those of TPI and HPI, and the first maxima of SMP and the second maxima of EfOM were close to the first maxima of HPI or TPI substances. Although the peak of fraction 1 of SMP (Figure 5-9-g) did not seem to appear in the bulk SMP EEM (Figure 5-9-f), it may be because this peak was hidden under the whole spectra due to the higher intensity of the first and second maxima, thus resulting in an obscuring of less intense spectra.

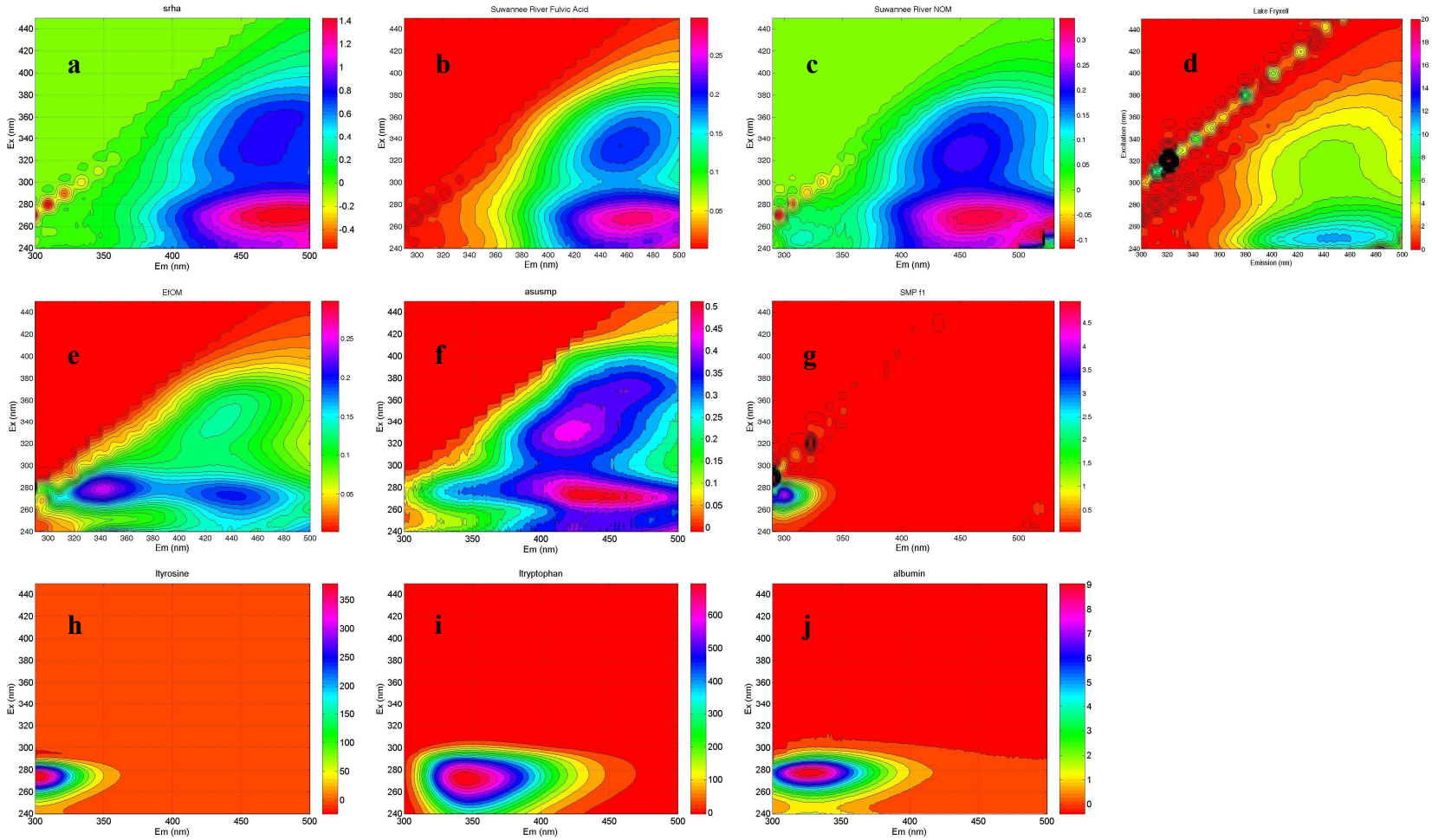


Figure 5-9 Examples of reference fluorophores: a-d for SRHA, SRFA, SRNOM and Lake Fryxell for reference NOM, e-g for isolated EfOM, SMP and SEC fraction 1 of SMP as autochthonous OM, h-j for L-tyrosine, L-tryptophan, and Albumin for protein-like OM

Table 5-5 Summary of fluorescence maxima locations of various different fluorophores in EEM

	Excitation & Emission wavelength (nm)						Stoke's shift for 1st maxima	Fluorescence Index	Type & descriptions of OM
	1st max		2nd max		3rd max				
SRHA	270	478	347	472			208	0.8643	River, Allochthonous
SRFA	268	467	333	455			199	0.9686	River, Allochthonous
SRNOM	269	468	324	450	249	315	199	0.9901	River, Allochthonous
Nordic FA	250	449	300	434			199	0.8763	Lake
SR HPOA	249	458	301	440			209	0.9249	River
SL HPOA	268	451	321	446			183	1.0575	Lake
SL TPIA	272	444	324	430			172	1.323	Lake
CRW TPIA	273	424	309	403			151	1.3416	River
CRW HPIA	274	424	328	421			150	1.3336	River
BL HPOA	269	448					179	1.2352	Lake
BL TPIA	271	431	320	410			160	1.4724	Lake
SMP (2005)	279	425	347	423			146	1.5897	Autochthonous
EfOM	279	343	272	440	339	430	64	1.4595	Autochthonous
First peak of SMP	273	300					27	-	Autochthonous
Albumin	277	328					51	-	Protein
L-tryptophan	272	345					73	-	Amino acid
L-tyrosine	274	303					29	-	Amino acid
91ST WW HPOA	246	415					169	1.1516	Wastewater EfOM
BDW HPOA	246	416	277	348			170	1.2507	Wastewater EfOM
BDW TPIA	291	408					117	1.3517	Wastewater EfOM
BDW XAD4	296	406					110	1.4275	Wastewater EfOM
SJW XAD4	264	370	319	394			106	1.3802	Wastewater EfOM
DN XAD4*	314	399					85	1.5582	Wastewater EfOM
HS XAD4*	280	399	318	404			119	1.6996	Wastewater EfOM
**component 1	270	455	360	455			185	-	Allochthonous
**component 2	295	343					48	-	Autochthonous
**component 3	325	401					76	-	Allochthonous
**component 4	275	326					51	-	Autochthonous
**component 5	240	397					157	-	Allochthonous

-: inappropriate to report that they are a single or pure fluorophores, *EEMs are included in Appendix, **: PARAFAC components introduced in Chapter 4

***Fluorescence vs. DOM properties*:** In order to investigate the fluorescence spectroscopic similarities/dissimilarities between NOM and EfOM, the relation of molecular weight of DOM and fluorescence wavelength was examined using isolates or fractions of NOM and EfOM. Previous studies have shown that blue-shifting of excitation and emission maxima as the molecular weight decreases [49, 50]. As a whole, since 50-60% of DOM in humic and fulvic acids is known to fluoresce [45], molecular weight by SEC-DOC chromatogram may not be appropriate to make a connection with fluorescence. For example, in the case of a chromatogram containing a large MW DOC peak, but no UV response, the molecular weight of substances involved in fluorescence would be overestimated, and this error would become greater for wastewater effluent with a high concentration of polysaccharide materials. Rather, SEC-UVA or SEC-fluorescence chromatogram may provide a more accurate understanding of the MW of fluorophores.

Figure 5-10 shows SEC-DOC-fluorescence of SRFA and EfOM as representative results of ~30 different isolates since it is too extensive to cover all results here. However, general observations and trends were mostly similar and can be generally described. Selections of fluorescence excitation and emission wavelengths in SEC-fluorescence detector were based on the most intense maxima and the second maxima positions, as derived from corresponding EEMs. In comparing SEC-fluorescence to SEC-DOC chromatograms, the maxima of the fluorescence peaks were shown at a later retention time than the maxima based on DOC. This same pattern of results occurred for most of the samples with different extents of delayed Rt, differing depending on samples (tens of seconds to hundreds seconds, ~3 min for SRFA); this has been also reported by other researchers [81, 101] although they did not make observations at several wavelengths. In the UV chromatogram, retardation of retention time was not observed as illustrated in Figure

5-6. Considering that the eluting time gap between the fluorescence and DOC detectors was only ~5 seconds based on the system configuration (e.g., tubing diameter & length between fluorescence cell and DOC detector, and flow rate), the delayed Rt in the fluorescence chromatogram indicates that the smaller the molecules, the more fluorescing (higher intensity). For SRFA, M_w comparison showed a greater value by fluorescence (2,850 Da) than by DOC (2,412 Da), and this was because the fluorescence chromatogram was more sharp (higher kurtosis) and narrow than the DOC chromatogram, yet, tailing in DOC chromatogram suggested fluorophores of DOC are of a smaller M_w than the intrinsic one. Therefore, the MWs of fluorophores in NOM (or at least for SRFA in this study) are more closely distributed than MWs of chromophores (light-absorbing moieties) and all of the carbon-containing moieties (i.e., MW ranges of fluorophores group < chromophores group < bulk DOM consisting group). In addition, a decreasing of fluorescence intensity of molecules at the different Ex/Em wavelengths, but at the same Rt (i.e., same MW), indicates that wide ranges of Ex/Em wavelengths in EEM are associated with a similar (group of) NOM with similar MW, furthermore, the first and second maxima of SRFA located at Ex/Em=268/467 nm and 333/455 nm (Table 5-5) were basically from the same NOM moiety.

On the other hand, fluorescence chromatograms of EfOM samples (Figure 5-10-bottom) exhibited multiple peaks at different retention time. Two peaks at Ex/Em=275/300 nm correspond to high fluorescence intensity in Zone-1 region and lower intensity in Zone-2 region, and this means that humic substances in EfOM also contribute to fluorescence in the protein-like peak with a minor fraction. By integration of peak area, of total fluorescence at 275/300 nm, the first peak area was responsible for 81%, and the second peak was 19%. At 300/410 nm fluorescence, high MW substances of Zone-1 did not show fluorescence, but humic substances

and low MW substances fluoresed. Their contributions to fluorescence at 300/410 nm were 75% and 25%, respectively. This suggests that low MW substances of EfOM may not be major fluorophores in the humic-like region in EEMs, and most of the EfOM fluorescence in the humic-like region is from fluorophoric moieties of humic substances.

In additional linking to the M_w, each peak was found out that two peaks at 275/300 nm were 11,330 Da and 430 Da, whereas, of two peaks of 310/400 nm chromatogram, one was similar to the second peak value of 275/300 nm chromatogram and the other showed tens of daltons.

Figure 5-11 demonstrates the locations of Ex/Em wavelengths maxima as well as M_w determined by SEC-fluorescence. NOM samples did not show their maxima emission wavelengths below 400 nm, and the Stoke's shift for the first maxima, in which the Stoke's shift ($\Delta\lambda$) is the difference in nanometers between the peak excitation and emission wavelengths (i.e., $\Delta\lambda=\lambda_{Ex}-\lambda_{Em}$), were always greater than 150 nm (i.e., $\Delta\lambda > 150$ nm, also refer Table 5-5) while EfOM or protein samples exhibited a shorter gap. NOM fractions separated by their polarities showed a tendency to have blue-shifted maxima compared to bulk NOM, showing that hydrophilic NOM is more shifted than hydrophobic NOM possibly due to the density of π conjugation and the functional groups. The HPI fraction of EfOM appears to have maxima at shorter wavelengths than the HPI or TPI of NOM. As aforementioned, some peaks of EfOM and SMP samples corresponded to TPI or HPI fractions. Through investigating the relations of MWs of organic matter with fluorescence excitation and emission wavelengths, based on their spectroscopic properties, it can be concluded that fluorescence wavelengths fundamentally depend on the chemical structures with specific bonding that can absorb and fluorescence, rather than the molecular weight or size of organic matter. However, it can be deduced that there is relevance the higher the MW of fluorophores, the longer the wavelengths fluorescence. Thus,

focusing on the blue-/red-shifting of fluorescence should be used for organic matters with similar origination (e.g., allochthonous, autochthonous, etc.) or with carefully prepared organic fractions either by MWs or by polarity, and it may not be appropriate for bulk organic matters.

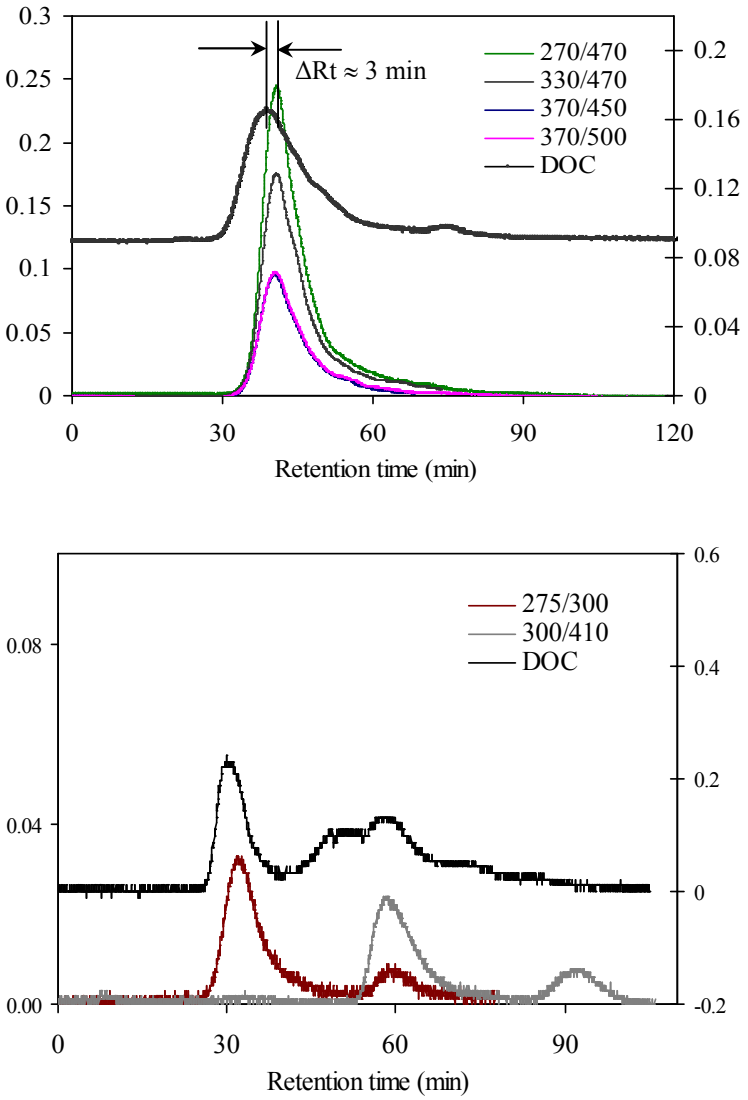


Figure 5-10 SEC-DOC-fluorescence chromatograms of SRFA (up) and EfOM (wastewater) samples (down); in EfOM chromatogram, straight bars and dotted bars mean segments of the peak integration ranges for 275/300 nm and 300/410 nm fluorescence chromatogram, respectively

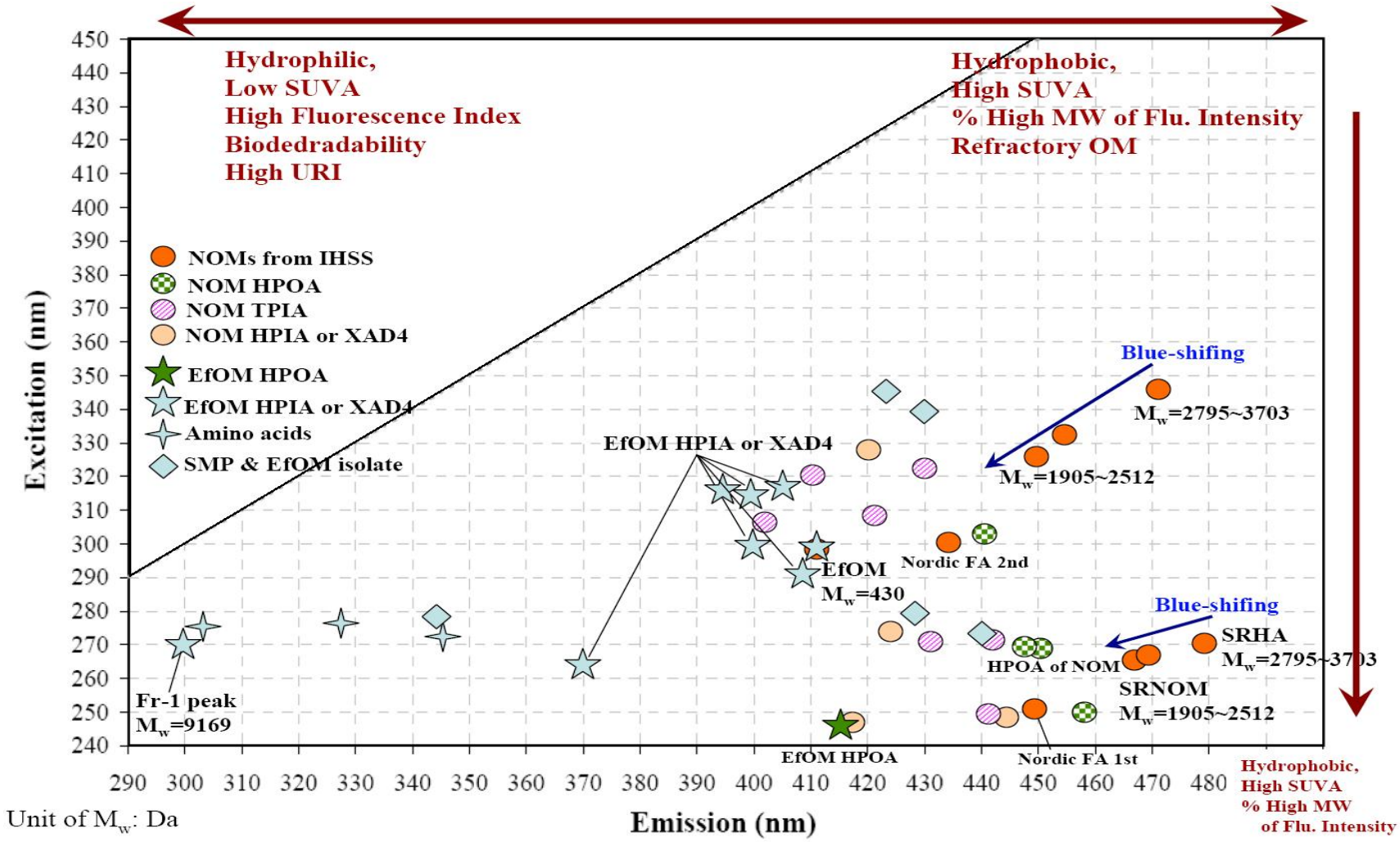


Figure 5-11 Peak maxima locations and MWs for different organic samples

Summary and conclusions

The followings conclusions are supported by the experimental results of this study.

1. In general, characterization of EfOM exhibits a low SUVA and increased HPI fraction in comparison with NOM. By fluorescence EEM, EfOM is demonstrated to have a higher FI than NOM.

2. SEC of EfOM showed a significant amount of high MW substances (polysaccharides for EfOM, and protein-like materials for SMPs) with weak UV absorbance for EfOM.

3. SEC-SUVA and SEC-URI showed lower SUVA values and higher URI for high MW substances and higher SUVA and lower URI for HS. Also, it was found that URI calculation showed differences by absorbance and by electrical signal at 210 and 254 nm, and URI could be used a surrogate indicator for the presence of organic nitrogen when a DON detector is not available.

4. SEC-fluorescence chromatograms indicate that low MW fluorophores had higher fluorescence intensity and fluorescence-sensitive moieties at certain MW were distributed more narrowly than substances responsible for DOC, indicating M_w by fluorescence chromatograms may provide more accurate information on fluorophores. Furthermore, at given excitation and emission wavelengths, higher MWs are more responsible for fluorescence intensity.

5. For NOM, MW of fluorophores was likely to be related to shifting of excitation and emission wavelengths: the higher MWs, the longer fluorescence wavelengths.

A Case Study of Effluent-dominant stream: the Passaic River Watershed, New Jersey[2]

Water sources

Wastewater effluents were collected from the Two Bridges Sewerage Authority's WWTP with an activated sludge process as a secondary treatment, and filtration as a tertiary treatment. Treated wastewater is discharged into the Pompton River. An upstream sample was taken from a location above the WWTP outfall site on the Pompton River. The Pompton River, impacted by wastewater effluent, is tributary to the Passaic River, which flows ~ 8.4 km until being used as influent by the Little Fall DWTP. Sampling was done twice in 2004 (February and August), and once more in 2005 (August).

Characteristics of NOM and EfOM

Table 5-6 summarizes trends showing the characteristics of NOM, EfOM and EfOM-impacted waters in the Northeast USA watershed. Compared to UPST (upstream NOM sample), WWTP (EfOM sample) exhibited relatively low SUVA values (1.63-2.13 L/mg-m, average: 1.90 L/mg-m), increased fraction of hydrophilic organic matter (32.5-38.6%, average: 34.8%), and higher FIs (1.397-1.433, average: 1.413). The low SUVA of wastewater implies that the DOC of EfOM is comprised of more non-aromatic (or less aromatic) organic carbon than NOM.

Organic matter derived from allochthonous (terrestrial plant origin) sources has lower FI values while organic matter derived from autochthonous (microbially derived, such as algae and

---

[2] A part of book chapter publication:

Seong-Nam Nam, Stuart W. Krasner, Gary L. Amy, (2007). Differentiating Effluent Organic Matter (EfOM) from Natural Organic Matter (NOM): Impact of EfOM on Drinking Water Sources, *Advanced Environmental Monitoring*, Chapter 20, p259-270, Springer-Verlag GmbH, Germany

bacteria) sources shows higher FI values. As a point of reference, FIs of the Suwannee River humic and fulvic acids, and pure EfOM (which was generated from a laboratory bench-scale bioreactor and isolated by an electrodialysis/reverse osmosis desalting process) were determined to be 0.864, 0.969, and 1.459, respectively. The FIs of EfOM were higher than those of NOM, which means that the properties of EfOM that distinguishes it from NOM are mainly microbial–in origin. The same trends of FIs were also observed in other wastewaters in this study (results not shown).

Table 5-6 Characterization results of upstream water, wastewater and downstream water

Sample ID	Sample Type	Sampled date	UVA_{254}, cm^{-1}	DOC, mg/L	SUVA, L/mg-m	HPO %	TPI %	HPI %	FI
UPST	upstream of WWTP	Feb, 2004	0.074	2.58	2.87	62.6	22.5	14.8	1.207
		Aug, 2004	0.149	5.22	2.85	44.9	27.9	27.2	1.160
		Aug, 2005	0.077	3.01	2.56	48.8	17.7	33.5	1.319
		Average	**0.100**	**3.60**	**2.76**	**52.1**	**22.7**	**25.2**	**1.229**
WWTP	WW effluent	Feb, 2004	0.124	7.61	1.63	39.8	21.6	38.6	1.422
		Aug, 2004	0.115	5.40	2.13	43.6	23.9	32.5	1.397
		Aug, 2005	0.120	6.17	1.94	40.5	26.2	33.2	1.433
		Average	**0.120**	**6.39**	**1.90**	**41.3**	**23.9**	**34.8**	**1.418**
INFL	downstream of WWTP	Feb, 2004	0.088	3.10	2.84	65.2	20.2	14.7	1.173
		Aug, 2004	0.158	5.12	3.09	51.9	22.3	25.8	1.150
		Aug, 2005	0.119	4.72	2.52	51.5	20.7	27.9	1.459
		Average	**0.122**	**4.31**	**2.82**	**56.2**	**21.1**	**22.8**	**1.261**

Fluorescence EEM measurements for upstream water, wastewater, and downstream of WWTP (influent of drinking water treatment plant) clearly showed different peaks between NOM and EfOM (Figure 5-11). Previous researchers working with fluorescence spectroscopy identified the tryptophan-like, protein-like peaks at Ex/Em = 275/340 nm and humic-like peaks at 260/380-460 nm and 350/420-480 nm [19, 102]. EEMs of wastewater and wastewater-impacted water exhibited the presence of protein-like substances at the range of excitation wavelength 260 to 290 nm and emission wavelength 320-370 nm, which were at similar location with other studies, and this protein-like peak may have originated from soluble microbial

products present in biologically treated wastewater.

In general, the SEC-DOC of surface waters and wastewaters is separated into three main peaks according to their molecular-weight distributions, as shown in Figure 5-13: the zone 1 peak represents large molecules, such as polysaccharides, proteins, and colloids, the zone 2 peak is attributed to humic substances, and the zone 3 peak corresponds to low-molecular-weight organic acids. Corresponding UVA_{254} chromatograms showed low absorbance in the zone 1 peak, and greatly increased absorbance in zone 2, with intermediate absorbance in the zone 3 peak, which indicates that macromolecular constituents in the zone 1 peak do not consist of aromatic rings or conjugated bonds, but saturated structures or linear ring systems that may be more labile and hydrophilic, although they are macromolecules. On the other hand, components (humic substances) in the zone 2 peak are likely to be hydrophobic and microbially resistant in spite of somewhat smaller molecules (relative to zone 1).

The SEC-DOC fractions for each peak from the SEC are summarized in Table 5-6. The percentage for each peak was determined by the ratio of integrated peak area to total peak area, and each percentage was converted to DOC (relative to the bulk water DOC). As can be seen in Table 5-7, EfOM had a higher percentage of polysaccharides, indicating a higher contribution to the DOC of the EfOM by microbial sources. In addition, the EfOM had a lower percentage of humic substances and a higher percentage of low-molecular-weight organic acids relative to the NOM.

