Fall 1-31-2009

Characterization of natural organic matter and precursors to trihalomethanes using spectral flourescence signatures

Krit Punburananon
New Jersey Institute of Technology

ABSTRACT

CHARACTERIZATION OF NATURAL ORGANIC MATTER AND PRECURSORS TO TRIHALOMETHANES USING SPECTRAL FLUORESCENCE SIGNATURES

by
Krit Punburananon

Disinfection is an essential process to kill pathogens (i.e., disease causing organisms) in source water during the production of drinking water. Chlorine is most widely used disinfectant because it is effective, affordable, and also provides chlorine residual to ensure that the water is safe through the distribution system. Nonetheless, chlorine reacts with Natural Organic Matter (NOM) and forms potentially carcinogenic Disinfection By-products (DBPs). The major chlorination DBPs are dominantly Trihalomethanes (THMs). However, not all organic compounds are equally reactive to THMs formation.

NOM in water samples collected from the Delaware & Raritan Canal and its tributaries (Central New Jersey) was isolated by resin adsorption into six fractions: Hydrophobic acid (HPOA), Hydrophobic neutral (HPON), Hydrophobic base (HPOB), Hydrophilic acid (HPIA), Hydrophilic neutral (HPIN), and Hydrophilic base (HPIB). HPIN, HPON, and HPOA were the major fractions in most of samples. Moreover, the fractions' seven-day THMs Formation Potentials (THMFP) were determined. HPOA was found to be the most reactive fraction to THMs formation in addition to being one of the most abundant fractions in the source water.

Additionally, the six fractions were also characterized by fluorescence spectroscopy to obtain three-dimensional fluorescence spectra. The spectra shape and peak locations are unique characteristics of organic compounds and also called Spectral

Fluorescence Signature (SFS). The SFS is the total sum of emission intensity of a sample at different excitation wavelengths, recorded as a matrix of fluorescent intensity in coordinates of excitation and emission wavelengths. Among the six fractions, HPOA spectra were large and the peak intensity was also high. Therefore, fluorescence spectroscopy could be a promising technique for characterization of HPOA fraction or THMs precursors in the source water.

Although a large number of intensities are related to THMs precursors, many of them are highly correlated by nature. Principle component analysis was then used to transform the fluorescence intensities into independent parameters called Principle Components (PCs). Best Subset Algorithm was performed to select the most important PCs for the prediction of THMFP by multiple linear regression. The prediction of THMFP using SFS is a rapid, inexpensive, reagent-free technique and thus can be used for optimization of water treatment processes.

CHARACTERIZATION OF NATURAL ORGANIC MATTER AND PRECURSORS TO TRIHALOMETHANES USING SPECTRAL FLUORESCENCE SIGNATURES

by
Krit Punburananon

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Environmental Engineering

Department of Civil and Environmental Engineering

January 2009

APPROVAL PAGE

CHARACTERIZATION OF NATURAL ORGANIC MATTER AND PRECURSORS TO TRIHALOMETHANES USING SPECTRAL FLUORESCENCE SIGNATURES

Krit Punburananon

12/10/08

Dr. Taha F. Marhaba, Dissertation Advisor — Date
Professor of Civil and Environmental Engineering, NJIT

12/10/08

Dr. Sima Bagheri, Committee Member — Date
Professor of Civil and Environmental Engineering, NJIT

12/10/08

Dr. Hsin-Neng Hsieh, Committee Member — Date
Professor of Civil and Environmental Engineering, NJIT

12/8/08

Dr. Methi Wecharatana, Committee Member — Date
Professor of Civil and Environmental Engineering, NJIT

12/3/08

Dr. Lee Lippincott, Committee Member — Date
Senior Research Scientist
Department of Environmental Protection, NJ

BIOGRAPHICAL SKETCH

Author: Krit Punburananon

Degree: Doctor of Philosophy

Date: January 2009

Undergraduate and Graduate Education:

- Doctor of Philosophy in Environmental Engineering,
  New Jersey Institute of Technology, Newark, NJ, 2009
- Master of Science in Petrochemical Technology,
  Chulalongkorn University, Bangkok, Thailand, 1999
- Bachelor of Engineering in Chemical Engineering,
  Kasetsart University, Bangkok, Thailand, 1997

Major: Environmental Engineering

Presentations and Publications:

Krit Punburananon and Taha Marhaba, Fractionation of natural organic matter and characterization of Trihalomethanes precursors using fluorescence spectroscopy. (Manuscript in preparation).

Krit Punburananon and Taha Marhaba, Prediction of dissolved organic carbon and Trihalomethanes formation potential of source water by fluorescence spectroscopy coupled with multivariate statistical analysis. (Manuscript in preparation).

Krit Punburananon and Taha Marhaba, "Characterization of Precursors to Trihalomethanes in Source Water by Multivariate Statistical Analysis of Fluorescence Spectra" Poster presentation at The 3rd Passaic River Symposium, Montclair, New Jersey, October 2008.

Krit Punburananon, Taha Marhaba, and Ashish Borgaonkar, "Characterization of THMs precursors by fractionation of natural organic matter and UV absorption" Poster presentation at The 235th American Chemical Society National Meeting, New Orleans, Louisiana, April 2008.

Watch your thoughts; they become words.
Watch your words; they become actions.
Watch your actions; they become habits.
Watch your habits; they become character.
Watch your character; it becomes your destiny.
Frank Outlaw

Remember, if you ever need a helping hand, you'll find one at the end of your arm.
As you grow older you will discover that you have two hands,
one for helping yourself and the other for helping others.
Audrey Hepburn

What you do speaks so loudly that I cannot hear what you say.
Ralph Waldo Emerson

Saying is one thing. Doing is another.
Montaigne

The secret of getting ahead is getting started.
The secret of getting started is breaking your complex overwhelming tasks
into small manageable tasks, and then starting on the first one.
Mark Twain

ACKNOWLEDGMENT

First of all, it is my pleasure to express my appreciation to Dr. Taha Marhaba, who not only served as my dissertation advisor but also introduced me a breathtaking research topic. I have learned tremendously over the course of my Ph.D. study at New Jersey Institute of Technology.

Furthermore, my deep indebtedness is given to Dr. Sima Bagheri, Dr. Hsin-Neng Hsieh, Dr. Methi Wecharatana, Dr. Lee Lippincott for participating in my committee and giving invaluable guidance beyond my academic pursuit.

Additionally, special thanks are also extended to Dr. Larisa Krishtopa and Dr. Yong Pu, from both of whom I have learned instrumental analysis and laboratory techniques crucial for my research.

Moreover, my acknowledgement to the Department of Civil and Environmental Engineering cannot be neglected. My accomplishment today would not have been possible without their financial support over five and a half years.

Last, but not least, I would like to express profound gratitude to my parents, siblings, and friends in Thailand for uncomplaining patience, perennial support, and enormous encouragement during my stay in USA.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES
(Continued)

LIST OF FIGURES
(Continued)

LIST OF ABBREVIATION

BOD	Biological oxygen demand
COD	Chemical oxygen demand
DBPs	Disinfection by-products
DOC	Dissolved organic carbon
DOM	Dissolved organic matter
Em	Emission
Ex	Excitation
GLR	General linear regression
HAAs	Haloacetic acids
HPIA	Hydrophilic acid
HPIB	Hydrophilic base
HPIN	Hydrophilic neutral
HPOA	Hydrophobic acid
HPOB	Hydrophobic base
HPON	Hydrophobic neutral
LRAA	Locational running annual average
MCL	Maximum contaminant level
MCLG	Maximum contaminant level goal
MRDL	Maximum residual disinfectant level
MRDLG	Maximum residual disinfectant level goals
MLR	Multiple linear regression
NOM	Natural organic matter

PC	Principle component
PCA	Principle component analysis
PCR	Principle component regression
POC	Particulate organic carbon
PRESS	Prediction sum of square
RAA	Running annual average
RMSE	Root mean square error
SFS	Spectral fluorescence signature
SUVA	Specific ultra violet absorbance
THMFP	Trihalomethanes formation potential
THMs	Trihalomethanes
TOC	Total organic carbon
TTHMs	Total Trihalomethanes
UV	Ultra violet
VIF	Variance inflation factor

CHAPTER 1

INTRODUCTION

1.1 History and Background

One hundred years ago, typhoid and cholera epidemics were common throughout American cities. Exposure to microbial contaminants such as bacteria, viruses, and protozoa (e.g., *Giardia lamblia* and *Cryptosporidium*) was likely the greatest remaining health risk management challenge for drinking water suppliers. Acute health effects from exposure to microbial pathogens are documented and associated illness can range from mild to moderate cases lasting only a few days to more severe infections that last several weeks. These cases may result in death for those with weakened immune systems (USEPA, 1998b; USEPA, 2006b; Krasner et al. 2006).

Over the past fifteen years, we have also learned that there are specific microbial pathogens, such as *Cryptosporidium*, that are highly resistant to traditional disinfection practices. In 1993, *Cryptosporidium* caused 400,000 people in Milwaukee to experience intestinal illness. More than 4,000 were hospitalized, and at least 50 deaths have been attributed to the disease. There have also been cryptosporidiosis outbreaks in Nevada, Oregon, and Georgia over the past several years (USEPA, 1998a; USEPA, 2006b). Therefore, disinfectants became an essential element of drinking water treatment because of the barrier they provide against waterborne disease-causing microorganisms.

Disinfection is a major factor in reducing these epidemics, and is an essential part of drinking water treatment today. It is a chemical process used in water systems to inactivate (or kill) pathogens (i.e., disease causing organisms) found in the source water (i.e., lake, river, reservoir, or ground water aquifer from which water is drawn and

treated). Disinfection through inactivation usually involves the use of disinfectants such as chlorine, ozone, and chlorine dioxide, and a combination of chlorine and ammonia (chloramines) may render many of these organisms harmless (Singer, 1999; Xie, 2007).

However, the disinfectants themselves can react with naturally-occurring materials in the water (e.g., humic and non-humic) to form unintended organic and inorganic by-products which may pose health risks (Bull, 1982). Disinfection by-products (DBPs) are formed when disinfectants (e.g., chlorine) react with NOM, and/or bromide/ iodide present in the source water. Different disinfectants produce different types or amounts of DBPs (Richardson et al. 2003; Sarai, 2006).

DBPs formed during disinfection with chlorine and chloramines are Total Trihalomethanes (TTHMs- chloroform, bromoform, bromodichloromethane, and dibromochloromethane) and haloacetic acids (HAA5 – monochloro -, dichloro -, trichloro -, monobromo -, dibromo -.) The amount of DBPs formation in drinking water can change from day to day, depending on the season, water temperature, amount of chlorine dosages, the amount of plant material in the water (Sharp et al. 2006).

Many of these DBPs have been shown to cause cancer and reproductive and developmental effects in laboratory animals. More than 200 million people consume water that has been disinfected in the United States. Because of the large population exposed, health risks associated with DBPs, even if small, need to be taken seriously. A major challenge for water suppliers is how to balance the risks from microbial pathogens and DBPs. It is important to provide protection from these microbial pathogens while simultaneously ensuring decreasing health risks to the population from DBPs (USEPA, 2006b).

1.2 Rules and Regulations for Disinfection By-products

The 1996 Safe Drinking Water Act (SDWA) amendments require USEPA to develop rules and standards for DBPs in drinking water. Published in December 1998, the Stage 1 Disinfectants/Disinfection By-products Rule (DBPR) required water systems to use treatment methods to reduce the formation of DBPs and to meet the standards. Maximum Contaminant Level (MCL), Maximum Contaminant Level Goal (MCLG) of TTHMs, HAA5, and others DBPs as well as their potential health effects are shown in Table 1.1 (USEPA, 1998b). In addition, MCLG for each of DBP and maximum residual disinfectant level goals (MRDLG) for each disinfectant listed in Tables 1.2 and 1.3, respectively (USEPA, 1998b).

While the Stage 1 DBPR is predicted to provide a major reduction in DBPs exposure, national survey data suggest that some customers may receive drinking water with elevated, or peak, DBPs concentrations even when their distribution system is in compliance with the Stage 1 DBPR. Some of these peak concentrations are substantially greater than the Stage 1 DBPR MCLs and some customers receive these elevated levels of DBPs on a consistent basis (Krasner et al. 2006). The Stage 2 DBPR sets new requirements provide for more consistent, equitable protection from DBPs across the entire distribution system and the reduction of DBP peaks. As in Stage 1, the Stage 2 DBPR focuses on monitoring for and reducing concentrations of two classes of DBPs: TTHM and HAA5. The concentrations of TTHM and HAA5 are monitored for compliance, but their presence in drinking water is representative of many other chlorination DBPs that may also occur in the water; thus, a reduction in TTHM and HAA5 generally indicates an overall reduction of DBPs.

The second provision of the Stage 2 DBPR is designed to address spatial variations in DBPs exposure through a new compliance calculation for TTHM and HAA5 MCLs. The MCL values remain the same as in the Stage 1. The Stage 1 DBPR running annual average (RAA) calculation allowed some locations within a distribution system to have higher DBPs annual averages than others as long as the system-wide average was below the MCL. The Stage 2 DBPR bases compliance on a locational running annual average (LRAA) calculation, where the annual average at each sampling location in the distribution system will be used to determine compliance with the MCLs (USEPA, 2006a).

The Stage 2 DBPR was released simultaneously with the Long Term 2 Enhanced Surface Water Treatment Rule (LT2) to address concerns about risk tradeoffs between pathogens and DBPs (USEPA, 2006b). The purpose of the LT2 rule is to reduce illness linked with the contaminant *Cryptosporidium* and other disease-causing microorganisms in drinking water. This rule applies to all public water systems that use surface water or ground water under the direct influence of surface water to treat reservoir discharge to inactivate 4-log virus, 3-log *Giardia lamblia*, and 2-log *Cryptosporidium*. These requirements are necessary to protect against the contamination of water that occurs in open reservoirs (USEPA, 2006b).

