Development of Advanced Characterization Techniques for Organic and Colloidal Material of Relevance to Ultrafiltration Membranes in Drinking Water Treatment

by

Fei Chen

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Civil Engineering

Waterloo, Ontario, Canada, 2016

AUTHOR'S DECLARATION

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

STATEMENT OF CONTRIBUTIONS

For Chapter 3, the experimental work, including pilot-scale facility installation, facility operation, as well as sample collection and measurement, was completed by previous students and collaborators of the NSERC Chair in Water Treatment group at the University of Waterloo. Writing, analysis of the dataset, including FEEM-PCA, fouling rate estimation, all relevant statistical analyses and interpretation presented in this chapter, were conducted by the author.

For Chapter 4, the experimental work, including model solution preparation and FEEM measurements, was completed by Dr. Leila Munla who was a PhD student with the NSERC Chair in Water Treatment group at the University of Waterloo at the time. Writing, analysis of the data, including FEEM-PCA and all relevant statistical analyses and interpretation as presented in this chapter, were conducted by the author.

For Chapter 5, water samples were provided by Ishita Rahman who was a M.A.Sc. student with the NSERC Chair in Water Treatment at the University of Waterloo at the time. She designed, built, and operated the biofiltration-ultrafiltration set-up where the samples were taken from. FEEM measurements, all related data analysis, writing and interpretation were conducted by the author.

For Chapter 6, the biofilter media EPS extract samples were provided by Ahmed El-Hadidy, a PhD student with NSERC Chair in Water Treatment group at the University of Waterloo at the time. The rest of the experimental work, including the development of the fractionation approach, sample collection and measurements, as well as data analysis presented in this chapter were performed by the author.

For Chapter 7, the pilot-scale facility was redesigned, constructed and operated by Ahmed El-Hadidy, a PhD student, and Brad Wilson, a M.A.Sc. student, both with the NSERC Chair in Water Treatment group at the University of Waterloo. The biofilter media biomass monitoring was conducted by Ahmed El-Hadidy. The remainder of the experimental work, including sample collection port configuration, sample collection and measurements, as well as data analysis presented in this chapter, were conducted by the author.

Abstract

Ultrafiltration (UF) membranes, which can effectively remove colloidal/particulate matter (C/P), macromolecules, and pathogenic microorganisms, have gained increasing acceptance in the drinking water industry. However, organic fouling becomes a significant obstacle in achieving widespread adoption of this advanced treatment technology as it is a major contributor to operational and maintenance problems. Biofiltration without prior coagulation has been identified as a "green" and effective pre-treatment technique to mitigate organic fouling and increase the sustainability of ultrafiltration operation. To achieve the adoption and enhance the commercialization potential of biofiltration as a pre-treatment alternative, there is a need to identify and develop approaches to better characterize the natural organic matter (NOM) components that contribute to UF fouling and develop a fundamental understanding of the fouling mitigation mechanism of biofiltration as a new UF pre-treatment process.

For the first phase of this research, a pilot-scale study of drinking water UF membrane fouling illustrated the utility of liquid chromatography-organic carbon detection (LC-OCD) and fluorescence excitation emission matrices (FEEM) combined with principal component analysis (PCA) for the NOM-C/P characterization and fouling behaviour investigation. Biopolymers and protein-like substances were identified as the key contributors to both hydraulically reversible and irreversible fouling. However, the lower molecular weight (MW) NOM fraction, humic-like substances, were shown to contribute little to UF fouling at low fouling rates. In addition, the capability of colloidal/particulate matter to alleviate the hydraulically irreversible fouling was highlighted. This study also illustrated the complementary nature of LC-OCD (size exclusion based) and FEEM (functional group composition based) with regard to NOM-C/P characterization. To further explore the physical significance of extracted principal components from FEEM-PCA, a model compound investigation using humic acid, bovine serum albumin, sodium alginate, and silica was conducted. It was found that both intrinsic fluorescence properties of a particular NOM/inorganic colloid component and their interactions with other components were included within the physical significance of each principal component (PC). Then, the differences in the principles and utility of PCA and other FEEM data analysis methods were compared. It is recommended that PCA should be used for the investigation of the possible effects of the NOM component of interest whereas PARAFAC and peak picking are more promising for investigation of the characteristics of NOM

components. In addition, FEEM-PCA is capable of extracting information from Rayleigh scattering regions which could be related to colloidal/particulate matter and this unique utility was shown to not be affected by the photomultiplier tube voltage settings which can influence the saturation of the signal. Meanwhile, the compositional characteristics of biopolymers (defined by LC-OCD) was studied using a newly developed approach, LC-OCD based fractionation was combined with NOM characterization techniques. For three different biopolymer sources (biofilter media biofilm extracellular polymeric substance (EPS) extracts, a municipal wastewater, and a river water), the lower MW biopolymer sized fractions was found to have higher protein content per biopolymer carbon mass and the normalized organic nitrogen (ON) concentration was shown to be a reliable parameter to distinguish different sources. The behaviour and kinetics of NOM-C/P component removal during biofiltration was investigated at pilot-scale to understand the mitigation mechanisms of biofiltration as a pre-treatment alternative for UF. Biofiltration was demonstrated to be a promising process to remove protein-like materials and biopolymers which were identified as the key UF foulants. Comparatively, biofiltration had limited performance for the control of lower MW NOM fractions such as humic substances and building blocks. Either first or second order kinetics was shown to be appropriate to model the removal of biopolymers and protein-like materials during biofiltration. Also, because of the interaction between protein-like materials and colloidal/particulate matter, the 0.45 μm pre-filtration with polymeric filter prior to FEEM measurements resulted in a reduction of protein-like fluorescence intensities, with less of an effect on humic-like substances. Additionally, the differences of the removal behavior and kinetics of NOM components between filtered and non-filtered samples were observed.

In summary, biofiltration was demonstrated to be a reliable drinking water treatment process to remove key UF foulants and can therefore be recommended as a promising pre-treatment alternative for UF fouling control. LC-OCD and FEEM based techniques could be used in combination for NOM-C/P characterization and monitoring to better investigate the behavior of various NOM-C/P components. The selection of FEEM data analysis methods should be determined based on the purpose of investigation. FEEM measurements on both filtered and non-filtered samples are recommended to extract more information on colloidal/particulate NOM components. Further investigations should be conducted to explore the properties of other fluorescent NOM fractions using the newly developed fractionation approach and to achieve a more definitive interpretation on the removal kinetics of various NOM-C/P components during biofiltration.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisors Drs. Peter M. Huck, Raymond L. Legge, and Sigrid Peldszus for their indispensable guidance and continuous support throughout my entire PhD studies.

To my PhD committee, Dr. Paul Westerhoff (School of Sustainable Engineering and the Built Environment, Arizona State University), Dr. Xianshe Feng (Department of Chemical Engineering, University of Waterloo), Dr. Wayne Parker (Department of Civil and Environmental Engineering, University of Waterloo), and Dr. Bryan Tolson (Department of Civil and Environmental Engineering, University of Waterloo), I am extremely grateful for your insightful feedback and guidance during my PhD studies.

I would also like to express my warm thankfulness to Monica Tudorancea, Ahmed El-Hadidy, Brad Wilson, Jangchuk Tashi, Sylwia Kolaska, and Katarzyna (Kasia) Jaszczyszyn for their help and assistance during the pilot facility operation and lab analysis. I would also like to thank Ms. Dana Herriman for her valuable support. My appreciation also extends to my current and previous NSERC Chair group colleagues. Special thanks to Dr. Stefan Huber, Andreas Balz, and Nico Seeleib from DOC-LABOR for their support with LC-OCD operation and hardware modification.

The author was financially supported by the Ontario Graduate Scholarship (OGS) during part of the PhD studies. Funding for this research was provided by the Ontario Research Fund (ORF) Research Excellence Program and the Natural Sciences and Engineering Research Council (NSERC) in the form of an Industrial Research Chair in Water Treatment at the University of Waterloo. Current Chair partners are listed at: http://www.civil.uwaterloo.ca/watertreatment/pandp/partners.asp.

Finally, I would like to express my sincerest appreciation for the unconditional and absolute love and support of my parents. Also, I would like to thank my wife for keeping up with the busy life of a PhD student and standing beside me throughout my career.

Dedication

I dedicate this work to my parents and wife.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AIC	Akaike information criterion
AMWD	Apparent molecular weight distribution
AOC	Assimilable organic carbon
ATP	Adenosine triphosphate
ATR-FTIR	Attenuated total reflection Fourier transform infrared
BB	Building blocks
BIC	Bayesian information criterion
BOM	Biodegradable organic matter
BW	Backwash
CER	Cation exchange resin
C/P	Colloidal/particulate matter
cTMP	Corrected transmembrane pressure
DBPs	Disinfection byproducts
DF	Dilution factor
DFA	Discriminant function analysis
DOC	Dissolved organic carbon
DOM	Dissolved organic matter
EBCT	Empty bed contact time
EfOM	Effluent organic matter
Em	Emission
EPS	Extracellular polymeric substances
ESI-FT-ICR-MS	Electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry
Ex	Excitation
FA	Fulvic acids
FEEM	Fluorescence excitation-emission matrix
FRI	Fluorescence regional integration
GAC	Granular activated carbon
GC/MS	Gas chromatography/mass spectrometry
GRW	Grand River water
HA	Humic acids

HOCs	Hydrophobic organic compounds
HPI	Hydrophilic
HPLC-SEC	High performance liquid chromatography-size exclusion chromatography
HPO	Hydrophobic
HPSEC	High performance size exclusion chromatography
HS	Humic substances
IRF	Irreversible fouling
LC-OCD	Liquid Chromatography – Organic Carbon Detection
MDL	Method detection limit
MLR	Multiple linear regression
MW	Molecular weight
MWCO	Molecular weight cut off
NMR	Nuclear magnetic resonance
NOM	Natural organic matter
O_3	Ozonation
OC	Organic carbon
OCD	Organic carbon detector
ON	Organic nitrogen
OND	Organic nitrogen detector
PA	Polyamide
PAC	Powdered activated carbon
PARAFAC	Parallel factor analysis
PBS	Phosphate buffer saline
PCs	Principal components
PCA	Principal component analysis
PES	Polyethersulfone
PMT	Photomultiplier tube
Protein-C/P	Protein-colloidal/particulate matter interaction
PS	Polysulfone
PVC	Polyvinyl chloride
PVDF	Polyvinylidene fluoride
RC	Regenerated cellulose

RF	Reversible fouling
RS	Rayleigh scattering
SAM	Self-assembled monolayer
SEC-OCD	Size-exclusion chromatography – organic carbon detector
SLC	Significance level criterion
SPR	Surface plasmon resonance
SSF	Slow sand filtration
SUVA	Specific UV absorbance
TEP	Transparent exopolymers
TMP	Transmembrane pressure
TOC	Total organic carbon
TRP	Trptophan
UF	Ultrafiltration
UV	Ultraviolet
UVD	UV (at 254 nm) detector
UVLS	Ultraviolet/visible light spectroscopy
VS	Volatile solids
WTP	Water treatment plant
XAD	Weak-base ion exchange

Chapter 1
Introduction

1.1 Problem Statement

Ultrafiltration membranes have been widely applied in place of traditional granular media filtration in the drinking water industry. However, as one of the most significant operating and maintenance problems for ultrafiltration processes, membrane fouling hinders the further adoption of this advanced technique in drinking water treatment. To increase the operational sustainability of ultrafiltration membranes and expand their applicability to more challenging waters, several pre-treatment processes have been developed and investigated. A bench-scale investigation conducted by Hallé *et al*. (2009) demonstrated the utility of biofiltration as a "green"/chemical free and promising pre-treatment technique, which can effectively mitigate the organic fouling in ultrafiltration membranes used for drinking water treatment. The ability of biofiltration to reduce UF membrane fouling was further validated at pilot-scale using the same source water (i.e. Grand River water) by Peldszus *et al*. (2012). These new findings considerably extend the current practice of biofiltration and provide a new pre-treatment alternative for the application of ultrafiltration. However, there are currently no guidelines for biofiltration design for UF fouling mitigation and no tools or approaches for performance monitoring and control of the biofiltration-ultrafiltration process.

Given the importance of various natural organic matter (NOM) fractions and colloidal/particulate matter (C/P) for fouling of ultrafiltration membranes (Lee *et al*., 2006; Jermann *et al*., 2008; Hallé *et al*., 2009; Peiris *et al*., 2011; Peldszus *et al*., 2011), the quantification and characterization of these potential foulants at the low concentrations encountered in surface water become the major obstacle in understanding the fouling mechanisms. The two most promising water matrix characterization techniques currently employed in studies of membrane fouling are liquid chromatography-organic carbon detection (LC-OCD) (Huber *et al*., 2011) and fluorescence excitation-emission matrices (FEEM) (Peiris *et al*., 2010a, b). FEEM can differentiate humic-like substances, protein-like substances, and colloidal/particulate matter whereas LC-OCD can be used to quantify biopolymers, humic substances (HS), building blocks of HS, and some lower molecular weight (MW) fractions such as low MW acids and neutrals.

There is an interest to identify/develop appropriate NOM-C/P characterization approaches to provide a fundamental understanding of NOM-C/P components that contribute to UF fouling and to further explore the fouling mitigation mechanisms of biofiltration as a new ultrafiltration pre-treatment alternative. These NOM-C/P techniques will contribute to a better understanding of ultrafiltration membrane fouling behaviour and of fouling mitigation mechanisms of biofiltration-ultrafiltration water treatment trains. This knowledge may form the basis for tools/approaches for biofilter design, and biofilter performance monitoring and optimization. Also, a further understanding and/or improvement of the current NOM-C/P characterization techniques could be important information to improve the design and operation of ultrafiltration membranes without pre-treatment or with coagulation pre-treatment in drinking water applications.

1.2 Research Objectives and Scope

The global goal of this research was to identify/develop the NOM-C/P characterization techniques to elucidate the contribution of NOM-C/P components to UF membrane fouling and to provide a fundamental understanding of key foulant(s) as well as of the fouling mitigation mechanisms of biofiltration as a pre-treatment technology for ultrafiltration membranes. The sub-objectives are summarized as follows:

- Review recent advancements in NOM-C/P characterization techniques used in ultrafiltration membrane fouling studies and the current understanding of NOM-C/P fouling behaviour.
- Evaluate the utility of existing characterization techniques and investigate the roles of NOM-C/P in ultrafiltration membrane fouling.
- Improve the understanding of these existing NOM-C/P techniques and develop new characterization techniques to provide more insight into key UF foulants
- Explore the removal behaviour of various NOM-C/P components during biofiltration and the mechanism of fouling mitigation by biofiltration as a pre-treatment alternative for membrane ultrafiltration processes using the improved and/or newly developed NOM-C/P characterization techniques.

1.3 Research Approach and Thesis Organization

This thesis is organized into eight chapters, including an Introduction (Chapter 1), Literature Review (Chapter 2), 5 results chapters (Chapters 3, 4, 5, 6, and 7) prepared in journal article format, and Conclusions (Chapter 8).

In Chapter 1, the research motivation, research objectives and scope as well as thesis structure are presented.

In Chapter 2, results of a comprehensive literature review of the fundamentals of NOM and the utility of NOM characterization techniques related to ultrafiltration fouling are presented. The fouling behavior of various NOM-C/P components in ultrafiltration is also reviewed and summarized. In addition, the fundamentals of biofiltration, a promising pre-treatment alternative for UF fouling mitigation, are presented.

In Chapter 3, data from a pilot-scale investigation of ultrafiltration fouling was used to evaluate the performance of FEEM combined with principal component analysis (PCA) and LC-OCD as tools for determining the fouling potential of NOM components in river water. Also, the contributions of various NOM-C/P components to ultrafiltration membrane fouling were investigated.

In Chapter 4, in order to improve the understanding of FEEM combined with PCA, the physical significance of principal components extracted from FEEMs of model solutions (humic acid, bovine serum albumin, sodium alginate, and silica) and their combinations was investigated.

In Chapter 5, the utility of various FEEM data analysis methods including peak picking, parallel factor analysis (PARAFAC) and PCA, were compared using natural river water as a source of the FEEM data and literature results. Recommendations for the selection of FEEM data analysis method for various scenarios were provided. The impact of photomultiplier tube (PMT) settings on the FEEM data analysis was also investigated.

In Chapter 6, an LC-OCD based size fractionation approach combined with other NOM fingerprinting techniques was developed to characterize the composition of the biopolymer fraction as defined by LC-OCD. In this chapter, biopolymers from biofilter media (anthracite and sand) extracellular polymeric substances (EPS) extracts, a municipal wastewater, and a river water were investigated.

In Chapter 7, a pilot-scale biofilter profiling study was conducted to investigate the behaviour of various NOM-C/P components during biofiltration and to understand the fouling mitigation mechanism of biofiltration as a UF pre-treatment alternative. Interpretation of these results was accomplished with an improved understanding of the various NOM-C/P components and of NOM characterization techniques. The concentration profiles of various NOM-C/P components along the depth of the pilot-scale biofilter as well as their removal kinetics during biofiltration were investigated.

Chapter 8, begins with a synthesis of the results, followed by a summary of the major conclusions and contributions from this research. Several recommendations are also made for future research. The structure of the thesis and relationship between the chapters are illustrated in Figure 1.1.

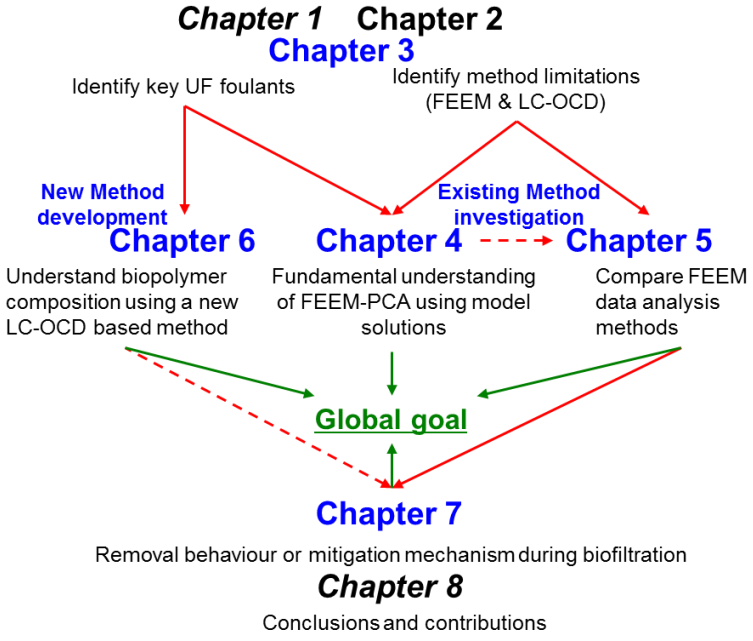


Figure 1.1 Thesis structure and relationship between chapters

Chapter 2
Literature Review

2.1 Introduction

The application of ultrafiltration (UF) membranes in place of traditional granular media filtration in drinking water treatment has gained considerable acceptance throughout the industry over the past decade. Ultrafiltration can effectively remove colloidal/particulate matter, organic macromolecules, pathogenic microorganisms (e.g. bacteria and viruses) in natural water with relatively low energy consumption (Aoustin *et al*., 2001; Jermann *et al*., 2008; Peter-Varbanets *et al*., 2011). Organic fouling is commonly recognized as the major obstacle to achieve further adoption of this advanced treatment technology in the drinking water industry (Amy, 2008; Krause and Obermayer, 2011). Two types of organic fouling were operationally defined according to backwashing practices, including hydraulically reversible and irreversible fouling (Amy, 2008; Peter-Varbanets *et al*., 2011). Both types of fouling can largely increase the maintenance cost and operational complexity and affect the productivity and product quality of the ultrafiltration processes (Zularisam *et al*., 2006). To maintain fluxes within the range of 50 – 100 $Lh^{-1}m^{-2}$, full-scale water treatment UF plants employ backwashing, disinfection and chemical cleaning to mitigate the membrane fouling (Peter-Varbanets *et al*., 2011). However, chemical cleaning for irreversible fouling reduction should be employed with minimum frequency because cleaning chemical reagents shorten membrane life and can result in disposal problems (Kimura *et al*., 2004).

Various natural organic matter (NOM) constituents, including biopolymers, humic substances (HS), protein-like and polysaccharide-like substances have been identified as potential foulants of ultrafiltration membranes via NOM-membrane interaction (Amy, 2008; Jermann *et al*., 2008; Peldszus *et al*., 2011; Rahman *et al*., 2014; Yamamura *et al*., 2014). NOM fractions can also interact with colloidal/particulate matter and divalent cations (e.g. Ca^{2+}), resulting in more complicated reversible and irreversible fouling behaviour (Jermann *et al*., 2007, 2008; Peiris *et al*., 2011; Peldszus *et al*., 2011).

To better understand the individual and combined behaviour of NOM constituents in the development of fouling layers, various NOM characterization techniques have been used in UF membrane fouling studies. However, due to the extremely complex composition of NOM, NOM characterization based

on individual constituents is unrealistic and therefore operationally defined chemical groups with similar chemical properties are used in NOM characterization and monitoring (Amy, 2008). Ion exchange (XAD) resin and membrane filtration using reverse osmosis and ultrafiltration were developed to fractionate the NOM constituents based on the hydrophobicity and molecular size of various NOM constituents, respectively (Kitis *et al*., 2002; Assemi *et al*., 2004; Matilainen *et al*., 2011). In recent years, some advanced techniques, including fluorescence spectroscopy, high performance size exclusion chromatography (HPSEC), and attenuated total reflection Fourier transform infrared spectrometry (ATR-FTIR), are gaining more acceptance for NOM characterization for membrane fouling investigations (Lee *et al*., 2006; Zularisam *et al*., 2007; Hallé *et al*., 2009; Peiris *et al*., 2010a, b; Zhao *et al*., 2011; Rahman *et al*., 2014; Shao *et al*., 2015; Siembida-Lösch *et al*., 2015).

To increase the operational sustainability of ultrafiltration membranes and expand their applicability to more challenging waters, several pre-treatment processes have been developed and investigated. Coagulation is one of the most commonly applied treatment technologies with respect to NOM removal and therefore has been widely practiced as a membrane pre-treatment alternative (Matilainen *et al*., 2010). A bench-scale investigation conducted by Hallé *et al*. (2009) demonstrated the utility of biofiltration (without any further pre-treatment) as a chemical free and promising pre-treatment technique, which can effectively mitigate the organic fouling in ultrafiltration membranes used for drinking water treatment. The ability of biofiltration to reduce UF membrane fouling was further validated at pilot-scale using the same source water (i.e. Grand River water) by Peldszus *et al*. (2012). These new findings from this novel application considerably extend the current practice of biofiltration and provide a new pre-treatment alternative for the application of ultrafiltration.

The objective of this chapter is to review the chemistry and fate of NOM in natural and drinking water systems. Advanced NOM characterization techniques which have been widely used in recent studies as well as their advantages over traditional NOM characterization techniques were summarized. Recent findings on membrane fouling mechanisms using advanced NOM characterization techniques were also comprehensively reviewed to identify the major NOM contributors to UF membrane fouling. Moreover, the advantages and disadvantages of advanced NOM characterization techniques as well as their potential for future application in membrane fouling

studies were evaluated. In addition, as a promising membrane pre-treatment alternative to coagulation, the fundamentals of biofiltration were also discussed in this review.

2.2 NOM Chemistry and Fate

2.2.1 Origin of NOM

NOM, which originates from the contact of dead and living organic matter with water throughout the hydrologic cycle (Egeberg *et al*., 2002), is a heterogeneous mixture of organic compounds (from largely aliphatic to highly colored aromatics) with different molecular weight (MW)/size and a broad spectrum of functional groups (e.g. aromatic, aliphatic, phenolic, and quinolic) (Baghoth *et al*., 2008; Lee *et al*., 2006; Matilainen *et al*., 2011; Stolpe *et al*., 2014; Zhou and Guo, 2015). Based on the origin of the NOM, it can be further classified into allochthonous and autochthonous NOM (Baghoth *et al*., 2008). Natural organic components with terrestrial origin such as terrestrial and vegetative debris is defined to be allochthonous NOM and organic matter with macrophite, algae, and microbial origin is defined to be autochthonous NOM, which is more biodegradable than the allochthonus form of the NOM (Leenheer and Croué, 2003; Lee *et al*., 2006; Zularisam *et al*., 2006; Baghoth *et al*., 2008). Original sources of the NOM, along with other three key factors, including physiochemical properties of water (e.g. temperature, ionic strength, pH, and major cation composition), surface chemistry of sediment sorbents, and the presence of other photoreactions and biogeochemical processes, will significantly affect the concentration, composition, and chemistry of the NOM (Leenheer and Croué, 2003).

2.2.2 Chemical Characteristics of NOM

Based on weak-base ion exchange (XAD) resin fractionation, aquatic NOM can be fractionated into three main categories, including hydrophobic, hydrophilic and transphilic fractions (Leenheer, 1981; Collins *et al*., 1986; Aiken *et al*., 1992; Lee *et al*., 2006; Zularisam *et al*., 2006). The hydrophobic fraction of NOM is mainly comprised of aromatic carbon with larger MW (e.g. phenolic structures and conjugated double bonds). Humic substances, including humic acids (HA) and fulvic acids (FA), are the main contributor to the aromatic NOM fraction (Aoustin *et al*., 2001). HA is more hydrophobic and has a lower solubility compared to FA (Aoustin *et al*., 2001). The hydrophilic fraction of NOM has more aliphatic carbon and nitrogenous compounds with lower MW (e.g. polysaccharide-like substances, carbohydrates, proteins, and amino acids) (Zularisam *et al*., 2007; Matilainen *et al*., 2011; Yamamura *et al*., 2014). The transphilic fraction of NOM is defined as

components with MW in between hydrophobic and hydrophilic fractions (Zularisam *et al*., 2007). Typically, hydrophobic, transphilic, and hydrophilic fraction represent 50%, 25%, and 25% of the dissolved organic carbon (DOC), respectively (Lee *et al*., 2006). Acids are dominant components within hydrophobic and transphilic fractions (Lee *et al*., 2006).

2.2.3 NOM Fate in Natural Aquatic and Drinking Water Systems

2.2.3.1 NOM Fate in Natural Aquatic System

As a significant carbon source and sink, NOM is ubiquitous in the natural aquatic systems such as surface waters, soil pore waters, and shallow groundwaters (Maurice *et al*., 2002; Leenheer and Croué, 2003). In natural environmental systems, NOM can act as carrier of heavy metals (Cabaniss and Shuman, 1988; Al-Sid-Cheikh *et al*., 2015), organic contaminants (Mazzei and Piccolo, 2015) and hydrophobic organic compounds (HOCs) (Murphy *et al*., 1990). Heavy metals (e.g. Cu) are mainly bound to NOM through complexation (Cabaniss and Shuman, 1988). Also, aquatic NOM can mediate natural biogeochemical and photoreaction processes which can regulate the physiochemical properties of surface waters such as levels of dissolved oxygen, oxidizing capacity, trace gas exchange, nutrient dynamics (e.g. nitrogen, phosphorus, sulfur), acidity and color (Gao and Zepp, 1998; Leenheer and Croué, 2003). Furthermore, previous studies reported that some microorganisms can utilize some NOM fractions (e.g. humic substances) as carbon and energy source (Sun *et al*., 1997; Hunt *et al*., 2000). The bioavailability of NOM is primarily dependent on the content of aliphatic carbon (Sun *et al*., 1997).

2.2.3.2 NOM Fate in Drinking Water Systems

NOM can cause aesthetic problems (e.g. odour, colour, and taste) of drinking water and generate carcinogenic disinfection byproducts (DBPs) by reacting with disinfectants/oxidants (e.g. chlorine) in traditional drinking water treatment processes (Aoustin *et al*., 2001; Kitis *et al*., 2002; Zheng *et al*., 2015; McKie *et al*., 2015). The dissolved and colloidal fractions of NOM are considered to be the major precursors of DBPs (Kitis *et al*., 2002). Zheng *et al*. (2015) reported that humics, building blocks and biopolymers could be important contributors to DBP formation. In addition, the presence of aquatic NOM can reduce the effectiveness of granular activated carbon and membrane filtration processes (Baghoth *et al*., 2011; Kennedy and Summers, 2015). Different components of NOM could be effectively removed by drinking water treatment processes such as coagulation (Matilainen *et al*., 2010) and biofiltration (Hallé *et al*., 2009).

2.3 NOM Isolation and Characterization Techniques

2.3.1 Traditional NOM Isolation and Characterization Techniques

2.3.1.1 Weak-base Ion Exchange (XAD) Resins

Traditionally, the most common technique for NOM constituent isolation are weak-base ion exchange (XAD) resins, which fractionate NOM into hydrophobic, transphilic, and hydrophilic fractions based on their affinities to different resins (Leenheer, 1981; Collins *et al*., 1986; Aiken *et al*., 1992; Carroll *et al*., 2000; Aoustin *et al*., 2001; Kitis *et al*., 2002; Matilainen *et al*., 2011). The International Humic Substances Society employed this technique as a standard method for fulvic acid (FA) and humic acid (HA) isolation (Matilainen *et al*., 2011). Two synthetic XAD resins were involved in this technique, including XAD-4 and XAD-8 (Collins *et al*., 1986; Aiken *et al*., 1992; Zularisam *et al*., 2006; Matilainen *et al*., 2011). Hydrophobic NOM fractions tend to adsorb onto XAD-8 resin and transphilic NOM fractions (weakly hydrophobic acid) are retained by XAD-4 (Aiken *et al*., 1992; Carroll *et al*., 2000). NOM fractions, which can pass through XAD-4 and XAD-8, are mainly hydrophilic (Carroll *et al*., 2000; Zularisam *et al*., 2006).

Even though XAD techniques have been widely accepted and employed in NOM characterization, various discrepancies and its time-consuming nature still limit the utility of this traditional technique. Matilainen *et al*. (2011) summarized the major disadvantages, including physiochemical alteration of NOM properties caused by extreme pH values and pH changes during fractionation, size-exclusion effects, contamination due to resin bleeding, and loss of NOM constituents because of irreversible adsorption onto resin. Moreover, the performance of the XAD resin is also significantly affected by operational conditions (Song *et al*., 2009).

2.3.1.2 Membrane Filtration

The most extensively used technique for aquatic NOM isolation is reverse osmosis (Maurice *et al*., 2002; Song *et al*., 2009; Matilainen *et al*., 2011). Ultrafiltration membranes (from 0.5 to 30 kDa) have also been used for NOM isolation and fractionation (Assemi *et al*., 2004). Membrane filtration was designed to fractionate NOM constituents based on their molecular weight/size (Assemi *et al*., 2004). However, the NOM fractionation using membrane filtration was demonstrated to be dependent on both molecular structure and size of NOM components (Assemi *et al*., 2004).

One of the major disadvantages is the alteration of NOM characteristics caused by changes in the inorganic component concentration and ionic strength during RO (Maurice *et al*., 2002; Song *et al*., 2009). Moreover, the performance of membrane filtration in NOM fractionation is affected by many other factors including membrane type, filtration scheme, pH and ionic strength of the solution, and initial concentration of the sample (Assemi *et al*., 2004). As a size selective separation technique, the user should be aware of these limitations when employing membrane filtration for NOM (especially humic substances) fractionation (Assemi *et al*., 2004).

2.3.1.3 Other General Parameters

Various general parameters, including total organic carbon (TOC)/dissolved organic carbon (DOC), UV_{254} absorbance, and specific UV absorbance (SUVA), have been used to estimate the overall amount of NOM constituents. However, these parameters cannot provide any specific information about the NOM composition and the amount for each fraction. TOC can be used to quantify the total amount of particulate, colloidal, and dissolved organic carbon. UV_{254} absorbance, which is associated with aromatic chromophores, can reflect the aromatic character of NOM constituents. SUVA, which is a ratio of UV_{254} absorbance and DOC, can reflect the aromaticity or hydrophobicity of NOM constituents (Weishaar *et al*., 2003). It should be noted that SUVA measurements are affected by various factors, including pH, nitrate, and iron (Weishaar *et al*., 2003). In addition, Filella (2014) summarized four NOM characterization methods for certain NOM fractions, including carbohydrates, thiols, transparent exopolymers (TEP), and humic substances, and concluded that results obtained from these methods are dependent on the standard used.

2.3.2 Advanced NOM Characterization Techniques

2.3.2.1 Liquid Chromatography – Organic Carbon Detection (LC-OCD)

2.3.2.1.1 Basics of LC-OCD

In drinking water treatment processes, especially ultrafiltration membrane filtration, the hydrophobicity and bioavailability of NOM constituents are significantly affected by molecular size or weight (Baghoth *et al*., 2011). The size-exclusion chromatography – organic carbon detector (SEC-OCD) technique was developed to separate NOM into various fractions based on molecular size or weight (Huber *et al*., 2011). LC-OCD, which is in essence a SEC-OCD technique employing the Grantzel thin-film UV-reactor as OCD, is a promising approach to fractionate NOM and quantify the

different fractions on the basis of their organic carbon content (Huber *et al*., 2011). The utility of this technique was further enhanced by combining it with an organic nitrogen detector (OND) and a UV (at 254 nm) detector (UVD) (Huber *et al*., 2011; Tran *et al*., 2015). The samples are pre-filtered using a flat polysulfone filter with a cut-off limit of 0.45 µm prior to chromatographic separation. More details regarding the design and physical description of the LC-OCD system can be found in Huber *et al*. (2011).

2.3.2.1.2 LC-OCD Data Interpretation

A customised software program (ChromCALC, DOC-LABOR, Karlsruhe, Germany) is typically used in conjunction with the LC-OCD to collect data and perform data analysis. ChromCALC firstly defines the boundaries between different NOM fractions through curve fitting using either Poisson or Exponential distributions and calculates the area units through the integration of peak areas under the pre-defined boundaries. The area units for different NOM fractions can then be converted to concentrations. Further details are described in DOC-Labor Dr. Huber (2010).

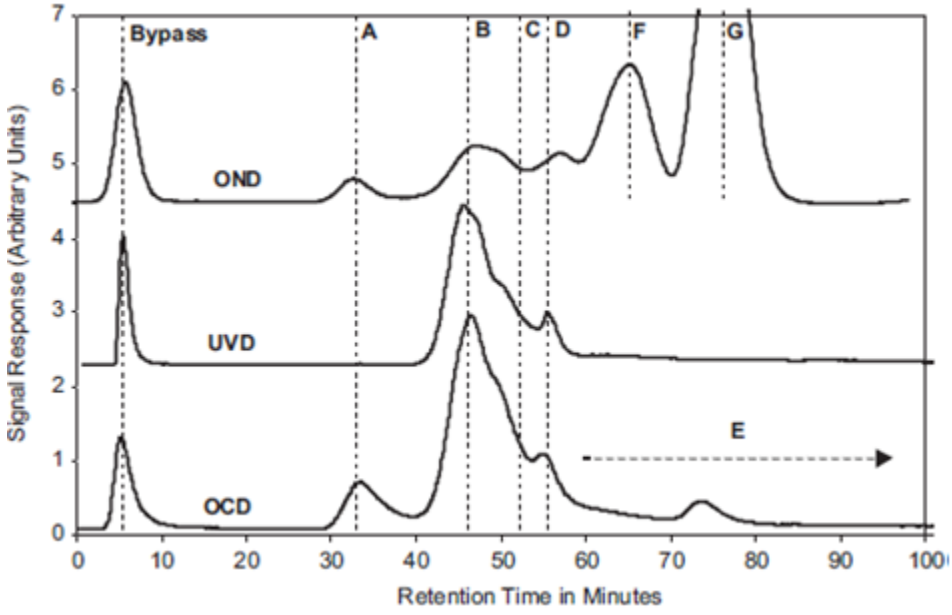


Figure 2.1 Chromatogram of a river water sample with organic carbon detector (OCD), UV (at 254 nm) detector (UVD), and organic nitrogen detector (OND) responses (Huber *et al*., 2011). A: biopolymer; B: humic substances; C: building blocks; D: low molecular weight acids; E: low molecular weight neutrals; F: nitrate (OND only); G: ammonium (OND only)

2.3.2.1.3 LC-OCD Application in NOM Characterization Related to UF Fouling Studies

As illustrated in Figure 2.1, LC-OCD can isolate and quantify five fractions, including biopolymers (e.g. polysaccharides, polypeptides, proteins, and amino sugars), humic substances (e.g. fulvic and humic acids), building blocks (e.g. hydrolysates of humic substances), low molecular-weight acids, and low molecular-weight neutrals (e.g. alcohols, aldehydes, ketones and amino acids) (Baghoth *et al*., 2011; Huber *et al*., 2011; Tran *et al*., 2015). LC-OCD has been successfully applied to characterize the NOM in various types of waters (e.g. surface water, groundwater, wastewater, storm water, seawater, and drinking water) (Baghoth *et al*., 2008; Hallé *et al*., 2009; Peter-Varbanets *et al*., 2011; Rahman *et al*., 2014; Siembida-Lösch *et al*., 2015).

2.3.2.2 Fluorescence Spectroscopy

2.3.2.2.1 Basics of Fluorescence Spectroscopy

Every atom or molecule can transit from a lower to an excited energy state through the absorption of a discrete quantum of light that has equivalent energy to the energy difference between the two energy levels (Senesi, 1990). The reverse transition from a higher to a lower energy level will result in the emission of radiant energy defined as luminescence, which can be further classified into fluorescence or phosphorescence according to the nature of excited state (Senesi, 1990; Lakowicz, 2006). Fluorescence refers to emission of a photon during the transition from the lowest vibrational level of the first excited singlet to various vibrational levels of the ground state (Senesi, 1990). Essentially, fluorescence is the reverse process of absorption by re-emitting the absorbed radiation by the molecule.

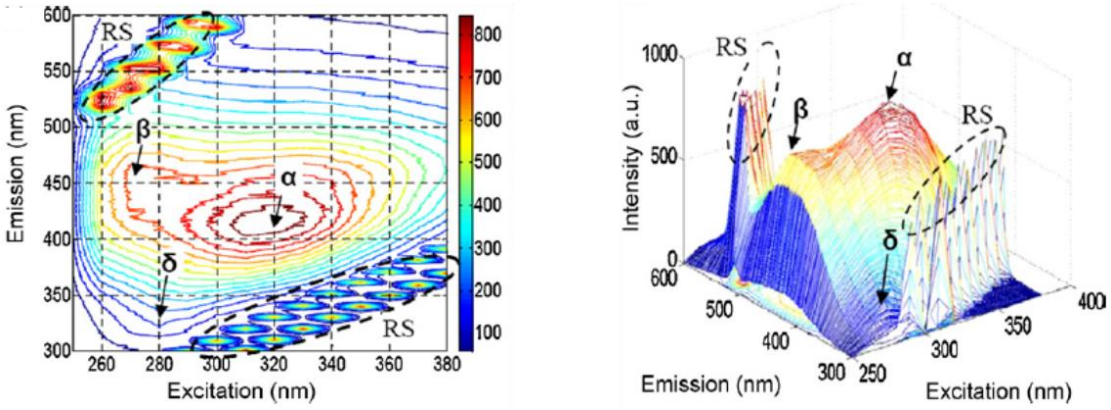
Fluorescence emission spectra are mainly dependent on the structural characteristics of the fluorophore and its surrounding environment (Lakowicz, 2006). Fluorophores can be mainly classified into two categories, including intrinsic and extrinsic fluorophores (Lakowicz, 2006). Intrinsic fluorophores are naturally occurring and include the aromatic amino acids, tryptophan, tyrosine, and phenylalanine (Lakowicz, 2006). Extrinsic fluorophores such as dansyl, fluorescein, and rhodamine, are used to provide fluorescence for samples by labeling the non- or low-fluorescent molecule or modify the spectral properties of the sample (Lakowicz, 2006).

Various factors such as concentration, cell path, absorptivity, and quantum yield will affect the fluorescence intensity at a given wavelength (Persson and Wedborg, 2001). Under the assumption that fluorophore interaction is negligible, the fluorescence intensity (F_{ij}) at a wavelength combination (excitation-λ_i. emission-λ_j) within an excitation emission matrix can be mathematically written as follows (Malinowsky, 1991; Boehme *et al*., 2004):

$$F_{ij} = \sum_{k=1}^{n} c_k \cdot X_{ik} \cdot Y_{kj}$$

where c_k is the concentration of the k^{th} fluorophore; X_{ik} represents the number of photons absorbed by the k^{th} fluorophore at a given excitation wavelength (λ_i); and Y_{kj} represents the fraction of emission produced by the k^{th} fluorophore at a given emission wavelength (λ_j).

Chen *et al*. (2003a) summarized three major fluorescence signal acquiring approaches, including emission scan at fixed excitation wavelengths, synchronous scan, and excitation-emission matrix (EEM). Fluorescence EEM is the most popular approach, which can provide more detailed information of both NOM and colloidal/particulate matter and improve the sensitivity of NOM characterization (Chen *et al*., 2003a; Henderson *et al*., 2009; Peiris *et al*., 2010a). Typical features of a 2D or 3D fluorescence EEM for river water are illustrated in Figure 2.2



(a) (b)

Figure 2.2 Typical (a) 2D and (b) 3D view of the fluorescence EEM for a river water (Peiris *et al*., 2010b). Peaks α and β correspond to humic-like substances. Peak δ and RS region refer to protein-like substances, and colloidal/particulate matter, respectively.

2.3.2.2.2 FEEM Data Interpretation

The traditional interpretation approaches for FEEM and quantification techniques are developed based on peak picking by using several peak positions from fluorescence spectra that may contain much more (e.g. 50 to > 10,000) wavelength-dependent fluorescence intensity data points (Chen *et al*., 2003a; Liu *et al*., 2011). Even though peak picking has been successfully applied in some cases, this technique cannot capture the heterogeneity of the different NOM fractions (Chen *et al*., 2003a). More importantly, the shift of fluorescence peak location with the compositional changes in the functional groups of NOM constituents has been widely reported (Senesi, 1990; Uyguner and Bekbolet, 2005; Lakowicz, 2006) and the fluorophore composition could be also very different for different NOM sources (Persson and Wedborg, 2001), supporting the importance of utilizing the full FEEMs (Chen *et al*., 2003a; Boehme *et al*., 2004; Peiris *et al*., 2010a) rather than individual peak positions.

Fluorescence regional integration (FRI) was developed by Chen *et al*. (2003a) and was successfully applied to interpret FEEM spectra for dissolved organic matter (DOM) and biomacromolecules (Chen *et al*., 2003a; Wang *et al*., 2009; Meng *et al*., 2011). In this approach, the volume beneath an EEM of a pre-determined region is integrated and normalized as the volumetric percentage of the full EEM spectra (Chen *et al*., 2003a). However, the determination of the regional boundaries requires modification according to the specific circumstances (Chen *et al*., 2003a). The key constraint to further application of this technique is the need for expert judgment of regional boundary delineation and interpretation of the physical meaning of each region.

The main challenges of utilizing the full fluorescence spectra come from the intrinsically multivariate nature of FEEMs, which are made of many variables (Legendre and Legendre, 1998). Multivariate statistical analysis techniques provide elegant means to explore and interpret all explanatory variables simultaneously, therefore preserving all the information from all dimensions of data structure and allowing for a more comprehensive data analysis (Stedmon *et al*., 2003). The application of a multivariate analysis instead of the peak-picking method can also avoid misinterpretation caused by variables that are not accounted for in the analysis and/or are biased by expert judgment.

The principal of multivariate analysis of three-way FEEMs (sample by excitation by emission) is to decompose the FEEMs of complex mixtures into their individual fluorescent components (Bro, 1997). The two most popular decomposition methods for three-way FEEMs include parallel factor analysis

(PARAFAC) (Bro, 1997; Kowalczuk *et al*., 2010; Baghoth *et al*., 2011; Zhang *et al*., 2011) and principal component analysis (PCA) (Persson and Wedborg, 2001; Peiris *et al*., 2010a; Peldszus *et al*., 2011).

Mathematically, PCA can decompose a two-way matrix (X) into bilinear components and a residual matrix (E) as follows (Eriksson *et al*., 2001):

$$X = \sum_{i=1}^{k} t_i \cdot p_i + E$$

Where k is the number of samples in the matrix X; t_i is the score vector of the principal components (PCs); and p_i is the corresponding loading vector.

PARAFAC can decompose a three-way matrix (Y) into trilinear components and a residual matrix (e_{ijk}) as follows (Bro, 1997):

$$Y_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} + e_{ijk};\ i = 1...I;\ j = 1...J;\ k = 1...K$$

Where F is the number of components; a_{if} represents the score matrix; b_{jf}, and c_{kf} are two loading matrices.

Prior to performing PCA, the three-way FEEMs are first unfolded into a two-way matrix and a variable is defined as each excitation-emission coordinate (Bro, 1997; Peiris *et al*., 2010a). Therefore, PCA has an intrinsic nature of a two-way decomposition method and FEEMs are decomposed into bilinear components with one score vector and one loading vector. As a three-way method, FEEM decomposition through PARAFAC will be made into trilinear components which consist of one score vector and two loading vectors (Bro, 1997).

Two-way PCA, which has more degrees of freedom and hence always fits data better, is a more complex and flexible model (Bro, 1997). Comparatively, PARAFAC is considered as a constrained version of PCA (Bro, 1997). Due to its uniqueness of solution, PARAFAC is capable of capturing the pure spectra of underlying fluorophores and considered to be superior to two-way PCA which suffers from intrinsic problem of rotational freedom (Bro, 1997; Stedmon *et al*., 2003; Bro, 2006). The solution uniqueness of PARAFAC also allows PARAFAC components to be used to estimate the real

fluorescent components identified by traditional physical chromatographic separations (Stedmon *et al*., 2003). Due to the intrinsic rotational freedom of two-way PCA, the physical significance of PCs must be verified by examining the similarity between folded loading matrices and original fluorescence spectra (Peiris *et al*., 2010a, b). Both techniques have shown to be promising to mathematically and uniquely identify fluorescent constituents without assumption of spectral shape and composition of fluorescent constituents.

2.3.2.2.3 FEEM Application in NOM Characterization Related to UF Fouling Studies

Another significant property for NOM is the composition of various structural and functional groups which can greatly affect the chemical, physical, biological, and polyelectrolytic properties of NOM constituents (Chen *et al*., 2002). Characterization of various functional groups will improve the understanding of the role of NOM constituents in complexation, adsorption, redox reactions, biodegradation, and immobilization processes, allowing for better fate and behaviour prediction of NOM constituents in drinking water treatment systems. Spectroscopic techniques can help extract valuable information with respect to the composition and characteristics of functional groups (Baghoth *et al*., 2011). FEEM-based techniques have been shown to be a promising technique to characterize NOM-colloidal/particulate matter (C/P) fractions from different sources such as natural surface water (e.g. river, lake), seawater, wastewater effluents, and effluents from drinking water treatment processes (e.g. coagulation, biofiltration. ozonation, sand filtration, ultrafiltration) (Baker, 2001; Persson and Wedborg, 2001; Chen *et al*., 2003a; Hudson *et al*., 2008; Wang *et al*., 2009; Peiris *et al*., 2010a, b; Baghoth *et al*., 2011; Shao *et al*., 2014; He and Hur, 2015; Lavonen *et al*., 2015). As illustrated in Figure 2.2 fluorescence spectroscopy using 3-dimensional FEEM can generally be used to identify various NOM fractions, including humic-like substances (e.g. terrestrial, anthropogenic, and marine origins), and protein-like substances (e.g. tryptophan-like and tyrosine-like) (Baker, 2001; Peiris *et al*., 2010a, 2011; Baghoth *et al*., 2011). In addition, Peiris *et al*. (2010a) also demonstrated that Rayleigh scattering (RS) regions in fluorescence spectra can be used to characterize the colloidal and particulate matter which have been identified as major membrane foulants (Jermann *et al*., 2008; Peiris *et al*., 2011).

2.3.2.2.4 Potential Application of Rayleigh Scattering in NOM Characterization

When a photon has insufficient energy and can only excite a molecule to a virtual energy state which is unstable, the molecule will quickly relax and emit light (Andrade-Eiroa *et al*., 2013). Rayleigh

scattering originates from this process and is an elastic scattering since the incident and emitted photons have entirely the same energy (Senesi, 1990; Andrade-Eiroa *et al*., 2013; Ham and MaHam, 2016). Compared with Rayleigh scattering which occurs due to the interaction between the molecules and excitation light, Tyndall scattering originates from the interaction between the particulate matter and excitation light (Zepp *et al*., 2004; Andrade-Eiroa *et al*., 2013). Since the light from both Rayleigh and Tyndall scattering has the same wavelength as the excitation light (Zepp *et al*., 2004; Andrade-Eiroa *et al*., 2013; Ham and MaHam, 2016), both forms of scattering are collected together with the fluorescence and hereinafter will be referred to as Rayleigh scattering. In addition to first order Rayleigh scattering which takes place when the excitation wavelength (Ex) is approximately equal to emission wavelength (Em), Rayleigh scattering can also take place when the Em is approximately equal to 2 times the Ex, which is referred to as second order scattering (Zepp *et al*., 2004). Rayleigh scattering is expected to linearly increase with the number of particles in the context of standard nephelometry (Ryan and Weber, 1982). Also, Rayleigh scattering can be affected by particle size which is related to the scattering light wavelength (Ryan and Weber, 1982). Compared to particles with smaller size, the efficiency of scattering is less for the larger particles (Ryan and Weber, 1982).

Due to the complexity of Rayleigh scattering, molecules, colloids and particles with different size ranges can impose an effect on this phenomenon. For natural organic matter in aquatic systems, the components can be operationally classified into dissolved and particulate organic matter based on 0.45 μm filtration (Aiken *et al*., 2011). Colloidal materials are defined to have a size range between dissolved and particulate matter (Aiken *et al*., 2011). The typical size range of colloidal materials is between 1 nm and 1μm and therefore they could be comprised of microparticles, macromolecules, and molecular aggregates (Buffle *et al*., 1998; Aiken *et al*., 2011). By definition, any inorganic or organic entity, which is sufficiently large to have supramolecular properties and sufficiently small to not to quickly gravity settle without further aggregation, could be considered as colloidal materials (Buffle *et al*., 1998).

According to the Rayleigh scattering mechanism, it is expected that inorganic colloidal/particulate matter can directly contribute to the Rayleigh scattering. The direct relationship between inorganic colloidal/particulate matter and Rayleigh scattering has been described in previous studies (e.g. Yoon and Lueptow, 2006). For particulate organic matter, Rayleigh scattering has been successfully utilized

previously to characterize and monitor precipitate or aggregate formation during complexation between NOM and various metal species (e.g. Ryan and Weber, 1982; Wu *et al*., 2004; Ohno *et al*., 2008). The dissolved organic matter (DOM) can further interact with inorganic colloidal/particulate matter (e.g. minerals) through physical and/or biogeochemical processes, which can further alter the heterogeneous precipitation of inorganic minerals (Aiken *et al*., 2011). In addition, the presence of DOM can also be involved in the aggregation and dissolution of colloid-size materials by modifying their surface chemistry (Aiken *et al*., 2011). These DOM involved processes will further influence the Rayleigh scattering efficiency and hence properties by changing the aggregate size. Elshereef *et al*. (2010) demonstrated that light scattering intensity can be directly affected by protein concentration. Therefore, information contained within the Rayleigh scattering is very complex and could be interpreted as an overall reflection of the characteristics of both inorganic and organic colloidal/particulate matter as well as the effects of their interactions with DOM and each other.

2.3.3 Other NOM Characterization Techniques for UF Fouling Studies

Attenuated total reflection Fourier transform infrared spectrometry (ATR-FTIR) can be used to characterize the functional chemistry of the NOM foulants which have deposited on the membrane (Howe *et al*., 2002; Kimura *et al*., 2004; Lee *et al*., 2006; Zularisam *et al*., 2007; Zhao *et al*., 2011). Humic substances, polysaccharide-like materials, protein-like substances, and silicon-containing minerals in natural waters (e.g. river and reservoir) were successfully identified using this technique (Zularisam *et al*., 2006; Zhao *et al*., 2011). The ATR-FTIR spectrum can be analyzed using double difference method developed by Chaufer *et al*. (2000). With respect to the membrane fouling investigation, a two-step procedure was suggested (Rabiller-Baudry *et al*., 2002). The residual water signal should be firstly removed from the recorded spectra and then the membrane contribution should be removed to underline the foulant contribution (Rabiller-Baudry *et al*., 2002). Compared to the LC-OCD and FEEM, ATR-FTIR has limited sensitivity in the identification of certain functional groups when broad overlapping bands occur due to the heterogeneity of natural water (Howe *et al*., 2002).

Flow field-flow fractionation was shown to be a very powerful tool for the characterization of different size groups of NOM (Stolpe *et al*., 2014; Zhou and Guo, 2015). Due to its wide applicability to samples with different status (e.g. liquid, gel and solid), nuclear magnetic resonance (NMR) spectroscopy could help understand the overall chemical composition of NOM and extract

information at the molecular level (Lam and Simpson, 2008; Mazzei and Piccolo, 2015). Similarly, electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (ESI-FT-ICR-MS) was also successfully applied to investigate the removal of dissolved organic matter with fluorophores/chromophores during the drinking water treatment at the molecular level (Lavonen *et al*., 2015).

2.4 Particle Retention and Theoretical Membrane Fouling Mechanisms

Various postulated particle retention mechanisms for dead-end mode filtration include cake filtration, deep bed filtration (or standard blocking), complete blocking, and intermediate blocking (Hermia, 1982; Huang *et al*., 2008; Tien and Ramarao, 2011). The deep-bed filtration mechanism refers to the constriction of membrane pores due to the deposition of penetrating particles with sufficiently small size (relative to the pore size). The complete blocking mechanism refers to the "sealing" of the pore openings due to the deposition of single particles with sizes greater than the pore size over individual pore openings, assuming that a cake layer is not formed and/or none of the particles are deposited onto the surface between the pores. The intermediate blocking mechanism refers to partial blocking of pore openings due to particle deposition and particle deposition on previously deposited particles. The cake filtration mechanism refers to the formation of a continuous cake layer with particles retained over the filter medium/surface. Compared to the other three pore blocking mechanisms, the hydraulic resistance is assumed to be generated by cake layer formation rather than pore structure changes (Huang *et al*., 2008). Although the above mechanisms have been used for fouling mechanism interpretation, cautions should be taken because membrane fouling is very complex process and these mechanisms were developed for particles. In addition, a resistance-in-series model has been developed to describe the permeate flux decline due to the increased resistance induced by various fouling phenomena such as solute adsorption and concentration polarization/gel layer formation (Chiang and Cheryan, 1986; Yeh and Cheng, 1993).

The hydraulically reversible fouling, which refers to the fouling that can be removed by hydraulic backwash, is usually accompanied by gel cake layer formation resulting from the particles retained on the membrane surface, which can be attributed to concentration polarization (Aoustin *et al*., 2001; Zularisam *et al*., 2006). The hydraulically irreversible fouling, which refers to the fouling that cannot be removed through hydraulic backwash and requires chemical cleaning, can generally be attributed to the foulant-membrane physicochemical interactions such as adsorption or pore plugging of

penetrating particles inside the membrane pore structure (Aoustin *et al*., 2001; Zularisam *et al*., 2006). The individual fouling mechanisms are more dependent on the specific foulant-membrane interactions in the initial stages and then more on the foulant-foulant interactions during the later stages of fouling (Huisman *et al*., 2000; Jermann *et al*., 2007). Jermann *et al*. (2007) also identified a transition from pore blocking to cake formation during the early stage of polysaccharide-induced reversible fouling. Furthermore, individual fouling mechanisms can mutually affect and interact with each other through hydrophobic and electrostatic interactions (Jermann *et al*., 2007).

Different fouling mechanisms can be triggered by the different NOM source characteristics (Zularisam *et al*., 2007). For a hydrophobic UF membrane, reservoir water with higher hydrophobic content (e.g. humic substances) tended to induce the cake layer deposition on a hollow-fiber membrane (68 kDa) whereas river water with a more hydrophilic characteristics was more likely to cause the foulant adsorption onto the membrane surface or within the pores (Zularisam *et al*., 2007). The characteristics of the membranes (e.g. type, pore size) can also affect the fouling mechanisms. Humic-like substances were not found in loosely attached foulants, implying that they tend to deposit onto the surface and/or inside the pores of a flat sheet membrane (20 – 100 kDa) (Jermann *et al*., 2007; Peiris *et al*., 2010b). However, Peldszus *et al*. (2011) reported that humic-like substances did not contribute to either reversible or irreversible fouling of hollow-fiber membrane with a molecular weight cut off (MWCO) of 400 kDa as they were passing through the membrane.

2.5 Application of Advanced NOM Characterization Techniques to Understand the UF membrane Fouling Mechanism

Individual and interactive behaviours of NOM fractions and colloidal/particulate matter in the fouling of different types of ultrafiltration membranes (e.g. hollow-fiber and flat-sheet) with different characteristics have been investigated (Aoustin *et al*., 2001; Kimura *et al*., 2004; Lee *et al*., 2006; Zularisam *et al*., 2007; Jermann *et al*., 2008; Hallé *et al*., 2009; Peiris *et al*., 2010b; Peldszus *et al*., 2011; Zhao *et al*., 2011). Findings regarding fouling behaviour as well as the relevant experimental conditions are summarized in Tables 2.1 and 2.2. The different characteristics of various types of membranes and their experimental conditions are reported to help interpret and understand fouling mechanisms. In addition, the differences of raw water quality and NOM characterization techniques are also reported.

The impacts of various factors (which have been included in Tables 2.1 and 2.2) on membrane fouling were discussed here. The membrane material can directly affect the hydrophobicity of the membrane and hence the extent of fouling (Lee *et al*., 2004; Jermann *et al*., 2007). Hydrophobic NOM fraction might cause more severe fouling of hydrophobic membrane through hydrophobic interactions compared to hydrophilic membrane dominated by adsorptive fouling (Lee *et al*., 2006). Additionally, PVDF and PES have more widespread application because they are chlorine and acid resistant (Huang *et al*., 2007). Full-scale water treatment plants predominantly employ hollow-fiber membranes rather than flat sheet membranes (Huang *et al*., 2007). Hydrodynamics of the membrane system plays an important role in membrane fouling (Jones and O'Melia, 2001; Huang *et al*., 2007). Compared to dead-end system, shear stress in cross-flow system can weaken the foulant-membrane and foulant-foulant interactions (Jermann *et al*., 2007). The operating flux, which is related to filtration pressure, can also significantly affect the fouling development through deforming the deposited foulants with different compressibilities, therefore changing the fouling reversibility (Zheng *et al*., 2010). Membrane fouling is expected to be more severe when permeate flux increases (Huang *et al*., 2007). NOM from different water sources (e.g. river, lake, EfOM) can trigger different particle retention/membrane fouling mechanisms and result in fouling layers with different characteristics (e.g. reversibility) (Huang *et al*., 2007; Zularisam *et al*., 2007). NOM from EfOM tend to cause greater total fouling of which the most part is more hydraulically reversible whereas allochthonous NOM will produce less total fouling but greater irreversible fouling (Huang *et al*., 2007). According to the information summarized in Tables 2.1 and 2.2, the individual and interactive behaviour of foulants were discussed together with various experimental factors in Sections 2.5.1 and 2.5.2.