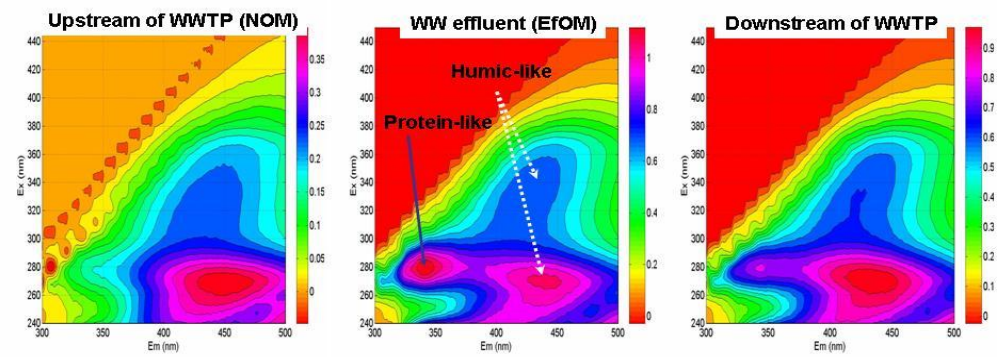


Figure 5-12 EEM contour maps of samples collected upstream of WWTP (left), wastewater effluent (middle), and downstream of WWTP (right) from a Northeast USA watershed. (x-axis: emission wavelengths of 290-500 nm, y-axis: excitation wavelengths of 240-450 nm)

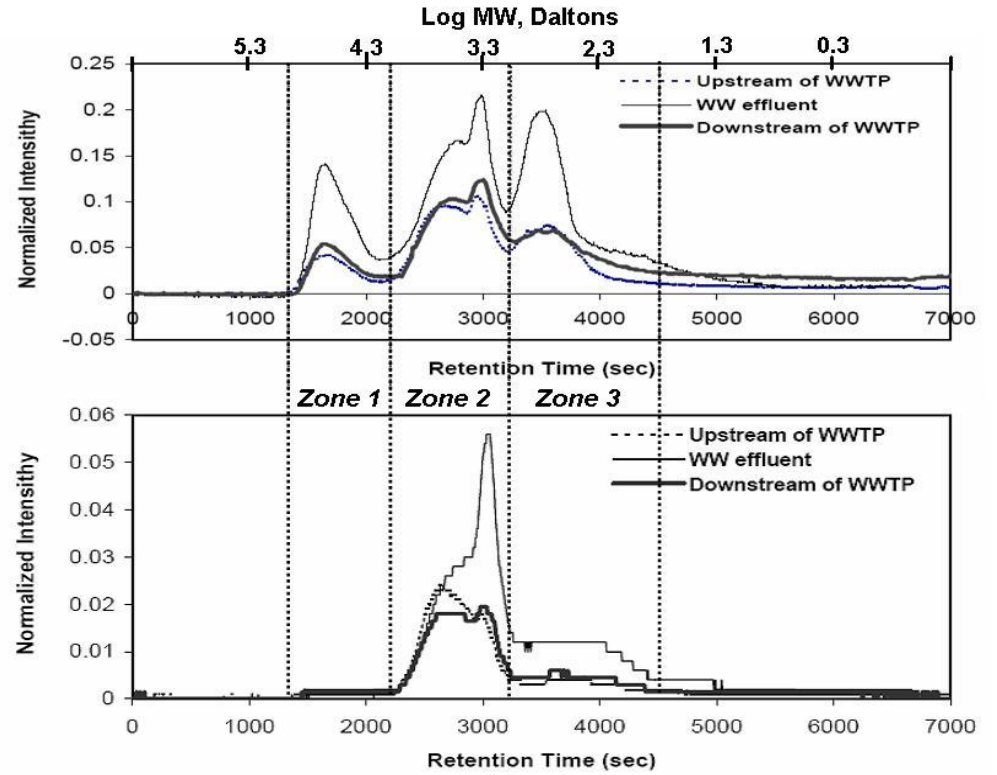


Figure 5-13 SEC of DOC fractions (upper) and UVA_{254} fractions (lower) for upstream water, wastewater, and downstream water from a Northeast USA watershed, and their molecular-weight (MW) distributions (UVA_{254} results for upstream water obscured by downstream water results)

Figure 5-14 shows the results of total dissolved carbohydrates measurements using the phenol-sulfuric acid method [103]. Higher levels of polysaccharides were present in EfOM, which was consistent with SEC-DOC data. Polysaccharides are one of the components of

bacterial cell walls and are released during the endogenous cell decay. Therefore, the primary characteristics of EfOM mostly reflect microbial sources.

Table 5-7 Characterization results of upstream water, wastewater, and downstream water from a Northeast USA watershed (SEC-DOC data from samples taken in February 2004) (PS, HS, and LMA denote polysaccharides, humic substances, low-molecular-weight organic acids)

Samples	DOC (mg/L)	%			as DOC (mg/L)		
		PS	HS	LMA	PS	HS	LMA
UPST (Upstream water)	2.58	10.4	46.0	43.5	0.27	1.19	1.12
WWTP (WW effluent)	7.61	7.7	57.7	34.6	0.58	4.39	2.63
INFL (Downstream water)	3.10	7.0	39.1	53.9	0.22	1.21	1.67

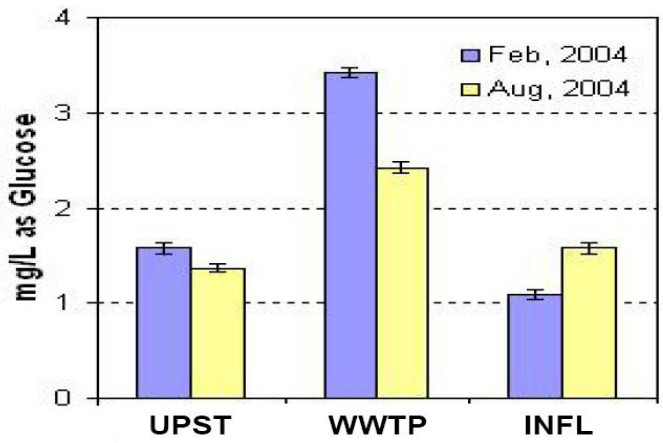


Figure 5-14 Total dissolved carbohydrates (polysaccharides) for upstream water (UPST), wastewater (WWTP), and downstream water (INFL) (n=3)

CHAPTER 6

ASSESSMENT OF CHARACTERISTICS OF NATURAL ORGANIC MATTER AND EFFLUENT ORGANIC MATTER USING MULTIVARIATE STATISTICAL ANALYSES

Introduction

Along with the scarcity of available drinking water sources and increase of wastewater reuse as a supplementary water source, appropriate water quality monitoring has become more of a concern. Especially, the discharge of municipal wastewater effluent to surface waters such as streams and lakes, and the subsequent use as a drinking water source gives rise to an increased importance of understanding reliable water quality information. Generally, the processes and technologies used in drinking water treatment are focused on removal of organic constituents (e.g., TOC or DOC). The choice of which treatment to use among various available processes is primarily dependent on the characteristics of organic matter in the water in addition to the cost of treatment or the specific water quality problems likely to be present. A priority of treatment should be the removal efficiency of organic matter, which is also related to the public health risks such as disinfection by-products (DBPs) formation or biostability in the distribution system. For example, it has been reported that organic matter impacted by wastewater showed less removal efficiency by conventional drinking water treatment due to the different properties of organic matter [104], and the occurrences and levels of DBP precursor materials and/or formation potentials were greatly influenced by characteristics of organic matter [32, 105, 106].

Multivariate statistical techniques have been applied to explore and evaluate the relationships between water quality and many other environmental factors [107-116]. Multivariate statistical approach, such as correlation analysis, principal component analysis (PCA) and multiple linear regression (MLR) analysis, is particularly advantageous when data come from a wide range of sources and comprise various forms which do not lend themselves to a simple theoretical treatment. However, using multivariate statistical analyses there have been an increasing number of research studies which have compared or assessed water qualities in diverse catchments affected by different land uses, storm events, human activities, etc.; most of these studies have been performed focusing on inorganic species (e.g., trace elements) and/or parameters determining wastewater quality (e.g., dissolved oxygen (DO), total suspended solids (TSSs) and biochemical/chemical oxygen demand (BOD and COD)) [113, 117, 118]. Several studies have dealt with fingerprinting of organic species in wastewater impacted streams, but these have been largely confined to probing chemical fragments using pyrolysis-GC/MS [107].

There have been few multidimensional datasets, which cover a variety of sampling sites including water treatment plants (DWTPs, WWTPs and WTPs) and rivers (i.e., watersheds scale), and in which diverse information on NOM and/or EfOM properties is implicitly contained. Therefore, this study was performed to analyze such as the dataset using multivariate statistical techniques and to understand/expand knowledge on the following assumptions: 1) wastewater EfOM has different characteristics from drinking water NOM, characteristics of which can be governed by one or several major discriminating or distinguishable factors, so that waters impacted (lightly or heavily) by wastewater EfOM may be classified differently from waters which are not (i.e., NOM); 2) characteristic similarities and discrepancies between NOM and EfOM can be identified by different analytical techniques. Specifically, the aim of this study is to

provide survey results of water quality parameters in terms of characteristics of NOM and EfOM over ~25 WTPs including drinking water treatment plants (DWTPs), wastewater reclamation plant (WRPs), wastewater treatment plants (WWTPs), and to apply multivariate statistical analyses (e.g., correlation analysis, principal component analysis, and multiple linear regression analysis) techniques in order to assess similarities and differences between NOM and EfOM, and to identify key parameters (or factors) providing the most information for differentiation of EfOM from NOM.

Materials and Methods

Sampling campaigns

Waters from U.S. DWTPs and WWTPs and rivers in 8 states were received from 21 WTPs (3 in Arizona, 3 in California, 3 in Colorado, 3 in Michigan, 2 in New Jersey, 2 in Ohio, 3 in Texas and 2 in Pennsylvania), and 5 rivers (1 in Arizona, 1 in California, 1 in Colorado, 1 in Michigan and 1 in New Jersey) for study periods from February 2004 to August 2005, bringing a total of 122 sampling points into the dataset.

Physicochemical parameters

For the selected sampling locations, nine (9) water quality parameters (e.g., electrical conductance, alkalinity, total ammonia, nitrite, and nitrate, DOC, UV_{254}, etc.) were measured according to the standard methods [15]. Although inorganic species are not of interest and out of major scope of this study which deals with organic matter, these were for the added purpose of showing impacts of wastewater discharge, and were not used for EfOM discriminating purposes in further statistical analyses such as PCA and MLR. Advanced NOM characterization

parameters were also analyzed using appropriate procedures and methods, and those parameters include HPO-, TPI- and HPI-fractions by XAD-8/-4 fractionation [41], fluorescence index (FI), humidification index (HIX), Phenol/Aromatic ratio (Phenol/Arom), C1 to C5 components by fluorescence excitation-emission matrix (EEM) [37, 43], amount of polysaccharide/humic substances/low molecular weight acid by high performance size-exclusion chromatography with DOC detection (HPSEC-DOC) [2, 42], biodegradability as 5-day biodegradable dissolved organic carbon ($BDOC_5$) [119], molar absorptivity at 280 nm (ε280) and dissolved organic nitrogen (DON) [13]. The resulted numbers of variables related to NOM/EfOM characteristics were nineteen (19) shown in Table 6-1, and detailed explanations on each procedure can be found in Chapter 2 (Experimental methods).

Dataset preparation and statistical analyses

In this study, 122 monitoring locations and 18 parameters explaining NOM characteristics among a total of 24 physicochemical parameters (Table 6-1) were used for statistical analyses. Data were *binned* into three distinct datasets based on the extent of EfOM content, i.e., group 1-upstream data of WWTPs which are not impacted by wastewater and are considered to be dominated by NOM; group 2-downstream data of WWTPs which are lightly or heavily impacted by wastewater; and group 3-WWTP effluent data with EfOM abundance. For the verification of assumption one, principal component analysis were performed for the separate datasets. For the second assumption, based on the results of box-and-whisker plots, multiple linear regression analysis was applied to assess how the relationships among properties of organic matter vary depending on water quality. All statistical analyses applied were performed using SPSS 15.0® for Windows (SPSS Inc.).

Prior to further statistical analyses, the normality of the frequency distribution was checked, and if the skewness and kurtosis is less than 2.5 in absolute values (which is approximately the p=0.01 level), the data are not significantly different from normality. Since water quality parameters and/or properties of organic matter are presented according to different scales (units) and wide ranges, it is necessary to perform the standardization. Otherwise, PCA or MLR may be influenced by the most intense variables. Thus, in this study, logarithmic scaling standardization was applied for some variables which showed high variations and skewness. Also, normality for variables was often violated because the overall database includes upstream, WWTP, and downstream data together, thus, standardization was applied for each separate dataset. Samples from mixed liquor tanks of WWTPs showed extremely high DOC and high hydrophilicity, and appeared to be outliers through the pre-checking steps; thus they were removed from the dataset because, for example, samples with a large distance to other points in the measurement space can pull the principal components toward it, but away from the direction of maximum variance [120]. Bivariate correlation analysis was performed to examine relationships between pairs of variables. Pearson correlation coefficients representing values between +1 and -1, with negative values inferring an inverse correlation, were considered to judge statistical significances of parameters at a 95% confidence level ($p<0.05$).

Principal component analysis (PCA) is one of the multivariate statistical analyses. The central idea of PCA is to reduce the dimensionality of a dataset consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the dataset [71]. The PCA is expressed by a least-square model (equation 6-1)

$$X = A \cdot F + E \tag{6-1}$$

X is the original data matrix, matrix A corresponding to factor loadings, matrix F is corresponds

to factor scores and matrix E contains the residual or error. The principal components (PCs) account for the maximum variance explaining all properties of original parameters. The first PC (PC1) explains most of the data variance, followed by PC2 which describes the next largest amount of data variance and is orthogonal to PC1. Such computations are repeated until the number of PCs becomes equal to the number of the original variables and PCs are orthogonal to each other, i.e., they are not correlated. But most of the data variance is usually contained in the first few PCs, thus a reduction of data dimensionality is possible. Hence PCs can be expressed as follows (equation 6-2):

$$PC_{jk} = \alpha_{j1}\chi_{k1} + \alpha_{j2}\chi_{k2} + ... + \alpha_{jm}\chi_{kn} \quad (6\text{-}2)$$

Where PC_{jk} is the value for principal component j for object k, α_{j1} is the loading of variable 1 on component j, χ_{k1} is the measurement value for variable 1 on object k and n is the total number of variables.

In this study, PCA involves the following major steps: (1) to compile the dataset without missing values or with missing values handling. Missing values handling were performed by inserting dummy values which were close to the mean of variables; (2) to calculate the covariance matrix or the correlation matrix of the original dataset (in cases where the units of variables are not uniform in scale, a correlation matrix would be appropriate [121]); (3) to find the eigenvalues and the corresponding eigenvectors of the dataset; (4) to establish the factor loading matrix and to perform the varimax rotation on the factor loading matrix in order to extract the principal components. In this step, the extraction criteria applied was that any components exhibiting an eigenvalue greater than one were retained [71, 121].

Multiple linear regression (MLR) analysis attempts to model the relationship between a dependent variable (predicted values) and two or more independent variables (predictors) by

fitting a linear equation to observed data. MLR is based on least-squares and the best-fitting line for the observed data is calculated by minimizing the sum-of-squares of differences of observed and predicted values. The MLR model can be denoted as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k + E \quad (6\text{-}3)$$

Where Y is the dependent variable, β_0 is the intercept, β_1 is the slope coefficient for the first explanatory variable, β_1 is the slope of coefficient for the second explanatory variable, k is the slope coefficient for the k^{th} explanatory variable, and E is the error (the remaining unexplained noise in the data).

Table 6-1 The water quality parameters associated with their abbreviations and units used in the study

Parameters*	Abbreviation	Unit
Organic parameters		
5-day aerobic BDOC fraction	$BDOC_5$	%
Carbohydrate concentration	Polysaccharide	mg/L as glucose
Chemical oxidation demand	COD	mg/L
Component 1, ... , Component 5 from PARAFAC	C1, C2, ... , C5	%
Dissolved organic carbon	DOC	mg/L
Dissolved organic nitrogen	DON	mg/L
Fluorescence index	FI	-
Humic substance peak in SEC-DOC	HS	%
Humidification index	HIX	-
Hydrophilic fraction	HPI	%
Hydrophobic fraction	HPO	%
Low molecular acid peak in SEC-DOC	LMA	%
Molar absorptivity at 280 nm	ε280	L/mole-cm
Phenol/Aromatic	Ph/Ar	-
Polysaccharide peak in SEC-DOC	PS	%
Specific UV absorbance at 254 nm	SUVA	L/mg-m
Transphilic fraction	TPI	%
UV absorbance at 254 nm	UV_{254}	cm^{-1}
Inorganic parameters**		
Alkalinity	Alk	mg/L as $CaCO_3$
Bromide	Br^-	mg/L
Dissolved inorganic nitrogen	DIN	mg/L as N
Electrical conductance	EC	μOhm/cm
Total hardness	TH	mg/L
Total Iodide	TI	mg/L

*Alphabetically ordered, **excluded in PCA and MLR.

Results and discussion

Table 6-2 summarizes briefly the minimum, maximum, mean values and standard deviation of the 28 measured variables for the three datasets. High variations (standard deviation), especially in wastewater-related samples, indicate the presence of variations in chemical composition among samples, possibly caused by diverse polluting sources, degree of treatment depending on the plants, and other environmental factors. High levels of inorganic and organic contents in downstream samples showed that the impact of wastewater discharge into rivers or streams causes not only increases of nutrients in pertinent locations, but also changes of physicochemical properties of organic matters. Increased inorganic species such as bromide and iodide may originate from major components of detergents and pharmaceuticals, and natural sources. Nitrogenous species (DIN and DON) are thought to be originated from breakdown and excretion products of nitrogen-containing compounds (e.g., proteins and urea), and residuals of biological (nitrification) treatment processes. In addition, high levels of several factors (e.g., polysaccharides, C2 and C4) that are closely related with microbial sources, and indices characterizing organic matter origin such as FI and Phenol/Arom help to explain that discharged wastewater influences domination of organic matter types in samples. Higher fractions of hydrophilic (HPI) or polysaccharides organic substances in wastewater and downstream samples were in contrast to upstream waters of WWTPs which showed higher fractions of hydrophobic or humic substances.

Table 6-2 Statistical descriptives of the analyzed parameters

	Upstream sites of WWTP (Group 1)					WW-impacted downstream sites (Group 2)					WWTP sites (Group 3)				
Parameters*	N	Min.	Max.	Mean	S.D.	n	Min.	Max.	Mean	S.D.	n	Min.	Max.	Mean	S.D.
Alk (mg/L as $CaCO_3$)	13	37	146	88	35	22	37	308	148	67	33	23	264	155	65
Br^- (mg/L)	7	0.02	0.21	0.06	0.07	17	0.02	0.97	0.25	0.28	13	0.08	0.42	0.20	0.08
C1 (%)	20	15.0	42.2	29.5	7.6	44	6.4	49.1	23.8	8.1	54	14.7	37.8	24.6	4.7
C2 (%)	20	2.4	16.8	10.5	3.3	44	0.0	28.2	17.9	7.0	54	0.0	26.1	18.6	6.1
C3 (%)	20	0.0	17.5	14.3	3.7	44	6.6	22.2	15.0	2.4	54	0.0	28.2	16.6	4.4
C4 (%)	20	6.4	26.2	14.5	5.1	44	6.8	34.0	21.6	9.1	54	6.3	71.7	24.1	11.2
C5 (%)	20	23.0	41.2	31.1	5.2	44	4.9	44.8	21.7	10.9	54	0.0	36.1	16.0	6.4
COD (mg/L)	17	1.0	24.0	7.7	5.7	41	0.5	36.7	18.5	12.0	45	3.0	194.7	28.7	27.1
DIN (mg/L as N)	16	0.01	3.31	0.97	0.98	37	0.41	36.19	9.47	10.79	44	1.32	46.47	16.14	10.59
DOC (mg/L)	20	1.77	7.52	3.21	1.37	46	0.51	14.30	6.04	3.99	54	0.36	71.81	9.54	9.45
DON (mg/L as N)	17	0.13	0.67	0.29	0.14	42	0.00	1.47	0.57	0.39	41	0.11	2.40	0.83	0.56
ε280 (L/mole-cm)	19	43	236	134	60	35	18	346	141	66	52	13	756	165	119
EC (μohm/cm)	13	216	1098	502	278	36	224	1191	821	242	33	712	2538	1204	406
Fluorescence index	20	1.16	1.46	1.27	0.08	46	0.90	1.71	1.39	0.15	54	1.18	1.74	1.52	0.11
$BDOC_5$ (%)	2	19.5	19.5	22.4	21.0	19	22.4	65.6	37.3	10.0	18	34.0	52.9	28.4	16.0
HPI (%)	20	0.15	0.46	0.33	0.08	46	0.15	0.59	0.38	0.10	52	0.23	0.69	0.41	0.09
HPO (%)	20	0.34	0.63	0.47	0.07	46	0.25	0.65	0.42	0.08	52	0.22	0.57	0.39	0.07
HS (%)	19	0.42	0.90	0.62	0.14	45	0.07	0.90	0.51	0.17	52	0.23	0.89	0.54	0.16
LMA (%)	19	0.08	0.48	0.29	0.13	45	0.00	0.90	0.33	0.17	52	0.00	0.63	0.31	0.15
PS (%)	19	0.02	0.17	0.08	0.05	45	0.03	0.35	0.17	0.09	52	0.00	0.33	0.14	0.08
TPI (%)	20	0.13	0.35	0.20	0.05	46	0.09	0.33	0.20	0.05	52	0.09	0.36	0.21	0.05
HIX	20	0.07	0.27	0.13	0.07	46	0.05	0.55	0.22	0.11	54	0.06	0.47	0.19	0.07
MOC	5	-2.41	1.75	-0.05	2.06	25	-2.38	1.90	0.10	1.55	40	-1.32	1.86	0.61	0.96
Phenol/Aromatic	20	3.73	10.05	7.19	1.49	46	2.56	16.13	9.40	2.66	54	6.42	19.85	9.75	2.35
Polysaccharide (mg/L as Glucose)	11	0.14	4.46	1.50	1.44	37	0.00	9.25	3.33	2.51	39	0.93	16.98	4.61	3.59
SUVA (L/mg-m)	20	0.88	2.84	1.85	0.61	46	0.79	3.29	1.80	0.52	54	0.47	3.97	1.75	0.53
TH (mg/L)	13	51	281	142	67	22	52	304	196	72	33	136	440	220	77.49
TI (mg/L)	10	0.0	22.0	9.0	7.7	16	0.0	42.0	17.4	13.7	20	18.0	91.0	49.4	17.39
UV_{254} (cm^{-1})	20	0.018	0.199	0.062	0.04	46	0.004	0.391	0.107	0.07	54	0.004	0.344	0.142	0.07

*Alphabetically ordered

Differentiation between EfOM versus NOM

***Exploratory data analysis by box plots*:** In order to identify appropriate components (variables) differentiating EfOM from NOM, PCA was used. Prior to doing the statistical analyses, box plots (box-and-whisker plots) of individual variables in the three dataset (i.e., group 1; upstream sites, group-2; wastewater-impacted downstream sites of WWTPs, group-3; wastewater effluent) were examined. The box-and-whisker plot is a very good technique to compare different group

variables or objects in relation to others. Every group is represented as a box in a diagram, and the center line of the box represents the group mean. The box represents the middle 50% of the cases between the 25^{th} and 75^{th} percentiles. The bar in the box represents the median value. The whiskers indicate the expected range of scores. Scores outside of this range are considered unusually high or low. Such scores, called outliers, are shown above and/or below the whiskers with asterisks and for very extreme scores are shown with circles in Figure 6-1.

Figure 6-1 shows several examples of box plots for variables related to the properties of organic matter in each bin. Three datasets exhibit clear distributions, for example, water quality parameters representing organic matter loadings (DOC, COD and DON) show higher concentrations of organics in WWTP effluents and WW-impacted downstream samples compared to upstream samples, indicating that WWTP effluents contain more organic nutrients and consequently contributed increased organic loadings to downstream sites. As parameters that can differentiate characteristics of EfOM from NOM, FI, fraction of HPO or HPI, fraction of polysaccharide substances, and the ratio of phenol and aromatic content (Phenol/Aromatic) were compared. Box plots provide information (e.g., shape of the distribution) on underlying distributions of parameters. For examples, box plots with long whiskers such as DON of WWTP effluents (bin "3") and WW-impacted downstream samples (bin "2") indicate that their distributions are skewed to high concentration; also, the DOC of WWTP effluents distribution is more resembles a normal distribution (i.e., less skewed) than DON of WWTP effluents. Overall, by inspecting these plots among the three datasets, box plots of WWTP effluents and WW-impacted downstream samples, compared to upstream samples, showed more skewed, but less kurtosised distributions (i.e., more broad through low to high concentration, as also shown in Table 6-2), indicating that the characteristics of organic matter in WWTP effluents and WW-

impacted downstream samples may be a conglomerate explained by diverse sources and by in-stream reactions. Moreover, water qualities or organic matter properties of WW-impacted downstream samples are likely deteriorated closely to WWTP effluents, apart from upstream samples. Box plots showing higher FI, HPI (%), polysaccharides and Phenol/Arom ratio in WWTP effluents and WW-impacted samples, but less hydrophobicity in the same sample groups suggest that organic matter constituting organic nutrients (e.g., DOC or DON) are also possibly related to hydrophilic and/or microbially-originated substances.