1.3 Objectives

Natural organic matter in source water quality is the major factor impacting DBPs formation following disinfection. The primary objective is to investigate the use of fluorescence spectroscopy for characterization of Trihalomethanes (THMs) precursors in

source water. The secondary objective is to investigate fluorescence spectroscopy coupled with multivariate statistical analysis as a rapid, inexpensive technique to determine the amount of THMs precursors and THMs Formation Potential (THMFP) of water samples. The outcome would help identify the potential water quality problems rapidly in advance and aid water utilities for source water management as well as drinking water treatment process optimization.

Table 1.1 MCL, MCLG, and Potential Health Effect of DBPs (USEPA, 1998b; USEPA, 2006a)

Disinfection by-products	**MCL (mg/L)**	**MCLG (mg/L)**	**Potential Health Effects**
Total trihalomethanes	0.080	N/A	Liver, kidney or central nervous system problems; increased risk of cancer
Haloacetic acids	0.060	N/A	Increased risk of cancer
Bromate	0.010	zero	Increased risk of cancer
Chlorite	1.0	0.8	Anemia; infants & young children: nervous system effects

Table 1.2 MCLG for Each DBP (USEPA, 2006a)

Disinfection by-products	MCLG (mg/L)
TTHMs	
• Chloroform	0.07
• Bromodichloromethane	zero
• Dibromochloromethane	0.06
• Bromoform	zero
HAA5	
• Monochloroacetic acid	0.07
• Dichloroacetic acid	zero
• Trichloroacetic acid	0.02
• Bromochloromethane	N/A
• Dibromochloromethane	N/A

Table 1.3 MRDLG for Disinfectants (USEPA, 1998b)

Disinfectant residual	MRDL (mg/L)	MRDLG (mg/L)	Potential Health Effects
Chlorine	4 (as Cl_2)	4	Eye/nose irritation; stomach discomfort.
Chloramines	4 (as Cl_2)	4	Eye/nose irritation; stomach discomfort, anemia
Chlorine dioxide	0.8	0.8	Anemia; infants & young children: nervous system effects

CHAPTER 2

LITERATURE REVIEW

2.1 Current Research

Current research has been focusing on the characterization of NOM by varied techniques to identify THMs precursors. NOM in natural water is a complex mixture of various hydrocarbon structures with attached functional groups (Leenheer and Croué, 2003), and its composition varies through out the year depending on the source of organics and environmental conditions (Rodriguez et al. 2004; Sharp et al. 2006). For this reasons, quick and accurate methods for water characterization must be developed to help water treatment plants cope with the dynamics of changing source water quality while maintaining drinking water quality standard. A variety of water characterization techniques and prediction models were reviewed in this chapter.

2.2 Organic Carbon Analysis

Total organic carbon (TOC) is an aggregate measurement used to quantify the presence of organic matter in aquatic systems. TOC can be divided into two fractions as dissolved organic carbon (DOC) and particulate organic carbon (POC). POC is the fraction of the TOC that is retained on a 0.45 micrometer (μm) porosity membrane (Leenheer and Croué, 2003). DOC is the organic carbon smaller than 0.45 μm in diameter. POC generally represents a minor fraction (below 10%) of the TOC. The proportion of POC increases with a river's size and flow rate, and DOC concentrations range from 0.1 milligrams per liter (mg/L) in groundwater to 50 mg/L in bogs (Thurman, 1985).

However, Chow et al. (2005) recommended that even a smaller pore size (0.10 μm or smaller) could be used. The operationally defined DOC fraction using 0.45 μm filter often contains heterogeneous organic carbon compounds that may lead to inconsistent results when evaluating THMFP. A finer pore size filter provides more homogeneous DOC properties and enables a better characterization of organic matter.

DOC is a complex mixture of aromatic and aliphatic hydrocarbon structures that have attached amide, carboxyl, hydroxyl, ketone, and various minor functional groups ranging in molecular weight from a few hundred to 100,000 Dalton. Heterogeneous molecular aggregates in natural waters increase DOC complexity (Leenheer and Croué, 2003). DOC concentrations not only depend on the nature of the watershed, but are also influenced by seasonal variations and POC inputs such as runoff or algae bloom. Most of the NOM is considered to be slowly biodegradable (Thurman, 1985; Rodriguez, 2004; Toroz and Uyak, 2005).

2.3 Fractionation of Dissolved Organic Matter by Resin Adsorption

Dissolved Organic Matter (DOM), measured as DOC, is commonly characterized by fractionation into chemically distinct categories via resin adsorption. This technique has been widely used to isolate hydrophobic fraction (humic substances e.g., humic and fulvic acids) from hydrophilic fraction (non-humic substance). This is the basis of a simple DOM analysis that determines the humic/nonhumic distribution of raw and treated waters. This isolation technique was first developed by Leenheer (1981) and later modified by Marhaba et al. (2003) to fractionate low level of dissolved organic carbon normally found in natural water into six chemically distinct fractions.

The six fractions of organic matter dissolved in water are classified by their physical properties and chemical structures as follows (Bengraïne and Marhaba, 2003a):

- Hydrophilic base (HPIB) contains amphoteric proteinaceous materials containing amino acids, amino sugars, peptides and proteins.
- Hydrophilic acid (HPIA) includes organic compound containing the hydroxyl acid group.
- Hydrophilic neutral (HPIN) are organic compounds made up of polysaccharides.
- Hydrophobic base (HPOB) is the basic portion of the humic substance retained by XAD-8 resin at normal pH (~7), which can be eluted by hydrochloric acid.
- Hydrophobic acid (HPOA) is a soil fulvic acid.
- Hydrophobic neutral (HPON) is a mixture of neutral hydrocarbon and carbonyl compounds.

Hydrophobic fractions are separated from source water into HPOA, HPON, HPOB by XAD-8 resin adsorption at pH of 2, 7, and 10, respectively. HPIB and HPIA are separated by adsorption on AG-MP-50 and WA-10 resin, respectively, at pH of 2. The last fraction left in the effluent of WA-10 resin is HPIN. The flow diagram of fractionation procedure is shown in the Figure 2.1.

2.4 Ultra Violet Absorption

Chromophores have the functional groups containing the electrons that are promoted when a molecule absorbs light (¶ → ¶*). In NOM molecules, the vast majority of the chromophores that absorb in the Ultra Violet (UV) region ($\lambda <$ 400 nm) are aromatic groups with various degrees and types of substitution, including mono- or poly-substituted phenols and various aromatic acids. Theses chromophores are associated primarily with the humic fraction of the NOM (Korshin et al. 1997). Surface water

absorption of both visible and UV light is widely attributed to the conjugated (¶) bonds present in NOM molecules, primarily humics, dissolved in the water.

Specific UV absorbance (SUVA or $SUVA_{254}$) is defined as the sample's UV absorbance at 254 nm divided by the DOC concentration of the solution. High SUVA waters are generally enriched in hydrophobic NOM, such as humic substances (Leenheer and Croué, 2003). Therefore, SUVA is a measure of the aromaticity of DOC and can be used to estimate the chemical nature of the DOC at a given location. Most research has focused on UV absorbance at 254 nm, which serves as a rough indicator of overall NOM concentration (Korshin et al. 1997; Leenheer and Croué, 2003).

The most important application of $SUVA_{254}$ is in water treatment process. It is used as criteria to determine if water utility will be required to use enhanced coagulation. Stage 1 DBPR mandates the use of enhanced coagulation or enhanced softening during water treatment to reduce TOC from water to reduce DBPs formation (USEPA 1998b; USEPA, 1999; Croué et al. 2000).

$SUVA_{254}$ is an average absorptivity for all the molecules that comprise the DOC in a water sample and has been used as a surrogate measurement for DOC aromaticity. The assumptions behind these criteria are that $SUVA_{254}$ is a good indicator of the humic fraction of the DOC, and coagulation is effective at removing the humic fraction which is DBPs precursors (Weishaar et al. 2003).

However, NOM in natural water is a complex mixture of various hydrocarbon structures that have attached functional groups. $SUVA_{254}$ of NOM is typically broad and nearly featureless, because the number of possible types of NOM is large and none possess an easily distinguishable spectrum (Leenheer and Croué, 2003).

$SUVA_{254}$ was strongly correlated with percent aromaticity in water sample with R^2 of 0.97, but it showed poor relationship with THMFP of neither filter whole water (R^2 = 0.54) nor isolate organic matter (R^2 = 0.41). This implied that the reactivity of NOM with chlorine was not strongly dependent on the aromaticity of the organic matter (Weishaar et al. 2003).

It should be noted that not all aromatic molecules react with chlorine to produce chloroform. Thus, some compounds that contribute to the overall UV absorbance may be inert with respect to THMFP, and some THMs precursor compounds may not contribute to the overall UV absorbance (Weishaar et al. 2003). In addition, Ates et al. (2007) also reported low SUVA water sample having high THMs formation and the organics in those samples had low molecular weight and low aromatic structure.

2.5 Fluorescence Spectroscopy

2.5.1 Background and Theory

Synchronous florescence spectroscopy, a novel developing technique to obtain the summation of fluorescence spectra at wide range of excitation wavelengths, provides much better sensitivity compared to SUVA (Leenheer and Croué, 2003). This fluorescence spectrum is a three-dimensional (e.g., Excitation (Ex) wavelength vs Emission (Em) wavelength vs Relative Intensity of response surface) having a wide range of Ex/Em matrix. This florescence spectrum is a "fingerprint" of natural organic compounds present in the water sample, called a Spectral Fluorescence Signature (SFS). Obtaining fluorescence spectrums are straightforward, but interpreting the data is still under research and being explored.

Molecules could be excited by absorbed radiation from high energy light source. These molecules are transferred from ground state to the excited state and then emitt the excess energy as radiation as they drop back to the ground state as shown in Figure 2.2 (Bourne, 2006). The radiation emitted is called molecular fluorescence (Robinson, 1995; Skoog et al. 1998). The radiation emitted or the fluorescence is of lower energy than the stimulating rations, and must be at a longer wavelength (Robinson, 1995; Skoog et al. 1998; Kebbekus and Mitra, 1998). The molecule absorbs at characteristic wavelength and emits a spectrum which is also a characteristic of the compound. In general fluorescence is a very sensitive technique, and its detection limits are typically an order of magnitude better than in UV absorption (Kebbekus and Mitra, 1998).

The most intense and the most useful fluoresecence is found in compound containing the aromatic functional group with low energy transition levels. Compounds containing aliphatic and alicyclic carbonyl structures or highly conjugated double-bond structures may also exhibit fluorescence, but the number of these is small compared with the number in the aromatic systems (Skoog et al. 1998).

Most un-substituted aromatic hydrocarbons can fluoresce in solution. The simple heterocyclics, such as pyridine, furan, thiophene, and pyrene do not; on the other hand, fused ring structure ordinary do. Substitution on benzene ring causes shifts in the wavelength of absorption maxima and corresponding changes in the fluorescence peaks. Influence of halogen substitution is striking. The decrease in fluorescence with increasing halogen atom was observed by Skoog et al. 1998.

The high energy light source is used to excite the samples. Mercury lamps are used for line source excitation and deuterium or xenon arc sources for continuous source.

A monochromator is used to select the wavelength for excitation and emission. The sample is irradiated by UV light, which is absorbed by the molecules. The molecules become excited and remit the fluorescence in all directions. The fluorescence intensity is measured at right angles to the radiation that excited the sample (Robinson, 1995). The simplified schematic diagram of a fluorescence spectroscopy is shown in Figure 2.3 (Bourne, 2006).

The organic molecule adsorbs at characteristic wavelengths and emits a spectrum which is characteristic or the finger print of the organic compound (Marhaba, 2000). This spectrum is called spectral fluorescence signature (SFS) and also called emission-excitation matrix (EEM), which is defined as is the total sum of emission spectra of a sample at different excitation wavelengths recorded as a matrix of fluorescent intensity in coordinate of excitation and emission wavelengths. This spectrum is a fingerprint for each organic compound (Marhaba and Lippincott, 2000; Marhaba and Pu, 2000).

Fluorescence intensities of SFS predominantly depend on the organic matter concentration, provided that other factors that affect fluorescence intensity (pH, metal ion interaction, etc.) remain relatively constant (Baker, 2001). Ionic strength in the range of 0-1 M KCl and humic substance concentration in the range of 5-100 mg/L had little effect on the fluorescence spectral characteristics of the humic substances, while pH caused a shift of peak locations. Absorbance correction was shown to be essential for accurate representation and comparison of the SFS of the humic substances at high concentrations (Mobed et al. 1996).

2.5.2 Relationship between Fluorescence Intensity and Concentration

The relationship between fluorescence intensity and sample concentration was reviewed by Robinson (1995) and Skoog et al. (1998). The intensity of fluorescence, F, is proportional to the radiant power of the excitation beam that is absorbed by the system. That is:

$$F = K' (P_o - P) \tag{2.1}$$

where P_o is the power of the beam incident upon the solution and P is its power after traversing a length, b, of the medium. The constant K' depends upon the quantum efficiency of the fluorescence process. In order to relate F to the concentration, c, of the fluorescing species, we write Beer's law in the form

$$\frac{P}{P_0} = 10^{-\varepsilon bc} \tag{2.2}$$

where ε is the molar absorptivity of the fluorescing molecules and εbc is the absorbance, A. By substitution of Equation (2.2) into Equation (2.1), we obtain

$$F = K'P_0(1 - 10^{-\varepsilon bc}) \tag{2.3}$$

The exponential term in Equation (2.3) can be expanded as a Maclaurin series to

$$F = K'P_0 [2.303\varepsilon bc - \frac{(2.303\varepsilon bc)^2}{2!} + \frac{(2.303\varepsilon bc)^3}{3!} ....] \tag{2.4}$$

Provided $2.303\varepsilon bc = A < 0.05$, all of the subsequent terms in the brackets are small with respect to the first; under these conditions, the maximum relative error caused by dropping all but the first term is 0.13%. Thus, we may write

$$F = 2.3\ K'\ \varepsilon bc\ P_o \tag{2.5}$$

or at constant P_o,

$$F = Kc \tag{2.6}$$

Thus, a plot of the fluorescence power of a solution versus concentration of the emitting species should be linear at low concentrations. When c becomes great enough so that the absorbance is larger than about 0.05, the higher order terms in Equation (2.4) become important and linearity is lost. Mobed et al. (1996) reported that the humic acid concentration in the range of 5-100 mg/L had little effect on the fluorescence spectral characteristics. Although less than a 1% of the aromatic moieties in NOM actually emit light as fluorophores, three-dimensional SFS is an attractive analytical tool because it is at least an order of magnitude more sensitive to NOM than UV absorbance (Leenheer and Croué, 2003).