Table 2.1 Summary of individual (indv.) and combined (com.) contribution (con.) of various NOM fractions to UF membrane fouling with relevant membrane characteristics and operating conditions – flat sheet membranes

Source	*Module* (θ: contact angle; ζ: zeta potential)				*Operation* (t: filtration time)		*Water Quality*			*Pre-trt.*	*NOM Char. Tech.*	*Irreversible fouling*		*Reversible fouling*	
	Type (Scale)	**Material (θ[a])**	**ζ[b] (pH)**	**MWCO (kDa)**	**TMP[c] (psi)**	**Mode (t: hrs)**	**Type (pH)**	**Turbidity (NTU)**	**DOC (mg/L)**			**Indv. Con.**	**Com. Con.[d]**	**Indv. Con.**	**Com. Con.[d]**
[1]	Flat sheet (Bench)	RC (HPI)	–16 (7)	100	Const. pressure 14.5	Dead end-s[e] (N/A)	Synthetic (7 – 8)	N/A	12.5	N/A	UVLS	HA	HA+Ca	FA; HPI	N/A
[2]	Flat sheet (Bench)	PA (45.5°)	– 17.0 (7)	8	Const. pressure 51	Dead end-s (N/A)	Synthetic & EfOM (5.7 – 7.3)	N/A	3.7 – 8.0	N/A	ATR-FTIR; HPSEC	Po; P	N/A	N/A	N/A
[3]	Flat sheet (Bench)	PES (HPO)	– 16.1 (5.4)	100	Const. pressure 7	Dead end (N/A)	Synthetic (7.5)	N/A	1 – 2	N/A	LC-OCD	HS	HS+Ca Po+Ca	Po	N/A
[4]	Flat sheet (Bench)	PES (HPO)	– 16.1 (5.4)	100	Const. pressure 7	Dead end (N/A)	Synthetic (7.5)	N/A	0.2 – 2	N/A	LC-OCD	HS	HS+C/P	C/P; Po	Po+C/P
[5]	Flat sheet (Bench)	PS/PES (72/80°)	N/A	20/60	Const. pressure 15	Cross flow (4.5)	River (7.8 – 8.4)	1.2–3.8	3.9 – 6.5	N/A	Fpca	P; HS	N/A	C/P	N/A
[6]	Flat sheet (Bench)	PES (HPI)	N/A	150	Const. pressure 14.5	Dead end-s (N/A)	Secondary effluent (6.6 – 7.5)	0.2–2.8	9.1 – 13.1	SSF	LC-OCD	BP	N/A	BP	N/A
[7]	Flat sheet (Bench)	PES (80°)	N/A	60	Const. pressure 15	Cross flow (4.5)	River (7.8 – 8.4)	1.2–3.8	3.9 – 6.5	N/A	Fpca	P; HS	HS+C/P P–C/P	C/P	N/A
[8]	Flat sheet (Bench)	PS (56.5°)	N/A	30	Const. pressure 29	Cross flow (200)	Reservoir (N/A)	N/A	0.97	N/A	ATR-FTIR Fpp	HS	N/A	P, HS, Po	N/A

[1]: Aoustin *et al*., 2001; [2]: Jarusutthirak and Amy, 2006; [3]: Jermann *et al*., 2007; [4]: Jermann *et al*., 2008; [5]: Peiris *et al*., 2010b; [6]: Zheng *et al*., 2010; [7]: Peiris *et al*., 2011; [8]: Zhao *et al*., 2011

Table 2.2 Summary of individual (indv.) and combined (com.) contribution (con.) of various NOM fractions to UF membrane fouling with relevant membrane characteristics and operating conditions – hollow fiber and other types of membranes

Source	*Module* (θ: contact angle; ζ: zeta potential)				*Operation* (t: filtration time)		*Water Quality*			*Pre-trt.*	*NOM Char. Tech.*	*Irreversible fouling*		*Reversible fouling*	
	Type (Scale)	**Material (θ[a])**	**ζ[b] (pH)**	**MWCO (kDa)**	**TMP[c] (psi)**	**Mode (t: day)**	**Type (pH)**	**Turbidity (NTU)**	**DOC (mg/L)**			**Indv. Con.**	**Com. Con.[d]**	**Indv. Con.**	**Com. Con.[d]**
[9]	Plate frame (Pilot)	PS (HPO)	N/A	750	Const. flux 8 – 21	Rotating membrane disk (149)	River (N/A)	1.8–15.4	1.4 – 2.4	N/A	Fpp; FTIR	Po; Fe; Mn	N/A	N/A	N/A
[10]	Hollow fiber (Bench)	PS (56°)	–27 (7)	68	Const. pressure 4.8	Dead end (120 min)	River/ Reservoir (~7.1)	N/A	6.7/ 8.04	N/A	ATR-FTIR	HPI	N/A	HS	N/A
[11]	Hollow fiber (Bench)	PVDF (HPO)	NEG (7)	400	Const. flux 57.5	Dead end (5)	River (8 – 8.4)	0.45–62	5 – 7	BF	LC-OCD	BP (cmp.)	N/A	BP (conc.)	N/A
[12]	Hollow fiber (Bench)	PVDF (HPO)	NEG (7)	400	Const. flux 57.5	Dead end (5)	River (8 – 8.4)	0.45–62	5 – 7	BF	Fpca	P	P+C/P	N/A	P+C/P
[13]	Hollow fiber (Bench)	PES (HPI)	N/A	150	Const. flux 50	Dead end (N/A)	EfOM (7.72±0.14)	N/A	6.8±0.9	O_3 PAC	LC-OCD (NP)	N/A	N/A	BP (conc.)	N/A

[9]: Kimura *et al*., 2004; [10]: Zularisam *et al*., 2007; [11]: Hallé *et al*., 2009; [12]: Peldszus *et al*., 2011; [13] Filloux *et al*., 2012

[a]If contact angle (θ) is not reported, the nature of membrane (HPO or HPI) will be reported.

[b]If zeta potential (ζ, mV) at pH = 7 is not reported, the surface charge of membrane (positive or negative) will be reported instead.

[c]If constant flux rather than constant pressure is applied, flux ($L/m^2/h$) is recorded.

[d]"+" refers to the synergistic relationship; "–" refers to the antagonistic relationship

[e]Dead end-s represents dead end mode with stirring

Membrane material:

PA: Polyamide; PES: Polyethersulfone; PS: Polysulfone; PVC: Polyvinyl chloride;
PVDF: Polyvinylidene fluoride; RC: Regenerated cellulose

NOM characterization (char.) techniques (tech.):

ATR-FTIR: Attenuated total reflection fourier transform infrared spectroscopy;
Fpca: FEEM-PCA; Fpp: FEEM peak picking;
HPSEC: High pressure size-exclusion chromatography;
LC-OCD (NP: No Pre-filtration): Liquid chromatography – organic carbon detector;
UVLS: Ultraviolet/visible light spectroscopy

Foulant names:

BP: Biopolymer; C/P: Colloidal and particulate matter; FA: Fulvic acids;
HA: Humic acid; HPI: Hydrophilic; HPO: Hydrophobic; HS: Humic substances;
P: Protein and/or protein-like substances;
Po: Polysaccharide and/or polysaccharide-like substances

Pre-treatment (trt.):

BF: Biofiltration; O_3: Ozonation; PCA: Powdered activated carbon; SSF: Slow sand filtration

Other abbreviations:

EfOM: Effluent organic matter; MWCO: Molecular weight cut off;
N/A: Not available; NEG: Negative

2.5.1 Individual Behaviour of Foulants

2.5.1.1 Biopolymers (including protein and polysaccharides)

Hallé *et al*. (2009) demonstrated that influent biopolymer concentration was related to the extent of hydraulically reversible fouling and biopolymer composition may have been critical for the level of hydraulically irreversible fouling for a PVDF hollow-fiber membrane. A strong exponential relationship between biopolymer content and reversible fouling of a hydrophilic hollow-fiber membrane was also revealed by Filloux *et al*. (2012). As an important component of the biopolymers, protein-like substances were shown to be a key contributor to the irreversible fouling for a hydrophobic hollow-fiber membrane with a MWCO of 400 kDa (Peldszus *et al*., 2011). For a tighter hydrophobic hollow-fiber membrane with a MWCO of 68 kDa, Zularisam *et al*. (2007) demonstrated that hydrophilic NOM fractions such as polysaccharide, polysaccharide-like substances, alcoholic compounds and aliphatic amide of protein groups are responsible for the membrane fouling. Furthermore, morphological analyses of membrane surface indicated that hydrophilic NOM fractions from river water tend to deposit onto the pore surface and cause irreversible fouling (Zularisam *et al*., 2007). Polysaccharides were also shown to be responsible for irreversible fouling of pilot-scale plate-frame membrane (Kimura *et al*., 2004).

For a hydrophilic flat-sheet membrane with a MWCO of 150 kDa, Zheng *et al*. (2010) found that the fouling layer formed by biopolymers from secondary effluent of a wastewater treatment plant will change from hydraulically reversible to irreversible in nature due to the alteration of fouling mechanisms from initial pore blocking to cake/gel filtration with more filtration cycles. However, it should be noted that the DOC of the membrane influent was much higher than in the other studies and the NOM composition could be also different. The conversion from reversible to irreversible fouling might not occur for other cases. Another hydrophilic flat-sheet study (MWCO = 100 kD) also reported that the fouling layer formed by hydrophilic NOM fraction was mainly reversible whereas hydrophilic NOM fractions were not found in irreversible fouling layer (Aoustin *et al*., 2001). For hydrophobic flat-sheet membranes with a MWCO of 100 kDa, polysaccharides from hydrophilic source water (e.g. river) tend to form the reversible fouling layer (Jermann *et al*., 2007, 2008). For either hydrophobic or hydrophilic flat-sheet membranes with a MWCO ranging from 8 to 60 kDa, protein-like substances from natural river water and polysaccharides in soluble microbial products generated from an activated sludge process were identified as major contributors to irreversible fouling (Jarusutthirak and Amy, 2006; Peiris *et al*., 2010b, 2011).

2.5.1.2 Humic Substances

For hydrophobic hollow-fiber membranes with a MWCO ranging from 68 to 400 kDa, humic substances that originated from river waters were found to be irrelevant to either reversible or irreversible fouling (Peldszus *et al*., 2011; Zularisam *et al*., 2007). However, humic substances that originated from reservoir waters were found to be responsible for the reversible fouling (Zularisam *et al*., 2007).

Humic-like substances in synthetic solution and natural waters, especially the hydrophobic fraction, are found to be responsible for irreversible fouling of hydrophobic flat-sheet membranes with a lower MWCO ranging from 20 to 100 kDa (Aoustin *et al*., 2001; Jermann *et al*., 2007, 2008; Peiris *et al*., 2010b, 2011; Zhao *et al*., 2011). Regarding reversible fouling, humic substances that originated from reservoir water were reported to play a role in the reversible fouling development (Zhao *et al*., 2011). In other studies, they were not related to the reversible fouling.

2.5.1.3 Colloidal/particulate Matter

The importance of colloidal/particulate matter (C/P) in reversible foulant layer development have been confirmed by synthetic solution and natural water (e.g. river, wastewater, and reservoir) studies for hydrophobic membranes using different C/P characterization techniques (Lee *et al*., 2006; Jermann *et al*., 2008; Peiris *et al*., 2010b, 2011). C/P alone exhibits complete reversible fouling (Jermann *et al*., 2008). The interaction between C/P and NOM fractions is discussed below.

2.5.1.4 Summary – Individual Foulant Behaviours

Synthetic solution studies without colloidal/particulate matter and natural water studies using LC-OCD and ATR-FTIR as NOM characterization techniques (Aoustin *et al*., 2001; Jermann *et al*., 2007, 2008; Zheng *et al*., 2010) reported that hydrophilic NOM constituents, including biopolymer, polysaccharides and protein, and fulvic acids, are major contributors to reversible fouling for both hydrophobic and hydrophilic flat-sheet membranes with negative charge and MWCO ranging from 100 to 150 kDa. But for both hydrophilic and hydrophobic flat-sheet membranes with a tighter MWCO ranging from 8 to 60 kDa, hydrophilic NOM fractions are more important for irreversible fouling (Jarusutthirak and Amy, 2006; Peiris *et al*., 2010b, 2011). For hollow-fiber and plate-frame membranes with a wide MWCO range (68 – 750 kDa), protein and polysaccharides, which are fractions of biopolymer, are mainly responsible for irreversible fouling (Kimura *et al*., 2004;

Zularisam *et al*., 2007; Peldszus *et al*., 2011). But biopolymer as a whole was found to be the major contributor to reversible fouling (Hallé *et al*., 2009).

For hydrophobic NOM fractions that originated from a hydrophilic source water (e.g. river), previous studies revealed the important role of humic substances in the irreversible fouling of flat-sheet membranes (Aoustin *et al*., 2001; Jermann *et al*., 2007, 2008; Peiris *et al*., 2010b, 2011; Zhao *et al*., 2011). In all these studies, humic substances always co-existed with hydrophilic NOM constituents (e.g. protein, polysaccharides) in the water matrix, suggesting the importance of the foulant-foulant interaction in irreversible fouling layer development. Humic substances appeared to be more readily attached irreversibly to the membrane surface and pores modified by the hydrophilic NOM constituents (less negatively charged and more hydrophilic). For hollow-fiber and plate-frame membranes, humic substances originated from hydrophilic source water were not considered as a major foulants (Kimura *et al*., 2004; Peldszus *et al*., 2011). Comparatively, humic substances from hydrophobic source water (e.g. reservoir) were the major contributors to reversible fouling of both hollow-fiber and flat-sheet membranes (Zularisam *et al*., 2007; Zhao *et al*., 2011).

2.5.2 Interactive Behaviour of Foulants

2.5.2.1 Interaction between NOM Fractions and Colloidal/particulate Matter

The co-existence of colloidal/particulate matter and NOM can further complicate the behaviour of various NOM foulants (Jermann *et al*., 2008; Peiris *et al*., 2011; Peldszus *et al*., 2011). Interactions between NOM and colloidal/particulate matter in the natural water matrix and during the fouling process were shown to play a significant role in reversible and irreversible fouling behaviour for PES flat sheet membranes (Jermann *et al*., 2008). This is attributed to the binding of NOM onto colloidal/particulate substances and the impact of NOM on particulate fouling-layer structure (Jermann *et al*., 2008). The combined contribution of NOM and colloidal/particulate matter to PES flat sheet and PVDF hollow-fiber membrane fouling was also reported by Peiris *et al*. (2011) and Peldszus *et al*. (2011), respectively.

As a completely reversible foulant, colloidal/particulate matter can form an irreversible fouling layer with reduced porosity when interacting with humic substances (Jermann *et al*., 2008). This can be explained by the tight attachment of colloidal/particulate matter onto the membrane surface via

bridging of humic substances absorbed onto the colloidal/particulate matter (Jermann *et al*., 2008). Peiris *et al*. (2011) also reported that the extent of irreversible fouling by humic substance-like substances was shown to be positively related to the increased accumulation of colloidal/particulate substances.

Comparatively, protein-like substances that are responsible for irreversible fouling were found to be antagonistically affected by colloidal/particulate matter with respect to the extent of irreversible fouling (Peiris *et al*., 2011). The extent of irreversible fouling by protein-like substances was shown to be negatively correlated with increased accumulation of colloidal/particulate substances (Peiris *et al*., 2011). Peiris *et al*. (2012) provided physical evidence for protein-colloidal/particulate matter interaction using surface plasmon resonance (SPR), which illustrates the impact of the content of colloidal/particulate matter on the extent of association between protein-like substances and self-assembled monolayers. This antagonistic phenomenon can be attributed to the formation of colloidal/particulate matter-induced reversible fouling layer which significantly reduce the protein concentration by protein adsorption onto colloidal/particulate matter and block the protein-membrane interaction sites on the membrane surface (Peiris *et al*., 2011, 2012). Moreover, Peldszus *et al*. (2011) proposed that protein-C/P fouling layer tend to transition from a reversible fouling layer to an irreversible fouling layer either with time or with increased protein content. The transition may occur when protein adsorption sites on protein-C/P fouling layer were saturated and protein-membrane interaction started to play a role, allowing protein-C/P fouling layer to irreversibly attach onto the membrane surface. With respect to polysaccharides, the irreversibility of fouling layer formed by polysaccharides-C/P only slightly increased compared to fouling layer formed by polysaccharides alone (Jermann *et al*., 2008).

2.5.2.2 Interaction between NOM Fractions and Divalent Cations

The presence of divalent cations (e.g. Ca^{2+}, Fe^{2+}, Mn^{2+}) with NOM constituents will promote the mutual influence of fouling mechanisms (Jermann *et al*., 2007; Kimura *et al*., 2004). Compared to the cake layer formed by polysaccharides without calcium, a tighter and less permeable gel layer can be formed by polysaccharides with calcium (Jermann *et al*., 2007). Binding of divalent cations with humic-like substances (e.g. humic and fulvic acids) and proteins can increase their deposition onto the membrane surface, therefore enhancing the membrane fouling (Jermann *et al*., 2007). This can be attributed to reduced repulsion between the negatively charged membrane and the negatively charged

humic substances due to the neutralization of its ionized functional groups by complexation with divalent cations (Aoustin *et al*., 2001; Jermann *et al*., 2007). Also, divalent cations can function as a bridge between the negatively charged membrane surface and negatively charged functional group (e.g. carboxyl groups) and/or molecules (e.g. humic acids) (Aoustin *et al*., 2001). The enhanced affinity of Ca-NOM aggregates onto the membrane will reduce the pore size through internal deposition of aggregates and eventually cause the irreversible fouling (Aoustin *et al*., 2001). In addition, with an initial increase in calcium concentration, decreased NOM (e.g. humic acid) rejection was observed, which can be attributed to the molecular size reduction (Aoustin *et al*., 2001).

2.6 Advantages and Disadvantages of Two Advanced NOM Characterization Techniques in Studying UF Fouling

2.6.1 Fluorescence Spectroscopy

Fluorescence spectroscopy has many attractive advantages, including low cost (relative), non-destructive nature of the technique, small amounts of sample required, simple experimental procedures, no or minimal requirement of sample pre-treatment, short signal acquisition time, high instrumental sensitivity, and the ability to provide information on composition, molecular structure and chemically functional characteristics of NOM (Chen *et al*., 2002; Lakowicz, 2006; Peiris *et al*., 2008, 2011; Baghoth *et al*., 2011; Tran *et al*., 2015). Moreover, fluorescence spectroscopy, which requires little to no pre-treatment, preserves the characteristics of the original water matrix and has the potential to be further developed as an on-line approach (Peiris *et al*., 2010a; Baghoth *et al*., 2011). In addition to these fundamental advantages, fluorescence-based techniques can also be used to characterize the combined contribution of NOM and colloidal/particulate matter at relatively low concentrations especially when combined with multivariate data analysis (Peiris *et al*., 2011).

Major challenges to NOM characterization using fluorescence spectroscopy include inner filtering effects, fluorescence quenching, and oxidant effects, which can interfere with the detection of target compounds (Henderson *et al*., 2009). Inner filtering refers to band shape distortion and/or emission quantum yield attenuation due to the absorption of excitation beam (primary inner-filtration) and emitted fluorescence photons (secondary inner-filtration) by the sample matrix prior to fluorescence intensity measurement (Ohno, 2002; Henderson *et al*., 2009). However, Hudson *et al*. (2008) demonstrated that inner filtering effects are insignificant when the concentration of the target NOM

fraction is low (below 25 mgC/L). Various physical and chemical factors, including temperature, pH, iodide, acrylamide, and metal ions, can induce the fluorescence quenching which causes the reduction of fluorescence intensity of target compounds (Lakowicz, 2006; Henderson *et al*., 2009). Henderson *et al*. (2009) suggested that the fluorescence quenching effect is controllable when the pH range and temperature are well monitored during sample measurement. Furthermore, Spencer *et al*. (2007) observed that fluorescence spectrophotometric parameters had little change at the typical range of pH levels in freshwater systems. Also, the metal ion quenching effect for water samples from natural and engineered systems is negligible (Henderson *et al*., 2009). The impacts of oxidants are unclear. Though Henderson *et al*. (2009) proposed that the fluorescence intensity may generally decrease under high chlorine doses and tryptophan-like fluorescence appears to be more susceptible to oxidant effects compared to other fluorophores.

2.6.2 LC-OCD

LC-OCD has various advantages, including high sensitivity, excellent reproducibility, small injection volume and minimal sample pre-treatment (Peiris *et al*., 2008; Huber *et al*., 2011). Even though the fractionation of NOM into 6 groups is based on convention, LC-OCD has found an appropriate compromise between practicability and degree of detail regarding the NOM characterization (Huber *et al*., 2011).

However, LC-OCD requires relatively long signal acquisition time, making it unsuitable for online application (Peiris *et al*., 2008). Also, LC-OCD requires the water samples to be filtered prior to measurement, which can cause the decrease in fluorescence intensity of protein-like substances (e.g. tryptophan) (Baker *et al*., 2007). Furthermore, LC-OCD is not suitable for measuring the colloidal/particulate matter due to the requirement for pre-filtering during sample preparation.

2.7 Fundamentals of Biofiltration

2.7.1 Introduction to Biofiltration

Biofiltration is a filtration process using biologically active media and is developed for the control of biodegradable organic matter (BOM) via natural biodegradation as well as particles via media filtration (Urfer *et al*., 1997). Biofiltration is an environmentally friendly technique, which is free of any chemical addition and can dramatically reduce the production of undesirable by-products and

waste (Hoefel *et al*., 2009; McDowall *et al*., 2009). Also, other advantages of biofiltration include little maintenance requirement, low infrastructure cost, and simple operation (Hoefel *et al*., 2009). Due to these attractive advantages and its promising performance, biofiltration has been widely employed in the drinking water industry for taste and order control (Hoefel *et al*., 2009; McDowall *et al*., 2009), limitation of bacterial regrowth in distribution system, trace contaminant removal (e.g. pharmaceuticals and endocrine disruptor) (Huck and Sozański, 2008) and reduction of disinfection by-product (DBP) formation potential through DBP precursor control (Niquette *et al*., 1999).

2.7.2 Biophysicochemical Processes during Biofiltration

The biofilm can be defined as the interface-associated aggregation of bacteria and other particulate material embedded in extracellular polymeric substances (EPS) which are mainly comprised of proteins and polymers (Wanner *et al*., 2006; Wilking *et al*., 2011). A biofilm system can be further divided into four compartments, including bulk liquid, mass transfer boundary layer, biofilm, and substratum (Wanner *et al*., 2006). Across the four biofilm compartments, various processes are involved in biofiltration, including mass transfer (e.g. substrate transport between the bulk liquid and the biofilm), mass transport (e.g. substrate diffusion within the biofilm), and mass transformation (e.g. substrate utilization, biofilm growth and decay) (Hozalski and Bouwer, 2001; Wanner *et al*., 2006). Mass transformation processes include microbial/cell growth, microbial loss processes (e.g. inactivation, predation and detachment/shear loss), and abiotic chemical transformations (e.g. solute precipitation) (Wanner *et al*., 2006). The cell growth rate can be modelled using the Monod equation (Bae and Rittmann, 1996; Wanner *et al*., 2006). Advection, molecular diffusion, and turbulent dispersion are three major types of mass transport processes (Wanner *et al*., 2006). Mass transport inside the biofilm is dominated by molecular diffusion (Wanner *et al*., 2006). Advection and turbulent dispersion dominate the mass transport outside the biofilm (Wanner *et al*., 2006). There are three major processes involved in mass transfer, including substrate exchange across the mass transfer boundary layer, cell attachment from the bulk liquid to the biofilm, and particulate component detachment from the biofilm to the bulk liquid (Wanner *et al*., 2006). Substrate exchange across the mass transfer boundary layer is considered to be the most important process, which links the bulk liquid (as the substrate source) with biofilm (Wanner *et al*., 2006).

2.7.3 Key Parameters of Biofiltration

The factors affecting the performance of biofiltration include temperature, empty bed contact time (EBCT), media surface area, media type, presence of other treatment processes (e.g. ozonation and chlorination) prior to biofiltration, concentration and characteristics of influent BOM, and effectiveness of backwashing (Urfer *et al*., 1997; Huck and Sozański, 2008). These operational and design parameters are directly related to the biofilm development in the biofilters and control the efficiency of BOM removal.

2.7.3.1 Temperature

Both the biochemical kinetics of BOM degradation (e.g. synthesis and activity of enzymes) as well as BOM transport/transfer processes are significantly affected by the water temperature (Urfer *et al*., 1997). In theory, higher water temperature will result in higher BOM removal in biofilters (Urfer *et al*., 1997). It is also reported that biofilters, especially anthracite-sand filters, can reach steady-state removal of glyoxal more rapidly at higher temperature (Krasner *et al*., 1993).

2.7.3.2 Empty Bed Contact Time (EBCT)

The EBCT can be defined as "the empty bed volume of the column (i.e. without media) divided by the volumetric flow rate of the feed solution" (Hozalski *et al*., 1995). The BOM removal can be improved by increasing EBCT (Krasner *et al*., 1993; Huck *et al*., 2000). However, this incremental benefit was reported not to be proportionally (Urfer *et al*., 1997; Huck and Sozański, 2008).

2.7.3.3 Media Type

The media commonly used in biofilters include sand, anthracite, and granular activated carbon (GAC) (Krasner *et al*., 1993; Holzalski *et al*., 1995). Compared to sand and anthracite, GAC can better promote biofilm attachment due to its macroporous structure and irregular surface (Urfer *et al*., 1997). However, a GAC filter, which has micropores smaller than the bacteria size, might have a smaller specific surface area for biofilm attachment compared to a sand filter (Urfer *et al*., 1997). Krasner *et al*. (1993) reported that the development of biological activity is more rapid in GAC filters which were also shown to have longer long-term stability (e.g. more resistant to intermittent chlorination, water quality change and shutdown). In mature biofilters, the BOM removal was found to be very similar for GAC/sand and anthracite/sand media when temperature is above 10 ^{o}C (Huck and Sozański, 2008). However, Huck and Sozański (2008) summarized that GAC media was shown to

have better performance at low temperature condition (< 5 ^{o}C). In addition, GAC is more expensive compared to other traditional media such as sand and anthracite.

2.7.3.4 Nature and Concentration of Influent BOM

Hozalski *et al*. (1995) reported that the biodegradability of organic carbon with higher aromaticity and other unsaturated moieties appeared to be lower. The removal of total organic carbon (TOC) was shown to be strongly affected by the organic carbon source and nature (Hozalski *et al*., 1995). UV-to-TOC ratio and apparent molecular weight distribution (AMWD) were suggested by Hozalski *et al*. (1995) as indicators of TOC removal potential. Specific UV absorbance (SUVA), which can be calculated as a ratio of UV_{254} absorbance and DOC, is also widely used to estimate the aromaticity or hydrophobicity of organic compounds (Weishaar *et al*., 2003). Different performance of biofiltration on the removal of different NOM fractions characterized by LC-OCD was also reported by Hallé *et al*. (2009).

2.7.3.5 Backwashing

Ahmad and Amirtharajah (1995) reported that the biomass is less prone to detach from media compared to non-biological particles. Therefore, backwashing strategies can be optimized to effectively restore the hydraulic capacity of the filter by removing the accumulated particulate/colloidal material without causing excessive loss of biomass (Holzalski *et al*., 1995; Urfer *et al*., 1997). Compared with water-only backwashing, backwashing with air scour does not necessarily improve the BOM removal efficiency (Huck *et al*., 2000). However, free chlorine in the backwash water can greatly impair the BOM removal capability of bench-scale biofilters (Huck and Sozański, 2008) whereas there was little impact observed on the performance of the full-scale biofilters (Huck and Sozański, 2008). This might be due to the fact that the biofilms in full-scale filters are more resilient or the chlorine residuals are relatively low.

2.7.3.6 Presence of Other Treatment Processes

The presence of other treatment processes can also directly affect the performance of biofiltration. For example, chlorination can suppress the biofilm development and hence decrease the BOM removal efficiency (Huck and Sozański, 2008). Comparatively, ozonation prior to biofiltration can promote the BOM removal by increasing the biodegradability of the influent organic carbon (Metz *et al*., 2006).

2.8 Research Gaps and Needs

2.8.1 Understanding of UF Fouling Mechanisms

The individual and combined contributions of NOM and colloidal/particulate matter to the hollow-fiber membrane fouling have only been investigated at bench scale (Zularisam *et al*., 2007; Hallé *et al*., 2009; Peldszus *et al*., 2011). To further use these findings to help optimize full-scale UF membrane operation and maintenance, the scale-up effects should be investigated. Therefore, pilot-scale fouling experiments should be conducted to further validate these bench-scale observations. In addition, biopolymers (defined by LC-OCD) have been identified as one of the key UF foulants (Hallé *et al*., 2009). However, the composition of this NOM component, which is defined based on the molecular size, is still unclear. Further exploration on biopolymer composition will help improve the understanding of its fouling behaviour.

2.8.2 Evaluation of NOM Characterization Techniques

LC-OCD and FEEM (combined with peak picking, PCA, and PARAFAC) have been demonstrated to be promising with respect to the NOM characterization. Previous investigations to compare the utility of these NOM characterization techniques have been conducted. Wassink *et al*. (2011) demonstrated that the content of humic substances from one river water estimated by FEEM-PCA is highly correlated with the concentration of humic substances quantified by LC-OCD. However, the difference of FEEM-PCA and LC-OCD in characterization of other NOM fractions was not sufficiently developed and the established relationship might be case-specific. The comparison between FEEM-PARAFAC components with LC-OCD fractions has been conducted by Baghoth *et al*. (2011) using water samples collected from treatment trains for two drinking water treatment plants. But FEEM-PCA was not involved in this study. Advantages and disadvantages of two-way PCA and three-way PARAFAC have been widely discussed based to their mathematical backgrounds (Bro, 1997, 2006; Stedmon *et al*., 2003). However, further comparison of their utility in FEEM-based NOM characterization is insufficient. Therefore, a thorough investigation is necessary to identify the different utilities of these advanced NOM characterization techniques.

2.8.3 Investigation on the Fouling Mitigation Mechanism of Biofiltration as a Pre-treatment Technique for Ultrafiltration

With the assistance of advanced NOM characterization techniques, bench- and pilot-scale investigations demonstrated that direct biofiltration without prior coagulation can effectively mitigate

the UF membrane fouling by reducing biopolymer/protein which is the key contributor to irreversible fouling (Hallé *et al*., 2009; Peldszus *et al*., 2011, 2012). However, the fouling mitigation mechanisms of biofiltration were only investigated at bench scale. Therefore, fouling mitigation mechanisms should be further studied at pilot scale to validate the bench-scale observations. Also, the removal behaviours of various NOM components along the biofilter media depths should be investigated to improve the understanding of their removal kinetics. This will greatly increase our understanding of the effect of scale-up for biofiltration-ultrafiltration treatment train and be useful for developing strategies to better monitor and optimize operation.

Chapter 3

Pilot-scale Investigation of Drinking Water Ultrafiltration Membrane Fouling Rates using Advanced Data Analysis Techniques

This chapter is based on a refereed journal article of the same title published in Water Research. The full reference is listed as follows:

Chen, F., Peldszus, S., Peiris, R.H., Ruhl, A.S., Mehrez, R., Jekel, M., Legge, R.L., Huck, P.M., 2014. Pilot-scale investigation of drinking water ultrafiltration membrane fouling rates using advanced data analysis techniques. Water Research 48, 508 – 518.

Summary

A pilot-scale investigation of the performance of biofiltration as a pre-treatment to ultrafiltration for drinking water treatment was conducted between 2008 and 2010. The objective of this study was to further understand the fouling behaviour of ultrafiltration at pilot scale and assess the utility of different foulant monitoring tools. Various fractions of natural organic matter (NOM) and colloidal/particulate matter of raw water, biofilter effluents, and membrane permeate were characterized by employing two advanced NOM characterization techniques: liquid chromatography – organic carbon detection (LC-OCD) and fluorescence excitation-emission matrices (FEEM) combined with principal component analysis (PCA). A framework of fouling rate quantification and classification was also developed and utilized in this study. In cases such as the present one where raw water quality and therefore fouling potential vary substantially, such classification can be considered essential for proper data interpretation. The individual and combined contributions of various NOM fractions and colloidal/particulate matter to hydraulically reversible and irreversible fouling were investigated using various multivariate statistical analysis techniques. Protein-like substances and biopolymers were identified as major contributors to both reversible and irreversible fouling, whereas colloidal/particulate matter can alleviate the extent of irreversible fouling. Humic-like substances contributed little to either reversible or irreversible fouling at low level fouling rates. The complementary nature of FEEM-PCA and LC-OCD for assessing the fouling potential of complex water matrices was also illustrated by this pilot-scale study.

3.1 Introduction

Due to their small footprints and reliable performance for colloidal/particulate matter and pathogenic microorganism removal, ultrafiltration (UF) membranes have gained considerable acceptance throughout the drinking water industry and are being increasingly used in place of traditional granular media filtration, in drinking water treatment. However, as one of the major challenges in the operation and maintenance of this advanced treatment technology, organic fouling can significantly increase the maintenance costs and operational complexity, and decrease productivity, therefore reducing the attractiveness for adoption of UF membranes in the drinking water industry. Regular maintenance of the full-scale UF membrane system usually employs backwashing, disinfection, and chemical cleaning as fouling mitigation measures (e.g. Peter-Varbanets *et al*., 2011).

Recent fouling studies, using either model solutions or natural waters, have revealed that colloidal/particulate matter and various fractions of natural organic matter (NOM) are major contributors to the organic fouling of UF membranes (Jermann *et al*., 2007, 2008; Hallé *et al*., 2009; Peiris *et al*., 2011; Peldszus *et al*., 2011; Peter-Varbanets *et al*., 2011). Various NOM characterization techniques have been adopted in UF fouling research and are gaining prominence with respect to the characterization and quantification of NOM fractions and colloidal/particulate matter. Two of the most promising NOM characterization techniques include liquid chromatography – organic carbon detection (LC-OCD) and fluorescence excitation and emission matrices (FEEM). LC-OCD can quantify five NOM fractions, including biopolymers (e.g. polysaccharides, proteins, and amino sugars), humic substances, building blocks, low molecular-weight acids and neutrals (Huber *et al*., 2011). FEEM combined with multivariate statistical analysis techniques, such as principal component analysis (PCA) and parallel factor analysis (PARAFAC), can further enhance the utility of the FEEM technique and were demonstrated to be an effective approach to indirectly quantify humic and fulvic acid-like substances, protein-like substances, and colloidal/particulate matter (Peiris *et al*., 2010a, b; Baghoth *et al*., 2011).

Several methods to quantify the hydraulically reversible and irreversible fouling rates of hollow-fibre low pressure membranes operated at constant flux have been proposed (Huang *et al*., 2008, 2009; Nguyen *et al*., 2011; Peldszus *et al*., 2011). Fouling rates can be quantified using process measurement data such as transmembrane pressure and flux. Furthermore, realistic cleaning practices

(e.g. automatic backwashing, maintenance cleaning) in full-scale membrane operation can significantly complicate the evaluation of fouling rates (Nguyen *et al*., 2011) and therefore should be taken into account in order to ensure the validity of fouling rate quantification.

The objective of this study was to employ both LC-OCD and FEEM-based techniques to characterize various potential UF membrane foulants in natural water, including NOM fractions and colloidal/particulate matter using data from a pilot scale study. First, a framework of fouling rate quantification and classification for pilot- and full-scale membrane operations was developed. This was followed by assessing the individual and combined contributions of different potential foulants to UF membrane fouling using LC-OCD data and PCA analysis of FEEM data.

3.2 Materials and Methods

3.2.1 Source Water

Grand River water (GRW, Southwestern Ontario, Canada), which is impacted by agricultural and municipal activities, was used as source water during this study. The chemical and physical water quality of GRW varied seasonally (e.g. total organic carbon (TOC): 5.8–8.2 mgC/L; Temperature: 0.7–25.3 ℃). Detailed characteristics of GRW during the study period can be found in Peldszus *et al*. (2012).

3.2.2 Pilot-scale Biofiltration-ultrafiltration Set-up

Between 2008 and 2010, an investigation of the performance of rapid biofiltration (without prior coagulation) as a pre-treatment to ultrafiltration was conducted at a pilot plant which was fed by GRW diverted from a full-scale water treatment plant in Southern Ontario, Canada. GRW was first filtered through roughing filters to remove coarser material and provide some reduction of periodically occurring turbidity peaks and then roughing filter effluents were continuously fed into three dual-media (anthracite/sand) biofilters in a downflow mode with a hydraulic loading of 5 m/h. The three biofilters had different bed depths and hence different empty bed contact times (EBCTs): 5, 10, and 15 min. The 5-minute and 10-minute EBCT filters (A and B, respectively) were operated in parallel. Filter C (also 5 min EBCT) was fed with the effluent of Filter B, to provide a total EBCT of 15 min. The biofilters and roughing filters were backwashed/cleaned at the same frequency.

The pilot-scale UF unit was equipped with a polyvinylidene fluoride (PVDF) hollow-fibre membrane module and was operated in a dead-end mode with constant flux. The nominal surface area and nominal pore size provided by the manufacturer (GE Water and Process Technologies) are 0.93 m^2 and 0.02 μm, respectively. Also, the UF module was operated at a temperature-corrected flux equivalent to 60 L/m^2/h at 20 ℃ in order to account for variations in water viscosity with temperature. This flux was determined at the beginning of each experiment using the temperature anticipated to be most prominent throughout the experiment. The flux was then kept at this value over the duration of the experiment. Because there was only one UF module installed in the pilot plant, either one of the biofilter effluents or GRW raw water was used as feed water to the UF module at any time. Each filtration cycle lasted 30 min and included a 1 min auto-backwash. A dilute chlorine solution at a concentration of 50 mg/L was used for the UF module maintenance cleaning which lasted approximately 20 min, and was performed approximately every 5–7 days before March 2009 and every 2 days after that unless the membrane fouling was too severe or the cleaning frequency was too high to allow for the development of irreversible fouling. Chemical cleaning of the UF module with a chlorine solution (500 mg/L) followed by a citric acid solution (1 g/L) was performed after completion of each ultrafiltration experiment. Further details pertaining to pilot plant operations can be found in Peldszus *et al*. (2012).

3.2.3 Analytical Methods

Various chemical and physical water quality parameters were monitored for GRW raw water, the biofilter effluents (or membrane feed), and the membrane permeate, including the on-line measurements of temperature and turbidity, and weekly measurements of pH, conductivity, total organic carbon (TOC), dissolved organic carbon (DOC), and UV absorbance following Standard Methods (2005) procedures 4500, 2510, 5310, 5310.2, and 5910, respectively. To further understand the characteristics of the NOM fractions and colloidal/particulate matter, FEEM and LC-OCD analyses were performed. LC-OCD can separate NOM into 5 major fractions based on their apparent molecular weight/size (Huber *et al*., 2011). Water samples were filtered using a 0.45 μm filter prior to LC-OCD analysis. LC-OCD analysis for this study focused on the biopolymer and humic substances. In the LC-OCD chromatogram, the biopolymers, which have a higher molecular weight/size, will elute before the humic substances. FEEM measurements were performed using a Varian Cary Eclipse spectrofluorometer (Palo Alto, CA). One FEEM sample consists of 14 emission spectra (captured at emission wavelength range: 300–600 nm) and each emission spectrum including 301 individual

emission intensity values is obtained at a particular excitation wavelength (250–380 nm, with 10 nm increments). In this study, the photomultiplier tube voltage and excitation/emission slit width were set at 750 V and 10 nm, respectively. A scan rate of 600 nm/min was used. Further details with respect to FEEM and LC-OCD measurement can be found in Peiris *et al*. (2010a, b) and Huber *et al*. (2011), respectively. During the study period, 278 FEEM samples and 80 LC-OCD samples were collected from GRW raw water, biofilter effluents, and membrane permeate.

3.2.4 Hydraulically Reversible and Irreversible Fouling Analysis

3.2.4.1 Fouling Rate Quantification

Transmembrane pressure (TMP) and flux before and after backwash (BW) were recorded on-line to monitor the fouling behaviour of the pilot-scale UF membrane. As temperature fluctuated during each ultrafiltration experiment, TMP readings for each experiment were corrected to the same corresponding temperature. For experiment j with corresponding temperature $T_{\text{ref},j}$, the TMP reading for ith filtration cycle at $T_{i,j}$ ℃ was corrected as follows (Hall é, 2009):

$$\text{cTMP}_{i,j} = \text{TMP}_{i,j} \times 1.025^{(T_{i,j} - T_{\text{ref},j})}$$

where $\text{TMP}_{i,j}$ represents the uncorrected TMP for ith filtration cycle of experiment j and $\text{cTMP}_{i,j}$ represents the corrected TMP for ith filtration cycle of experiment j.

These pilot-scale ultrafiltration experiments were designed to challenge the UF membrane unit and the duration of individual experiments ranged from 8 to 57 days. Each experiment included several maintenance cleanings and these were used to further separate each experiment into what is defined here as fouling cycles (i.e. the period between two consecutive maintenance cleanings). Shutdown and sharp flux increases were also used as end points of fouling cycles. The fouling cycle needs to be distinguished from the filtration cycle which includes 30 min permeation followed by backwash and sparging. Hence one fouling cycle consisted of numerous filtration cycles. Reversible and irreversible fouling rates were then quantified using TMP data for each fouling cycle. Rates were calculated for the entire fouling cycle and also for a 3 day period within each fouling cycle (i.e. sampling day and the days before and after sampling). This was followed by a comparison between results for the 3 day period to results for the overall fouling cycle (see Section 3.2.4.3). The irreversible fouling (IRF) rate (unit: kPa/day) was quantified as the slope/increase rate of the temperature-corrected TMP after BW

for the entire time period of each fouling cycle and for the 3 day period around sampling in each fouling cycle. The reversible fouling (RF) rate (unit: kPa/min) was first calculated for each filtration cycle as the difference between corrected TMP before and after BW divided by 30 min (duration of each filtration cycle). The RF rate (unit: kPa/min) of each fouling cycle was then calculated as the average of RF rates of all filtration cycles within each fouling cycle and also as the average of RF rates of all filtration cycles within a 3 day period in each fouling cycle.

3.2.4.2 Fouling Rate Classification

Peldszus *et al*. (2011) demonstrated that relationships between foulants and fouling rates can be very different at different levels of reversible and irreversible fouling. This indicates that the role of foulants in the fouling process may differ at different degrees of membrane fouling. Therefore, it is necessary to classify fouling rates into different categories to better understand the individual and combined contributions of various foulants to membrane fouling. The fouling rate of the entire fouling cycle was used for the classification of an individual fouling cycle. Clustering analysis and discriminant function analysis (DFA) were used to help explore the structure of the fouling rate dataset comprised of both reversible and irreversible fouling rates of all 25 fouling cycles. Analyses were performed using the software R 2.13.1 (The R Foundation for Statistical Computing). Prior to the clustering analysis and DFA, the fouling rate dataset was auto-scaled to make the variables (i.e. RF and IRF rates) compatible so that the units and magnitudes of the variables were comparable. Auto-scaled fouling rates were only used for the clustering and DFA analysis. All other evaluations listed in Sections 3.2.4.3 and 3.2.6 used the actual fouling rates as calculated in Section 3.2.4.1.

Clustering analysis identifies categories which can be used to describe the data structure by allocating similar objects into the same category (Legendre and Legendre, 1998). In this study, average agglomerative clustering analysis (Legendre and Legendre, 1998) was conducted using both reversible and irreversible fouling rates as clustering criteria. The performance of clustering analysis was evaluated using the cophenetic correlation coefficient, which measures the similarity between the dissimilarities in the original dataset and dissimilarities estimated from the dendrogram (Borcard *et al*., 2011). The magnitude of this coefficient should be close to 1 when the clustering results represent the dissimilarities in the original dataset (Borcard *et al*., 2011). After classifying fouling cycles into different categories, the validity of their classification was also evaluated using DFA. DFA performs

an ordination of samples to maximize the possible separation among the pre-assigned groups on the new axes (Legendre and Legendre, 1998).

3.2.4.3 Comparison between Fouling Rates of Whole Fouling Cycle and 3-day Duration

The duration of fouling cycles ranged from 0.4 to 9.9 days. Fouling rates for these cycles might not be representative of conditions at the time LC-OCD or fluorescence samples were taken, when water quality fluctuates substantially during a fouling cycle. This is less likely the case when the fouling cycle duration is kept relatively concise. Therefore, in this study, it was assumed that water quality remained relatively stable within 3 days, which was supported by stable on-line turbidity data of the membrane influents (e.g. biofilter effluents and raw water) for a 3-day period (LC-OCD/FEEM sampling day, the day before and the day after the sampling day) of each fouling cycle. More importantly, the selection of a 3-day duration is to ensure sufficient time for the development of irreversible fouling and associated TMP increase which was especially important when the fouling rate was relatively low. Both reversible and irreversible fouling rates were calculated using TMP data recorded within the 3-day period of each fouling cycle, including the sampling day, the day before and after the sampling day. If the day before or after the sampling day were not available, TMP data for 2 days or less were used to calculate the IRF and RF rates. Then, a Pearson's correlation analysis of reversible/irreversible fouling rates of the whole fouling cycle vs. the fouling rates for 3-day duration was conducted to investigate whether 3-day duration fouling rates can accurately represent the fouling rates of the whole fouling cycle.

3.2.5 Fluorescence Data Pre-treatment and Principal Component Analysis (PCA)

Principal component analysis (PCA) was performed to identify trends in the complex dataset which is comprised of 278 FEEM samples and hence is intrinsically a 3D matrix. Therefore, prior to performing PCA, each FEEM sample with 4214 (301×14) excitation and emission coordinate points was reorganized into one row for all 278 samples and this rearrangement procedure generated a 278×4214 (2D) data matrix (X). To ensure the equal prominence of each intensity reading captured within the EEMs, this X matrix was auto-scaled by adjusting it to zero mean and unit variance. Then, PCA with random subset cross validation was conducted on the auto-scaled X matrix to identify statistically significant principal components and generate score and loading plots. PCA analysis was performed using PLS Toolbox 6.2.1 (Eigenvector Research, Inc., Manson, WA) within the MATLAB

7.11.0 computational environment (MathWorks, Natick, MA). Detailed procedures are described in Peiris *et al*. (2010a, b).

3.2.6 Evaluation of Individual and Combined Contributions of NOM-colloidal/particulate Matter to UF Membrane Fouling

Fouling rates calculated using TMP data for 3-day periods within each fouling cycle were used for the evaluation of individual and combined contributions of NOM fractions and colloidal/particulate matter to UF membrane fouling.

The individual contributions of NOM fractions and colloidal/particulate matter to UF membrane fouling were assessed by conducting regression analysis between PC scores or LC-OCD concentration for each foulant and irreversible/reversible fouling rates at different levels. Step-wise multiple linear regression (MLR) analysis was performed to investigate the individual and combined contributions of various NOM fractions to reversible and irreversible fouling separately (Rawlings *et al*., 1998). In this study, response variables include reversible or irreversible fouling rates, whereas explanatory variables include PC scores of various foulants and their interaction terms. The interaction term between variable 1 (V1) and variable 2 (V2) can be mathematically expressed as $V1 \times V2$. An F-test and t -test were used to assess the overall significance of all explanatory variables and the significance of any individual explanatory variables, respectively. Step-wise MLR as well as F-test and t -test were conducted using MATLAB 7.11.0 and R 2.13.1 (The R Foundation for Statistical Computing). Due to the small size of LC-OCD-fouling rate sample pairs, these statistical techniques were only applied to FEEM-fouling samples.

3.3 Results and Discussion

3.3.1 Fouling Rate Analysis

3.3.1.1 Fouling Rate Quantification and Classification

The reversible and irreversible fouling rates were calculated for the entire fouling cycle. The fouling rates of the entire fouling cycle ranged from 0.39 kPa/day to 35.61 kPa/day for irreversible fouling and from 0.07 kPa/min to 0.65 kPa/min for reversible fouling (Table A1). According to Figure 3.1a, all fouling cycles were classified into three main clusters/categories by performing clustering analysis

on reversible and irreversible fouling rates of the whole fouling cycle. Three distinct clusters, including low-, intermediate-, and high-level groups, reflect the extent of overall (both reversible and irreversible) fouling behaviour of different fouling cycles in the corresponding groups. The cophenetic correlation coefficient is 0.940, suggesting that the clustering using average linkage can well preserve the original dissimilarities of the original matrix. Each category was well isolated from other categories according to the cophenetic distance/height between two categories in the dendrogram (Figure 3.1a). Then, discriminant function analysis was performed using three groups pre-assigned by clustering analysis (Group 1 – high; Group 2 – intermediate; Group 3 – low) and the ordination of fouling cycles in these groups is illustrated in Figure 3.1b. Low-, intermediate-, and high-level groups were well isolated from each other, confirming the validity of categories identified from clustering analysis.

By examining the clustering dendrogram (Figure 3.1a) and raw fouling rate data (Table A1), A1, A3, and R1 were found to deviate from other samples in the low-level fouling group and A8, A9, and C10 were identified to be different from other samples in the intermediate-level fouling group. A8 and A9 were samples collected from the same fouling cycle and hence they have the same RF and IRF rates of the whole fouling cycle. The fouling cycles containing A8, A9, and C10 were found to have much lower IRF rates compared to other intermediate-level fouling cycles. With respect to the low-level category, three fouling cycles (containing samples A1, A3, and R1 in the same dendrogram branch) have relatively higher RF rates compared to the other ones in the low-level group. Therefore, fouling cycles containing these six samples (i.e. A1, A3, A8, A9, C10, and R1) were taken out when defining the boundaries of low-, intermediate-, and high-level fouling rates. Their RF and IRF rates were re-classified after the boundaries were established using regular fouling cycles only. The boundaries for both RF and IRF between two categories were defined as the average of the minimum value of the higher-level category and the maximum value of the lower-level category (Figure A1). If the RF or IRF rate is above the corresponding boundary, the RF or IRF rate of the fouling cycle is considered to be at the higher level.

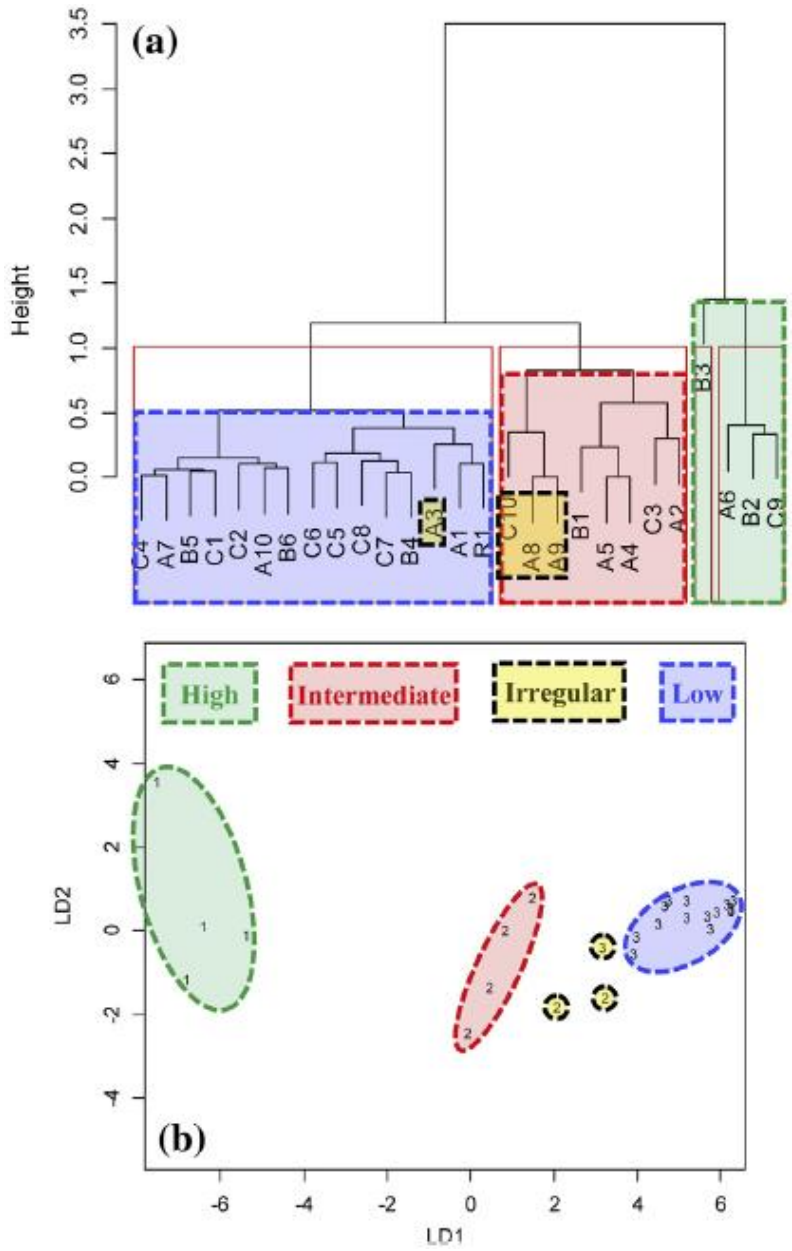


Figure 3.1 Classification of fouling cycles using (a) clustering analysis and (b) discriminant function analysis of reversible and irreversible fouling rates of the whole fouling cycle. Letter-number combinations are fouling cycles containing FEEM or LC-OCD samples. Letter indicates the types of membrane influent samples such as effluents of biofilter A–C or raw water (R). All fouling cycles were classified into low-level (blue box/ellipse), intermediate-level (red box/ellipse), high-level (green box/ellipse), and irregular (yellow box/ellipse) groups. Group 1 to 3 in DFA plot refer to high-, intermediate-, and low-level groups, respectively.

Re-classification results indicate that all fouling cycles containing these six samples have low-level IRF rates. Fouling cycles containing A1 and R1 have also low-level RF rates whereas three fouling cycles containing samples A3, A8, A9, and C10 have intermediate-level RF rates. Discriminant function analysis also isolated these latter three fouling cycles from the regular low- and intermediate-level fouling cycles as shown in Figure 3.1b. The fouling cycle containing A3 (Group 3 in yellow) is situated between the low- (Group 3) and intermediate-level (Group 2) fouling groups. Similarly, the fouling cycles containing C10, A8 and A9 (Group 2 in yellow) are also different from other intermediate-level fouling cycles. They are the only three fouling cycles having low-level IRF rates but intermediate-level RF rates, making them relatively irregular compared to other regular fouling cycles. Table 3.1 summarized the RF and IRF rate ranges as well as irregular samples for each category.

Table 3.1 Classification of fouling cycles.

Fouling cycles	IRF rate (kPa/day)		RF rate (kPa/min)		Irregular samples
	Max	Min	Max	Min	
High-level	35.61	22.40	0.65	0.57	None
Intermediate-level	10.74	5.94	0.39	0.27	C10, A8 & A9
Low-level	3.99	0.39	0.15	0.07	A3

Note: irregular samples were not included in RF and IRF rate ranges.

3.3.1.2 Impact of Fouling Cycle Duration on Fouling Rate Quantification

As noted previously, fouling rates for a duration of 3 days were calculated to ensure that the calculated fouling rates were representative for the point in time at which water quality samples (e.g. LC-OCD and FEEM) were taken, assuming that water quality was relatively stable within a 3-day period and irreversible fouling had developed. However, we also investigated whether fouling rates of a 3-day duration are representative of the fouling rates for the whole fouling cycle. If not, different fouling mechanisms and/or other operational issues (e.g. sudden water quality change) might play a role in data interpretation and further investigation would be required. As illustrated in Figure 3.2, the high correlation between RF or IRF fouling rates calculated using TMP data from the whole fouling cycle and for a 3-day duration, for regular fouling cycles, show that a 3-day duration can accurately reflect the extent of fouling for the whole fouling cycle. Furthermore, this also suggests that the 3-day duration fouling rates will follow the classification criteria based on the fouling rates for the whole fouling cycle.

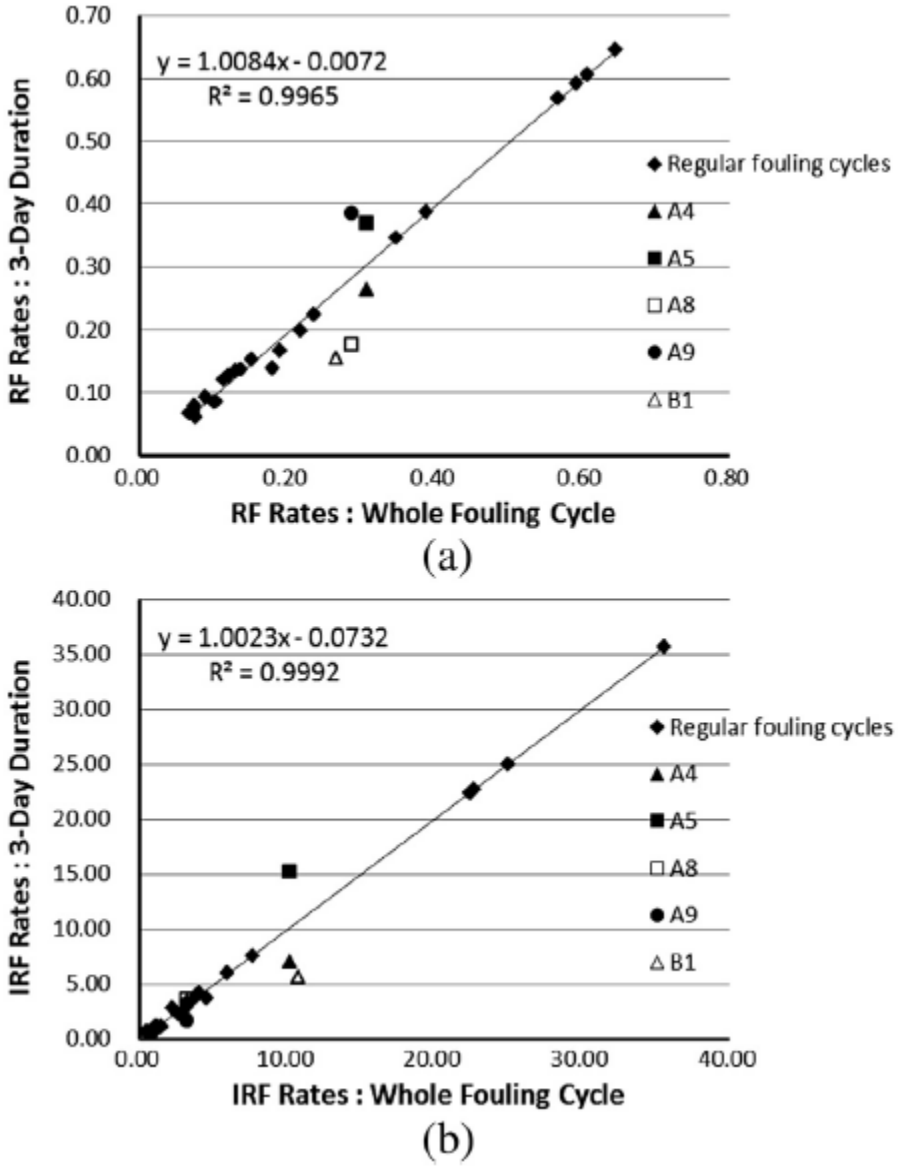


Figure 3.2 Comparison between the fouling rates calculated using TMP data of the whole fouling cycle vs. a 3-day duration. (a) Reversible fouling (RF) rates and (b) irreversible fouling (IRF) rates. R^2 is calculated only based on regular fouling cycles.