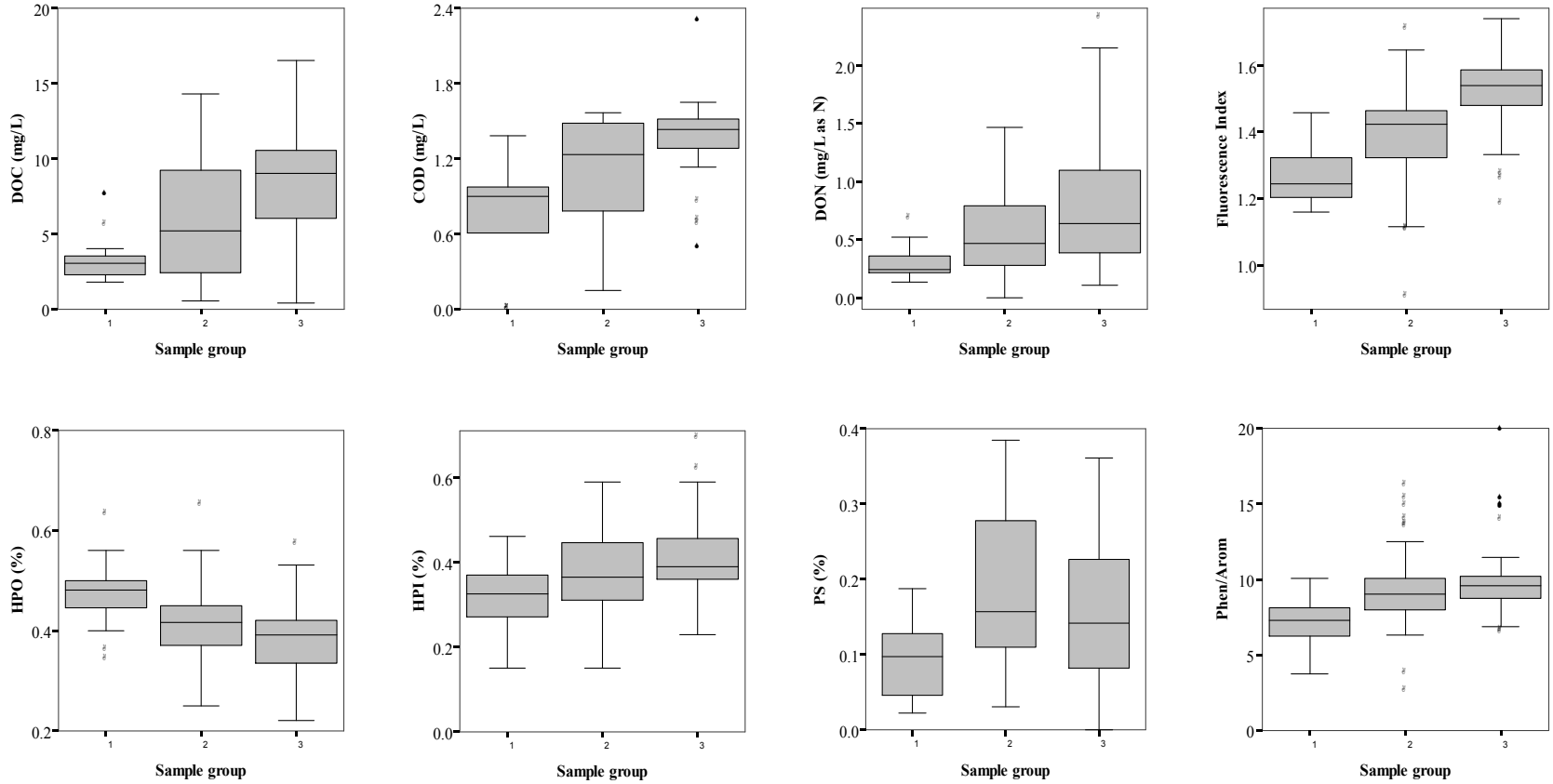


Figure 6-1 Box plots of parameters for organic pollution loadings (DOC, COD and DON) and for organic characteristics (FI, HPO, HPI, PS and Phenol/Aromatic); 1=Upstream (n=20), 2=WW-impacted (n=46), and 3=WWTP effluents (n=54)

Figure 6-2 shows box plots of five components identified by PARAFAC analysis. As determined in Chapter 4, component 1 and component 5, which are humic-like and allochthonous fluorophores, show higher presence in upstream samples, while component 2 and component 4, which are thought to be microbially-derived and autochthonous fluorophores exhibited a higher presence in WWTP effluents and WW-impacted downstream sites. Component 3 did not show significant differences among the three datasets. Here, one thing that should be pointed out is that component 2 and component 4 shows broader distributions than WWTP effluents, implying that amounts of these two components may be easily influenced by the environmental and/or geospatial conditions of downstream locations. That is, although component 2 and 4 can be easily (bio-) degradable (as shown in Chapters 3 and 7) during the course of natural purification in the downstream sites (streams or rivers, etc), insufficient travel times for degradation may result in higher amount of these components in downstream locations; also, for example, if the watersheds experiences inputs of fertilizers, rainfall events, stormwater runoff or algal blooming, etc., dominance of these components may vary. As Figure 6-2 shows, bin 2 (WW-impacted downstream samples) displays broad distributions for most components.

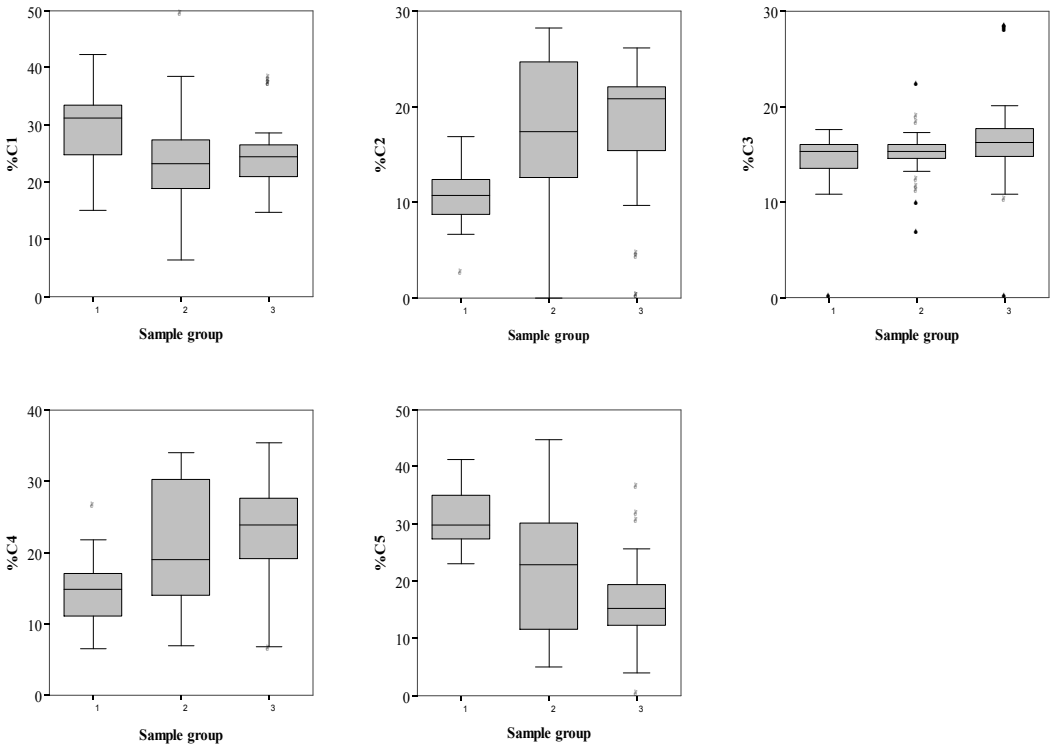


Figure 6-2 Box plots of PARAFAC components (component 1 through component 5); 1=Upstream (n=20), 2=WW-impacted (n=46), and 3=WWTP effluents (n=54)

Table 6-3 Correlation matrices-Upstream of WWTP (NOM dominant)

	Alk	TH	ε280	EC	Tot I	COD	DIN	DON	HIX	Ph/Ar	DOC	HPO	TPI	HPI	PS	HS	LMA	Br-	UV_{254}	SUVA	PolyS	FI	%C1	%C2	%C3	%C4	%C5
Alk	1																										
TH	**0.953**	1																									
ε280	-0.337	-0.016	1																								
EC	0.538	**0.752**	0.498	1																							
Tot I	**0.803**	**0.758**	-0.331	0.180	1																						
COD	0.412	0.548	0.396	**0.598**	0.614	1																					
DIN	0.150	0.249	0.006	0.508	-0.032	0.172	1																				
DON	0.479	0.481	-0.030	0.386	0.479	**0.757**	0.425	1																			
HIX	-0.049	-0.129	-0.363	-0.291	0.468	-0.091	-0.039	-0.066	1																		
Ph/Ar	0.200	0.412	0.537	**0.692**	-0.111	**0.842**	0.195	0.408	0.007	1																	
DOC	0.006	0.106	0.498	0.345	-0.117	**0.717**	0.276	**0.745**	-0.397	0.562	1																
HPO	-0.229	0.021	**0.690**	**0.585**	-0.686	0.242	0.261	-0.047	-0.484	0.418	0.257	1															
TPI	-0.121	-0.245	-0.358	-0.297	-0.354	-0.420	0.206	0.047	-0.430	**-0.668**	-0.060	0.080	1														
HPI	0.227	0.114	-0.309	-0.273	0.658	0.058	-0.324	-0.005	**0.646**	0.078	-0.173	**-0.797**	**-0.664**	1													
PS	**0.584**	**0.635**	-0.195	0.496	0.219	0.550	0.506	0.576	-0.068	0.372	0.246	0.013	-0.066	0.011	1												
HS	-0.307	-0.304	0.401	-0.309	0.452	0.196	-0.190	0.072	-0.099	0.076	0.384	-0.193	-0.143	0.231	-0.286	1											
LMA	0.147	0.153	-0.442	0.264	-0.387	-0.345	0.503	-0.170	0.214	-0.137	-0.432	0.169	0.185	-0.232	0.279	**-0.783**	1										
Br-	**0.810**	**0.900**	0.178	**0.863**	**0.902**	**0.699**	0.563	**0.733**	-0.219	0.570	0.457	0.278	-0.285	-0.145	0.647	-0.251	0.153	1									
UV_{254}	-0.396	-0.315	**0.741**	0.046	-0.406	0.420	0.196	0.481	-0.377	0.352	**0.904**	0.284	0.089	-0.275	-0.059	0.483	-0.443	0.083	1								
SUVA	-0.553	-0.327	**0.946**	0.265	**-0.771**	0.303	0.051	-0.022	-0.402	0.507	0.563	**0.709**	-0.169	-0.431	-0.306	0.365	-0.372	-0.049	0.698	1							
PolyS	**0.638**	**0.694**	-0.102	0.621	0.233	0.519	0.051	0.126	0.140	**0.643**	0.007	0.041	-0.601	0.431	**0.646**	-0.231	0.262	**0.800**	-0.295	-0.262	1						
FI	0.382	0.543	0.546	**0.631**	0.622	0.533	0.334	0.443	0.105	0.429	0.260	0.232	-0.412	0.082	0.297	0.024	-0.092	**0.940**	-0.004	0.138	0.147	1					
%C1	-0.407	-0.393	-0.041	-0.138	**-0.860**	-0.461	0.099	-0.374	**-0.666**	-0.355	0.013	0.341	0.552	**-0.599**	-0.112	-0.030	0.207	-0.314	0.243	0.230	-0.221	-0.566	1				
%C2	0.234	0.360	0.278	0.479	0.474	**0.795**	0.362	**0.715**	0.305	**0.721**	0.570	0.098	-0.506	0.231	0.538	-0.021	-0.067	0.594	0.332	0.189	0.317	**0.651**	**-0.614**	1			
%C3	-0.049	0.159	**0.685**	0.497	**-0.777**	**0.825**	0.098	0.395	-0.071	**0.925**	**0.673**	0.462	-0.540	-0.029	0.167	0.345	-0.401	0.361	0.563	**0.705**	0.368	0.357	-0.293	**0.643**	1		
%C4	0.144	-0.017	-0.381	-0.360	0.672	-0.213	-0.146	0.047	0.706	-0.339	-0.419	-0.548	0.003	0.430	-0.249	-0.147	0.021	-0.184	-0.436	-0.501	-0.534	0.135	-0.705	0.086	-0.339	1	
%C5	0.280	0.156	-0.439	-0.209	0.396	-0.566	-0.331	-0.519	-0.193	**-0.597**	**-0.628**	-0.318	0.180	0.126	-0.076	-0.121	0.174	-0.161	**-0.636**	-0.574	0.151	-0.265	0.349	**-0.725**	**-0.717**	-0.086	1

Table 6-4 Correlation matrices-Downstream of WWTP (EfOM lightly or heavily impacted)

	Alk	TH	ε280	EC	Tot I	COD	DIN	DON	HIX	Ph/Ar	DOC	HPO	TPI	HPI	PS	HS	LMA	Br-	UV_{254}	SUVA	$BDOC_5$	PolyS	FI	%C1	%C2	%C3	%C4	%C5
Alk	1																											
TH	**0.693**	1																										
ε280	-0.007	-0.230	1																									
EC	**0.665**	**0.883**	0.045	1																								
Tot I	**0.672**	**0.878**	**-0.523**	**0.640**	1																							
COD	0.419	0.084	-0.050	0.084	**0.627**	1																						
DIN	**0.639**	**0.512**	0.347	**0.513**	**0.515**	**0.536**	1																					
DON	0.050	-0.070	0.107	-0.049	0.418	**0.594**	**0.518**	1																				
HIX	0.379	0.280	-0.490	0.298	0.080	-0.031	0.264	0.107	1																			
Ph/Ar	0.256	0.482	0.274	**0.590**	0.371	0.243	0.304	-0.052	0.084	1																		
DOC	0.163	-0.210	0.222	-0.072	0.497	**0.515**	0.446	**0.756**	-0.249	-0.037	1																	
HPO	-0.471	-0.108	0.132	-0.081	-0.156	-0.137	-0.244	-0.153	-0.350	0.003	-0.068	1																
TPI	0.072	-0.079	0.312	0.011	-0.172	0.290	0.171	0.209	-0.045	0.209	**0.503**	-0.012	1															
HPI	0.372	0.114	-0.232	0.051	0.214	-0.016	0.134	0.063	0.320	-0.108	-0.169	**-0.882**	-0.456	1														
PS	0.392	0.050	0.206	0.126	0.113	0.395	**0.535**	0.339	0.317	0.278	0.404	-0.048	0.480	-0.192	1													
HS	-0.175	-0.057	0.279	-0.175	0.198	-0.108	-0.162	-0.030	**-0.510**	0.043	-0.055	0.392	-0.271	-0.217	-0.322	1												
LMA	0.008	0.035	-0.332	0.136	-0.259	-0.067	-0.048	-0.127	0.389	-0.172	-0.128	-0.388	0.056	0.319	-0.122	**-0.901**	1											
Br-	**0.782**	**0.654**	0.157	**0.681**	**0.864**	-0.351	0.223	-0.326	**0.586**	**0.567**	-0.355	-0.083	-0.160	0.126	0.276	-0.195	0.086	1										
UV_{254}	-0.011	-0.302	0.343	-0.086	0.198	**0.510**	**0.534**	**0.795**	-0.045	0.048	**0.893**	-0.009	0.483	-0.215	**0.526**	-0.155	-0.079	-0.308	1									
SUVA	**-0.536**	-0.325	**0.639**	-0.143	**-0.613**	-0.019	0.195	0.145	-0.269	0.271	-0.045	0.363	0.109	-0.375	-0.023	0.236	-0.237	-0.323	0.219	1								
$BDOC_5$	**0.998**	**-0.998**	-1.000	**0.650**	-0.720	0.391	**0.738**	-0.025	**0.540**	-0.194	**0.619**	-0.493	0.195	0.202	**0.589**	-0.384	0.256	1.000	**0.623**	-0.433	1							
PolyS	**0.696**	0.468	0.056	0.384	**0.859**	**0.506**	0.449	0.446	-0.007	0.139	**0.670**	-0.158	0.272	0.019	**0.505**	-0.361	0.150	-0.183	**0.700**	-0.164	-0.026	1						
FI	**0.638**	**0.732**	0.056	**0.771**	**0.658**	0.193	**0.645**	0.247	0.271	**0.612**	0.151	-0.237	0.338	0.045	0.357	-0.166	0.009	0.281	0.232	0.063	-0.052	0.456	1					
%C1	-0.381	-0.475	0.239	**-0.508**	-0.274	0.066	-0.410	-0.103	**-0.781**	**-0.505**	0.007	0.258	-0.176	-0.146	-0.274	0.297	-0.188	**-0.621**	-0.081	0.130	-0.444	-0.031	**-0.686**	1				
%C2	**0.556**	0.260	0.324	0.342	0.041	0.088	**0.684**	0.389	0.423	0.219	**0.527**	-0.372	0.349	0.172	**0.594**	-0.415	0.157	**0.807**	**0.541**	-0.146	**0.586**	0.394	0.432	**-0.501**	1			
%C3	0.096	0.221	0.374	0.356	-0.216	-0.027	0.241	0.190	-0.221	0.335	0.037	-0.249	0.011	0.238	-0.025	0.072	-0.067	-0.214	0.098	0.492	-0.234	0.096	0.353	0.003	0.114	1		
%C4	**0.628**	0.493	-0.155	**0.548**	0.118	0.253	**0.605**	0.212	**0.748**	0.264	0.028	-0.423	0.363	0.210	**0.511**	**-0.533**	0.327	0.337	0.159	-0.185	**0.551**	0.287	**0.666**	**-0.655**	0.452	-0.075	1	
%C5	-0.486	-0.139	-0.435	-0.337	0.163	-0.327	**-0.615**	-0.383	-0.223	-0.038	-0.401	0.412	-0.363	-0.208	**-0.611**	0.477	-0.216	-0.186	-0.465	0.026	**-0.584**	-0.438	-0.383	0.067	**-0.682**	-0.270	**-0.603**	1

Table 6-5 Correlation matrices-WWTP (EfOM dominant)

	Alk	TH	ε280	EC	Tot I	COD	DIN	DON	HIX	Ph/Ar	DOC	HPO	TPI	HPI	PS	HS	LMA	Br-	UV_{254}	SUVA	$BDOC_5$	PolyS	FI	%C1	%C2	%C3	%C4	%C5
Alk	1																											
TH	0.467	1																										
ε280	-0.312	0.332	1																									
EC	0.399	**0.701**	0.173	1																								
Tot I	0.148	-0.216	0.221	0.100	1																							
COD	0.310	-0.211	0.014	-0.223	-0.060	1																						
DIN	0.009	**-0.563**	-0.204	-0.359	0.409	0.245	1																					
DON	0.275	-0.163	-0.161	-0.031	0.305	**0.660**	0.397	1																				
HIX	0.160	-0.336	-0.494	-0.191	0.294	-0.043	0.259	0.232	1																			
Ph/Ar	0.114	-0.142	-0.209	0.001	0.123	0.022	0.255	0.145	**0.519**	1																		
DOC	0.295	-0.199	-0.227	-0.174	0.100	**0.670**	0.273	**0.748**	0.085	0.027	1																	
HPO	-0.259	0.020	0.414	-0.209	-0.431	-0.141	-0.202	-0.331	-0.339	0.004	-0.417	1																
TPI	-0.242	-0.034	0.198	0.217	-0.206	-0.192	-0.091	-0.243	-0.237	-0.049	-0.266	0.124	1															
HPI	0.366	0.016	-0.411	0.012	0.421	0.228	0.195	0.368	0.379	0.018	0.459	**-0.804**	**-0.686**	1														
PS	0.426	-0.031	-0.335	-0.213	0.165	0.155	0.160	0.393	0.468	0.393	0.426	-0.429	-0.293	**0.475**	1													
HS	-0.217	0.033	0.320	-0.240	0.153	-0.030	-0.194	-0.309	-0.385	-0.287	-0.286	0.438	0.047	-0.349	-0.395	1												
LMA	0.024	0.002	-0.145	0.340	-0.213	-0.108	0.010	-0.001	0.096	0.066	-0.024	-0.238	0.142	0.092	-0.133	**-0.744**	1											
Br-	0.435	-0.188	0.050	0.170	0.571	0.557	0.464	0.437	-0.119	-0.196	0.476	-0.179	-0.486	0.374	0.055	-0.453	0.027	1										
UV_{254}	0.041	-0.032	0.113	-0.181	0.174	**0.511**	-0.022	0.366	-0.292	-0.155	**0.695**	0.008	-0.190	0.105	0.029	0.260	-0.358	0.457	1									
SUVA	-0.326	0.232	0.463	-0.062	-0.091	-0.222	-0.420	-0.437	-0.482	-0.295	**-0.528**	**0.575**	0.106	-0.479	**-0.542**	**0.708**	-0.404	-0.450	0.226	1								
$BDOC_5$	0.299	-0.382	**-0.697**	-0.334	0.104	**0.695**	0.486	**0.581**	**0.525**	0.488	**0.776**	-0.293	-0.188	0.312	**0.714**	-0.102	-0.498	**-**	**0.773**	**-0.538**	1							
PolyS	0.364	-0.015	-0.394	-0.088	0.094	0.132	0.091	0.139	0.297	-0.082	0.111	-0.345	-0.383	0.461	0.099	-0.358	0.351	0.214	0.001	-0.110	0.227	1						
FI	0.418	0.248	0.154	0.170	-0.045	0.192	0.278	0.338	0.179	**0.509**	0.162	-0.153	-0.046	0.142	0.341	-0.459	0.239	-0.091	-0.139	-0.375	0.306	0.101	1					
%C1	-0.036	0.245	0.375	0.120	-0.028	-0.138	-0.157	-0.143	**-0.730**	-0.461	-0.175	0.253	0.040	-0.200	-0.317	0.380	-0.184	0.080	0.230	0.486	-0.385	-0.144	-0.138	1				
%C2	-0.078	-0.044	-0.145	-0.065	0.103	-0.295	0.365	-0.089	0.434	0.377	-0.440	0.065	0.185	-0.162	0.014	-0.141	0.144	**-0.833**	**-0.588**	-0.060	0.302	0.105	0.366	-0.354	1			
%C3	0.014	0.322	0.140	0.232	-0.082	-0.120	0.260	-0.252	**-0.509**	-0.146	-0.217	0.357	0.173	-0.360	-0.443	0.326	-0.065	0.149	0.069	0.349	-0.428	-0.186	-0.068	**0.700**	0.085	1		
%C4	0.194	-0.220	-0.049	-0.145	0.123	0.419	-0.092	0.395	0.446	0.237	**0.504**	-0.421	-0.203	0.425	**0.538**	-0.389	0.045	0.044	0.130	**-0.502**	**0.600**	0.053	0.365	**-0.564**	-0.224	**-0.802**	1	
%C5	-0.129	-0.083	-0.164	0.004	-0.143	-0.270	-0.205	-0.266	-0.238	-0.274	-0.202	0.227	0.041	-0.190	-0.391	0.277	-0.020	0.353	0.060	0.301	-0.460	0.061	**-0.760**	0.048	-0.255	0.104	**-0.540**	1

Principal Component Analysis

Principal components (PCs) were extracted by the varimax rotation principal component method. PCA allows a clustering of variables or a grouping of objects based on their differences/similarities which would reflect on variations of organic matter properties among datasets.

A graphical representation of the first two component loadings is provided in Figure 6-3. The first component (PC1) accounted for 70.9% and the second component (PC2) explained about 13.1% of the total variance in the dataset (123 sampling sites). These two components together showed 84.0% the explained variations, and going from the third to fifth PCs explained an additional 14.7%. Since most of the data variance was explained by the first two PCs, meaning that most of the information contained in the dataset is explainable using two PCs, the interpretation of the data variance and graphical representation will be enough with these components. As delineated in Figure 6-3, samples were distinguished by forming three clusters characterized by properties of organic matter in the samples, and the grouping-lines were made based on the virtual diagonal lines crossing from 1 of PC1 to 1 of PC2, and from -1 of PC1 to -1 of PC2. On the whole, the first PC separated samples according to NOM-dominant samples versus EfOM-related samples (i.e., either EfOM-impacted samples or wastewater effluent). Furthermore, the EfOM-related samples were separated according to the extent of EfOM dominance, i.e., "wastewater effluents/dominant streams" (so called, highly EfOM-impacted samples), and "samples having wastewater impacts" (less EfOM-impacted samples) by the second PC. From the perspective of the first PC, the NOM-dominant samples showed positive loading, whereas the EfOM-related samples (WWTP effluents and EfOM-impacted downstream samples) had negative loadings. In consideration that the first PC showed a significant difference

in explained variation compared to the second PC (i.e., 70.9% vs. 13.1%), one can presume that the first PC has the greatest contribution to the clustering, especially a separation between NOM-dominant samples and EfOM-related samples, meaning that the first PC is a factor discriminating between NOM and EfOM properties (i.e., the first PC contains information represented by parameters such as SUVA, hydrophobicity/hydrophilicity and fluorescence index). On the other hand, EfOM-related samples (WWTP effluents and EfOM-impacted downstream samples) showed positive loadings along the second PC, which showed more positive for less EfOM-impacted downstream samples than WWTP effluents (i.e., highly EfOM-impacted samples), while NOM-dominant samples showed negative loadings. Based on careful investigations of box plots for organic parameters, information held by the second PC is thought to be related to the redox status, mean oxidation number of carbon, or humidification (index), which all are attributable to a degree of oxidation of organic matter [122-125].

Among EfOM-impacted samples, there exists an overlapping area between the WWTP effluents and wastewater impacted downstream streams/rivers samples, which may indicate that characteristics of organic matter in highly EfOM-impacted downstream samples are close to WWTP effluents. Moreover, it was observed that some samples in the less EfOM-impacted group, samples from the lower courses of downstream rivers and from the end process in downstream DWTPs, were clustered to the group of NOM-dominant samples. Therefore, overall, a transition of EfOM-like organics to NOM-like organics in terms of properties of organic constituents may vary depending on the extent of degradation of EfOM-like organics during wastewater treatment processes or other natural purification processes in the downstream river/stream.