2.6 Application of Fluorescence Spectroscopy

2.6.1 Applications of SFS for Dissolved Organic Matter Characterization

Synchronous fluorescence spectroscopy provides highly detailed information that can be used to identify fluorescent compounds present in complex mixtures. Fluorophores are associated with the humic portion of NOM (Coble, 1996). Two distinct classes of fluorophores are generally discussed, the humic-like fluorophores and the protein-like fluorophores. Each type of organic compounds has a unique range of excitation and emission wavelength. Coble (1996) reported a variety of marine organic compounds characterized by fluorescence spectroscopy in the Table 2.1. The composition of organics in water is depending on the origin of the water samples.

Dissolved organic contents were fractionated by Marhaba (2000) into six fractions: HPOA, HPOB, HPON, HPIA, HPIB, and HPIN using three types of resins: XAD-8, AG-MP-50, and WA-10. The SFSs of six fractions are summarized in Table 2.2

(Marhaba, 2000). This showed that each six fractions has a unique SFS, and this information can be used to characterize organic fractions in water.

Chen et al. (2003) characterized standard Suwannee River fulvic acid and 15 hydrophobic or hydrophilic acid, neutral, and base DOM fractions by fluorescence spectroscopy. The peaks for each organic substance on the Ex/Em matrix are presented in Figure 2.4. The matrixes are divided in to five regions: Region I Aromatic protein, Region II Aromatic protein II, Region III Fulvic acid like, Region IV Soluble microbial by-product-like, Region V Humic acid-like (Chen et al. 2003).

2.6.2 Applications of SFS for Characterization of Waste Water Contamination to Natural Water

SFS is a fingerprint of organic compounds, and then it could be potentially used to identify the source of organics in water. Ma et al. (2001) isolated DOM from natural water by XAD-8 resin (Supelco, Belfonte, PA) and found that the fulvic acid fractions predominated in natural waters and accounted for 54-68% of the total amount of dissolved organic carbon, whereas the humic acid and hydrophilic fractions constituted, respectively, 13-29% and 9-30% of the total DOC. In addition, they also reported that the effluent of wastewater was almost devoid of humic acid and the hydrophilic fractions fraction exceeded fulvic acid. These differences in composition of each organic fraction could be characterized by shape and intensity of the SFS.

Recent advances in fluorescence spectroscopy have started to differentiate potential sewerage-related river pollutants from natural water sources. Baker (2002b) has demonstrated the use of fluorescence excitation-emission matrix to differentiate organic waste from NOM. Fluorescence spectroscopy was applied to five neighboring rivers, including one that is impacted by wastewater from a large tissue mill. Two samples were

taken from the tissue mill effluent and the impacted river. These two samples with significantly higher fluorescence intensity were dominated by tryptophan fluorescence and a fluorescence center possibly due to the presence of fluorescent whitening agents from tissue mill effluent. In contrast, the three other samples exhibited lower fluorescence intensities typical of river systems with tryptophan (sewage), humic-like (peat derived color), and fulvic-like NOM.

Beside contamination of industrial waste, the fluorescence spectroscopy can identify the contamination of non-point source farm or agriculture waste. Baker (2002a) applied this technique to investigated silage liquor, pig and cattle slurry, and sheep barn waste. All farm wastes exhibited high intensities of fluorescence that can be attributed to the protein tryptophan. Silage liquor was characterized by a very high fluorescence intensity with an initial tryptophan : fulvic-like fluorescence intensity ratio of >20. Cattle and pig slurries exhibited a lower tryptophan : fulvic-like fluorescence intensity ratio (~2–5) and lower tryptophan fluorescence intensity, and tyrosine fluorescence was also observed. Sheep barn wastes had the lowest tryptophan : fulvic-like fluorescence intensity ratios (~0.5–4.0). The ratios of tryptophan : fulvic-like fluorescence intensity observed from the farm wastes investigated are significantly higher than those observed in the majority of river waters, suggesting that farm waste pollution events could leave a signature in river waters due to their distinctively high protein fluorescence intensity.

In addition, Baker (2003) also applied fluorescence spectroscopy to identify organic contamination from point-source municipality sewer discharges. It was shown that samples contaminated with sewage had both high tryptophan fluorescence intensity

and ammonia. The highest levels of tryptophan and ammonia were observed in the summer, probably because of the low base flow in the river during this time of the year.

These evidences suggest that SFS is a useful tool for detecting, monitoring, and controlling the impact of organic waste contamination from industrial sources, point sources, and non-point sources for the watershed of study (Baker, 2002a; Baker, 2002b; Baker, 2003).

2.6.3 Application of SFS for Characterization of Waste Water

Fluorescence analysis has shown a potential use for sewage treatment plants in an attempt to provide an on-line monitoring method superior to that of using UV absorption at 254 nm. Reynolds (2002) showed that the fluorescence emission spectra (at Ex wavelength = 280 nm and Em wavelength = 350 and 440 nm) of sewage samples were observed in all samples. Fluorescence data at 280/350 Ex/Em wavelength correlated well with corresponding BOD, COD and TOC values (R^2 values ranging between 0.93 and 0.98). Moreover, fluorescence intensity at 280/440 Ex/Em wavelength has been demonstrated to correlate with non-biodegradable dissolved organic matter (COD-BOD). Within the treatment process, untreated sewage was shown to have higher fluorescence intensity than treated sewage.

The fluorescence spectra of domestic waste water, treated effluent of a waste water treatment plant, and receiving river water were analyzed by Lee and Ahn, 2004. Ex/Em wavelengths of 220/350 nm and 270/350 nm for protein-like fluorescence and 240/450 nm and 340/450 nm for humic-like fluorescence were suggested as fluorescence peaks. The protein-like fluorescence peaks showed better correlation between COD values and fluorescence intensities than the humic-like fluorescence peaks.

Baker and Inverarity (2004) studied the relationship between tryptophan-like fluorescence intensity and chemical water quality of river water from the River Tyne, NE England. The results showed statistically significant relationships between nitrate, phosphate, ammonia, BOD and dissolved oxygen. Tryptophan-like fluorescence intensity at the 280 nm excitation/350 nm emission wavelength fluorescence center correlated with both phosphate ($r = 0.80$) and nitrate ($r = 0.87$), whereas tryptophan-like fluorescence intensity at the 220 nm excitation/350 nm emission wavelength center correlated with BOD ($r = 0.85$), ammonia ($r = 0.70$) and dissolved oxygen ($r = -0.65$). The strongest correlations were between tryptophan-like fluorescence intensity and nitrate and phosphate (Baker and Inverarity, 2004).

2.7 Disinfection By-Products

2.7.1 Disinfection By-products Formation and Their Behaviors

The disinfection of water using chlorine is a common practice used to destroy microorganisms and to ensure the capability to maintain disinfection by carrying a residual chlorine concentration in distribution systems, thus protecting water from microorganism re-growth. Chlorine is extremely efficient and relatively inexpensive. However, chlorine reacts with the NOM contained in natural water and generates THMs and other DBPs (Sarai, 2006).

The operational parameters that influence THMs formation in the distribution systems are chlorine dose, water temperature, pH, and exposure to disinfectant contact time of water within the system. It has been concluded that, in terms of water quality, the fulvic and humic fractions of organic matter constitute major precursors to THMs

(Christian et al. 1990). The relative contributions to THMs production come from the humic fraction rather than fulvic fraction since the former reacts more readily with chlorine.

According to Milot et al. (2000), the rate of THMs production increases three-time per unit pH. THMs level was higher with increasing temperature. The rate of THMs formation at 2°C was normally 60-70% less than observed at 22°C (Knocke et al. 1986). Hence, THMs levels in summer and fall were, on average, about five times higher than in winter, whereas average HAAs in spring were about four times higher than in winter. THMs increased and stabilized in the most distant regions of the distribution system, whereas HAAs begin to increase, and then decrease mainly due to a reduction of dichloroacetic acid. This decrease was significantly higher in warm waters (higher than 15 °C) than in cold waters (lower than 15 °C), which led to the hypothesis of microbial degradation of HAAs as water approaches the system extremities (Rodriguez et al. 2004).

Kim and Yu (2005) reported that THMFP was highly influenced by the hydrophobic fraction, whereas haloacetic acid formation potential (HAAFP) depended more on the hydrophilic fraction. However the hydrophobic fraction is removed more than the hydrophilic fraction through conventional water treatment (Marhaba and Van, 2000; Sharp et al. 2006). Therefore residual hydrophilic NOM after conventional treatment needs to be removed to reduce HAA formation (Kim and Yu, 2005).

2.7.2 Kinetic of Trihalomethanes Formation

Panyapinyopol et al. (2005) fractionated DOM into six fractions and reported that that the main organic matter components were HPOA and HPIN, which were also the two most important contributors of THMFP. HPOB and HPIB were the most active precursors, and

these two base fractions also provided relatively high SUVA. Kinetic study of THMs formation for six fractions showed linear dependencies between the level of each organic fraction and the formation potential of THMs, which suggests the reactions between the organic fraction and chlorine during the chlorination are first order.

The chlorination of HPON and HPIN fractions led to the formation of mostly chloroform, while other organic fractions formed both chloroform and bromodichloromethane. The chloroform formation was generated mostly from the HPOB fraction, while bromodichloromethane formation was mainly from HPIB. The chlorination of an individual organic fraction resulted in a higher level of THMFP than that of the raw water and mixed fractions, indicating an inhibitory effect between the organic species (Panyapinyopol et al. 2005a,b).

In addition, the relationship of NOM molecular weight and chlorine demand to THMs formation was also reported by Gang et al. (2003). THMs formation increased as a function of chlorine consumption. A larger number of conjugate bounds in NOM fraction lead to high chlorine demand for oxidation and substitution thus yielding higher THMs formation. As the molecular weight of NOM decreased, THMs yield increased. The possible reason is that the halogenated intermediates formed from the smaller molecular weight DOCs decompose easily, which could favor formation of more THMs.

2.7.3 Reduction of Trihalomethanes Formation

THMs can be reduced by removing their precursors, the NOM in reclaimed industrial waste water. Musikavong et al. (2005) studied the removal of NOM by coagulation process using alum and ferric chloride. Jar-test experiments were conducted using separate alum and ferric chloride dosages from 10 to 80 mg/L at pH conditions from 5 to

6.5. The obtained results showed that TOC was gradually reduced from the average value of about 6.1 mg/L to a level of about 4.0 mg/L by alum and ferric chloride dosages of approximately 40 mg/L. Moreover, DOC were reduced from an average value of 5.1 mg/L to a level of about 4.0 mg/L by alum and ferric chloride dosages of approximately 40 mg/L.

In addition, chlorine demands at one-day reaction were the same as those of 7-day demands with a correlation coefficient of 0.98 (n = 10, correlation significant at the 0.01 level). Maximum THMFP percentage removal of 25% and 28% by using alum and ferric chloride dosages of about 80 mg/L at pH 5.5 and 5 were obtained, respectively (Musikavong et al. 2005).

2.7.4 Prediction of Trihalaomethanes Formation

Edzwald et al. (1985) first introduced UV as surrogate parameters for TOC and THMs. General linear regression (GLR) models were developed from samples collected at Grasse River and Glenmore Reservoir in New York State. However, these models were season and watershed empirical relationships. The models from two different water sources were different as shown below.

Models for Grasse River:

$$TOC = 19.8\,(UV) + 0.65 \qquad R^2 = 0.93 \tag{2.7}$$

$$THMs = 1573.2\,(UV) + 38.2 \qquad R^2 = 0.94 \tag{2.8}$$

$$THMs = 75.8\,(TOC) + 13.6 \qquad R^2 = 0.91 \tag{2.9}$$

Model for Glenmore Reservoir:

$$TOC = 28.1\,(UV) - 0.3 \qquad R^2 = 0.71 \tag{2.10}$$

$$THMs = 1596.1\,(UV) + 49.3 \qquad R^2 = 0.82 \tag{2.11}$$

$$THMs = 45.1\ (TOC) + 113.4 \qquad R^2 = 0.72 \tag{2.12}$$

Additionally, THMs prediction models could also be developed using chlorine incubation conditions as predictor variables. The general equation is usually in the forms of exponential power function (Sohn et al. 2004).

$$THMs = k(DOC)^a(Br^-)^b(Temp)^c(Cl_2)^d(pH)^e(time)^f \tag{2.13}$$

The independent variables (i.e., DOC, Br^-, Temp, Cl_2, pH, and time) correspond to DOC (mg/L), bromide concentration (μg/L), temperature (^oC), chlorine dose (mg/L), chlorination pH, and chlorine reaction time (hr) respectively, and the exponent k, a, b, c, d, e, and f are empirical constant. It is noted that the model for coagulated water from treatment plant or distribution system may not include pH and temperature, as they are controlled (Sohn et al. 2004). The THMs prediction models for raw water and coagulated water were developed by Amy et al. (1998). For raw water, the empirical expression is:

$$THMs = 10^{-1.385}(DOC)^{1.098}(Cl_2)^{0.152}(Br^-)^{0.068}(Temp)^{0.609}(pH)^{1.601}(Time)^{0.261},$$

$R^2 = 0.90.$ (2.14)

For coagulated water, the empirical expression is:

$$THMs = 3.296(DOC)^{0.801}(Cl_2)^{0.152}(Br^-)^{0.261}(Time)^{0.264},\ R^2 = 0.87 \tag{2.15}$$

Uyak et al. (2005) showed that among measured parameters both pH (r = 0.963) and temperature (r = 0.921) were found to be the parameters of the highest statistical significance as predictors for THMs formation.

$$THMs = 7.07x\ 10^{-2}\ (TOC + 3.2)^{1.314}\ (pH - 4.0)^{1.496}\ (Dose - 2.5)^{-0.197}\ (Temp + 10)^{0.724}$$

$R^2 = 0.986$ (2.16)

Toroz and Uyak (2005) showed that in the water distribution network, THMs levels correlated with TOC, temperature and Cl_2 residue. The regression model for THMs was developed as follows:

$$THMs = 11.967\ (TOC)^{0.398}\ (Temp)^{0.158}\ (C1)^{0.702}, \qquad R^2 = 0.909 \qquad (2.17)$$

In addition, UV can be used interchangeably with DOC in the model (Amy et al. 1998).

$$THMs = 0.42(UV)^{0.482}(Cl_2)^{0.339}(Br^-)^{0.023}(Temp)^{0.617}(pH)^{1.609}(Time)^{0.261},\ R^2 = 0.71 \qquad (2.18)$$

2.8 Prediction Models Using SFS

2.8.1 Univariate Regression Model

DOM was fractionated by Goslan et al. (2004) using resin adsorption techniques into hydrophobic (fulvic and humic acid fractions) and hydrophilic (acid and non-acid fractions) components. The pH of each sample was adjusted to seven. The raw water was diluted by deionized water to five concentrations and the SFSs of samples at varied concentrations were used for model calibration. The prediction of hydrophilic fraction concentrations may be less accurate due to their lower intensity of fluorescence when compared with the more hydrophobic fractions. The predicted values of the hydrophobic concentrations were within 10% error for a particular source water of study. However, the model was unable to predict fraction concentrations for waters from a different watershed (Goslan et al. 2004).