Five samples, A4, A5, A8, A9 and B1, were however found not to follow the trend (Figure 3.2). All 5 samples were collected at the early or end stage of their corresponding fouling cycles. For these five samples fouling rates of 3-day duration cannot reflect the fouling characteristics of the whole cycle. The TMP plots for these fouling cycles show curvature (Figure A2), suggesting that they have

different fouling mechanisms compared to the other cycles. Therefore, fouling rates calculated for these 5 irregular samples were not comparable to other samples and hence were not included in the following analysis.

3.3.1.3 Generalized Framework for Fouling Rate Quantification and Classification

The framework of fouling rate quantification and classification developed and introduced above can be further generalized and applied to other pilot- and full-scale membrane fouling monitoring studies to help understand the structure of a fouling-rate dataset and the level/extent of fouling experienced by the membrane facility, although in cases where membrane feed water quality varies less than in the present study, such distinct classifications may not necessarily be present. The five steps of this framework can be summarized as follows:

1) Define the length of a fouling cycle using operational practices (e.g. maintenance cleaning, shutdown, flux adjustment).

2) Quantify the RF and IRF rates for each fouling cycle.

3) Conduct the clustering analysis on both RF and IRF rates to classify fouling cycles into categories with different extents of reversible and irreversible fouling (e.g. low, intermediate, high).

4) Identify irregular fouling cycles, if any, by examining clustering results and original data, e.g. irregularly high RF rate when IRF rate is low. Define boundaries for each fouling category excluding irregular samples. Re-classify RF and IRF rates of irregular samples into appropriate categories.

5) Confirm the classification results using discriminant function analysis.

Performing such a classification procedure can be very important in investigations where membrane fouling mechanisms are being investigated, because they may be different, depending on the (raw water quality related) level of fouling occurring.

3.3.2 Individual Contribution of NOM-colloidal/particulate Matter Fractions to UF Membrane Fouling

3.3.2.1 FEEM-based Technique

Using the FEEM-PCA technique, three statistically significant principal components (PCs), which captured 83.4% of the total variance in the X data matrix, were successfully extracted. By comparing loading plots of these three PCs (Figure A3) with original FEEM spectra of Grand River water and

the literature (Chen *et al*., 2003a; Sierra *et al*., 2005), PC1 (58.5%, Ex/Em ~ 320 nm/415 nm and 270 nm/460 nm), PC2 (14.8%, first and second order Rayleigh scattering), and PC3 (10.1%, Ex/Em ~280 nm/330 nm) were determined to be related to humic-like substances, colloidal/particulate matter, and protein-like substances, respectively. PC1, PC2 and PC3 scores of each sample were subsequently used to represent the content of the corresponding foulants. FEEM-PCA results in this pilot study were consistent with previous studies using the same source water (Peiris *et al*., 2010a, b; Peldszus *et al*., 2011).

The individual contribution of three potential foulants, including colloidal/particulate matter, humic- and protein-like substances, to both hydraulically reversible and irreversible fouling was investigated by conducting regression analysis between their PC scores and irreversible/reversible fouling rates, respectively. There are 16 parallel data pairs of FEEM samples and fouling (RF or IRF) rates.

With respect to reversible fouling, Figure 3.3 a–c illustrated that protein-like substances seem to contribute to low-level reversible fouling, whereas colloidal/particulate matter and humic-like substances appeared to play a role in more severe fouling situations. However, these trends need to be confirmed as there are only 3 or 4 data points available for high and intermediate fouling rates, respectively. Figure 3.3d suggests that the low-level irreversible fouling rate was well correlated with the content of protein-like substances. In addition, both humic- and protein-like substances appeared to play an important role when irreversible fouling became severe (Figure 3.3d–e). Comparatively, the content of colloidal/particulate matter was shown to have a consistently negative correlation with both low- and high-level irreversible fouling rates (Figure 3.3f), suggesting that colloidal/particulate matter can potentially reduce the extent of irreversible fouling.

The importance of protein-like substances in irreversible fouling of hollow-fibre UF membranes was therefore highlighted by this pilot-scale study. These findings corroborate the results from earlier bench-scale studies with shorter duration (Peldszus *et al*., 2011) and are consistent with results from a bench-scale flat-sheet UF membrane study using the same source water (Peiris *et al*., 2010b) and a hollow-fibre UF membrane study using secondary effluents from wastewater treatment (Haberkamp

et al., 2011). Moreover, the important role of protein-like substances in low-level reversible fouling was also revealed in this pilot-scale study.

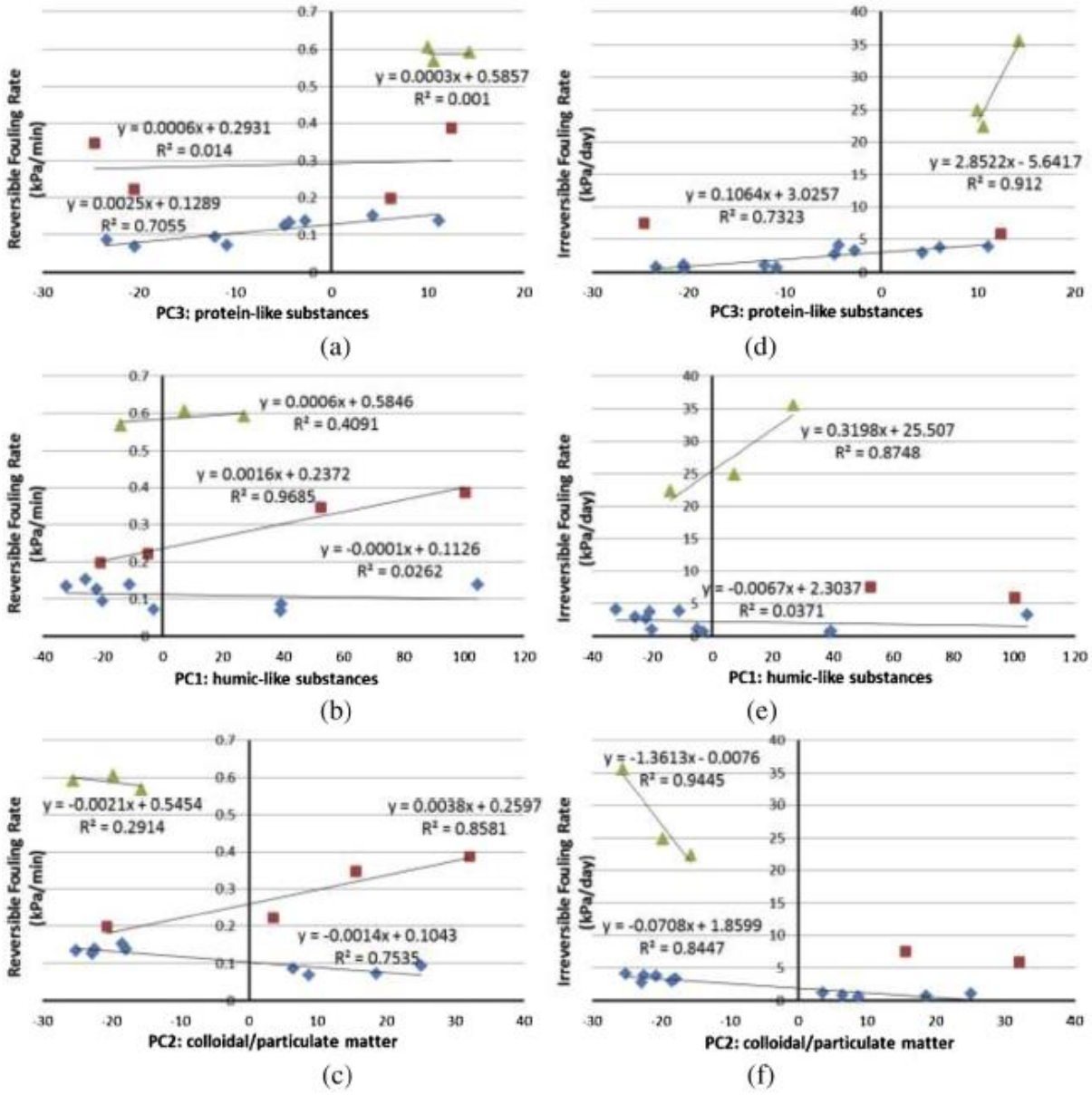


Figure 3.3 Impacts of the colloidal/particulate matter, protein- and humic-like substances (semi-quantified using the FEEM-PCA technique) on the reversible (a–c) and irreversible fouling (d–f) of UF membranes. Reversible and irreversible fouling rates were classified into low-level (blue diamond), intermediate-level (red square), and high-level (green triangle) groups according to clustering analysis results.

It is important to note that the roles of NOM fractions and colloidal/particulate matter in membrane fouling, and by inference the fouling mechanisms, were shown to be different at different levels of both reversible and irreversible fouling (Figure 3.3a–f). This is consistent with previous studies (Peldszus *et al*., 2011), confirming the necessity of using techniques such as clustering analysis and DFA to classify fouling rates into different categories, prior to further interpretation. Fouling rate data interpretation without classification ignores the different fouling behaviours at different stages of fouling layer development, which is inappropriate and might result in a false understanding of fouling mechanisms.

3.3.2.2 LC-OCD-based Technique

80 LC-OCD samples were collected for 16 sampling events with each sampling event consisting of 5 samples: GRW raw water, three biofilter effluents, and membrane permeate. However, as noted earlier, the feed water to the UF module in any given experiment could only be GRW raw water or one of the three biofilter effluents. As a result, only one LC-OCD sample for each sampling event can be used to investigate the impact of NOM characteristics of the feed water on the membrane fouling behaviour. In addition, due to the limited availability of the LC-OCD instrument, only 8 LC-OCD samples are available in parallel with fouling rate data. Also, one data pair was available for each intermediate- and high-level reversible/irreversible fouling situation, respectively. Therefore, only the individual contribution of foulant species to low-level RF or IRF could be discussed. The relationship between the LC-OCD concentration of NOM fractions, including humic substances and biopolymers, and reversible and irreversible fouling rates is illustrated in Figure 3.4. According to Figure 3.4a and b, the concentration of humic substances is not correlated with low-level reversible fouling rates whereas biopolymer concentration has a good correlation with low-level reversible fouling rates. With respect to irreversible fouling, the biopolymer concentration was found to be well correlated with low-level irreversible fouling rates (Figure 3.4c), indicating its important role in the development of irreversible fouling. Comparatively, the correlation between humic substances and irreversible fouling rates was very poor (Figure 3.4d).

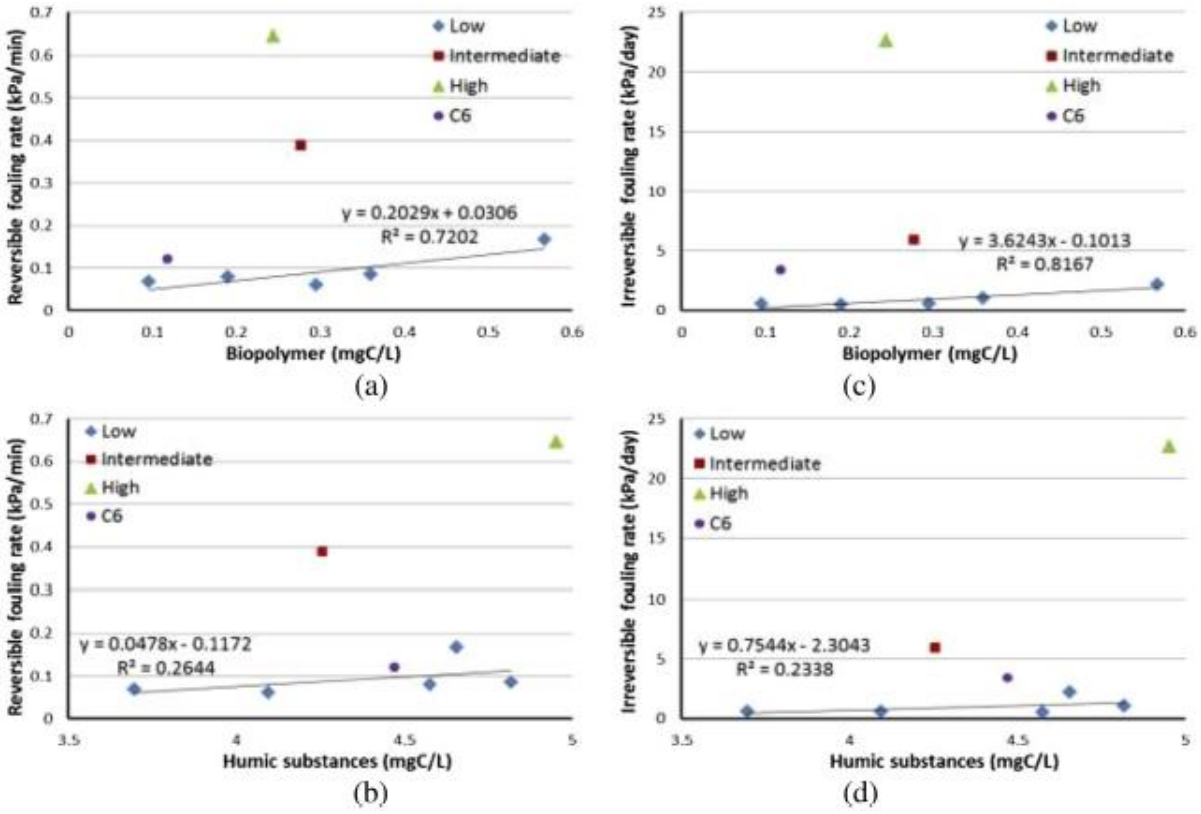


Figure 3.4 Impacts of the biopolymer and humic substances (quantified using the LC-OCD technique) on the reversible and irreversible fouling of UF membranes. Reversible and irreversible fouling rates were classified into low-level (blue diamond), intermediate-level (red square), and high-level (green triangle) groups according to clustering analysis results. The irregular sample (C6) is shown as a purple circle and is not included in the trend line.

Regarding the significant contribution of biopolymers to reversible fouling and the role of humic substances in the reversible and irreversible fouling, results from this pilot-scale study corroborated previous bench-scale studies (Hallé *et al*., 2009; Peldszus *et al*., 2011). However, the correlation between biopolymer concentration and irreversible fouling has not been reported for previous studies. Different conclusions have been drawn with respect to the contribution of biopolymers to IRF. Hallé *et al*. (2009) suggested that the extent of hydraulically reversible fouling was related to the concentration of biopolymers whereas irreversible fouling was postulated to be related to the composition of biopolymers. The difference between this previous study and the study here could be at least in part attributable to the different biopolymer concentration ranges for the two studies. In the

investigation by Hallé *et al*. (2009), biopolymer concentration ranged from 0.03 to 0.26 mgC/L whereas this pilot-scale study had a much wider biopolymer concentration range, from 0.19 to 0.57 mgC/L, indicating that the pilot-scale study reflected larger fluctuations of water quality and hence a more representative biopolymer concentration range. Furthermore, the composition of biopolymers in the two studies may have been different and this could potentially be another reason for the different findings. It should be noted that Hallé *et al*. (2009) did not conduct a rigorous fouling rate classification when investigating the individual contribution of biopolymer to the membrane fouling, which might also have contributed to the different findings between the two studies. As the LC-OCD dataset is relatively small, further studies are required to confirm this observation.

When estimating the irreversible fouling rate over the 3-day duration (slope of temperature-corrected TMP after BW vs. time) for the fouling cycle containing C6, a very unusual residual pattern (the difference between the predicted TMP and observed TMP vs. predicted TMP) of the data was observed, compared to the other fouling cycles within the same ultrafiltration experiment (immediately before and after this fouling cycle). An increase in variance of the residuals was identified and this violates the constant-variance assumption for regression analysis, suggesting other factors such as instrumental error might play a role. Therefore, C6 was excluded from the above analysis.

3.3.2.3 Comparison between LC-OCD- and FEEM-based Techniques

The FEEM-based technique groups NOM into various fractions based on their different structural and functional properties, which can significantly affect the chemical, physical, and polyelectrolytic behaviours of NOM constituents (Chen *et al*., 2003b; Baghoth *et al*., 2011). However, the LC-OCD technique separates NOM into different fractions based on their apparent molecular size or weight (Huber *et al*., 2011). FEEM and LC-OCD also differ in that FEEM is able to characterize colloidal/particulate matter by interpreting the Rayleigh scattering regions, whereas for LC-OCD the larger fraction of the colloidal/particular matter is removed by pre-filtration through a 0.45 μm filter during sample preparation.

With respect to humic-like substances, both LC-OCD and FEEM techniques revealed that for this study they do not contribute to either reversible or irreversible fouling when these are at a low level.

LC-OCD illustrated that biopolymers, which are comprised of polysaccharides and proteins or protein-like material, are a major contributor to low-level reversible and irreversible fouling. FEEM results were consistent with LC-OCD results in that FEEM demonstrated that protein-like substances were important in both low-level reversible and irreversible fouling. In general, the two advanced NOM characterization techniques provided consistent results regarding the contribution of humic substances and biopolymers (or the constituent fractions) to reversible and irreversible fouling.

3.3.3 Combined Contribution of NOM-colloidal/particulate Matter Fractions to UF Membrane Fouling

Step-wise multiple linear regression (MLR) was performed to investigate individual and possible combined contributions of foulants to membrane fouling. Six explanatory variables, including humic-like substances (PC1), colloidal/particulate matter (PC2), protein-like substances (PC3), humic-protein interaction (PC1 × PC3), humic-colloidal/particulate matter (C/P) interaction (PC1 × PC2), and protein-C/P interaction (Protein-C/P, PC3 × PC2), were included in the step-wise MLR analysis. Akaike information criterion (AIC), Bayesian information criterion (BIC), and significance level criterion (SLC) based on the *p*-value of an F-statistic are three commonly used criteria for the step-wise variable selection process (Rawlings *et al*., 1998). The optimal model should minimize AIC or BIC value (Rawlings *et al*., 1998). In this study, step-wise MLR was performed using each of the three criteria to validate each other. The optimal models for reversible (RF) and irreversible fouling (IRF) rates generated from step-wise MLR analysis based on all three criteria are summarized as follows:

$$RF = 0.0078 \times \text{Protein} + 0.2720 \quad (1)$$

$$IRF = 0.5142 \times \text{Protein} - 0.0239 \times \text{Protein-C/P} + 6.6438 \quad (2)$$

Table 3.2 Results of significance test of optimal models (RF – Equation (1); IRF – Equation (2)) from step-wise multiple linear regression (MLR).

	p-value (t-test)		*p-value (F-test)*	R^2	*Adjusted* R^2
	Protein	Protein-C/P			
RF	0.025 (+)	N/A	0.025	0.31	0.26
IRF	0.004 (+)	0.032 (−)	0.006	0.54	0.47

Note: Numbers are bolded if $p < \alpha$ (0.05).
"+" or "–" is the sign of each explanatory parameter in the regression model.

As illustrated above, the same optimal models were identified for both RF and IRF rates using all three criteria. Other statistics (i.e. R^2, adjusted R^2) as well as the results of significance tests (i.e. *t* -test and F-test) for the optimal models are summarized in Table 3.2.

Compared to R^2, an adjusted R^2 rescales R^2 by degrees of freedom and as such is a better statistic for the comparison between models with different numbers of explanatory variables (Rawlings *et al*., 1998). It should be noted that the adjusted R^2 is not very high for optimal models, which is not surprising and can most likely be attributed to random noise of both response and explanatory data. Another possible reason for the relatively low adjusted R^2 could be that other significant explanatory variable(s) or interactions (e.g. no means of quantifying polysaccharides at these low concentrations) in addition to the ones identified (i.e. NOM characteristics) are missing, which only highlights the complexity of the fouling process.

According to step-wise MLR results, humic substances and colloidal/particulate matter were determined to be insignificant in all situations, suggesting they were not directly related to the reversible and irreversible fouling of the hollow-fibre UF membrane used in this pilot study. The contribution of protein-like substances to the RF and IRF rate were shown to be significant, confirming the importance of protein-like substances in both reversible and irreversible fouling. These observations are consistent with the observations from the individual foulant contribution investigation in Section 3.3.2.1 (Figure 3.3). Humic-like substances have an insignificant correlation with both reversible and irreversible fouling rates at low level compared to protein-like substances.

The protein-C/P interaction was found to be significantly related to the IRF rate, whereas it had insignificant relationship with the RF rate. The negative correlation between the interaction term (Protein-C/P = PC3 × PC2) and irreversible fouling rates suggested that the co-existence of protein-like substances and colloidal/particulate matter can alleviate the extent of IRF fouling. Physical evidence of the interaction between protein-like substances and colloidal/particulate matter was demonstrated by Peiris *et al*. (2012) using surface plasmon resonance (SPR). They demonstrated how the protein- and colloidal/particulate-like matter interactions contributed to reduced physical level protein-like matter binding on self-assembled monolayer surfaces that were specifically designed to

promote strong protein binding. Peiris *et al*. (2012) provides the physical evidence for the negative correlation between irreversible fouling rates and C/P content. Therefore, it is reasonable to believe that the protein-C/P interaction can reduce the impact of protein-like substances available to cause irreversible fouling by masking the sites available for protein-like substances to interact with the UF membrane surface. However, the impact of protein-C/P interaction on the formation of the reversible fouling layer is limited.

3.4 Conclusions

This pilot-scale investigation employed both LC-OCD and FEEM-PCA to characterize various foulants and understand their individual and combined contributions to UF membrane fouling in a drinking water context. Earlier bench-scale results were corroborated by this pilot study. Results presented here underscore the value of LC-OCD and FEEM-PCA for characterising NOM and their contributions to membrane fouling, thereby providing insights into fouling mechanisms. The major conclusions can be drawn as follows:

1) The necessity and validity of fouling rate classification were illustrated by this study. A framework combining clustering analysis and DFA was proposed to help understand the fouling rate data structure and classify the fouling rates into categories with different RF and IRF fouling extents. Although such classification might not be required in situations where raw water quality (and therefore fouling potential) were less variable, it would be prudent in general for investigators to subject their fouling data to a classification analysis prior to undertaking detailed interpretations. This would help to identify the possibility that different fouling mechanisms might be at work on different sampling days.

2) Individual foulant contribution investigation using FEEM-PCA and LC-OCD techniques highlighted the importance of protein-like substances and biopolymers in reversible and irreversible fouling of UF membranes.

3) The combined foulant contribution investigation suggested that protein-like substances and colloidal/particulate matter are positive and negative contributors to UF membrane fouling, respectively. By this we mean that colloidal/particulate matter can alleviate the extent of membrane fouling.

4) Both the individual and combined foulant contribution investigation found that humic-like substances do not contribute to either reversible or irreversible fouling for the membrane used at low

fouling rates. Though at higher fouling rates there is some limited evidence that humic-like substances contribute to reversible and irreversible fouling.

5) This pilot-scale study confirmed the results from previous bench-scale work using the same polymeric hollow-fibre UF membrane and illustrates the potential of FEEM-PCA and LC-OCD as effective tools for evaluating the fouling potential of complex water matrices. These techniques are also effective tools for optimizing a biofiltration-ultrafiltration treatment train, therefore minimizing the organic fouling and enhancing the application potential of UF membranes in the drinking water industry.

Chapter 4
Physical Significance of Principal Components Extracted from Fluorescence Excitation-Emission Matrices of Individual and Combined Water Model Solutions

Summary

Fluorescence excitation-emission matrices (FEEM) have been widely used in the natural organic matter (NOM) characterization and quantification. This study investigates the physical significance of principal components (PCs) extracted from FEEM combined with principal component analysis (PCA) and the impacts of NOM-NOM and NOM-inorganic colloid interactions on PC scores using four model compounds: bovine serum albumin, humic acids, sodium alginate and silica. Three statistically significant principal components were identified and determined to be related to humic-like substances (PC1), inorganic colloids and protein-like substances (PC2) and protein-like substances (PC3). Each PC can provide information regarding the fluorescence properties of inorganic colloids and a particular NOM component as well as their interactions with other NOM components and inorganic colloids. In addition, silica and alginate were found to contribute positively to PC1. PC2 is shown to be related to both protein and inorganic colloids. The presence of humic acid can further complicate the Rayleigh scattering pattern (PC2). The antagonistic effect between bovine serum albumin and silica on PC3 was also observed. This study illustrates that FEEM-PCA can be potentially applied to investigate the NOM-NOM and NOM-inorganic colloid interactions.

4.1 Introduction

Due to the presence of intrinsic fluorophores in natural organic matter (NOM), fluorescence excitation emission matrices (FEEMs) have been widely used to characterize various NOM fractions in natural water, including humic-like substances (e.g. fulvic and humic acids) and protein-like substances (Sierra *et al.*, 2005; Baghoth *et al.*, 2011; Chen *et al.*, 2014). Comparatively, the Rayleigh scattering has rarely been explored and used in natural water investigations. Rayleigh scattering measurements at certain excitation-emission wavelength pairs (e.g. 350/350 nm, 400/400 nm, and 500/500 nm) have been used to characterize the complexation between metal and water soluble soil organic matter (Ohno *et al.*, 2008) and monitor the metal-induced precipitation of soil fulvic acid and dissolved organic matter (DOM) (Ryan and Weber, 1982; Ohno and Cronan, 1997). Peiris *et al.*

(2010a) employed the entire light scattering region (1st and 2nd order Rayleigh scattering) to characterize the colloidal/particulate matter in surface water.

The utility of FEEM-based technique has been further enhanced by combining it with multivariate statistical techniques, such as parallel factor analysis (PARAFAC) (Baghoth *et al.*, 2011; Ishii and Boyer, 2012; Sanchez *et al.*, 2014; Shutova *et al.*, 2014) and principal component analysis (PCA) (Peiris *et al.*, 2010a; Chen *et al.*, 2014). PARAFAC and PCA (with matrix pre-treatment) can decompose the three-way FEEMs into trilinear and bilinear components, respectively (Bro, 1997; Eriksson *et al.*, 2001). A PCA model can fit data better but suffers from the intrinsic problem of rotational freedom (Bro, 1997). That is to say, even though the pure spectra of the target fluorophores can be reflected by the loadings in a bilinear decomposition, the physical significance of loadings cannot be interpreted or the pure spectra cannot be actually identified without external information (Bro, 1997). As a constrained version of PCA, a PARAFAC model has unique solutions and therefore is mathematically superior to PCA (Bro, 1997). In practice, the true underlying spectra of a trilinear dataset can be found through PARAFAC under the assumption that the signal-to-noise ratio and the selection of components are appropriate (Bro, 1997). In addition, Bro (1997) recommended that Rayleigh scattering should be avoided in PARAFAC due to its non-multilinear nature and therefore Rayleigh scattering were commonly removed prior to PARAFAC analysis (Andrade-Eiroa *et al.*, 2013). Compared to FEEM-PARAFAC, FEEM-PCA has gained increasing acceptance for NOM characterization with respect to drinking water treatment because of its ability to accommodate the Rayleigh scattering as a potential indicator of colloidal/particulate matter, which is directly relevant to various treatment issues (e.g. membrane fouling). This advantage has been highlighted in the literature (Andrade-Eiroa *et al.*, 2013).

In previous surface water studies, three statistically significant principal components (PCs) have been successfully extracted and were attributed to humic-like substances, colloidal/particulate matter, and protein-like substances (Peiris *et al.*, 2010a; Peiris *et al.*, 2011; Peldszus *et al.*, 2011; Chen *et al.*, 2014). The physical significance of these principal components (PCs) was determined by qualitatively comparing peak positions between the loading plots of extracted PCs and FEEM results for natural water as found in the literature (Peiris *et al.*, 2010a). This qualitative evaluation did not take into

account the impacts of NOM-NOM (both fluorescent and non-fluorescent) and NOM-inorganic colloid interactions on each PC.

In addition, the physical significance of the Rayleigh scattering-related PC is still not clear. In standard nephelometry, for a particle at a defined size the increase in particle number will cause a linear increase of Rayleigh scattering (Ryan and Weber, 1982). Yoon and Lueptow (2006) successfully developed a fluorescence excitation-emission spectroscopy-based technique to measure the concentration of Ludox colloidal silica with diameters between 7 and 22 nm. Also, the formation of larger-sized aggregates/particles, which have lower light scattering efficiency compared to smaller particles due to their size relative to the light wavelength, can further complicate changes in the Rayleigh scattering region (Ryan and Weber, 1982; Ohno and Cronan, 1997). For natural water with various NOM components and colloidal/particulate matter, the physical significance of Rayleigh scattering was rarely studied. Peiris *et al.* (2010a) attributed the Rayleigh scattering-related PC to the colloidal/particulate matter based on the empirical observations that the increase in colloidal/particulate matter enhanced the intensity of Rayleigh scattering regions. Being natural water investigations the composition of the colloidal/particulate matter is complex. Furthermore, it has been reported that protein can also have a direct impact on the light scattering intensity (Elshereef *et al.*, 2010). Therefore, there is a need to further investigate the contribution of both inorganic colloid and NOM-inorganic colloid interaction in natural water to the Rayleigh scattering region.

To-date the physical significance of PCs extracted through PCA of natural water FEEMs have been interpreted qualitatively. The major objective of this study is to quantitatively determine the physical significance of each PC extracted from FEEM-PCA analysis of NOM and inorganic colloids, and to investigate the impacts of NOM-NOM and NOM-inorganic colloid interactions on PC scores. This was achieved by performing FEEM measurements on defined systems of solutions of individual NOM model compounds and silica and mixtures thereof. In addition, the physical significance of the Rayleigh scattering region was further explored.

4.2 Materials and Methods

4.2.1 Model Solution Preparation

Four model compounds, including bovine serum albumin (surrogate for protein-like materials), humic acids (surrogate for humic substances), sodium alginate (surrogate for polysaccharides), and silica (surrogate for inorganic colloids), were selected to investigate the interaction between NOM fractions and inorganic colloids. Sodium alginate, bovine serum albumin, humic acid, and silica (Ludox® HS-30 colloidal silica) were obtained from Sigma Aldrich (St. Louis, MO). The physiochemical properties of the four model compounds are summarized in Table 4.1. Model solution preparation followed a 3^4 full factorial design with 4 factors/model compounds each at 3 concentrations. Bovine serum albumin, humic acid and alginate were at three concentrations of 0, 2.5, and 5 mg/L and silica was at 0, 100, and 200 mg/L. Stock solutions were made by dissolving the appropriate amount of each of the 4 model compounds in ultrapure water (MilliQ) without further purification and were used for model solution preparation. The stock solutions were not stored for more than 2 weeks at 4 ^{o}C. The TOC was additive for the model solutions with a mixture of model compounds. The pH of all model solutions ranged from 5.7 to 6.3. Model solutions were analyzed within 24 h once they were prepared.

Table 4.1 Physiochemical properties of four model compounds

Model compound	*Type*	*Carbon Content (%)*	*Shape*	*Charge*	*Size (kDa)*	*Average Hydrodynamic Diameter (nm)*	*Dimensions (nm)*
S: Silica	Inorganic	N/A	Round[c]	Negative (pH_{PZC}~3)[c]	N/A	~9[e]	N/A
B: Bovine serum albumin	Protein	0.46[a] 0.47[b]	N/A	pH_{IEP} = 4.7[c]	67[a]	7.6[a]	14×4×4[a]
H: Humic acid	Humic substance	0.31[a] 0.4[b]	Globular[c]	Negative (pH_{PZC}<3)[c]	70[a]	7.3[a]	7.3×7.3×7.3[a]
A: Sodium alginate	Poly-saccharides	0.29[b]	Extended random coil[c,d]	Negative[c,d]	30 – 100[b]	~0.01 – 1 μm[f]	N/A

[a]Xiao *et al.*, 2009; [b]Katsoufidou *et al.*, 2007; [c]Tang *et al.*, 2011; [d]Buffle *et al.*, 1998;

[e]From manufacturer Sigma Aldrich; [f]van de Ven *et al.*, 2008

N/A: not available

4.2.2 Fluorescence Measurements and Data Analysis

Fluorescence excitation emission matrices (FEEMs) for the model solutions were measured using a Varian Cary Eclipse Fluorescence Spectrofluorometer (Palo Alto, CA). Replicates for 16 of the model solutions were collected. A detailed description for acquisition of reproducible FEEM measurements can be found in Peiris *et al.* (2009). In this study, the FEEMs were obtained at a photomultiplier tube voltage of 750 V and slit widths of 10 nm for both excitation and emission. The excitation and emission wavelengths ranges were 250 – 380 nm and 300 – 600 nm, respectively. Each FEEM is a three dimensional matrix and contains 4,214 excitation and emission coordinate points which can be further re-arranged into one row for each FEEM using the rearrangement procedure in Peiris *et al.* (2010a). This pre-treatment procedure generated a 110 ×4,214 fluorescence data matrix (X). Principal component analysis (PCA) was then performed on the auto-scaled X data matrix. The detailed procedure is provided in Peiris *et al.* (2010a). The FEEM-PCA was conducted using PLS Toolbox 6.2.1 (Eigenvector Research, Inc., Manson, WA) within the MATLAB 7.11.0 computational environment (MathWorks, Natick, MA).

4.2.3 Investigation of Physical Significance of Principal Components

The multiple linear regression (MLR) with significance test was performed to explore the physical significance of each principal component. In addition, the impacts of NOM-NOM and NOM-inorganic colloid interaction on the PC scores were also investigated by comparing the PC scores of different model solutions. The response variable of the MLR models for each PC (i.e. PC1 or PC2 or PC3) are the PC scores obtained from FEEM-PCA analysis and explanatory variables include concentrations of the 4 model compounds in the different model solutions. Since the physical evidence for the interaction between protein-like substances and colloidal/particulate matter has been described (Peiris *et al.*, 2012), the interaction term between bovine serum albumin and silica was also included as an explanatory variable in the MLR models for PCs which are related to protein-like substances and colloidal/particulate matter. This interaction term between variables x_i and x_j can be expressed as the product of x_i and x_j (Rawlings *et al.*, 1998). Both R^2 and adjusted R^2 were estimated for each regression model. Adjusted R^2 rescales R^2 by degrees of freedom and is a better indicator when comparing the regression models with different numbers of explanatory variables compared to R^2 (Rawlings *et al.*, 1998). In addition, a t-test was conducted to assess the significance of any individual explanatory variables in each regression model. These statistical analyses were conducted using MATLAB 7.11.0.

Since R^2 (or adjusted R^2) represents the proportion of the total variation of the response variable explained by the explanatory variables (Rawlings *et al.*, 1998), the adjusted R^2 was used in combination with statistical significance of each explanatory variable to determine the model subset size. The full models including all explanatory variables and subset models that can account for a major proportion of the adjusted R^2 for the full model (i.e. this indicates that the subset model has similar performance as the full model and explanatory variable(s) included in the subset model are more important), are summarized in Table 4.2. The explanatory variable, which is statistically significant and can account for the highest proportion of adjusted R^2 of the full model, is determined as the most important contributor to the corresponding PC. For other statistically significant contributors which contribute much less to the adjusted R^2 of the full model, their effects on the fluorescent properties of the most important contributor were understood by comparing the corresponding PC scores.

4.3 Results and Discussion

4.3.1 Qualitative Investigation of Physical Significance of Principal Components

The proportion of the total variance explained by PC1, PC2, and PC3 are 68.81%, 15.60%, and 7.35%, respectively. The variance captured by each of the additional PCs is less than 2.5%, indicating that these PCs are not statistically significant. Therefore, FEEM-PCA analysis successfully extracted three statistically significant PCs, which captured 91.8% of the total variance in the raw FEEM data matrix.

Table 4.2 Illustration of contributions of model solution components and their interactions to each individual principal component (PC) using multiple linear regression (MLR)

	Model solution composition	R^2	*Adjusted* R^2	*P value*				
				A	*B*	*H*	*S*	*BSI*
PC1	ABHS	0.955	0.952	Y **(+)**	N **(–)**	**Y (+)**	Y **(+)**	**–**
	H only	0.927	0.926	**–**	**–**	**Y (+)**	**–**	**–**
PC2	ABHS + BSI	0.766	0.750	N **(+)**	Y **(+)**	Y **(–)**	**Y (+)**	N **(+)**
	ABHS	0.759	0.746	N **(+)**	Y **(+)**	Y **(–)**	**Y (+)**	**–**
	B + S + BSI	0.732	0.722	**–**	Y **(+)**	**–**	**Y (+)**	N **(+)**
	B + S	0.726	0.719	**–**	Y **(+)**	**–**	**Y (+)**	**–**
	S only	0.598	0.593	**–**	**–**	**–**	**Y (+)**	**–**
	B only	0.118	0.106	**–**	**Y (+)**	**–**	**–**	**–**
PC3	ABHS + BSI	0.546	0.516	N **(+)**	Y **(+)**	N **(+)**	**Y (–)**	N **(+)**
	ABHS	0.525	0.499	N **(+)**	**Y (+)**	N **(+)**	Y **(–)**	**–**
	B + S + BSI	0.532	0.514	**–**	Y **(+)**	**–**	**Y (–)**	N **(+)**
	B + S	0.509	0.497	**–**	**Y (+)**	**–**	Y **(–)**	**–**
	B only	0.366	0.358	**–**	**Y (+)**	**–**	**–**	**–**
	S only	0.153	0.142	**–**	**–**	**–**	**Y (–)**	**–**

A: Alginate; B: Bovine serum albumin; H: Humic acid; S: Silica

BSI: bovine serum albumin -silica Interaction (BSI = bovine serum albumin conc. × Silica conc.)

Y: significant (the p-values smaller than 0.05 which indicates explanatory variable is significant)

Note: The explanatory variable with the lowest p-values for each model are bolded

The details regarding p-values can be found in Table B1.

N: non-significant

"+" or "–" in parentheses is the sign of each explanatory parameter in the regression model

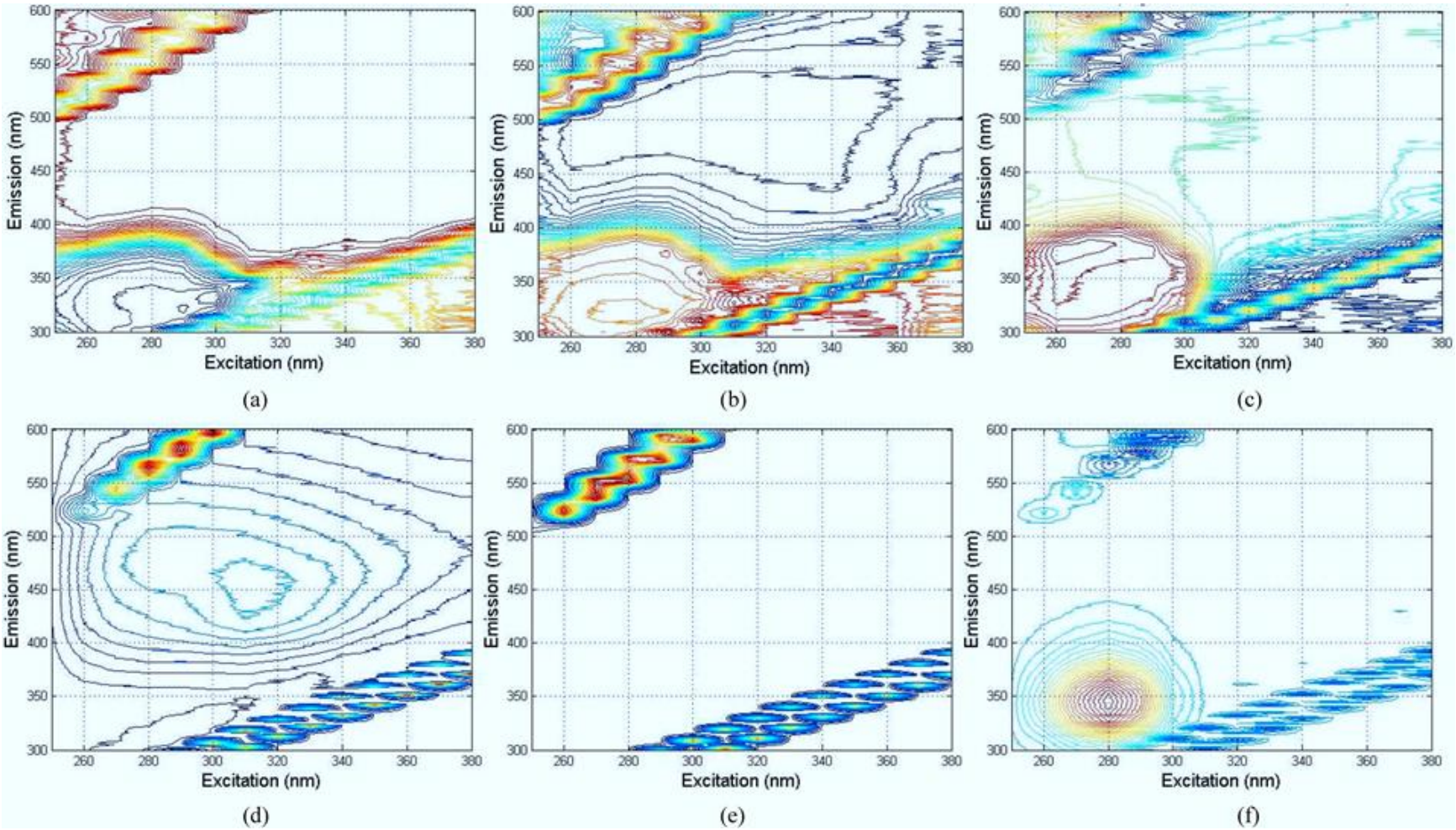
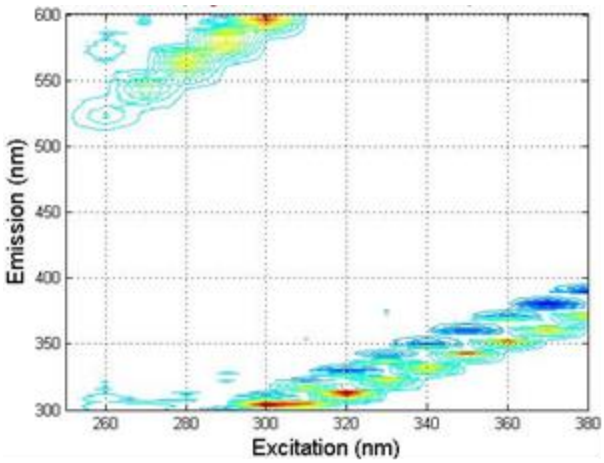


Figure 4.1



(g)

Figure 4.1 Loading plots of (a) PC1, (b) PC2, and (c) PC3 and original FEEM plots of model solutions with (d) humic acid (5 mg/L), (e) silica (200 mg/L), (f) bovine serum albumin (5 mg/L), and (g) alginate (5 mg/L)

The loading plots for three PC were compared with the original FEEM plots of the single-compound model solutions with humic acid (5 mg/L), bovine serum albumin (5 mg/L), and silica (200 mg/L) to qualitatively investigate the physical significance of each PC (Figure 4.1). The loading plot for PC1 has a distinct peak at the same region where the FEEM peak related to humic acid is situated, indicating that PC1 is related to humic acid. In the loading plot of PC2, distinct peak arrays are observed at the same regions where the 1st and 2nd order Rayleigh scattering regions are situated in the FEEM plots of model solutions. Even though Rayleigh scattering regions can be observed in all 4 model compounds, silica has much higher intensities in the Rayleigh scattering regions compared to the other 3 model compounds and hence is likely to be the major contributor to PC2 (Figure B1). In addition, another peak in the PC2 loading plot is situated at the same region where the FEEM peak related to bovine serum albumin occurs. Therefore, PC2 appears to be related to both silica and bovine serum albumin. The loading plot for PC3 has a distinct peak at the same region where the FEEM peak related to bovine serum albumin occurs, suggesting PC3 can be mainly attributed to the bovine serum albumin.

As a non-fluorescent compound, the FEEM plots of alginate-only model solution (both 2.5 and 5 mg/L) have low intensity values in both Rayleigh scattering and fluorescence regions (Figure B1), suggesting alginate is not directly associated with any of the PCs although the impact of interaction between alginate and other model compounds on the PC scores will be discussed later.

4.3.2 Quantitative Investigation of Physical Significance of PC1

The results of MLR analysis and significance test are summarized in Table 4.2. In Table 4.2, humic acid is seen to be a significant positive contributor to PC1 in both MLR models. Furthermore, the adjusted R^2 (0.93) of the model with humic acid is only slightly smaller than the adjusted R^2 (0.95) obtained with the original model with 4 model compounds, suggesting PC1 is dominated by humic acid. The results from the quantitative investigation of physical significance of PCs corroborated with the qualitative investigation. In addition, silica and alginate were found to be statistically significant and make a positive contribution to PC1. This is also illustrated in Figure 4.2 where PC1 scores increase with silica and alginate concentration even though this synergistic effect could be low in some cases. Regarding the synergistic effect between humic acid and alginate, the increase of PC1 score is more obvious at higher humic acid concentrations (i.e. 5 mg/L) and the PC1 score increased more at higher alginate concentrations (Figure 4.2a). For the synergistic effect of humic acid-silica at both humic acid concentrations (i.e. 2.5 and 5 mg/L), the increase of PC1 score was shown to be more distinct as the silica concentration increased (Figure 4.2b).

The reasons for the humic acid-silica and humic acid-alginate synergistic effect on PC1 score were unclear. Adsorption of humic acid onto silica has been reported (Jekel, 1986; Fairhurst *et al.*, 1995; Liang *et al.*, 2011) and suspended silica colloids can be stabilized by humic acid (Jekel, 1986). Liang *et al.* (2011) summarized that both cation bridging and hydrogen bonding can contribute to the humic acid-SiO_2 adsorption. The potential conformational changes of humic acid molecules upon adsorption might be responsible for the humic acid-silica synergistic effect. The interaction between alginate and humic acid was investigated by Myat *et al.* (2014) using liquid chromatography and molecular dynamics simulations. They reported that the water-mediated divalent ion bridging is likely to be the dominating mechanism for the alginate-humic acid interaction at a neutral pH conditions (Myat *et al.*, 2014). Also, the UV absorption at 254 nm was enhanced for the high molecular weight peak region when alginate-humic acid complexation occurred, revealing the potential conformational change of humic acid molecules during the complexation since the alginate has no UV_{254} absorbance (Myat *et al.*, 2014). In addition, the UV_{254} increase was shown to be more significant for alginate-humic acid mixture in electrolyte solutions (i.e. NaCl, $CaCl_2$, KCl, and $MgCl_2$) compared to mixtures in deionized water (i.e. no significant UV_{254} increase), indicating this effect is ionic environment-dependent (Myat *et al.*, 2014). These results from Myat *et al.* (2014) suggest that the alginate-humic

acid synergistic effect on PC1 scores could also be possibly induced by the conformational change of humic acid due to the alginate-humic acid complexation mediated by cations.

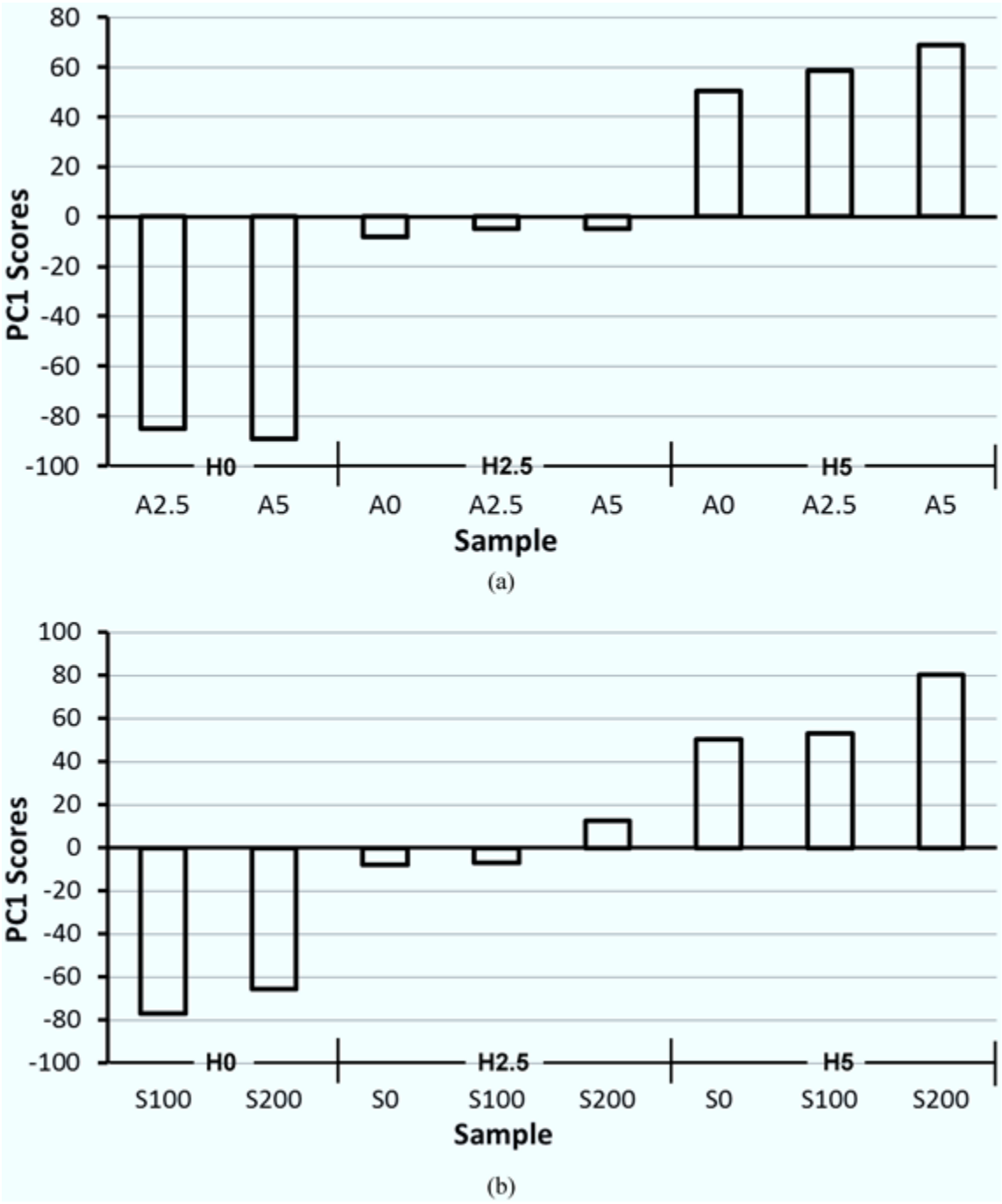
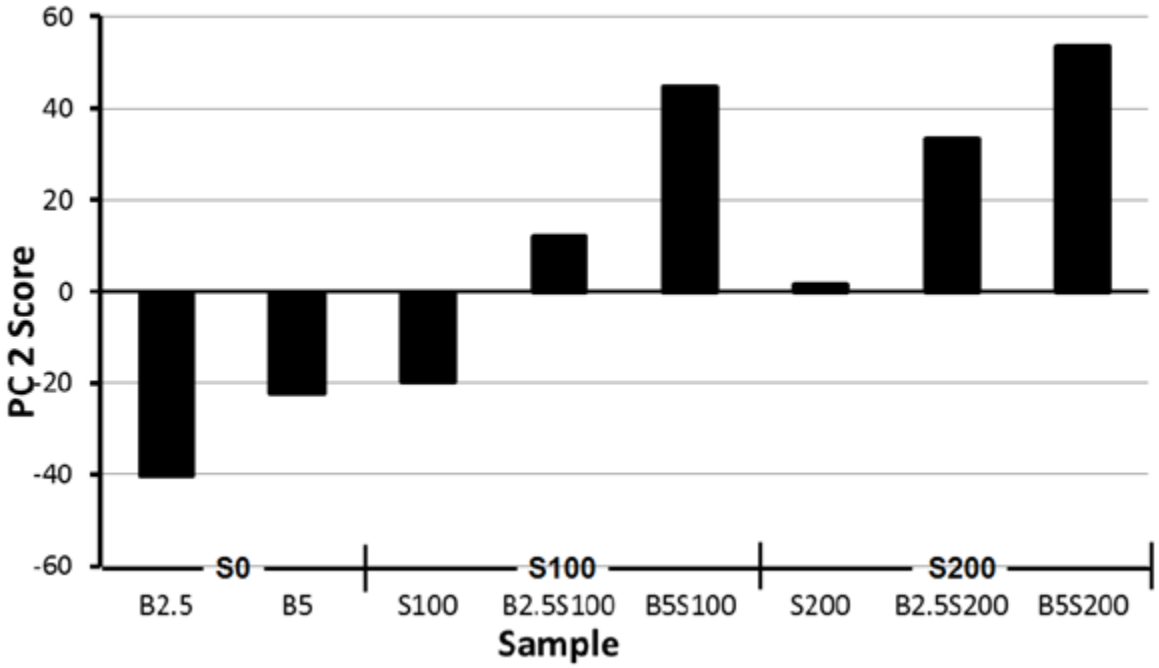
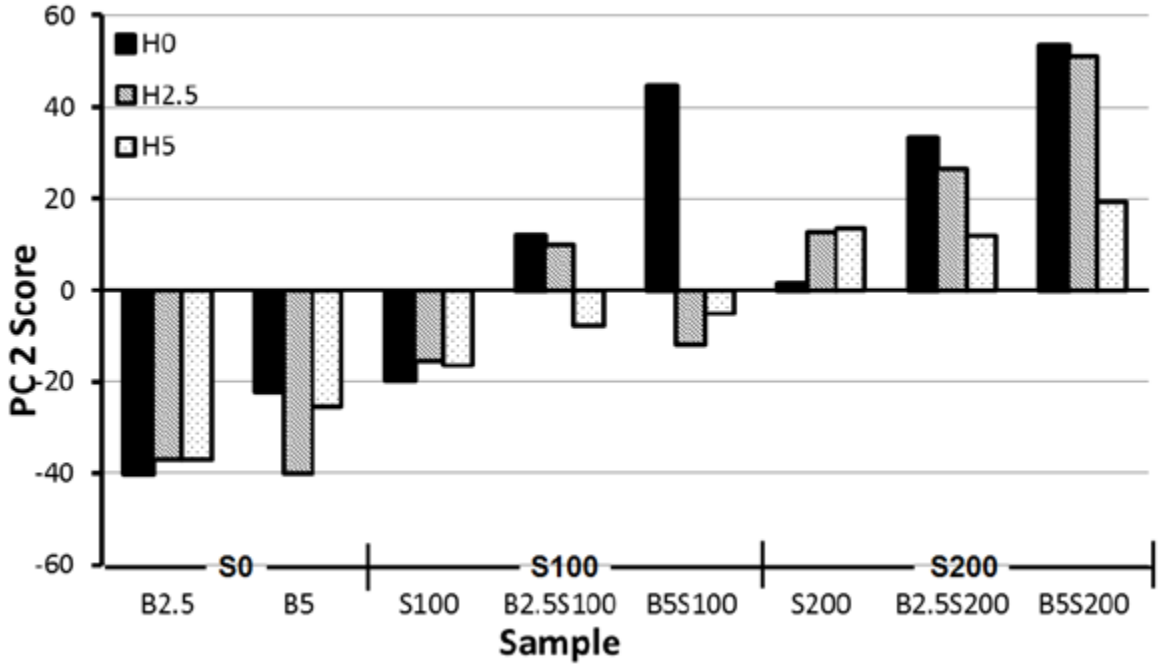


Figure 4.2 Illustration of the synergistic effect of (a) humic acid (H)-alginate (A) and (b) humic acid (H)-silica (S) on PC1 scores



(a)



(b)

Figure 4.3 Illustration of the synergistic effect of (a) bovine serum albumin (B)-silica (S) on PC2 scores and (b) the impact of humic acid (H) on the PC2 scores dominated by B-S system

4.3.3 Quantitative Investigation of Physical Significance of PC2

Due to the co-existence of Rayleigh scattering peak arrays and protein-like peak in PC2 loading plots, the interaction term between bovine serum albumin and silica (BSI) was included in the MLR analysis for PC2. As seen in Table 4.2, the MLR model with bovine serum albumin and silica has a slightly lower adjusted R^2 compared to the full model with 4 model compounds and BSI (0.72 out of 0.75). Silica is shown to be a significant positive contributor and contribute most to adjusted R^2 (0.59 out of 0.75). Another significant explanatory variable, bovine serum albumin, is the second most important positive contributor to PC2 (Table 4.2). This suggests that silica is the dominant contributor to PC2. As illustrated in Figure 4.3a, PC2 scores increase with silica concentration with and without the presence of bovine serum albumin that is corroborated with the results reported by Yoon and Lueptow (2006). Within the concentration range of 0 – 300 mg/L SiO_2 at pH = 7, the fluorescence intensities of colloidal silica particle suspensions at Ex/Em ~ 308 nm/318 nm were reported to be linearly correlated with the silica concentration ($R^2 > 0.98$) (Yoon and Lueptow, 2006). This linear relationship was observed for 7 nm, 12 nm, and 22 nm (in average diameter) colloidal silica particles with either positive or negative charge (Yoon and Lueptow, 2006).

In addition, the effect of bovine serum albumin on PC2 is also clearly shown in Figure 4.3a. The samples with both bovine serum albumin and silica had higher PC2 scores compared to the samples with bovine serum albumin or silica only. For a given silica concentration, increasing the bovine serum albumin concentration substantially increased the PC2 response. A similar synergistic effect between two compounds was also observed at each given bovine serum albumin concentration. Also, it should be noted that bovine serum albumin at a concentration of 5 mg/L had a similar PC2 score as 100 mg/L silica. Elshereef *et al.* (2010) reported that an increase in protein concentration will cause a linear increase of light scattering intensity in the concentration range from 0 to 4 g/L, and it might be assumed that this linear relationship may also apply at the much lower protein concentrations as reported in this study. In addition, other factors, such as protein-protein interactions, presence of covalently formed aggregates, and average molecular weight of protein, can also affect the light scattering intensity and pattern for protein solutions (Elshereef *et al.*, 2010). These findings confirmed that bovine serum albumin is a direct contributor to Rayleigh scattering region and explains the co-existence of bovine serum albumin-related peak and silica-related Rayleigh scattering region in the PC2 loading plot (Figure 4.1).

Another statistically significant variable, humic acid, was shown to contribute negatively to PC2; however, the negative contribution to PC2 is not consistent in all cases. As illustrated in Figure 4.3b, humic acid made a positive contribution to PC2 scores of single compound solutions (e.g. B2.5, S100 and S200) but contributed negatively to PC2 scores of model solutions when in combination with silica and bovine serum albumin or high concentrations of bovine serum albumin (e.g. B5). Therefore, humic acid is a negative contributor to PC2 overall and is likely to contribute negatively to PC2 in the context of natural water with complex composition.

The above findings contribute to better understanding the origins of the Rayleigh scattering (RS) region and the physical significance of PC2. In previous studies (Peiris *et al.*, 2010a, 2010b, 2011), the physical significance of PC2 was qualitatively attributed to the colloidal/particulate matter, which is defined by the authors as a mixture of inorganic colloids and NOM-like material (Peiris *et al.*, 2010b). This study confirmed that both inorganic colloid (e.g. silica) and protein (e.g. bovine serum albumin) can directly affect the light scattering intensity. Therefore, the Rayleigh scattering region in FEEM can be attributed to both inorganic colloids and protein. Additionally, the presence of humic acids can further complicate the light scattering pattern.