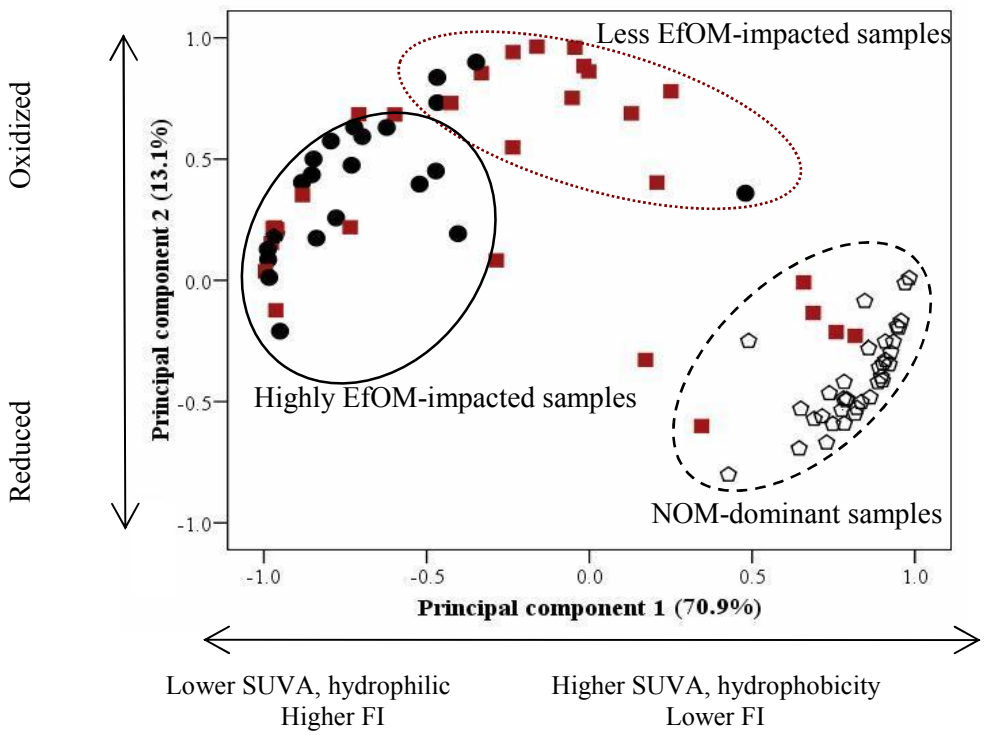


Figure 6-3 Principal component loading plot of upstream and wastewater EfOM containing samples. Values in the parentheses are the explained variations for each component; hollow pentagon (☞-NOM-dominant samples, dark circle (●)-wastewater effluent, red rectangle (■)-wastewater-impacted samples

Multiple Linear Regression Analysis

As previously mentioned, the MLR model assumes that some dependent variable, 'Y', responds to other independent variables (predictors), X_i, which should not be inter-correlated and should have significant and separate effects on the value of dependent variables. Given the reliable model developed, it is possible not only to predict future levels of the dependent variable, but also to assess the contributions of each predictor on the model, in which the scales of values would be expressed.

In this study, bulk DOC can be used as a differentiating factor because DOC is an integration of various types of organic matter and its property varies depending on the dominance of NOM or EfOM. Hence, prediction of DOC was investigated for the same group of

data. Meanwhile, based on the PCA, as the presumable parameter determining direction of PC1 which showed the most variations between NOM-*dominant* samples and EfOM-impacted samples, the hydrophilicity (HPI) was selected, and in MLR was used in assessing this as a dependant variable in the model depending on the datasets of NOM-dominant samples, and EfOM-dominant samples consisted of WW effluent and WW-impacted samples. All independent variables for the model developed were selected by the stepwise method, and the decision of regression coefficients as well as selection of independent variables was made on the basis of ANOVA results.

Table 6-6 shows the MLR results between the NOM-dominant dataset and the EfOM-impacted datasets. For DOC, both datasets show that DOC prediction is most highly related to UV_{254} values, and different coefficients (19.541 and 14.212, respectively) are the indications of different contributions of UV-absorbing constituents to DOC depending on samples, suggesting that $DOC_{NOM\text{-}dominant}$ is more related to UV_{254} than $DOC_{EfOM\text{-}impacted}$. Moreover, DOC in the NOM-dominant sample dataset exhibited that HPO materials of NOM are responsible for DOC, whereas DOC-causing materials in EfOM-impacted sample dataset are likely to reflect HPI and TPI properties (i.e., non-HPO). Overall, bulk DOC prediction equations confirm that organic matter contributions to DOC are different and polarities are the most differentiating parameters between NOM-dominant and EfOM-impacted samples. The prediction of HPI properties showed $HPI_{NOM\text{-}dominant}$ is greatly influenced by negative log values of UV_{254}. On the other hand, the HPIs of EfOM-impacted samples are explained by the amount of polysaccharide constituents and the degree of humidification index (HIX). As seen in clustering by PCA (Figure 6-3), EfOM-related samples were grouped according to be oxidized/reduced, which relates to more humidification processes.

Table 6-6 MLR results of DOC and HPI predictions

Dependent Variables	Regression equation	R^2	F	*p*
DOC	$DOC_{NOM\text{-}dominant}=10^{-1.047} \times HPO^{1.311} \times UV_{254}{}^{19.541}$	0.975	119.32	0.000
	$DOC_{EfOM\text{-}impacted}=10^{0.536} \times HPI^{1.400} \times TPI^{2.288} \times UV_{254}{}^{14.212}$	0.926	59.834	0.000
HPI	$HPI_{NOM\text{-}dominant}=10^{-0.714} \times DOC^{0.609} \times UV_{254}{}^{-13.175}$	0.975	85.400	0.000
	$HPI_{EfOM\text{-}impacted}=10^{-0.139} \times DOC^{0.354} \times PS^{0.648} \times HIX^{4.203}$	0.939	102.22	0.000

Summary and Conclusions

Multivariate statistical analyses were applied to investigate and explain the characteristic similarities/discrepancies of EfOM compared to NOM, using three-binned dataset. Explanatory data analysis by box plots exhibited clear distributions of higher organic loadings in WWTP effluents and WW-impacted samples indicating contributions of EfOM from WWTP to downstream sites.

PCA showed that samples were separated mostly by two PCs. PC1 contributed greatly to clustering samples depending on the extent of dominance of NOM or EfOM, meaning that PC1 is a discriminating factors between NOM and EfOM properties. PC2 separated samples depending on a degree of oxidation status of organic matter more to be oxidized than NOM-dominant samples.

MLR was applied to assess the similarities/differences between NOM and EfOM, and to identify key factors/parameters providing the most information for differentiation of EfOM versus NOM. By prediction of DOC which was selected as a differentiating factor HPO/UV-absorbing material are more contributing to DOC in NOM-dominant samples, whereas organics with HPI/TPI properties are likely to contribute to DOC in EfOM samples. Hydrophilicity in NOM-dominant was shown to be greatly influenced by negative log values whereas HPO

properties shown in EfOM-impacted samples are more explainable by its constituent/compositions and the degree of natural processes such as oxidation/reduction/humidification.

In conclusions, the statistical assessment of NOM characteristic parameters using multivariate analyses seem to provide real assessment of water sources or each site, and to determine and identify the role of EfOM or its impact to water quality classification. Thus water quality monitoring data intended for wastewater impact as well as characteristic comparisons between NOM and EfOM would offer better information if the performances are incorporated with chemometric approaches.

CHAPTER 7‡

TREATABILITY OF WASTEWATER-IMPACTED WATER BY A CONVENTIONAL DRINKING WATER TREATMENT (COAGULATION) AND BY BIODEGRADATION (BANK FILTRATION) AS A PRETREATMENT

Introduction

Wastewater reuse has been increasing as an alternative source for regions with limited water supplies. Indirect potable reuse of wastewater occurs when the treated wastewater is discharged into rivers or lakes that are used downstream as drinking water sources. EfOM present in biologically treated wastewater consists of refractory NOM derived from drinking water sources, soluble microbial products from bacteria in the activated sludge, and trace levels of synthetic organic compounds or DBPs from domestic and/or industrial use [59, 126, 127]. In earlier research on EfOM properties, it was proposed that hydrophilic acids were the most abundant fraction in EfOM, accounting for 32-74% of the DOM, followed by aquatic humic substances [128]. Molecular weight distributions (MWDs) of EfOM by ultrafiltration and gel permeation chromatography indicated that EfOM consists of considerably more high MW compounds (possibly humic substances, proteins and polysaccharides), but exhibits a broad MW spectrum of <500 daltons to >50,000 daltons in biological effluents [69, 129-131]. However, despite

‡ A part of book chapter publication:

Seong-Nam Nam, Stuart W. Krasner, Gary L. Amy, (2007). Differentiating Effluent Organic Matter (EfOM) from Natural Organic Matter (NOM): Impact of EfOM on Drinking Water Sources, *Advanced Environmental Monitoring*, Chapter 20, p259-270, Springer-Verlag GmbH, Germany

increased studies on EfOM, it is poorly understood compared to the understanding of NOM, due to its complex and heterogeneous composition and varied sources such as domestic, industrial, or agricultural origins, moreover, it has not been well assessed in terms of what impact wastewater effluent organic matter (EfOM) will have on drinking water treatment plants in terms of efficiency of organics removal. The removal of organic matter is a critical process in reducing disinfection by-products (DBPs). Since the physico-chemical natures of EfOM is known to be different from that of NOM, EfOM may have different impacts/efficiencies on removal of DOC, increase disinfectant or coagulant dose, and increase the formation of DBPs of health and regulatory concern. Therefore, a better understanding of EfOM will enable the drinking water treatment plant receiving EfOM-*impacted* or *–dominated* waters to improve and/or select other treatment approaches to efficiently remove EfOM.

As a representative treatment process for NOM removal, coagulation-flocculation is a well-established drinking water treatment process and has a potential to remove a significant amount of DBP precursors. In coagulation using hydrolyzing metal salts (e.g., ferric chloride or alum), NOM is removed through several mechanisms such as charge neutralization, adsorption, and sweep floc (entrapment) [78, 132]. The degree of removal performance by coagulation varies depending on the coagulant type, coagulation dosage, pH, amount of dissolved organic matter as well as physicochemical properties of the raw waters [133]. Several studies have concluded that iron based coagulants are more effective than alum, leading to lower an overall DBP formation potentials (DBPFPs) by better DOC removal [134, 135]. The initial pH before coagulation determines the character of NOM in solution (i.e., protonation/deprotonation of the functional groups in NOM) as well as the instantaneous polymeric metal-hydroxide species formed by coagulant addition [136]. Coagulation in the acidic pH range (e.g., 5-6) may go through a co-

precipitation by the charge neutralization process while in a higher pH range (e.g., 7-8) coagulation may be via adsorption of organic matter onto metal hydroxide precipitates. Using a typical dosage of the ferric salts, the best precipitation of NOM has been observed at pH 4-6 [134, 137, 138]. A general consensus it has been known that higher MW and hydrophobic substances are more easily removed than lower MW and hydrophilic counterparts [139].

The primary aim of the experiments was to evaluate the impacts of EfOM-receiving waters on the conventional drinking water treatment process (coagulation) based on understanding of the characteristics of EfOM, differentiated from NOM. The secondary objective was to assess a biodegradation of EfOM. Biodegradation occurs in associated with river bank filtration (RBF) for potable purposes, and soil aquifer treatment (SAT) for indirect wastewater reuse [59, 140, 141]. Comparisons of treatability or sustainability of EfOM removal by environmental treatment processes will be reported in conjunction with several multivariate statistical analyses such as analysis of variance (ANOVA) and parallel factor analysis (PARAFAC).

Materials and Methods

Source of raw water samples

For the first set of experiments, the selected wastewater effluent (as a EfOM sample) was collected from a WWTP (operated by Two Bridges Sewerage Authority, New Jersey, EPA Region II) which applies nitrification (2-stages of aeration), pressure filtration followed by chlorination; an upstream sample was collected from the Pompton River, which is upstream of the WWTP (as a NOM sample) in August, 2005.

For the second set of experiments, a wastewater effluent and a sample upstream of a WWTP were taken from Philadelphia's Northeast Water Pollution Control Plant (Pennsylvania, EPA

Region III) using the activated sludge process, also in August, 2005.

Coagulation procedures of NOM/EfOM mixtures

In order to investigate the impact of EfOM on drinking water treatment plants, the coagulation-flocculation process was selected as a representative drinking water treatment process for a removal of NOM. Treatability experiments were conducted with a series of NOM/EfOM mixtures using jar-test apparatus. Mixtures of water upstream of the WWTP (i.e., NOM) and the wastewater (WW) effluent (i.e., EfOM) were prepared with 100/0, 75/25, 50/50, 25/75, and 0/100 (NOM %/EfOM %) ratios. Predetermined doses of ferric chloride ($FeCl_3$) using $FeCl_3 \cdot 6H_2O$ (Fishers Scientific Co.) were added in the range of 2 to 40 mg/L (extended to 60 mg/L for a second set of experiments) and these dosages correspond to 0.69 to 13.77 mg/L as Fe^{3+} (extended to 20.66 mg/L for a second set of experiments). Jar testing was carried out for 1 min at 100 rpm for rapid mixing, followed by 15 min at 20 rpm for flocculation, and 20 min for (quiescent) settling. Modification of initial pH was not applied. The final post-coagulation pH (data included in Appendix) was recorded (as well as the initial pH). Treatability of NOM and EfOM was evaluated in terms of DOC removal efficiency. Hereby, EfOM-*dominated* waters are defined as waters mixed with 50/50 or 25/75 or 0/100 (NOM %/EfOM % in volumes) ratios, and EfOM-*impacted* water correspond to the 75/25 mixture.

Biodegradation

Biodegradation experiment of EfOM was performed in terms of 5-day biodegradable dissolved organic carbon ($BDOC_5$), a parameter often used in drinking water treatment to assess biostability. Bioactive sand acclimated with wastewater effluent was used in BDOC

determinations. A 5-day time period would be representative of (shorter-term) residence time in RBF and SAT.

Statistical analysis

Analysis of variance (ANOVA) was used to evaluate statistical significances between the effect of the ratio of EfOM as well as the coagulation pH in addressing the treatability of EfOM.

Results and discussion

Raw water characteristics

The general water characteristics of NOM/EfOM mixtures are summarized in Table 7-1 (for case 1, NJ site) and Table 8-2 (for case 2, PA site). The specific UV absorbance at 254 nm (SUVA) was seen to decrease with an increase in EfOM mixing ratio, UVA_{254} and DOC level, indicating that organic matter with hydrophilic and non-aromatic characters increased relatively in the EfOM-augmented samples. For the NJ site, it was noteworthy that there was a low alkalinity for EfOM-*dominant* waters.

Table 7-1 Water quality parameters of NOM/EfOM mixture samples for New Jersey (NJ) sites

Mixture samples (NOM/EfOM)	**pH**	**UVA_{254},** cm^{-1}	**DOC,** mg/L	**SUVA,** L/mg-m	**Conductivity,** µS/cm	**TDS,** mg/L	**Alkalinity,** mg/L as $CaCO_3$
100/0	8.26	0.077	3.01	2.56	451	216	80
75/25	7.94	0.088	3.80	2.31	511	245	60
50/50	7.50	0.099	4.59	2.15	552	264	40
25/75	7.53	0.109	5.38	2.03	678	326	28
0/100	7.43	0.12	6.17	1.94	822	397	28

Table 7-2 Water quality parameters of NOM/EfOM mixture samples for Pennsylvania (PA) sites

Mixture samples (NOM/EfOM)	**pH**	**UVA_{254},** cm^{-1}	**DOC,** mg/L	**SUVA,** L/mg-m	**Alkalinity,** mg/L as $CaCO_3$
100/0	7.93	0.054	1.94	2.78	68
75/25	8.20	0.074	3.81	1.94	60
50/50	8.11	0.103	5.70	1.80	64
25/75	7.90	0.130	7.45	1.74	68
0/100	8.20	0.163	9.15	1.78	76

Treatability of EfOM-impacted or -dominated water: Case study 1-New Jersey sample waters

Removal of DOC by coagulation

Figure 7-1 shows the DOC removal efficiencies of NOM/EfOM mixtures by coagulation. As shown, the DOC removals from adding 40 mg/L of coagulant were 54 and 57% in the upstream sample and the EfOM-*impacted* water, respectively, whereas the DOC in the EfOM-*dominated* waters was removed by 38 to 46%.

Based on previous researchers' studies related to coagulation of NOM, it is known that the coagulation process preferentially removes hydrophobic organic matter over hydrophilic organic matter. The fluorescence EEM spectra of samples through the coagulation process support this trend (Figure 7-2). As the NOM water was increasingly mixed with more of the EfOM, the protein-like peak became clearer and its fluorescence intensity increased as well (upper figures in Figure 7-2). From the coagulation process, humic-like materials were significantly removed, but the protein-like materials showed lower removals.

Considering that EfOM is more hydrophilic than NOM according to previously presented characteristics of EfOM (Chapter 5), these results are reasonable. In addition, as the upstream water was mixed with more of the EfOM, the bulk concentration of DOC increased (Table 7-1). For NJ samples, when 40 mg/L of coagulant was added, the coagulant to DOC ratio was 13 mg/mg for the upstream (100/0) sample, whereas the ratio was only 6.5 mg/mg for the EfOM (0/100) sample. The presence of EfOM can also increase the coagulant demand. On a constant coagulant/DOC ratio (e.g., ~6.5 mg/mg), coagulant doses of 20 and 40 mg/L would be needed for the upstream and EfOM samples, respectively. The use of 20 mg/L of coagulant on the upstream sample or 40 mg/L of coagulant on the EfOM sample both resulted in similar DOC

removal efficiencies (~40%, Figure 7-1).

These results mean that the presence of EfOM in a drinking water source may have a negative impact on the efficiency of the drinking water treatment process due to the different properties of EfOM with respect to NOM, especially in terms of DOC (DBP precursor) removal and/or coagulant demand.

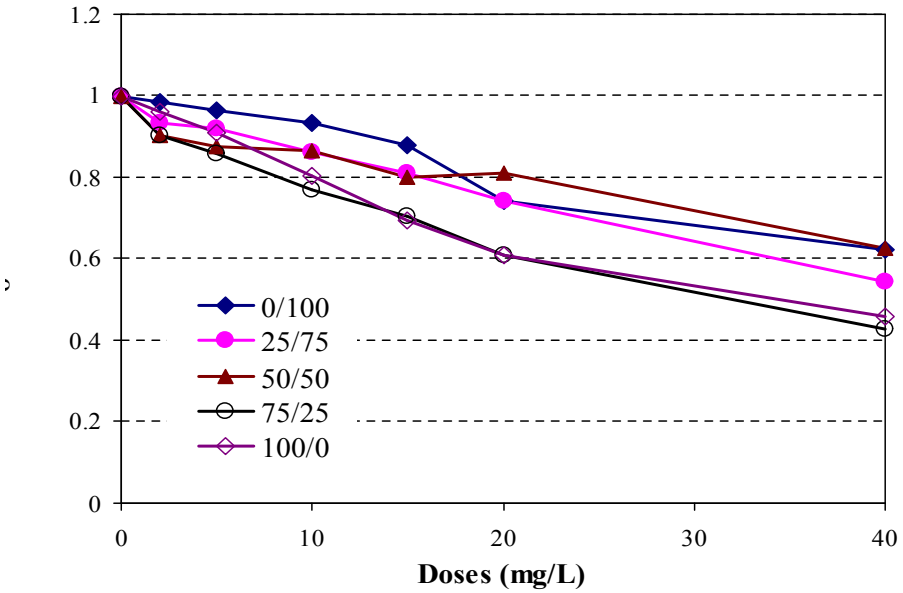


Figure 7-1 Removal of DOC in NOM/EfOM mixtures by coagulation for NJ samples

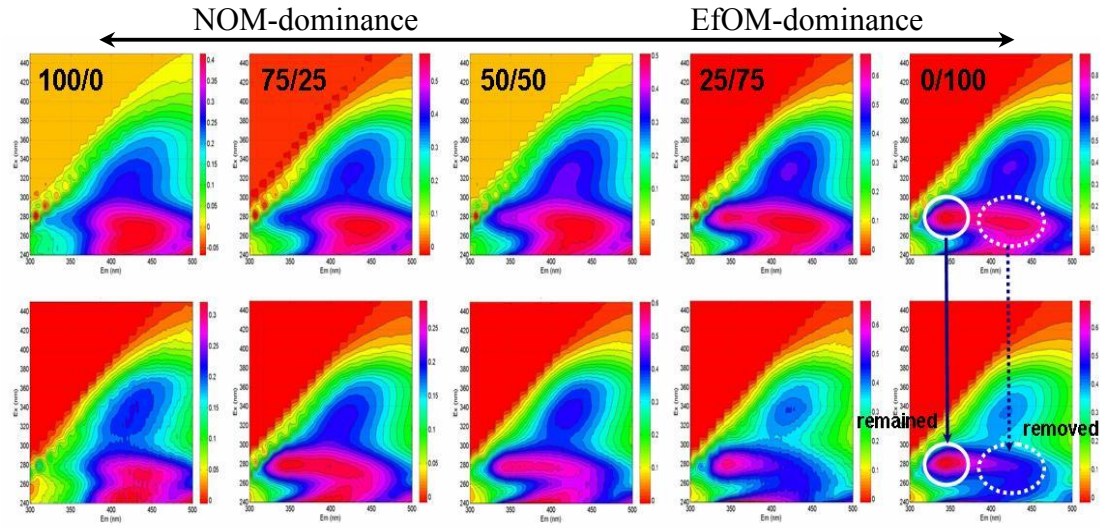


Figure 7-2 EEM contour maps before coagulation (upper) and after coagulation with 40 mg/L dose (lower) (x-axis: emission wavelength of 290-500 nm, y-axis: excitation wavelength of 240-450 nm)

Coagulation resulted in increased values of fluorescence index (FI) of the NOM/EfOM mixtures. Allochthonous materials, which are sources of humic-like materials, were preferentially removed over autochthonous organic matter, which caused an increase in the relative portion of organic matter from autochthonous sources (Figure 7-3).

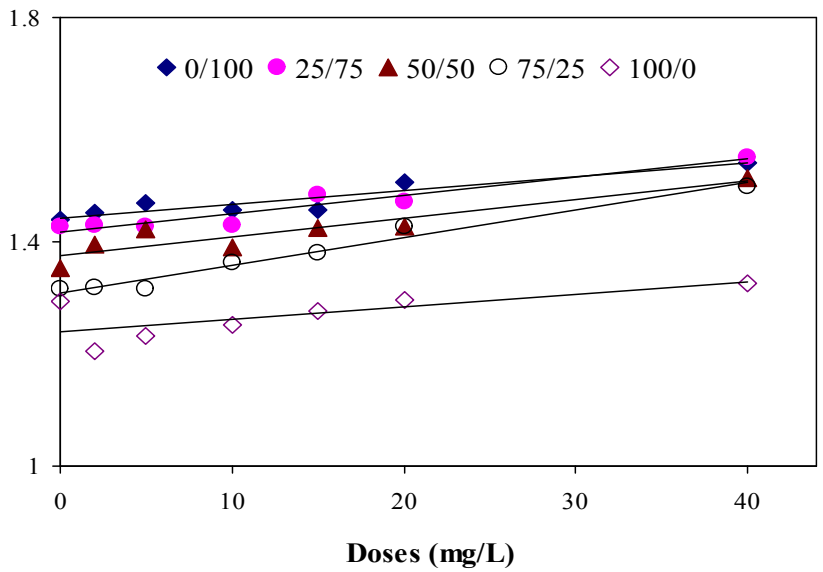


Figure 7-3 Changes in FI values of the NOM/EfOM mixtures by the coagulation process (fluorescence sample preparation: DOC dilution to ~1 mg/L with pH 2.7-2.8 adjusted 0.01M KCl using concentrated HCl)

Biodegradation of EfOM

Figure 7-4 and 7-5 show the EEM contour plots and SEC-DOC of EfOM through aerobic BDOC tests. In contrast to the coagulation process, biodegradation preferentially removed protein-like substances. In Figure 7-5, degradation of organic carbon in the polysaccharide peak was more complete and faster, whereas humic substances were slowly and incompletely biodegraded.

After a 21-day BDOC test, the polysaccharide peak almost completely disappeared. Because large-sized components of the polysaccharide peak are saturated or linear structures, with fewer double bonds, bacteria were more easily able to assimilate these structures compared to humic substances with unsaturated, branched or ring structures. Aerobic $BDOC_5$ showed a linear relation with DOC, the hydrophilic organic matter fraction, the amount of polysaccharides of raw waters, as well as the concentration of dissolved organic nitrogen (DON). However, it was inversely related to SUVA values of samples (results not shown). These results are representative of trends expected for a river bank filtration with dominantly aerobic conditions.

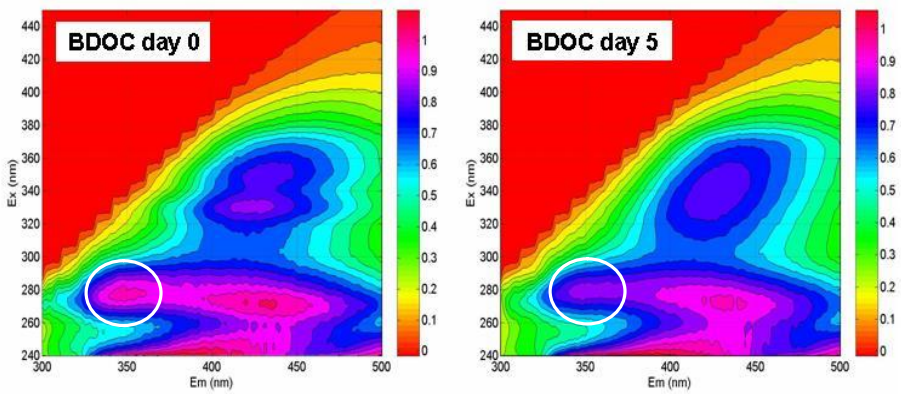


Figure 7-4 Changes in EEM contour maps by aerobic $BDOC_5$ test of EfOM (data from wastewater sample taken in August 2005) (x-axis: emission wavelength of 290-500 nm, y-axis: excitation wavelength of 240-450 nm)

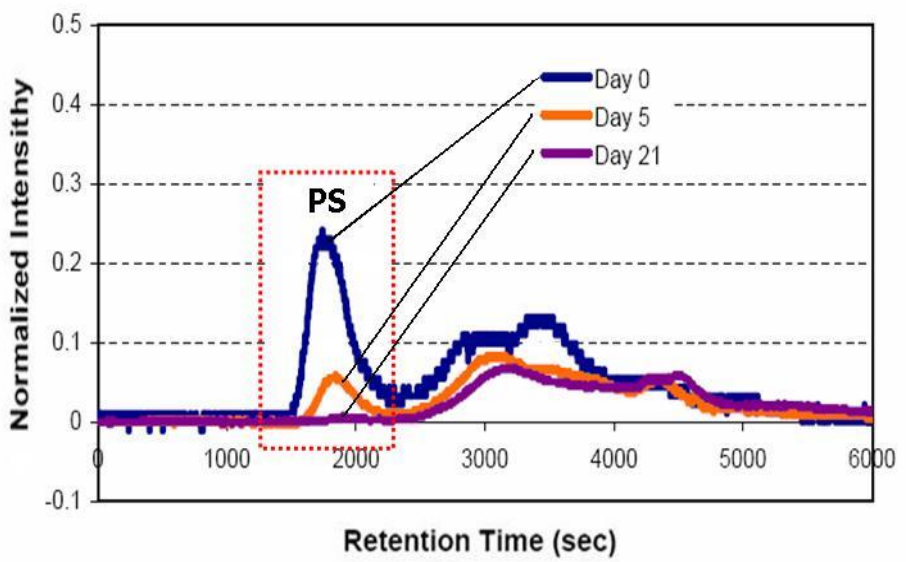


Figure 7-5 SEC-DOC of aerobic BDOC test of EfOM for 5 and 21 days (data used here were for a sample collected from a river in South eastern USA (EPA region 9) in February 2005, which was EfOM-*dominated*

Treatability of EfOM-impacted or -dominated water: Case study 2-Pennsylvania sample waters

General water quality parameters of NOM/EfOM mixtures for PA samples are summarized in Table 7-2. In addition to Table 7-2, UV scans were conducted (Figure 7-6-a and -b). Light absorption of humic substances at λ>250 nm is attributed to aromatic chromophores in the molecules [142, 143], and the absorbance at 254 or 280 nm is used as an indicator of the concentration of NOM in water. Although UV spectra of NOM having an exponential decrease are typically featureless and do not give much information about the characteristics of NOM, however, as shown in Figure 7-6-a, in which the absorbance increases as the EfOM ratio in the mixtures increases. The profiles of specific UV absorbance, defined as $UVA_{\lambda i}$ to DOC ratio ($UVA_{\lambda i}$/DOC multiplied by 100), provide an indication of a degree of aromaticity of unsaturated bonds (e.g., carbon to carbon) as well as the profiles of molar absorptivity ($M^{-1}cm^{-1}$) analogous to SUVA. It has been reported that SUVA has a strong correlation with the contents of aromatic carbon by ^{13}C nuclear magnetic resonance (^{13}C-NMR) [144]. Tambo *et al.*[145] characterized correlations between MWs of organic matter and SUVA, showing that the higher molecular weight (HMW) humic substances such as humic acids and lower molecular weight compounds such as fulvic acids exhibited SUVA values of 3-5 and 2 L/mg·m, respectively. Figure 7-6-b represents profiles of SUVA of the NOM/EfOM mixtures, showing that NOM has greater SUVA than other EfOM mixed waters, and there are distinguishable differences in the range of 250-350 nm, which may indicate that chemical structures/compositions absorbing these ranges of light energy are underlining as major distinction between NOM and EfOM. The opposite trends between UVA and SUVA in NOM/EfOM mixtures profiles imply that organic materials in samples have different chemical properties.