2.8.2 Multiple Linear Regression

In many applications, one response from an instrument is related to the concentration of a single chemical component. This is referred to as univariate calibration because only one instrument response is used per sample. Multivariate calibration is the process of relating

multiple responses from an instrument to properties of a sample (Babee et al. 1998). The samples could be, for example, a mixture of chemical components and the goal is to predict the concentration levels of the different chemical components from their SFS.

Multivariate calibration offers several advantages over univariate approaches (Babee et al. 1998):

- It is possible to analyze for multiple components simultaneously.
- Multiple redundant measurements can also provide improved precision in prediction. Statistics show that repeating a measurement n times and calculating a mean value result in a factor of $\sqrt{n}$ reduction in the standard deviation of the mean.
- Redundant measurements in multivariate calibration also facilitate fault detection or the presence of interference.

The general linear regression is defined as

$$Y_n = b_o + b_1X_1 + b_2X_2 + \ldots + b_mX_m + e \tag{2.19}$$

Or it can be written in matrix form as

$$\mathbf{Y}_{nx1} = \mathbf{X}_{nxm}\mathbf{b}_{mx1} + \mathbf{e}_{nx1} \tag{2.20}$$

Where **Y** is response e.g., TOC, THMFP; **X** is the fluorescence spectra; **e** is error; n is number of samples or observation; m is number of predictor variable. The regression coefficient matrix, **b**, can be solved by least square method (Kutner et al. 2004). There are three cases: $m > n$, $m = n$, and $m < n$ (Geladi and Kowalski, 1986).

(1) $m > n$. There are more predictor variables than samples. In this case, there are an infinite number of solutions for b to fit the system of equation.

(2) $m = n$. The number of samples and variable are equal. This gives a unique solution.

(3) $m < n$. There are more samples than variables. Regression coefficient can be solved.

Multiple Linear regression (MLR) is the process of calculating the line that best describes the correlation between variables. This method is based on the flowing assumptions (Tranter, 2000):

- The relationship is a straight line.
- The predictor variables are independent.
- The response variable, y, has error that is normal distributed.
- The error is independent from y and x.

Marhaba and Pu (2000) developed MLR model using such parameters as peak intensity, spectral slopes and areas of SFS. Isolation and fractionation of NOM into six hydrophobic and hydrophilic substances was performed on sampling locations within three water treatment plants in New Jersey. Fluorescence spectroscopy was performed on the original non-fractionated samples and fractionated samples.

Model expressions were:

$$C = -0.0074 + 0.0003548\ Area + 3.317\ Slope - 0.00445\ Slope * Area + 0.10875\ HPOA + 0.004\ HPOB + 0.16475\ HPON + 0.80225\ HPIA - 0.225\ HPIB \quad (2.21)$$

The R^2 was 0.924 (R^2 adjusted was 0.884), indicating good correlation between the dependent variable (i.e., DOC), and the independent variables (i.e., slope, area, and fractions' concentration) (Marhaba and Pu, 2000).

2.8.3 Principal Component Regression

One of the problems with MLR is that predictors (variable x) in the methods may be correlated with each other such as when using the whole spectra as predictors. The fluorescence spectra at different excitation/emission wavelength are usually highly correlated by nature, and then general linear regression become inapplicable. This

problem is called multicollinearity (Tranter, 2000), and it causes the regression coefficient not to have any effects of the predictor variable on the response but only a marginal or partial effect. The often-used index to justify co-linearity is the correlation coefficient (Kutner et al. 2004). If the correlation coefficients between predictor variables are more than 0.7, multicollinearity can be a problem (Field, 2005).

A possible solution to this problem is to generate the principal components (PCs) from highly correlated predictors. The PCs are guaranteed to be orthogonal to each other (a matrix whose transpose equals its inverse) and as they are formed from linear combinations of original variables. The orthogonal property means that the model is particularly simple with no interaction terms. In the other words, this technique is to transform a big set of correlated variables to a small set of uncorrelated principal components (Tranter, 2000; Brereton, 2003). This data transformation process is called Principal Component Analysis (PCA).

PCs are calculated in such a way that the first component explains as much information as possible, and each PC is independent. The number of PC included the reduced data set is determined by eigenvalue of that PC. The first PC has the highest eigenvalue (Liu et al. 2003; Brereton, 2003, Minitab, 2004). Based on Kaiser's Criteria, all PCs with eigenvalue more than one should be included in the reduced data set. However, others may suggest to retain the PCs with eigenvalues more than 0.7 (Field, 2005). The number of PCs is ideally equal to the number of significant components in the sample (Geladi and Kowalski, 1986; Brereton, 2003).

The main applications of PCA are to reduce the number of variables and to detect structure in the relationships between variables. General linear regression can be

developed on PCs instead of original predictor variables, X. This two-step regression (PCA and MLR) is called Principal Component Regression (PCR). The general form of PCR models is shown below.

$$Y = b_0 + b_1PC_1 + b_2PC_2 + ... + b_nPC_n \quad (2.22)$$

The principal component obtained from PCA is a function of the original predictor variable. Transform PC to X and the regression becomes:

$$Y = b_0 + b_1X_1 + ... + b_mX_m \quad (2.19)$$

PCR has advantages and disadvantages (Beebe et al. 1998; Thermo Electron Corporation, 2006a), as outlined below:

PCR Advantages:

- It does not require wavelength selection; usually the whole spectrum can be used.
- Larger number of wavelengths gives averaging effect, making model less susceptible to spectral noise.
- It can be used for very complex mixtures.
- It could sometimes be used to predict samples with constituents (contaminants) not present in the original calibration mixtures.

PCR Disadvantages:

- Calculations are slower than most classical methods.
- Although the number of dimensions is decreased, it is still possible to become inundated with plots. The utility of PCA is limited with data sets with high inherent dimensionality, because multiple pair wise or three-dimensional plots must be used to visualize the data.
- Optimization requires some knowledge of PCA; models are more complex to understand and interpret.
- No guarantee that PC vectors directly correspond to constituents of interest.

- Generally, a large number of samples are required for accurate calibration.
- Collecting calibration samples can be difficult; must avoid collinear constituent concentrations.

Table 2.1 Major Fluorescent Components in Excitation-Emission Matrix (Coble, 1996)

Component type	Range of excitation (nm)	Range of emission (nm)
Fulvic–like, Humic like	330-355	420-480
Humic like	250-260	380-480
Marine humic-like	310-320	380-420
Tyrosine-like, protein-like	270-280	300-320
Tryptophan-like, protein-like or phenol-like	270-280	320-350

Table 2.2 Characteristics of Six DOM Fractions' Peaks (Marhaba, T.F., 2000)

Fraction	Major peak location		Minor peak location	
	Ex (nm)	Em (nm)	Ex (nm)	Em (nm)
HPIA	225-237	345-357	273	357-369
HPIB	225-237	357-369	273-285	357-381
HPIN	225	609-621	-	-
HPOA	237-249	417-429	297-309	417-429
HPOB	225-237	369-381	273-285	369-381
HPON	225-237	309-321	-	-

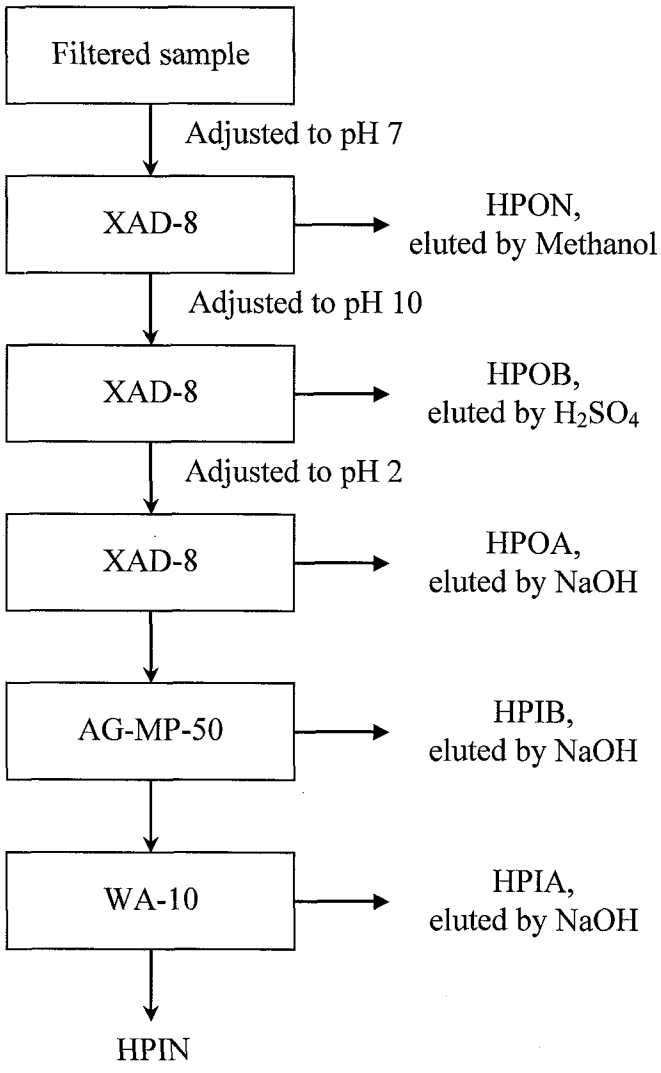


Figure 2.1 Flow diagram of fractionation procedure.

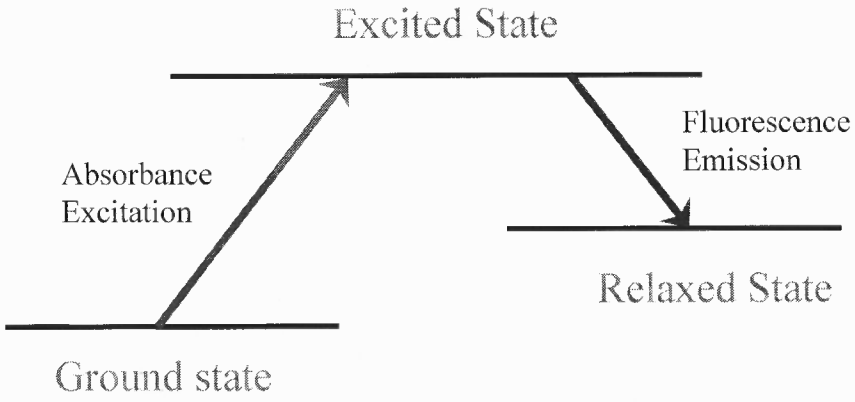


Figure 2.2 Absorption to an excited state and fluorescence emission.

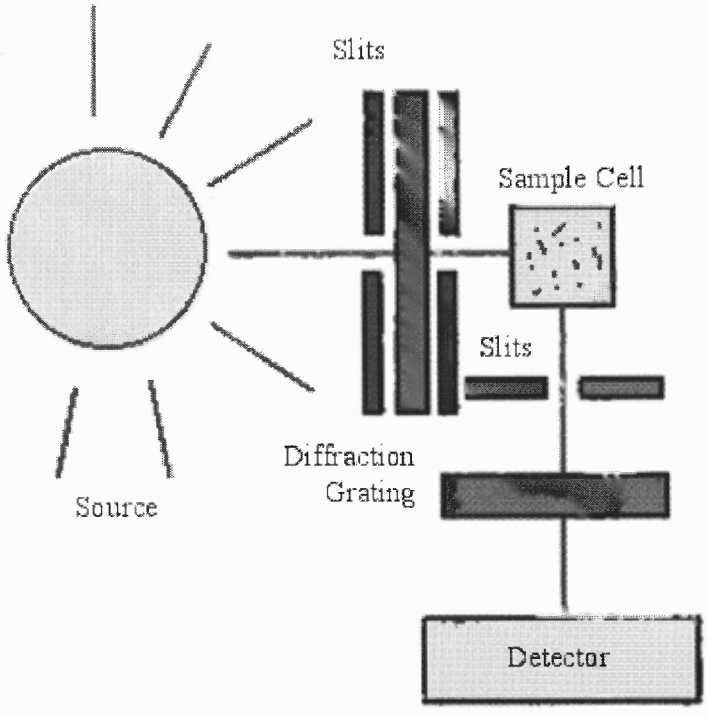


Figure 2.3 Schematic diagram of a fluorescence spectroscopy.