4.3.4 Quantitative Investigation of Physical Significance of PC3

From Table 4.2 it is seen that bovine serum albumin is the dominant positive contributor to PC3 and confirmed by the qualitative relationship that PC3 is directly related to bovine serum albumin (Figure 4.1). The low adjusted R^2 can likely be attributed to the lack of other significant explanatory variables in addition to the model compound concentration and the high variability of PC3. For a given silica concentration, increasing the concentration of bovine serum albumin significantly increased the PC3 response (Figure 4.4). Also, silica is shown to be a significant negative contributor to PC3 (Table 4.2). It should be noted that the PC2 and PC3 models are very different from each other. In the PC2 model, bovine serum albumin and silica are both significant positive contributors, indicating both compounds tend to directly contribute to light scattering. However, in the PC3 model, bovine serum albumin and silica are positive and negative contributors to PC3, respectively. This indicates that bovine serum albumin is directly related to PC3 but silica tends to affect the PC3 through the interaction with bovine serum albumin in a negative way. As illustrated in Figure 4.4, when the concentration ratio between silica and bovine serum albumin (silica:bovine serum albumin) $\geq$ 40:1 (there is sufficient silica available in the solution), PC3 scores of bovine serum albumin-silica solution were smaller

compared to bovine serum albumin alone at both bovine serum albumin concentrations, reflecting the existence of a bovine serum albumin-silica interaction which has been widely reported in the literature (Hlady *et al.*, 1986; Norde and Favier, 1992; Clark *et al.*, 1994; Larsericsdotter *et al.*, 2005). Also, the decrease of PC3 score after adding silica is greater when the bovine serum albumin concentration is lower (i.e. 2.5 mg/L) but the PC3 score decrease at the bovine serum albumin of 2.5 mg/L did not appear to be more significant for a higher silica concentration (i.e. 200 mg/L). As a soft protein with low conformational stability, bovine serum albumin tends to rearrange its molecular structure upon adsorption (Bos *et al.*, 1994; Clark *et al.*, 1994). Furthermore, the adsorption behaviour of bovine serum albumin is not strongly affected by the electrostatic interactions or the surface charge of the sorbent because adsorption can induce a gain in conformational entropy which can act as the predominant driving force for bovine serum albumin adsorption (Norde and Favier, 1992; Bos *et al.*, 1994; Nakanishi *et al.*, 2001).

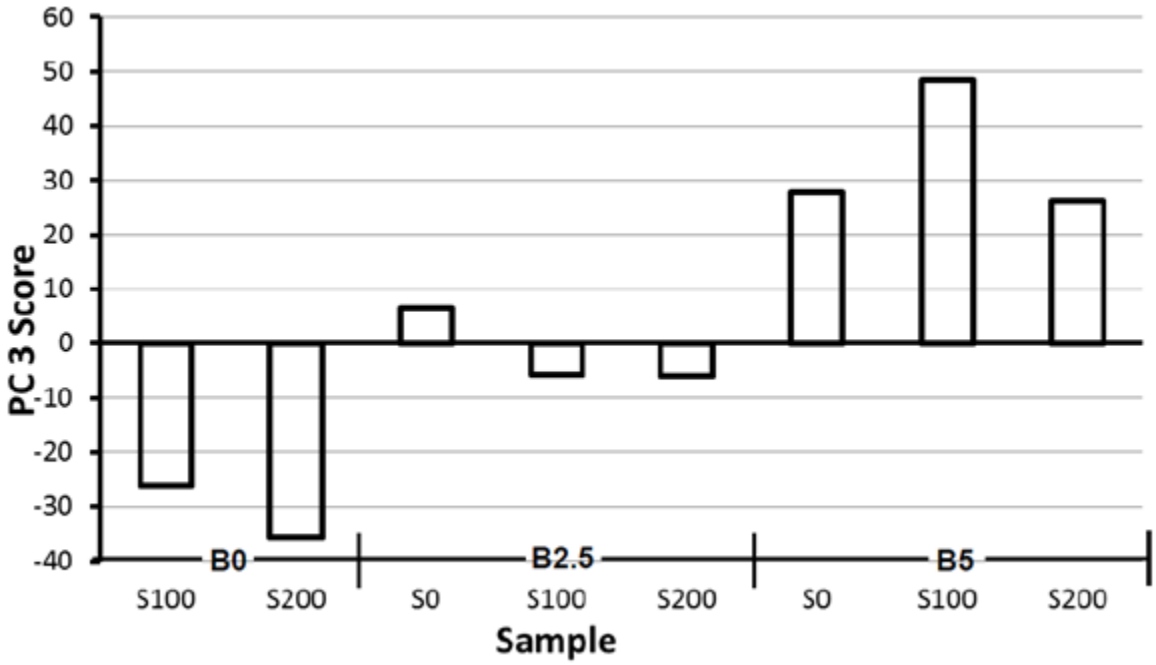


Figure 4.4 Illustration of the antagonistic effect of bovine serum albumin (B)-silica (S) on PC3 scores

In this study, the antagonistic effect between bovine serum albumin and silica on PC3 is in agreement with the results of Clark *et al.* (1994). Compared with the unadsorbed bovine serum albumin (bovine serum albumin-only), a significant decrease (55%) of intrinsic fluorescence intensity (normalized with respect to the unadsorbed protein) was observed when bovine serum albumin in solution was

adsorbed onto the silica nanoparticles (Clark *et al.*, 1994). In their study the pH was adjusted to be close to the isoelectric point of the bovine serum albumin, thereby minimizing the electrostatic interactions between bovine serum albumin and silica and hence the formation of large aggregates which can increase scattering. They attributed the decrease in the intrinsic fluorescene intensity of the adsorbed bovine serum albumin to the conformational change upon the adsoprtion. It is likely that the two trptophan (TRP) residues (TRP-134 and TRP-212) located in subdomains 1-C and 2-AB were positioned at the sorbent (i.e. silica) surface by the adsorption orientation of bovine serum albumin (Clark *et al.*, 1994). Larsericsdotter *et al.* (2005) confirmed that bovine serum albumin is adsorbed onto silica with an orientation such that Domain 2 is placed at the sorbent surface. Also, the fluorescence quantum yield emited by adsorbed bovine serum albumin onto hydrophilic amorphous silica was found to be lower compared to the bovine serum albumin alone solution (Hlady *et al.*, 1986). The authors also proposed that conformational changes resulting from adsoprtion were possibly the reason for this phenomenon (Hlady *et al.*, 1986).

In addtiion, Peiris *et al.* (2012) demonstrated that the binding of protein-like substances onto a self-assembled monolayer (SAM) which was desinged to enhance strong protein binding was reduced due to the interaction between protein-like substances and colloidal/particulate matter extracted from natural river water. This phenomenon is likely induced by the ability of colloidal/particulate matter to mask the sites available for protein-like substances to bind onto the SAM (Peiris *et al.*, 2012). Their work illustrated the physical evidence for the interaction between protein-like substances and inorganic colloids contained in natural colloidal/particulate matter.

Note also that PC3 scores of bovine serum albumin-silica solution (i.e. B5S100) is higher (Figure 4.3) than the bovine serum albumin-only solution (i.e. B5). This synergistic effect between bovine serum albumin and silica on PC3 only occurred when the concentration ratio between silica and bovine serum albumin was low (silica:bovine serum albumin = 20:1). The reason for this phenomenon is unclear and more fundamental studies are required.

4.3.5 Implications of Model Solution Study for Natural Water Systems

In this model solution study, PC2 as an example was shown to be a complicated term which could be a direct indicator of both inorganic colloids/particulates and proteins. Even though silica is determined to be the dominant contributor, the presence of proteins as well as humic acids can further

complicate the scattering pattern and hence the interpretation of PC2. Depending on the source of natural water, PC2 is likely to be dominated by the inorganic colloidal/particulate matter because the concentration of protein-like materials in natural water can be very low and its direct contribution to scattering is much less compared with inorganic colloids. However, it should be noted that the interaction of inorganic colloidal/particulate matter with humic-like substances and protein-like materials can also been reflected by the PC2. In general, the extracted PCs contain the overall information of the corresponding dominant contributors, including their intrinsic fluorescent properties as well as their fluorescent property change (e.g. due to conformational change) through the interaction with other NOM components and inorganic colloidal/particulate matter.

4.4 Conclusions

This study quantitatively investigated the physical significance of principal components (PCs) extracted from the PCA of FEEMs for four model compounds (i.e. humic acids, bovine serum albumin, sodium alginate, and silica) and their mixtures by performing multiple linear regression. In addition, the impact of NOM-NOM and NOM-inorganic colloid interactions on the PC scores was also investigated. The major conclusions are:

Each principal component extracted from FEEM-PCA provides information regarding the fluorescence properties of a particular NOM component and its interaction with other NOM components and inorganic colloids. The key contributors to each principal component were humic acid for PC1, silica for PC2 (i.e. Rayleigh scattering and protein-like substances) and protein-like substances for PC3. Silica and alginate were found to be two statistically significant positive contributors to PC1. Both inorganic colloids and protein were found to be direct contributors to Rayleigh scattering (PC2) and their co-existence with humic acid can further complicate the light scattering pattern. Proteins can physically interact with inorganic colloids and induce the antagonistic effects between two compounds which can impact the PC3 scores. FEEM-PCA has the potential to be used in the investigation of NOM-NOM and NOM-inorganic colloid interactions.

Chapter 5
A Comparison of Data Analysis Methods for Fluorescence Excitation Emission Matrix (FEEM) Characterization of Natural Organic Matter

Summary

The objectives of this study were to compare the utility of various FEEM data analysis methods for natural organic matter (NOM) characterization and to investigate the potential impacts of photomultiplier tube (PMT) voltage setting on the FEEM data analysis results. Water samples from a pilot-scale biofiltration-ultrafiltration treatment train were collected and FEEMs were obtained at 4 different PMT voltage settings. Principal component analysis (PCA) and peak picking were performed to analyze these FEEM datasets. Peak picking had consistent performance on fluorophore characterization at 4 different PMT conditions and intensity values of selected excitation emission coordinates increased with increasing PMT voltage as expected. PMT voltages settings were shown to have little impact on FEEM-PCA results. Especially, saturation issue of Rayleigh scattering at higher PMT was successfully accommodated by FEEM-PCA. Based on the information obtained from this study and literature results, the nature and utility of these two FEEM data analysis methods were compared with parallel factor analysis (PARAFAC). It is recommended that FEEM-PCA should be employed for the monitoring and prediction of effects posed by natural organic matter (NOM) components as well as colloidal/particulate matter which can be characterized using Rayleigh scattering by FEEM-PCA. FEEM-PARAFAC and peak picking have superior performance in the characterization of the fluorophores alone since both approaches cannot extract the information from Rayleigh scattering region together with fluorophores.

5.1 Introduction

Natural organic matter (NOM) has been widely reported to be related to various drinking water treatment issues such as formation of carcinogenic disinfection by-product (e.g. Zheng *et al*., 2015), performance deterioration of treatment processes (e.g. Baghoth *et al*., 2011), and promotion of bacterial regrowth in drinking water distribution networks (e.g. Jacangelo *et al*., 1995). To effectively characterize the complex physicochemical properties of NOM and monitor its behaviour in the drinking water treatment system, various promising characterization approaches have been developed and successfully employed, including XAD resin fractionation (e.g. Leenheer, 1981; Maurice *et al*.,

2002), liquid chromatography – organic carbon detection (LC-OCD) (Huber *et al.*, 2011), and fluorescence excitation emission matrices (FEEM) (e.g. Baghoth *et al.*, 2011). The FEEM technique has gained more acceptances and is widely recommended for NOM characterization because of its relatively low cost, simple measurement procedure, high sensitivity, minimal sample pre-treatment, non-destructive nature and minimum modification of original NOM properties (Chen *et al.*, 2003a; Peiris *et al*., 2008; Baghoth *et al*., 2011).

Peak picking, which obtains intensity values at certain excitation and emission wavelengths, has been widely used (e.g. Her *et al.*, 2003; Bieroza *et al.*, 2011) in NOM characterization studies. However, it has been pointed out that this method cannot reflect the heterogeneity of aromatic dissolved organic matter (DOM) because this method only uses a few data points instead of entire FEEM spectra for quantification (Chen *et al.*, 2003a). Chen *et al.* (2003a) proposed a new approach named fluorescence regional integration (FRI) which can use entire FEEM spectra for quantification. In this method, FEEM spectra were delineated into five regions which were quantified separately by integrating the volume of each FEEM region (Chen *et al.*, 2003a). However, this approach is not a "variance-based" method which cannot delineate the regions "automatically" based on the heterogeneity present across different FEEM regions. Instead, the delineation of regions is still based on the dominant peak locations and regional boundaries could vary depending on the interest of study (Chen *et al.*, 2003a). Therefore, the nature of this approach is similar to that of peak picking. The key difference is that peak picking uses one Ex/Em coordinate for quantification of a certain fluorophore whereas FRI employs an Ex/Em region for quantification. Recently, multivariate analysis methods, including parallel factor analysis (PARAFAC) and principal component analysis (PCA), have been gaining more popularity for FEEM data analysis (Peiris *et al.*, 2010a; Bieroza *et al.*, 2011; Murphy *et al.*, 2011; Ishii and Boyer, 2012; Shutova *et al*., 2014; Peleato and Andrews, 2015a) because these approaches can determine the fluorophore and other components "automatically" by conducting statistical analysis using the entire FEEM spectra and heterogeneity among different FEEM regions.

Comparisons among these FEEM data analysis methods have been conducted by several studies. The mathematical differences between PCA and PARAFAC have been comprehensively discussed in Bro (1997). PARAFAC is essentially a constrained version of a two-way PCA which has more flexibility and can fit data better than PARAFAC (Bro, 1997). PARAFAC is considered to be mathematically

superior because the PARAFAC model is more interpretable and has the solution uniqueness (Bro, 1997). However, it should be noted that all PARAFAC models have to be built based on FEEMs without the Rayleigh scattering (Bro, 1997) which has been shown to contain information about colloidal/particulate matter (Peiris *et al.*, 2010a). The utility comparison between PARAFAC and PCA has been conducted in several studies. Peleato and Andrews (2015b) compared the performance of PARAFAC and PCA regarding the NOM monitoring along a water treatment train. Also, the prediction power between components extracted from PCA and PARAFAC has been compared in Omrani *et al.* (2012) and Peleato and Andrews (2015a). Findings from these studies are discussed in more detail in later sections. For the comparison between PARAFAC and peak picking, previous studies have reported that PARAFAC scores were well correlated with intensities obtained through peak picking (Bieroza *et al.*, 2011; Shutova *et al.*, 2014). These studies proposed that the sophisticated data processing could hinder the wide application of FEEM for online monitoring even though the FEEM data acquisition becomes more efficient due to the advancement of spectrofluorometric instrumentation (Shutova *et al.*, 2014). Shutova *et al.* (2014) have attempted to use PARAFAC to identify fluorescence excitation emission wavelengths which can be used for online NOM monitoring (Shutova *et al.*, 2014), re-emphasizing the value of peak picking as a promising FEEM data analysis tool.

Since each approach has its own advantages and limitations, there is a need to comprehensively discuss the utility of various FEEM data analysis approaches regarding the NOM characterization and to provide FEEM users with guidance on the selection of an approach in various situations. It has been noticed that a knowledge gap exists regarding the comparison between FEEM-PCA and peak picking. Comparison between these two methods can provide more insight on how PCA works and the physical meaning of PCA components. FEEM-PCA is capable of modeling Rayleigh scattering regions which can be saturated when using the optimized instrument settings for the fluorophores of interest. It is therefore of interest to investigate the impacts of instrumentation settings on FEEM-PCA and whether these will affect the application of FEEM-PCA. Also, the performance of peak picking at different instrument settings will provide valuable information for the situations where two fluorophores of interest have dramatically different concentrations in the monitored water. The discussion in this study will focus on the FEEM-PCA, FEEM-PARAFAC, and peak peaking.

The major objective of this study was to investigate the impacts of PMT voltage settings on the performance of peak picking and FEEM-PCA. Based on the synthesis of the results from this study and information from the literature pertaining to PARAFAC, a comprehensive comparison of the utility between these two FEEM data analysis methods and FEEM-PARAFAC for NOM characterization was conducted and recommendations on the approach selection in different situations were provided.

5.2 Methods and Materials

5.2.1 Water Sample Sources

A pilot-scale setup, including two dual media (sand and anthracite) biofilters followed by an ultrafiltration membrane unit, was operated using Saugeen River water (Ontario, Canada) to investigate the performance of direct biofiltration as an ultrafiltration pre-treatment alternative. The major proportion of dissolved organic carbon was humic-like substances (Rahman, 2013). Raw water, effluents from roughing filters and biofilters as well as membrane permeate were collected for FEEM measurements from February, 2013 to May, 2013. More experimental and setup operation details can be found in Rahman (2013).

5.2.2 FEEM Measurements

Collection of FEEMs was conducted using a Cary Eclipse fluorescence spectrophotometer (Agilent Technologies, Canada). Excitation and emission wavelengths ranged from 250 to 380 nm (with a sequential increment of 10 nm) and 300 to 600 nm (with an increment of 1 nm at 600 nm/min scan rate), respectively. Both the excitation and emission slit width were set to 10 nm. FEEMs were measured at varied photomultiplier tube (PMT) voltages of 550 V, 600 V, 650 V, and 700 V. Milli-Q FEEMs measured at the same instrument settings were used for the blank subtraction to eliminate water Raman scattering and background noise (Peiris *et al.*, 2010a). The number of collected samples including replicates (i.e. re-measurement of a sample) for some samples at PMT voltage of 550 V, 600V, 650 V, and 700 V are 62, 61, 63, and 69, respectively. More details regarding FEEM measurement procedure can be found in Peiris *et al.* (2009). The collected water samples were measured without 0.45 μm pre-filtration to capture the Rayleigh scattering inherent to the unaltered sample and FEEM measurements were completed within 2 days after sample collection.

5.2.3 FEEM Data Analysis Methods

5.2.3.1 Peak Picking

In this study, intensities at Ex/Em = 270 nm/460 nm (Sierra *et al*., 2005; Peiris *et al*., 2008), Ex/Em = 320 nm/415 nm (Coble *et al*., 1990; Sierra *et al*., 2005), and Ex/Em = 280 nm/330 nm (Chen *et al*., 2003a; Peiris *et al*., 2010a) were used for the semi-quantification of humic acid (HA)-like substances, fulvic acid (FA)-like substances, and protein-like materials, respectively.

5.2.3.2 Principal Component Analysis (PCA)

PCA was conducted on each FEEM dataset collected for each PMT voltage. Overall, 4 separate PCA were conducted, one for each PMT setting. For the FEEM of each sample, 4214 excitation and emission coordinates were re-arranged into one row using the procedure provided by Peiris *et al.* (2010a). This procedure generated 4 data matrices (sample number×4214) for 4 different PMT voltages conditions. Then, each data matrix was auto-scaled and analyzed using PCA analysis. For each PCA model, all samples within each data matrix were utilized together to extract the information. The PCA analysis was performed using PLS Toolbox 6.2.1 (Eigenvector Research, Inc., Manson, WA). More details regarding the data pre-treatment and PCA procedure can be found in Peiris *et al.* (2010a). In this study, two outlier water samples were not included in the data analysis because of measurement error (e.g. intensity at PMT = 700 V, which should be higher than the intensity at PMT = 650 V, is equal to the intensity at PMT = 650 V).

5.2.3.3 Comparison of FEEM Data Analysis Results

R^2 value obtained from linear regression analysis was used to compare the FEEM data analysis results. R^2 can reflect the proportion of the total variance of the one variable that can be explained by the other variable (Rawlings *et al*., 1998). In this study, R^2 value was employed to interpret the proportion of the variance (e.g. fluctuation) of intensity values (or PC scores) at one PMT setting that can be explained by the intensity values (or PC scores) at the other PMT setting. Similarly, this parameter was also used to interpret the proportion of the total variance of scores of a certain PC that can be explained by the corresponding intensity values. These results were summarized in Tables 5.1 and 5.2.

5.3 Results and Discussions

5.3.1 Impacts of PMT Setting on FEEM Peak Picking Results

Figure C1a illustrated the FA intensity values of all samples at 4 different PMT settings. As expected, intensities at each selected Ex/Em coordinate increased with PMT voltages. Table 5.1 summarized the relationship of FEEM data analysis results (i.e. intensities or PC scores) between each pair of PMT voltages settings. The relationship of each parameter between two PMT settings (e.g. PMT550 vs. PMT600) was evaluated using R^2 value. For each selected Ex/Em coordinate, intensity values were well correlated between each pair of PMT voltages (Table 5.1). The high R^2 values indicate that different NOM fractions could be well detected and quantified even at the lowest PMT voltage condition. For HA-, FA-, and protein-like peak intensities, the R^2 values between intensities obtained at the PMT voltage of 700 V and the other PMT conditions (i.e. 550 V, 600 V, and 650 V) were shown to be slightly lower (Table 5.1), suggesting the variance contained within the dataset collected at 700 V and other PMT voltage settings might be slightly different.

Table 5.1 Relationships of FEEM data analysis results (intensities from peak picking and PC scores from FEEM-PCA) between each pair of PMT settings. The relationship is evaluated using R^2 value

	PMT550			*PMT600*		*PMT650*
	PMT600	PMT650	PMT700	PMT650	PMT700	PMT700
PC1 score	0.997	0.992	0.975	0.998	0.980	0.983
HA intensity	0.989	0.984	0.946	0.994	0.951	0.949
FA intensity	0.986	0.981	0.949	0.988	0.958	0.971
PC2 score	0.996	0.974	0.948	0.984	0.962	0.986
PC3 score	0.982	0.972	0.944	0.989	0.957	0.977
PC4 score	0.939	0.918	0.894	0.985	0.954	0.958
Protein intensity	0.993	0.989	0.978	0.991	0.971	0.982

It is noted that protein-like intensity values of some samples were very close to zero at PMT voltage of 550 V, indicating that the instrument was reaching the detection capacity of protein-like materials at this PMT voltage condition and measurement conducted at lower PMT voltage might not be feasible. The above observations demonstrated that PMT settings will not dramatically affect the peak picking results when the instrumentation settings are sufficiently sensitive for the detection of fluorophores(s). However, in order to ensure a reasonable performance of peak picking, it is crucial to

optimize the instrumentation settings to detect fluorophores at both high and low concentrations when they co-existed in the water samples and both fluorophores were components of interest. In addition, Rayleigh scattering region had saturation issues to different extents with increase of PMT voltage. Peak picking cannot handle this problem (i.e. intensities from saturated region are not representative) and therefore information contained in this region cannot be utilized.

5.3.2 Impacts of PMT Setting on FEEM-PCA Results

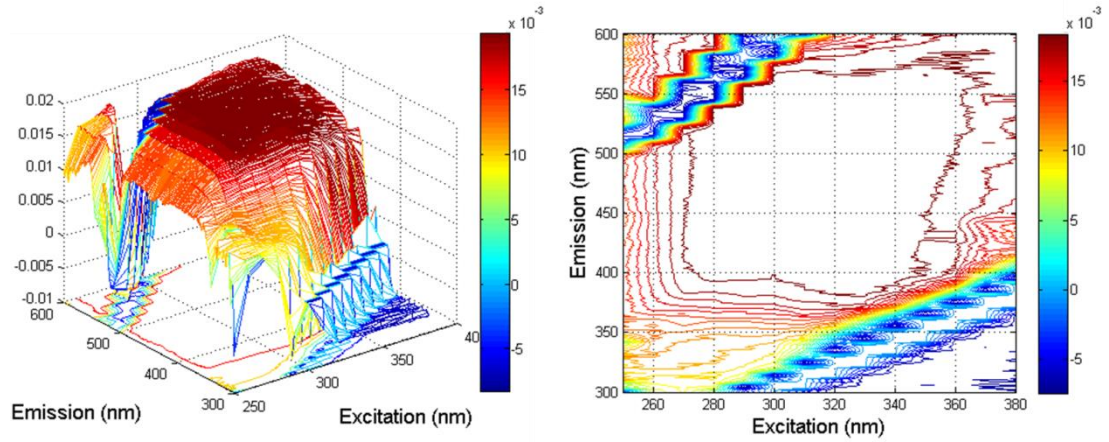
5.3.2.1 Principal Components (PCs) at Different PMT Settings

For the datasets collected at 4 different PMT settings, 4 PCAs were performed, one for each PMT setting. All 4 PCAs extracted four major principal components (PCs) which were similar in nature when comparing them among the PCAs from the different PMT settings. As an example the loading plots of the four PCs extracted from FEEMs measured at PMT voltage of 700 V are illustrated in Figure 5.1.

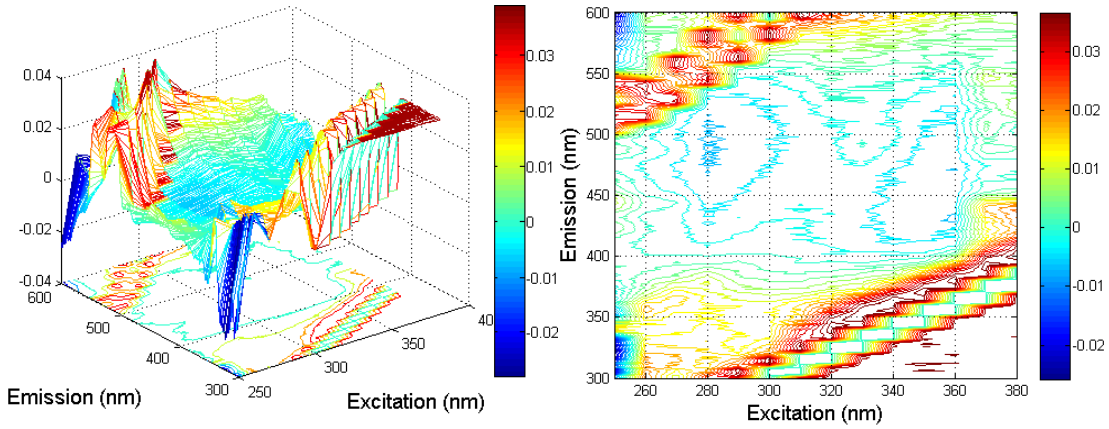
For PC1, the loading plots were dominated by peaks located at coordinates similar to those of the fluorescence peaks of HA- (Ex/Em = 270 nm/460 nm) and FA-like (Ex/Em = 320 nm/415 nm) materials (Coble *et al*., 1990; Sierra *et al*., 2005; Peiris *et al*., 2008) (Figure 5.1a), indicating that PC1 is mainly related to the humic-like substances. The percentage of total variance captured by PC1 were 60.96%, 62.35%, 62.65%, and 64.87% for the PCAs of the FEEMs collected at PMT voltages of 550 V, 600 V, 650 V, and 700 V, respectively.

In PC2 loading plots (Figure 5.1b), the presence of two arrays of peaks at the same locations where the Rayleigh scattering is situated suggested that PC2 is associated with colloidal/particulate matter (Peiris *et al*., 2010a). Percent variances captured by PC2 for datasets measured at PMT voltages of 550 V, 600 V, 650 V, and 700 V were 19.34%, 18.66%, 17.06%, and 15.75%, respectively.

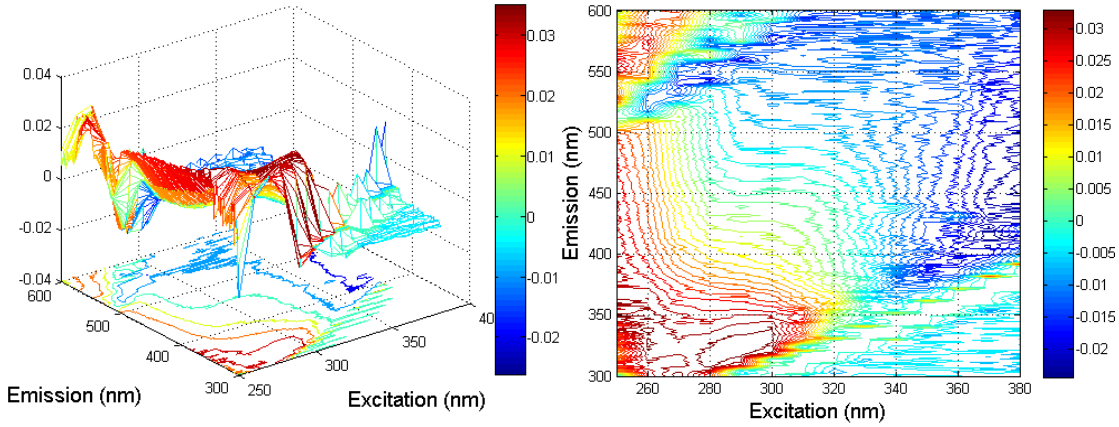
PC3 loading plots demonstrated a distinct peak at the same region where the fluorescence peak related to protein-like materials (Ex/Em = 280 nm/330 nm) occurs (Chen *et al*., 2003a; Peiris *et al*., 2010a) (Figure 5.1c), indicating that PC3 is mostly related to protein-like materials. For the PCAs of FEEMs collected at PMT voltages of 550 V, 600 V, 650 V, and 700 V, the percent variances captured by PC3 were 9.13%, 9.53%, 10.77%, and 9.49%, respectively.



(a)

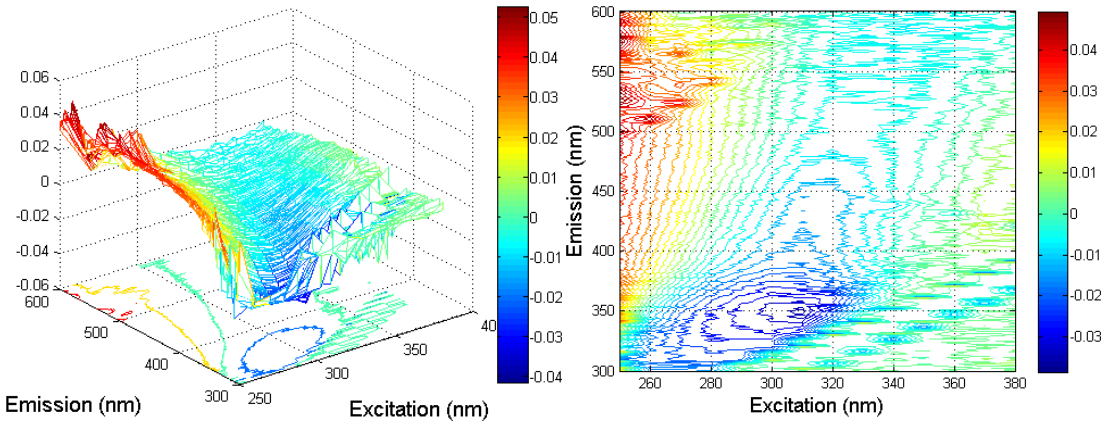


(b)



(c)

Figure 5.1



(d)

Figure 5.1 Loading plots for (a) PC1, (b) PC2, (c) PC3, and (d) PC4 extracted from the PCA for the FEEMs collected at PMT voltage of 700 V

In the PC4 loading plot (Figure 5.1d), a depression or valley was present at a similar location to where the protein-like materials related peak is situated, suggesting that PC4 might also be associated with protein-like materials. However, the percent variances explained by PC4 were relatively low compared to the other three PCs. Percent variances captured by PC4 were 4.94%, 4.98%, 4.86% and 4.64% for the 4 different PCAs from the datasets collected at PMT voltages of 550 V, 600 V, 650 V, and 700 V, respectively.

5.3.2.2 Impacts of PMT Setting on the Variance Captured by PCs

For PC1 which is related to humic-like substances, the increase of HA- and FA-like responses with increasing PMT voltages (Figure C1a) resulted in a slight increase of percent variance captured by PC1. This is likely to be attributed to the reduction of signal noise as HA- and FA-like responses increased with increasing PMT voltage. Similarly, percent variance explained by PC3 (i.e. protein-like materials) also slightly increased with increasing PMT voltages up to PMT setting of 650 V, and similar reasons for this increase can be put forth. The relatively low percent variance captured by PC3 at PMT voltage of 700 V might be attributed to the slight difference between the variance contained within the dataset collected at PMT voltage of 700 V and other PMT voltage conditions as illustrated by the R^2 values in Table 5.1. In this study, the signal response of protein-like materials at the lowest PMT setting was sufficiently high for the extraction of a significant PC3 instead of being identified as background noise, indicating that the measurement methods using the selected instrument settings

were sufficiently sensitive. This explained why the percent variance increase due to PMT setting change was fairly small and therefore relatively unimportant in this study. This also demonstrated that FEEM-PCA could tolerate the measurement noise changes due to the change of PMT voltage conditions.

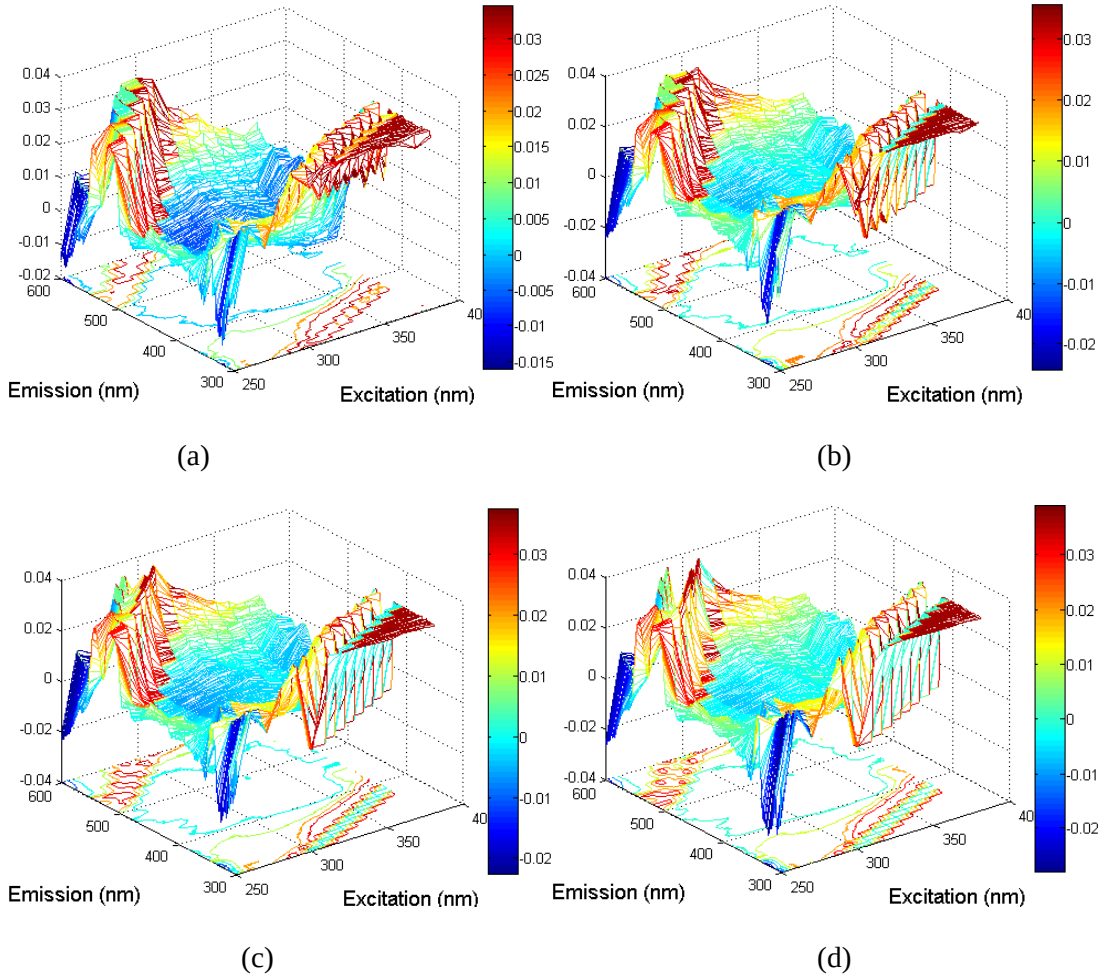


Figure 5.2 Loading plots for PC2 extracted from FEEMs collected at PMT voltage of (a) 550 V, (b) 600 V, (c) 650 V and (d) 700 V

The percent variances explained by PC2 (Rayleigh scattering related) had a slight decreasing trend with increasing PMT voltage. This can be attributed to the increasing level of saturation of first-order Rayleigh scattering signals (Rayleigh scattering on the right side as illustrated in Figure 5.2) with increasing PMT voltage. This is illustrated in Figure 5.2, where the depth of the valley on the right side, which reflects the extent of the saturation of the first-order Rayleigh scattering, increased with

increasing PMT setting. However, the decrease in percent variance captured by PC2 with increasing PMT was relatively small, suggesting that saturated Rayleigh scattering regions can still reflect the larger portion of information contained in this region.

As mentioned above, the percent variances captured by PC4 of the PCAs at the 4 different PMT voltage conditions were very similar. Even though the loading plots indicated PC4 was possibly related to protein-like materials, the interpretation should be cautious due to the low percent variance that can be explained by this component.

5.3.2.3 Impacts of PMT Setting on the PC Scores

The impacts of PMT voltage conditions on the PC scores were extremely dissimilar to intensity values. Instead of increasing with PMT voltage, scores of the four PCs were shown to be essentially the same when comparing the PCs among the 4 PCAs obtained using data with different PMT voltage conditions (Figure C1b). For the same PC, PC scores were highly correlated between each pair of PMT voltages (Table 5.1). For each PC, correlation of scores for the dataset collected at PMT voltage of 700 V and other PMT voltage condition were slightly lower, which corroborated with the observations for intensity values. The above findings implied that PMT voltage settings had no effect on FEEM-PCA results when the instrument setting is sufficiently sensitive for the detection of fluorophores. It should be also noted that the level of correlation between each pair of PMT voltage settings for PC4 was relatively lower compared to those of the other PCs. This can be attributed to that PC4 is more sensitive to noise due to low percent variance captured by this PC, confirming that the information provided by PC4 is less reliable and should be interpreted with caution.

In addition, the high R^2 values for PC2 scores between each pair of PMT voltages (Table 5.1) confirmed that the information loss due to saturation of first-order Rayleigh scattering is negligible. It is challenging to optimize the instrument settings for both fluorophores and Rayleigh scattering regions simultaneously since the priority of instrumentation optimization is usually on ensuring that the fluorophores are well characterized. This is especially important for the low fluorophore concentration conditions which are encountered in drinking water related NOM studies. A high PMT voltage condition will be required to capture the fluorophores but this causes the saturation of the Rayleigh scattering region to increase. Therefore, the information contained within the Rayleigh scattering has to be discarded when peak picking is used. Alternatively, the Rayleigh scattering can be

measured at its optimized instrument setting in addition to the measurement of the fluorophores at their optimized settings. However, this will make the FEEM monitoring tedious and time-consuming. This study demonstrated though that FEEM-PCA can successfully accommodate the Rayleigh scattering saturation issue without the requirement of optimizing the instrument setting separately for fluorophores and the scattering regions i.e. one set of measurements is sufficient to derive meaningful results.

5.3.3 Comparison of FEEM Data Analysis Methods

5.3.3.1 Comparison between FEEM-PCA and Peak Picking

The relationships between PC scores and intensity values at each PMT voltage conditions are summarized in Table 5.2. PC1 scores were shown to be highly correlated with FA-like intensities at each PMT voltage condition. Similarly, HA-like intensities were also well correlated with PC1 scores but the R^2 value is smaller compared to the one for FA-like intensities and PC1 scores. The higher correlation for FA-like intensities can be attributed to the dominance of FA-like peak in the humic-like substance region. In both cases, the R^2 values were very consistent for 4 different PMT voltage settings. The high correlation between corresponding PC scores and intensity values of humic-like substances was expected because the loading plots of PC1 illustrated that PC1 is capable of extracting the information contained within the humic-like region in the FEEM spectra. This indicates that FEEM-PCA and peak picking are expected to have similar performance regarding the quantification of humic-like substances including both HA- and FA-like substances. However, due to the capability of characterizing HA- and FA-like responses simultaneously, peak picking is superior to FEEM-PCA when composition characterization of humic-like substances is the main focus of the study.

Table 5.2 Relationships between peak intensities and PC scores at different PMT settings. The relationship is evaluated using R^2 value.

	Protein Intensity			*HA Intensity*	*FA Intensity*
	PC2 score	PC3 score	PC4 score	PC1 score	PC1 score
PMT550	0.21	0.38	0.13	0.87	0.96
PMT600	0.22	0.34	0.17	0.86	0.96
PMT650	0.17	0.40	0.17	0.87	0.96
PMT700	0.15	0.39	0.12	0.88	0.95

From the loading plots (Figure 5.1) PC2 was determined to be related to Rayleigh scattering which can be employed to characterize the colloidal/particulate matter (Peiris *et al.*, 2010a). In Chapter 4, protein-like materials were shown to be also a direct contributor to the Rayleigh scattering related principal component which is dominated though by contributions from inorganic colloid. Nevertheless, the relationship between PC2 scores and protein-like intensity was evaluated. In this study, the correlation between PC2 scores and protein-like intensities was weak for all 4 different PMT voltage conditions, suggesting protein-like materials in the selected water source had limited impact on the Rayleigh scattering. Also, PC4, which was determined to be related to the protein-like peak region and captured only a low percent variance, was shown to have low correlation with protein-like intensities. This can be attributed to that PC4 is more sensitive to noise because of the low percent variance explained by this PC. As discussed above, FEEM-PCA had better performance compared to peak picking regarding the characterization of colloidal/particulate matter. FEEM-PCA can successfully extract the information contained within the Rayleigh scattering region without being affected by the saturation of scattering peaks when high PMT voltage was required for the reproducible measurement of fluorophores at low concentrations.

Compared to PC2 and PC4, PC3 scores and protein-like intensity values were found to have somewhat higher R^2 values but these were still relatively low (Table 5.2). This is not surprising since PC3 only explained approximately 10% of total variance in each case and the presence of other trends in addition to protein-like peak region were clearly shown in the loading plots. In other words, if PC3 is solely related to the protein-like peak, a reasonably well correlation between PC3 scores and protein-like intensities would be expected. The presence of other trends might contain the information regarding other physical and/or statistical interactions between protein-like peak and other regions of FEEM spectra. Therefore, PC3 extracted a synthesis of information including protein-like materials and their interactions with other fluorophores and colloidal/particulate matter. For the characterization of protein-like materials, FEEM-PCA will have superior performance than peak picking if the interest of study focused on the monitoring of the effects induced by protein-like materials since the information regarding the interactions between protein-like materials and other components would be also provided by PC scores. However, if the study focused on the kinetics and characteristics of protein-like materials, peak picking is considered to be more appropriate because intensity values will only include the information related to protein-like materials.

5.3.3.2 Comparison between FEEM-PCA and FEEM-PARAFAC

Previous studies have reported that PARAFAC scores were highly correlated with the intensity values derived from peak picking evaluations (Bieroza *et al.*, 2011; Shutova *et al.*, 2014). In this study, a high correlation between PC scores and intensity values were only found for PC1 which is characterized as humic-like substances and explained the largest variance of the FEEMs. The correlation of intensity values with scores of protein-like material related PC3, which explained the much lower variance of the FEEMs compared to PC1, is much weaker. A key difference between this study and PARAFAC studies is that FEEM-PCA analyses include Rayleigh scattering regions whereas these have to be removed for PARAFAC analyses. This means that the FEEM datasets for FEEM-PCA and PARAFAC are not mathematically identical. Therefore, it is not surprising that PCs extracted from PCA contained different information (i.e. variance) compared to PRAFAC components. This is probably one factor contributing to the difference in correlation between peak picking and PCA vs. peak picking and PARAFAC. In addition, there are also other fundamental differences in approach between PCA and PARAFAC which also contributes to these differences.

Peleato and Andrews (2015b) have conducted a direct comparison between PCA and PARAFAC by removing the Rayleigh scattering regions from their data sets. The physical meaning of PARAFAC components identified corresponded to the pure spectra of humic-like substances (i.e. HA- and FA-like) and protein-like material (i.e. tryptophan peak). Comparatively, the extracted PCs were characterized by humic-like substances and protein-like materials as well as their combination (Peleato and Andrews, 2015b). The humic-like substance-related components extracted from PARAFAC (C1) and PCA (PC1) had similar loading plots (physical meaning) and therefore had relatively similar score trends (Peleato and Andrews, 2015b) for 4 different raw water datasets. However, since the loading plots (physical meaning) for protein-like material related components extracted from PARAFAC (C2) and PCA (PC3) were quite different, their score trends did not corroborate well with each other and C2 and PC3 had very different removal behavior in the drinking water treatment train investigated (Peleato and Andrews, 2015b). The authors attributed the different protein-like trends between PC3 and C2 to that PCA and PARAFAC extracted distinct NOM constitutes for the protein-like region (Peleato and Andrews, 2015b). Therefore, a direct comparison between PARAFAC and PCA scores seems only be valid when the components extracted from the two methods have similar physical meaning.

The extracted PCs from PCA essentially reflected FEEM spectra patterns with rotational freedom whereas PARAFAC components reflected the pure component spectra. The PCs are hence indicative of major trends of a certain fluorophore possibly with contributions of minor trends of other fluorophore (and Rayleigh scattering if included) as illustrated in Peleato and Andrews (2015b) and this study. When a PC explains the majority of the variance inherent in FEEMs and is dominated by the major trends of a fluorophore, then the corresponding PC scores are expected to be well correlated with the intensity values of said fluorophore (e.g. PC1 in this study). However, when a PC explains a lower amount of variance (i.e. more sensitive to noise) and interactions were also important (e.g. PC3 in this study, major trends of protein-like fluorophore are not dominating the minor trends), the corresponding PC scores are not expected to be well correlated with the intensity values of the fluorophore contributing to the PC. Peleato and Andrews (2015b) also highlighted that PARAFAC component scores were shown to be more variable compared to PC scores from PCA although the physical meaning of PARAFAC components is more specific, and they attributed this observation to the increased noise or sensitivity inherent to PARAFAC. These observations confirmed that components extracted from PCA and PARAFAC are mathematically/intrinsically different for the same fluorophore peak region. In practice, PCs are orthogonal to each other and can include additional information regarding the physical and/or statistical interactions with other fluorophores as well as Rayleigh scattering. They will have stronger power for the prediction and monitoring of the overall effects posed by the fluorophores compared to PARAFAC as illustrated in Peleato and Andrews (2015a). However, if a study focuses on the investigation of the characteristics of fluorophores, PARAFAC would be superior.

In addition, compared to two-way PCA as employed in this study, the PARAFAC model can generate components which can be referred to the pure underlying spectra of real fluorophores based on loadings of emission and excitation (Bieroza *et al.*, 2011). Therefore, the interpretation of the components extracted by three-way PARAFAC is easier compared to the two-way PCA (Bieroza *et al.*, 2011, Peleato and Andrews, 2015b). However, deriving a valid and robust PARAFAC model is a non-trivial task because the model robustness is dramatically affected by the data trilinearity (i.e. more efforts on data pre-processing) and the selected number of components (Bieroza *et al.*, 2011). Especially, the determination of the number of valid PARAFAC components is very time-consuming due to a lack of an effective diagnostic tool aiding in the component selection (Bieroza *et al.*, 2011). Compared to the supervised PARAFAC method, PCA approach is unsupervised and can extract the

valid components in a faster manner. This could be a distinct advantage for the analysis of FEEMs which are from similar sources and have a relatively uniform pattern of fluorophores (Bieroza *et al.*, 2011).

5.3.3.3 Recommendations for the Selection of FEEM Data Analysis Methods for Different Applications

Both peak picking and PARAFAC can extract the information of pure spectra of fluorophores. Therefore, both approaches would be appropriate for investigations on fluorophore characteristics and kinetics. Since peak picking and PARAFAC have little difference regarding the fluorophore identification and their quantification as discussed above, peak picking should be considered over the application of the more sophisticated PARAFAC due to its simplicity and efficient manner (Bieroza *et al.*, 2011). This is especially valid for the online monitoring of water samples from the same or similar source. Since common PARAFAC components could be extracted from different water sources (Murphy *et al*., 2011; Shutova *et al.*, 2014; Peleato and Andrews, 2015b), PARAFAC would be a promising tool to help identify the excitation emission coordinate(s) for the fluorophore monitoring using peak picking. This is especially helpful and important when a fluorophore has a low concentration in the water matrix. In addition, both approaches were illustrated to not be suitable for the characterization of Rayleigh scattering.

Due to the orthogonal nature of PCs extracted from FEEM-PCA and their capability of reflecting physical and/or statistical interactions between the dominating fluorophore (major trend) and other components (fluorophores and/or Rayleigh scattering) in addition to pure spectra information of the dominating fluorophore, FEEM-PCA is superior to PARAFAC and peak picking for the investigation on the overall effects caused by the fluorophore/Rayleigh scattering. The capability of characterizing the Rayleigh scattering which is related to colloidal/particulate matter provided additional value to PCA compared to PARAFAC and peak picking. This is especially beneficial since the instrumentation optimization was not required to be conducted separately for the fluorophores and the Rayleigh scattering. FEEM-PCA may also be able to assist with identification of peak locations of fluorophores although the interpretation should be more cautious because the physical meaning of the extracted PCs reflects more than the pure spectra of the fluorophore.

5.4 Conclusions

This study investigated the impacts of PMT voltage conditions on the performance of two FEEM data analysis methods, including peak picking and FEEM-PCA. A utility comparison of these two methods with PARAFAC was also conducted based on the results from this study and reports in the literature. The major conclusions drawn from this study are summarized as follows:

- Increase of PMT voltage resulted in an increase of intensities at pre-selected peak locations. The high correlation of intensities between each pair of PMT voltage settings indicated that the performance of peak picking was not dramatically affected by the PMT settings if fluorophores could be detected. However, high PMT voltage will cause the saturation of Rayleigh scattering regions, making interpretation of the information contained in this region through peak picking impossible.
- FEEM-PCA results were not affected by the PMT voltage settings and PC scores were shown to be essentially the same at different PMT voltages. Information contained within the Rayleigh scattering regions was successfully extracted by PC2 even when the Rayleigh scattering regions were saturated.
- FEEM-PARAFAC and peak picking were shown to be promising approaches for the characterization of fluorophores whereas both techniques are not suitable for investigations on the information contained within the Rayleigh scattering region and on interactions among fluorophores and/or with Rayleigh scattering.
- Compared with PARAFAC and peak picking, FEEM-PCA was superior in its ability to incorporate Rayleigh scattering region in the analyses and was better able to monitor overall effects induced by the fluorophores due to the orthogonal properties of PCs and the additional information contained within each PC.

Chapter 6
Characterization of the Biopolymer NOM Fraction from Different Sources using a LC-OCD-based Size Fractionation Approach Combined with FEEM Analysis

Summary

Biopolymers from different NOM sources, including a biofilm extracellular polymeric substances (EPS) extract from pilot-scale biofilter support media, secondary effluent from a wastewater treatment plant and natural river water, were size-fractionated using a liquid chromatography–organic carbon detection (LC-OCD)-based approach and the collected isolated sized fractions were further characterized using fluorescence excitation emission matrices (FEEM) and LC-OCD. This approach was taken to better understand the composition of the biopolymers and to confirm the attribution of some of the FEEM peaks. The organic nitrogen (ON) concentration of the biopolymers (LC-OCD based) was found not to be a good indicator of the FEEM protein intensities whereas the normalized ON concentration (i.e. biopolymer-organic nitrogen per carbon mass ratio, ON:OC) was found to be a good indicator of different biopolymer sources. Biopolymer sized fractions with lower molecular weights were shown to have higher normalized FEEM protein intensities for different biopolymer sources and therefore are likely to contain more protein-like materials. This biopolymer fractionation approach has the potential for use in characterizing other fluorescent NOM fractions.

6.1 Introduction

Aquatic natural organic matter (NOM), which is generated from the contact of living and dead organic materials with water through various hydrological processes, is ubiquitous in surface and ground water systems (Egeberg *et al.*, 2002; Maurice *et al.*, 2002). NOM composition is complex and influenced by various factors, such as origin, physiochemical properties of the water, properties of natural sorbents in the water, and the presence of other reactive processes in the local environment (Leenheer and Croué, 2003).

NOM is a heterogeneous mixture of organic compounds with a wide range of molecular weights (MW), various functional groups, varying charge, hydrophobicity, and status (dissolved vs. particulate/colloidal) (Leenheer and Croué, 2003; Shutova *et al.*, 2014) all of which is making NOM characterization very challenging. Traditional NOM characterisation techniques have a variety of

limitations, they either measure bulk properties (e.g. dissolved organic carbon (DOC), and specific UV absorbance (SUVA)) or they alter the physiochemical properties of the NOM (e.g. XAD resin and serial membrane filtration) (Maurice *et al.*, 2002; Song *et al.*, 2009; Matilainen *et al.*, 2011; Shutova *et al.*, 2014; Chen *et al.*, 2015). Recently, liquid chromatography–organic carbon detection (LC-OCD) and fluorescence excitation emission matrices (FEEM) have gained increasing acceptance for NOM characterization due to their ability to provide NOM compositional information with minimum modification of the NOM during the analysis (Huber *et al*., 2011; Ishii and Boyer, 2012; Shutova *et al.*, 2014). LC-OCD provides information on the organic carbon (OC), organic nitrogen (ON), and UV properties of five different NOM size groups with the following attribution: biopolymers, humics, building blocks, low MW acids and neutrals (Huber *et al*., 2011). FEEM provides compositional information based on the presence of fluorescent functional groups in the NOM structure and is capable of characterizing intrinsically fluorescent NOM groups that include the humic- and protein-like materials (Peiris *et al.*, 2010a; Baghoth *et al.*, 2011; Shutova *et al.*, 2014). Since these two techniques are based on different physical principles the combination can provide complementary information that is useful in describing compositional and physicochemical properties of the various NOM components.

Being the important NOM components with larger MW, biopolymers as measured by LC-OCD and a fluorescent component (i.e. protein-like substances) as measured by FEEM have been identified as important contributors to low pressure membrane fouling (Hallé *et al.*, 2009; Chen *et al.*, 2014). However, the composition of the biopolymers is rarely studied and poorly understood. The concept of the biopolymers was proposed by Huber *et al*. (2011) and is defined on the basis of size exclusion chromatography. The biopolymers are attributed to consist of a mixture of polysaccharides and nitrogen-containing compounds (e.g. proteins and amino sugars) with a molecular weight (MW) of 10 kDa or higher (Huber *et al.*, 2011). The organic carbon (OC) and organic nitrogen (ON) detectors associated with LC-OCD can provide overall information on the composition of biopolymers. In addition to these overall information, there is an interest to extract more detailed information on biopolymer composition and compare the compositional properties of biopolymers from different sources such as natural surface water, biofilm extracellular polymeric substances (EPS), and wastewater. To better understand their composition, the biopolymers were isolated into 3 sized fractions (hereafter named as biopolymer sized fractions) using LC-OCD. These biopolymer sized fractions were then subjected to LC-OCD and FEEM analysis. A similar concept has been employed

previously to investigate the fluorescence properties of humic acid and its sized fractions (Halim *et al.*, 2013) albeit using a re-dissolved and concentrated humic acid extract (50 mg/L). Three humic acid sized fractions were isolated using preparative high performance size exclusion chromatography, freeze-dried and then characterized by FEEM (Halim *et al.*, 2013). This size fractionation approach required improvement in order to characterize biopolymers 'in situ' at the low concentrations encountered in natural or drinking water treatment systems. The fractionation approach presented herein does not concentrate the original sample nor does it concentrate/freeze dry the isolated fractions, mainly because currently available concentration methods (e.g. freeze-drying and rotary vacuum evaporation) likely cause alteration of nature and/or content (Chen *et al.*, 2015) especially of the protein-like constituents. Spencer *et al.* (2007) have also illustrated that protein-like fluorophores had a lower stability during freeze/thaw cycles compared to humic-related fluorophores.

The main objective of this study was to provide an improved fundamental understanding of the composition of the biopolymers by isolating biopolymers into three sized fractions using LC-OCD. The properties of these biopolymer sized fractions were then examined individually using further LC-OCD analysis with OC, ON and UV detectors and also FEEM. Biopolymers were analyzed in three very different sources, including a river water, a municipal wastewater (secondary effluent), and biofilm EPS from support media of a pilot-scale drinking water biofilter to ascertain differences and similarities among biopolymers.

6.2 Methods and Materials

6.2.1 Biopolymer Sources

6.2.1.1 Biofilm EPS Extract from Biofilter Support Media

The pilot-scale biofilter was located at a full-scale drinking water treatment plant (Southwestern Ontario, Canada) and consisted of a dual-media (sand and anthracite) filter with Grand River water (GRW) as the feed water. The biofilter had been acclimated for more than six months prior to sampling. The biofilm EPS was extracted according to El-Hadidy et al. (in preparation). Two media sampling events were conducted on 17-April-2014 and 08-May-2014, respectively. For each sampling event, the anthracite and sand media were collected from 10 cm and 30 cm depth, respectively. Immediately after sampling anthracite and sand media samples were separated by sieving due to the mixing of two media types at the sampling depth, then rinsed using biofilter

effluent to remove loosely attached biofilm and stored over night at 4 ℃ until extraction. The media was rinsed with the extraction buffer (phosphate buffer saline (PBS), 6 mM, pH = 7) and extracellular polymeric substances (EPS) were extracted using 3 g wet media, 3 g rinsed CER (Dowex Marathon C cation exchange resin, Sigma Aldrich, USA), and 20 mL PBS mixed in sterilized 45 mL polypropylene centrifuge tubes which were shaken horizontally in an incubated shaker (350 rpm and 15^{o}C) for 1 hr. Then, the samples were centrifuged at 5000 ×g for 20 minutes and the supernatant was filtered through a 0.2 μm polyethersulfone syringe filter to isolate the extracted EPS from any remaining biofilm bacteria. The EPS extract was stored at 4 ^{o}C until fractionation.

6.2.1.2 Municipal Wastewater (Secondary Effluent)

The secondary effluent from a local wastewater treatment plant (Southwestern Ontario, Canada) was used as a biopolymer source representative of a municipal wastewater system. The wastewater treatment plant is a Class III facility which includes grit removal, primary clarifiers, aeration tank, secondary clarifiers, phosphorous removal, UV disinfection and sludge treatment. Two secondary effluent samples (after the UV disinfection) were taken on 14-April-2014 and 28-July-2014, respectively, and filtered through a 0.45 μm polyethersulfone filter (Pall Corporation, Supor®, USA) prior to LC-OCD analysis, FEEM analysis and fractionation.