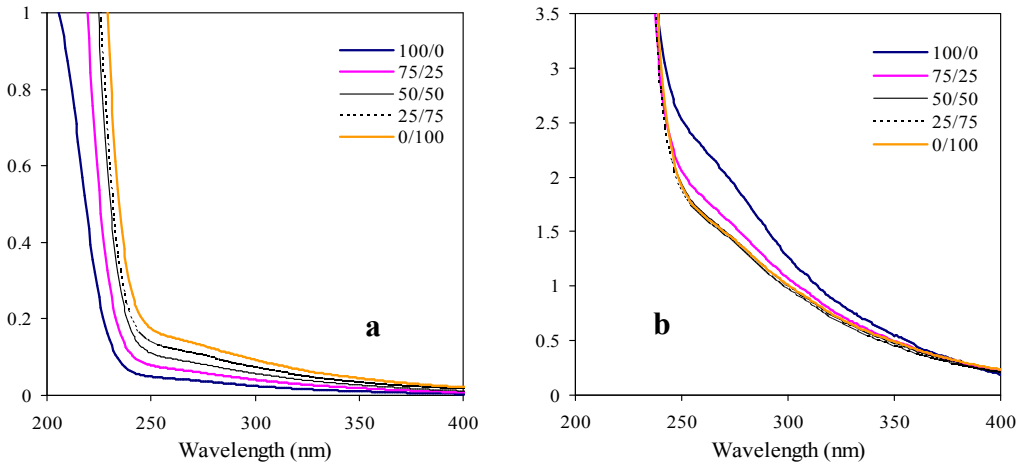


Figure 7-6 UV spectrum (a) and specific UV absorbance (b) profiles of NOM/EfOM mixtures (%/% in volume) before coagulation

Removal of DOC by coagulation

Similar to the results for the NJ sites, an increase of the coagulant doses led to an increase in DOC removal. After the coagulation application of up to 60 mg/L of coagulant, pH was lowered to 5.65 for NOM (100%/0%), and 5.80 for EfOM-*dominant* waters (0%/100%). Decrease of removal efficiency by restabilization due to excessive coagulants was not observed for the given experimental condition. At 60 mg/L dosage, the highest DOC removal (60% removal) was observed in the NOM (100%/0%) sample and the least removal (40 % removal) in the EfOM (0%/100%) (Figure 7-7). As shown earlier in Table 7-2 and Figure 7-6-b, the NOM (100%/0%) sample exhibited higher SUVA values in 250-350 nm regions than that of EfOM-*impacted* or –*dominant* waters. Changes of SUVA for each water show noticeable decreases in no or less EfOM-*impacted* waters, whereas there are small changes in more EfOM-*impacted* waters.

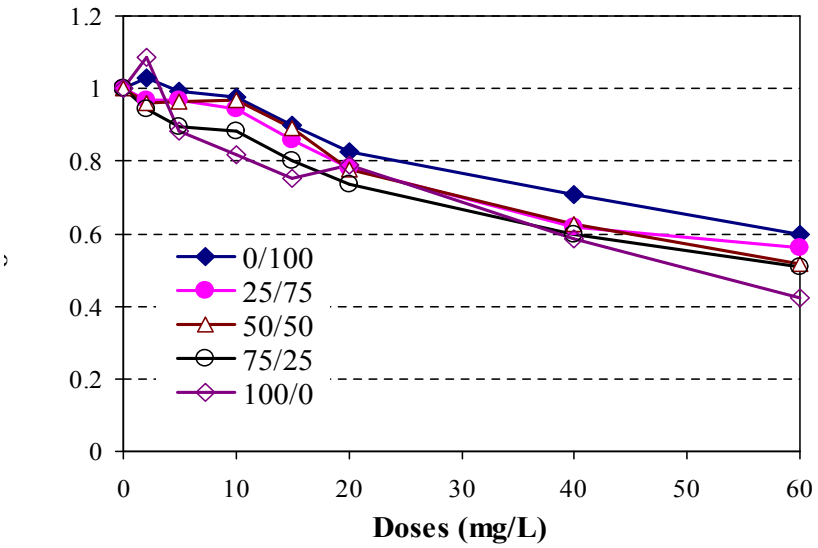


Figure 7-7 DOC removal efficiency by coagulation for NOM/EfOM mixtures from the PA sites

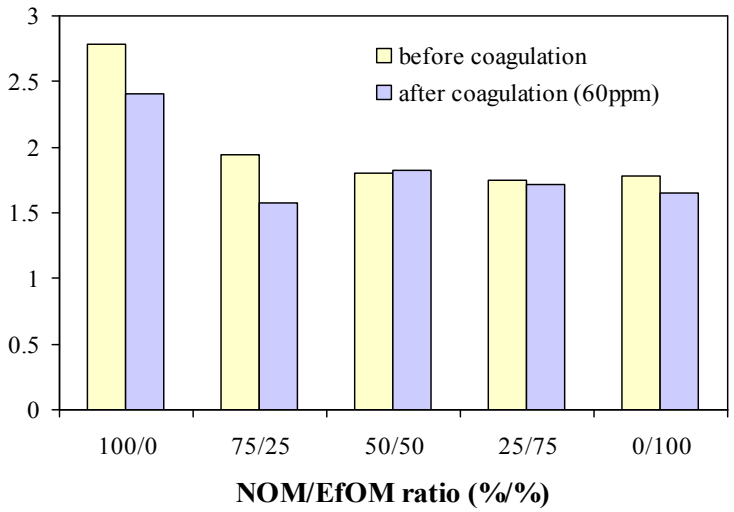


Figure 7-8 SUVA change by coagulation (dose 60 mg/L) for NOM/EfOM mixtures from PA sites

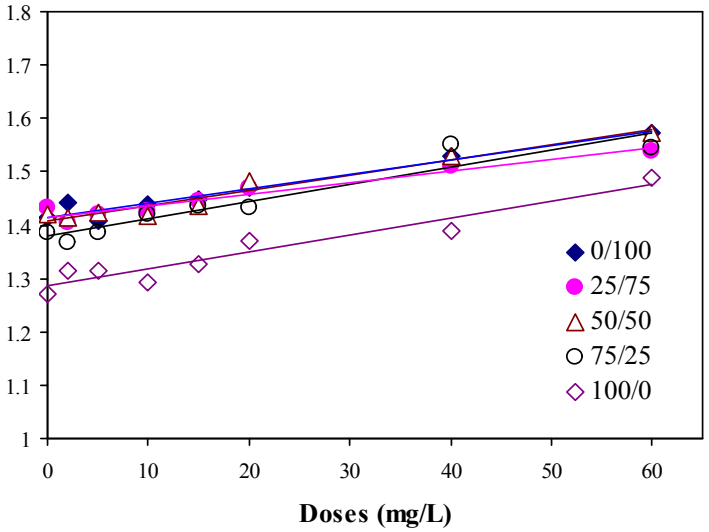


Figure 7-9 Fluorescence Index change by coagulation for NOM/EfOM mixtures from PA sites

HPSEC-DOC-UVA-Differential DOC chromatogram

Figure 7-10 and 7-11 depict the molecular weight distributions of NOM (100%/0% in NOM/EfOM), and EfOM-*dominant* samples before and after coagulation with a 60 mg/L dose. As shown, the molecular weight distributions range from small to a few tens of thousands daltons, and greater amount of humic substances and low molecular weight (LMW) acids were observed in EfOM-*dominant* waters along with tailing effect of LMW molecules. Differential DOC chromatogram was calculated by subtracting the normalized $DOC_{after\ coagulation}$ response from the normalized $DOC_{before\ coagulation}$ response to see which fractions of NOM were most removed. The result of NOM waters (100%/0%) showed that the higher and hydrophobic MW substances were removed more significantly than the LMW substances. Even with high doses of coagulant, the LMW substances showed minor removal. Krasner and Amy [137] reported that certain fractions of NOM, i.e., hydrophobic and higher MW NOM, were more effectively removed than any other fractions, and some waters with a dominance of LMW NOM were harder to treat with coagulation. They also observed that there was a preferential removal of

NOM that absorbs UV light, indicating that aromatic NOM is preferentially removed by coagulation. The results of the NOM sample (Figure 7-10 and 7-11) to a large extent agreed with their study. In SEC-UVA chromatograms with various wavelength detections revealed that humic substances with the highest UV absorptivity at 280 nm were removed preferentially, whereas LMW components with the highest UV absorptivity at 210 nm and 230 nm, and fair absorptivity at 280 nm were unlikely to be removed by coagulation. However, these trends should be interpreted with the fact that a high nitrate in low DOC waters may interfere with interpretation of the UV-based results because nitrate is not removed by coagulation.

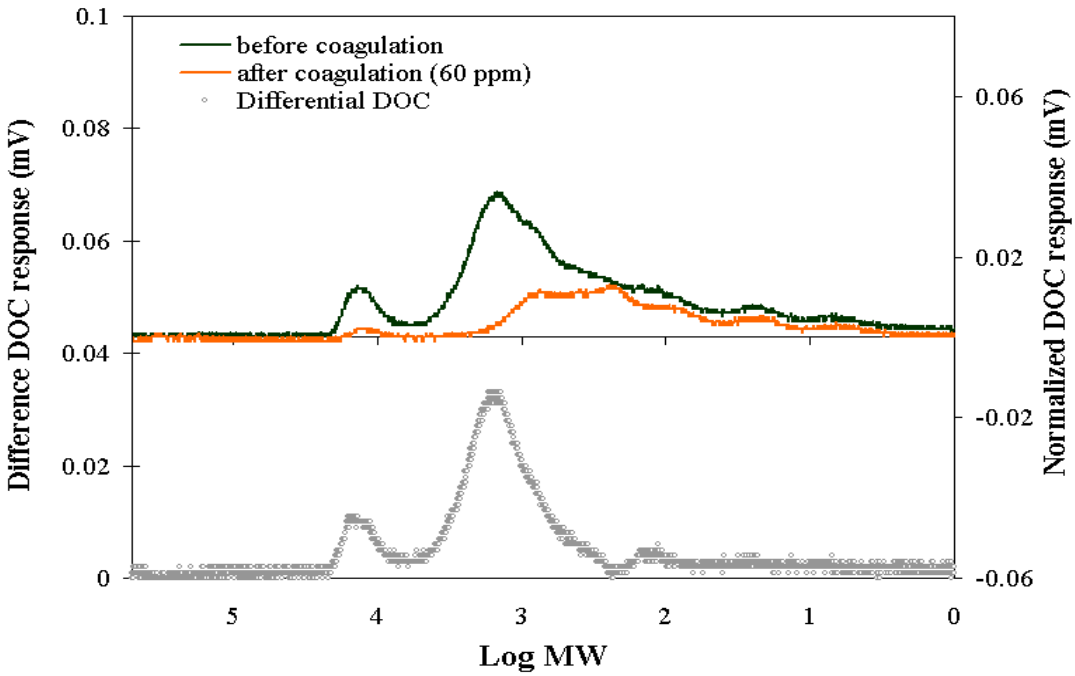


Figure 7-10 HPSEC-DOC-Differential DOC chromatograms for before coagulation and after coagulation with 60 mg/L for NOM (100%/0% in NOM/EfOM) sample (PA sites). Differential DOC was calculated by subtracting the normalized $DOC_{after\ coagulation}$ response from the normalized $DOC_{before\ coagulation}$ response; TSK HW-50S column (2×25 cm), Na_2SO_4 eluent with phosphate buffer (pH: 6.8, ionic strength: 0.1 M), flow rate 1 mL/min

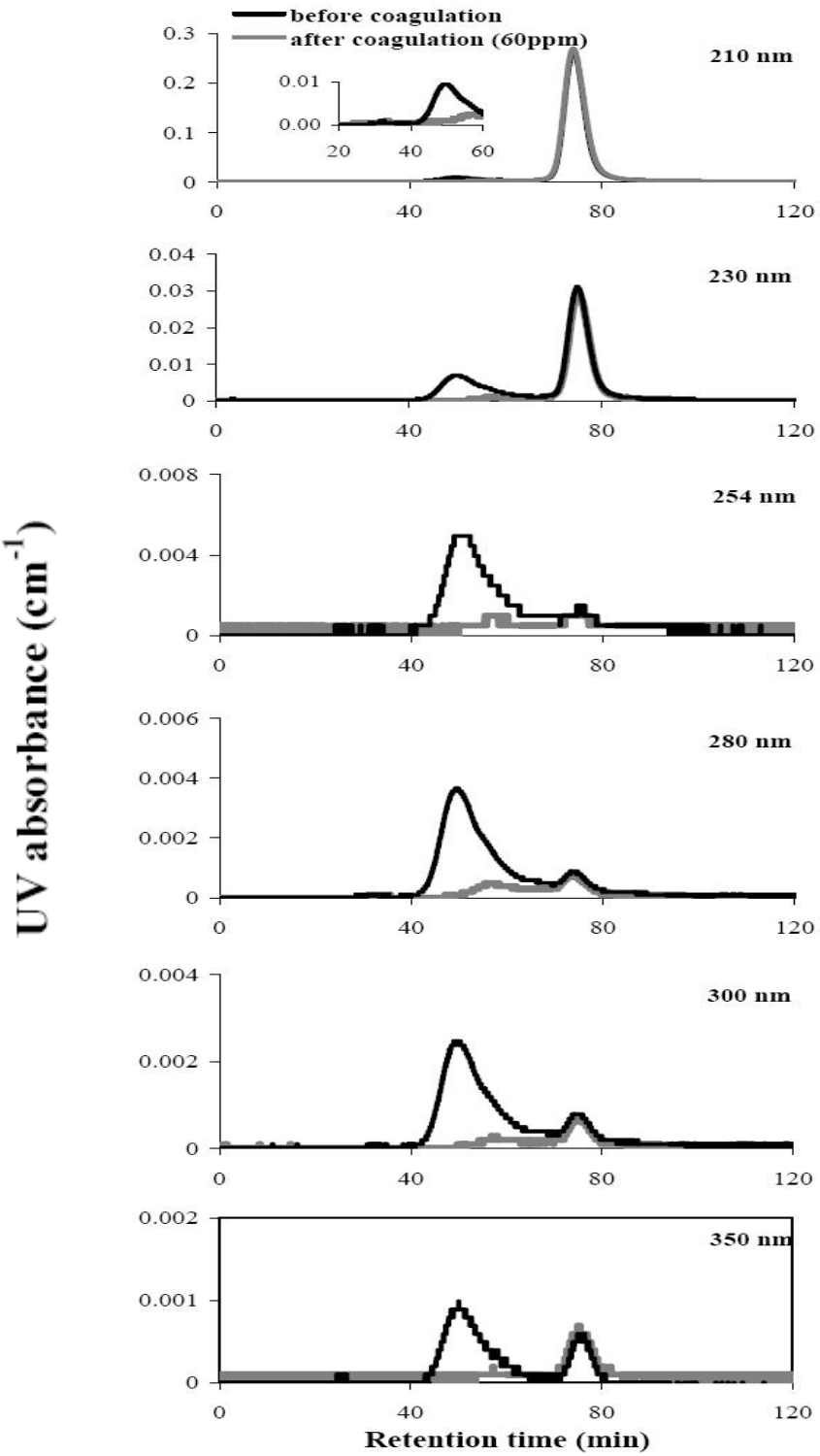


Figure 7-11 HPSEC-UVA chromatograms for samples before coagulation and after coagulation with multiple wavelengths for NOM (100%/0% in NOM/EfOM) sample (PA sites). Doses of 60 mg/L; TSK HW-50S column (2×25 cm), Na_2SO_4 eluent with phosphate buffer (pH: 6.8, ionic strength: 0.1 M), flow rate 1 mL/min

For EfOM-*dominant* water, on the other hand, the SEC-DOC chromatogram (Figure 7-12) showed a high amount of LMW substances with MWs of about 2,700 daltons and below, although the SEC-UVA chromatogram (Figure 7-14) did not exhibit this. It was observed that a significant amount of LMW substances was removed in addition to HMW substances of polysaccharides and humic substances, with a total removal corresponding to ~ 50% of the DOC, based on the integrated area of the differential DOC chromatogram. It appeared that because LMW substances in the EfOM-*dominant* water had somewhat different characteristics compared to those in NOM, either by changes of composition or of chemical nature while passing through biological treatment processes. Also, the large dose of $FeCl_3$ used (60 mg/L) provided removal through a sweep coagulation mechanism.

Changes in molecular weight by coagulation were investigated in terms of weight average molecular weight (M_w) (Table 7-3). It was found that NOM water had higher M_w than EfOM-*dominant* water, implying that the high input effect of EfOM containing LMW substances results in shifting the weight average molecular weight to a lower value. In addition, the M_w after coagulation decreased, and a greater decrease (i.e., greater shifting to low M_w) was shown in NOM water (1,835 to 554 daltons) compared to EfOM-*dominant* water (1,136 to 654 daltons). These results confirm that coagulation preferentially removes HMW substances, and smaller molecules are more difficult to be removed, in agreement with other coagulation research [146].

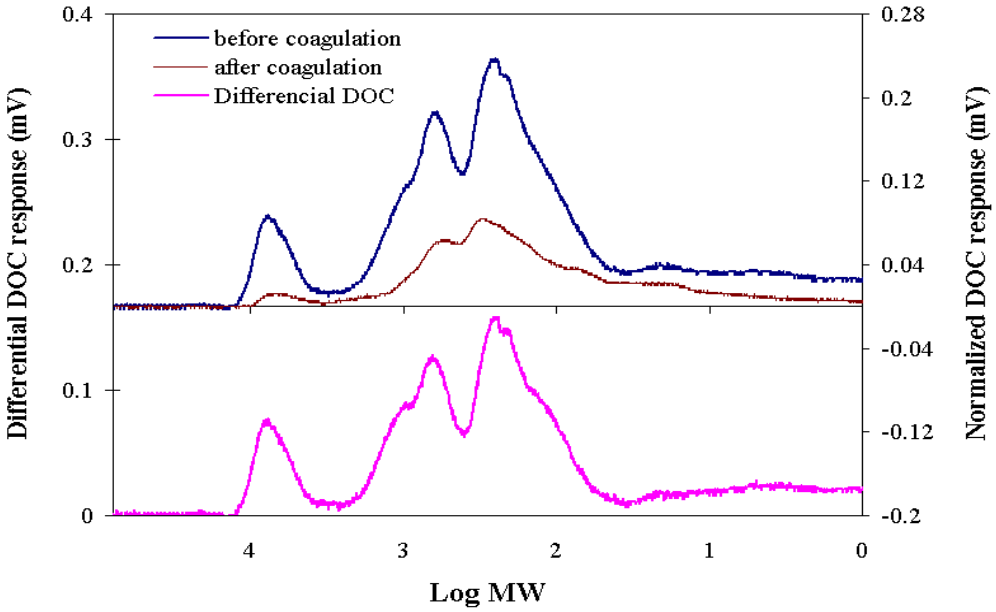


Figure 7-12 HPSEC-DOC-Differential DOC chromatograms for before coagulation and after coagulation with 60 mg/L for EfOM-*dominant* (25%/75% in NOM/EfOM) sample. Differential DOC was calculated by subtracting the normalized $DOC_{after\ coagulation}$ response from the normalized $DOC_{before\ coagulation}$ response; TSK HW-50S column (2×25 cm), Na_2SO_4 eluent with phosphate buffer (pH: 6.8, ionic strength: 0.1 M), flow rate: 1 mL/min)

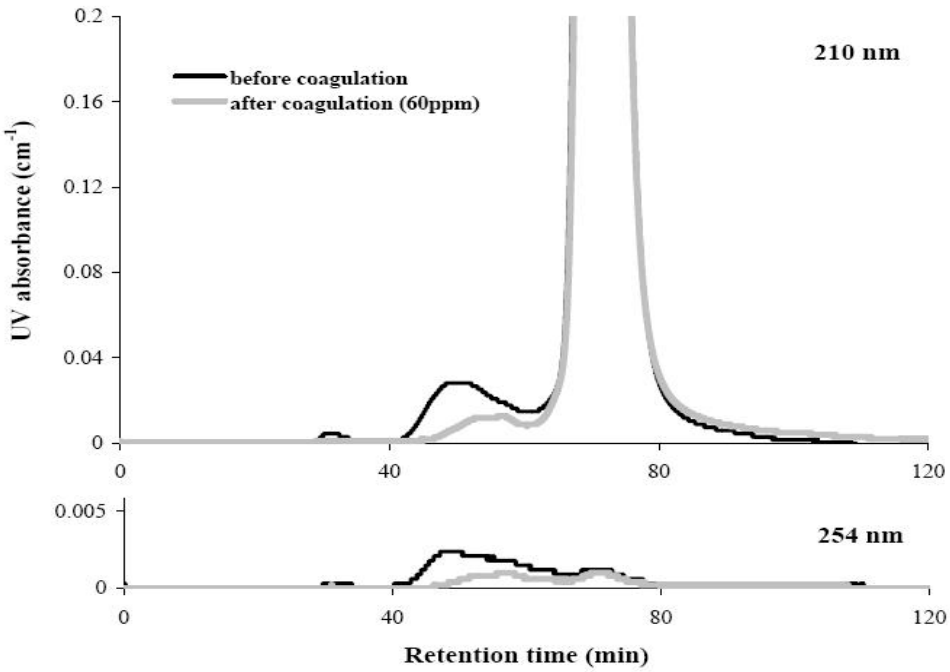


Figure 7-13 HPSEC-UVA chromatograms for before coagulation and after coagulation with multiple wavelengths for NOM (25%/75% in NOM/EfOM) sample (PA sites). Doses of 60 mg/L; TSK HW-50S column (2×25 cm), Na_2SO_4 eluent with phosphate buffer (pH: 6.8, ionic strength: 0.1 M), flow rate: 1 mL/min)

Table 7-3 Changes of M_w by coagulation in NOM and EfOM-*dominant* waters

	100%/0%	25%/75%
Before coagulation	1,835	1,136
After coagulation (60 mg/L)	554	654

Unit: dalton

Differential fluorescence excitation-emission matrix (EEM)

Figure 7-14 represents the differential EEMs which were plotted by subtracting the EEM with coagulation from the initial (before coagulation) EEM. In the differential EEM, a positive number in fluorescence intensity indicates removal by coagulation, thus the greater the positive values are, the more preferential removal that occurs in the wavelength regions. As shown, most of removals took place in the regions of 265-275 nm/430-460 nm of Ex/Em wavelengths, and 310-330 nm/410-440 nm regions for the second most responsive removals, whereas substances in protein-like regions of 270-290 nm of excitation and 330-350 nm of emission wavelengths were not removed by coagulation. The differential EEMs resemble the one of Suwannee River references NOM (introduced in Chapter 5) in both contour shapes and regions, and this means that EEMs of the hydrophobic substances with moderately HMW substances (i.e., humic substances, introduced as Zone-2 peak in Chapter 5) would correspond to the differential EEM generated by the "coagulation" process. Also, based on the results using fluorescence techniques, the SRNOM-like substances (or humic-like) are easier to control by coagulation.