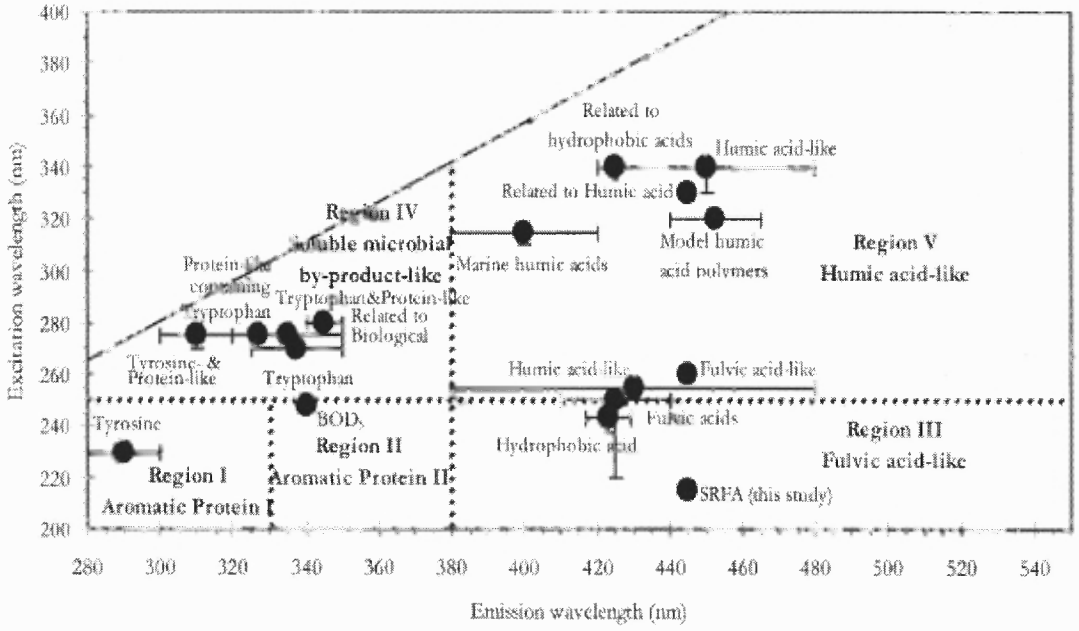


Figure 2.4 Location of SFS peaks on Ex/Em matrix. (Chen et al. 2003).

CHAPTER 3

EXPERIMENTAL

3.1 Research Overview

In this research, water samples were collected from various locations along the Delaware & Raritan Canal and its tributaries. The samples were characterized by TOC, resin adsorption, UV absorbance, and fluorescence spectroscopy. THMFP analyses were performed on all original samples and fractions. Multivariate statistical models were developed to predict the concentration of organic matter and potential of THMs formation using SFSs.

3.2 Samples Collection and Preparation

Water samples were collected between October 2006 and June 2008 following rain events by Middlesex Water Company and the NJIT research team. Figure 3.1 shows the sampling locations along Delaware & Raritan Canal and its tributaries. Samples were collected from locations that were suspected to be sources of problematic organics (i.e., outfalls) and from various locations along the canal. The samples were kept in the refrigerator at 4 °C and filtered through Nylon 0.45 μm membranes (Advantec MFS Inc., Pleasanton, CA) within 24 h of collection.

3.3 Samples Characterizations

Water samples were passed through 0.45 μm filter to remove organic particle before characterization. The experimental diagram is shown Figure 3.2.

3.3.1 Organic Carbon Analysis

DOC was determined via Phoenix 8000 TOC analyzer (Tekmar Dohrmann, Cincinnati, OH) using the UV/Persulfate oxidation method (Standard method 5310 C). The instrument error were controlled to within ± 4% RSD by running a 5 mg/L standard every fifth sample. Three replicates of samples were performed with a variation limit within ± 5%. All fraction samples were pH adjusted and diluted to reduce the contribution of eluent chemicals to DOC if necessary. Mili-Q water (Millipore Corp., Bedford, MA) was used for all dilutions, sample preparation, and final glassware washing.

3.3.2 UV Absorption

A double beam UV/Visible spectrophotometer, Varian DMS 300 UV/Visible spectrophotometer, (Varian Inc, Palo Alto, CA) was used in this study to measure UV absorbance at 254 nm (Standard method 5910 B). SUVA at 254 nm ($SUVA_{254}$) was determined as the sample's UV absorbance at 254 nm divided by the DOC of the sample.

3.3.3 Fractionation of Dissolved Organic Matter

The fractionation procedure was first developed by Leenheer (1981) and modified by Marhaba et al. (2003). The diagram of this complex time-consuming procedure is shown in Figure 3.3. The general procedure is discussed below.

The amount of XAD-8 (Superlite™ XAD-8, SUPELCO, Supelco Park, Bellefonte, PA) resin was determined according to Leenheer and Croué (2003) with a capacity factor of 50 (K'=50) and a porosity of 0.60. XAD-8 resin was intensively refined with 0.1N NaOH for 24 hour and sequentially extracted with acetone and hexane for another 24 h in a set of Soxhlet extraction apparatus. The XAD-8 resin was transferred into columns (2.5 cm×120 cm, Kontes, Vineland, NJ) in slurry of methanol. The packed

resin was rinsed with two times 2.5 bed volumes of 0.1N each NaOH first, then H_2SO_4, and finished with Milli-Q water until the conductivity and DOC of the effluents are below 10 μs/cm and 0.2 mg/l, respectively. This resin cleanup is necessary to eliminate any impurities brought during the resin manufacturing process.

Hydrophobic neutral (HPON) was the first fraction to be fractionated. The water sample pH was adjusted around 7±0.2 and then filtered by gravity through the XAD-8 resin bed with a flow rate less than 12 bed volumes/h. Sample solution constrained inside resin bed was displaced with 1 bed volume of Milli-Q water and discarded. The column was turned up side down and the resin is air-retrieved, stored, and dried in a desiccators. The HPON fraction was extracted with methanol which is removed later by rotary vacuum dryer.

The operation for hydrophobic base (HPOB) and hydrophobic acid (HPOA) is similar to that of HPON. The sample effluent after HPON was de-protonated to pH10 with 10 N NaOH, then loaded through the second XAD-8 resin column. The fraction was collected with 0.25 bed volume of 0.1N H_2SO_4, followed by 1.5 bed volumes of 0.01N H_2SO_4 at a flow rate less than 2 bed volumes/h, forming a total of 1.75 bed volumes of this fraction, HPOB. The effluent of the second XAD-8 resin column was acidified to pH 2 with concentrated H_2SO_4, loaded on the third XAD-8 resin column. Elution of HPOA was conducted using 0.25 bed volumes of 0.1N NaOH followed by 1.25 bed volumes of 0.01N NaOH at no great than 2 bed volumes/h. H_2SO_4 instead of HCl was used for pH adjustment due to the chloride interference with the UV/Persulfate oxidation of carbon. The removal of hydrophobic substances was concluded after runs of the three XAD-8 resin columns in series.

The hydrophilic base (HPIB) fraction was separated by AG-MP-50 cationic resin (BIO-RAD, Hercules, CA). This fraction was eluted by 1.0 N NaOH. The service flow rate and the resin regeneration are at no greater than 5 and 2 bed volumes/h, respectively.

Diaion WA 10 (SUPELCO), a weak anion exchange resin, was the final resin applied to isolate HPIA. Effluents after HPIB fractionation were put through the WA 10 resin for the HPIA. This fraction was eluted by 1.5 bed volume of 0.1 N NaOH and 0.5 bed volume of 0.01 N NaOH. The service flow and elution rates are 8 and 4 bed volumes/h, respectively.

3.3.4 Fluorescence Spectroscopy

The Hitachi F-4500 fluorescence spectrophotometer (Tokyo, Japan) equipped with 150 W ozone free Xenon lamp was used for the fluorescence measurements. The samples was kept in a one-cm quartz cuvette of four ml volume sample size and excited from 225 to 525 nm wavelengths with six-nm bandwidth. The SFS matrix consisted of fluorescence intensity responses recorded between Em wavelengths 231 to 633 nm also with six-nm bandwidth. This results in 51 Ex wavelengths x 68 Em wavelengths SFS matrix with a total of 3468 fluorescence intensities.

3.3.5 Trihalomethanes Formation Potential

Chorine demand test was carried out following standard method 2350 B to find the appropriate amount of chlorine dosing for disinfection of raw water and maintain the free chlorine residual at 3-5 mg/L. Samples were incubated at pH 7, in the dark, at 25 ^{o}C for seven days according to standard method 5710 B. THMs were measured by Varian GC-ECD 3400 using EPA standard method 551.1.

Standard concentrations of four THMs: chloroform ($CHCl_3$), bromodichloromethane ($CHBrCl_2$), dibromochloromethane ($CHBr_2Cl$), bromoform ($CHBr_3$), were prepared at 5, 10, 20, 50, 80, 100, 200, 500, 1000, 2000 μg/L. Highest purity of 99.99% Methyl *tert*-butyl ether (MTBE) (Sigma Aldrich, Inc) was the solvent used for extraction. Bromofluorobenzene and decafluorobiphenyl (Restek, Bellefonte, PA) were used as internal standard and surrogate, respectively. Other chemicals are reagent grade supplied by Thermo Fisher Scientific, Inc. The procedure of THMs measurement can be outlined as follows.

- The samples were adjusted to a pH of 7 ± 0.2 using H_2SO_4 and NaOH before incubation.
- One ml of pH-seven buffer solution was added to 68 ml incubation bottle followed by appropriate amount of standard sodium hypochlorite solution obtained previously from chlorine demand test. The incubation bottle was filled up completely with the sample.
- Samples were incubated in the dark at 25 °C for THMs formation potential.
- After seven days, ten ml of sample are removed from incubation bottle. The sample was dechlorinated using sodium sulfite (Na_2SO_3), and their pHs were adjusted to pH around 4.5 to 5.5 by the addition of buffer salt.
- Buffer salt for pH 4.5 to 5.5 was prepared by a mixture of two g Na_2HPO_4, 198 g KH_2PO_4, and 1.2 g Na_2SO_3. The salt was washed by methanol four times, acetone two times, and MTBE two times to remove the impurity.
- Add 50 μL of surrogate solution, three ml of MTBE, 20 g of Na_2SO_4 to the sample bottle. Shake vigorously for four minutes.
- THMs were extracted from the sample into MTBE phase. Two μL of MTBE were sampled and measured by Varian GC-ECD 3400.

3.4 Development of Prediction Models

Three-dimension fluorescence spectra were obtained to develop the prediction models for TOC and THMFP using multivariate statistical techniques performed by Minitab statistical software version 15 (Minitab Inc, State College, PA). Fluorescence intensities are highly correlated by nature. One assumption of MLR is that predictor variables must be independent or un-correlated (Kutner et al. 2004). Therefore, MLR cannot be performed directly to fluorescence spectra. Fluorescence intensities on the selected area were analyzed by PCA to transform a large number of fluorescence intensities into a few un-correlated PCs. The number of PCs that were used for further analyses were determined by their eigenvalues. The PCs with eigenvalues larger than one were kept, and the rest were discarded (Field, 2005).

Multi-linear regression was used to fit the best equations using the following criteria: R-square, adjusted R-square, Mallows Cp, and error standard deviation. The best possible models should have (1) high R-square, adjusted R-square, (2) Mallows Cp closest to number of predictor variables and constant, and (3) low prediction error. Using the similar process, DOC and THMFP prediction models were developed from SFS.

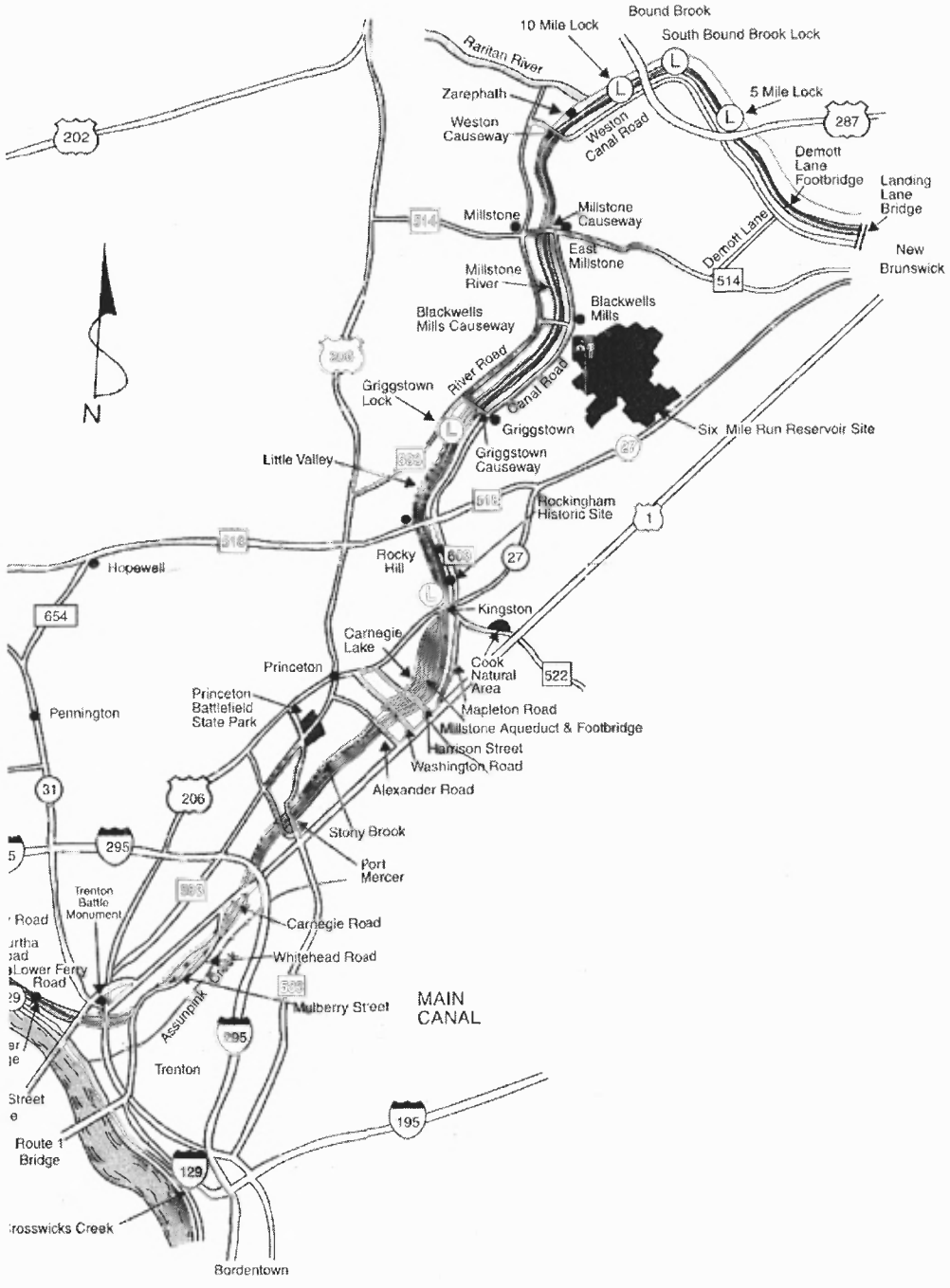


Figure 3.1 Sampling locations at Delaware & Raritan Canal and its tributaries. (Delaware & Raritan Canal State Park, 2008)