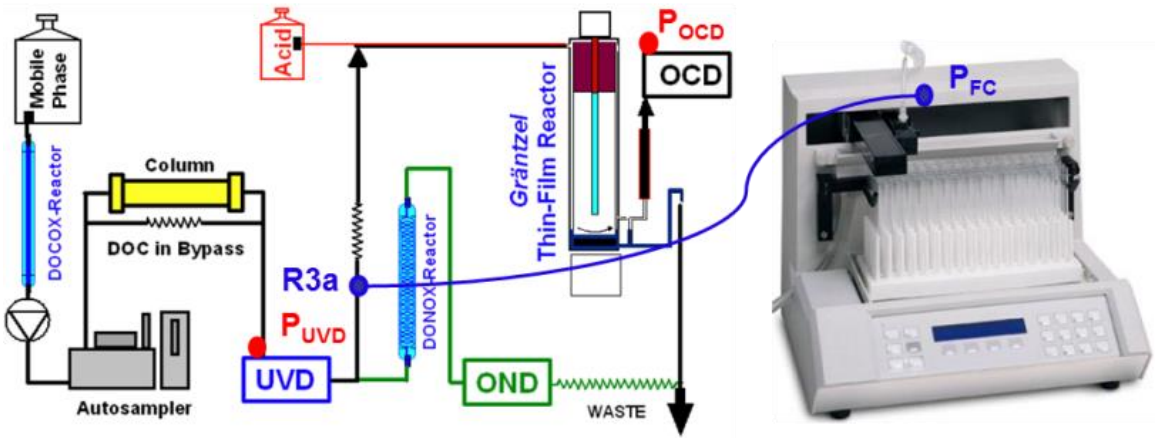
6.2.1.3 Grand River Water (GRW)

Grand River water (GRW, Southwestern Ontario, Canada), which is an intensively affected area by various municipal and agricultural activities, was selected as a biopolymer source representative of a natural surface water system. Two GRW samples were collected from the full-scale treatment plant intake which feed the pilot-scale biofilter on 02-June-2014 and 09-July-2014, respectively, and filtered through a 0.45 μm polyethersulfone filter (Pall Corporation, Supor®, USA) prior to LC-OCD analysis, FEEM analysis and fractionation.

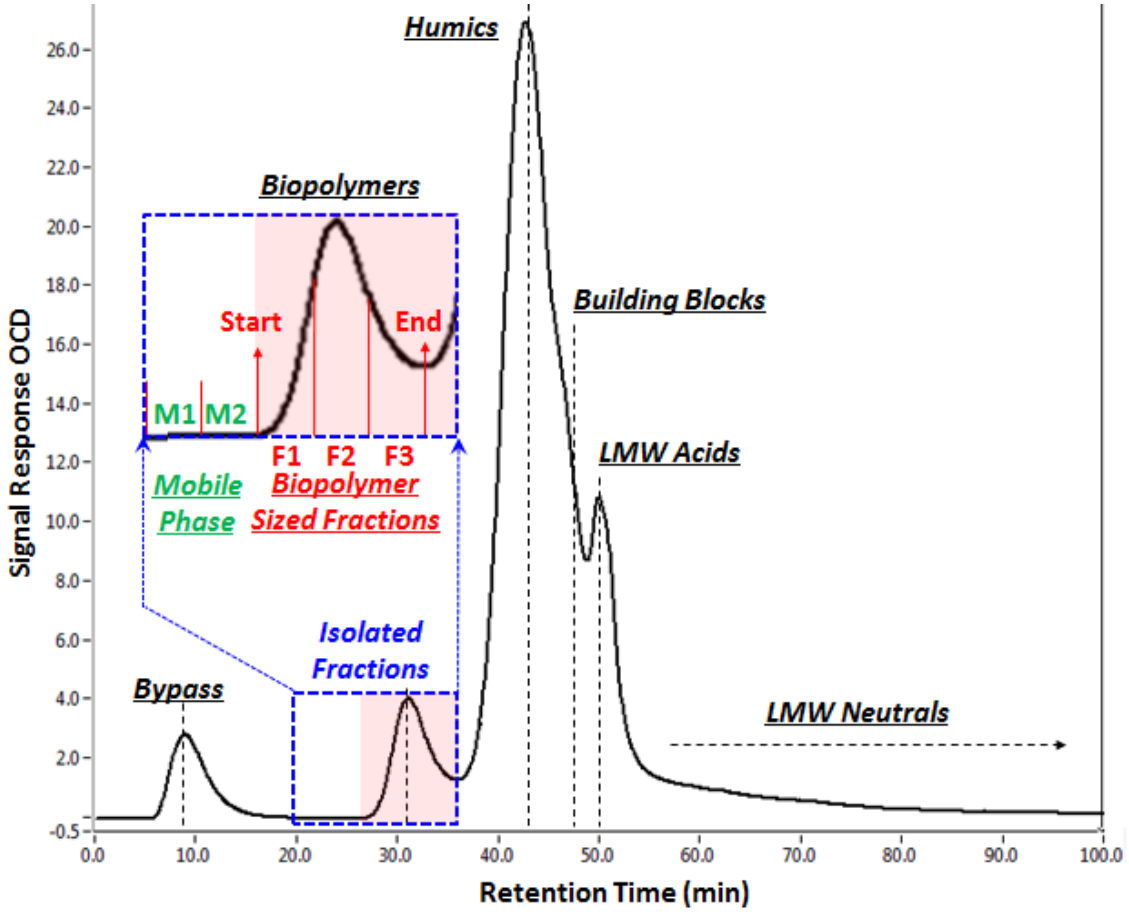
6.2.2 LC-OCD-based Fractionation Approach

6.2.2.1 Instrumentation and Overall Fractionation Procedure

A LC-OCD system equipped with a fraction collector (Gilson FC 203B, Gilson Inc., USA) (Figure 6.1) was used to perform the isolation and concurrent size fractionation of the biopolymers from different sources.



(a)



(b)

Figure 6.1 (a) Schematic diagram of the LC-OCD based size fractionation system and (b) representative chromatogram of Grand River water

OCD: organic carbon detector; OND: organic nitrogen detector; UVD: UV detector

LMW: low molecular weight

Table 6.1 Biopolymer characteristics of bulk samples from different biopolymer sources

	DOC (μg/L)	***Biopolymer -OC (μgC/L)***	***Biopolymer -ON (μgN/L)***	***Normalized ON Conc.***	***FEEM -Protein***	***Normalized FEEM protein***	***Biopolymer -Start(min)***	***Biopolymer -Peak (min)***	***Biopolymer -End (min)***
AnthraBE1	1916	456	125	0.27	54.01	0.118	27.16	31.69	35.25
AnthraBE2	3311	649	210	0.32	129.41	0.199	27.33	31.75	34.46
SandBE1	2120	737	204	0.28	82.85	0.112	27.08	31.42	36.19
SandBE2	3697	1075	321	0.30	175.06	0.163	27.33	31.37	35.75
GRW1	5610	494	52	0.11	48.68	0.099	26.83	31.57	36.47
GRW2	5824	403	48	0.12	52.12	0.129	26.92	31.11	35.88
2ndEffluent1	7081	1032	180	0.17	320.29	0.310	26.74	30.98	37.83
2ndEffluent2	9163	1186	191	0.16	330.99	0.279	26.84	30.71	36.58

AnthraBE: Anthracite Biofilm EPS (1:17-APR-2014; 2: 08-MAY-2014);

SandBE: Sand Biofilm EPS (1:17-APR-2014; 2: 08-MAY-2014);

GRW: Grand River Water (1:02-JUN-2014; 2: 09-JUL-2014);

2ndEffluent: Secondary Effluent (1:14-APR-2014; 2: 28-JUL-2014)

The overall procedure consisted of the following: analysis of 3 mL of sample by the LC-OCD to determine the biopolymer peak start and end times (Table 6.1); calculation of the collection time windows for the biopolymer sized fractions (Section 6.2.2.2); followed by LC based fractionation, where 3 mL of sample were injected and biopolymer size fractions were collected at the pre-set collection time windows; repeat of fractionation to collect sufficient volume for each biopolymer sized fraction; and lastly characterization of the collected biopolymer sized fraction by LC-OCD and FEEM (Section 6.2.3).

The LC-OCD was equipped with a weak cation exchange chromatography column (TSK HW50S, 3000 theoretical plates, Tosoh, Japan) (Huber *et al.*, 2011). A 0.45 μm polysulfone in-line filter (SARTORIUS®, Germany) just prior to the column provided added protection from particles. Phosphate buffer (i.e. 2.5 g/L KH_2PO_4 and 1.5 g/L $Na_2HPO_4 \times 2H_2O$, pH=6.58) at a flow rate of 1.1 mL/min was used as eluent. Details regarding the fundamentals and operation of the LC-OCD system can be found in Huber *et al.* (2011). Based on a reproducibility test with seven injections (injection volume = 1 mL) of Grand River Water sample (DOC=5.3 mgC/L and filtered with seven different polymeric filters), biopolymer-OC and -ON concentration were 356.57±10.31 μg/L (Rep%=standard deviation/average= 2.89%) and 40.13±1.16 μg/L (Rep%=2.90%) for the LC-OCD system, respectively. Based on a method detection limit (MDL) test with seven injections of Grand River water sample (DOC=5.3 mgC/L), the MDL of biopolymer-OC (dilution factor(DF)=25, Rep%=7.14%) and -ON (DF=5, Rep%=7.68%) were determined to be 3 μg/L and 2 μg/L, respectively.

6.2.2.2 Determination of Collection Time Window for Biopolymer Sized Fractions

As illustrated in Figure 6.1, the OC-based elution time (at P_{OCD}) had to be corrected to the UV-based elution time (at P_{UVD}) which had to be further corrected to the fraction collection port of LC-OCD system (i.e. R3a) and then the fraction collection port of the fraction collector (i.e. P_{FC}). The biopolymer elution times were corrected as follows to obtain the corresponding collection time windows for the fraction collector:

$$T_{FC} = T_{OCD} - T_{OCD\text{-}UVD} + T_{UVD\text{-}R3a} + T_{R3a\text{-}FC} \quad (1)$$

T_{FC}: collection time for fraction collector

T_{OCD}: OC-based elution time

$T_{OCD-UVD}$: time difference between P_{OCD} and P_{UVD} which was determined as the elution time difference of the bypass peak between the OC and UV signal.
$T_{UVD-R3a}$: time difference between P_{UVD} and R3a which was provided by the manufacturer.
T_{R3a-FC}: time difference between R3a and P_{FC} which was determined based on the tubing size and eluent flow rate.

Each of the biopolymer sized fractions was 1/3 of the overall biopolymer peak width (Figure 6.1) and collection time for the first biopolymer sized fraction (F1) began with the corrected biopolymer peak start time (see Equation 1). Overall, five fractions of equal size (time per tube) were collected for each sample including M1 and M2 (i.e. background mobile phase after bypass peak and prior to start of the biopolymer peak, respectively) as well as F1 to F3 (i.e. biopolymer sized fractions) (Figure 6.1). Each sample was fractionated two times using 3 mL injection volume to collect sufficient sample volume for each fraction for further analyses.

6.2.3 Characterization of Bulk Samples and Collected Fractions

6.2.3.1 LC-OCD Based Characterization

Prior to fractionation, each sample was analyzed by LC-OCD to quantify the biopolymer content of the unfractionated sample (OC and ON based) and to determine the fraction collection time windows for the 3 biopolymer sized fractions. Aliquots (2 mL) of each of the five collected fractions were also analyzed by the LC-OCD.

6.2.3.2 FEEM Analysis

Similar to the LC-OCD analysis, all unfractionated samples and aliquots of all fractions including mobile phase fractions and biopolymer sized fractions for each sample were characterized using FEEM. Milli Q water was also measured for each unfractionated sample and corresponding fractions. A Cary Eclipse fluorescence spectrophotometer (Agilent Technologies, Canada) was used. All samples and fractions were measured using a quartz cuvette (10 mm width) with 4 optical windows. The slit widths for both excitation and emission were set at 10 nm. Excitation wavelength ranged from 250 to 380 nm with a sequential increment of 10 nm and emission wavelength ranged from 300 to 600 nm with a data interval of 1 nm. Photomultiplier tube (PMT) voltage was set to 650 V for all unfractionated samples and to 950 V for all fractions due to their low concentrations. The above settings were adapted from an optimization study conducted by Peiris *et al.* (2009). An intensity at an

Ex/Em pair of 280 nm/330 nm (Peiris *et al*., 2010a) was used as the peak intensity for protein-like material for both bulk and sized fraction samples. Based on a reproducibility test with seven measurements (PMT=650V, Excitation and Emission slit width = 10 nm) of Grand River water sample (DOC=5.3 mgC/L and filtered with seven different 0.45 µm polyethersulfone filters), the difference (Rep%) of protein-like peak intensity values were 1.43% (std=0.72) for the spectrofluorometer.

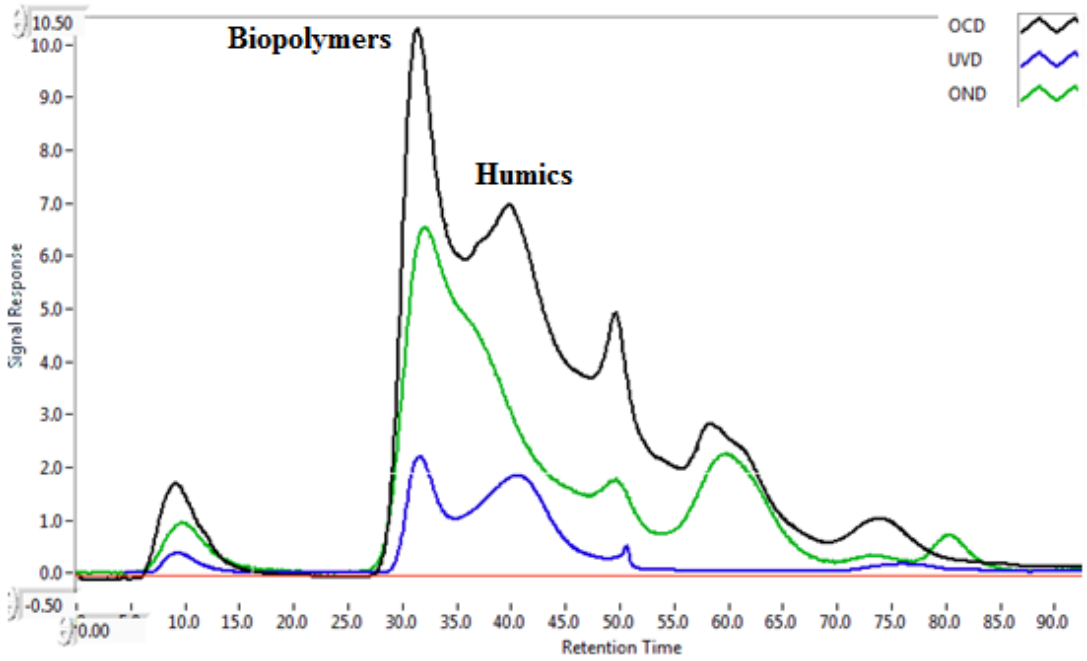
6.3 Results and Discussions

6.3.1 General NOM and Biopolymer Characteristics for Different Sources

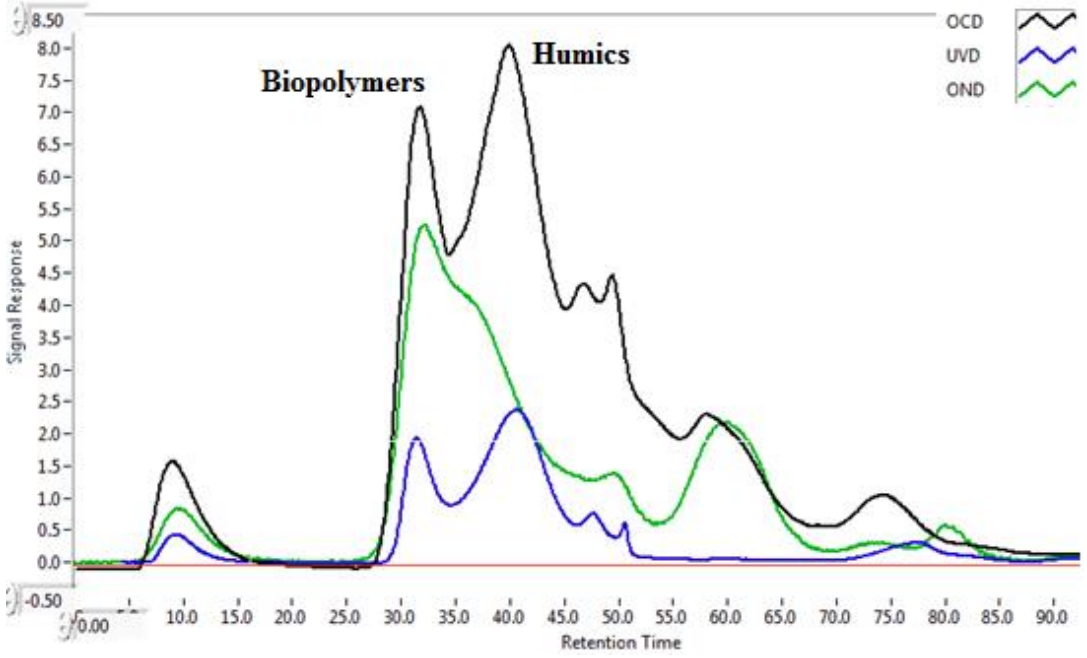
6.3.1.1 LC-OCD Based Characterization

Secondary effluent (pH=7.5 and 7.4) and biofilm EPS extracts (pH=7.0) had the highest and lowest DOC concentration among three biopolymer sources, respectively (Table 6.1). As shown in Figure 6.2, humics had much larger OC peak than biopolymers for secondary effluent and GRW (pH=8.1 and 8.2), suggesting that their high DOC is mainly attributed to the presence of high concentration of humics. On the contrary, biopolymers of biofilm EPS extracts had similar or larger OC peak than humic-like substances. In addition, biofilm EPS extracts had more complicated composition of LMW compounds compared to the other two biopolymer sources.

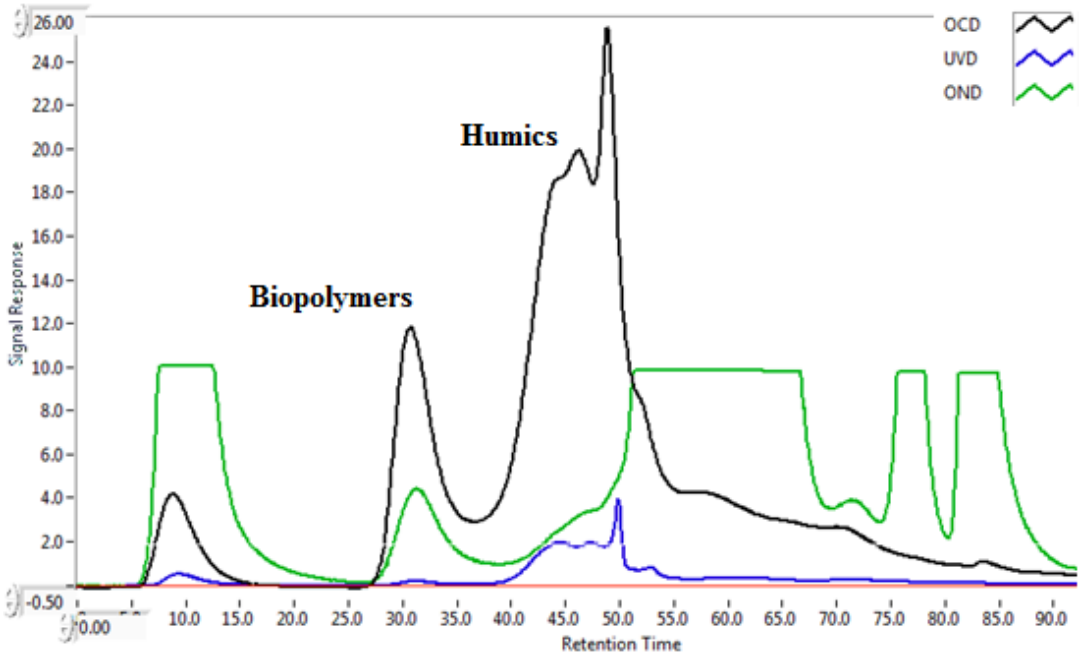
Biopolymers from different sources contained both OC and ON compounds, which is consistent with previous NOM characterization studies using similar samples (Frølund *et al.*, 1996; Her *et al.*, 2003). Grand River water (GRW) had the lowest organic carbon (OC) and organic nitrogen (ON) concentration for the biopolymers amongst the different biopolymer sources even though the GRW had higher DOC concentration than the media biofilm EPS extracts (Table 6.1). Comparatively, secondary effluent had the highest DOC concentration and biopolymer-OC concentration amongst three different biopolymer sources; however, secondary effluent samples had similar biopolymer-ON concentrations compared to media biofilm EPS extracts. The biopolymer ON:OC ratio (hereafter defined as normalized ON concentration) was also different for the biopolymers from different sources (Table 6.1). Media biofilm EPS extracts had the highest normalized ON concentration amongst the different biopolymer sources. Secondary effluent had a slightly higher normalized ON concentration than GRW samples.


Biopolymers
Humics
OCD
UVD
OND
Signal Response
Retention Time

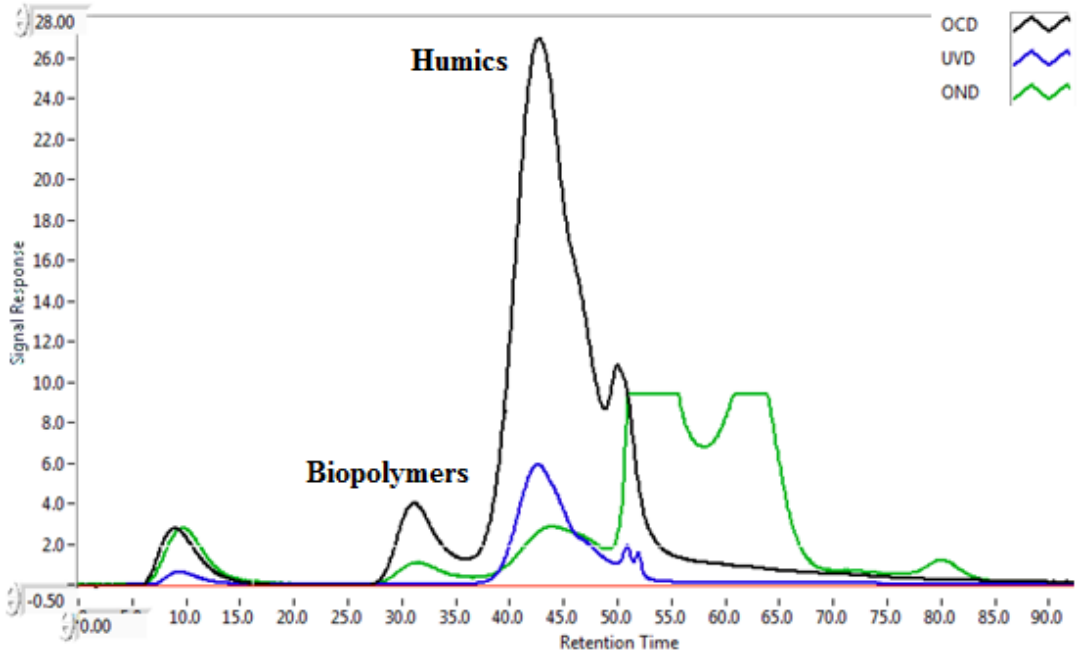
(a)


Humics
Biopolymers
OCD
UVD
OND
Signal Response
Retention Time

(b)



(c)



(d)

Figure 6.2 LC-OCD chromatograms of (a) sand media biofilm EPS extract (08-MAY-2014), (b) anthracite media biofilm EPS extract (08-MAY-2014), (c) secondary effluent (28-JULY-2014), and (d) Grand River water (GRW) (09-JULY-2014)

OCD: organic carbon detector; OND: organic nitrogen detector; UVD: UV detector

Additionally, the biopolymers in biofilter biofilm EPS extracts, both sand and anthracite, had distinct UV responses whereas a much lower UV response was observed for the biopolymers from secondary effluent and GRW (Figure 6.2), illustrating the differences in biopolymer characteristics between biofilter biofilm EPS extracts and other biopolymer sources. These observations were corroborated with the results reported by previous studies. Her *et al*. (2008) showed that the bovine serum albumin as a protein surrogate and asparagine as an amino acid surrogate, which have high density of functional groups associated with aliphatic components, had a low UVA_{254}. Comparatively, for the soluble microbial products and tightly bound EPS extracted from wastewater sludge and fractionated by SEC, the fraction with MW over 20 kDa had high UV_{254} absorbance (Tsai *et al.*, 2008).

As illustrated in Figure 6.2 and Table 6.1, the retention times for the start (27.03 ± 0.23 min) and apex of the bulk biopolymer peaks (OC signal) were relatively consistent among different biopolymer sources; however, the retention time at the end of the bulk biopolymer peak was different (36.05 ± 1.00 min) depending on the source, reflecting the different characteristics of biopolymers from different sources. Therefore, the biopolymer peak width (i.e. collection window) and hence time per fraction (3.01 ± 0.40 min) were different for different biopolymer sources, which is mainly attributed to the variation of retention time at the end of bulk biopolymer peak. Since the biopolymer fractionation was performed based on OC signal, the comparison of three biopolymer sized fractions across different sources was conducted after normalizing the parameter of interest to the carbon mass of corresponding sized fraction. The reproducibility of the LC-OCD fractionation process was validated by comparing the UVD chromatograms of LC-OCD injections for regular (collection time window determination) and fractionation (sized fraction collection) modes. During fractionation the eluting fractions were collected immediately after the UV detector and the OCD signal was therefore not available for validation (Figure 6.1). The UVD chromatograms were very consistent between the two fractionation injections and the regular sample injection (i.e. UV peaks elute at the identical time windows), confirming the reproducibility of LC-OCD based fractionation procedure.

6.3.1.2 FEEM Based Characterization

Secondary effluent and GRW had the highest and lowest protein-like peak intensities amongst the different biopolymer sources, respectively (Table 6.1 and Figure D1) and protein-like peak intensities were similar for both sampling dates for these samples. In contrast, protein-like peak intensities differed substantially between the two sampling events for the biofilm EPS extracts of both media. This can be attributed to the different biofilm status at different time period during the media

acclimation, indicating that the biofilm had not reached steady-state which was also confirmed by the changes in cell number and ATP measurements in another long term study using the same sample (El-Hadidy *et al*., in preparation).

6.3.1.3 Comparison between LC-OCD and FEEM Based Characterization

The concentration of biopolymer-ON did not correlate well with FEEM protein-like peak intensities when all biopolymer sources are included (R^2 = 0.22). Compared to media biofilm EPS extract samples, secondary effluent samples had a lower biopolymer-ON concentration but much higher protein-like peak intensities (Table 6.1), indicating that biopolymer-ON in the secondary effluent has a higher proportion of organic nitrogen that can fluoresce. Hence, the correlation between biopolymer-ON and FEEM protein-like peak intensities was much improved when excluding secondary effluent samples (R^2 = 0.87). This implies that the overall fluorescence properties of biopolymer-ON in the GRW samples and the biofilm EPS extracts were similar. Therefore, the correlation between biopolymer-ON and FEEM protein-like peak intensities depends on the nature of the biopolymer sources and biopolymer-ON cannot necessarily be used as a direct indicator of fluorescent protein-like materials.

6.3.2 Characteristics of Biopolymer Sized Fractions from Different Sources

6.3.2.1 LC-OCD Based Characterization

All re-injected fractions of the background mobile phase, M1 and M2 had very low or no OC and ON responses in the biopolymer region (Figure D2a). OC and ON responses started to appear in the biopolymer region when injecting fraction F1 (Figure 6.1) of all samples, confirming the validity of the fraction collection windows. Both F2 and F3 (Figure 6.1) fractions had distinct OC and ON peaks in the biopolymer region and the peak elution times of F1, F2, and F3 fell within the elution time range of the biopolymers in the corresponding bulk samples (e.g. Figure 6.2 and Figure D2b). Biopolymers in the re-injected fractions F1 and F2 showed tailing as they have now been isolated from humics. However, a shoulder situated on the right side of the biopolymer peak was consistently observed in all F3 fractions of all sources, suggesting that F3 is likely a mixture of biopolymer and humics. This is reasonable since the collection end point of F3 was set to the lowest point between the humics and biopolymer peaks in the bulk samples which were not baseline separated. It should also be noted that the shoulder is more pronounced for the biofilm EPS extract samples (Figures 6.2a and 6.2b) which had much poorer separation between the biopolymers and humics in the bulk samples

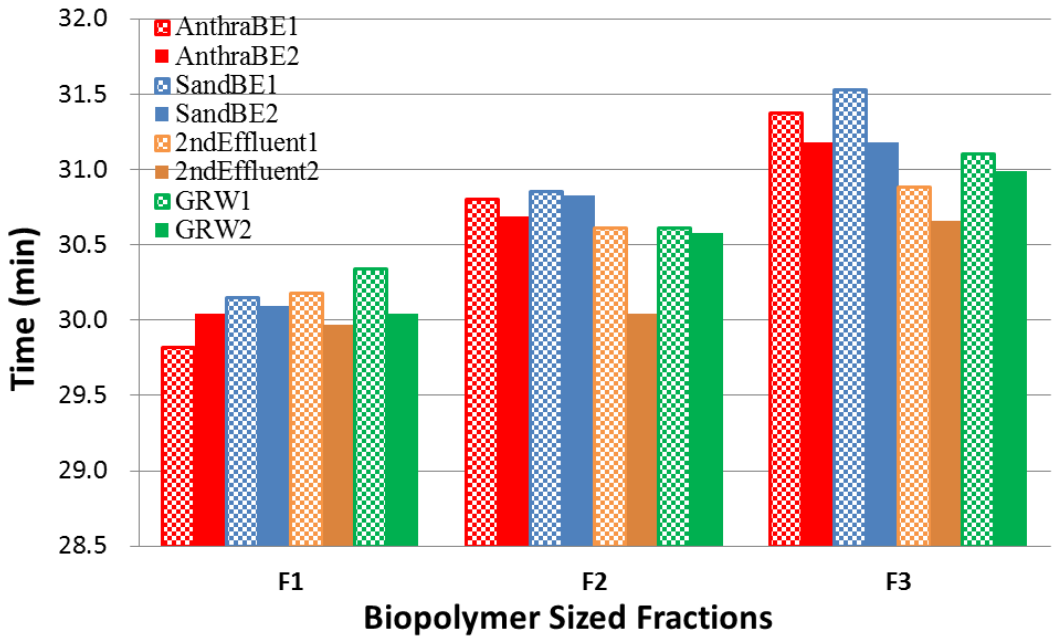
compared to secondary effluent and GRW samples (Figures 6.2c and 6.2d). In addition, F2 for one anthracite media biofilm EPS extract showed a distinct peak in the humics region. This can possibly be attributed to the poorer separation of biopolymer and UV-containing compounds with similar MW as the humics in the corresponding bulk sample. This co-eluted mixture could possibly be better isolated during the LC-OCD analysis of the re-injected sized fraction F2. When estimating the OC and ON-based biopolymer concentrations for this particular F2, the humics peak was not integrated as part of the biopolymers. Otherwise, the entire biopolymer peak including the shoulder was integrated to estimate the biopolymer concentration.

As shown in Figure 6.3a, biopolymer peak elution time increased in the order of F1-F2-F3 for all sources as expected, indicating that the MW/size of biopolymer sized fractions decreases in the same order. This confirmed the success of fractionation procedure. The average and standard deviation of peak elution time for F1, F2, and F3 were 30.08 ± 0.15, 30.63 ± 0.26, and 31.11 ± 0.27, respectively.

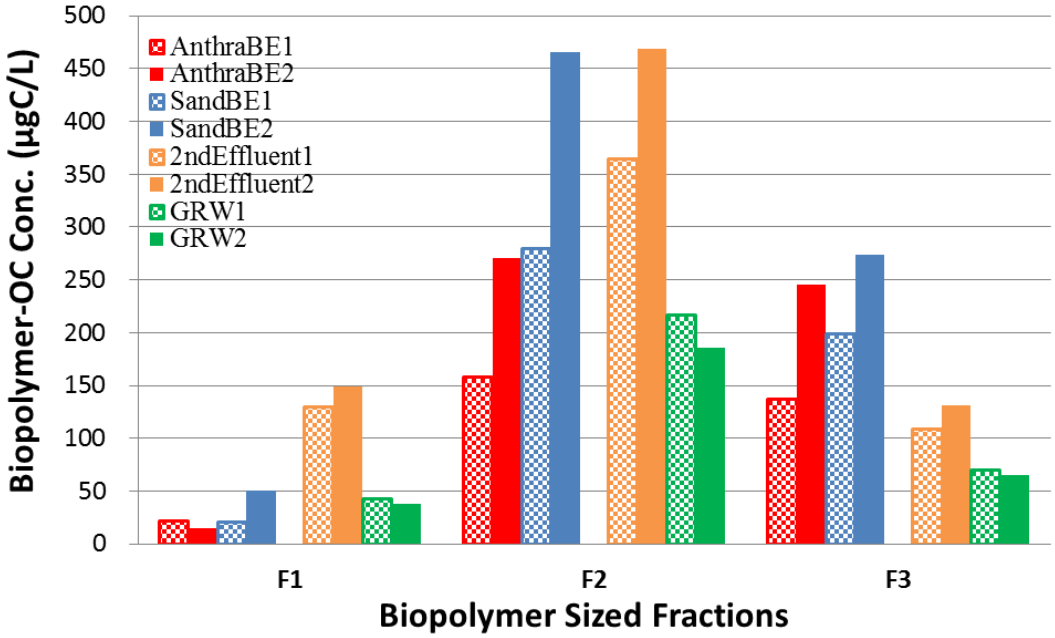
Similar to biopolymers in bulk samples, all three biopolymer sized fractions from different sources contained both OC and ON compounds. However, different sources had different distribution of OC and ON mass into three sized fractions, which is affected by the shape of the biopolymer peak. The differences in the OC and ON concentrations of the biopolymers in the three sized fractions are illustrated in Figures 6.3b and 6.3c. Unlike GRW/secondary effluent samples, the biopolymer concentration of biofilter media biofilm EPS is affected by the extraction process parameters such as extracted volume and media mass. Therefore, the reported biopolymer concentration of each media biofilm EPS sample is the concentration for the biofilm EPS extract sample and is specific to the extraction process used in this study. Therefore, it is not appropriate to directly compare the concentration between media biofilm EPS extract and GRW/secondary effluent samples. Since F2 encompassed the largest portion of peak area of the biopolymers, it is not surprising that F2 had the highest biopolymer-OC concentration amongst the three biopolymer sized fractions for all sources. Also, for the biofilter biofilm EPS extracts, the biopolymer-OC concentration of F1 was much lower compared to F2 and F3 (Figure 6.3b). Comparatively, for secondary effluent and GRW samples, the biopolymer-OC concentrations of F1 and F3 were relatively similar but much lower compared to F2 (Figure 6.3b). It is also noted that F2 and F3 of the anthracite biofilm EPS extract were more similar to each other compared to the sand media biofilm EPS extracts. The same trend was observed for the ON concentrations of biopolymer for the three sized fractions (Figure 6.3c). The above observations

illustrate the different biopolymer-OC and -ON characteristics of three biopolymer sized fractions across different sources, which can affect the shape of the biopolymer peak. In addition, when including all sized fraction samples from different biopolymer sources, biopolymer-OC and -ON were reasonably correlated (R^2 = 0.805). However, the correlation between biopolymer-OC and -ON was improved when including sized fractions from biofilter biofilm extract samples only (R^2 = 0.956). The same observation was found when sized fractions from secondary effluent and GRW samples only (R^2 = 0.952) were included. For sized fractions of different sources, a poorer correlation between OC and ON content indicates that the change of ON with OC behaved in a different way for different sources, reflecting a different biopolymer composition regarding the OC and ON content. The above observations illustrate that the LC-OCD fractionation approach can successfully extract information regarding the composition of the biopolymers in the sized fractions isolated from different sources.

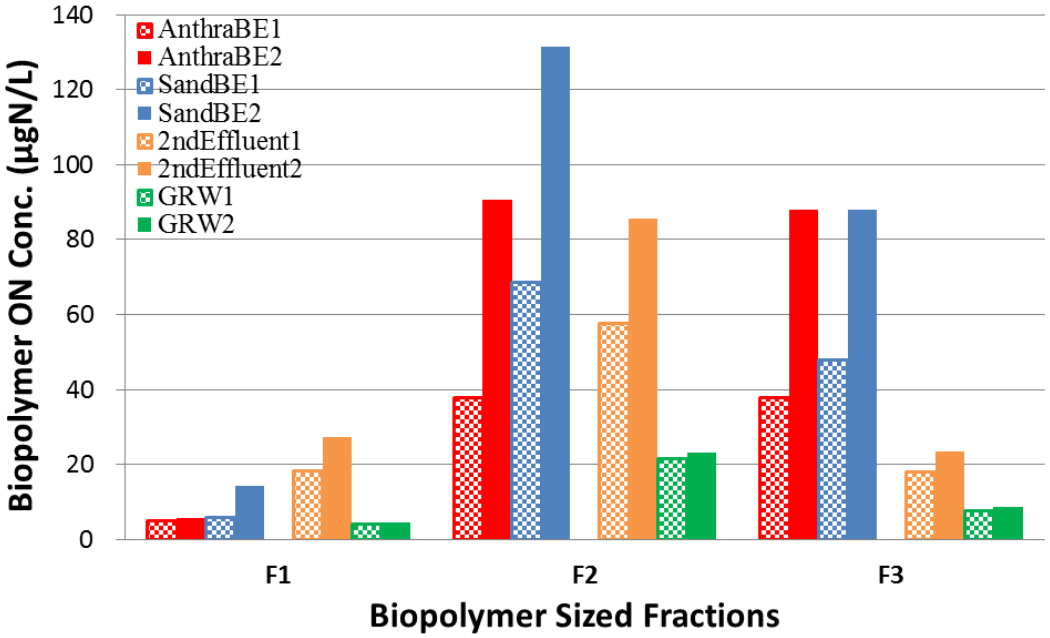
The normalized ON concentration for the three sized fractions from the different biopolymer sources is shown in Figure 6.3d. A reproducibility test with seven injections of GRW samples (DOC=5.3 mgC/L, DF=5, and filtered with seven different polymeric filters), the normalized ON concentration values were 0.11±0.01. Therefore, a normalized ON concentration difference of 0.01 was considered to be a significant difference. Biofilter biofilm EPS extracts for both sand and anthracite and the GRW had highest and lowest normalized ON concentration, respectively. This was also observed for the bulk samples (Table 6.1). However, the three sized fractions had different normalized ON concentration compared to their corresponding bulk samples, suggesting that the LC-OCD based characterization of the bulk samples provides mainly overall characteristics of the biopolymer. In addition, the trends for the normalized ON concentration among three sized fractions of media biofilm EPS extract samples were different between sand and anthracite media. For the same media type (i.e. sand or anthracite), biofilm EPS extract samples from two different sampling events had different normalized ON concentration trend among three sized fractions. These differences might be caused by the different biofilm status in two different media types at different time period during the acclimation. Different normalized ON concentration trends were also observed for two secondary effluent samples, which can be attributed to the biopolymer composition change of secondary effluents. Comparatively, normalized ON concentration consistently increase in the order F1-F2-F3 for two GRW samples in this study, suggesting that higher MW biopolymer sized fraction (i.e. F1) had less ON content per carbon mass than lower MW biopolymer sized fractions for GRW samples.


32.0
31.5
31.0
30.5
30.0
29.5
29.0
28.5
Time (min)
AnthraBE1
AnthraBE2
SandBE1
SandBE2
2ndEffluent1
2ndEffluent2
GRW1
GRW2
F1
F2
F3
Biopolymer Sized Fractions

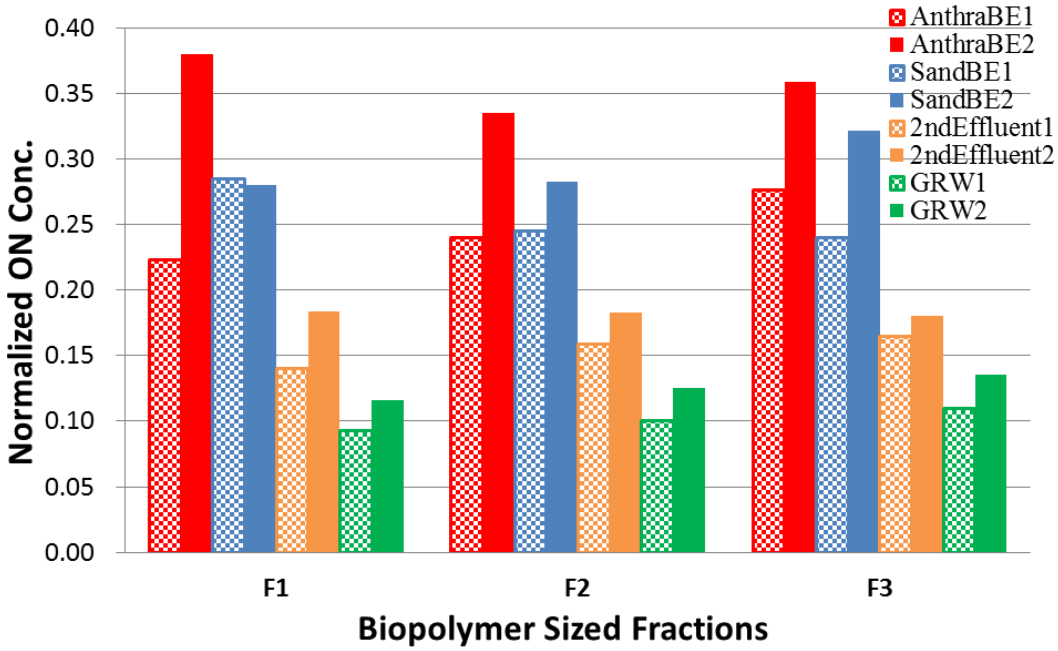
(a)


500
450
400
350
300
250
200
150
100
50
0
Biopolymer-OC Conc. (µgC/L)
AnthraBE1
AnthraBE2
SandBE1
SandBE2
2ndEffluent1
2ndEffluent2
GRW1
GRW2
F1
F2
F3
Biopolymer Sized Fractions

(b)



(c)



(d)

Figure 6.3 LC-OCD based characterization of re-injected isolated biopolymer fractions (a) biopolymer peak elution time, (b) biopolymer-OC, (c) biopolymer-ON and (d) normalized ON concentration (AnthraBE: Anthracite biofilm EPS; SandBE: Sand biofilm EPS; 2ndEffluent: Secondary Effluent; GRW: Grand River water)

Regarding the UV responses, biopolymer sized fractions of media biofilm EPS extract samples had a distinct UV peak in the biopolymer region as was also observed for the corresponding bulk samples (Figures 6.2a and 6.2b). In contrast, biopolymer sized fractions of secondary effluent and GRW samples had very low UV responses in the biopolymer region, illustrating the different UV-related fingerprints/characteristics of biopolymer sized fractions from different sources.

6.3.2.2 FEEM Based Characterization

The protein-like peak intensities of both M1 and M2 were shown to be relatively low compared to the corresponding biopolymer sized fractions (especially F2 and F3) (Figures D3 and D4). For all different biopolymer sources, the protein-like peak started to appear in F1. More distinct protein-like responses were observed in F2 and F3 for each biopolymer source. The presence of protein-like peak in the biopolymer sized fractions of different sources indicates that protein-like materials are important components of biopolymers, which is also reported by other studies using similar samples (Frølund *et al.*, 1996; Christy *et al.*, 1999; Her *et al.*, 2003). The protein-like peak intensities of biofilter biofilm EPS extracts increased in the order of F1-F2-F3. Even though F2 had higher biopolymer-OC concentration than F3 for biofilm EPS extracts, F3 still had higher protein-like peak intensities than F2, implying that lower MW sized fraction of biofilm EPS extracts tends to have more fluorescent protein-like materials. However, F2 of secondary effluent and GRW had higher protein-like peak intensities than F1 and F3, which were found to be relatively similar. This also illustrates the compositional differences of biopolymers among different sources regarding how the bulk content of protein-like material is distributed across three different biopolymer sized fractions.

6.3.2.3 Comparison between LC-OCD and FEEM Based Characterization

The above LC-OCD and FEEM results are corroborated with observations from previous studies. A high performance liquid chromatography-size exclusion chromatography (HPLC-SEC) based technique was developed by Her *et al.* (2003) to characterize the dissolved organic matter (DOM) from lake water and a wastewater secondary effluent. The authors reported that their DOM fraction with high apparent MW (over 10000 g/mol) is a mixture of polysaccharide- and protein-like materials (Her *et al.*, 2003). Earlier studies employing pyrolysis gas chromatography/mass spectrometry (GC/MS) also revealed that biopolymers from aquatic NOM are predominantly comprised of polysaccharides, amino sugars, proteins, polyphenolic compounds, lignins, tannins, DNA and polyhydroxybutyrates (Christy *et al.*, 1999; Vilgé-Ritter *et al.*, 1999). Flemming and Wingender (2010) summarized that EPS are comprised of polysaccharides, proteins, nucleic acids, lipids and

humic substances. In addition, Frølund *et al*. (1996) showed that the major components of CER extracted EPS from activated sludge include protein (46-52% of volatile solids (VS)), humic compounds (18-23% of VS), carbohydrates (17% of VS), uronic acids and DNA. Even though these studies focused on the characterization of biopolymers in bulk samples, they did confirm the LC-OCD and FEEM observations presented here for biopolymers in the bulk samples and for the biopolymer sized fractions.

FEEM protein-like peak intensities (Fi - M2, i = 1, 2, 3) show a reasonable correlation with biopolymer-ON concentrations for sized fraction samples (R^2 = 0.73), which is different from bulk samples (R^2 = 0.22) (Figures 6.3c and D3). The correlation was improved when secondary effluent and GRW samples only (R^2 = 0.96) were included. The relationship between FEEM protein-like peak intensities and biopolymer-ON was also investigated for each biopolymer sized fraction from different sources. F3 (R^2 = 0.908) had a much stronger correlation compared with F1 (R^2 = 0.826) and F2 (R^2 = 0.735), suggesting that the fluorescence properties of biopolymer-ON in F3 samples (with lower MW) are more similar among different sources compared to the larger MW sized fractions F1 and F2.

Different conclusions can be made regarding the similarity of fluorescence properties of biopolymer-ON among different sources when comparing bulk (GRW and media biofilm EPS extract samples were more similar) and sized fractions (GRW and secondary effluent samples were more similar). This emphasizes that characterization of bulk samples can merely provide an understanding of the overall fluorescence properties of biopolymer-ON and can be misleading when different sized fractions of biopolymers had different characteristics and/or a certain sized fraction played a more important role than other sized fractions under certain circumstances. In contrast, characterization of the sized fractions can provide more detail regarding the biopolymer composition and more insight into understanding of the characteristics of each sized fraction. This also explains why the FEEM protein-like peak intensities and biopolymer-ON were only weakly correlated for bulk samples.

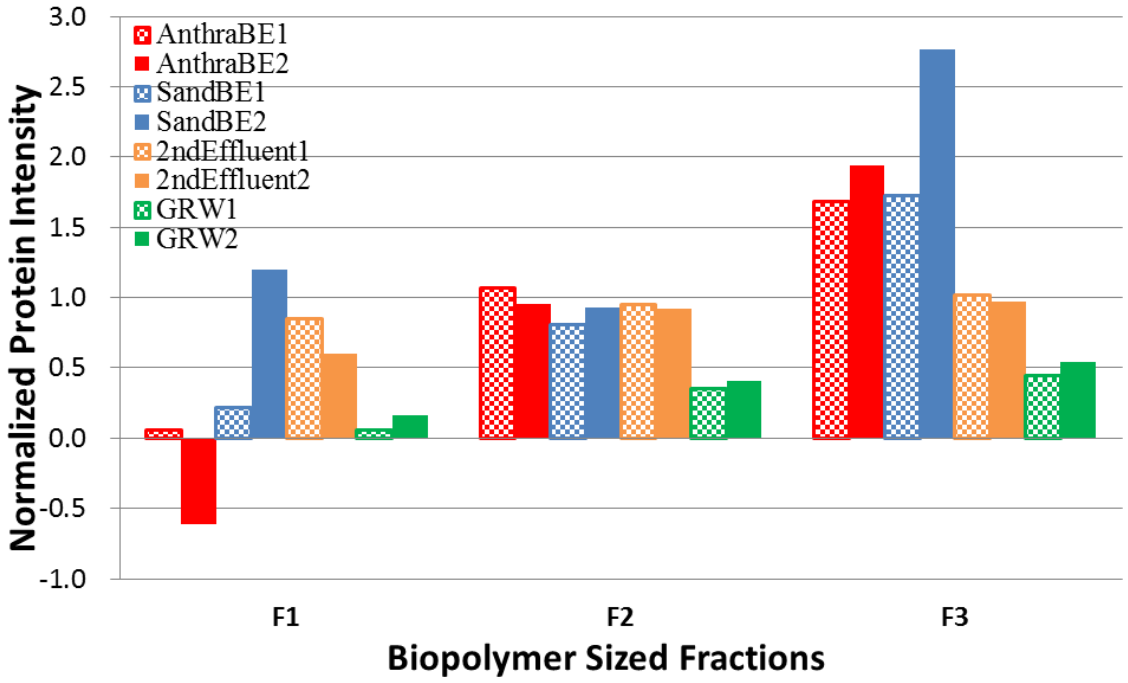


Figure 6.4 Normalized FEEM protein intensities (i.e. (F_i – M2) ÷ biopolymer-OC_i, i = 1, 2, 3) of biopolymer sized fractions

6.3.3 Normalized FEEM Protein-like Peak Intensity

A new parameter, coined the normalized FEEM protein-like peak intensities ($F_{i,\ Normalized}$), is proposed and can be calculated as follows:

$$F_{i,\ Normalized} = (F_i - M_2)/\text{biopolymer-}OC_i,\ i = 1, 2, 3$$

where $F_{i,\ Normalized}$ is the normalized FEEM protein-like peak intensity, F_i is the FEEM protein-like peak intensity of Fraction i, M_2 is the FEEM protein-like peak intensity of M_2, and biopolymer-OC_i is the OC concentration of biopolymer of Fraction i.

This new parameter can be used to evaluate the fluorescence properties of organic carbon of biopolymer and provide information regarding the protein content (where the FEEM protein-like peak intensities is used) per carbon mass of biopolymer. As shown in Figure 6.4, F3 consistently had the highest normalized FEEM protein-like peak intensities among three biopolymer sized fractions for all different sources. Except for one biofilm EPS extract for sand, F2 always had a higher normalized

FEEM protein-like peak intensity than F1. It can be concluded that the biopolymer sized fraction F1 with a higher MW has a lower normalized FEEM protein-like peak intensity than the sized fraction F3 with a lower MW for all biopolymer sources, suggesting that the biopolymer-OC has stronger protein-like fluorescence in the sized fraction with lower MW. Amongst various biopolymer sources, secondary effluent samples had a much higher normalized FEEM protein-like peak intensity in F1 than the other two sources except for one sand biofilm EPS extract sample (Figure 6.4). However, secondary effluent samples had similar and much lower normalized FEEM protein-like peak intensities in F2 and F3 compared to media biofilm EPS extract samples, respectively. GRW samples had the lowest normalized FEEM protein-like peak intensities in both F2and F3 amongst various biopolymer sources. In addition, for GRW samples, F1 was found to be mainly comprised of non-fluorescent compounds and protein-like materials are relatively more abundant in F2 and F3. Comparatively, all three biopolymer sized fractions of secondary effluent samples had consistently high levels of protein-like fluorescence per carbon mass, implying the biopolymer-OC contained protein-like materials in all different sized fractions. Anthracite biofilm EPS extracts had a very low content of protein-like materials per carbon mass in F1 as did GRW samples. However, biopolymer-OC in F2 and F3 of anthracite biofilm EPS extract samples had much higher content of protein-like materials than GRW samples. Similar trend were found for the sand biofilm EPS extract samples except for the high level of protein-like material in the F1 of one sample, which might be attributed to the different biofilm status at different acclimation stages.

The relationship between normalized ON concentration and the normalized FEEM protein-like peak intensities were also investigated for sized fractions. The correlation between two parameters was weak (R^2 = 0.421) when sized fractions from all biopolymer sources were included, whereas the correlation was improved when biofilter biofilm EPS extracts were excluded (R^2 = 0.689). The correlation of two parameters for each sized fraction followed the order of F1 (R^2 = 0.121), F2 (R^2 = 0.510), F3 (R^2 = 0.832). Therefore, the fluorescence and ON properties of biopolymer sized fractions from different sources were more comparable to each other in F3, suggesting that biopolymer-ON in F3 is likely mainly comprised of protein-like material. The weak correlation between two parameters for F1 and F2 can be attributed to the compositional difference of biopolymer sized fraction among different biopolymer sources (i.e. the protein-like fluorescence might not be necessarily abundant in the biopolymer-ON of these two sized fractions).

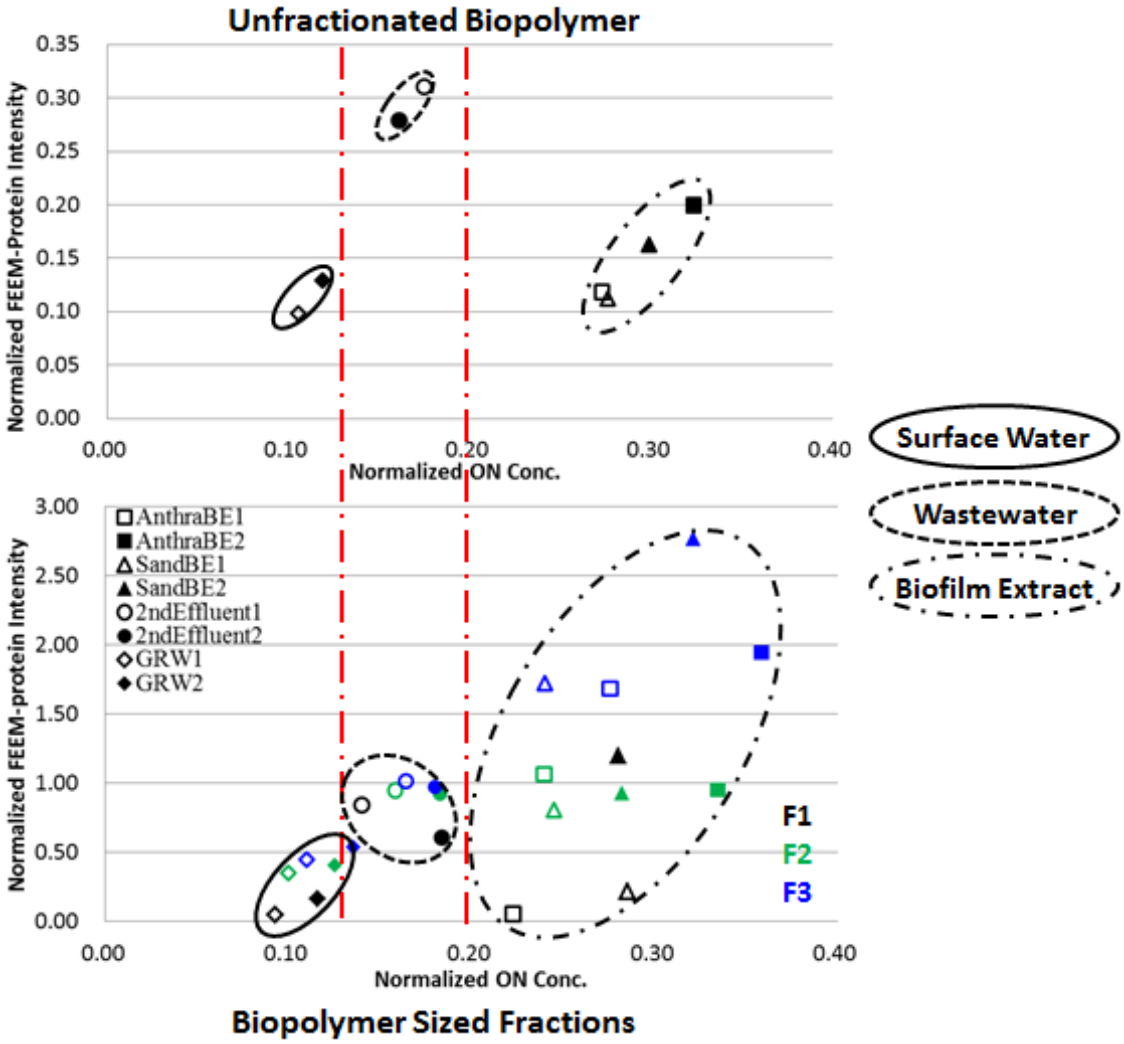


Figure 6.5 Relationship between normalized ON concentration and normalized FEEM protein intensities of bulk and sized fraction samples from different biopolymer sources.

Note: F1 of AnthraBE2 has a negative value of normalized FEEM protein intensity and therefore was not plotted

Additionally, the relationship between normalized ON concentration and normalized FEEM protein-like peak intensities can be used to map the fluorescence and ON fingerprints of biopolymer from different sources as illustrated in Figure 6.5. The sized fractions from three different biopolymer sources clustered into three separate groups with different normalized ON concentration. However, it is noted that three biopolymer sized fractions (i.e. F1, F2, and F3) of GRW and secondary effluent samples tend to cluster together (i.e. the difference of normalized FEEM protein-like peak intensities

is relatively small within each group) whereas three sized fractions of biofilter biofilm EPS extracts tend to be more scattered interpreted as the difference of normalized FEEM protein-like peak intensities is more distinct within each group. Figure 6.5 helps distinguish different biopolymer sources based on their normalized ON concentration and demonstrates the different protein-like properties based on fluorescence for different biopolymer sized fractions within each source. For the characterization of the bulk samples, the normalized ON concentration and normalized protein-like peak intensities were shown to be uncorrelated when all biopolymer sources were included, whereas there appeared to be a good correlation when including media biofilm EPS extract samples only or GRW and secondary effluent samples only. Similar to the sized fraction samples, bulk samples from different biopolymer sources clustered into three separate groups based on their normalized ON concentration (Figure 6.5). It is also noted that the bulk and sized fraction samples for each biopolymer source had a similar normalized ON concentration range. Regarding the normalized FEEM protein intensity, sized fraction samples were much more scattered compared with the bulk samples (Figure 6.5). As mentioned previously, the sized fractions within each source group were scattered across the ranges of normalized ON concentration and normalized FEEM protein-like peak intensities. Comparatively, bulk samples could be considered as an overall representation of fluorescence and ON characteristics of the corresponding sized fraction samples.

6.4 Conclusions

In this study, a LC-OCD based size fractionation approach was combined with advanced NOM analytical methods (i.e. LC-OCD and FEEM) to characterize different biopolymer sized fractions (fractionated from biopolymers) from various biopolymer sources, including biofilter media biofilm EPS extracts, secondary municipal effluent, and natural river water from an agricultural and industrially impacted region. The differences of biopolymer characteristics based on UV, ON, OC and fluorescence for different sources were demonstrated for both a bulk as well as sized fractions. The following conclusions can be drawn:

- Organic nitrogen (ON) concentration did not necessarily correlate with FEEM protein-like peak intensities for both biopolymers in the unfractionated sample and in sized fractions.
- Different biopolymer sources were well distinguished by the normalized ON concentration in the unfractionated sample and in the biopolymer sized fractions.

- Higher normalized FEEM protein-like peak intensities were observed for the lower molecular weight (MW) biopolymer sized fractions from all biopolymer sources, suggesting more protein-like materials are present in these fractions.
- This LC-OCD based fractionation approach provided more insight into the compositional characteristics of the biopolymers and the NOM characteristics of different sized fractions in addition to the overall information extracted from the regular characterization of unfractionated samples.
- Similar approach would be useful to characterize other fluorescent NOM fractions, such as the humic substances, to better understand their composition and properties.