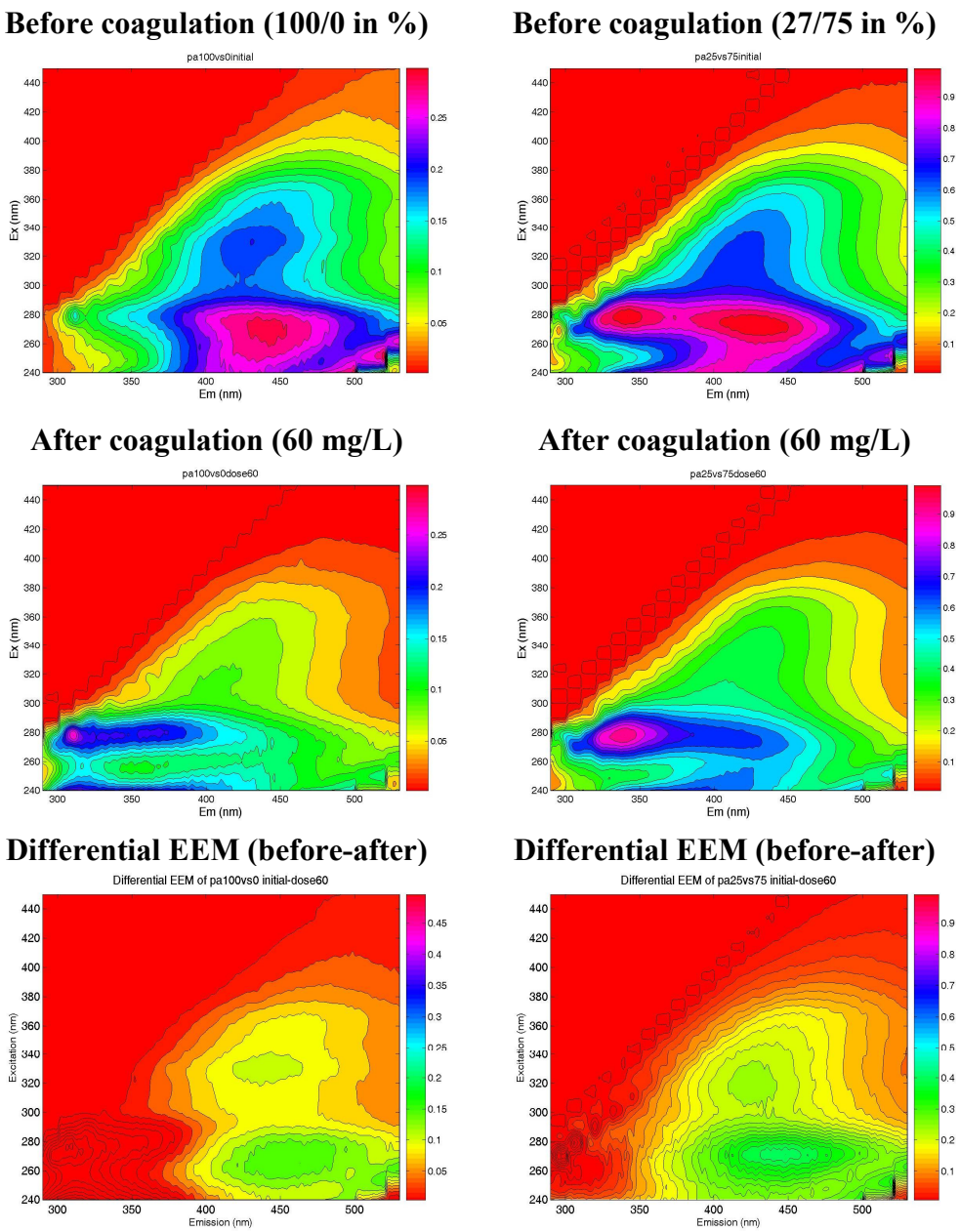


Figure 7-14 The differential EEM contour maps of NOM/EfOM mixtures (PA sites) for before coagulation and after coagulation with dose of 60 mg/L (left column: 100/0, and right column: 25/75 in %/%)

Analysis of Variance (ANOVA)

As mentioned earlier, some of waters had initially low alkalinity whose buffering capacity might be exhausted by adding coagulants at doses as high as 40 or 60 mg/L, possibly causing unfavorable coagulation conditions compared to other waters. Hence, by using an ANOVA test,

the different effects of factors (e.g., % EfOM, coagulant dosage, pH, etc.) operating simultaneously on a result was examined to decide which effect is statistically significant, and to estimate the contribution of parameters to the variability of results.

As a result, for both cases of water samples (NJ and PA sites), the extent of wastewater mixing ratio (for NJ site, F=7.076, p=0.011 and, for PA site, F=8.523, p=0.007) and coagulant dosage (for NJ site, F=11.786, p=0.003, and for PA site, F=13.537, p=0.000) turned out to be statistically significant on DOC removal efficiency, whereas the (final) pH was not statistically important (for NJ site, F=2.669, p=0.111, and for PA site, F=1.775, p=0.201). This result confirms that different characters of EfOM from NOM have a deteriorating impact on drinking water treatment efficiency when water containing EfOM is used as a drinking water source in DWTP. A similar result was also observed by Musikavong *et al*. (2005) who investigated the THMFP along with DOC removal for reclaimed water, reporting no differences of controlled pH and uncontrolled pH in DOC removal efficiency [147].

Therefore, it can be concluded that the overall characteristics of organic matter (e.g., hydrophobicity, hydrophilicity, etc.) changed by wastewater blending is a more important factor to control water quality by drinking water treatment processes rather than pH, although pH affects both the state of ferric hydrolysis products and ionization of functional groups (i.e., -COOH and -OH) of organic matter under the given experimental conditions of this study.

Application of PARAFAC to coagulation and biodegradation

In order to better understand the impact of EfOM on DW treatment efficiency, the fluorescence EEM results from treatability experiments were re-analyzed using the 5 components N&A PARAFAC model (introduced in Chapter 3). As shown in previous chapters, component 1 and component 5 have hydrophobic natures whereas component 2 and component 4 showed hydrophilic properties. Component 3 appears to be of an intermediate nature.

Figure 7-15 depicts the fractional changes of 5 components in NOM/EfOM mixtures. As the ratio of EfOM blending with NOM increases, the bulk DOC of the mixtures increased. However, relative proportions of component 1 and component 5 for the bulk DOC were diminished, while proportions from component 2 and component 4 increased. The fraction of component 3 was mostly the same regardless of the mixing ratio.

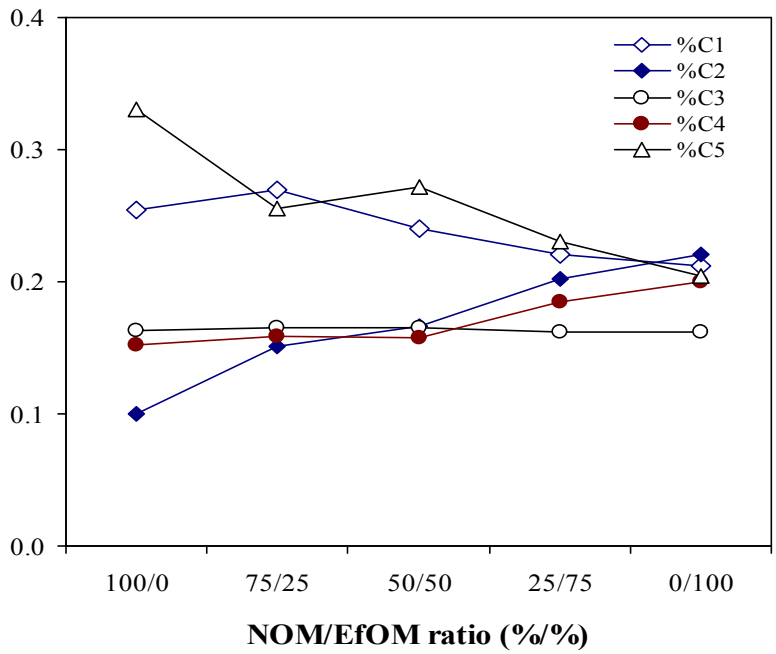


Figure 7-15 Fractional changes of components by increasing the EfOM blending ratio (NJ sites)

Figure 7-16 shows fractional changes of 5 components by coagulation at the various coagulant doses and various mixing ratios. A figure on the bottom of the right column depicts the

regression slope values of the fractional changes of each component shown in the previous 5 sub-figures (100/0, 75/25, 50/50, 25/75 and 0/100 in %/%). What the slope can tell us is which components are preferentially removed by coagulation, and which are not (these results are shown in the lower right comer of Figure 7-16). Hence, a negative value means "preferentially" removed, and the more negative components, the more easily removed. A positive value does not necessarily mean "no removal", because the fractional changes were based on the initial fraction of each component, but rather it indicates "unfavourable removal. Fractions of each component signify their dominances out of 1 (or 100 percent) when assuming the summation of each component accounts for all fluorophores in the sample. For instance, although the fraction of component 1 decreased in the 0/100 mixture compared to the 100/0 mixture, the amount of organic carbon (DOC) accounted by component 1 was greater than that of the 100/0 mixture since the bulk DOC of the 0/100 mixture was greater than the 100/0 mixture. Component 1 and 5 showed negative slope values, while components 2 and 4 exhibited positive slopes. But, component 1 exhibits the greatest negative value, indicating that this component can be removed more preferentially than component 5. In the same way, the negative magnitude of the component 2 slope was greater than component 4, which means component 2 is the hardest to remove with coagulation, followed by component 4. Interestingly, component 3 removals remained mostly the same, even though its DOC was gradually decreased by coagulation. Therefore, this does not mean that component 3 was not removed during coagulation, but rather its removal was stable albeit it might not be "preferentially" removed. Based on these results, the order of coagulation preference is: component 1 > component 5 > component 3 > component 4 > component 2.

In this study, treatability assessment of EfOM in combination with coagulation and

PARAFAC was only explainable in terms of fluorophores; additionally, biodegradability of each component was evaluated by a 5-day aerobic BDOC test with the wastewater (i.e., 0/100 mixture), taken on the same date (August 22, 2006) (Figure 7-17). The $BDOC_5$ of the wastewater was 0.86 mg/L which corresponds to ~14% of the (bulk) initial DOC. Figure 7-17 shows the biodegradability of components and their contributions to $BDOC_5$. As shown, contrary to results of coagulation, 38% of the $BDOC_5$ resulted from component 4, and component 2 and 4 contributed to about 60% of the total $BDOC_5$. Again, as stated earlier, the BDOC simulation is a representative of river bank filtration under aerobic conditions.

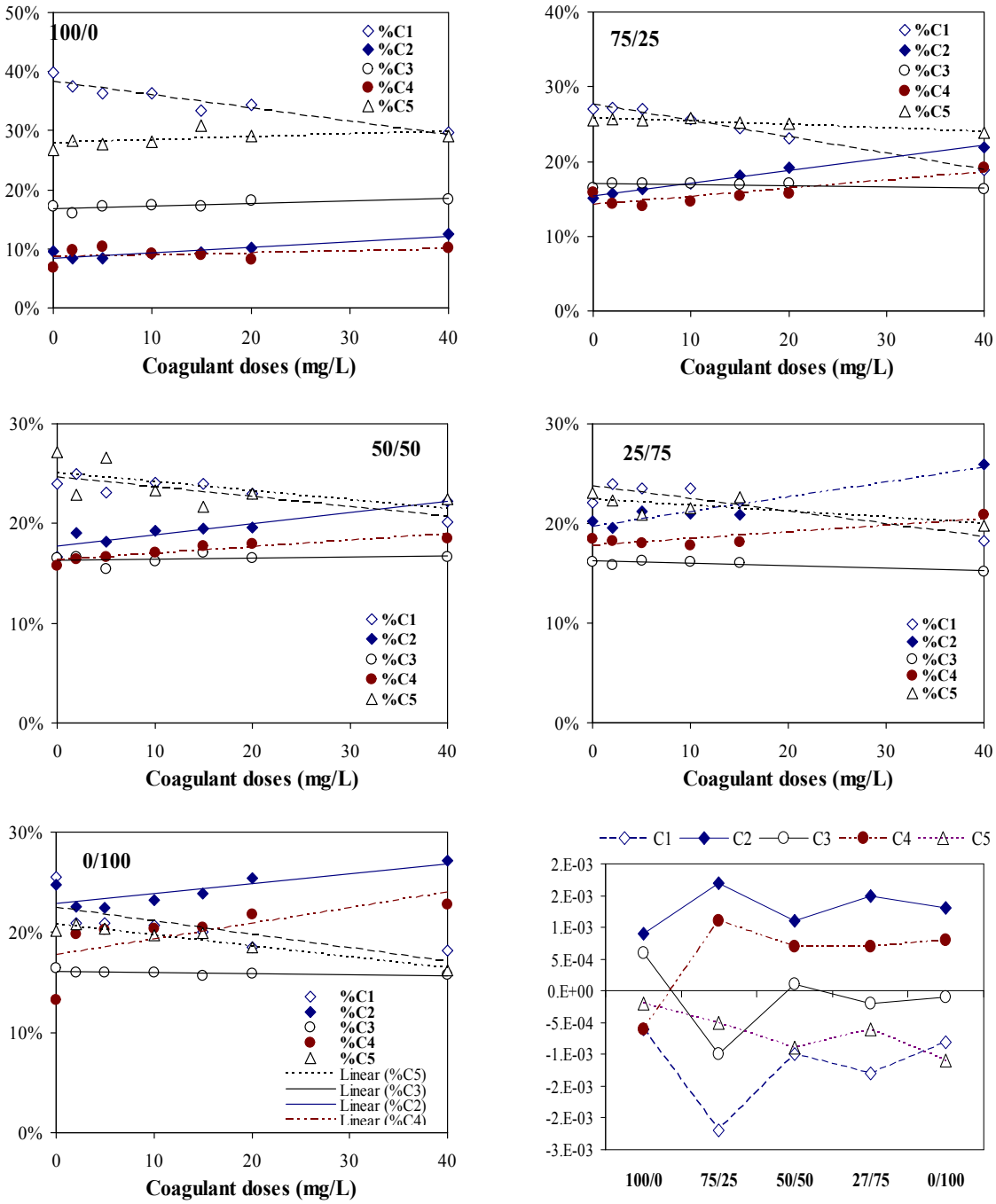


Figure 7-16 Fractional and concentration changes of each component for NOM/EfOM mixtures (NJ sites) for before coagulation and after coagulation; lower-right graph shows regression slopes at dose of 60 mg/L (left column: 100/0, and right column: 25/75 in %/%)

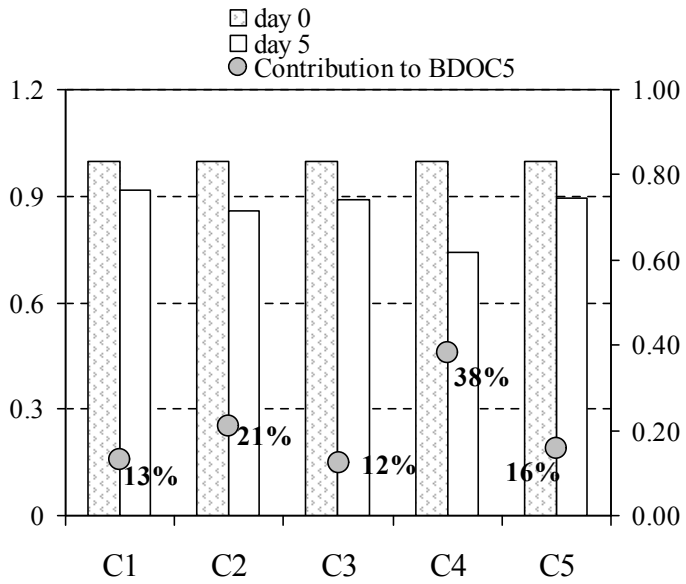


Figure 7-17 Contribution of each components to biodegradation ($BDOC_5$) for WW TP effluent from NJ site (NOM/EfOM=0%/100%)

Conclusion and implications

The following conclusions from the experiments can be drawn:

(1) With the same coagulant dose, DOC removals in EfOM were less efficient than NOM, showing increased FI values and decreased SUVA values after coagulation; these results were attributed to the higher hydrophilic character of EfOM (where the coagulation process removed hydrophobic organic matter more favorably) and due to an increase in coagulant demand.

(2) In coagulation, EEMs and SEC results showed significant disappearance of the humic-like peak, and preferential removal of higher MW and UV absorbing substances. In contrast, biodegradation of EfOM showed preferential removal of hydrophilic substances (significant disappearance of polysaccharide peak in SEC, and faster decrease of protein-like peak in EEM).

(3) ANOVA showed statistically that a change in the characteristics of organic matter by wastewater impact is a more important factor than pH.

(4) PARAFAC elucidated that component 1 and component 5 having hydrophobic properties

are dominant in NOM whereas component 2 and component 4 are dominant in EfOM and these are responsible for increased hydrophilic properties of EfOM. In coagulation, component 1 and component 5 were removed preferentially, while component 4 and component 2 showed a higher contribution to $BDOC_5$.

This study elucidates that EfOM has distinguishable properties that are different from NOM, and which may have a negative impact on drinking water treatment plants using conventional (coagulation) processes. Therefore, when EfOM-*impacted* waters are used as a indirect or direct potable source, other additional treatment processes or environmental buffers (e.g., biological filtration, river bank filtration, soil aquifer treatment, etc.) may need to be considered in order to remove hydrophilic components (proteins and polysaccharides). In addition, since the property of organic matter impacted by EfOM may be varied depending on wastewater, the treatability of EfOM may show significant differences and an understanding of source EfOM will play an important role in finding optimum treatment processes or conditions.

CHAPTER 8

CONCLUSIONS AND FUTURE RESEACH PERSPECTIVES

Conclusions

The study in chapter 3 was performed to identify wastewater fingerprinting signatures using fluorescence EEM combined with parallel factor analysis (PARAFAC). In this chapter was proposed the wastewater EfOM-focused 5 components PARAFAC model which was developed by using 423 EEMs from diverse water and wastewater sources along with wastewater impacted sites. Along with that, the model showed the fraction of component 2 or component 4 as wastewater fingerprint(s) evaluating the wastewater influence to the pertinent water. In a later part of this chapter, a case study estimating wastewater impacts into a watershed was presented and the result showed the availability of PARAFAC components as a way of tool to evaluate wastewater influence at the aspect of "EfOM amount", not merely by means of wastewater "flow".

Studies in chapter 4 were initiated with several hypotheses that some of 5 components proposed through the PARAFAC model (chapter 3) may be aggregates or a complex group of fluorophores, and if so, those components may be resolved into sub- or separated fluorophores (components) when they are fitted to the built-in 13 components PARAFAC model. The results showed likelihood to be 4 out of 5 components (except component 3) as combinations or aggregates merged with 2 or 3 fluorophores in the built-in 13 components' PARAFAC model.

Additional comparisons through further statistical analyses support the research hypotheses.

In chapter 5, EfOM and NOM were characterized by incorporations of multiple analytical techniques such as HPSEC-DOC-UVA-fluorescence detection, and fluorescence EEM. The results in this chapter provide in-depth approaches and understandings on properties of organic matters, especially at the MWD perspectives how MWDs of organic matters might be linked with or to spectroscopic parameters of UVA, SUVA, URI, and fluorescence-sensitivity.

Chapter 6 presented differentiation of EfOM versus NOM using multivariate statistical analyses such as Box-and-Whisker plots, principal component analysis (PCA) and multiple linear regression (MLR) analysis. Statistical analyses discriminated or classified EfOM and NOM by its hydrophobic or hydrophilic properties. MLR results showed hydrophilicity of EfOM as the most representatively distinguishable characteristic of EfOM differentiating from NOM. This chapter shows that HPI as the key discriminating parameter of EfOM from NOM, and organic matter with HPI or TPI properties as the most contributing to DOC of EfOM.

In chapter 7, a potent negative impact of EfOM onto the drinking water treatment processes was addressed in terms of DOC removal, explaining that the less treatability of wastewater-impacted water was mainly because of different characteristics between EfOM and NOM. As one way of strategies for dealing with the negative impact of wastewater, the biodegradation of EfOM was proposed, and RBF for drinking water and SAT for wastewater could be one of selections.

Suggestions and recommendations for future work

Results from this dissertation showed that EfOM, as complex mixtures of allochthonous and autochthonous organic substances, has different properties to NOM, which differences may

negatively affect the (conventional) drinking water treatment processes in terms of treatabilities.

As the concerns on the global warming and climate changes have been increasing, the future potential effects on the water quality would be required to be further described, since the more severe rainfall or drought events, or rises of water temperature happen, the more water reuse/reclamation for the potable purpose are demanded, and presumably the current understanding and discussions on characteristics and behaviours of wastewater-derived organics may also be changed.

Accordingly, at the increased wastewater reuse perspective for portable purpose, future study should be recommended to further understanding on microbial organic constituents such as extracellular polymeric substances (EPS) as one of important components of EfOM. Specific investigations of dissolved organic nitrogen, amino acids, polysaccharides, etc., may be recommended to better understanding.

Regarding DBP issues, wastewater EfOM may also have different characteristics, especially in reactivity with a disinfectant such as chlorine, resulting in different DBP formation trends or yields. Thus, the further study on the DOM fractions versus reactivity with disinfectants should be performed. Consequently, it will make it possible to evaluate which fractions of EfOM are the most contributing class to DBPs formation or must be removed in priority of potable purpose of wastewater reuse.

Meanwhile, since it is expected for synthetic organic compounds such as EDCs and PPCPs to be more important issues in water reuse, especially at the aspects of concentrations and types of those contaminants in WWTP effluents, therefore, it is recommended to include them in the study of EfOM understanding.

Nanoparticles (NPs), engineered NPs (e.g., TiO_2, SiO_2, etc.) that are derived from the human

activities, are becoming of concerns throughout the world. Little information on physical-chemical interactions of NPs with NOM or EfOM is available, and consequently very limited understanding on the impacts to water treatment processes, whether it would be adverse or benign. Interactions with existing organic matters of nanoparticles and/or their behaviours in water treatment processes are recommended to investigate.

REFERENCES OF CHAPTERS

1. *Standard Methods in the Examination of Water and Wastewater*. 20TH ed. 1998 APHA, AWWA and WEF: Washington, D.C.
2. Her, N., Amy, G.L., Foss, D., Cho, J, *Optimization Method for Detecting and Characterization NOM by HPLC - Size Exclusion Chromatography with UV and On-Line DOC detection.* Environmental Science and Technology, 2002. **36**(5): p. 1069-1076.
3. Lee, N., *Natural Organic Matter (NOM) Fouling of Low-pressure (MF and UF) membranes: Identification of Foulants, Fouling Mechanisms, and Evaluation of Pretreatment*. 2003, Ph.D. Dissertation, University of Colorado at Boulder: Boulder, CO.
4. Aiken, G.R., McKnight, D.M., Thorn, K.A., Thurman, E.M., *Isolation of hydrophilic organic acids from water using nonionic macroporous resins.* Organic Geochemistry, 1992. **18**(4): p. 567-573.
5. Aiken, G.R., *Chloride interference in the analysis of dissolved organic carbon by the wet oxidation method.* Environmental Science and Technology, 1992. **26**(12): p. 2435-2439.
6. Westerhoff, P., Wen Chen, W., Esparza, M., *Organic Compounds in the Environment: Fluorescence Analysis of a Standard Fulvic Acid and Tertiary Treated Wastewater.* Journal of Environmental Quality, 2001. **30**: p. 2038-2046.
7. McKnight, D.M., Boyer, E.W., Westerhoff, P.K., Doran, P.T., Kulbe, T., Andersen, D.T., *Spectrofluorometric characterisation of dissolved organic matter for indication of precursor organic material and aromaticity.* Limnology Oceanography, 2001. **46**: p. 38-48.
8. Zsolnay, A., Baigar, E., Jimenez, M., Steinweg, B., Saccomandi, F., *Diffentiating with Fluorescence Spectroscopy the Source of Dissolved Organic Matter in Soils subjected to Drying.* Chemosphere, 1999. **34**(1): p. 45-50.
9. Cox, L., Celis, R., Hermosin, M.C., Cornejo, J., Zsolnay, A., Zeller, K., *Effect of Organic Amendments on Herbicide Sorption as Related to the Nature of the Dissolved Organic Matter.* Environmental Science and Technology, 2000. **34**(21): p. 4600-4605.
10. Ohno, T., *Fluorescence Inner-Filtering Correction for Determining the Humidification Index of Dissolved Organic Matter.* Environmental Science and Technology, 2002. **36**(4): p. 742-746.
11. Allgeier, S.C., Summers, R.S., Jacangelo, J.G., Hatcher, V.A., Moll, D.M., Hooper, S.M., Swertfeger, J.W., Green, R.B. . *A Simplified and Rapid Method for Biodegradable Dissolved Organic Carbon Measurement*. in *AWWA Water Quality Technology Conference*. 1996. Boston, MA.
12. Vogt, R.D., Akkanen, J., Andersen, D.O., Bruggemann, R., Chatterjee, B., Gjessing, E., Kukkonen, J.V.K., Larsen, H.E., Luster, J., Paul, A., Pflugmacher, S., Starr, M., Steinberg, C.E.W., Schmitt-Kopplin, P., Zsolnay, A., *Key site variables governing the functional characteristics of Dissolved Natural Organic Matter (DNOM) in Nordic forested catchments.* Aquatic Sciences, 2004. **66**: p. 195-210.
13. Lee, W., Westerhoff, P., *Dissolved Organic Nitrogen Measurement Using Dialysis Pretreatment.* Environmental Science and Technology, 2005. **39**(3): p. 879-884.
14. Lee, W., *Occurence, Molecular Weight and Treatability of Dissolved Organic Nitrogen*. 2005, Ph.D. Dissertation, Arizona State University: Tempe, Arizona.
15. *Standard Methods in the Examination of Water and Wastewater*. 20th ed. 1998, Washington, D.C.: APHA, AWWA and WEF.