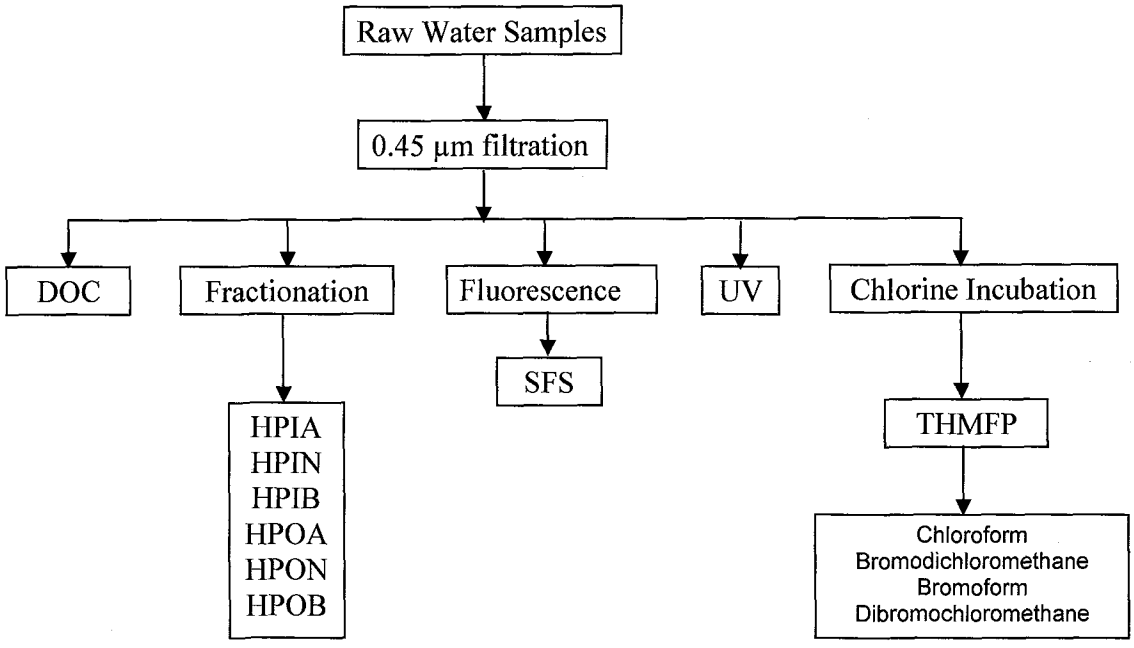


Figure 3.2 Experimental diagram.

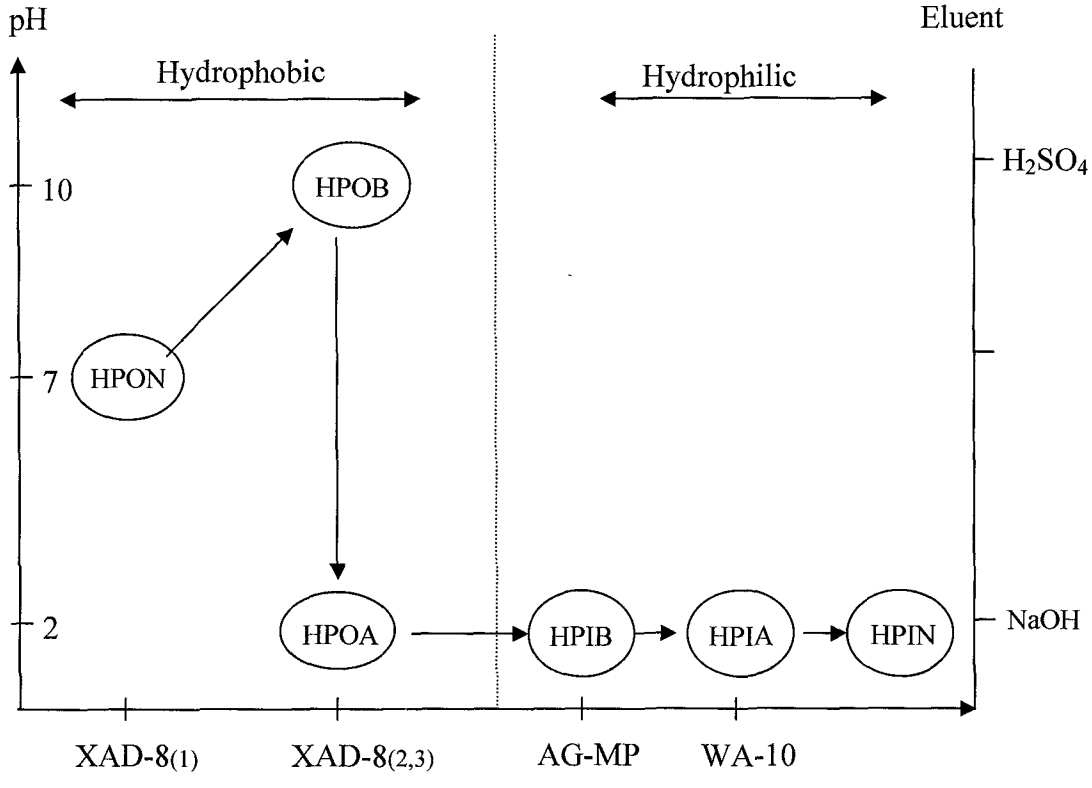


Figure 3.3 Fractionation procedures by resin adsorption (Pu, 2005).

CHAPTER 4

CHARACTERIZATION OF TRIHALOMETHANES PRECURSORS BY RESIN ADSORPTION

4.1 Introduction

DOM can be categorized into two groups: hydrophobic and hydrophilic, and each group can be divided into 3 sub-groups: acid, neutral and base. Theses six fractions are HPOA, HPON, HPOB, HPIA, HPIN, and HPIB. While DOM fractionation technique provides detailed information about types of DOM, it is expensive, time consuming and not appropriate for routine monitoring of problematic organics in source water.

In this study, samples were collected at various locks and locations along the canal, at outfalls or discharge to the canal near Blackwells Mills and Port Mercer, and also at marsh area near Port Mercer. The map of sampling area is shown in Figure 3.1. The outfall samples have larger DOC than the canal samples. For this reason, the outfall samples were diluted by the canal samples to obtain samples with a potential range of DOC possibly found in the canal.

For all samples, water quality parameters such as pH, conductivity, and DOC were determined. Additionally, they were characterized by UV absorption, fluorescence spectroscopy, and also incubated for determination of THMFP. However, due to time limitation, only a few samples were selected for characterization by resin adsorption. THMFP were measured by GC-ECD using USEPA method 551.1, and the ranges of samples' THMFP at various locations are shown in Figure 4.1.

4.2 Characteristics of Water Samples

Samples were collected from the canal, outfall and marsh area, but only fractionation samples are presented in this chapter. The water quality parameters are shown in Table 4.1(a) for samples collected at outfalls & marsh area and in Table 4.1(b) for samples collected from the canal.

Two outfalls along the canal were selected for samples collection: one at Blackwells Mills and the other at Port Mercer. The outfall at Port Mercer was larger, and the samples had higher DOC on the same day of collection. Port Mercer outfall then became the only outfall in later samples collections. The outfall samples basically contained runoff generated after rainfall event. Runoff pick ups soil contaminants such as petroleum, pesticides (in particular herbicides and insecticides), or fertilizers, if any, before storm water discharged to the canal.

As shown in Table 4.1(a), the outfall samples usually had DOC greater than 2 mg/L and SUVA greater than 2 L/mg-m. This indicated that water utilities may require the treatment of the raw water by enhanced coagulation (USEPA, 1999) when the water quality from the outfalls influences the canal. This causes the water utilities to use more chemicals for their treatment processes, and finally increase throughput of the water treatment. Figure 4.1 shows THMFP of the samples collected from the canal and outfalls. The outfall samples had higher THMFP (up to 910 μg/L) especially at Port Mercers. Depending on the flow rate, the discharge from the outfalls could dramatically increase THMFP of the water in the canal especially after rainfall event.

The samples collected from the marsh area near Port Mercer had substantial high DOC, conductivity, and THMFP compared to other samples. Although there are no

discharges from the marsh area directly to the canal, the samples gave us information about DOM at this location.

The six-fraction distribution is shown in Table 4.2 in terms of fraction concentrations and in Table 4.3 in forms of percentage using the subscript "a" for outfalls & marsh area samples, and the subscript "b" for water samples collected directly from the canal. These four tables show high spatial and temporal variation of six-fraction concentrations in the samples collected at this watershed. However, useful information could be drawn from this data. HPOA and HPON were the most abundant fractions in samples collected at the outfalls or marsh area at 34.82% and 27.50% on average, respectively, whereas HPIN (32.36% on average) was the highest fraction in the samples collected from the canal. HPOA contains natural fulvic and humic acids, and HPON contains tannin or lignin derived from plant degradation (Leenheer, 2004). HPIN could include proteins such as tyrosine or sugars (Chen et al. 2003; Leenheer, 2004).

A very high HPOA fraction in the samples collected at the outfalls suggested that the discharges could increase THMs formation of raw water, because HPOA is the major THMs precursor (Rook, 1975; Singer, 1999; Leenheer, 2004). However, humic fractions (hydrophobic fractions) are well removed by coagulation processes (Marhaba and Van, 2002; Weishaar et al. 2003).

4.3 Reactivity to THMs Formation

After resin elution, fractions concentration were diluted by Mili-Q water to 1 mg/L, and then incubated for determination of THMFP. The reactivity of each fraction to THMs formation is shown in Table 4.4. THMs formation could be reported into two units: μg

THMs per liter of sample (μg/ L) or μg THMs/ mg of organic carbon (μg/mg). In this case, the fractions had a made-up DOC concentration of 1 mg/L for consistency and comparison, and therefore THMs formation of the fraction in μg/mg and that in μg/L were the same in this case.

It should be noted that the results in Table 4.4 are different from the work reported by Panyapinyopol et al. (2005). The discrepancy could be from different laboratory techniques and sources of DOM in the samples. First, the samples were collected from different watershed and therefore the fractions could have different composition as well as reactivity to THMs formation. In addition, H_2SO_4 was used for elution in this work while HCl was used by Panyapinyopol et. al. (2005). The presence of chloride ion in the sample could enhance THMs formation during incubation (Munch and Hautman, 1995).

HPOA was the most reactive among six fractions producing THMs at 132 μg/mg. HPIN and HPON had moderate activity at 81 and 73 μg/mg, respectively. These three fractions were not only the three most reactive fractions but also the major three fractions in most of the samples. However, THMFP of the pure fractionated DOM could deviate from THMFP of the original mixed fractions in the samples as reported by Chow (2006) and Chow et al. (2006). The humic substance is operationally desorbed from XAD-8 resin with 0.1N NaOH. The recovery process is likely to be affected by a number of factors, such as the pH and ionic strength, water–resin volume ratio and possibly minor variations of the characteristics of the resin. It was indicated by Chow (2006) that alkaline extractable humic substances had lower THMFP during chlorination. The lower THMs formation was probably attributed to chemical changes during alkaline extraction

process in such a way that it decrease in aromaticity of humic substances (Pokorna et al. 2001). Aromatic carbon is generally considered as a major reactive moiety in forming THMs during chlorination (Norwood et al. 1987). These alterations may destroy some THMs reactive sites during the extraction process.

4.4 Chapter Summary

The six fractions' concentrations of natural water vary depending on its source. HPOA is the major fraction in rainfall runoff, whereas it is HPIN for the water in the canal. As HPOA is the most reactive fraction to THMs formation as well as the most abundant fraction in the runoff, the contamination of rainfall runoff to source water would increase THMFP of source water.