Chapter 7
Kinetics of Natural Organic Matter (NOM) Removal during Biofiltration Using NOM Characterization Techniques

Summary

To better understand biofiltration, concentration profiles of various natural organic matter (NOM) components throughout a pilot-scale biofilter were monitored and investigated using liquid chromatography – organic carbon detection (LC-OCD) and fluorescence excitation and emission matrices (FEEM). During each sampling event, water samples were collected from six ports at different media depths of the biofilter. In this study, the biofilter is demonstrated to be a promising technique for the removal of the biopolymers (i.e. large molecular weight (MW) NOM components as characterized by LC-OCD) and FEEM protein-like materials. However, NOM components with relatively lower MW, including humic substances, building blocks, low MW acids and neutrals, were shown to be much more recalcitrant to biofiltration. The above findings corroborated previous studies in the context of drinking water treatment. The removal kinetics investigation revealed that FEEM protein-like materials, which are presumably also contained within the biopolymers, in both filtered and non-filtered samples (0.45 μm filter) have different removal kinetics compared to the LC-OCD-based biopolymers, illustrating the complementary nature of LC-OCD and FEEM techniques. As the NOM components with higher percent removal, LC-OCD biopolymers (both organic carbon and nitrogen) and FEEM protein-like materials (in both filtered and non-filtered samples) were shown to follow either the first or second order kinetics. Due to the low percent removal, the performance of three kinetic models was not distinguishable for humic-like substances as characterized either by LC-OCD or FEEM. Further, more detailed investigation should be undertaken in the future. In addition, various NOM components as characterized by FEEM, including humic acid (HA)-, fulvic acid (FA)-, and protein-like materials, were shown to have different removal behaviours and/or kinetics for filtered (using a 0.45 μm filter) and non-filtered samples. This illustrates the value of performing FEEM characterization on both non-filtered and filtered samples.

7.1 Introduction

Natural organic matter (NOM), which can act either as a significant carbon sink or source, is ubiquitous in natural aquatic systems such as surface waters, soil pore waters, and shallow groundwater (e.g. Maurice *et al.*, 2002; Leenheer and Croué, 2003). The concentration and

characteristics of aquatic NOM including composition and physciochemical properties, are significantly impacted by the sources, the presence of biogeochemical and/or photo-reactive processes, surface properties of sediment sorbents, and physicochemical properties of the water in the aquatic system (Leenheer and Croué, 2003; Baghoth *et al.*, 2008). Due to the presence of aquatic NOM, various drinking water-related issues can arise including adverse aesthetic problems (e.g. Aoustin *et al.*, 2001), formation of carcinogenic disinfection by-products by reacting with disinfectants/oxidants (e.g. Kitis *et al.*, 2002), introduction of contaminants into the drinking water treatment processes with NOM as a carrier (e.g. Leenheer and Croué, 2003; Matilainen *et al.*, 2010), deterioration of the performance of granular activated carbon and of membrane filtration (e.g. Hallé *et al.*, 2009; Baghoth *et al.*, 2011), pipeline corrosion (e.g. Broo *et al.*, 1999), and bacterial regrowth in distribution systems (e.g. Jacangelo *et al.*, 1995; Basu and Huck, 2004).

NOM is a group of heterogeneous macromolecules with aromatic and aliphatic hydrocarbon structures of varying molecular weight/size, hydrophobicity, and functional groups (Leenheer and Croué, 2003; Matilainen *et al.*, 2010). Traditional bulk NOM characterization techniques either cannot capture differences among different NOM groups or can only characterize one specific group of compounds (e.g. assimilable organic carbon (AOC), UV254 absorbance and specific UV absorbance). XAD resin fractionation technique has been widely employed to isolate and characterize various NOM groups with different hydrophobicity (Leenheer, 1981; Matilainen *et al.*, 2011). However, this approach is very time-consuming and has several limitations such as alteration of NOM characteristics due to extreme pH conditions and contamination caused by resin bleeding (Matilainen *et al.*, 2011). Recently, liquid chromatography – organic carbon detection (LC-OCD) and fluorescence excitation emission matrices (FEEM) have been gaining increasing acceptance as NOM characterization tools (Hallé *et al.*, 2009; Peiris *et al.*, 2010b; Huber *et al.*, 2011; Baghoth *et al.*, 2011; Lavonen *et al.*, 2015). LC-OCD, which is essentially a size-exclusion-based chromatography technique, separates the NOM constituents into five major method-defined groups based on their molecular size: biopolymers (e.g. polysaccharides, proteins and amino sugars), humic substances (e.g. fulvic and humic acids), building blocks (hydrolysates of humic substances), low molecular-weight (MW) acids and neutrals (Huber *et al.*, 2011). FEEM characterizes the NOM constituents based on their functional group/fluorophore composition without separating constituents by size. Several groups, humic acids, fulvic acids and protein-like substances, can be semi-quantified by either applying peak picking to key coordinates of the FEEM (Baker *et al.*, 2007) or by combining the

FEEM analysis with multivariate statistical analysis (Baghoth *et al.*, 2011; Peiris *et al.*, 2010b). In addition, pre-filtration with polymeric filters was shown to have impacts on the fluorescence intensities of protein-like materials (Baker *et al.*, 2007). The effect of this phenomenon on the NOM monitoring for a drinking water treatment train has rarely been studied.

Biofiltration, which is a biologically active media filtration process, was initially developed to reduce biodegradable organic matter (BOM) through natural biodegradation, and in many cases, biofiltration also serves to remove particles through media filtration (e.g. Urfer *et al.*, 1997). The biodegradation process is mediated by the biofilm, an interface/media-associated layer of bacteria cells embedded within extracellular polymeric substances (Wanner *et al.*, 2006; Wilking *et al.*, 2011). In addition to BOM removal, biofiltration has also been used to mitigate other drinking water problems, including the removal of trace contaminants such as pharmaceuticals and endocrine disruptors (Huck and Sozański, 2008; Zearley and Summers, 2012; Hallé *et al.*, 2015), taste and odour compounds (Hoefel *et al.*, 2009), and various NOM components (Basu and Huck, 2004; Hallé *et al.*, 2009). Hallé *et al.* (2009) reported that the biopolymer fraction from natural river water is more biodegradable than humic substances and that substantial removal of biopolymers can be achieved through biofiltration. In the presence of other readily biodegradable compounds and sufficient acclimation time, Basu and Huck (2004) reported that the removal of humic substances in a synthetic solution (TOC: 1.6 to 4.5 mg/L; 65% humic acid, 15% formate, 10% acetate, and 10% formaldehyde on a carbon mass basis) could also be significant. Biofiltration performance for NOM removal is affected by various operational parameters, including empty bed contact time (EBCT), temperature, acclimation time, and the presence of other treatment processes (e.g. ozonation) (Hozalski *et al.*, 1999; Basu and Huck, 2004; Huck and Sozański, 2008; Hallé *et al.*, 2009). Even though the ability of biofiltration to remove NOM constituents has been reported, differences in removal behaviour for various NOM components at different bed depths and removal kinetics during biofiltration have not been studied.

The main purpose of the present study was to investigate the removal behaviour and kinetics of various NOM components throughout a pilot-scale biofilter. The utility of two advanced NOM characterization techniques, namely LC-OCD and FEEM, to assist in this evaluation was assessed. In addition, the effect of filtering water samples using 0.45 μm polymeric filters on the outcome of the NOM analysis using FEEM was also investigated.

7.2 Methods and Materials

7.2.1 Feed Water Source

The pilot-scale biofilter was installed in a full-scale water treatment plant (WTP, Southwestern Ontario, Canada). Grand River water (GRW), which is municipally and agriculturally impacted, was the feed water in this study. Raw GRW was initially retained in storage reservoirs for several days and then pumped to the full-scale WTP as the influent stream for the full-scale treatment train. The feed water of the pilot-scale facility was directly diverted from this influent stream.

7.2.2 Pilot-scale Biofiltration Set-up and Sampling Procedure

7.2.2.1 Pilot-scale Biofiltration Set-up

As illustrated in Figure 7.1, the feed water was pre-filtered using a roughing filter (RF) to reduce peaks in the raw water turbidity prior to biofiltration. The RF used crushed gravel as the media and was operated in down flow mode. The RF effluent was then fed into a dual-media, anthracite and sand, biofilter using a magnetic sealed centrifugal pump to avoid contamination. The biofilter was operated in down flow mode at a constant hydraulic loading rate of 3.1 m/h. Due to height limitations at the pilot facility, the biofilter media was loaded into two separate columns operated in series, Column I and Column II with 80 (20 cm anthracite over 60 cm of sand) and 40 cm (sand only) media depths, respectively. The media in the two columns had been acclimated for more than 8 months prior to the current study. At the hydraulic loading rate of 3.1 m/h, the total empty bed contact time (EBCT) overall for the two columns together was 23 min. During the study, the biofilter was backwashed 3 times (Monday, Wednesday, and Friday) every week to maintain the targeted hydraulic loading rate of 3.1 m/h. The backwash procedure included an air scouring at collapse pulsing condition followed by a water backwash with 50% bed expansion volume. More experimental details regarding the pilot-scale biofilter and backwash can be found in Wilson (2015). The targeted hydraulic loading rate was successfully maintained during the majority of the study period except that the hydraulic loading rate dropped to 2.2 m/h in one sampling event because of a high turbidity peak in Grand River (Table E1).

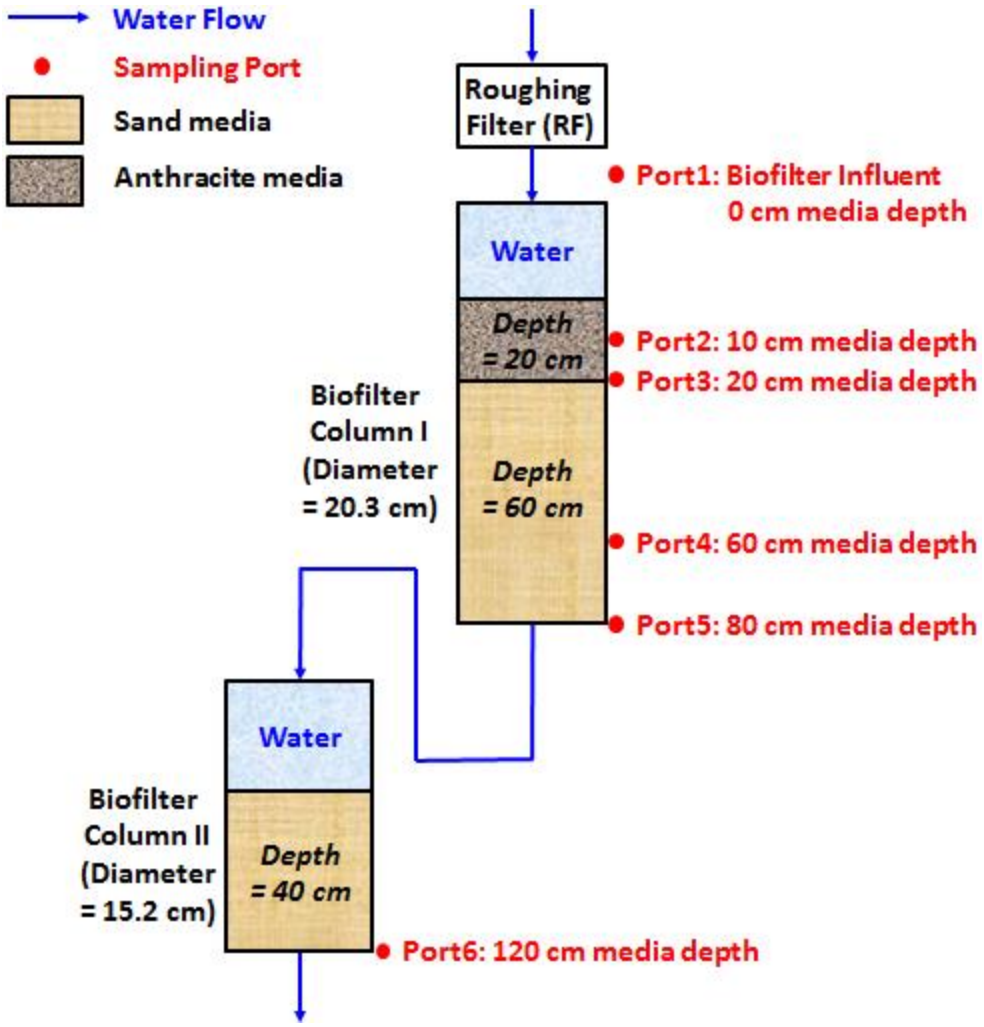


Figure 7.1 Schematic of biofilter pilot-scale set-up including location and depth of sampling ports

7.2.2.2 Location of Sampling Ports and Sampling Procedure

There were six sampling ports for water sample collection throughout the depth of the biofilter as shown in Figure 7.1 and the corresponding EBCT for each port is summarized in Table E2. Five sampling events were completed from August to September in 2014. The sampling events were consistently scheduled on the day following the second biofilter backwash during the week (Thursday). Four sampling events (Aug-14, Aug-28, Sept-04, and Sept-11) were conducted at a hydraulic loading rate of 3.1 m/h and one sampling event (Sept-18) was conducted at a reduced hydraulic loading at a hydraulic loading rate of 2.2 m/h due to a high turbidity event in Grand River water. The water samples were collected in clean glass bottles and analyses were initiated within 1 h of collection. At least 500 mL of water was wasted to flush the sampling ports prior to sample collection.

7.2.3 NOM Characterization Techniques

7.2.3.1 Liquid Chromatography – Organic Carbon Detection (LC-OCD)

The LC-OCD system employed a weak cation exchange column (i.e. polymehacrylate based, TSK HW 50S, Toso, Japan) equipped with three different detectors: organic carbon detector (OCD), organic nitrogen detector (OND), and UV (254 nm) detector (UVD) (Huber *et al.*, 2011). The OC and ON properties of various NOM components were characterized and quantified using a software program provided by the manufacturer (ChromCALC, DOC-LABOR, Karlsruhe, Germany). Prior to LC-OCD analysis, all water samples were pre-filtered through a 0.45 μm polymeric filter (polyethersulfonate, Pall Corporation, Supor®). Details regarding the physical design and description of the LC-OCD system can be found in Huber *et al.* (2011). In this study, five NOM components, including biopolymers, humics, building blocks, Low MW acids and neutrals, were quantified using LC-OCD. Also, DOC was determined using the bypass DOC of the LC-OCD.

7.2.3.2 Fluorescence Excitation Emission Matrices (FEEM)

FEEM analyses were conducted using a Cary Eclipse fluorescence spectrophotometer (Agilent Technologies, Canada) and a quartz cuvette with 4 optical windows. Excitation (Ex) wavelengths ranged from 250 to 380 nm with a sequential increment of 10 nm and emission (Em) wavelengths range from 300 to 600 nm with a scan rate of 600 nm/min (i.e. data interval is 1 nm). The photomultiplier tube (PMT) voltage was set at 650 V and slit widths for both excitation and emission were set to 10 nm. FEEM measurements were performed on both non-filtered and filtered water samples from six sampling ports for each sampling event. For both filtered (0.45 μm filter) and non-filtered samples, peak intensities at Ex/Em = 270 nm/460 nm, Ex/Em = 320 nm/415 nm and Ex/Em = 280 nm/330 nm were used to semi-quantify the humic acid (HA)-, fulvic acid (FA)-, and protein-like content, respectively.

7.2.4 Biofilter Active Biomass Estimation

The biofilter media active biomass was estimated using adenosine triphosphate (ATP) as an indicator. For samples from the media interface at 20 cm filter depth, sand and anthracite media were separated from each other by sieving them gently. Loosely attached biomass was removed by rinsing all media sample using biofilter effluent. 1 g of media and 5 mL of commercial cell lysis buffer (Ultralyse 7, LuminUltra Technologies, NB, Canada) were mixed in a 15 mL sterile polypropylene centrifuge tube using a vortex mixer for 20 s. After 10 min, 20 μL of sample and 100 μL of the constructed luciferase

of the Promega ENLITEN ATP kit (Promega Corporation, WI, USA) were added to a 96 well microplate. The ATP analysis was performed using a GloMax® 96 Microplate Luminometer along with five different ATP standards prepared using the same lysis.

During the study, media samples were collected at 20 cm (media interface), 60 cm and 85 cm filter depth (Figure 7.1) on Aug-07, Aug-21, Sept-11, and Oct-02. ATP was measured for all media samples except for the 60 cm sample from Aug-07. These media samples were taken to reflect the overall biological activity throughout the sampling period and therefore media sampling dates were not necessarily consistent with water sampling dates. In addition, the ATP measurements were also conducted for the roughing filter media.

7.2.5 Kinetics Analysis of NOM Removal by Biofiltration

Zero-, first-, and second-order kinetic models were used to fit the experimental data. The equations used for kinetics analysis are summarized as follows:

Zero order kinetics: $[A] = -kt + [A]_0$
First order kinetics: $ln[A] = -kt + ln[A]_0$
Second order kinetics: $1/[A] = kt + 1/[A]_0$

where k is the reaction rate constant, t is time and [A] is concentration or intensity of component of interest.

The best-fit reaction order(s) and the corresponding rate constants were determined for various NOM components, including LC-OCD based components and FEEM based components. The residuals and standard error were examined for the kinetics analysis. Differences in the reaction kinetics and rate constants for filtered and non-filtered FEEM-based components were also compared.

7.3 Results and Discussions

7.3.1 Active Biomass of Pilot-scale Biofilter

Profiles of the estimations of media active biomass for the pilot-scale biofilter during the study period are given in Figure 7.2. The biofilter media was observed to be bio-active throughout the filter depth,

confirming the success of media acclimation. At the media interface at 20 cm depth, the sand media generally had a higher ATP value than the anthracite media. This may be in part due to the different media diameters and therefore the ratio of specific surface areas between sand and anthracite media (i.e. sand:anthracite = 2.2:1, calculated based on the effective size and uniformity coefficient provided by media supplier). Other factors could also be present and might play a role. In addition, the active biomass of the biofilter media (both sand and anthracite) generally increased as the study period progressed, indicating that development of biomass has not reached a steady state. However, the minimum biomass to maintain the steady-state BOM removal could have reached during the study period, which can be confirmed by the consistent percent removal of DOC and biodegradable NOM components in this study (more details can be found in the following sections). This trend was more distinct for the upper media compared to the lower media at depths of 60 and 85 cm. It should also be noted that the active biomass of biofilter sand media decreased with increasing depth, indicating that the active biomass profile was consistently continuous throughout the entire biofilter even though the biofilter was comprised of two separate columns due to height limitations at the pilot-plant facility. In addition, the active biomass of the preceding roughing filter media was always at a low level (i.e. less than 10% of the biomass of biofilter media), confirming that the biological activities mainly occurred within the biofilter.

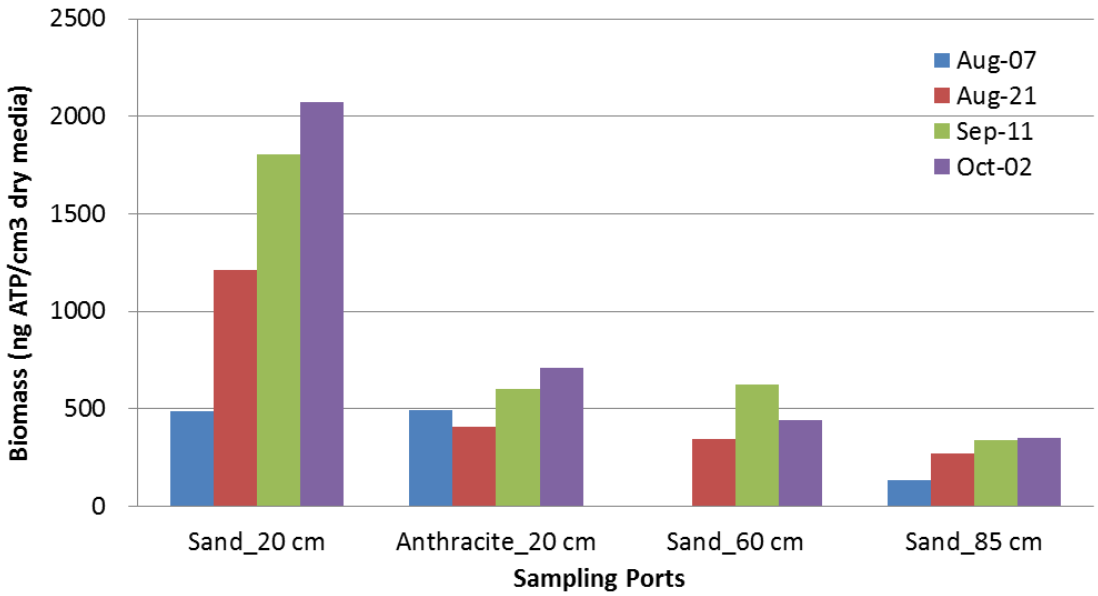


Figure 7.2 Biomass ATP based profile of the biofilter during the study period

7.3.2 Concentration and Intensity of Various NOM Components in the Biofilter

7.3.2.1 Characteristics of Various NOM Components in Biofilter Influent

The overall organic content as indicated by the DOC can be further fractionated into five NOM fractions by LC-OCD. As the largest MW component, biopolymers are considered to be comprised of polysaccharides, protein, and amino sugars (Huber *et al.*, 2011). Protein-like materials could also be characterized using FEEM based approach. Humic substance, which is the second largest MW component fractionated by LC-OCD, could be characterized using HA- and FA-like responses in FEEM as well. However, it should be noted that LC-OCD and FEEM characterize the NOM components based on molecular size and functional group composition, respectively. Therefore, LC-OCD and FEEM based NOM components with similar definition (e.g. LC-OCD humics vs. FEEM-HA and FA) are not necessarily entirely the same regarding the nature and composition. Other lower MW components, including building blocks as well as low MW acids and neutrals, were only investigated using LC-OCD in this study.

The concentrations of DOC and humics (i.e. LC-OCD humics, FEEM-FA and FEEM-HA) remained stable for the first 3 sampling events and increased for the 4^{th} and 5^{th} sampling events (Figure 7.3) for the biofilter influent. It is not surprising that DOC and humics had similar trend since humic substances are the major source of DOC in GRW. The biopolymer organic carbon and nitrogen concentrations were relatively similar during the first four sampling events but decreased to a lower level during the 5^{th} sampling event. For filtered (with 0.45 μm filter) samples, the FEEM protein peak intensities in the biofilter influent were relatively similar during the study period, however, non-filtered samples showed more variability across the 5 sampling events. The difference in trends between the LC-OCD biopolymers and FEEM protein illustrates the complementary nature of the two techniques, allowing a better understanding of NOM characteristics from different perspectives (molecular size vs. fluorophore composition). The differences between filtered and non-filtered FEEM protein might be attributed to the presence of colloidal/particulate protein-like material in the biofilter influent which is removed during 0.45 μm filtration. In terms of the lower MW NOM groups, the concentration of building blocks generally increased during the study period whereas the concentration of low MW neutrals in biofilter influent was relatively stable (Table E3). The concentrations of low MW acids were consistently found to be very low or could not be detected in the biofilter influent (Table E3).

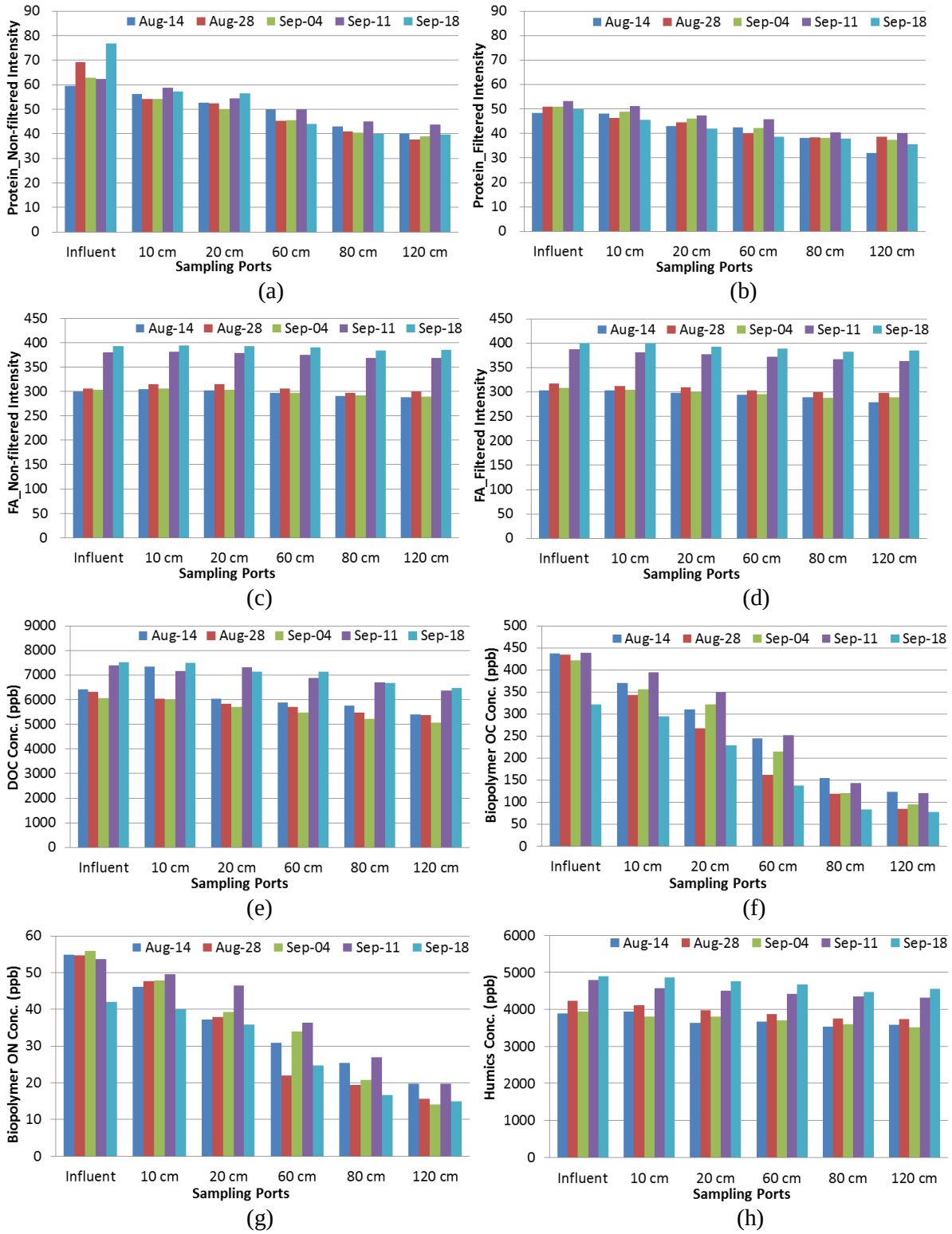


Figure 7.3 Biofilter concentration and FEEM intensity profiles of (a) FEEM-protein (non-filtered), (b) FEEM-protein (filtered), (c) FEEM-FA (non-filtered), (d) FEEM-FA (filtered), (e) DOC, (f) LC-OCD-biopolymer (organic carbon), (g) LC-OCD-biopolymer (organic nitrogen), and (h) LC-OCD-humics
Note: sampling event on Sept-18 was conducted at a reduced hydraulic loading

7.3.2.2 Concentration and Intensity Profiles of Various NOM Components in the Biofilter

As illustrated in Figure 7.3, the intensities/concentrations of FEEM-protein in both non-filtered and filtered samples (Figures 7.3a and 7.3b) as well as LC-OCD-biopolymers (organic carbon and nitrogen, Figures 7.3f and 7.3g), were observed to consistently decrease with increasing media depth in the biofilter. Comparatively, the intensities/concentrations of humic substances, including FEEM-FA in both non-filtered and filtered samples (Figures 7.3c and 7.3d), FEEM-HA in both filtered and non-filtered samples (Table E3), and LC-OCD-humics (Figure 7.3h), did not change substantially as a function of filter depth for each sampling event. It should also be noted that FEEM-FA and -HA intensities of non-filtered samples increased slightly from Port 1 (i.e. biofilter influent) to Port 2 (i.e. 10 cm media depth from top, anthracite layer); however, this trend was not observed for the filtered samples. Baker *et al*. (2007) found that the surface of some materials can potentially sorb humic-like materials characterized by FEEM. Therefore, the slight increase in the FEEM peak intensities of humic-like materials in non-filtered samples may result from humic-like materials adsorbed onto colloidal/particulate matter which were retained in the anthracite layer. Similar to the humic substances, the concentrations of low MW neutrals did not change dramatically along the media depth (Table E3). In addition, the lower MW NOM groups, LC-OCD building blocks and low MW acids had no clear trend in their concentration profiles (Table E3). As the overall/bulk parameter combining the contents of various NOM components, DOC concentrations (as measured by the LC-OCD instrument) generally decreased with the increasing media depth (Figure 7.3e). The high DOC concentration at Port 2 in the first sampling event was due to the high concentration of lower MW compounds (i.e. building blocks and low MW neutrals) from unknown sources.

7.3.3 Removal Behaviour of Various NOM Components in the Biofilter

7.3.3.1 Removal Behaviour of Various NOM Components at the Target Hydraulic Loading Rate

The percent removal of FEEM protein-like materials in both filtered and non-filtered samples (Figures 7.4a and 7.4b) as well as LC-OCD biopolymers, i.e. organic carbon and nitrogen (Figures 7.4f and 7.4g), consistently increased with the increasing bed depth. For non-filtered samples from different sampling events, higher FEEM protein peak intensities in the biofilter influent resulted in a higher percent removal deeper in the biofilter (i.e. a steeper FEEM protein intensity profile). However,

for filtered (with 0.45 μm filter) samples, the percent removal at the downstream/deeper sampling ports did not depend on the biofilter influent FEEM protein peak intensities.

The percent removal of humic-like materials, including FEEM-FA in both filtered and non-filtered samples (Figures 7.4c and 7.4d), FEEM-HA in both filtered and non-filtered samples (Table E3), and LC-OCD humics (Figure 7.4h), were consistently low but generally increased with increasing media depth. LC-OCD low MW neutrals had a similar removal behavior during biofiltration (Table E3). However, LC-OCD building blocks (BB) and low MW acids did not have any distinct removal trend (Table E3). The percent removal of DOC generally increased with increasing media depth (Figure 7.4e). The differences of percent removal between DOC and NOM components illustrated the benefits of monitoring various NOM components instead of general parameters such as DOC. It is also noted that the overall percent removal of DOC at Port 6 was relatively consistent across different sampling events (Figure 7.4e).

LC-OCD biopolymers, measured using both the organic carbon and nitrogen detectors, were found to have higher percent removal compared to FEEM protein-like materials in both filtered and non-filtered samples. This trend is more distinct for the sampling ports (Ports 4, 5 and 6) deeper in the column. The differences in removal behavior between LC-OCD biopolymers and FEEM protein-like materials during biofiltration are reasonable because protein-like materials are only one of the key components in the biopolymers. However, the FEEM protein-like materials might not be entirely the same as the proteins contained within LC-OCD biopolymers. The percent removal of LC-OCD humics was slightly higher compared to the FEEM-HA and -FA in both filtered and non-filtered samples. This can be attributed to the difference in NOM characterization approaches where LC-OCD is size exclusion based and FEEM is fluorophore based. The two techniques characterize the same NOM group (i.e. humic-like material) in different ways. Not all components captured by LC-OCD are necessarily chemically functionalized in a way that would allow them to be fluorescent. As a result, they are not picked up by FEEM. Also, not all of the mixture of HA and FA will quantitatively fluoresce in the same way, hence their removal will not be captured quantitatively in the same way by FEEM in contrast to following the carbon components by LC-OCD. This illustrates the complementary nature of LC-OCD and FEEM regarding NOM characterization.

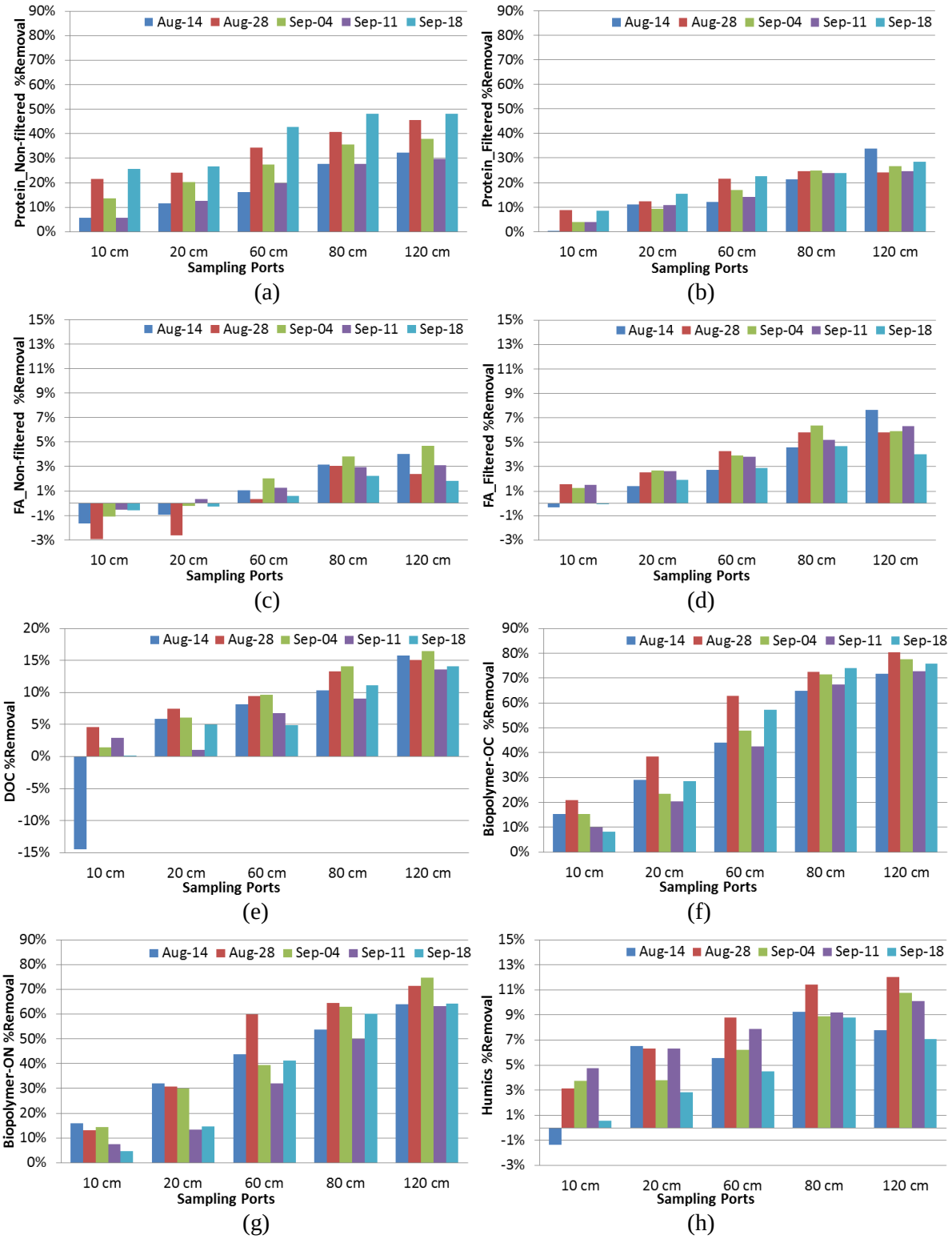


Figure 7.4 Biofilter profiles of (a) FEEM-protein (non-filtered), (b) FEEM-protein (filtered), (c) FEEM-FA (non-filtered), (d) FEEM-FA (filtered), (e) DOC, (f) LC-OCD-biopolymer (organic carbon), (g) LC-OCD-biopolymer (organic nitrogen), and (h) LC-OCD-humics

Note: sampling event on Sept-18 was conducted at a reduced hydraulic loading

7.3.3.2 Removal Behaviour under the Reduced Hydraulic Loading Conditions

Under reduced hydraulic loading conditions (i.e. 2.2 m/h), the percent removal of FEEM protein-like materials in non-filtered samples was higher compared to those under the operating conditions with a hydraulic loading rate of 3.1 m/h (Figure 7.4). This trend was less distinct for FEEM protein-like materials in filtered samples and was not observed for the LC-OCD biopolymers for both OC and ON, indicating that reduced hydraulic loading due to high turbidity peak had a more dramatic effect on particle removal and therefore the protein-like materials absorbed onto colloidal/particulate matter. Comparatively, the percent removal of humic-like materials, i.e. LC-OCD humics, FEEM-HA and -FA, were consistently lower at the reduced hydraulic loading condition compared to the 3.1 m/h operation mode even though their concentrations were higher in the biofilter influent. This could possibly be attributed to the different composition of humic-like substances because of the high turbidity event in Grand River water and the presence of larger amount of colloidal/particulate matter.

7.3.4 Kinetics of NOM Component Removal

7.3.4.1 Reaction Order

A bioreactor having an axial Peclet number (Pe) greater than 40 could be regarded as a plug flow reactor (Bailey and Ollis, 1986). Zhang (1996) performed calculations to show that the axial hydrodynamic dispersion could be neglected and hence the pilot-scale GAC contactors and biofilters used to provide data for his modelling could be treated as plug flow reactors. This assumption has been confirmed for the pilot-scale biofilter used in the present study by estimating the axial Peclet number which was at least two orders of magnitude greater than 40. Therefore the biofilter is regarded to be a plug flow reactor for the kinetics analysis in this study. In addition, the exponential decrease of biomass ATP (Figure 7.2) along the media depth of the pilot-scale filter provided further qualitative validation of this assumption.

Because the R^2 values of zero-, first-, and second-order kinetic models for various NOM components were very similar, for ease of interpretation they are presented in categories in Table 7.1 (The actual R^2 values can be found in Table E4). Examples of kinetics plots are illustrated in Figure E1. The residual plots (e.g. Figure E2) had no indication of systematic trends for kinetics analysis, although it should be noted that results from these residual plots should not be over-interpreted since there are only 6 data points for each kinetics analysis.

In most cases, the organic carbon (OC) and organic nitrogen (ON) concentrations of LC-OCD biopolymers followed first and second order kinetics better than zero order kinetics. For biopolymer-OC, first order kinetics was the best-fit model in most cases (4 sampling events). However, first order kinetics did not work better than the second order kinetics for biopolymer-ON. In addition, a slight difference of R^2 between the first and second order models indicates that the order of biodegradation kinetics can possibly switch between first and second orders and further investigation is required to make more definitive conclusions. In general the biopolymer-ON and -OC had similar removal kinetics (either first or second order) in the biofilter during the study period. Similar to biopolymer-OC and -ON, first and second order kinetics performed better than the zero order kinetics in most cases for FEEM protein-like materials for both filtered (with 0.45 μm filter) and non-filtered samples. It is noted that second order kinetics were always a better fit than first order. This observation was much more consistent for FEEM protein-like materials compared to the biopolymer-OC and -ON, which might be attributed to that the composition of biopolymers as characterized by LC-OCD is based on size-exclusion and is more complex compared to the FEEM protein-like materials as characterized by FEEM. Therefore, it can be concluded that LC-OCD biopolymers (both OC and ON) and FEEM protein-like materials (in both filtered and non-filtered samples) followed the similar kinetics (either first or second order) even though difference in the best-fit model was noticed for these two NOM groups.

For humic-like substances, including LC-OCD humics, FEEM-HA and -FA, the performance of three different kinetic models were not distinguishable due to the low percent removal of these NOM components. Since other lower MW NOM components, including building blocks, low MW neutrals and acids, had either very low percent removals or showed no clear trend in their concentration profiles and removal behaviour, their removal kinetics were not investigated in this study.

Table 7.1 Summary of R^2 categories (Category I: 0.95 to 1; Category II: 0.90 to 0.94; Category III: 0 to 0.89) of zero-, first-, and second-order kinetics models for various NOM components
Note: sampling event on Sept-18 was conducted at a reduced hydraulic loading

	Zero order kinetics				
	Aug-14	**Aug-28**	**Sep-04**	**Sep-11**	**Sep-18**
LC-OCD_BP-OC	II	III	II	I	II
LC-OCD_BP-ON	II	III	II	***I***	II
FEEM-Protein_NF	II	III	III	II	III
FEEM-Protein_F	***II***	III	II	III	III
FEEM-HA_NF	***III***	III	***III***	***III***	III
FEEM-HA_F	I	I	III	III	III
FEEM-FA_NF	III	III	II	II	***III***
FEEM-FA_F	***I***	III	III	II	III
LC-OCD_Humics	III	III	II	III	***III***
	First order kinetics				
	Aug-14	**Aug-28**	**Sep-04**	**Sep-11**	**Sep-18**
LC-OCD_BP-OC	***I***	I	***I***	***I***	***II***
LC-OCD_BP-ON	I	I	***I***	I	***I***
FEEM-Protein_NF	I	III	III	II	III
FEEM-Protein_F	II	III	I	II	III
FEEM-HA_NF	III	III	III	III	III
FEEM-HA_F	I	I	III	III	III
FEEM-FA_NF	III	III	II	II	III
FEEM-FA_F	I	III	III	II	III
LC-OCD_Humics	III	III	II	III	III
	Second order kinetics				
	Aug-14	**Aug-28**	**Sep-04**	**Sep-11**	**Sep-18**
LC-OCD_BP-OC	I	***I***	II	II	II
LC-OCD_BP-ON	***I***	***I***	II	I	I
FEEM-Protein_NF	***I***	***II***	***II***	***II***	***III***
FEEM-Protein_F	II	***III***	***I***	***II***	***II***
FEEM-HA_NF	III	***III***	III	III	***III***
FEEM-HA_F	***I***	***I***	***III***	***III***	***III***
FEEM-FA_NF	***III***	***III***	***II***	***II***	III
FEEM-FA_F	I	***III***	***III***	***II***	***III***
LC-OCD_Humics	***III***	***III***	***II***	***III***	III

BP: Biopolymer; OC: Organic Carbon; ON: Organic Nitrogen;
NF: Non-Filtered; F: Filtered; HA: Humic Acid; FA: Fulvic Acid
NOTE: The R^2 of best-fit kinetics is bolded and italicized for each NOM component for each sampling event.

Table 7.2 Summary of first-order reaction rate constants (min^{-1}) for various NOM components

Note: sampling event on Sept-18 was conducted at a reduced hydraulic loading

	BP-OC	***BP-ON***	***Protein_NF***	***Protein_F***
Aug-14	0.054	0.041	0.016	0.017
Aug-28	0.069	0.055	0.023	0.011
Sep-04	0.067	0.057	0.019	0.014
Sep-11	0.059	0.043	0.015	0.012
Average[a]	0.062	0.049	0.018	0.014
Sep-18	0.047	0.034	0.018	0.0091

[a]The average is calculated using the rate constants at the regular hydraulic loading excluding Sep-18.

BP: Biopolymer; OC: Organic Carbon; ON: Organic Nitrogen;

NF: Non-Filtered; F: Filtered

Previous biofiltration studies reported that removals of non-volatile organic carbon, AOC, biodegradable DOC can be approximated using a first-order kinetics in a biological drinking water treatment process such as biofilters and GAC contactors (Huck and Anderson, 1992; Huck *et al*., 1994). In addition, the pseudo first-order kinetics was also used to model the removal of secondary substrates during biofiltration, including emerging contaminants (i.e. pharmaceuticals and insecticide, sand and anthracite media) (Hallé *et al*., 2015) as well as taste and odour compounds (i.e. geosmin and MIB, sand media) (Ho *et al*., 2007). Similarly, Andreozzi *et al*. (2006) employed pseudo first order kinetics to model the removal of aromatic compounds through aerobic biodegradation. This modeling approach is considered to be reasonable when the concentration of primary substrate (i.e. NOM) is considerably higher than that of the secondary substrates (Ho *et al*., 2007). In the current study, the first-order kinetics was also shown to be promising to model the removal of biodegradable NOM components (especially for organic carbon) during biofiltration. Also, the current study illustrated that the removal of biopolymers with larger MW and protein-like materials will also follow second order kinetics under certain circumstances, suggesting the presence of other preferred type of primary substrates (e.g. low MW biodegradable components) in the biofilter influent. Since first and second order kinetic models cannot be well distinguished from each other in this study and the first-order model has been widely reported to be promising for biofiltration kinetics (e.g. Huck and Anderson, 1992; Huck *et al*., 1994), the first-order rate constants were reported in Table 7.2.

7.3.4.2 Reaction Rate Constants

Table 7.2 summarizes the reaction rate constants estimated based on first-order kinetic models for various NOM components. The ranges of standard error for the first-order rate constant across 5 sampling events were 0.005-0.006, 0.003-0.007, 0.002-0.004, and 0.002-0.003 min^{-1} for biopolymer-OC, biopolymer-ON, protein-like materials in non-filtered samples, and protein-like materials in filtered samples, respectively. In general, these standard errors ranged from about 6% to 24% of the estimated values. Under 3.1 m/h operating conditions, first order reaction rate constants of biopolymer-OC and -ON were relatively higher compared to the rate constants under reduced hydraulic loading conditions. This might imply that all of the biodegradable components of biopolymers have been essentially removed at the higher hydraulic loading (i.e. shorter EBCT). Similar trends were observed for FEEM protein-like materials in filtered samples whereas this was not observed for FEEM protein-like materials for non-filtered samples. In addition, higher rate constants of biopolymer-OC generally corresponded to higher rate constants of biopolymer-ON. Higher rate constants of FEEM protein-like material removal in non-filtered samples did not necessarily correspond to higher rate constants of FEEM protein-like material removal in filtered samples.

Due to the low percent removal of LC-OCD humics compared to LC-OCD biopolymer-OC, it is expected that the first order reaction rate constants of LC-OCD humics were much smaller compared to the LC-OCD biopolymer-OC (e.g. Figure E1b). The same trends were observed between the FEEM protein- and humic-like materials (i.e. HA- and FA-like substances) for both filtered and non-filtered samples (e.g. Figure E1b). Also, each humics-related component, as characterized either by LC-OCD or FEEM, had higher first order reaction rate constants under the 3.1 m/h operating conditions compared to the reduced hydraulic loading conditions. In addition, higher rate constants of FEEM-HA and -FA in filtered samples generally corresponded to higher rate constants in non-filtered samples.

7.3.5 Comparison of FEEM Measurements of Non-filtered and Filtered Samples for Biofilter Monitoring and Kinetics Investigation

The effect of pre-filtration of water samples using 0.45 μm polymeric filter on the FEEM based NOM characterization results was also explored in this study. As shown in Figures 7.3a and 7.3b, FEEM protein peak intensities of non-filtered biofilter influent samples were more variable across the

different sampling events compared to the filtered biofilter influent samples. This is likely caused by the presence of colloidal/particulate protein-like materials in the non-filtered samples. In contrast, the patterns of FEEM-FA of biofilter influent samples from different sampling events were very similar between filtered and non-filtered samples (Figures 7.3c and 7.3d), indicating that the colloidal/particulate matter had little impact on the FEEM-FA in this study. The same pattern was observed for FEEM-HA.

Non-filtered water samples consistently have higher protein-like fluorescence peak intensities compared to the filtered samples. In addition, the intensity difference between filtered and non-filtered samples is more distinct for the upper ports (Figures 7.3a and 7.3b). This is attributed to the fact that the water samples collected from upper ports are less treated by the biofilter and hence may contain more colloidal/particulate protein-like material. For HA- and FA-like materials, the fluorescence peak intensity differences resulting from the 0.45 μm pre-filtration are relatively insignificant compared to protein-like substances. These observations are corroborated with the results of Baker *et al.* (2007) who attributed differences in filtration effects between protein- and humic-like materials to their different interactions with colloidal/particulate materials. Protein-like materials tend to be associated with large particles and hence are readily retained by the filter whereas humic-like materials are more likely to pass through the filter (Baker *et al.*, 2007). In addition, Peiris *et al*. (2012) illustrated that there are interactions between protein and colloidal/particulate matter at the physical level, providing the explanation for the fluorescence intensity differences between filtered and non-filtered water samples. The 0.45 μm pre-filtration will remove the protein-like materials that have interacted with colloidal/particulate matter, resulting in the reduction in the protein-like fluorescence of the filtered water samples.

As previously discussed, the FEEM-characterized NOM components in non-filtered and filtered water samples were shown to have different removal behaviour and kinetics along the biofilter. It is noted that the percent removal of FEEM protein peak intensities in non-filtered samples were generally higher than the filtered samples (Figures 7.4a and 7.4b). Comparatively, the percent removals of FEEM-HA and -FA intensities in non-filtered samples were consistently, but slightly lower, compared to the filtered samples (Figures 7.4c and 7.4d). Also, the differences between protein- and humic-like materials were observed regarding the correspondence of reaction rate

constants between filtered and non-filtered samples (Table 7.2). These differences in removal behaviour and kinetics between filtered and non-filtered samples are believed to be caused by the presence of colloidal/particulate materials which have been removed by 0.45 μm pre-filtration. This illustrates the value of conducting FEEM measurements on both filtered and non-filtered water samples. This is especially important when colloidal/particulate NOM components play an important role in the treatment process of interest.

7.4 Conclusions

In this study, various NOM components in water samples from six different sampling ports of a pilot-scale biofilter were characterized using LC-OCD and FEEM-based techniques. The removal behaviour and kinetics of these NOM components were investigated. The following conclusions can be drawn:

- Biofilters can effectively remove protein-like materials and biopolymers with large molecular weight (MW) but have limited performance for the removal of lower MW NOM components such as humic substances, building blocks, low MW acids and neutrals.
- LC-OCD biopolymer and FEEM protein-like materials (in both non-filtered and 0.45 μm filtered samples) were shown to have different removal behaviours and kinetics during biofiltration. This demonstrates the value of performing NOM characterization using both techniques. Each technique captures different components of the NOM revealing their different behaviours during biofiltration, which would not otherwise be identified if only one technique were used.
- Since NOM components could adsorb onto colloidal/particulate matter, the removal behaviours and kinetics of various NOM components characterized by the FEEM based technique were demonstrated to be different between non-filtered and filtered samples (0.45 μm filters), illustrating the value of conducting FEEM measurements on both non-filtered and filtered samples.
- NOM components with higher percent removal, including LC-OCD biopolymers and FEEM protein-like materials, were observed to follow either first or second order kinetics whereas the three kinetic models were indistinguishable for NOM components with low percent removal.

Chapter 8
Conclusions and Recommendations

8.1 Conclusions

The major objective of this research was to identify and develop NOM-C/P characterization approaches to provide more insight into the role of NOM-C/P components in ultrafiltration (UF) membrane fouling as well as their physiochemical properties and to investigate the UF fouling mitigation mechanisms of biofiltration as a pre-treatment alternative. Initially, FEEM-PCA and LC-OCD were shown to be promising approaches for NOM-C/P fouling behaviour investigations and the assessment of the fouling potential of complex water matrices (Chapter 3). The importance of biopolymers, protein-like substances and colloidal/particulate matter in UF fouling was also shown in this chapter. Unlike FEEM-PARAFAC which can extract the information from the underlying pure spectra of FEEMs, the value FEEM-PCA was always underestimated due to its relatively complex physical meaning and hence worse interpretability of the extracted principal components. In Chapter 4, to better understand the FEEM-PCA approach and the physical significance of the extracted principal components, advantage was taken of a model compound study that had been acquiring FEEMs using humic acid, bovine serum albumin, sodium alginate and silica as surrogates for the NOM-C/P components. The results from this study were analyzed using FEEM-PCA. Then, in order to further explore the differences in nature and utility between FEEM-PCA and other FEEM data analysis methods such as peak picking and FEEM-PARAFAC, a comparison of various approaches used in the literature for the analysis of FEEM results was conducted in Chapter 5. The recommendations on the selection of FEEM data analysis under different circumstances were provided. In addition, the impact of photomultiplier tube voltage (PMT) on the performance of FEEM-PCA and peak picking was also explored. This is especially crucial for FEEM-PCA as this is the only available method which can characterize fluorophores and colloidal/particulate matter using Rayleigh scattering at the same time. The purpose of Chapter 6 was to gain a better understanding of the composition and characteristics of biopolymers which are defined on the size exclusion basis and act as the key UF foulant. A LC-OCD based fractionation approach combined with NOM characterization techniques was developed to provide a more comprehensive understanding of biopolymers (defined by LC-OCD) from different sources. Based on the improved understanding of NOM-C/P characterization approaches and the characteristics of NOM-C/P components from previous studies, in Chapter 7, a pilot-scale biofiltration investigation on the behaviour and kinetics of

the removal of these key UF foulants (i.e. biopolymers and protein-like substances) as well as other NOM-C/P components was conducted to explore the mitigation mechanisms of biofiltration as a pre-treatment process of ultrafiltration.

The major conclusions drawn from this dissertation are summarized as follows:

8.1.1 Behaviour of Various NOM-C/P Components during Ultrafiltration Fouling

In the pilot-scale ultrafiltration fouling study, biopolymers and protein-like substances, were found to be key contributors to both hydraulically reversible and irreversible fouling. Comparatively, humic-like substances with lower MW were reported to have little contribution to UF membrane fouling when fouling rates were at a low level. It was also found that the presence of colloidal/particulate matter aided in the reduction of hydraulically irreversible fouling. LC-OCD and FEEM-PCA were shown to be promising techniques regarding NOM characterization. Due to the crucial role of various NOM-C/P components in UF fouling and their complex physicochemical properties, this study highlights the importance of understanding the existing techniques and developing new techniques to better explore the nature of these NOM-C/P components with respect to UF fouling.

8.1.2 Understanding of FEEM based NOM Characterization Technique

8.1.2.1 Physical Significance of Extracted PCs from FEEM-PCA Analysis of Model Solutions

PC1, PC2, and PC3 were attributed to humic-like substances, inorganic colloids and protein-like substances, and protein-like substances, respectively. The most important contributor to PC2 was determined to be inorganic colloids. Each PC reflected the fluorescence properties of a certain NOM/inorganic colloid component and the fluorescent property change of this particular component due to its interactions with other NOM components and inorganic colloid. This illustrated the major difference of the physical significance of the PCA components compared to PARAFAC components which are related to the pure spectra of FEEMs.

8.1.2.2 Comparison of FEEM Data Analysis Methods

FEEM-PCA and peak picking were shown to perform similarly with respect to content estimation of the humic-like substances whereas peak picking was capable of providing additional information on the composition of humic-like substances. Regarding the characterization of colloidal/particulate

matter using Rayleigh scattering, FEEM-PCA was shown to be superior compared to peak picking and FEEM-PARAFAC. For protein-like materials, the principal components extracted from FEEM-PCA were complex and were characterized by information on protein-like components and interactions, which could be beneficial for monitoring the overall effect posed by protein-like materials on treatment processes (e.g. membrane fouling). Comparatively, peak picking and FEEM-PARAFAC would be more appropriate for studies where the focus is on the nature of protein-like materials. FEEM-PCA is not impacted by instrumentation settings (i.e. photomultiplier tube voltage) and there is no need to separately optimize the settings for the fluorophores and Rayleigh scattering regions, that is, saturation of this region is tolerable.

8.1.2.3 FEEM Measurements of Filtered and Non-filtered Samples

The comparison of FEEM measurements of filtered and non-filtered samples illustrated that 0.45 μm pre-filtration with a polymeric filter resulted in the reduction of protein-like fluorescence peak intensities, reflecting the interaction between protein-like materials and colloidal/particulate matter. The effect of 0.45 μm pre-filtration was shown to be less important for HA- and FA-like materials. In addition, the behaviour and kinetics of removal of various NOM components were found to be different between filtered and non-filtered water samples. FEEM monitoring using non-filtered samples was shown to be beneficial for extracting additional information on the characteristics of the colloidal/particulate NOM components and their interactions with dissolved components. This is especially important when the information contained in the Rayleigh scattering (e.g. colloidal/particulate matter) region would be of interest to the study.

8.1.3 Complementary Nature between LC-OCD and FEEM as NOM Characterization Techniques

LC-OCD is intrinsically a size exclusion based chromatography technique which can provide information regarding the molecular weight (MW) distribution of various NOM components whereas FEEM is only suitable for extracting information regarding the composition related to functional groups of fluorescent NOM-C/P components. Characterizing NOM-C/P components using both techniques provided a more comprehensive understanding of their properties from different aspects including molecular size and functional group properties. The benefits of this complementary nature were especially well demonstrated using the LC-OCD based fractionation approach combined with FEEM based technique.

8.1.4 Development of a LC-OCD Based Fractionation Approach Combined with NOM Characterization Techniques for Investigation of the Compositional Characteristics of NOM Biopolymers

An LC-OCD-based fractionation approach combined with a NOM characterization technique was successfully developed and used to characterize the composition of NOM biopolymers from different sources. The normalized FEEM protein intensities, that is protein content per biopolymer carbon mass determined by LC-OCD, were demonstrated to be higher for the biopolymer-sized fractions with lower molecular weights. The normalized organic nitrogen (ON) concentration, that is the ON concentration per biopolymer carbon mass, was found to be a useful indicator for distinguishing different biopolymer sources.

8.1.5 Behaviour and Kinetics of NOM-C/P Removal during Biofiltration

Protein-like materials and biopolymers with high MW (identified as key UF foulants) were effectively removed by biofiltration, whereas lower MW NOM components (e.g. humic substances and building blocks) were more recalcitrant to biofiltration. Because of the low percentage removal, the performance of three kinetic models was not distinguishable for lower MW NOM components. For biopolymers and protein-like materials with higher percentage removal, the kinetics of their removal during biofiltration could follow either first or second order kinetics.

8.2 Recommendations

Based on the findings from this research, several recommendations are made for future research related to NOM characterization and drinking water treatment as well as industrial applications:

- The characteristics of other fluorescent NOM fractions (e.g. humics) should be further explored using the newly developed LC-OCD based fractionation technique.
- Characterization and monitoring of NOM-C/P components should be conducted using techniques of a complementary nature (e.g. LC-OCD and FEEM) which will provide more insight into NOM-C/P properties and behaviour in natural systems as well as drinking water treatment processes.
- Both intrinsic fluorescent properties of dominant inorganic colloids/NOM components and the effect of interactions among inorganic colloids and NOM component on the fluorescent properties should be considered when interpreting FEEM-PCA results.

- Considering the presence of colloidal/particulate NOM fraction (e.g. protein-like material) and interactions of NOM components with colloidal/particulate matter, performing FEEM measurements on both filtered and non-filtered samples is useful for ensuring more reliable characterization and measurement of NOM-C/P components. FEEM measurement on non-filtered samples is recommended for the situation when the colloidal/particulate materials play an important role in treatment processes.
- To achieve a more definitive conclusion regarding the kinetics of the removal of various NOM components during biofiltration, more comprehensive and detailed investigation should be undertaken. For example, more sampling ports could be configured and used along the media depth.
- Due to the excellent performance on the control of key UF foulants in Grand River water (GRW), biofiltration can be recommended as an effective and promising pre-treatment alternative to the mitigation of ultrafiltration fouling for source waters with similar NOM-C/P characteristics as GRW.

References

Ahmad, R., Amirtharajah, A., 1995. Detachment of biological and nonbiological particles from biological filters during backwashing. Proceedings of AWWA Annual Conference, Anaheim, California.

Aiken, G.R., Hsu-Kim, H., Ryan, J.N., 2011. Influence of dissolved organic matter on the environmental fate of metals nanoparticles, and colloids. Environmental Science and Technology 45(8), 3196 – 3201.