16. Khiari, D.S., Ventura. R., Barrett, S.E. *Occurrence of iodo-trihalomethanes in drinking water*. in *217th American Chemical Society National Meeting, 21-25 March 1999, Washington, D.C.: American Chemical Society*. 1999. Anaheim, CA.
17. Baker, A., *Fluorescence Excitation-Emission Matrix Characterization of River Waters Impacted by a Tissue Mill Effluent.* Environmental Science and Technology, 2002. **36**(7): p. 1377-1382.
18. Mobed, J.J., Hemmingsen, S. L., Autry, J. L., McGown, L. B. , *Fluorescence characterization of IHSS humic substances: Total luminescence spectra with absorbance correction.* Environmental Science and Technology, 1996. **30**(10): p. 3061-3065.
19. Coble, P.G., *Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy.* Marine Chemistry, 1996. **51**(4): p. 325-346.
20. Santos, E.B.H., Filipe, O.M., Duarte, R.M.B.O., Pinto, H., Duarte, A.C., *Fluorescence as a Tool for Tracing the Organic Contamination from Pulp Mill Effluents in Surface Waters.* Acta Hydrochimica Hydrobiology, 2000. **28**(7): p. 364-371.
21. Yuan Yan, H.L., and M. L. Myrick, *Fluorescence Fingerprint of Waters: Excitation-Emission Matrix Spectroscopy as a Tracking Tool.* Applied Spectroscopy, 2000. **54**(10): p. 1539-1542.
22. Holbrook, D.R., Yen, J.H., Grizzard, T.J., *Characterizing natural organic material from the Occoquan Watershed (Northern Virginia, US) using fluorescence spectroscopy and PARAFAC.* Science of the Total Environment, 2006. **361**: p. 249-266.
23. Hall, G.J., Clow, K.E., Kenny, J.E., *Estuarial Fingerprinting through Multidimensional Fluorescence and multivariate analysis.* Environmental Science and Technology, 2005. **39**(19): p. 7560-7567.
24. Cory, R.M., McKnight, D.M., *Fluorescence spectroscopy reveals ubiquitous prescence of oxidized and reduced quinones in dissolved organic matter.* Environmental Science and Technology, 2006. **39**(21): p. 8142-8149.
25. Stedmon, C.A., Markager, S., Bro, R., *Tracing dissolved organic matter in aquatic environments using a new approach to fluorescence spectroscopy.* Marine Chemistry 2003. **82**: p. 239-254.
26. Stedmon, C.A., Markager, S., *Tracing the production and degradation of autochthonous fractions of dissolved organic matter by fluorescence analysis.* Limnology Oceanography, 2005. **50**(5): p. 1415-1426.
27. Christensen, J., Becker, M.E., Frederiksen, C.S., *Fluorescence spectroscopy and PARAFAC in the analysis of yogurt.* Chemometrics and Intelligent Laboratory Systems, 2005. **75**: p. 201-208.
28. JiJi, R.D., Andersson, G.G., Booksh, K.S., *Application of PARAFAC for calibration with excitation-emission fluorescence spectra of three classes of environmental pollutants.* Journal of Chemometrics, 2000. **14**: p. 171-185.
29. Beltran, J.L., Guiteras, J., Ferrer, R., *Parallel factor analysis of partially resolved chromatographic data: Determination of polycyclic aromatic hydrocarbons in water samples.* Journal of Chromatography A, 1998. **802**: p. 263-275.
30. Young, M.S., Uden, P.C., *Byproducts of the Aqueous Chlorination of Purines and Pyrimidines.* Environmental Science and Technology, 1994. **28**(9): p. 1755-1758.
31. Scully, F.E., Dean Howell Jr., G., Kravltz, R., Jewel, J.T., Hahn, V., Speed, M., *Proteins in Natural Waters and Their Relation to the Formation of Chlorinated Organics during Water Disinfection.* Environmental Science and Technology, 1988. **22**(5): p. 537-542.

32. Krasner, S.W., Westerhoff, P., Chen, B., Amy, G., Nam, S.N. *Contribution of Wastewater to DBP formation: Case-Study of an Effluent-Impacted River*. in *American Water Works Association Water Quality Technology Conference*. Nov. 6-10, 2005. Quebec City, Canada.
33. Hall, G.J., Kenny, J.E., *Estuarine water classification using EEM spectroscopy and PARAFAC-SIMCA.* Analytica Chimica Acta, 2007. **581**: p. 118-124.
34. Brereton, R.G., *Chemometrics : data analysis for the laboratory and chemical plant*. 2003, Chichester, West Sussex, England: John Wiley & Sons, Ltd.
35. Bro, R., *PARAFAC. Tutorial and applications.* Chemometrics and Intelligent Laboratory Systems, 1997. **38**: p. 149-171.
36. Smilde, A., Bro, R., Geladi, P., *Multi-way Analysis with Applications in the Chemical Sciences*. 2004, Chichester, England: John Wiley & Sons Ltd.
37. Lakowicz, J.R., *Principles of fluorescence spectroscopy*. 1983, New York: Plenum Press. xiv, 496.
38. Schulman, S.G., DI, Q.Q., Juchum, J., *Organic Chemistry Applications of Fluorescence Spectroscopy*, in *Encyclopedia of Spectroscopy and Spectrometry, Volume 2*, G. Tranter, Holmes, J., Lindon, J., Editor. 2000, Academic Press.
39. Stedmon, C.A., Markager, S., *Resolving the variability in dissolved organic matter fluorescence in a temperate estuary and its catchment using PARAFAC analysis.* Limnology Oceanography, 2005. **50**(2): p. 686-697.
40. Andersson, C.A., Rasmus Bro, *The N-way toolbox for MATLAB.* Chemometrics and Intelligent Laboratory Systems, 2002. **52**: p. 1-4.
41. Thurman, E.M., Malcolm, R.L., *Preparative isolation of aquatic humic substances.* Environmental Science and Technology, 1981. **15**: p. 463-466.
42. Huber, S.A., Frimmel, F.H., *Direct Gel Chromatographic Characterization and Quantification of Marine Dissolved Organic Carbon Using High-Sensitivity DOC Detection.* Environmental Science and Technology, 1994. **24**(6): p. 1194-1197.
43. McKnight, D.M., Boyer, E.W., Westerhoff, P.K., Doran, P.T., Kulbe, T., Andersen, D.T., *Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity.* Limnology Oceanography, 2001. **46**(1): p. 38-48.
44. Murphy, K.R., Ruiz, G.M., Dunsmuir, W.T., Waite, T.D., *Optimized Parameters for Fluorescence-Based Verification of Ballast Water Exchange by Ships.* Environmental Science and Technology, 2006. **40**(7): p. 2357-2362.
45. Leenheer, J.A., Croue, J.P., *Aquatic organic matter.* Environmental Science and Technology, 2003. **37**(1): p. 19A-26A.
46. Coble, P.G., Castillo, C.E.D., Avril, B., *Distribution and optical properies of CDOM in the Arabian Sea during the 1995 Southwest Monsoon.* Deep-Sea Research Part II, 1998. **45**: p. 2195-2223.
47. Parlanti, E., Worz, K., Geoffroy, L., Lamotte, M., *Dissolved organic matter fluorescence spectroscopy as a tool to estimate biological activity in a coastal zone submitted to anthropogenic inputs.* Organic Geochemistry, 2000. **31**: p. 1765-1781.
48. Yau, W.W., Kirkland, J.J., Bly, D.D., *Modern Size-Exclusion Liquid Chromatography, Practice of Gel Permeation and Gel Filtration Chromatography*. 1979, New York: John Wiley & Sons.

49. Ewald, M., Berger, P., Visser, S.A., *UV-Visible Absorbance and Fluorescence Properties of Fulvic Acids of Microbial Origin as Functions of their Molecular Weights.* Geoderma, 1988. **43**: p. 11-20.
50. Senesi, N., *Molecular and quantitative aspects of the chemistry of fulvic acid and its interactions with metal ions and organic chemicals: Part II. The fluorescence spectroscopy approach.* Analytica Chimica Acta, 1990. **232**: p. 77-106.
51. Sharma, A., Schulman, S.G., *Introduction to Fluorescence Spectroscopy*. 1999, New York: Wiley.
52. Birnbaum, L.S., Staskal, D.F., *Brominated Flame Retardants: Cause for Concern?* Environmental Health Perspectives, 2004. **112**(1): p. 9-17.
53. Sollars, C.J., Peters, C.J., Perry, R. *Bromide in urban runoff—water quality considerations*. in *Effects of Waste Disposal on Groundwater and Surface Water(Proceedings of the Exeter Symposium)*. July 1982.
54. Krasner, S.W., Westerhoff, P., Chen, B., Amy, G., Nam, S.-N., Chowdhury, Z. K., Sinha, S., Rittmann, B. E., *Contribution of Wastewater to DBP Formation*. 2008, AWWA Research Foundation (AwwaRF): Denver, Colorado., USA.
55. Zhou, J.L., Liu, R., Wilding, A., Hibberd, A., *Sorption of Selected Endocrine Disrupting Chemicals to Different Aquatic Colloids.* Environmental Science and Technology, 2007. **41**: p. 206-213.
56. Holbrook, R.D., Love, N.G., Novak, J.T., *Biological Wastewater Treatment and Estrogenic Endocrine Disrupting Compounds: Importance of Colloid Organic Carbon.* Practice Periodical of Hazardous, Tocic, and Radioactive Waste Management, 2003. **7**(4): p. 289-296.
57. Yu, Z., Huang, W., *Competitive Sorption between 17α-Ethinyl Estradiol and Naphthalene/Phenanthrene by Sediments.* Environmental Science and Technology, 2005. **39**(13): p. 4878-4885.
58. Osenbruck, K., Glaser, H.R., Knoller, K., Weise, S.M., Moder, M., Wennrich, R., *Sources and transport of selected organic micropollutants in urban groundwater underlying the city of Halle (Saale), Germany.* Water Research, 2007. **41**: p. 3259-3270.
59. Drewes, J.E., Reinhard, M., Fox, P., *Comparing microfiltration-reverse osmosis and soil-aquifer treatment for indirect potable reuse of water.* Water Research, 2003. **37**(15): p. 3612-3621.
60. Goel, S., Hozalski, R.M., Bouwer, E.J., *Biodegradation of NOM: effect of NOM source and ozone dose.* Journal of American Water Works Association, 1995. **87**(1): p. 90-105.
61. Edwards, M., Benjamin, M.M., *Transformation of NOM by Ozone and its Effect on Iron and Aluminum Solubility.* Journal of American Water Works Association, 1992. **84**(6): p. 56-66.
62. Chen, W., Westerhoff, P., Leenheer, J., A., Booksh, K., *Fluorescence Excitation-Emission Matrix Regional Integration to Quantify Spectra for Dissolved Organic Matter.* Environmental Science and Technology, 2003. **37**: p. 5701-5710.
63. Coble, P.G., Green, S.A., Blough, N.V., Gagosian, R.B., *Characterization of dissolved organic matter in the Black Sea by fluorescence spectroscopy.* Nature, 1990. **348**: p. 432-435.
64. Cory, R.M., *Redox and Photochemical Reactivity of Dissolved Organic Matter in Surface Waters*. 2005, Ph.D. Dissertation, University of Colorado, Boulder: Boulder, CO.

65. Andersson, C.M., Bro, R., *Practical aspects of PARAFAC modeling of fluorescence excitation-emission data.* Journal of Chemometrics, 2003. **17**: p. 200-215.
66. Mopper, K., Schultz, C.A., *Fluorescence as a possible tool for studying the nature and water column distribution of DOC components.* Marine Chemistry, 1993. **41**(1-3): p. 229-238.
67. Reynolds, D.M., Ahmad, S.R., *Rapid determination of biochemical oxygen demand values of wastewater using fluorescence technique.* Water Research, 1997. **31**: p. 2012-2018.
68. Manka, J., Rebhun, M., Mandelbaum, A., Bortinger, A., *Characterization of organics in secondary effluents.* Environmental Science and Technology, 1974. **8**(12): p. 1017-1020.
69. Amy, G.L., Bryant Jr., C.W., Belyani, M., *Molecular weight distribution of soluble organic matter in various secondary and tertiary effluents.* Water Science and Technology, 1987. **19**: p. 529-538.
70. Nam, S.-N., Krasner, S.W., Amy, G.L., *Differentiating Effluent Organic Matter (EfOM) from Natural Organic Matter (NOM): Impact of EfOM on Drinking Water Sources*. Advanced Environmental Monitoring. 2007, Heidelberg, Germany: Springer-Verlag GmbH.
71. Jolliffe, I.T., *Principal Component Analysis*. 2nd Edition ed. 2002: Springer-Verlag, New York.
72. Chin, Y.-P., Aiken, G., O'Loughlin, E., *Molecular Weight, Polydispersity, and Spectroscopic Properties of Aquatic Humic Substances.* Environmental Science and Technology, 1994. **28**(11): p. 1853-1858.
73. Drewes, J.E., Fox, P., *Behavior and characterization of residual organic compounds in wastewater used for indirect potable reuse.* Water Science and Technology, 1999. **40**(4-5): p. 391-398.
74. Schimtt, D., Taylor, H.E., Aiken, G.R., Roth, A., Frimmel, F.H., *Influence of Natural Organic Matter on the Adsorption of Metal Ions onto Clay Minerals* Environmental Science and Technology, 2002. **36**: p. 2932-2938.
75. Jarvis, P., Jefferson, B., Parsons, S.A., *How the Natural Organic Matter to Coagulant Ratio Impacts on Floc Structural Properties.* Environmental Science and Technology, 2005. **39**(22): p. 8919-8924.
76. Reckhow, D.A., Singer, P.C., Malcolm, R.L., *Chlorination of Humic Materials: Byproducts Formation and Chemical Interpretations.* Environmental Science and Technology, 1990. **24**(11): p. 1655-1664.
77. Schnoor, J.L., Nitzschke, J.L., Lucas, R.D., Veenstra, J.N., *Trihalomethane Yields as a Function of Precursor Molecular Weight.* Environmental Science and Technology, 1979. **13**(9): p. 1134-1138.
78. Vilge-Ritter, A., Maison, A., Boulange, T., Rybacki, D., Bottero, J.-Y., *Removal of Natural Organic Matter by Coagulation-Flocculation: A Pyrolysis-GC-MS Study.* Environmental Science and Technology, 1999. **33**(17): p. 3027-3032.
79. Westerhoff, P., Aiken, G., Amy, G., Debroux, J. , *Relationships between the structure of natural organic matter and its reactivity towards molecular ozone and hydroxyl radicals.* Water Research, 1999. **33**(10): p. 2265-2276.
80. Siddiqui, M.S., Amy, G.L., Murphy, B.D., *Ozone enhanced removal of natural organic matter from drinking water sources.* Water Research, 1997. **31**(12): p. 3098-3106.

81. Abbt-Braun, G., Lankes, U., Frimmel, F.H., *Structural characterization of aquatic humic substances-The need for a multiple method approach.* Aquatic Sciences, 2004. **66**: p. 151-170.
82. Gjessing, E.T., Alberts, J.J., Bruchet, A., Egeberg, P.K., Lydersen, E., McGown, L.B., Mobed, J.J., Mυnster, U., Pempkowiak, J., Perdue, M., Ratnawerra, H., Rybacki, D., Takacs, M., Abbt-Braun, G., *Multi-Method Characterization of Natural Organic Matter isolated from Water: Characterisation of Reverse Osmosis-Isolates from Water of Two Semi-Identical Dystrophic Lakes Basins in Norway.* Water Research, 1998. **32**(10): p. 3108-3124.
83. Jarusutthirak, C., Amy, G., Croue, J.-P., *Fouling Characteristics of Wastewater Effluent Organic Matter (EfOM) Isolates on NF and UF Membranes.* Desalination, 2002. **145**: p. 247.
84. Shon, H.K., Vigneswaran, S., Snyder, S.A., *Effluent Organic Matter (EfOM) in Wastewater Constituents, Effects, and Treatment.* Critical Reviews in Environmental Science and Technology, 2006. **36**: p. 327-374.
85. Frimmel, F.H., Abbt-Braun, G., *Basic characterization of reference NOM from Central Europe-similarities and differences.* Environmental International, 1999. **25**(2-3): p. 191-207.
86. Debroux, J.-F., *The physical-chemical and oxidant reactive properties of effluent organic matter (EfOM) intended for potable reuse* 1998, University of Colorado, Boulder. p. xviii, 227 leaves.
87. Drewes, J.E., Croue, J.-P., *New approaches for structural characterization of organic matter in drinking water and wastewater effluents.* Water Science and Technology: Water Supply, 2002. **2**(2): p. 1-10.
88. Her, N., Amy, G.L., Park, H.-R., Song, M., *Characterizing algogenic organic matter (AOM) and evaluating associated NF membrane fouling* Water Research, 2004. **38**(6): p. 1427-1438.
89. Her, N., *Identification and characterization of foulants and scalants on NF membrane*. 2002, Ph.D. Dissertation, University of Colorado at Boulder: Boulder, CO.
90. Perminova, I.V., Frimmel, Fritz. H., Kovalevskii, Dmitrii. V., Braun, Gudrun. Abbt., Kudryavtsev, Alexey. V., Hesse, Sebastian., *Development of a predictive model for calculation of molecular weight of humic substances.* Water Research, 1998. **32**(3): p. 872-881.
91. Savitky, A., Golay, M.J.E., *Smoothing and Differentiation of Data by Simplified Least Squares Procedures.* Analytical Chemistry, 1964. **36**(8): p. 1627-1637.
92. Dubin, P.L., *Aqueous size-exclusion chromatography*. Journal of Chromatogrphy Library. Vol. 40. 1988, Amsterdam: Elsevier.
93. Her, N., Amy, G., Foss, D., Cho, J., *Variations of Molecular Weight Estimation by HP-Size Exclusion Chromatogrphy with UVA versus Online DOC Detection.* Environmental Science and Technology, 2002. **36**: p. 3393-3399.
94. Jarusutthirak, C., Amy, G., *Role of Soluble Microbial Products (SMP) in Membrane Fouling and Flux Decline.* Environmental Science and Technology, 2006. **40**(3): p. 969-974.
95. Skoog, D.A., Leary, J.J., *Principles of Instrumental Analysis*. 4th ed. 1992: Saunders College Publishing. pp.150-170.

96. Vance, G.F., David, M.B., *Chemical characteristics and acidity of soluble organic substances from a northern hardwood forest floor.* Geochimica et Cosmochimica Acta, 1991. **55**: p. 3611-3625.
97. Gauglitz, G., Vo-Dinh, T., ed. *Handbook of Spectroscopy*. 2003, WILEY-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany.
98. Young, K.C., *Utilization of Natural Organic Matter (NOM) Substrates by Bacteria*. 2005, Ph.D. Dissertation, University of Notre Dame: Notre Dame, Indiana.
99. Zhou, Q., Cabaniss, S.E., Maurice, P., *Considerations in the the use of high-performance size exclusion chromatography (HPSEC) for determining molecular weights of aquatic humic substances.* Water Research, 2000. **34**(14): p. 3505-3514.
100. Hayase, K., Tsubota, H., *Sedimentary humic acid and fulvic acid as fluorescent organic materials.* Geochimica et Cosmochimica Acta, 1985. **49**: p. 159-163.
101. Specht, C.H., Frimmel, F.H., *Specific Interactions of Organic Substances in Size-Exclusion Chromatography.* Environmental Science and Technology, 2000. **34**(11): p. 2361-2366.
102. Coble, P.G., Castillo, C.E.D., Avril, B., *Distribution and optical properties of CDOM in the Arabian Sea during the 1995 Southwest Monsoon.* Deep-Sea Research Part II, 1998. **45**: p. 2195-2223.
103. Dubois, M., Gilles, K.A., Hamilton, J.K., Rebers, P.A., Smith, F., *Colorimetric Method for Determination of Sugars and Related Substances.* Analytical Chemistry, 1956. **28**(3): p. 350-356.
104. Nam, S.-N., Krasner, S.W., Amy, G.L., ed. *Differentiating Effluent Organic Matter (EfOM) from Natural Organic Matter (NOM): Impact of EfOM on Drinking Water Sources*. Advanced Environmental Monitoring, ed. Y.J. Kim, Platt, U. 2007, Chapter 20, p259-270, Springer-Verlag GmbH: Heidelberg, Germany.
105. Nam, S.-N., Krasner, S., Amy, G. *Relating Natural Organic Matter (NOM) and Effluent Organic Matter (EfOM) Properties to Disinfection By-Products (DBP) Formation*. in *American Water Works Association Annual Conference and Exposition*. June 11-15, 2006. San Antonio, Texas.
106. Hua, G., Reckhow, D.A., *Characterization of Disinfection Byproduct Precursors Based on Hydrophobicity and Molecular Size.* Environmental Science and Technology, 2007. **41**(9): p. 3309-3315.
107. Sirivedhin, T., Gray, K.A., *Part I. Identifying anthropogenic markers in surface waters influenced by treated effluents: a tool in potable water reuse* Water Research, 2005. **39**(6): p. 1154-1164.
108. Singh, K.P., Malik, A., Mohan, D., Sinha, S., Singh, V.K., *Chemometric data analysis of pollutants in wastewater-a case study.* Analytica Chimica Acta, 2005. **532**: p. 15-25.
109. Mendiguchia, C., Moreno, C., Galindo-Riano, M.D., Garcia-Vargas, M., *Using chemometric tools to assess anthropogenic effects in river water A case study: Guadalquivir River (Spain).* Analytica Chimica Acta, 2004. **515**: p. 143-149.
110. Mihailov, G., Simeonov, V., Nikolov, N., Mirinchev, G., *Multivariate statistical assessment of the pollution sources along the stream of Kamchia River, Bulgaria.* Water Science and Technology, 2005. **51**(11): p. 37-43.
111. Shine, J.P., Ika, R.V., Ford, T.E., *Multivariate Statistical Examination of Spatial and Temporal Patterns of Heavy Metal Contamination in New Bedford Harbor Marine Sediments.* Environmental Science and Technology, 1995. **29**(7): p. 1781-1788.

112. Tauler, R., Barcelo, D., Thurman, E.M., *Multivariate Correlation between Concentrations of Selected Herbicides and Derivatives in Outflows from Selected U.S. Midwestern Reservoirs.* Environmental Science and Technology, 2000. **34**(16): p. 3307-3314.
113. Vega, M., Pardo, R., Barrado, E., Deban, L., *Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis.* Water Research, 1998. **32**(12): p. 3581-3592.
114. Ouyang, Y., *Evaluation of river water quality monitoring stations by principal component analysis.* Water Research, 2005. **39**: p. 2621-2635.
115. Ouyang, Y., Nkedi-Kizza, P., Wu, Q.T., Shinde, D., Huang, C.H. , *Assessment of seasonal variations in surface water quality.* Water Research, 2006. **40**: p. 3800-3810.
116. Kowalkowski, T., Zbytniewski, R., Szpejna, J., Buszewski, B., *Application of chemometrics in river water classification.* Water Research, 2006. **40**: p. 744-752.
117. Lambrakis, N., Antonakos, A., Panagopoulos, G., *The use of multicomponent statistical analysis in hydrogeological environmental research.* Water Research, 2004. **38**: p. 1862-1872.
118. Astel, A., Biziuk, M., Przyjazny, A., Namiesnik, J., *Chemometrics in monitoring spatial and temporal variations in drinking water quality.* Water Research, 2006. **40**: p. 1706-1716.
119. Allgeier, S.C., Summers, R.S., Jacangelo, J.G., Hatcher, V.A., Moll, D.M., Hooper, S.M., Swertfeger, J.W., Green, R.B. *A Simplified and Rapid Method for Biodegradable Dissolved Organic Carbon Measurement*. in *AWWA Water Quality Technology Conference*. 1996. Boston, MA.
120. Lavine, B.K., *Clustering and Classification of Analytical Data*, in *Encyclopedia of Analytical Chemistry*, M.Ó. Robert A., Editor, John Wiley & Sons Ltd.: Chichester, UK.
121. Einax, J.W., Zwanziger, H.W., Geiβ, S., *Chemometrics in Environmental Analysis*. 1997: WILEY-VCH Verlag GmbH.
122. Cory, R.M., McKnight, D.M., *Fluorescence spectroscopy reveals ubiquitous presence of oxidized and reduced quinones in dissolved organic matter.* Environmental Science and Technology, 2006. **39**(21): p. 8142-8149.
123. Hunt, J.F., Ohno, T., *Characterization of fresh and decomposed dissolved organic matter using excitation-emission matrix fluorescence spectroscopy and multivariate analysis.* Journal of Agricultural and Food Chemistry, 2007. **55**(6): p. 2121-2128.
124. Vogel, F., Harf, J., Hug, A., Rohr, P.R.V., *The Mean Oxidation Number of Carbon (MOC)-A Useful Concept For Describing Oxidation Process.* Water Research, 2000. **34**(10): p. 2689-2702.
125. Chen, B., Nam, S.-N., Westerhoff, P., Krasner, S.W., Amy, G., *Fate of effluent organic matter and DBP precursors in an effluent-dominated river: a case study of wastewater impact on downstream water quality*. 2008, submitted to Water Research.
126. Crook, J., MacDonald, J.A., Trussell, R.R., *Potable use of reclaimed water.* Journal of American Water Works Association, 1999. **91**(8): p. 40-49.
127. Namkung, E., Rittmann, B.E., *Soluble Microbial Products (SMP) Formation Kinetics by Biofilms.* Water Research, 1986. **20**(6): p. 795-806.
128. Imai, A., Fukushima, T., Matsushige, K., Kim, Y., Choi, K., *Characterization of dissolved organic matter in effluents from wastewater treatment plants.* Water Research, 2002. **35**: p. 859-870.

129. Manka, J., Rebhun, M., *Organic groups and molecular weight distribution in tertiary effluents and renovated waters.* Water Research, 1982. **16**(4): p. 399-403.
130. Grady Jr., C.P.L., Kirsch, E.J., Koczwara, M.K., Trgovcich, B., Watt, R.D., *Molecular weight distributions in activated sludge effluent.* Water Research, 1984. **18**(239-246).
131. Chudoba, J., *Inhibitory effect of refractory organic compounds produced by activated sludge microorganisms on microbial activity and flocculation.* Water Research, 1985. **19**: p. 197-200.
132. Sharp, E.L., Parsons, S.A., Jefferson, B., *Seasonal variations in natural organic matter and its impact on coagulation in water treatment.* Science of the Total Environment, 2006. **363**: p. 183-194.
133. Edwards, G.A., Amirtharajah, A., *Removing color caused by humic acids.* Journal of American Water Works Association, 1985. **77**(3): p. 50-57.
134. Chadik, P.A., Amy, G.L., *Removing Trihalomethane Precursors from Various Natural Waters by Metal Coagulant.* Journal of American Water Works Association, 1983. **75**(10): p. 532-536.
135. Knocke, W.R., Conley, L., Van Benschoten, J.E., *Impact of dissolved organic carbon on the removal of iron during water treatment.* Water Research, 1992. **26**(11): p. 1515-1522.
136. Amirtharajah, A., O'Melia, C.R., *Coagulation processes: Destabilization, Mixing, and Flocculation*, in *Water Quality and Treatment: A Handbook of Community Water Supplies*, F.W. Pontius, Editor. 1990, McGraw-Hill: New York.
137. Krasner, S.W., Amy, G., *Jar-tests evaluations of enhanced coagulations.* Journal of American Water Works Association, 1995. **87**(10): p. 93-107.
138. Qin, J.-J., Oo, M.H., Kebre, K.A., Knops, F., Miller, P., *Impact of coagulation pH on enhanced removal of natural organic matter in treatment of reservoir water.* Separation and Purification Technology, 2006. **49**: p. 295-298.
139. Randtke, S.J., *Organic Contaminant Removal by Coagulation and Related Process Combinations.* Journal of American Water Works Association, 1988. **80**(5): p. 40-56.
140. Quanrud, D.M., Arnold, R.G., Wilson, L.G., Gordon, H.J., Graham, D.W., Amy, G.L., *Fate of organics during column studies of soil aquifer treatment.* Journal of Environmental Engineering, 1996. **122**(4): p. 314-321.
141. Drewes, J.E., Fox, P., *Fate of natural organic matter (NOM) during groundwater recharge using reclaimed water.* Water Science and Technology, 1999. **40**(9): p. 241-248.
142. Traina, S.J., Novak, J., Smeck, N.E., *An ultraviolet absorbance method of estimating the percent aromatic carbon content of humic acids.* Journal of Environmental Quality, 1990. **19**(1): p. 151-153.
143. Novak, J.M., Mills, G.L., Bertsch, P.M., *Estimating the percent aromatic carbon in soil and aquatic humic substances using ultraviolet absorbance spectrometry.* Journal of Environmental Quality, 1992. **21**(1): p. 144-147.
144. Croue, J.-P., Debroux, J.-F., Amy, G.L., Aiken, G.R., Leenheer, J.A., *Natural Organic Matter: Structural Characteristics and Reactive Properties. In Formation and Control of Disinfection By-Products in Drinking Water*, P. Singer, Editor. 2000, American Water Works Association Research Foundation: Denver, CO. p. 65-93.
145. Tambo, N., Kamei, T. , *Coagulation and flocculation on water quality matrix.* Water Science and Technology, 1998. **37**(10): p. 31-41.