Table 4.1(a) Water Quality of Outfalls and Marsh Area Samples

Date	Sample Description	DOC (mg/L)	UV (1/cm)	SUVA (L/mg-m)	pH	Conductivity (µS)	THMFP (µg/L)
2/4/08	Blackwells Mills Outfall	3.8561	0.0783	2.03	6.97	529	65
2/4/08	Port Mercer Outfall	7.9995	0.2433	3.04	6.10	429	495
2/14/08	Blackwells Mills Outfall	3.3667	0.0948	2.25	7.10	524	294
2/14/08	Port Mercer Outfall	5.5914	0.2002	3.04	6.71	709	202
5/13/08	Port Mercer Outfall	10.0464	0.3060	3.17	7.06	523	910
5/13/08	Port Mercer Marsh Area	15.1121	0.4688	2.91	7.30	1080	1445
6/2/08	Port Mercer Outfall	11.8700	0.4807	3.95	7.05	543	1135
6/2/08	Port Mercer Marsh Area	14.8567	0.6195	4.03	7.66	998	1300

Table 4.1(b) Water Quality of the Samples along the Canal

Date	Sample Description	DOC (mg/L)	UV (1/cm)	SUVA (L/mg-m)	pH	Conductivity (µS)	THMFP (µg/L)
10/18/06	Raw water	3.5317	0.0150	0.43	N/A	530	N/A
3/8/07	Raw water	5.5685	0.0223	0.40	N/A	622	N/A
5/16/07	Raw water	2.7708	0.0423	1.28	N/A	630	170
8/6/07	Lock10	2.9185	0.0490	1.68	N/A	N/A	152
8/6/07	Lock11	2.9149	0.0510	1.75	N/A	N/A	146
8/6/07	Lock12	2.8104	0.0480	1.71	N/A	N/A	147
8/6/07	Raw Water	2.6449	0.0350	1.32	N/A	N/A	184
8/6/07	Plant Effluent	1.6993	0.0180	1.06	N/A	N/A	178
11/1/07	Raw water	2.4834	0.0521	2.10	N/A	520	217
2/4/08	Blackwells Mills before outfall	4.7575	0.1218	2.56	6.91	526	290
2/4/08	Blackwells Mills after outfall	5.0284	0.1657	3.29	6.90	512	307
2/4/08	Port Mercer before outfall	6.4205	0.1860	2.90	6.89	321	474
2/4/08	Port Mercer after outfall	6.8816	0.1837	2.67	6.98	368	485
5/13/08	Port Mercer after outfall	3.7241	0.1066	3.06	7.4	515	192
6/2/08	Port Mercer before outfall	2.7602	0.1011	3.51	7.12	502	292

Table 4.2(a) Six Fractions' Concentration of the Outfalls and Marsh Area Samples

Date	Sample Description	DOC (mg/L)	HPOA (mg/L)	HPON (mg/L)	HPOB (mg/L)	HPIA (mg/L)	HPIN (mg/L)	HPIB (mg/L)
2/4/08	Blackwells Mills Outfall	3.8561	1.2789	0.4923	0.3605	0.2557	1.212	0.2567
2/4/08	Port Mercer Outfall	7.9995	1.0288	4.4201	0.8632	0.1490	1.2974	0.2410
2/14/08	Blackwells Mills Outfall	3.3667	0.6350	1.1935	0.0506	0.0878	1.119	0.2808
2/14/08	Port Mercer Outfall	5.5914	0.4268	2.9351	1.0438	0.0598	1.0522	0.0737
5/13/08	Port Mercer Outfall	10.0464	5.6486	0.8015	0.0912	2.3179	0.809	0.3782
5/13/08	Port Mercer Marsh Area	15.1121	7.6362	3.3057	0.8651	1.9854	1.1261	0.1936
6/2/08	Port Mercer Outfall	11.8700	6.0951	2.5796	1.0082	1.2551	0.7126	0.2194
6/2/08	Port Mercer Marsh Area	14.8567	7.1288	1.8493	1.9728	2.3944	0.7659	0.7455

Table 4.2(b) Six Fractions' Concentration of the Samples along the Canal

Date	Sample Description	DOC (mg/L)	HPOA (mg/L)	HPON (mg/L)	HPOB (mg/L)	HPIA (mg/L)	HPIN (mg/L)	HPIB (mg/L)
10/18/06	Raw water	3.5317	1.1660	0.3265	0.2592	1.0487	0.6913	0.0400
3/8/07	Raw water	5.5685	1.9086	1.4643	0.0221	0.0823	1.9315	0.1597
5/16/07	Raw water	2.7708	0.5997	0.2357	0.3279	0.1214	1.4813	0.0048
8/6/07	Lock10	2.9185	0.4856	0.3645	0.2200	0.3484	1.0337	0.4663
8/6/07	Lock11	2.9149	0.2962	0.4438	0.6522	0.1629	1.0898	0.2700
8/6/07	Lock12	2.8104	0.1324	0.4661	0.2468	0.3745	0.9573	0.6333
8/6/07	Raw water	2.6449	0.3231	0.6566	0.3176	0.1819	1.0548	0.1109
8/6/07	Plant effluent	1.6993	0.4230	0.1578	0.0048	0.0632	0.9176	0.1329
11/1/07	Raw water	2.4834	0.2062	0.0542	0.1220	0.5222	1.3496	0.2292
2/4/08	Blackwells Mills before outfall	4.7575	1.3265	0.9303	0.3111	0.3700	1.0827	0.7369
2/4/08	Blackwells Mills after outfall	5.0284	1.2248	1.5035	0.4411	0.5890	1.0672	0.2028
2/4/08	Port Mercer before outfall	6.4205	0.7756	2.8808	1.0076	0.6093	1.0439	0.1033
2/4/08	Port Mercer after outfall	6.8816	0.8969	3.1431	0.7074	0.8124	1.1418	0.1800
5/13/08	Port Mercer after outfall	3.7241	0.4510	0.4801	0.1738	1.5958	0.8475	0.1759
6/2/08	Port Mercer before outfall	2.7602	0.6482	0.8219	0.2807	0.1631	0.6349	0.2114

Table 4.3(a) Percentage of Six Fractions of the Outfalls and Marsh Area Samples

Date	Sample Description	%HPOA	%HPON	%HPOB	%HPIA	%HPIN	%HPIB
2/4/08	Blackwells Mills Outfall	33.17	12.77	9.35	6.63	31.43	6.66
2/4/08	Port Mercer Outfall	12.86	55.26	10.79	1.86	16.22	3.01
2/14/08	Blackwells Mills Outfall	18.86	35.45	1.50	2.61	33.24	8.34
2/14/08	Port Mercer Outfall	7.63	52.49	18.67	1.07	18.82	1.32
5/13/08	Port Mercer Outfall	56.23	7.98	0.91	23.07	8.05	3.76
5/13/08	Port Mercer Marsh Area	50.53	21.88	5.72	13.14	7.45	1.28
6/2/08	Port Mercer Outfall	51.35	21.73	8.49	10.57	6.00	1.85
6/2/08	Port Mercer Marsh Area	47.98	12.45	13.28	16.12	5.16	5.02
	Average	34.83	27.50	8.59	9.38	15.80	3.91

Table 4.3(b) Percentage of Six Fractions of the Samples along the Canal

Date	Sample Description	%HPOA	%HPON	%HPOB	%HPIA	%HPIN	%HPIB
10/18/06	Raw water	33.02	9.24	7.34	29.69	19.57	1.13
3/8/07	Raw water	34.27	26.30	0.40	1.48	34.69	2.87
5/16/07	Raw water	21.64	8.51	11.83	4.38	53.46	0.17
8/6/07	Lock10	16.64	12.49	7.54	11.94	35.42	15.98
8/6/07	Lock11	10.16	15.23	22.37	5.59	37.39	9.26
8/6/07	Lock12	4.71	16.58	8.78	13.33	34.06	22.53
8/6/07	Raw water	12.22	24.83	12.01	6.88	39.88	4.19
8/6/07	Plant effluent	24.89	9.29	0.28	3.72	54.00	7.82
11/1/07	Raw water	8.30	2.18	4.91	21.03	54.34	9.23
2/4/08	Blackwells Mills before outfall	27.88	19.55	6.54	7.78	22.76	15.49
2/4/08	Blackwells Mills after outfall	24.36	29.90	8.77	11.71	21.22	4.03
2/4/08	Port Mercer before outfall	12.08	44.87	15.69	9.49	16.26	1.61
2/4/08	Port Mercer after outfall	13.03	45.67	10.28	11.81	16.59	2.62
5/13/08	Port Mercer after outfall	12.11	12.89	4.67	42.85	22.76	4.72
6/2/08	Port Mercer before outfall	23.48	29.78	10.17	5.91	23.00	7.66
	Average	18.59	20.49	8.77	12.50	32.36	7.29

Table 4.4 THMs Formation of Six Fractions (n=6).

Fractions	DOC (mg/L)	THMFP (µg/mg or µg/L)	Standard Deviation (µg/mg or µg/L)
HPOA	1	132	8
HPON	1	73	7
HPOB	1	11	3
HPIA	1	67	3
HPIN	1	81	6
HPIB	1	43	3

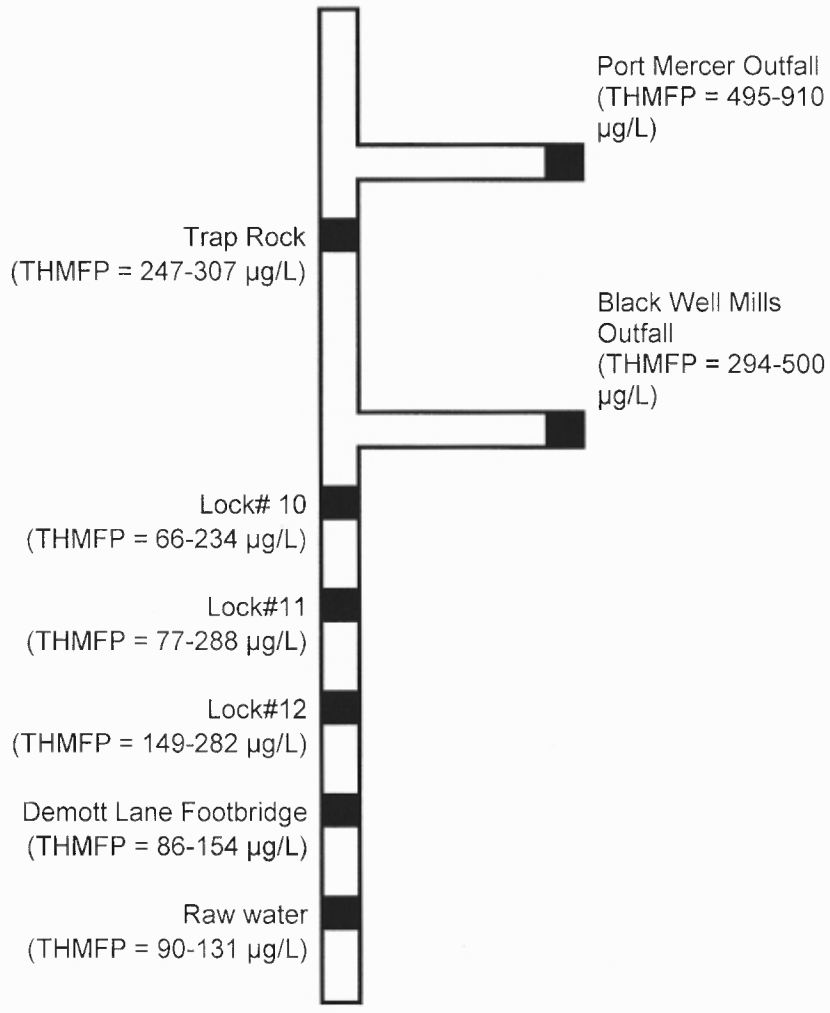


Figure 4.1 The map of sampling locations and the samples' THMFPs along Delaware & Raritan Canal and its tributaries.

CHAPTER 5

CHARACTERIZATION OF TRIHALOMETHANES PRECURSORS BY FLUORESCENCE SPECTROSCOPY

5.1 Introduction

Synchronous fluorescence spectroscopy is a technique used to obtain the summation of fluorescence spectra of varied organic compounds in water samples at a wide range of excitation and emission wavelengths. This fluorescence spectrum (i.e., SFS of a water sample) is a "fingerprint" of natural organic compounds present in source water. The peak location of the spectrum, on the Ex/Em matrix, was used to identify the type of organic compounds (Chen et al. 2003). The intensities could indicate the compounds' concentrations (Robinson, 1995; Skoog et al. 1998). However, using SFS for DOM characterization could be a complicated process due to the interference of water scattering and the overlap of the spectrum of a variety of DOM in the samples.

Three types of scatterings may interfere with the interpretation of fluorescence spectrum: Rayleigh, Raman, and Tyndall scattering. Tyndall scatter involves reflection from colloid as well as particulate matter present in the sample, while Rayleigh and Raman scatters originate from an interaction between the excitation light and water molecules (Zepp et al. 2004). Both scattering phenomena occur when a molecule has been excited to a higher energy state and quickly relaxes, emitting lights (scatters) of equal or higher wavelength (Larsson et al. 2007). Raman scatter is at a certain energy difference from the first order Rayleigh scatter. This energy difference is dependent upon the media in the sample.

Each sample was collected from different locations and time. Therefore, it could have various fractions' concentration as well as DOM composition. In this chapter, a sample was collected for demonstration of using resin adsorption and fluorescence spectroscopy for characterization of DOM in source water. First, DOM was isolated into six fractions by resin adsorption and was characterized by fluorescence spectroscopy. The selection of desirable SFS region on Ex/Em matrix for DOM characterization was discussed. Also, the application of SFS for identification of DOM and THMs precursors were demonstrated.

5.2 Fractionation Results

A sample, collected at the intake of Middlesex Water Company in October 2006, was selected for demonstration of characterization by resin adsorption and fluorescence spectroscopy. The sample had TOC and DOC of 3.6350 mg/L and 3.4957 mg/L, respectively. DOM was isolated into six fractions as shown in Table 5.1. The pie chart of six fractions contribution to the total DOC is shown in Figure 5.1. HPOA, HPIA, and HPIN were the major fractions in this samples accounting for 33.36%, 30.00%, and 19.78% of total DOC, respectively. HPIB is usually present at lowest concentration, and it is only 0.0040 mg/L in this sample.

5.3 Fluorescence Scatters

Figure 5.2(a) shows three dimensional Rayleigh scatters and fluorescence spectra of the sample. Rayleigh scatter lies on a diagonal line of the Ex/Em matrix. The first order Rayleigh scatter line is centered at the Em wavelength equal Ex wavelength, and second

order Rayleigh scatter at the Em wavelength equal twice the Ex wavelength. The first-order Rayleigh had a very high intensity, larger than 5000 au, while the peak intensity of the sample DOM was less than 2000 au. The second-order Rayleigh scatter also had higher intensity than DOM spectra but smaller than the first-order scatter.