Aiken, G.R., McKnight, D.M., Thorn, K.A., Thurman, E.M., 1992. Isolation of hydrophilic organic acids from water using nonionic macroporous resins. Organic Geochemistry 18(4), 567 – 573.

Al-Sid-Cheikh, M., Pédrot, M., Dia, A., Guenet, H., Vantelon, D., Davranche, M., Gruau, G., Delhaye, T., 2015. Interactions between natural organic matter, sulfur, arsenic and iron oxides in re-oxidation compounds within riparian wetlands: NanoSIMS and X-ray adsorption spectroscopy evidences. Science of the Total Environment 515-516, 118 – 128.

Amy, G., 2008. Fundamental understanding of organic matter fouling of membranes. Desalination 231(1-3), 44 – 51.

Andrade-Eiroa, Á., Canle, M., Cerdá, V., 2013. Environmental applications of excitation-emission spectrofluorimetry: an in-depth Review I. Applied Spectroscopy Reviews 48, 1 – 49.

Andreozzi, R., Cesaro, R., Marotta, R., Pirozzi, F., 2006. Evaluation of biodegradation kinetic constants for aromatic compounds by means of aerobic batch experiments. Chemosphere 62(9), 1431 – 1436.

Aoustin, E., Schäfer, A.I., Fane, A.G., Waite, T.D., 2001. Ultrafiltration of natural organic matter. Separation and Purification Technology 22-23, 63 – 78.

Assemi, S., Newcombe, G., Hepplewhite, C., Beckett, R., 2004. Characterization of natural organic matter fractions separated by ultrafiltration using flow field-flow fractionation. Water Research 38(6), 1467 – 1476.

Bae, W. and Rittmann, B.E., 1996. A structured model of dual-limitation kinetics. Biotechnology and Bioengineering 49(6), 683 – 689.

Baghoth, S.A., Maeng, S.K., Salinas Rodr ǵuez, S.G., Ronteltap, M., Sharma, S., Kennedy, M., Amy, G.L., 2008. An urban water cycle perspective of natural organic matter (NOM): NOM in drinking water, wastewater effluent, storm water, and seawater. Water Science and Technology: Water Supply 8(6), 701 – 707.

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

Bailey, J.E., Ollis, D.F., 1986. Biochemical Engineering Fundamentals. 2nd Edition, McGraw-Hill, New York, p. 572.

Baker, A., 2001. Fluorescence excitation-emission matrix characterization of some sewage-impacted rivers. Environmental Science and Technology 25(5), 948 – 953.

Baker, A., Elliott, S., Lead, J.R., 2007. Effects of filtration and pH perturbation on freshwater organic matter fluorescence. Chemosphere 67(10), 2035 – 2043.

Basu, O.D., Huck, P.M., 2004. Removal of humic acid by biofiltration. Water Science and Technology: Water Supply 4(4), 147 – 154.

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

Boehme, J., Coble, P., Conmy, R., Stovall-Leonard, A., 2004. Examining CDOM fluorescence variability using principal component analysis: seasonal and regional modeling of three-dimensional fluorescence in the Gulf of Mexico. Marine Chemistry 89(1-4), 3 – 14.

Borcard, D., Gillet, F., Legendre, P., 2011. Numerical Ecology with R. Springer, New York, pp. 53 – 197.

Bos, M.A., Shervani, Z., Anusiem, A.C.I., Giesbers, M., Norde, W., Kleijn, M., 1994. Influence of the electric potential of the interface on the adsorption of proteins. Colloids and Surfaces B: Biointerfaces 3(1-2), 91 – 100.

Bro, R., 1997. PARAFAC. Tutorial and applications. Chemometrices and Intelligent Laboratory Systems 38(2), 149 – 171.

Bro, R., 2006. Review on multiway analysis in chemistry – 2000-2005. Critical Reviews in Analytical Chemistry 36(3-4), 279 – 293.

Broo, A.E., Berghult, B., Hedberg, T., 1999. Drinking water distribution – the effect of natural organic matter (NOM) on the corrosion of iron and copper. Water Science and Technology 40(9), 17 – 24.

Buffle, J., Wilkinson, K.J., Stoll, S., Filella, M., Zhang, J., 1998. A generalized description of aquatic colloidal interactions: the three-colloidal component approach. Environmental Science and Technology 32(19), 2887 – 2899.

Cabaniss, S.E., Shuman, M.S., 1988. Copper binding by dissolved organic matter: I. Suwannee river fulvic acid equilibria. Geochim Cosmochim Acta 52(1), 185 – 193.

Carroll, T., King, S., Gray, S.R., Bolto, B.A., Booker, N.A., 2000. The fouling of microfiltration membranes by NOM after coagulation treatment. Water Research 34(11), 2861 – 2868.

Chaufer, B., Rabiller-Baudry, M., Bouguen, A., Labbé, J.P., Quémerais, A., 2000. Spectroscopic characterization of zirconia coated by polymers with amine groups. Langmuir 16(4), 1852 – 1860.

Chen, F., Peldszus, S., Peiris, R.H., Ruhl, A.S., Mehrez, R., Jekel, M., Legge, R.L., Huck, P.M., 2014. Pilot-scale investigation of ultrafiltration membrane fouling rates using advanced data analysis techniques. Water Research 48, 508 – 518.

Chen, J., Gu, B., LeBoeuf, E.J., Pan, H., Dai, S., 2002. Spectroscopic characterization of the structural and functional properties of natural organic matter fractions. Chemosphere 48(1), 59 – 68.

Chen, W., Westerhoff, P., Leenheer, J.A., Booksh, K., 2003a. Fluorescence excitation-emission matrix regional integration to quantify spectra for dissolved organic matter. Environmental Science and Technology 37(24), 5701 – 4710.

Chen, J., LeBoeuf, E.J., Dai, S., Gu, B., 2003b. Fluorescence spectroscopic studies of natural organic matter fractions. Chemosphere 50 (5), 639 – 647.

Chen, B., Westerhoff, P., Zhang, L., Zhu, A., Yang, X., Wang, C., 2015. Application of pretreatment methods for reliable dissolved organic nitrogen analysis in water – a review. Critical Reviews in Environmental Science and Technology 45(3), 249 – 276.

Chiang, B.H., Cheryan, M., 1986. Ultrafiltration of skimmilk in hollow fibers. Journal of Food Science 51(2), 340 – 344.

Christy, A., Bruchet, A., Rybacki, D., 1999. Characterization of natural organic matter by pyrolysis/GC-MS. Environment International 25(2-3), 181 – 189.

Clark, S.R., Billsten, P., Elwing, H., 1994. A fluorescence technique for investigating protein adsorption phenomena at a colloidal silica surface. Colloids and Surfaces B: Biointefaces 2(5), 457 – 461.

Coble, P.G., Green, S.A., Blough, N.V., Gagosian, R.B., 1990. Characterization of dissolved organic matter in the Black Sea by fluorescence spectroscopy. Nature 348(6300), 432 – 435.

Collins, M.R., Amy, G.L., Steellnk, C., 1986. Molecular weight distribution, carboxylic acidity, and humic substances content of aquatic organic matter: implications for removal during water treatment. Environmental Science and Technology 20(10), 1028 – 1032.

DOC-Labor Dr. Huber, 2010. LC-OCD-OND Manual Software Model 8. Version 2010-08-12. pp. 1 – 65.

Egeberg, P.K., Christy, A.A., Eikenes, M., 2002. The molecular size of natural organic matter (NOM) determined by diffusivimetry and seven other methods. Water Research 36(4), 925 – 932.

El-Hadidy, A.M., Van Dyke, M.I., Chen, F., Peldszus, S., Huck, P.M., Improved approach for the characterization of attached biomass and extracellular polymeric substances on granular media from a drinking water biological active filter. In preparation.

Elshereef, R., Budman, H., Moresoli, C., Legge, R.L. 2010. Probing protein colloidal behavior in membrane-based separation processes using spectrofluorometric Rayleigh scattering data. Biotechnology Progress 26(3), 772 – 780.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., Wold, S., 2001. Multi- and Megavariate Data Analysis, Principles and Applications. Umetrics Academy, Umea, Sweden. p. 533.

Fairhurst, A.J., Warwick, P., Richardson, S., 1995. The influence of humic acid on the adsorption of europium onto inorganic colloids as a function of pH. Colloids and Surfaces A: Physicochemical and Engineering Aspects 99(2-3), 187 – 199.

Filella, M., 2014. Understanding what we are measuring: Standards and quantification of natural organic matter. Water Research 50, 287 – 293.

Filloux, E., Gallard, H., Croué, J.P., 2012. Identification of effluent organic matter fractions responsible for low-pressure membrane fouling. Water Research 46(17), 5531 – 5540.

Flemming H.C., Wingender, J., 2010. The biofilm matrix. Nature Reviews Microbiology 8, 623 – 633.

Frølund, B., Palmgren, R., Keiding, K., Nielsen, P.H., 1996. Extraction of extracellular polymers from activated sludge using a cation exchange resin. Water Research 30(8), 1749 – 1758.

Gao, H., Zepp, R.G., 1998. Factors influencing photoreactions of dissolved organic matter in a coastal river of the southeastern United States. Environmental Science and Technology 32(19), 2940 – 2946.

Haberkamp, J., Ernst, M., Paar, H., Pallischeck, D., Amy, G., Jekel, M., 2011. Impact of organic fractions identified by SEC and fluorescence EEM on the hydraulic reversibility of ultrafiltration membrane fouling by secondary effluents. Desalination and Water Treatment 29 (1-3), 73 – 86.

Halim, M., Spaccini, R., Parlanti, E., Amezghal, A., Piccolo, A., 2013. Differences in fluorescence properties between humic acid and its size fractions separated by preparative HPSEC. Journal of Geochemical Exploration 129, 23 – 27.

Hallé, C., 2009. Biofiltration in Drinking Water Treatment: Reduction of Membrane Fouling and Biodegradation of Organic Trace Contaminants. PhD thesis. Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, Ontario, Canada.

Hallé, C., Huck, P.M., Peldszus, S., 2015. Emerging contaminant removal by biofiltration: temperature, concentration and EBCT impacts. Journal of American Water Works Association 107(7), E364 – E379.

Hallé, C., Huck, P.M., Peldszus, S., Haberkamp, J., Jekel, M., 2009. Assessing the performance of biological filtration as pretreatment to low pressure membranes for drinking water. Environmental Science and Technology 43(10), 3878 – 3884.

Ham, B.M., MaHam, A., 2016. Analytical Chemistry: a Chemist and Laboratory Technician's Toolkit. John Wiley and Sons, Inc., Hoboken, New Jersey, pp. 247 – 248.

He, W., Hur, J., 2015. Conservative behavior of fluorescence EEM-PARAFAC components in resin fractionation processes and its applicability for characterizing dissolved organic matter. Water Research 83, 217 – 226.

Henderson, R.K., Baker, A., Murphy, K.R., Hambly, A., Stuetz, R.M., Khan, S.J., 2009. Fluorescence as a potential monitoring tool for recycled water systems: A review. Water Research 43(4), 863 – 881.

Her, N., Amy, G., McKnight, D., Sohn, J. Yoon, Y., 2003. Characterization of DOM as a function of MW by fluorescence EEM and HPLC-SEC using UVA, DOC, and fluorescence detection. Water Research 37 (17), 4295 – 4303.

Hermia, J., 1982. Constant pressure blocking filtration laws – application to power-law non-Newtonian fluids. Transactions of the Institution of Chemical Engineers 60(3), 183 – 187.

Her, N., Amy, G., Sohn, J., von Gunten, U., 2008. UV absorbance ratio index with size exclusion chromatography (URI-SEC) as an NOM property indicator. Journal of Water Supply: Research and Technology-AQUA 57(1), 35 – 44.

Hlady, V., Reinecke, D.R., Andrade, J.D., 1986. Fluorescence of adsorbed protein layers I. quantification of total internal reflection fluorescence. Journal of Colloid and Interface Science 111(2), 555 – 569.

Ho, L., Hoefel, D., Bock, F., Saint, C.P., 2007. Biodegradation rates of 2-methylisoborneol (MIB) and geosmin through sand filters and in bioreactors. Chemosphere 66(11), 2210 – 2218.

Hoefel, D., Ho, L., Monis, P.T., Newcombe, G., Saint, C.P., 2009. Biodegradation of geosmin by a novel Gram-negative bacterium: isolation, phylogenetic characterization and degradation rate determination. Water Research 43(11), 2927 – 2935.

Howe, K.J., Ishida, K.P., Clark, M.M., 2002. Use of ATR/FTIR spectrometry to study fouling of microfiltration membranes by natural waters. Desalination 147(1-3), 251 – 255.

Hozalski, R.M., Bouwer, E.J., 2001. Non-steady state simulation of BOM removal in drinking water biofilters: model development. Water Research 35(1), 198 – 210.

Hozalski, R.M., Bouwer, E.J., Goel, S., 1999. Removal of natural organic matter (NOM) from drinking water supplies by ozone-biofiltration. Water Science and Technology 40(9), 157 – 163.

Hozalski, R.M., Goel, S., Bouwer, E.J., 1995. TOC removal in biological filters. Journal of the American Water Works Association 87(12), 40 – 54.

Huang, H., Lee, N., Young, T., Gary, A., Lozier, J.C., Jacangelo, J.G., 2007. Natural organic matter fouling of low-pressure, hollow-fiber membranes: Effects of NOM source and hydrodynamic conditions. Water Research 41(17), 3823 – 3832.

Huang, H., Young, T.A., Jacangelo, J.G., 2008. Unified membrane fouling index for low pressure membrane filtration of natural waters: principles and methodology. Environmental Science and Technology 42(3), 714 – 720.

Huang, H., Young, T., Jacangelo, J.G., 2009. Novel approach for the analysis of bench-scale, low pressure membrane fouling in water treatment. Journal of Membrane Science 334 (1-2), 1 – 8.

Huber, S.A., Balz, A., Abert, M., Pronk, W., 2011. Characterisation of aquatic humic and non-humic matter with size-exclusion chromatography – organic carbon detection – organic nitrogen detection (LC-OCD-OND). Water Research 45(2), 879 – 885.

Huck, P.M., Anderson, W.B., 1992. Quantitative relationships between the removal of NVOC, chlorine demand and AOX formation potential in biological drinking water treatment. VOM WASSER, 78, 281 – 303.

Huck, P.M., Coffey, B.M., Amirtharajah A., Bouwer, E.J. 2000. Optimizing Filtration in Biological Filters. AWWA Research Foundation and American Water Works Association, Report No. 90793. Denver, Colorado, USA.

Huck, P.M., Sozański, M.M., 2008. Biological filtration for membrane pre-treatment and other applications: towards the development of a practically-oriented performance parameter. Journal of Water Supply: Research and Technology – AQUA 57(4), 203 – 224.

Huck, P.M., Zhang, S., Price, M.L., 1994. BOM removal during biological treatment: a first-order model. Journal of American Water Works Association 86(6), 61 – 71.

Hudson, N., Baker, A., Ward, D., Reynolds, D.M., Brunsdon, C., Carliell-Marquet, C., Browning, S., 2008. Can fluorescence spectrometry be used as a surrogate for the biochemical oxygen demand (BOD) test in water quality assessment? An example from South West England. Science of the Total Environment 391(1), 149 – 158.

Huisman, I.H., Prádanos, P., Hernández, A., 2000. The effect of protein-protein and protein-membrane interactions on membrane fouling in ultrafiltration. Journal of Membrane Science 179(1-2), 79 – 90.

Hunt, P., Parry, J.D., Hamilton-Taylor J., 2000. Further evidence of elemental composition as an indicator of the bioavailability of humic substances to bacteria. Limnology and Oceanography 45(1), 237 – 241.

Ishii, S.K., Boyer, T.H., 2012. Behavior of reoccurring PARAFAC components in fluorescent dissolved organic matter in antural and engineered systems: a critical review. Environmental Science and Technology 46(4), 2006 – 2017.

Jacangelo, J.G., Demarco, J., Owen, D.M., Randtke, S.J., 1995. Selected processes for removing NOM: an overview. Journal of American Water Works Association 87(1), 64 – 77.

Jarusutthirak, C., Amy, G., 2006. Role of soluble microbial products (SMP) in membrane fouling and flux decline. Environmental Science and Technology 40(3), 969 – 974.

Jekel, M.R., 1986. The stabilization of dispersed mineral particles by adsorption of humic substances. Water Research 20(2), 1543 – 1554.

Jermann, D., Pronk, W., Meylan, S., Boller, M., 2007. Interplay of different NOM fouling mechanisms during ultrafiltration for drinking water production. Water Research 41(8), 1713 – 1722.

Jermann, D., Pronk, W., Kägi, R., Halbeisen, M., Boller, M., 2008. Influence of interactions between NOM and particles on UF fouling mechanisms. Water Research 42(14), 3870 – 3878.

Jones, K.L., O'Melia, C.R., 2001. Ultrafiltration of protein and humic substances: effect of solution chemistry on fouling and flux decline. Journal of Membrane Science 193(2), 163 – 173.

Katsoufidou, K., Yiantsios, S.G., Karabelas, A.J., 2007. Experimental study of ultrafiltration membrane fouling by sodium alginate and flux recovery by backwashing. Journal of Membrane Science 300(1-2), 137 – 146.

Kennedy, A.M., Summers, R.S., 2015. Effect of DOM size on organic micropollutant adsorption by GAC. Environmental Science and Technology 49(11), 6617 – 6624.

Kimura, K., Hane, Y., Watanabe, Y., Amy, G., Ohkuma, N., 2004. Irreversible membrane fouling during ultrafiltration of surface water. Water Research 38(14-15), 3431 – 3441.

Kitis, M., Karanfil, T., Wigton, A., Kilduff, J.E., 2002. Probing reactivity of dissolved organic matter for disinfection by-product formation using XAD-8 resin adsorption and ultrafiltration fractionation. Water Research 36(15), 3834 – 3848.

Kowalczuk, P., Cooper, W.J., Durako, M.J., Kahn, A.E., Gonsior, M., Young, H., 2010. Characterization of dissolved organic matter fluorescence in the South Atlantic Bight with use of

PARAFAC model: Relationships between fluorescence and its components, absorption coefficients and organic carbon concentrations. Marine Chemistry 118(1-2), 22 – 36.

Krasner, S.W., Sclimenti, M.J., Coffey, B.M., 1993. Testing biologically active filters for removing aldehydes formed during ozonation. Journal of the American Water Works Association 85(5), 62 – 71.

Krause, S., Obermayer, A., 2011. Predictability of fouling-potential of raw water for ultrafiltration membranes. Water Science and Technology: Water Supply 11(4), 481 – 489.

Lakowicz, J.R., 2006. Principles of Fluorescence Spectroscopy. 4th Edition. Springer. pp. 1 – 89.

Lam, B., Simpson, A.J., 2008. Direct ^{1}H NMR spectroscopy of dissolved organic matter in natural waters. Analyst 133(2), 263 – 269.

Larsericsdotter, H., Oscarsson, S., Buijs, J., 2005. Structure, stability, and orientation of BSA adsorbed to silica. Journal of Colloid and Interface Science 289(1), 26 – 35.

Lavonen, E.E., Kothawala, D.N., Tranvik, L.J., Gonsior, M., Schmitt-Kopplin, P., Köhler, S.J., 2015. Tracking changes in the optical properties and molecular composition of dissolved organic matter during drinking water production. Water Resarch 85, 286 – 294.

Lee, N, Amy, G., Croué, J.P., Buisson, H., 2004. Identification and understanding of fouling in low-pressure membrane (MF/UF) filtration by natural organic matter (NOM). Water Research 38(20), 4511 – 4523.

Lee, N., Amy, G., Croué, J.P., 2006. Low-pressure membrane (MF/UF) fouling associated with allochthonous versus autochthonous natural organic matter. Water Research 40(12), 2357 – 2368.

Leenheer, J.A., 1981. Comprehensive approach to preparative isolation and fractionation of dissolved organic carbon from natural waters and wastewaters. Environmental Science and Technology 15(5), 578 – 587.

Leenheer, J.A., Croué, J.P., 2003. Characterizing dissolved aquatic organic matter. Environmental Science and Technology 37(1), 18A-26A.

Legendre, P., Legendre, P., 1998. Numerical Ecology. 2nd Edition. Elsevier Science, Amsterdam. pp. 1 – 626.

Liang, L., Luo, L., Zhang, S., 2011. Adsorption and desorption of humic and fulvic acids on SiO_2 particles at nano- and micro-scales. Colloids and Surfaces A: Physicochemical and Engineering Aspects 384(1-3), 126 – 130.

Liu, T., Chen, Z., Yu, W., You, S., 2011. Characterization of organic membrane foulants in a submerged membrane bioreactor with pre-ozonation using three-dimensional excitation-emission matrix fluorescence spectroscopy. Water Research 45(5), 2111 – 2121.

Malinowsky, E.R., 1991. Factor Analysis in Chemistry. Wiley, New York. p. 350.

Matilainen, A., Vepsäläinen, M., Sillanpää, M., 2010. Natural organic matter removal by coagulation during drinking water treatment: A review. Advances in Colloid and Interface Science 159(2), 189 – 197.

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

Maurice, P.A., Pullin, M.J., Cabaniss, S.E., Zhou, Q., Namjesnik-Dejanovic, K., Aiken, G.R., 2002. A comparison of surface water natural organic matter in raw filtered water samples, XAD, and reverse osmosis isolates. Water Research 36(9), 2357 – 2371.

Mazzei, P., Piccolo, A., 2015. Interactions between natural organic matter and organic pollutants as revealed by NMR spectroscopy. Magnetic Resonance in Chemsitry 53(9), 667 – 678.

McDowall, B., Hoefel, D., Newcombe, G., Saint, C.P., Ho, L., 2009. Enhancing the biofiltration of geosmin by seeding sand filter columns with a consortium of geosmin-degrading bacteria. Water Research 43(2), 433 – 440.

McKie, M.J., Tylor-Edmonds, L., Andrews, S.A., Andrews, R.C., 2015. Engineered biofiltration for the removal of disinfection by-product precursors and genotoxicity. Water Research 81, 196 – 207.

Meng, F., Zhou, Z., Ni, B.J., Zheng, X., Huang, G., Jia, X., Li, S., Xiong, Y., Kraume, M., 2011. Characterization of the size-fractionated biomacromolecules: Tracking their role and fate in a membrane bioreactor. Water Research 45(15), 4661 – 4671.

Metz, D.H., Pohlman, R.C., Vogt, J., Summers, R.S., 2006. Removal of MIB and geosmin by full scale biological sand filters. In: Gimbel, R., Graham, N.J.D., Collins, M.R. (Eds.), Recent Progress in Slow Sand and Alternative Biofiltration Processes. IWA Publishing, London, UK, pp. 352 – 359.

Murphy, E.M., Zachara, J.M., Smith, S.C., 1990. Influence of mineral-bound humic substances on the sorption of hydrophobic organic compounds. Environmental Science and Technology 24(10), 1507 – 1516.

Murphy, K.R., Hambly, A., Singh, S., Henderson, R.K., Baker, A., Stuetz, R., Khan, S.J., 2011. Organic matter fluorescence in municipal water recycling schemes: toward a unified PARAFAC model. Environmental Science and Technology 45(7), 2909 – 2916.

Myat, D.T., Stewart, M.B., Mergen, M., Zhao, O., Orbell, J.D., Gray, S., 2014. Experimental and computational investigations of the interactions between model organic compounds and subsequent membrane fouling. Water Research 48, 108 – 118.

Nakanishi, K., Sakiyama, T., Imamura, K., 2001. Review on the adsorption of proteins on solid surfaces, a common but very complicated phenomenon. Journal of Bioscience and Bioengineering 91(3), 233 – 244.

Nguyen, A.H., Tobiason, J.E., Howe, K.J., 2011. Fouling indices for low pressure hollow fiber membrane performance assessment. Water Research 45 (8), 2627 – 2637.

Niquette, P., Prévost, M., Merlet, N. & Lafrance, P., 1999 Influence de facteurs contrôlant l'enlèvement de la demande en chlore et de précurseurs de sous-produits de chloration dans les filtres biologiques. Water Research 33(10), 2329 – 2344.

Norde, W., Favier, J.P., 1992. Structure of adsorbed and desorbed protein. Colloids and Surfaces 64(1), 87 – 93.

Ohno, T., Cronan, C.S., 1997. Comparative effects of ionic- and nonionic-resin purification treatments on the chemistry of dissolved organic matter. International Journal of Environmental Analytical Chemistry 66(2), 119 – 136.

Ohno, T., 2002. Fluorescence inner-filtering correction for determining the humification index of dissolved organic matter. Environmental Science and Technology 36(4), 742 – 746.

Ohno, T., Amirbahman, A., Bro, R., 2008. Parallel factor analysis of excitation-emission matrix fluorescence spectra of water soluble soil organic matter as basis for the determination of conditional metal binding parameters. Environmental Science and Technology 42(1), 186 – 192.

Omrani, H., Barnes, J.A., Dudelzak, A.E., Loock, H., Waechter, H., 2012. Fluorescence excitation-emission matrix (EEM) spectroscopy and cavity ring-down (CRD) absorption spectroscopy of oil-contaminated jet fuel using fiber-optic probes. Analyst (137), 2782 – 2790.

Peiris, B. R.H., Hallé, C., Haberkamp, J., Legge, R.L., Peldszus, S., Moresoli, C., Budman, H., Amy, G., Jekel, M., and Huck, P.M., 2008. Assessing nanofiltration fouling in drinking water treatment using fluorescence fingerprinting and LC-OCD analyses. Water Science and Technology: Water Supply 8(4), 459 – 465.

Peiris, B.R.H., Budman, H., Moresoli, C., Legge, R.L. 2009. Acquiring reproducible fluorescence spectra for very low concentrations of dissolved organic matter. Water Science and Technology 60(6), 1385 – 1392.

Peiris, R.H., Hallé, C., Budman, H., Moresoli, C., Peldszus, S., Huck, P.M., Legge, R.L., 2010a. Identifying fouling events in a membrane-based drinking water treatment process using principal component analysis of fluorescence excitation-emission matrices. Water Research 44(1), 185 – 194.

Peiris, R.H., Budman, H., Moresoli, C., and Legge, R.L., 2010b. Understanding fouling behavior of ultrafiltration membrane processes and natural water using principal component analysis of fluorescence excitation-emission matrices. Journal of Membrane Science 357(1-2), 62 – 72.

Peiris, R.H., Budman, H., Legge, R.L., Moresoli, C., 2011. Assessing irreversible fouling behavior of membrane foulants in the ultrafiltration of natural water using principal component analysis of fluorescence excitation-emission matrices. Water Science and Technology: Water Supply 11(2), 179 – 185.

Peiris, R.H., Ignagni, N., Budman, H., Moresoli, C., Legge, R.L., 2012. Characterizing natural colloidal/particulate-protein interactions using fluorescence-based techniques and principal component analysis. Talanta 99, 457 – 453.

Peldszus, S., Benecke, J., Jekel, M., Huck, P.M., 2012. Direct biofiltration pre-treatment for fouling control of ultrafiltration membranes. Journal of American Water Association 104(7), 430 – 445.

Peldszus, S., Hallé, C., Peiris, R.H., Hamouda, M., Jin, X., Legge, R.L., Budman, H., Moresoli, C., Huck, P.M., 2011. Reversible and irreversible low-pressure membrane foulants in drinking water treatment: identification by principal component analysis of fluorescence EEM and mitigation by biofiltration pretreatment. Water Research 45(16), 5161 – 5170.

Peleato, N.M., Andrews, R.C., 2015a. Comparison of three-dimensional fluorescence analysis methods for predicting formation of trihalomethanes and haloacetic acids. Journal of Environmental Sciences, 159 – 167.

Peleato, N.M., Andrews, R.C., 2015b. Contributions of spatial, temporal, and treatment impacts on natural organic matter character using fluorescence-based measures. Water Science and Technology: Water Supply 15(3), 589 – 598.

Persson, T., Wedborg, M., 2001. Multivariate evaluation of the fluorescence of aquatic organic matter. Analytica Chimica Acta 434(2), 179 – 192.

Peter-Varbanets, M., Margot, J., Traber, J., Pronk, W., 2011. Mechanisms of membrane fouling during ultra-low pressure ultrafiltration. Journal of Membrane Science 377(1-2), 42 – 53.

Rabiller-Baudry, M., Le Maux, M., Chaufer, B., Begoin, L., 2002. Characterisation of cleaned and fouled membrane by ART-FTIR and EDX analysis coupled with SEM: application to UF of skimmed milk with a PES membrane. Desalination 146(1-3), 123 – 128.

Rahman, I., 2013. Direct Biofiltration and Nutrient (Phosphorus) Enhancement for Polymeric Ultrafiltration Membrane Fouling Control. M.A.Sc. Thesis. Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, Ontario, Canada.

Rahman, I., Ndiongue, S., Jin, X., Van Dyke, I., Anderson, W.B., Huck, P.M., 2014. Fouling of low-pressure membranes during drinking water treatment: effect of NOM components and biofiltration pretreatment. Water Science and Technology: Water Supply 14(3), 453 – 460.

Rawlings, J.O., Pantula, S.G., Dickey, D.A., 1998. Applied Regression Analysis: a Research Tool, 2nd Edition. Springer, New York, pp. 213 – 227.

Ryan, D.K., Weber, J.H., 1982. Fluorescence quenching titration for determination of complexing capacities and stability constants of fulvic acid. Analytical Chemistry 54(6), 986 – 990.

Sanchez, N.P., Skeriotis, A.T., Miller, C.M., 2014. A PARAFAC-based long-term assessment of DOM in a multi-coagulant drinking water treatment scheme. Environmental Science and Technology 48(3), 1582 – 1591.

Senesi, N., 1990 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 232, 77 – 106.

Shao, S., Liang, H., Qu, F., Yu, H., Li, K., Li, G., 2014. Fluorescent natural organic matter fractions responsible for ultrafiltration membrane fouling: Identification by adsorption pretreatment coupled with parallel factor analysis of excitation-emission matrices. Journal of Membrane Science 464, 33 – 42.

Shutova, Y., Baker, A., Bridgeman, J., Henderson, R.K., 2014. Spectroscopic characterisation of dissolved organic matter changes in drinking water treatment: from PARAFAC analysis to online monitoring wavelengths. Water Research 54, 159 – 169.

Siembida-Lösch, B., Anderson, W.B., Wang, Y., Bonsteel, J., Huck, P.M., 2015. Effect of ozone on biopolymers in biofiltration and ultrafiltration processes. Water Research 70, 224 – 234.

Sierra, M.M.D., Giovanela, M., Parlanti, E., Soriano-Sierra, E.J., 2005. Fluorescence fingerprint of fulvic and humic acids from varied origins as viewed by single-scan and excitation/emission matrix techniques. Chemosphere 58 (6), 715 – 733.

Song, H., Orr, O., Hong, Y., Karanfil, T., 2009. Isolation and fractionation of natural organic matter: evaluation of reverse osmosis performance and impact of fractionation parameters. Environmental Monitoring and Assessment 153(1-4), 307 – 321.

Spencer, R.G.M., Bolton, L., Baker, A., 2007. Freeze/thaw and pH effects on freshwater dissolved organic matter fluorescence and absorbance properties from a number of UK locations. Water Research 41(13), 2941 – 2950.

Standard Methods, 2005. Standard Methods for the Examination of Water and Wastewater, 21st Edition. American Public Health/American Water Works Association/Water Environment Federation, Washington, DC, USA.

Stedmon, C.A., Markager, S., Bro, R., 2003. Tracing dissolved organic matter in aquatic environments using a new approach to fluorescence spectroscopy. Marine Chemistry 82(3-4), 239 – 254.

Stolpe, B., Zhou, Z., Guo, L., Shiller, A., 2014. Colloidal size distribution of humic- and protein-like fluorescent organic matter in the northern Gulf of Mexico. Marine Chemistry 164, 25 – 37.

Sun, L., Perdue, E.M., Meyer, J.L., Weis, J., 1997. Use of elemental composition to predict the bioavailability of dissolved organic matter in a Georgia river. Limnology and Oceanography 42(4), 714 – 721.

Tang, C.Y., Chong, T.H., Fane, A.G., 2011. Colloidal interactions and fouling of NF and RO membranes: a review. Advances in Colloid and Interface Science 164(1-2), 126 – 143.

Tien, C., Ramarao, B.V., 2011. Revisiting the laws of filtration: An assessment of their use in identifying particle retention mechanisms in filtration. Journal of Membrane Science 383 (1-2), 17 – 25.

Tran, N.H., Ngo, H.H., Urase, T., Gin, K.Y.H., 2015. A critical review on characterization strategies of organic matter for wastewater and water treatment processes. Bioresource Technology 193, 523 – 533.

Tsai, B.N., Chang, C.H., Lee, D.J., 2008. Fractionation of soluble microbial products (SMP) and soluble extracellular polymeric substances (EPS) from wastewater sludge. Environmental Technology 29(10), 1127 – 1138.

Urfer, D., Huck, P.M., Booth, S.D.J., Coffey, B.M., 1997. Biological filtration for BOM and particle removal: a critical review. Journal of the American Water Works Association 89(12), 83 – 92.

Uyguner, C.S., Bekbolet, M., 2005. Evaluation of humic acid photocatalytic degradation by UV-vis and fluorescence spectroscopy. Catalysis Today 101(3-4), 267 – 274.

van de Ven, W.J.C., van't Sant, K., Pünt, I.G.M., Zwijnenburg, A., Kemperman, A.J.B., van der Meer, W.G.J., Wessling, M., 2008. Hollow fiber dead-end ultrafiltration: Influence of ionic environment on filtration of alginates. Journal of Membrane Science 308(1-2), 218 – 229.

Vilgé-Ritter, A., Masion, A., Boulangé, T., Rybacki, D., Bottero, J.Y., 1999. Removal of natural organic matter by coagulation-flocculation: a pyrolysis-GC-MS study. Environment Science and Technology 33(17), 3027 – 3032.

Wang, Z., Wu, Z., Tang, S., 2009. Characterization of dissolved organic matter in a submerged membrane bioreactor by using three-dimensional excitation and emission matrix fluorescence spectroscopy. Water Research 43(6), 1533 – 1540.

Wanner, O., Eberl, H.J., Morgenroth, E., Noguera, D.R., Picioreanu, C., Rittmann, B.E., van Loosdrecht, M.C.M., 2006. Mathematical Modeling of Biofilms. Report of the IWA Biofilm Modeling Task Group, Scientific and Technical Report No. 18, IWA Publishing, London.

Wassink, J.K., Andrews, R.C., Peiris, R.H., Legge, R.L., 2011. Evaluation of fluorescence excitation-emission and LC-OCD as methods of detecting removal of NOM and DBP precursors by enhanced coagulation. Water Science and Technology: Water Supply 11(5), 621 – 630.

Weishaar, J.L., Aiken, G.R., Bergamaschi, B.A., Fram, M.S., Fujii, R., Mopper, K., 2003. Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon. Environmental Science and Technology 37(20), 4702 – 4708.

Wilking, J.N., Angelini, T.E., Seminara, A., Brenner, M.P., Weitz, D.A., 2011. Biofilms as complex fluids. Materials Research Society Bulletin 36(5), 385 – 391.

Wilson, B., 2015. Impact of Biofilter Backwashing on the Biofiltration/ultrafiltration Process. M.A.Sc. Thesis. Department of Civil and Environmental Engineering, University of Waterloo, Waterloo, Ontario, Canada.

Wu, F.C., Mills, R.B., Evans, R.D., Dillon, P.J., 2004. Kinetics of metal-fulvic acid complexation using a stopped-flow technique and three-dimensional excitation emission fluorescence spectrophotometer. Analytical Chemistry 76(1), 110 – 113.

Xiao, K., Wang, X., Huang, X., Waite, T.D., Wen, X., 2009. Analysis of polysaccharide, protein and humic acid retention by microfiltration membranes using Thomas's dynamic adsorption model. Journal of Membrane Science 342(1-2), 22 – 34.

Yamamura, H., Okimoto, K., Kimura, K., Watamabe, Y., 2014. Hydrophilic fraction of natural organic matter causing irreversible fouling of microfiltration and ultrafiltration membranes. Water Research 54, 123 – 136.

Yeh, H.M., Cheng, T.W., 1993. Resistance-in-series for membrane ultrafiltration in hollow fibers of tube-and-shell arrangement. Separation Science and Technology 28(6), 1341 – 1355.

Yoon, Y., Lueptow, R.M., 2006. Concentration of colloidal silica suspensions using fluorescence spectroscopy. Colloids and Surfaces A: Physicochemical and Engineering Aspects 277(1-3), 107 – 110.

Zearley, T.L., Summers, R.S., 2012. Removal of trace organic micropollutants by drinking water biological filters. Environmental Science and Technology 46(17), 9412 – 9419.

Zepp, R.G., Sheldon, W.M., Moran, M.A., 2004. Dissolved organic fluorophores in southeastern US coastal waters: correction method for eliminating Rayleigh and Raman scattering peaks in excitation-emission matrices. Marine Chemistry 89(1-4), 15 – 36.

Zhang, S., 1996. Modeling Biological Drinking Water Treatment Processes. PhD thesis. Department of Civil Engineering, University of Alberta, Edmonton, Alberta, Canada.

Zhang, Y., Yin, Y., Feng, L., Zhu, G., Shi, Z., Liu, X., Zhang, Y., 2011. Characterizing chromophoric dissolved organic matter in Lake Tianmuhu and its catchment basin using excitation-emission matrix fluorescence and parallel factor analysis. Water Research 45(16), 5110 – 5122.

Zhao, B.Q., Huang, C.P., Chen, S.Y., Wang, D.S., Li, T., Qu, J.H., 2011. Fouling analysis of polysulfone ultrafiltration membranes used for drinking water treatment. Water Science and Technology: Water Supply 11(6), 668 – 674.

Zheng, D., Andrews, R.C., Andrews, S., Taylor-Edmonds, L., 2015. Effects of coagulation on the removal of natural organic matter, genotoxicity and precursors to halogenated furanones. Water Research 70, 118 – 129.

Zheng, X., Ernst, M., Huck, P.M., Jekel, M., 2010. Biopolymer fouling in dead-end ultrafiltration of treated domestic wastewater. Water Research 44(18), 5212 – 5221.

Zhou, Z., Guo, L., 2015. A critical evaluation of an asymmetrical flow field-flow fractionation system for colloidal size characterization of natural organic matter. Journal of Chromatography A 1399, 53 – 64.

Zularisam, A.W., Ismail, A.F., Salim, R., 2006. Behaviours of natural organic matter in membrane filtration for surface water treatment – a review. Desalination 194(1-3), 211 – 231.

Zularisam, A.W., Ismail, A.F., Salim, M.R., Sakinah, M., Hiroaki, O., 2007. Fabrication, fouling and foulant analyses of asymmetric polysulfone (PSF) ultrafiltration membrane fouled with natural organic matter (NOM) source waters. Journal of Membrane Science 299(1-2), 97 – 113.

Appendix A
Supplementary Data for Chapter 3 (i.e. fouling rates, fouling rate classification, fouling cycle plots, and loading plots of principal components (PCs))

Table A1 Summary of reversible and irreversible fouling rates for both the whole fouling cycle and 3-day duration.

Sample	*Whole Fouling Cycle*			*3-Day Duration*			*Category*
	Duration (day)	IRF Rate (kPa/day)	RF Rate (kPa/min)	Duration (day)	IRF Rate (kPa/day)	RF Rate (kPa/min)	F/L
B2	0.72	24.956	0.607	0.72	24.956	0.607	F
B3	0.39	35.607	0.592	0.39	35.607	0.592	F
A6	1.43	22.399	0.567	1.43	22.399	0.567	F
C9	0.44	22.658	0.646	0.44	22.658	0.646	L
C3	1.92	7.609	0.347	1.92	7.609	0.347	F
A2	1.14	5.943	0.388	1.14	5.943	0.388	F&L
A10	6.94	1.238	0.100	2.52	1.047	0.085	L
B5	7.01	1.003	0.075	2.44	0.588	0.061	L
C6	3.21	3.174	0.114	1.84	3.406	0.121	L
C7	2.84	3.986	0.131	2.14	4.139	0.134	F
C5	2.69	2.228	0.120	1.62	2.791	0.126	F
C8	1.69	3.160	0.152	1.30	2.919	0.152	F
C1	8.32	0.870	0.067	2.87	0.605	0.067	F&L
C2	9.87	0.391	0.102	2.50	0.775	0.086	F
B4	0.78	3.890	0.137	0.78	3.890	0.137	F
A1	4.79	3.430	0.179	2.59	3.343	0.138	F
B6	2.78	1.079	0.088	1.40	1.067	0.094	F
C4	2.84	0.573	0.072	2.24	0.651	0.071	F
R1	3.93	2.619	0.189	2.52	2.222	0.167	L
A7	5.66	0.469	0.074	2.39	0.495	0.079	L
C10	4.90	1.433	0.236	3.00	1.133	0.223	F
A3	1.81	4.529	0.217	1.20	3.740	0.199	F
A5	3.91	10.140	0.307	1.61	15.222	0.368	L
A4	3.91	10.140	0.307	2.30	6.998	0.263	F
B1	4.25	10.736	0.266	1.34	5.645	0.155	F&L
A8	7.06	3.179	0.287	1.62	3.606	0.177	L
A9	7.06	3.179	0.287	1.44	1.683	0.384	F

F: Fluorescence samples available; L: LC-OCD samples available

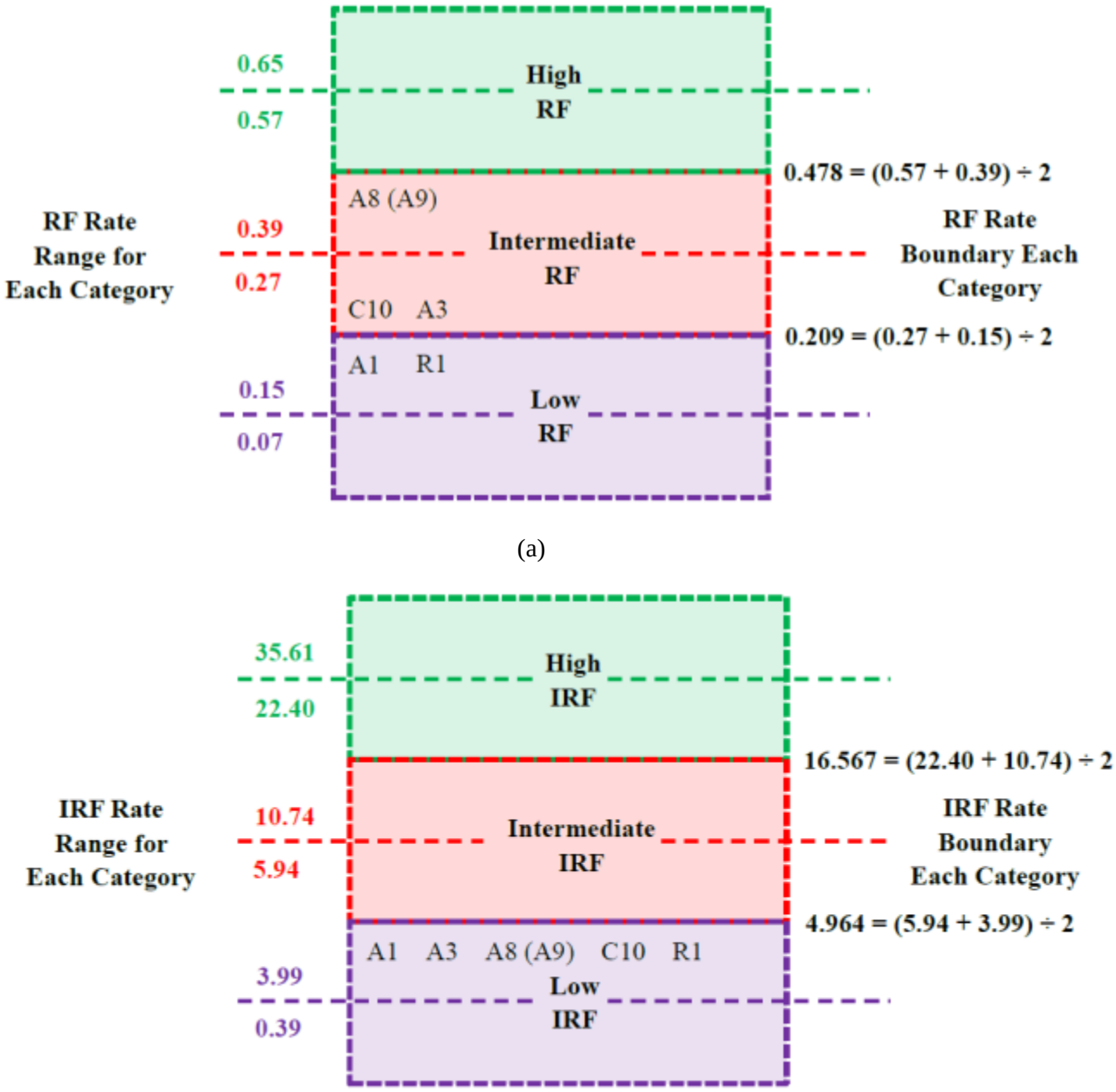
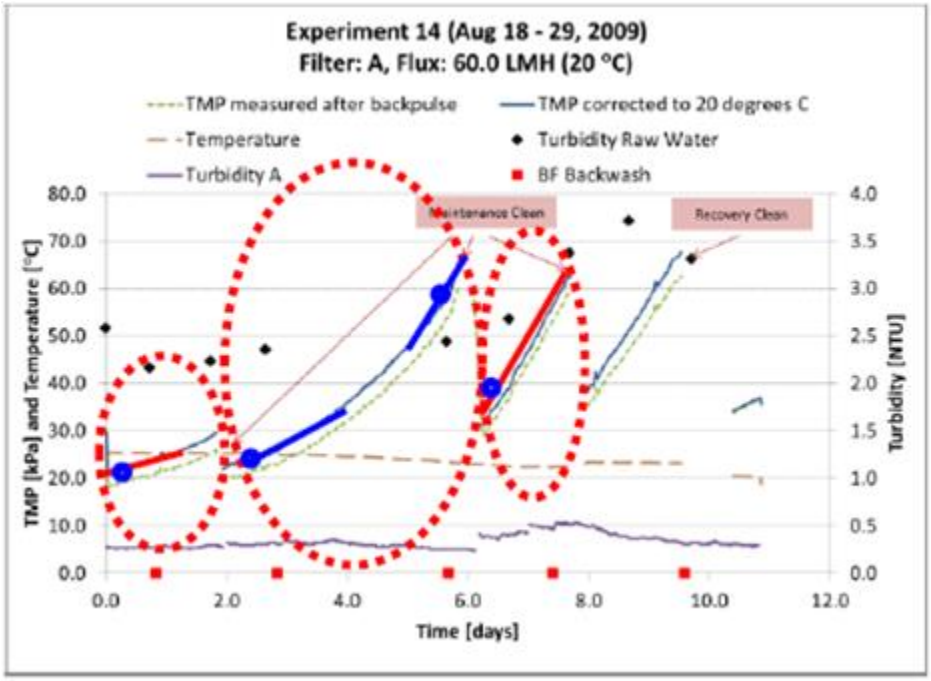
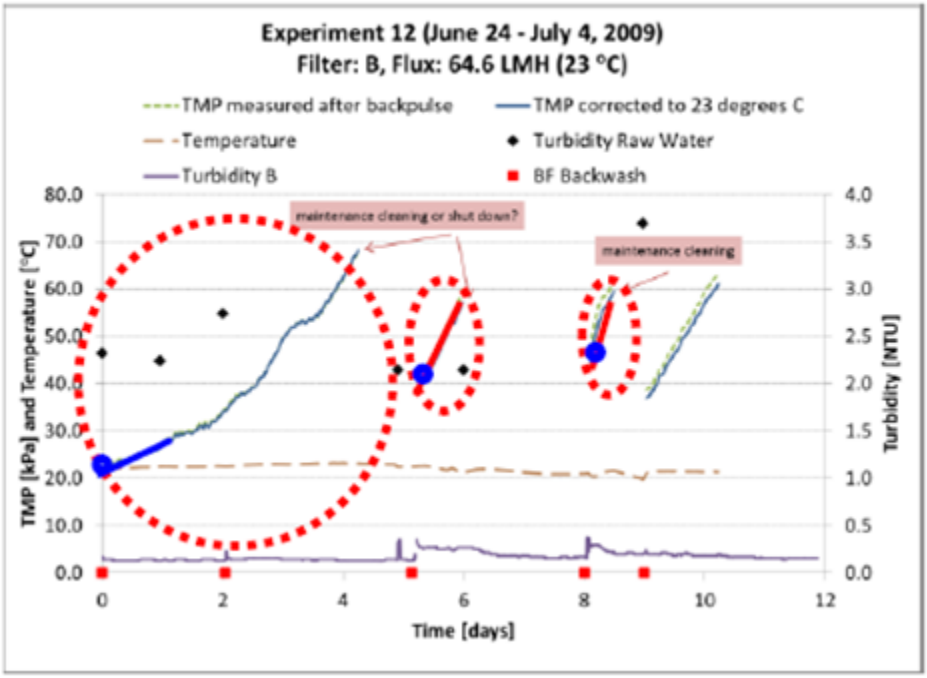


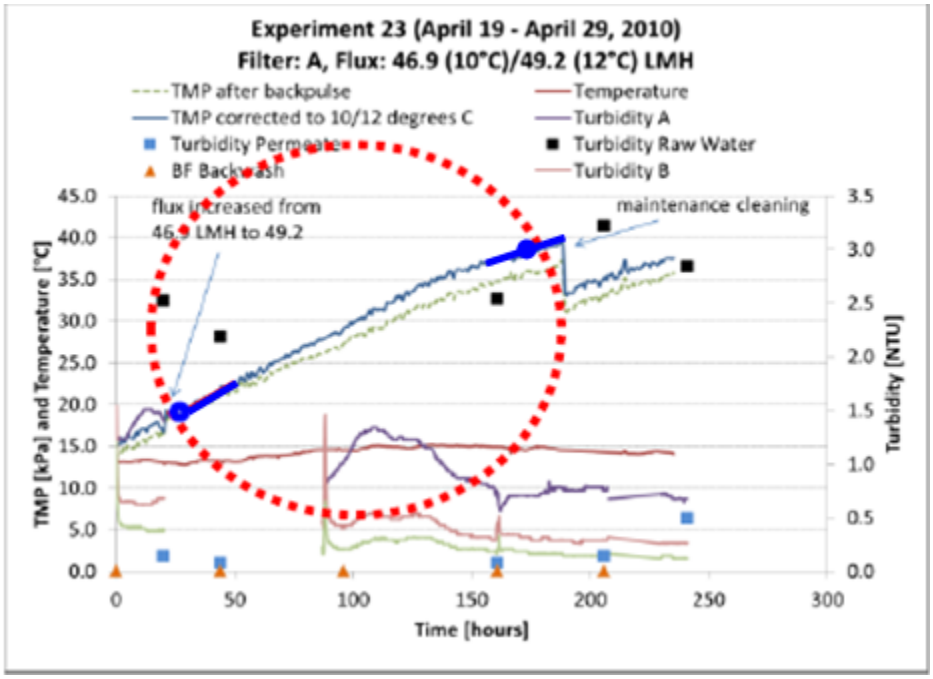
Figure A1 Classification of (a) reversible fouling (RF) rates (kPa/min) and (b) irreversible fouling (IRF) rates (kPa/day). Numbers in color (i.e. green-high, red-intermediate, purple-low) are highest and lowest fouling rates in each category.


Experiment 14 (Aug 18 - 29, 2009)
Filter: A, Flux: 60.0 LMH (20 °C)
TMP measured after backpulse
TMP corrected to 20 degrees C
Temperature
Turbidity Raw Water
Turbidity A
BF Backwash
Maintenance Clean
Recovery Clean
TMP [kPa] and Temperature [°C]
Turbidity [NTU]
Time [days]

(a)

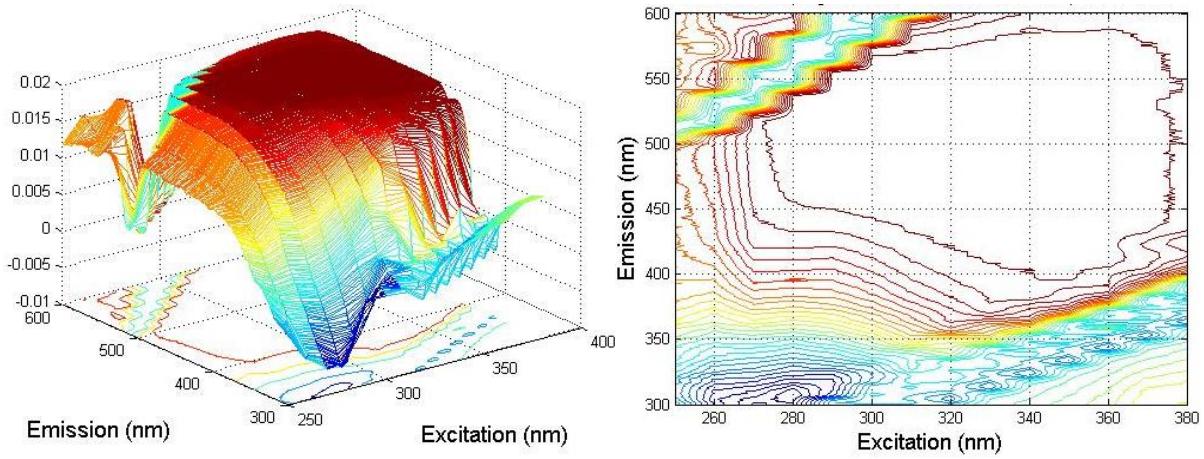

Experiment 12 (June 24 - July 4, 2009)
Filter: B, Flux: 64.6 LMH (23 °C)
TMP measured after backpulse
TMP corrected to 23 degrees C
Temperature
Turbidity Raw Water
Turbidity B
BF Backwash
maintenance cleaning or shut down?
maintenance cleaning
TMP [kPa] and Temperature [°C]
Turbidity [NTU]
Time [days]

(b)

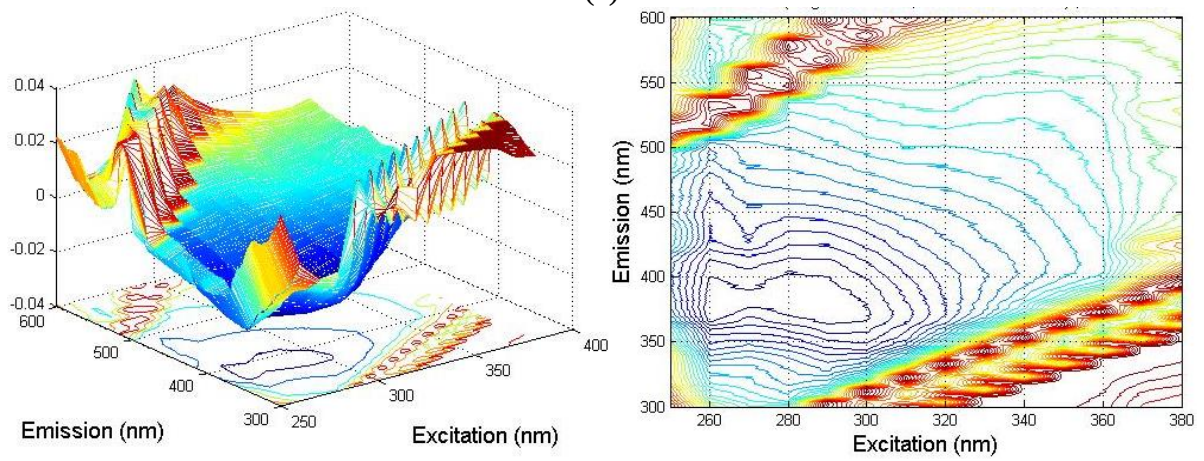


(c)

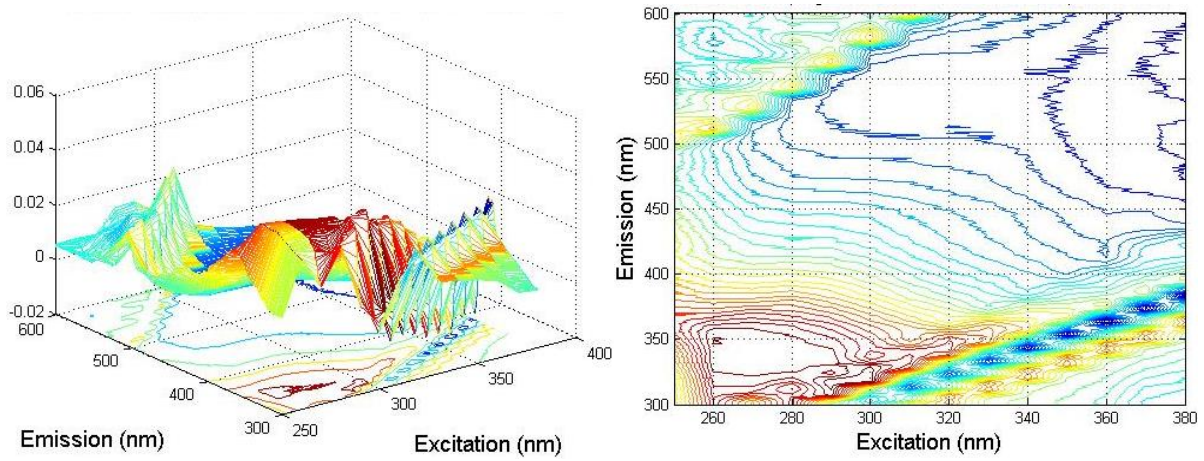
Figure A2 Fouling cycle plots for (a) A4 and A5, (b) B1, and (c) A8 and A9. Irregular samples were shown as blue lines.



(a)



(b)



(c)

Figure A3 Loading plots of (a) PC1 (58.5%): humic-like substances, (b) PC2 (14.8%): colloidal/particulate matter, and (c) PC3 (10.1%): protein-like substances.