146. Ratnaweera, H., Hiller, N., Bunse, U., *Comparison of the coagulation behavior of different Norwegian aquatic NOM sources.* Environmental International, 1999. **25**(2-3): p. 347-355.
147. Musikavong, C., Wattanachira, S., Marhaba, T.F., Pavasant P., *Reduction of organic matter and trihalomethane formation potential in reclaimed water from treated industrial estate wastewater by coagulation.* Journal of Hazardous Materials B, 2005. **127**: p. 58-67.

APPENDICES

This section provides supplementary appendices of the entire chapters in order to promote better understanding on the technical findings if necessary.

APPENDIX OF CHAPTER 3

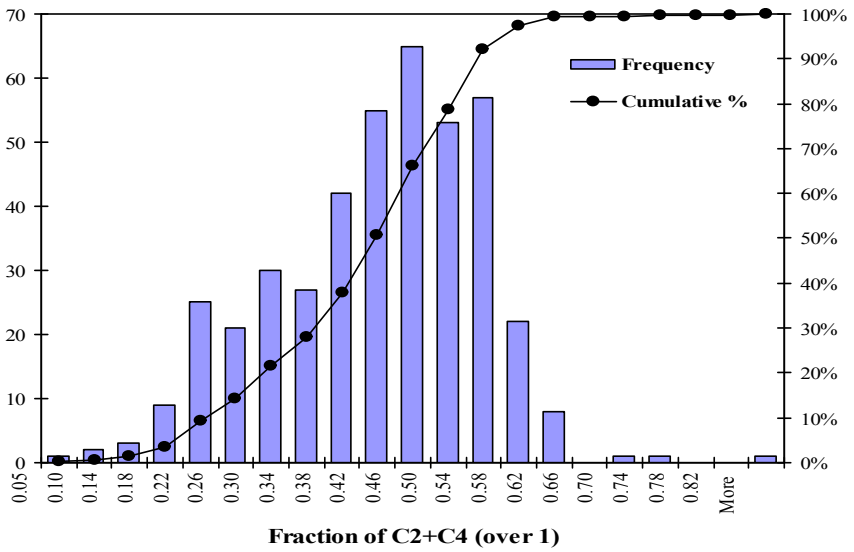


Figure 3A-1 Cumulative of sum of component 2 and component 4 fractions over 423 EEM samples

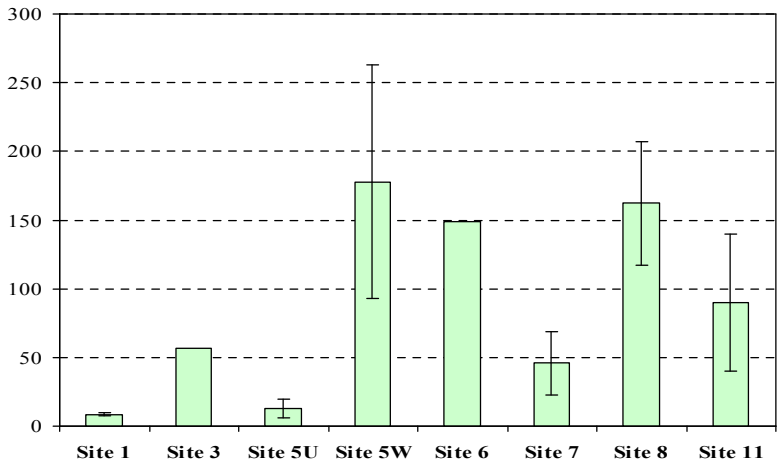


Figure 3A-2 Variations of primidone at each sampling sites in the South Platte watershed (plotted values are mean of three sampling campaign of Feb. and Sep. in 2004, April in 2005) – data provided by MWDSC (Stuart W. Krasner)

APPENDIX OF CHAPTER 4

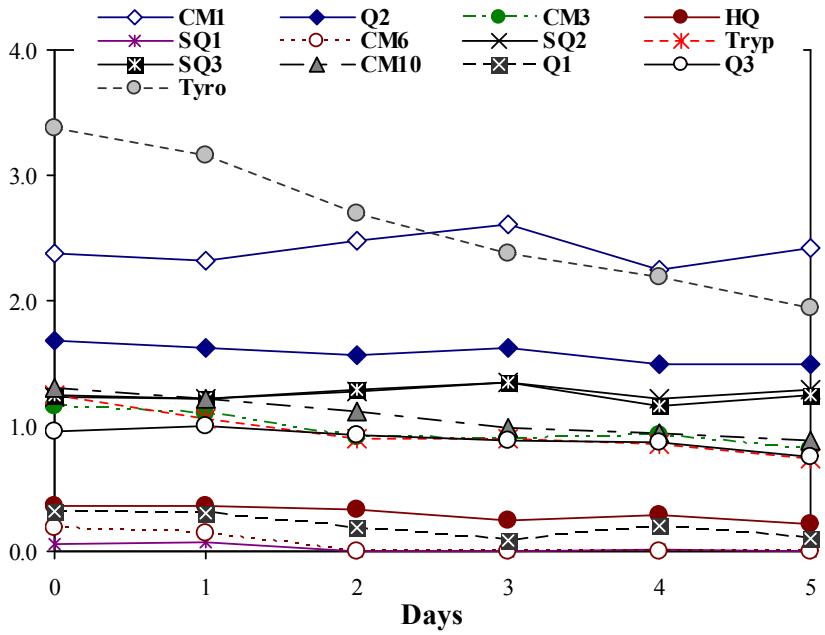


Figure 4A-1 DOC changes of C&M components by 5-day anaerobic biodegradation for the EfOM-*dominant* stream in the Southern East of US (taken in Feb, 2005)

APPENDIX OF CHAPTER 5

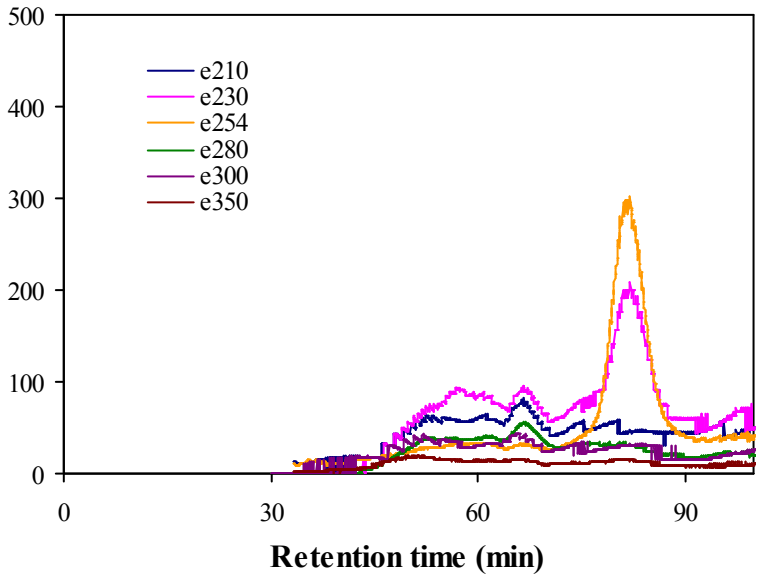


Figure 5A-1 Molar absorptivity chromatogram of EfOM; Chromatogram ahead of Rt ~30 min because the absorbance was low

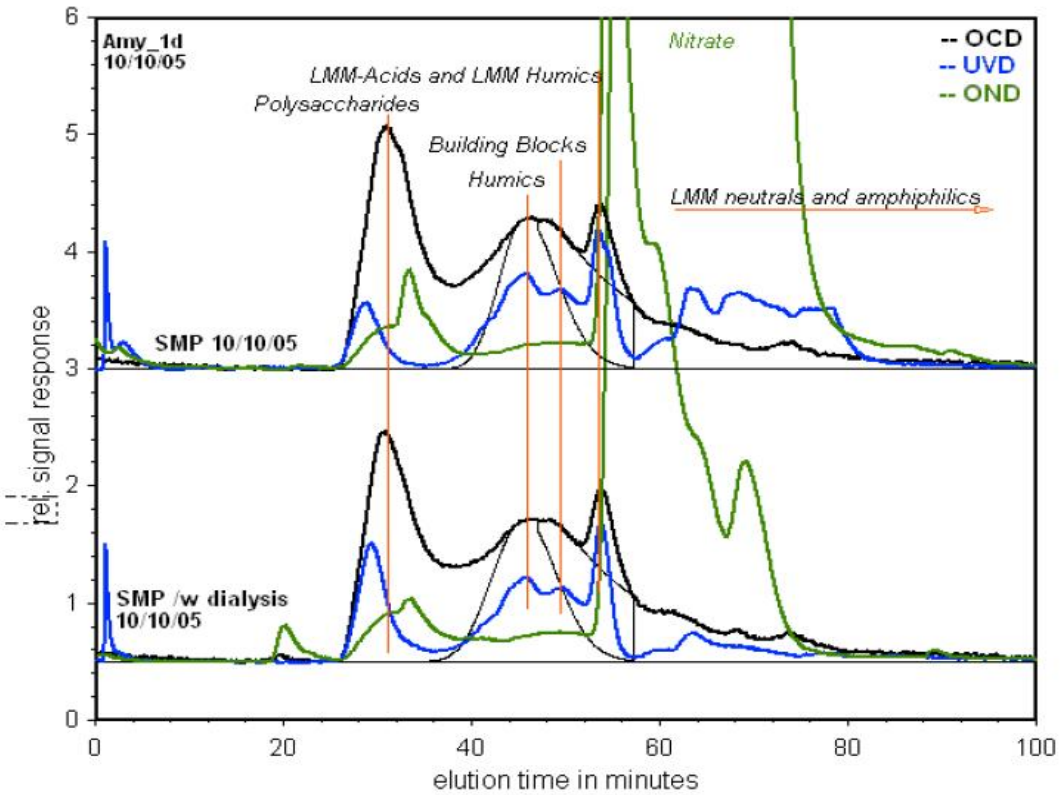


Figure 5A-2 LC-OCD/UVD/OND chromatograms of SMP gone through the 100 Da dialysis (measured in DOC-LABOR, Dr. Stefan Huber, in October, 2005): This SMP were from the same SBR to one shown in chapter 4, but the different sampling date may be slightly different chromatogram to one in chapter 4. Also, the detailed eluent compositions and buffer may cause the different elution chromatograph as well as retention time. The calibration of MW that the DOC-LABOR have applied at the time of measurement was $M=-0.292\times RT + 20.3271$ where M means molecular weight and RT is the elueted time (min), and this calibration is set-up using two references of IHSS Suwannee River humic and fulvic acid, therefore, the calibration is only valid for humic substances using this system

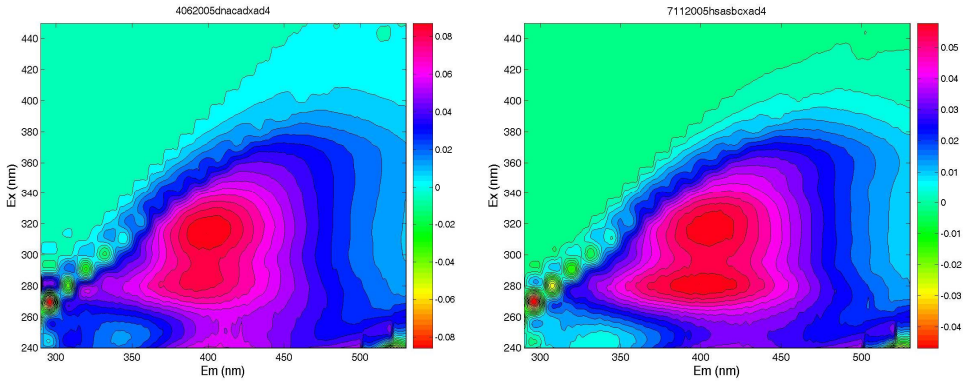


Figure 5A-3 EEMs of hydrophilic fraction of wastewater effluent (left: DN XAD4 and right: HS XAD4)

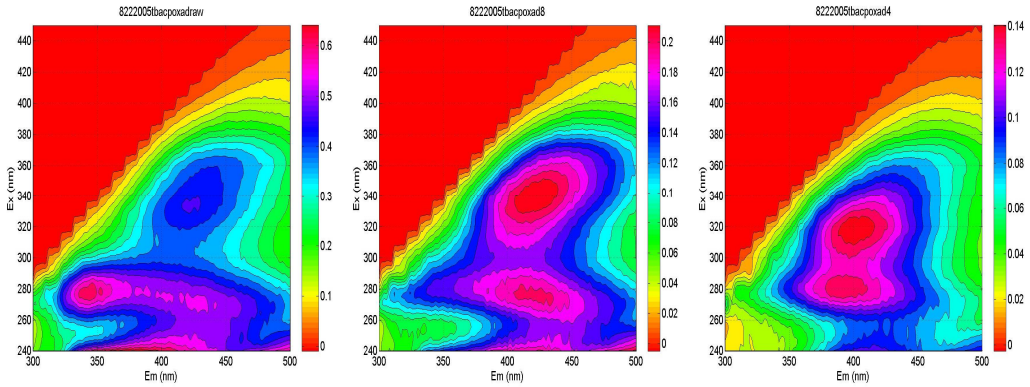


Figure 5A-4 EEMs of hydrophilic fraction of wastewater effluent (left: DN XAD4 and right: HS XAD4)

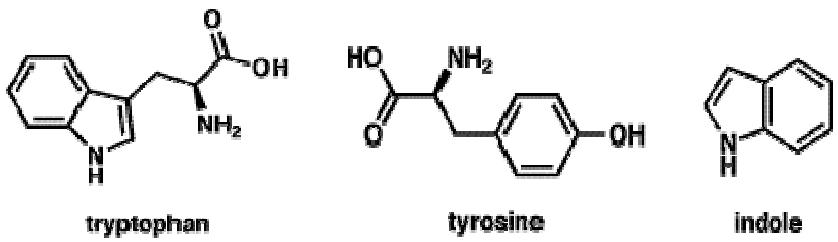


Figure 5A-4 Chemical structures of amino acids and indole moiety

APPENDIX OF CHAPTER 6

Table 6A-1 Summary of Pearson correlation coefficients among variables (values only for p <0.05)

	Upstream				WW impacted				WWTP			
	+		-		+		-		+		-	
TH	Alk	0.95			Alk	0.69			EC	0.7	DIN	-0.56
EC	TH	0.75			Alk	0.67			TH	0.7		
					TH	0.88						
TI	Alk	0.8			Alk	0.67	e280	-0.52				
	TH	0.76			TH	0.88						
Br-	Alk	0.81			Alk	0.78						
	TH	0.9			TH	0.65						
	EC	0.86			EC	0.68						
	TI	0.9			TI	0.86						
	COD	0.7			HIX	0.59						
	DON	0.73			Ph/Ar	0.57						
DIN					Alk	0.64						
					TH	0.51						
					EC	0.51						
					TI	0.52						
					COD	0.54						
DON	COD	0.76			COD	0.59			COD	0.66		
					DIN	0.52						
COD	EC	0.6			EC	0.64						
Ph/Ar	EC	0.69			EC	0.59			HIX	0.52		
	COD	0.84										
DOC	COD	0.72			COD	0.52			COD	0.67		
	DON	0.75			DON	0.76			DON	0.75		
HPO	e280	0.69										
	EC	0.59										
TPI			Ph/Ar	-0.67	DOC	0.5						
HPI	HIX	0.65	HPO	-0.8			HPO	-0.88			HPO	-0.8
			TPI	-0.66							TPI	-0.69
PS	Alk	0.58			DIN	0.54			HPI	0.48		
	TH	0.64										
HS							HIX	-0.51				
LMA							HS	-0.9			HS	-0.74
UV_{254}	e280	0.74			COD	0.51			COD	0.51		
	DOC	0.9			DIN	0.53			DOC	0.7		
					DON	0.8						
					DOC	0.89						
					PS	0.53						
SUVA	e280	0.95	TI	-0.77	e280	0.64	Alk	-0.54	HPO	0.58	DOC	-0.53
	HPO	0.71					TI	-0.61	HS	0.71	PS	-0.54
$BDOC_5$					Alk	1	TH	-1	COD	0.7	e280	-0.7
					EC	0.65			DON	0.58	SUVA	-0.54
					DIN	0.74			HIX	0.53		
					HIX	0.54			DOC	0.78		
					DOC	0.62			PS	0.71		

					PS	0.59			UV_{254}	0.77		
					UV_{254}	0.62						
PolyS	Alk	0.64			Alk	0.7						
	TH	0.69			TI	0.86						
	Ph/Ar	0.64			COD	0.51						
	PS	0.65			DOC	0.67						
	Br-	0.8			PS	0.51						
					UV_{254}	0.7						
FI	EC	0.63			Alk	0.64			Ph/Ar	0.51		
	Br-	0.94			TH	0.73						
					EC	0.77						
					TI	0.66						
					DIN	0.65						
					Ph/Ar	0.61						
%C1			TI	-0.86			EC	-0.51			HIX	-0.73
			HIX	-0.67			HIX	-0.78				
			HPI	-0.6			Ph/Ar	-0.51				
							Br-	-0.62				
							FI	-0.69				
%C2	COD	0.8	%C1	-0.61	Alk	0.56	%C1	-0.5			Br-	-0.83
	DON	0.72			DIN	0.68					UV_{254}	-0.59
	Ph/Ar	0.72			DOC	0.53						
	FI	0.65			PS	0.59						
					Br-	0.81						
					UV_{254}	0.54						
					$BDOC_5$	0.59						
%C3	e280	0.69	TI	-0.78					%C1	0.7	HIX	-0.51
	COD	0.83										
	Ph/Ar	0.93										
	DOC	0.67										
	SUVA	0.71										
	%C2	0.64										
%C4					Alk	0.63	HS	-0.53	DOC	0.5	SUVA	-0.5
					EC	0.55	%C1	-0.66	PS	0.54	%C1	-0.56
					DIN	0.61			$BDOC_5$	0.6	%C3	-0.8
					HIX	0.75						
					PS	0.51						
					$BDOC_5$	0.55						
					FI	0.67						
%C5			Ph/Ar	-0.6			DIN	-0.62			FI	-0.76
			DOC	-0.63			PS	-0.61			%C4	-0.54
			UV_{254}	-0.64			$BDOC_5$	-0.58				
			%C2	-0.73			%C2	-0.68				
			%C3	-0.72			%C4	-0.6				

APPENDIX OF CHAPTER 7

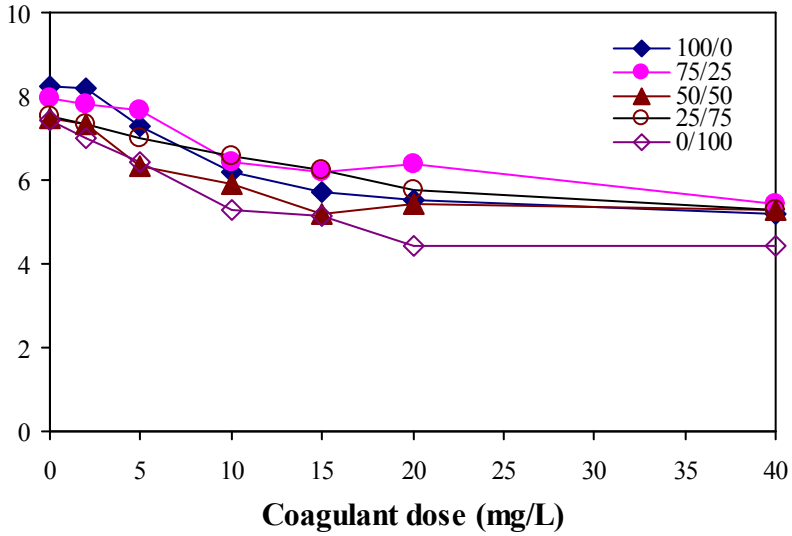


Figure 7A-1 Final pHs after coagulation for NJ sites

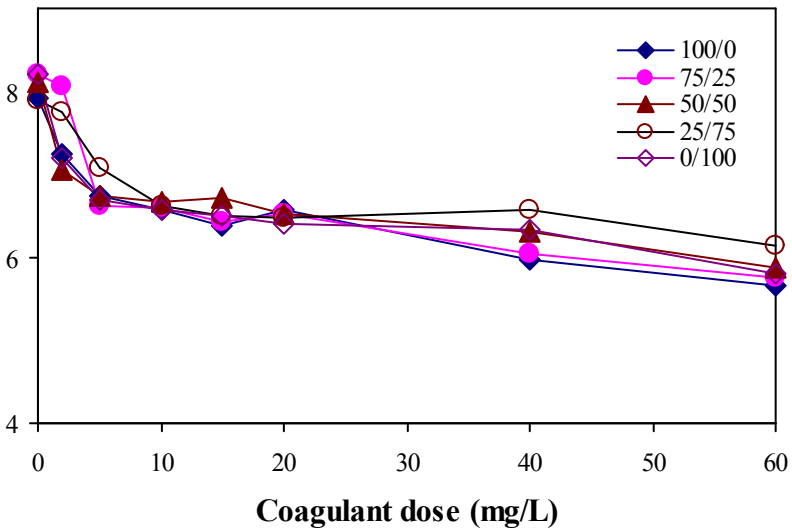


Figure 7A-2 Final pHs after coagulation for PA sites

Table 7A-1 pH changes by coagulation for NOM/EfOM mixtures of NJ and PA samples

	NOM/EfOM mixing ratio for NJ site (%/%)				
Dose (mg/L)	100/0	75/25	50/50	25/75	0/100
0	8.26	7.94	7.5	7.53	7.43
2	8.18	7.81	7.35	7.34	7.02
5	7.3	7.67	6.34	7.01	6.41
10	6.2	6.43	5.92	6.56	5.3
15	5.72	6.21	5.2	6.23	5.12
20	5.51	6.37	5.42	5.74	4.43
40	5.2	5.44	5.3	5.3	4.45
	NOM/EfOM mixing ratio for PA site (%/%)				
Dose (mg/L)	100/0	75/25	50/50	25/75	0/100
0	7.93	8.2	8.11	7.9	8.2
2	7.25	8.07	7.06	7.76	7.19
5	6.74	6.63	6.75	7.08	6.7
10	6.57	6.59	6.66	6.63	6.57
15	6.39	6.42	6.71	6.51	6.49
20	6.58	6.53	6.53	6.47	6.4
40	5.97	6.04	6.31	6.57	6.32
60	5.65	5.76	5.88	6.13	5.8

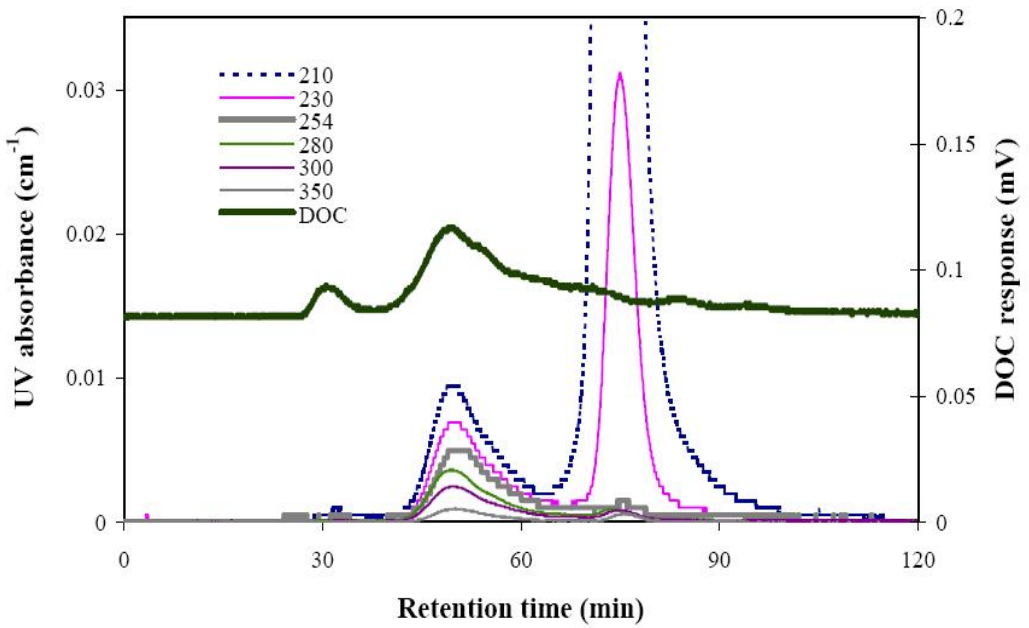


Figure 7A-3 HPSEC-DOC-UVA chromatogram of NOM (100/0 in %) for PA site; TSK HW-50S column (2×25 cm), Na_2SO_4 eluent with phosphate buffer (pH: 6.8, ionic strength: 0.1 M), flow rate 1 mL/min)