Appendix B Supplementary Data for Chapter 4 (i.e. model solution FEEM plots and outputs of multiple linear regression)

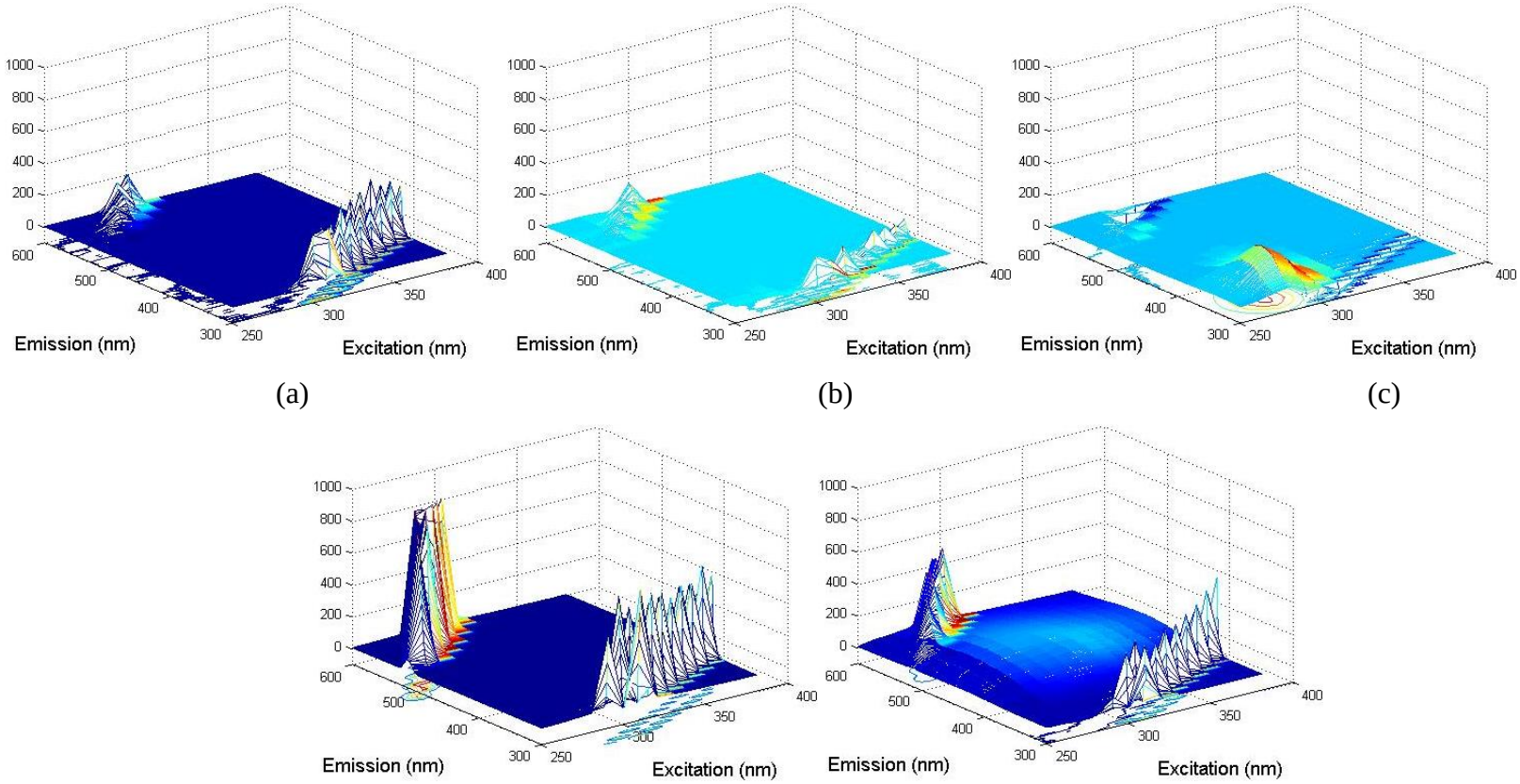


Figure B1 Original FEEM plots of model solutions with (a) alginate (2.5 mg/L) only, (b) alginate (5 mg/L) only, (c) BSA (5 mg/L) only, (d) silica (200 mg/L) only, and (e) HA (5 mg/L) only

Table B1 Illustration of contributions of model solution components and their interactions to each individual principal component (PC) using multiple linear regression (MLR)

	Model solution composition	R^2	*Adjusted* R^2	*P value*				
				A	*B*	*H*	*S*	*BSI*
PC1	ABHS	0.955	0.952	**0.002(+)**	0.224**(–)**	***7.666×10^{-52}*(+)**	**1.027×10^{-7}(+)**	–
	H only	0.927	0.926	–	–	***5.021×10^{-46}*(+)**	–	–
PC2	ABHS + BSI	0.766	0.750	0.896**(+)**	**0.007(+)**	**0.002(–)**	***1.510×10^{-10}*(+)**	0.142**(+)**
	ABHS	0.759	0.746	0.868**(+)**	**2.327×10^{-8}(+)**	**0.002(–)**	***4.273×10^{-22}*(+)**	–
	B + S + BSI	0.732	0.722	–	**0.008(+)**	–	***4.665×10^{-10}*(+)**	0.186**(+)**
	B + S	0.726	0.719	–	**6.193×10^{-8}(+)**	–	***3.016×10^{-21}*(+)**	–
	S only	0.598	0.593	–	–	–	***4.189×10^{-17}*(+)**	–
	B only	0.118	0.106	–	***0.002*(+)**	–	–	–
PC3	ABHS + BSI	0.546	0.516	0.279**(+)**	**0.002(+)**	0.292**(+)**	***2.942×10^{-5}*(–)**	0.065**(+)**
	ABHS	0.525	0.499	0.267**(+)**	***8.779×10^{-11}*(+)**	0.279**(+)**	**1.088×10^{-5}(–)**	–
	B + S + BSI	0.532	0.514	–	**0.0018(+)**	–	***2.258×10^{-5}*(–)**	0.058**(+)**
	B + S	0.509	0.497	–	***9.949×10^{-11}*(+)**	–	**9.287×10^{-6}(–)**	–
	B only	0.366	0.358	–	***2.837×10^{-9}*(+)**	–	–	–
	S only	0.153	0.142	–	–	–	***3.417×10^{-4}*(–)**	–

A: Alginate; B: Bovine serum albumin; H: Humic acid; S: Silica

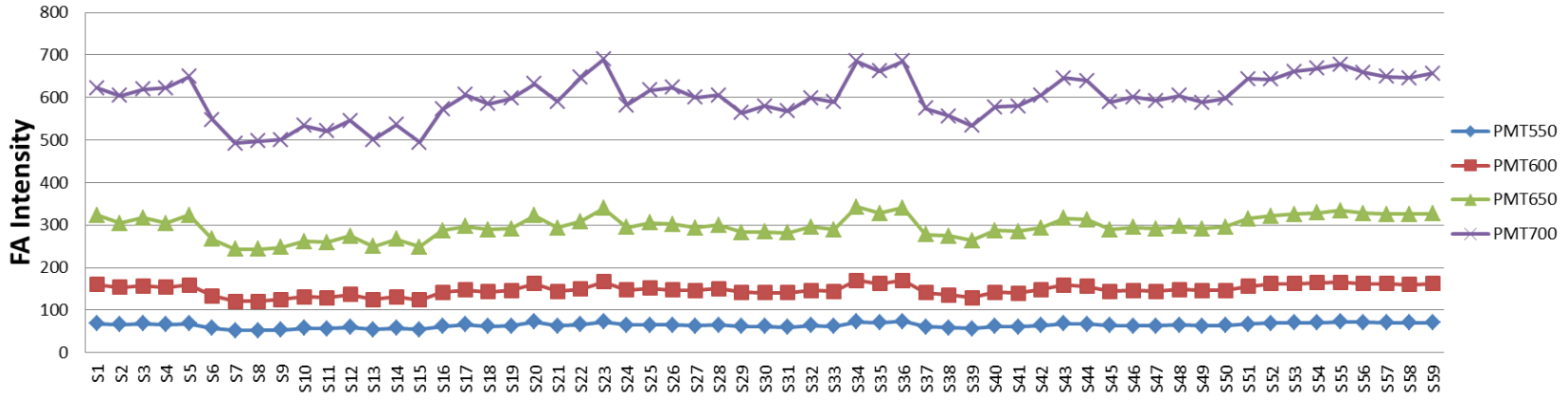
BSI: bovine serum albumin -silica Interaction (BSI = bovine serum albumin conc. × Silica conc.)

Note: the p-values smaller than 0.05 (indicating explanatory variable is significant) are bolded

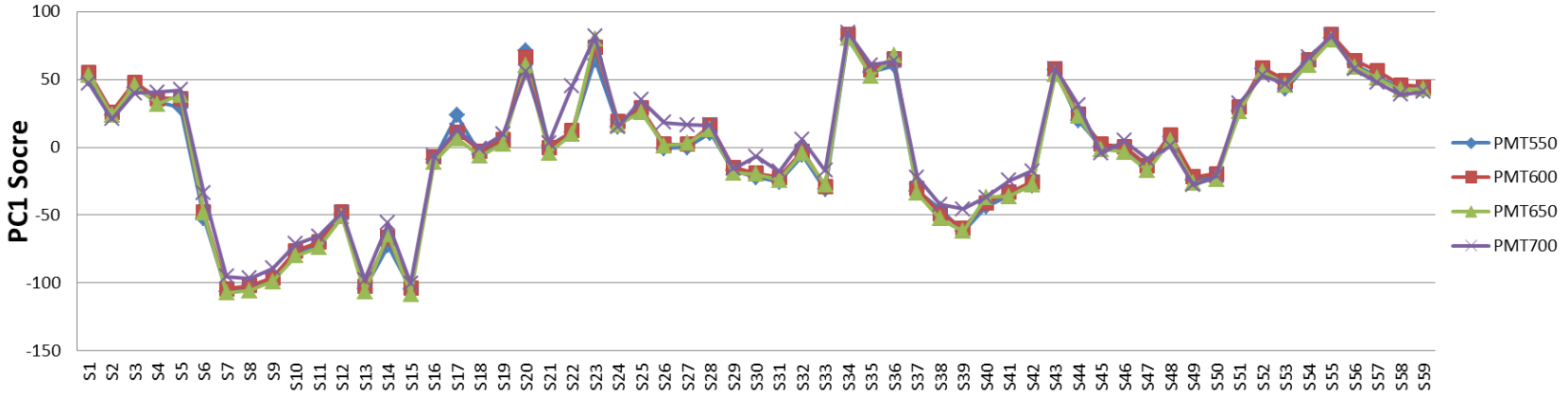
The lowest p-values for each model are bolded and italicized

"+" or "–" in parentheses is the sign of each explanatory parameter in the regression model

Appendix C Supplementary Data for Chapter 5 (i.e. comparison between fulvic acid-like fluorescence intensity and PC1 scores)



(a)



(b)

Figure C1 Illustration of FA intensities and PC1 scores for all samples

Appendix D
Supplementary Data for Chapter 6 (i.e. FEEM plots of samples from different sources, comparison of LC-OCD chromatograms among background, unfractionated biopolymers and biopolymer sized fraction samples, comparison of FEEM protein intensity between background and biopolymer sized fraction samples, and FEEM plots of background and biopolymer sized fraction samples)

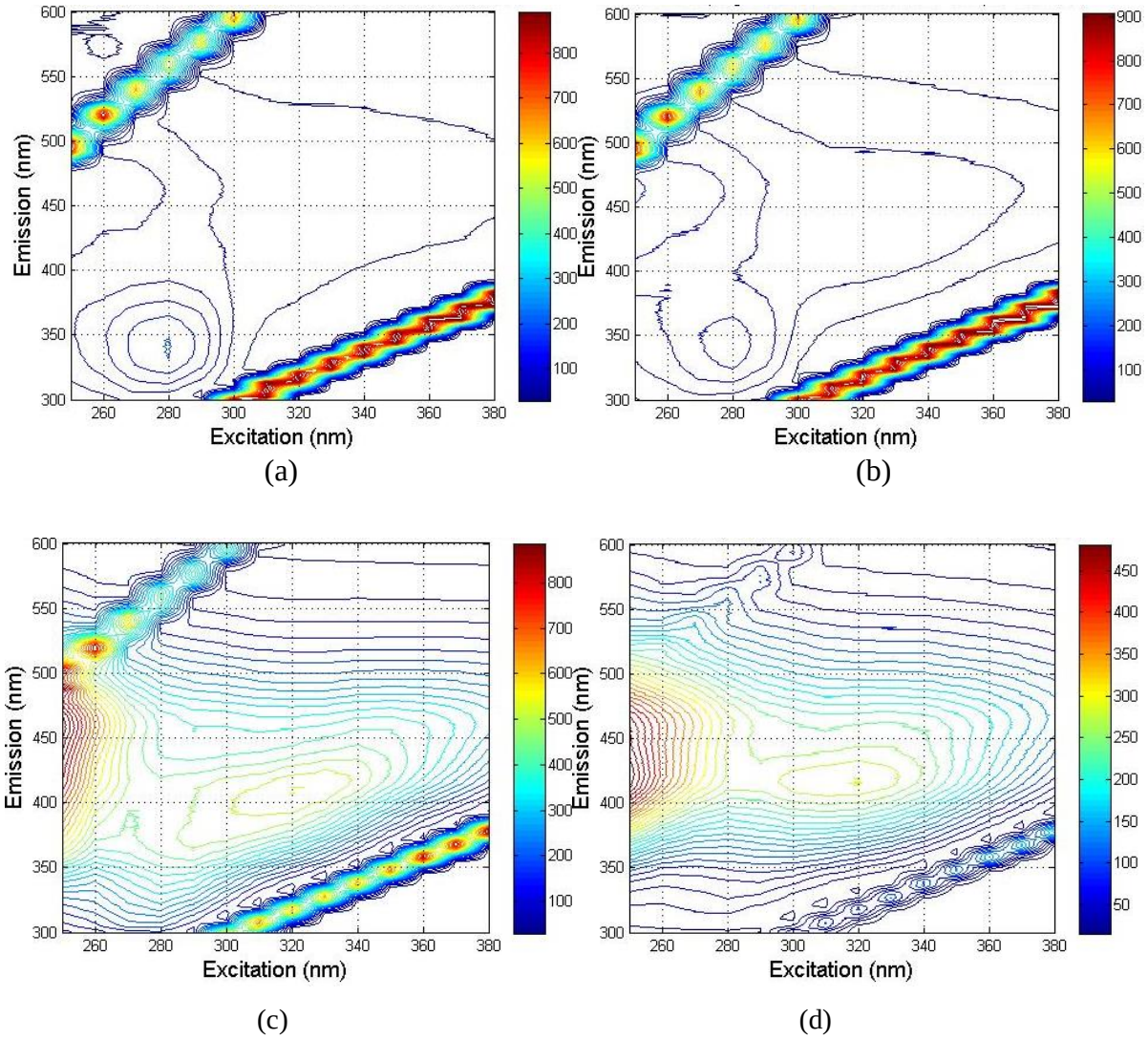
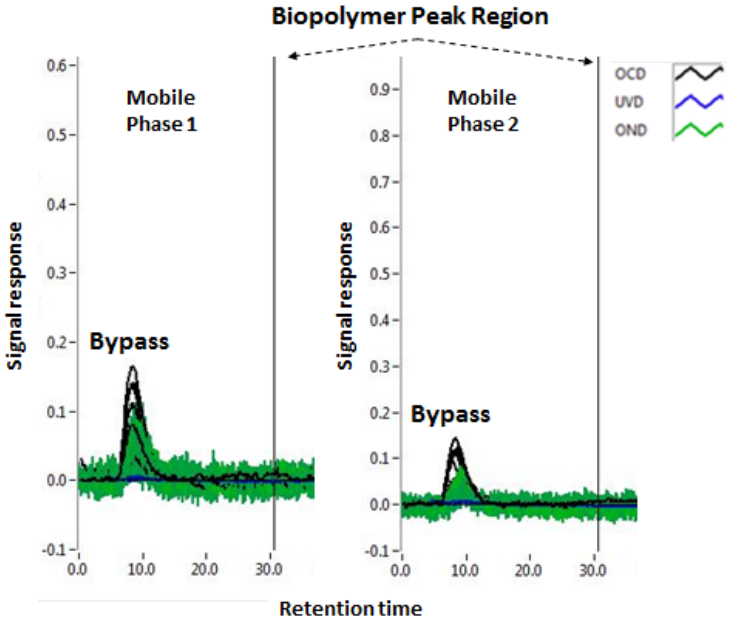
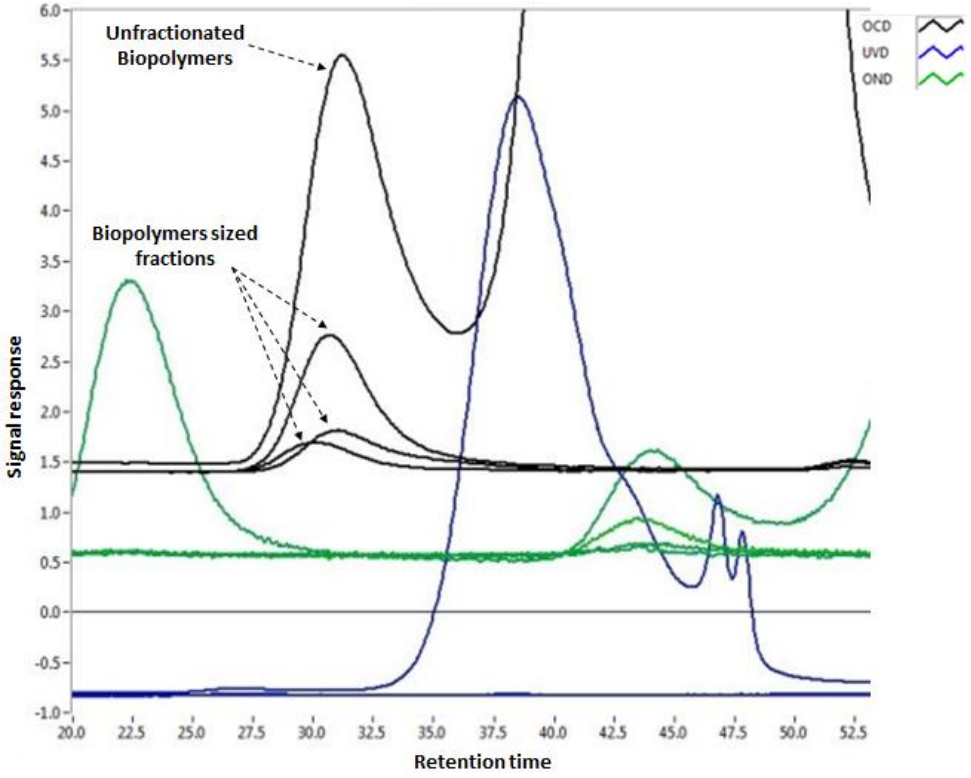


Figure D1 FEEM plots of (a) sand media biofilm EPS extract, (b) anthracite media biofilm EPS extract, (c) secondary effluent, and (d) Grand River water (GRW)



(a)



(b)

Figure D2 Example LC-OCD chromatograms for (a) background mobile phase of all sources and (b) the comparison between unfractionated biopolymers and three biopolymer sized fractions of Grand River water (i.e. GRW2)

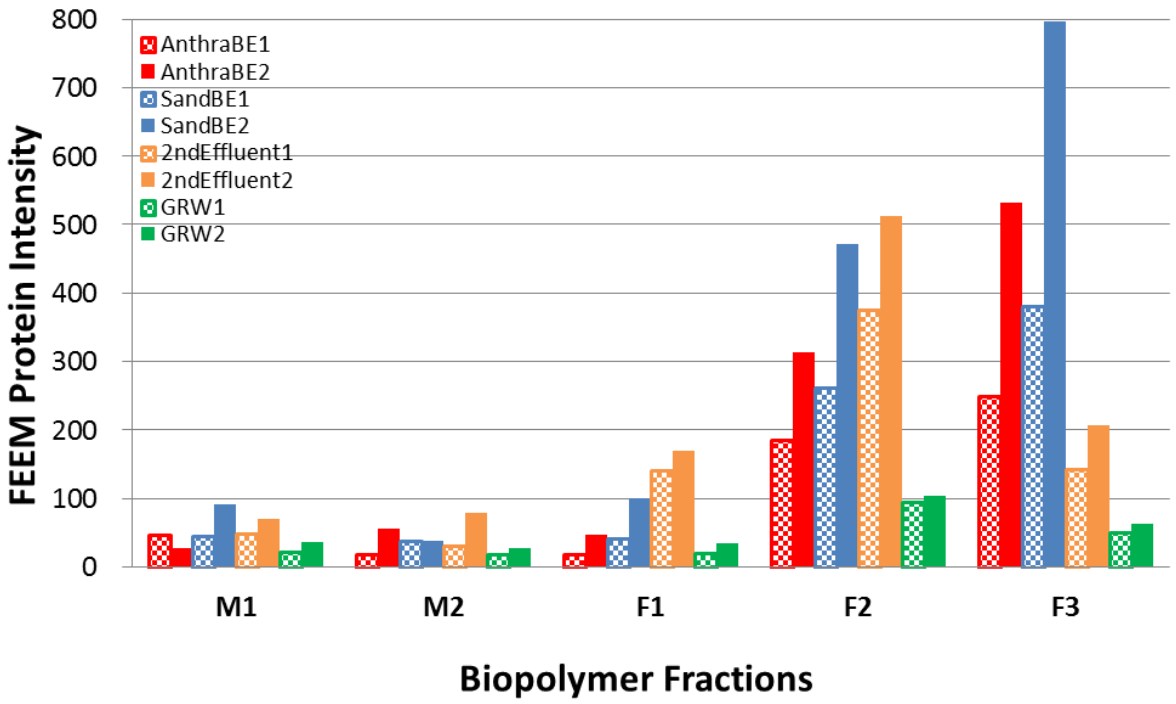


Figure D3 FEEM protein intensities of background mobile phase (M_i - MQ, i = 1, 2) and biopolymer sized fractions (F_i - MQ, i = 1, 2, 3)

AnthraBE: anthracite biofilm EPS; SandBE: sand biofilm EPS; 2ndEffluent: secondary effluent

GRW: Grand River water

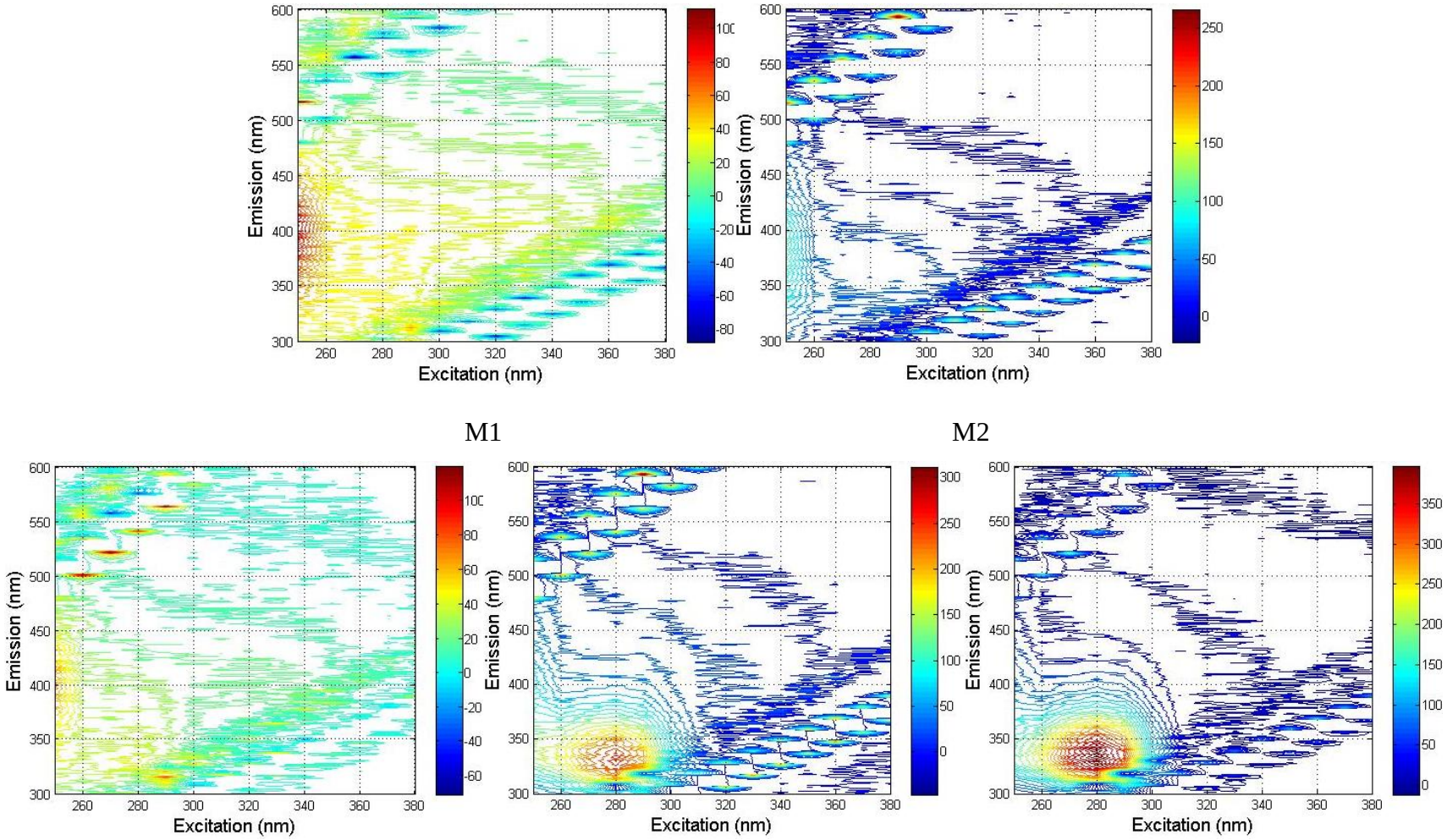


Figure D4 FEEM plots of background mobile phase (M_i - MQ, i = 1, 2) and biopolymer sized fractions (F_i - MQ, i = 1, 2, 3) of sand biofilm EPS extract samples (i.e. SandBE1)

Appendix E
Supplementary Data for Chapter 7 (i.e. turbidity and water temperature of biofilter influent, calculated EBCTs at different depths, concentration and percent removal of NOM components, Summary of R^2 for three kinetics models, kinetics plots for three models, example statistical outputs of kinetics analysis)

Table E1 Turbidity and water temperature of biofilter influent from different sampling events

Note: sampling event on Sept-18 was conducted at a reduced hydraulic loading

	Turbidity (NTU)	*Water Temperature (^oC)*
Aug-14	3.20	22.6
Aug-28	3.10	23.3
Sep-04	4.11	23.3
Sep-11	4.75	21.0
Sep-18	9.72	17.3

Table E2 Calculated empty bed contact time (EBCT, unit: min) at different depths under the regular and reduced hydraulic loading conditions

	Depth (cm)	*EBCT (min)*	
		Regular	Reduced (Sept-18)
Port 1	0 (Influent)	0.00	0.00
Port 2	10	1.95	2.78
Port 3	20	3.89	5.56
Port 4	60	11.67	16.68
Port 5	80	15.57	22.24
Port 6	120	23.35	33.36

Table E3 Concentration and percent removal of various NOM components

Sampling on AUG-14

	FEEM Intensity						*LC-OCD Concentration (µg/L)*						
	Protein_NF	Protein_F	FA_NF	FA_F	HA_NF	HA_F	Bypass DOC	BP-OC	BP-ON	Humics	BB	Neutrals	Acids
0 cm	59.47	48.41	299.32	302.62	341.17	346.07	6417	437	54.8	3885	743	459	0
10 cm	56.15	48.22	304.31	303.66	342.13	341.25	7348	370	46.1	3937	971	993	0
20 cm	52.56	43.09	302.07	298.35	339.52	340.21	6040	310	37.3	3632	821	451	57
60 cm	49.85	42.60	296.21	294.25	338.74	331.16	5895	245	30.8	3670	774	426	34
80 cm	42.95	38.09	289.95	288.75	331.92	330.45	5753	154	25.4	3526	778	419	57
120 cm	40.23	32.04	287.37	279.46	330.80	322.77	5408	123	19.7	3583	594	378	0
	Percent removal												
10 cm	5.6%	0.4%	-1.7%	-0.3%	-0.3%	1.4%	-14.5%	15.3%	15.9%	-1.3%	-30.7%	-116.3%	N/A
20 cm	11.6%	11.0%	-0.9%	1.4%	0.5%	1.7%	5.9%	29.1%	31.9%	6.5%	-10.5%	1.7%	N/A
60 cm	16.2%	12.0%	1.0%	2.8%	0.7%	4.3%	8.1%	43.9%	43.8%	5.5%	-4.2%	7.2%	N/A
80 cm	27.8%	21.3%	3.1%	4.6%	2.7%	4.5%	10.3%	64.8%	53.6%	9.2%	-4.7%	8.7%	N/A
120 cm	32.4%	33.8%	4.0%	7.7%	3.0%	6.7%	15.7%	71.9%	64.1%	7.8%	20.1%	17.6%	N/A

F: filtered; NF: non-filtered

FA: fulvic acid-like; HA: humic acid-like; BP: biopolymers; BB: building blocks

Table E3 Sampling on AUG-28

	FEEM Intensity						*LC-OCD Concentration (µg/L)*						
	Protein_NF	Protein_F	FA_NF	FA_F	HA_NF	HA_F	Bypass DOC	BP-OC	BP-ON	Humics	BB	Neutrals	Acids
0 cm	69.09	50.94	306.07	317.13	349.26	355.27	6315	434	54.7	4238	877	511	0
10 cm	54.14	46.43	315.07	312.20	353.64	352.33	6027	343	47.6	4106	813	466	0
20 cm	52.49	44.65	314.09	309.09	353.17	350.96	5846	267	37.9	3971	863	471	2
60 cm	45.30	39.98	305.12	303.65	346.96	346.25	5717	161	22	3866	814	421	0
80 cm	40.98	38.41	296.75	298.77	344.80	344.16	5477	119	19.4	3753	840	401	0
120 cm	37.61	38.68	298.84	298.68	344.67	338.99	5362	85	15.7	3728	833	426	0
	Percent removal												
10 cm	21.6%	8.9%	-2.9%	1.6%	-1.3%	0.8%	4.6%	21.0%	13.0%	3.1%	7.3%	8.8%	N/A
20 cm	24.0%	12.3%	-2.6%	2.5%	-1.1%	1.2%	7.4%	38.5%	30.7%	6.3%	1.6%	7.8%	N/A
60 cm	34.4%	21.5%	0.3%	4.2%	0.7%	2.5%	9.5%	62.9%	59.8%	8.8%	7.2%	17.6%	N/A
80 cm	40.7%	24.6%	3.0%	5.8%	1.3%	3.1%	13.3%	72.6%	64.5%	11.4%	4.2%	21.5%	N/A
120 cm	45.6%	24.1%	2.4%	5.8%	1.3%	4.6%	15.1%	80.4%	71.3%	12.0%	5.0%	16.6%	N/A

F: filtered; NF: non-filtered

FA: fulvic acid-like; HA: humic acid-like; BP: biopolymers; BB: building blocks

Table E3 Sampling on SEP-04

	FEEM Intensity						*LC-OCD Concentration (μg/L)*						
	Protein_NF	Protein_F	FA_NF	FA_F	HA_NF	HA_F	Bypass DOC	BP-OC	BP-ON	Humics	BB	Neutrals	Acids
0 cm	62.83	50.86	302.76	307.90	341.72	349.20	6066	421	55.9	3945	840	477	15
10 cm	54.28	48.94	306.13	304.00	343.82	342.78	5980	356	47.9	3797	859	455	48
20 cm	50.04	46.18	303.48	299.69	340.40	338.70	5700	322	39.2	3796	812	424	64
60 cm	45.56	42.23	296.69	295.92	334.75	335.61	5483	215	33.9	3699	799	421	18
80 cm	40.52	38.22	291.26	288.27	330.29	330.29	5212	120	20.7	3594	762	400	14
120 cm	39.02	37.29	288.61	289.76	333.54	333.12	5065	94	14.1	3520	780	392	13
	Percent removal												
10 cm	13.6%	3.8%	-1.1%	1.3%	-0.6%	1.8%	1.4%	15.4%	14.3%	3.8%	-2.3%	4.6%	-220.0%
20 cm	20.4%	9.2%	-0.2%	2.7%	0.4%	3.0%	6.0%	23.5%	29.9%	3.8%	3.3%	11.1%	-326.7%
60 cm	27.5%	17.0%	2.0%	3.9%	2.0%	3.9%	9.6%	48.9%	39.4%	6.2%	4.9%	11.7%	-20.0%
80 cm	35.5%	24.9%	3.8%	6.4%	3.3%	5.4%	14.1%	71.5%	63.0%	8.9%	9.3%	16.1%	6.7%
120 cm	37.9%	26.7%	4.7%	5.9%	2.4%	4.6%	16.5%	77.7%	74.8%	10.8%	7.1%	17.8%	13.3%

F: filtered; NF: non-filtered

FA: fulvic acid-like; HA: humic acid-like; BP: biopolymers; BB: building blocks

Table E3 Sampling on SEP-11

	FEEM Intensity						*LC-OCD Concentration (μg/L)*						
	Protein_NF	Protein_F	FA_NF	FA_F	HA_NF	HA_F	Bypass DOC	BP-OC	BP-ON	Humics	BB	Neutrals	Acids
0 cm	62.24	53.16	379.74	387.10	424.29	425.24	7381	439	53.6	4798	835	505	0
10 cm	58.73	51.10	381.69	381.20	427.21	424.16	7166	395	49.6	4569	967	500	11
20 cm	54.42	47.40	378.40	376.91	425.78	418.29	7305	349	46.5	4495	934	481	43
60 cm	49.87	45.71	374.90	372.34	423.34	415.27	6884	252	36.4	4419	886	471	16
80 cm	45.02	40.53	368.65	366.97	417.52	413.77	6715	143	26.9	4356	850	436	49
120 cm	43.82	40.12	368.05	362.65	419.57	413.89	6379	120	19.7	4314	828	438	15
	Percent removal												
10 cm	5.6%	3.9%	-0.5%	1.5%	-0.7%	0.3%	2.9%	10.0%	7.5%	4.8%	-15.8%	1.0%	N/A
20 cm	12.6%	10.8%	0.4%	2.6%	-0.4%	1.6%	1.0%	20.5%	13.2%	6.3%	-11.9%	4.8%	N/A
60 cm	19.9%	14.0%	1.3%	3.8%	0.2%	2.3%	6.7%	42.6%	32.1%	7.9%	-6.1%	6.7%	N/A
80 cm	27.7%	23.8%	2.9%	5.2%	1.6%	2.7%	9.0%	67.4%	49.8%	9.2%	-1.8%	13.7%	N/A
120 cm	29.6%	24.5%	3.1%	6.3%	1.1%	2.7%	13.6%	72.7%	63.2%	10.1%	0.8%	13.3%	N/A

F: filtered; NF: non-filtered

FA: fulvic acid-like; HA: humic acid-like; BP: biopolymers; BB: building blocks

Table E3 Sampling on SEP-18

	FEEM Intensity						*LC-OCD Concentration (μg/L)*						
	Protein_NF	Protein_F	FA_NF	FA_F	HA_NF	HA_F	Bypass DOC	BP-OC	BP-ON	Humics	BB	Neutrals	Acids
0 cm	76.94	49.84	392.11	400.49	447.03	455.27	7519	321	42	4893	956	487	0
10 cm	57.25	45.59	394.30	400.72	455.45	458.75	7506	295	40	4866	947	496	87
20 cm	56.39	42.10	393.26	392.72	451.69	451.41	7145	229	35.8	4755	881	454	9
60 cm	43.98	38.62	389.88	388.95	449.82	444.58	7148	137	24.7	4674	831	441	43
80 cm	39.94	37.98	383.30	381.83	445.66	444.93	6680	83	16.7	4462	931	422	10
120 cm	39.83	35.60	384.98	384.45	445.61	445.04	6463	78	15	4548	777	409	10
	Percent removal												
10 cm	25.6%	8.5%	-0.6%	-0.1%	-1.9%	-0.8%	0.2%	8.1%	4.8%	0.6%	0.9%	-1.8%	N/A
20 cm	26.7%	15.5%	-0.3%	1.9%	-1.0%	0.8%	5.0%	28.7%	14.8%	2.8%	7.8%	6.8%	N/A
60 cm	42.8%	22.5%	0.6%	2.9%	-0.6%	2.3%	4.9%	57.3%	41.2%	4.5%	13.1%	9.4%	N/A
80 cm	48.1%	23.8%	2.2%	4.7%	0.3%	2.3%	11.2%	74.1%	60.2%	8.8%	2.6%	13.3%	N/A
120 cm	48.2%	28.6%	1.8%	4.0%	0.3%	2.2%	14.0%	75.7%	64.3%	7.1%	18.7%	16.0%	N/A

F: filtered; NF: non-filtered

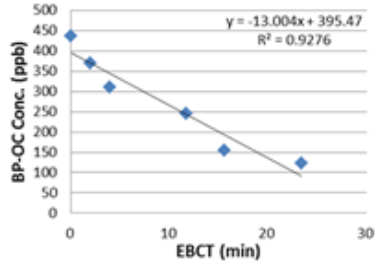
FA: fulvic acid-like; HA: humic acid-like; BP: biopolymers; BB: building blocks

Table E4 Summary of R^2 of zero-, first-, and second-order kinetic models for various NOM components

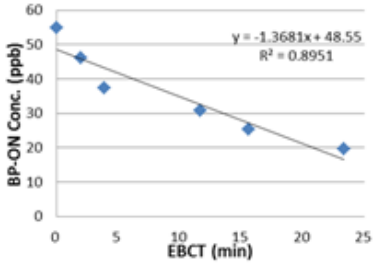
	Zero order kinetics				
	Aug-14	**Aug-28**	**Sep-04**	**Sep-11**	**Sep-18**
LC-OCD_BP-OC	0.93	0.88	0.94	0.95	0.90
LC-OCD_BP-ON	0.90	0.88	0.94	***0.99***	0.94
FEEM-Protein_NF	0.94	0.81	0.84	0.91	0.74
FEEM-Protein_F	***0.93***	0.81	0.94	0.89	0.85
FEEM-HA_NF	***0.88***	0.70	***0.75***	***0.69***	0.39
FEEM-HA_F	0.97	0.99	0.72	0.78	0.72
FEEM-FA_NF	0.86	0.63	0.92	0.90	***0.78***
FEEM-FA_F	***0.97***	0.88	0.86	0.94	0.80
LC-OCD_Humics	0.61	0.86	0.93	0.76	***0.80***
	First order kinetics				
	Aug-14	**Aug-28**	**Sep-04**	**Sep-11**	**Sep-18**
LC-OCD_BP-OC	***0.96***	0.98	***0.97***	***0.96***	***0.94***
LC-OCD_BP-ON	0.96	0.95	***0.96***	0.99	***0.96***
FEEM-Protein_NF	0.95	0.88	0.89	0.93	0.81
FEEM-Protein_F	0.93	0.83	0.95	0.90	0.88
FEEM-HA_NF	0.88	0.70	0.75	0.69	0.39
FEEM-HA_F	0.97	0.99	0.73	0.78	0.72
FEEM-FA_NF	0.87	0.63	0.93	0.90	0.78
FEEM-FA_F	0.97	0.89	0.86	0.94	0.80
LC-OCD_Humics	0.61	0.87	0.93	0.77	0.80
	Second order kinetics				
	Aug-14	**Aug-28**	**Sep-04**	**Sep-11**	**Sep-18**
LC-OCD_BP-OC	0.95	***0.99***	0.94	0.94	0.94
LC-OCD_BP-ON	***0.98***	***0.99***	0.92	0.95	0.95
FEEM-Protein_NF	***0.95***	***0.93***	***0.92***	***0.94***	***0.86***
FEEM-Protein_F	0.92	***0.84***	***0.95***	***0.91***	***0.91***
FEEM-HA_NF	0.88	***0.70***	0.75	0.69	***0.39***
FEEM-HA_F	***0.98***	***0.99***	***0.73***	***0.78***	***0.73***
FEEM-FA_NF	***0.87***	***0.63***	***0.93***	***0.90***	0.77
FEEM-FA_F	0.97	***0.89***	***0.86***	***0.94***	***0.80***
LC-OCD_Humics	***0.61***	***0.88***	***0.94***	***0.78***	0.79

BP: Biopolymer; OC: Organic Carbon; ON: Organic Nitrogen;
NF: Non-Filtered; F: Filtered; HA: Humic Acid; FA: Fulvic Acid
NOTE: The R^2 of best-fit kinetics is bolded and italicized for each NOM component from each sampling event.

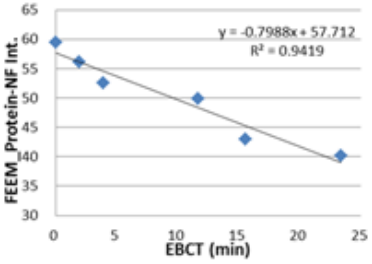
LC-OCD Biopolymer (BP)-OC



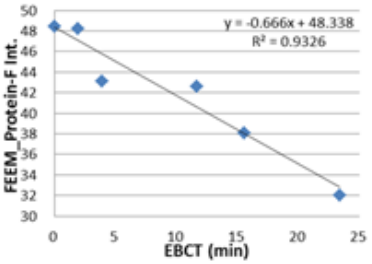
LC-OCD BP-ON



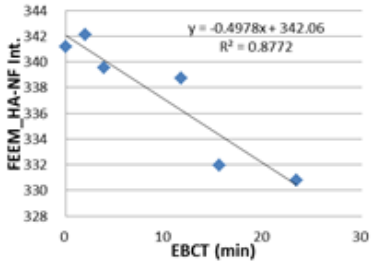
FEEM Protein_Nonfiltered (NF)



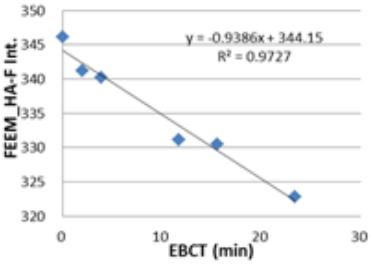
FEEM Protein_Filtered (F)



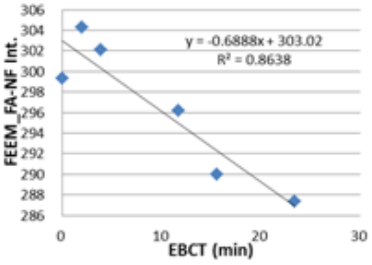
FEEM HA_NF



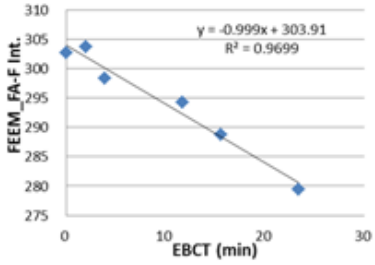
FEEM HA_F



FEEM FA_NF



FEEM FA_F



LC-OCD Humics (HS)

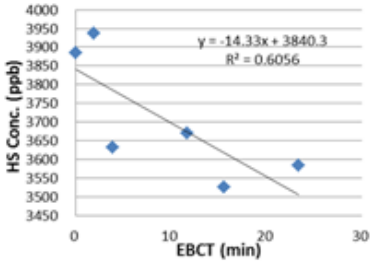
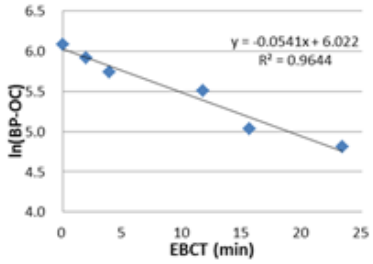
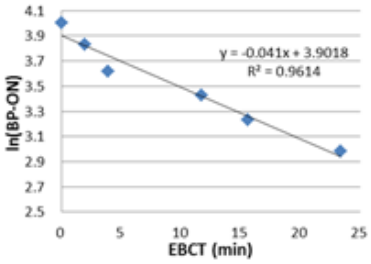


Figure E1 (a)

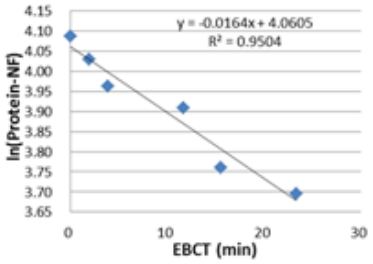
LC-OCD Biopolymer (BP)-OC



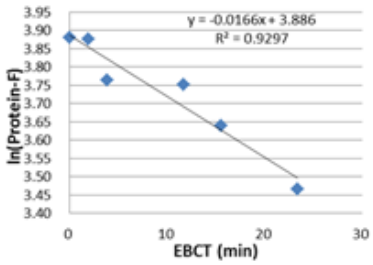
LC-OCD BP-ON



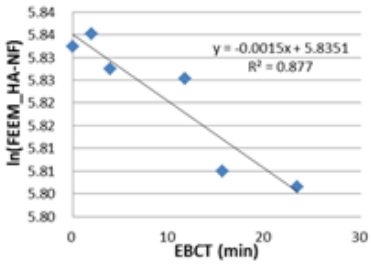
FEEM Protein_Nonfiltered (NF)



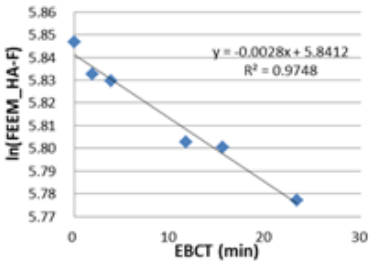
FEEM Protein_Filtered (F)



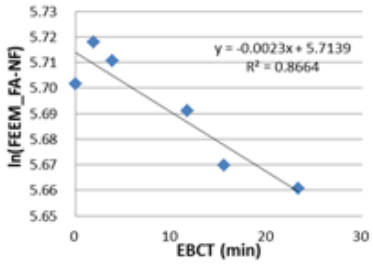
FEEM HA_NF



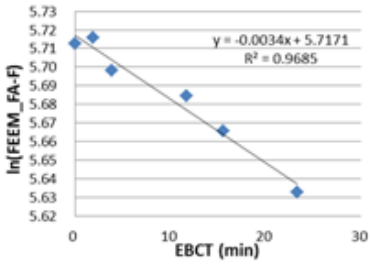
FEEM HA_F



FEEM FA_NF



FEEM FA_F



LC-OCD Humics (HS)

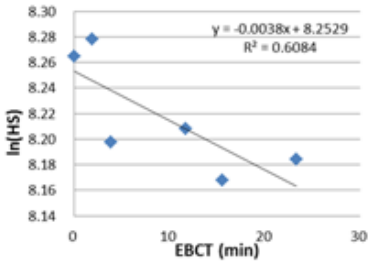
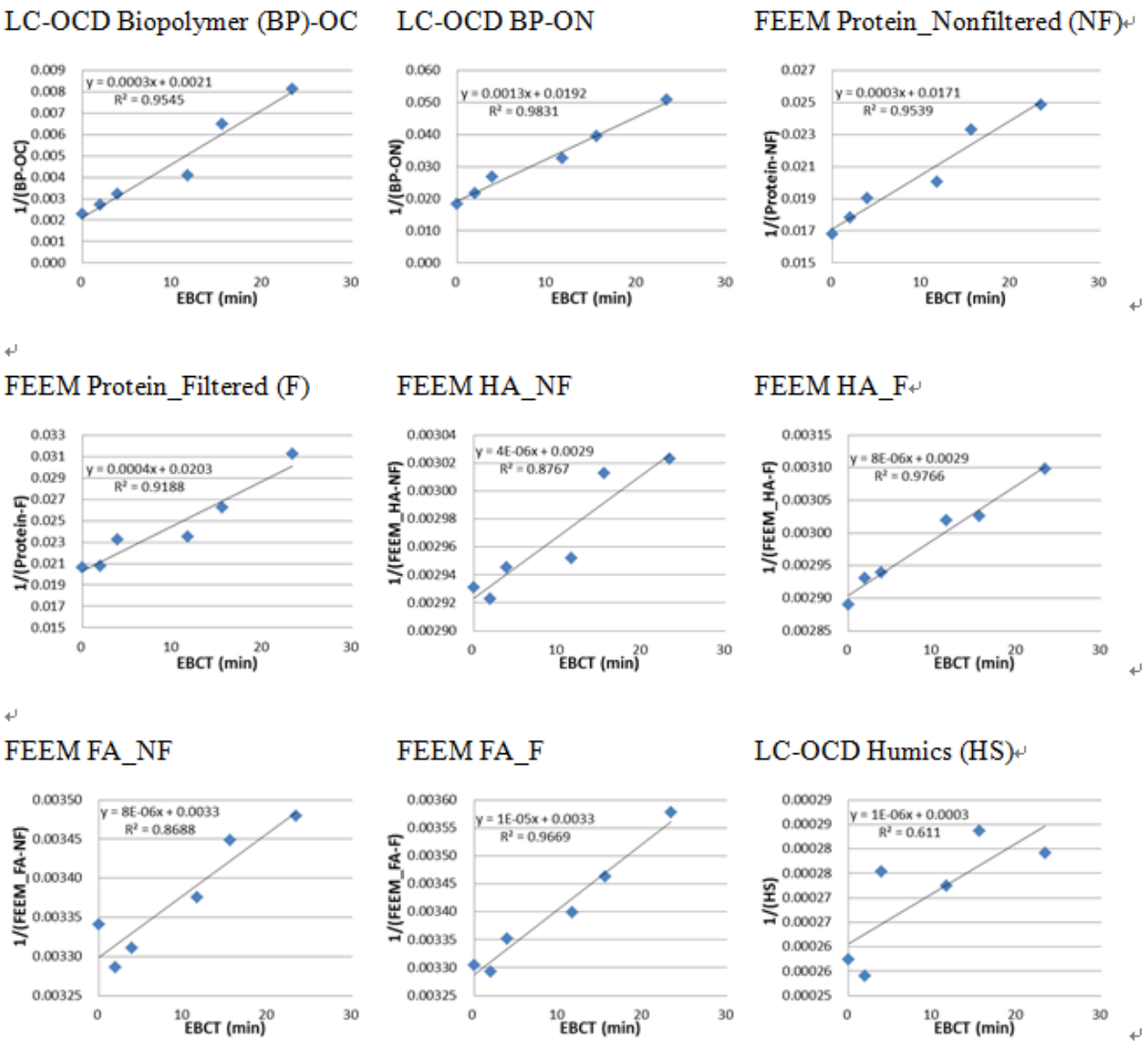


Figure E1 (b)



(c)

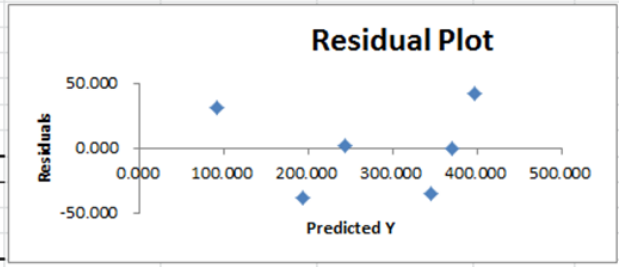
Figure E1 (a) zero-order, (b) first-order, and (c) second-order kinetics plots for sampling event on Aug-14

Regression Statistics	
Multiple R	0.96
R Square	0.93
Adjusted R Square	0.91
Standard Error	36.87
Observations	6

ANOVA

	df	SS	MS	F	Significance F
Regression	1	69680.839	69680.839	51.255	0.002
Residual	4	5437.994	1359.498		
Total	5	75118.833			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	395.466	22.768	17.369	0.000	332.250	458.681	332.250	458.681
X Variable 1	-13.004	1.816	-7.159	0.002	-18.048	-7.961	-18.048	-7.961



RESIDUAL OUTPUT

Observation	Predicted Y	Residuals	Standard Residuals
1	395.466	41.534	1.259
2	370.162	-0.162	-0.005
3	344.859	-34.859	-1.057
4	243.646	1.354	0.041
5	193.040	-39.040	-1.184
6	91.827	31.173	0.945

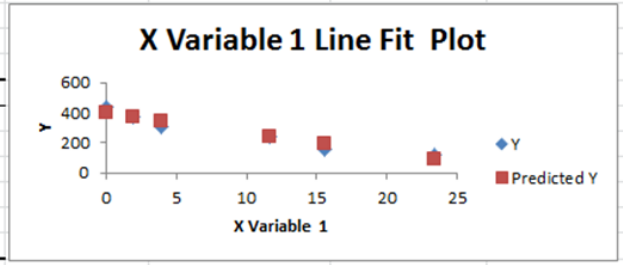
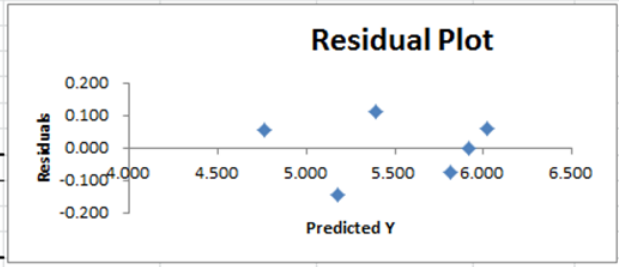


Figure E2 (a)

Regression Statistics	
Multiple R	0.98
R Square	0.96
Adjusted R Square	0.96
Standard Error	0.11
Observations	6

ANOVA

	df	SS	MS	F	Significance F
Regression	1	1.205	1.205	108.233	0.000
Residual	4	0.045	0.011		
Total	5	1.250			



	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	6.022	0.065	92.418	0.000	5.841	6.203	5.841	6.203
X Variable 1	-0.054	0.005	-10.404	0.000	-0.069	-0.040	-0.069	-0.040

RESIDUAL OUTPUT

Observation	Predicted Y	Residuals	Standard Residuals
1	6.022	0.058	0.614
2	5.917	-0.003	-0.035
3	5.812	-0.075	-0.794
4	5.391	0.111	1.172
5	5.180	-0.143	-1.517
6	4.759	0.053	0.561

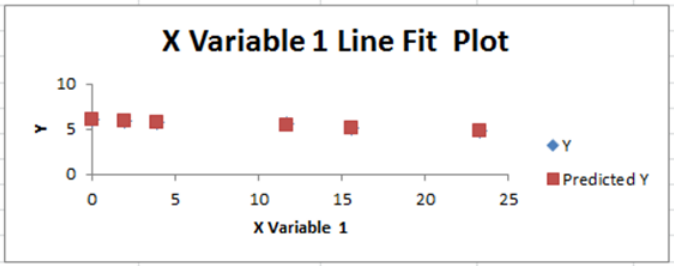


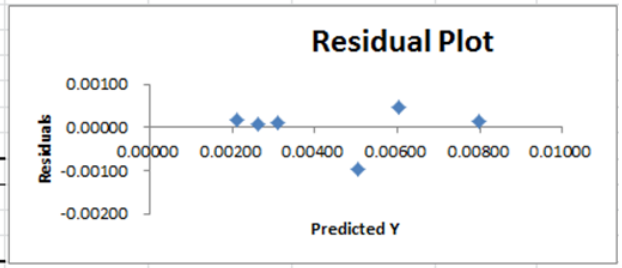
Figure E2 (b)

Regression Statistics	
Multiple R	0.98
R Square	0.95
Adjusted R Square	0.94
Standard Error	0.00
Observations	6

ANOVA

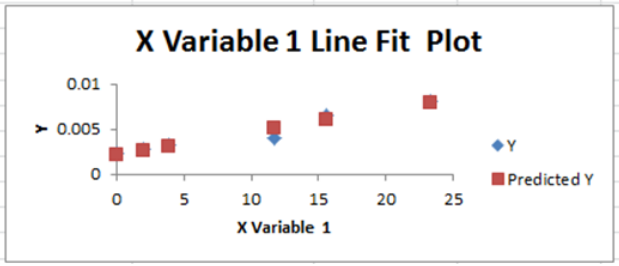
	df	SS	MS	F	Significance F
Regression	1	2.584E-05	2.584E-05	83.913	0.001
Residual	4	1.232E-06	3.079E-07		
Total	5	2.707E-05			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	0.00213	0.00034	6.22163	0.00340	0.00118	0.00308	0.00118	0.00308
X Variable 1	0.00025	0.00003	9.16038	0.00079	0.00017	0.00033	0.00017	0.00033



RESIDUAL OUTPUT

Observation	Predicted Y	Residuals	Standard Residuals
1	0.00213	0.00016	0.31511
2	0.00262	0.00008	0.16826
3	0.00311	0.00012	0.24048
4	0.00506	-0.00097	-1.96209
5	0.00603	0.00046	0.93388
6	0.00798	0.00015	0.30435



(c)

Figure E2 Statistical outputs for (a) zero-order, (b) first-order, and (c) second-order kinetics models of biopolymer-OC from sampling event on Aug-14

Appendix F

Reproducibility of FEEM and LC-OCD measurements

Grand River water (GRW, DOC=5.3 mgC/L) was collected in a glass bottle from the pilot-scale facility located within a full-scale drinking water treatment plant (Southwestern Ontario, Canada) and used to test the reproducibility of FEEM and LC-OCD measurements. For filtered sample measurement, the GRW was firstly pre-filtered using seven different 0.45 μm polyethersulfone filters and the filtered samples were collected into seven different glass vials. Then, these seven filtered samples were measured by the LC-OCD and FEEM at the same condition and using the same instrumental settings, respectively. The detailed instrumental settings for reproducibility test can be found in Sections 6.2.2.1 and 6.2.3.2 in Chapter 6 for LC-OCD and FEEM, respectively. The reproducibility test of non-filtered sample measurement was also conducted for FEEM analysis. The non-filtered GRW samples from the collection bottle were further transferred into seven different glass vials and these seven non-filtered samples were measured by FEEM following the same procedure as mentioned above. The reproducibility test results from the above mentioned experiments were summarized in Tables F1 and F2 for FEEM and LC-OCD, respectively.

Table F1 Reproducibility results for FEEM measurements

Filtered	*Protein* Ex./Em.=280/330 nm	*Fulvic Acid* Ex./Em.=320/415 nm	*Humic Acid* Ex./Em.=270/460 nm
GRW_Filter1_PMT650_SW10	50.62	300.28	345.99
GRW_Filter2_PMT650_SW10	50.58	301.25	344.41
GRW_Filter3_PMT650_SW10	50.01	299.30	340.93
GRW_Filter4_PMT650_SW10	49.82	302.93	343.82
GRW_Filter5_PMT650_SW10	49.42	298.56	343.76
GRW_Filter6_PMT650_SW10	49.24	298.70	342.82
GRW_Filter7_PMT650_SW10	48.66	299.89	341.84
Standard Deviation	0.72	1.55	1.68
Average	49.76	300.13	343.37
Standard Deviation/Average	1.44%	0.52%	0.49%
Non-filtered	*Protein*	*Fulvic Acid*	*Humic Acid*
GRW_Vial1_PMT650_SW10_S	75.36	292.88	336.32
GRW_Vial2_PMT650_SW10_S	74.99	294.07	337.16
GRW_Vial3_PMT650_SW10_S	75.21	291.94	335.37
GRW_Vial4_PMT650_SW10_S	76.28	292.91	334.19
GRW_Vial5_PMT650_SW10_S	74.11	291.96	334.53
GRW_Vial6_PMT650_SW10_S	75.18	294.63	337.09
GRW_Vial8_PMT650_SW10_S	74.86	295.74	337.63
Standard Deviation	0.65	1.42	1.36
Average	75.14	293.45	336.04
Standard Deviation/Average	0.86%	0.48%	0.40%

Table F2 Reproducibility results for LC-OCD measurements (unit: μg/L)

Sample name	*Bypass DOC*	*Biopolymer-OC*	*Biopolymer-ON*	*Humics*	*Building Blocks*	*Neutrals*	*Acids*
GRW_Bottle1_DF1	5363	363	41.2	3116	619	447	29
GRW_Bottle2_DF1	5350	360	38.6	3151	599	443	22
GRW_Bottle3_DF1	5317	359	38.8	3030	697	430	21
GRW_Bottle4_DF1	5261	369	41.6	3132	621	436	26
GRW_Bottle5_DF1	5350	359	39.8	3115	631	438	27
GRW_Bottle6_DF1	5174	338	40.9	3095	639	419	27
GRW_Bottle7_DF1	5144	348	40	3110	603	424	22
Standard deviation	89.58	10.31	1.16	38.27	32.82	10.09	3.13
Average	5279.86	356.57	40.13	3107.00	629.86	433.86	24.86
Standard Deviation/Average	1.70%	2.89%	2.90%	1.23%	5.21%	2.33%	12.60%

OC: organic carbon; ON: organic nitrogen