First order Rayleigh scatter is the emitted light, which has the same energy as the exciting light. This process, known as elastic scattering, is the most probable one, and, consequently, it has the highest intensity (Zepp et al. 2004, Rinnan et al. 2005). The second order Rayleigh is an inelastic process, in which a molecule relaxes to a different vibrational energy level of the ground state than the original one. If the molecule relaxes to a higher vibrational level, the emitted light has a lower frequency and higher wavelength than the exciting light (Larsson et al. 2007).

The two dimensional top-view of Figure 5.2(a) is shown in Figure 5.2(b) illustrating the location of both scatters and the area of DOM spectra on Ex/Em matrix. This figure also shows the contour line of SFS, the line of the same SFS intensity, which easily locates the scatters and peak intensity of the sample.

5.4 SFS Post-Processing

As the scatters were very large compared to SFS of the DOM in the sample, the peak and shape of DOM spectra can be observed only in a particular angle. For this reason, scatters were removed as shown in Figure 5.3(a). For a better observation, the figure was rotated to another angle and its intensity axis was scaled at 2000 au maximum as shown in Figure 5.3(b). Additionally, the area where emission wavelengths were less than excitation

wavelength should be excluded from the study since by theory emission energy must be less than excitation energy.

Figure 5.3(c) shows the SFS in Figure 5.3(b) in two-dimensions (top view). The contour line of SFS showed that the water sample had two peak locations at Ex/Em matrix of 249/423 and 315/423 (nm/nm) with the fluorescence intensity of 1889 and 1034 au, respectively. These peaks were related to the major organic compounds in the sample.

5.5 Coefficient of Variation of SFS Intensity

Coefficient of variation (CV) is defined as standard deviation divided by the average (Anderson et al. 2002). It is a measure of dispersion of data. If the data have high CV, it indicates high dispersion or low precision. CV is usually reported as a percentage (%) by multiplying by 100. Ten duplicates of SFSs were obtained for the original water sample and then % CV was computed for intensity values at all Ex/Em data points of SFS.

The %CV of SFS of the water sample in three dimensions is shown in Figure 5.4(a), and in two dimensions (top-view) in Figure 5.4(b). Two areas on the Ex/Em matrix have shown high %CV or low precision of fluorescence intensity measurement. These areas are in the low excitation wavelengths regions from 225 to 255 nm and high emission wavelength regions from 531 to 633 nm. The darkened region in Figure 5.4(b) shows the area of % CV larger than 10%. Although not shown herein, %CVs of fluorescence intensity for six fractions were also obtained, and they all have the same characteristics (i.e., having high % CV on the area of low excitation and high emission wavelength). It is not a routine to obtain 10 duplicate SFSs for each sample. The

objective of this work is to specify the useful area of high precision of fluorescence intensity for DOM characterization and quantitative analysis.

In summary, the useful SFS region for DOM characterization should have the following characteristics: 1) Not overlap with the scatters, 2) The Em wavelengths longer than Ex wavelengths, and 3) Having good precision of fluorescence intensity or the area between 225-489 nm of excitation wavelengths and 291-603 nm of emission wavelengths.

5.6 Characterization of the Water Sample and Its Six Fractions Using SFS

All the six fractions were eluted from the resins and their concentrations were diluted to 1 mg/L for consistency and comparison; then characterized by fluorescence spectroscopy. Two and three dimensional SFSs of HPOA, HPON, HPOB, HPIA, HPIN, HPIB were illustrated in Figures 5.5 to 5.10, respectively. The peak intensity, peak location, and % CV of the intensity are reported in Table 5.2. The fluorescence peak intensities of the six fractions at the same concentration were in the following order: HPIB ~ HPOA > HPON > HPOB >> HPIN ~ HPIA. Figures 5.5 to 5.10 and Table 5.2 show that at the same concentrations, each fraction possesses unique characteristics of having a unique SFS shape, intensity, peak location on Ex/Em matrix. The peak location was the characteristic of each organic fraction. The variety of organic types were characterized by fluorescence spectroscopy and their peak locations were reported by Chen et al. (2003).

It should be noted that peaks location in Table 5.2 is slightly different from ones in Table 2.2 reported earlier by Marhaba, 2000. This is probably because SFSs were obtained from different samples and possibly different set-up of fluorescence photometer.

Figure 5.3(c) shows that the water sample had two peak locations at Ex/Em matrix of 249/423 and 315/423 (nm/nm) with the fluorescence intensity of 1889 and 1034 au, respectively. The first peak location was near the location of fulvic acid, whereas the second peak location around Ex/Em of 315 /423 can be linked humic acid (Chen et al. 2003). It can be concluded that natural fulvic and humic acid were the major fluorophores in the water sample.

Three-dimensional SFS of 1 mg/L HPOA fraction is shown in Figure 5.5(a) and its two dimensional SFS is shown in Figure 5.5(b). The peak location of HPOA was at Ex/Em matrix of 255/423 and 309/423 nm/nm with the intensities of 888 and 544 au, respectively. These two peak locations of HPOA fraction were also close to the peak locations of the original sample. This implies that HPOA could be the dominant fraction in this sample. This conclusion is confirmed by resin adsorption results as shown in Table 5.1.

SFS of 1 mg/L HPON is presented in Figure 5.6(a) and Figure 5.6(b) in two and three dimensions, respectively. The peak location was at 243/391 Ex/Em matrix at fluorescence intensity of 617 au. The peak location of HPON did not match to peaks of any compounds reported by Chen et al. (2003). Leenheer (2004) reported that HPON fraction was related to tannin or lignin derived compounds of organic matter found after plant degradation.

SFS of 1 mg/L HPOB fraction is illustrated in Figure 5.7(a) in three dimensions, and in Figure 5.7(b) in two dimensions. HPOB had only one peak location at Ex/Em matrix of 237/405 nm/nm with the fluorescence intensity of 318 au. The HPOB fraction

was reported to be amine or amide having aromatic characteristics (Leenheer, 2004; Kanokkantapong et al. 2005).

Figure 5.8(a) and (b) show SFSs of 1 mg/L HPIA in three and two dimensions, respectively. The peak was observed around 237/387 Ex/Em matrix at fluorescence intensity of 263 au. HPIA was reported to be polyuronic acids (Leenheer, 2004) or low molecular weight carboxylic acids (Kanokkantapong et al. 2005).

SFSs of 1 mg/L HPIN is demonstrated in Figure 5.9(a) and (b). Only one peak was observed at Ex/Em matrix of 225/303 nm/nm at fluorescence intensity of 301 au. The location was near that of tyrosine or aromatic proteins (Chen et al. 2003). Tyrosine is one of amino acids mostly found in milk or cheese. Leenheer reported that HPIN was composed of some kinds of sugars (Leenheer, 2004).

Figure 5.10(a) and (b) show SFSs of 1 mg/L HPIB fraction. Two peaks were observed at 231/309 and 267/369 on Ex/Em matrix at fluorescence intensity of 624 and 884 au, respectively. The first peak was close to tyrosine or aromatic protein and the second peak was close to peak of tryptophan or protein like biological compounds (Chen et al. 2003). Although the intensity of HPIB was high compared to that of other fractions, HPIB was found at low concentration in this sample. Hence, the peak of HPIB cannot be observed in the total SFS of the sample.

DOC, SFS peak location, and peak intensity are reported in Table 5.2. At the same concentration of 1 mg/L, the HPOA, HPIB, and HPON had relatively high intensity. Because HPOA was the highest fraction and it is also very sensitive to fluorescence spectroscopy, only the peak of HPOA was clearly observed in SFS of the sample. Although the sample may contain some protein-like organic compounds (e.g.,

HPIN or HPIB), their SFSs were not significant compared to SFS of fulvic and humic acid in the sample.

5.7 Chapter Summary

SFS is a characteristic of each group of organic compounds, as each fraction had unique chemistry, shape, peak intensity, and locations on Ex/Em matrix. The useful SFS on Ex/Em matrix for characterization of THMs precursors should not include the area of scatters, the area where the Ex wavelengths longer than Em wavelengths, and the area of low precision of fluorescence intensity. After post-processing to remove undesirable regions, the SFSs can be used to identify organic compound classes in the sample.

HPOA is the major THMs precursor, the most reactive fraction to THMs formation, and also very sensitive to fluorescence spectroscopy. Consequently, fluorescence intensities in the region of HPOA spectra could be related to THMs formation.

This chapter has demonstrated the application of resin adsorption and fluorescence spectroscopy for characterization of DOM in natural water. Similarly, these procedures could also be used for characterization of water samples from other watersheds.

Table 5.1 DOC of the Six Fractions

Samples	DOC (mg/L)
Hydrophobic	
▪ Acid	1.1660
▪ Neutral	0.3265
▪ Base	0.2592
Hydrophilic	
▪ Acid	1.0487
▪ Neutral	0.6913
▪ Base	0.0040
Total organic carbon	3.4957

Table 5.2 Fluorescence Peak Intensity and Their Locations for the Water Sample and Its Six Fractions

Samples	DOC (mg/L)	Peak intensity (au)	Peak location (Ex/Em, nm/nm)	%CV
Water sample	2.7708	1889 1034	249/423 315/423	10.73% 2.50%
HPOA	1	888 544	255/423 309/423	9.86% 2.78%
HPON	1	617	243/391	12.97%
HPOB	1	318	237/405	24.09%
HPIA	1	263	237/387	16.79%
HPIN	1	301	225/303	18.68%
HPIB	1	624 884	231/309 267/369	18.48% 5.78%

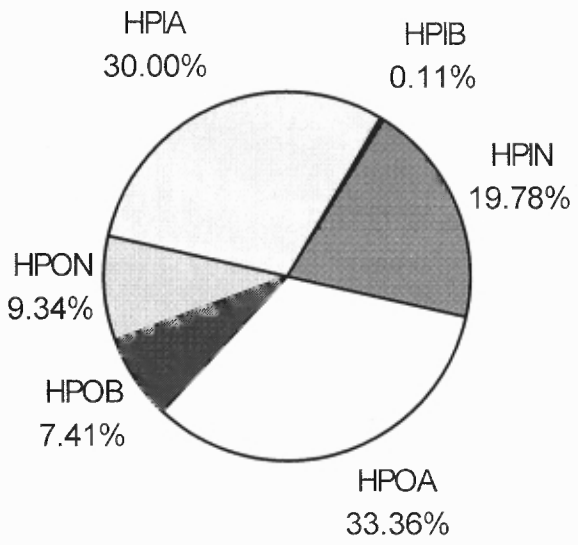


Figure 5.1 Pie chart of the six fractions as percent of total DOC.

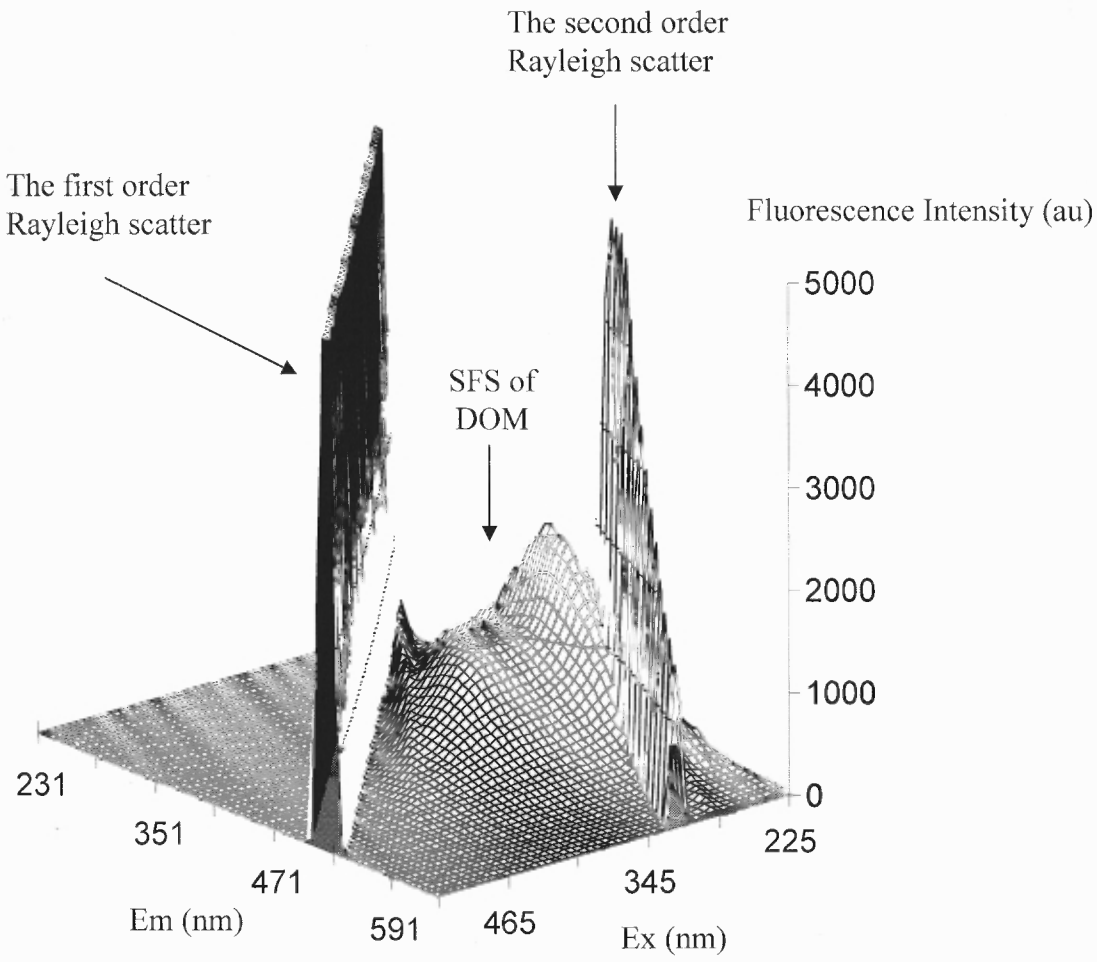


Figure 5.2(a) Three dimensional SFS of the water sample shows intensities of the first and second order Rayleigh scatters.

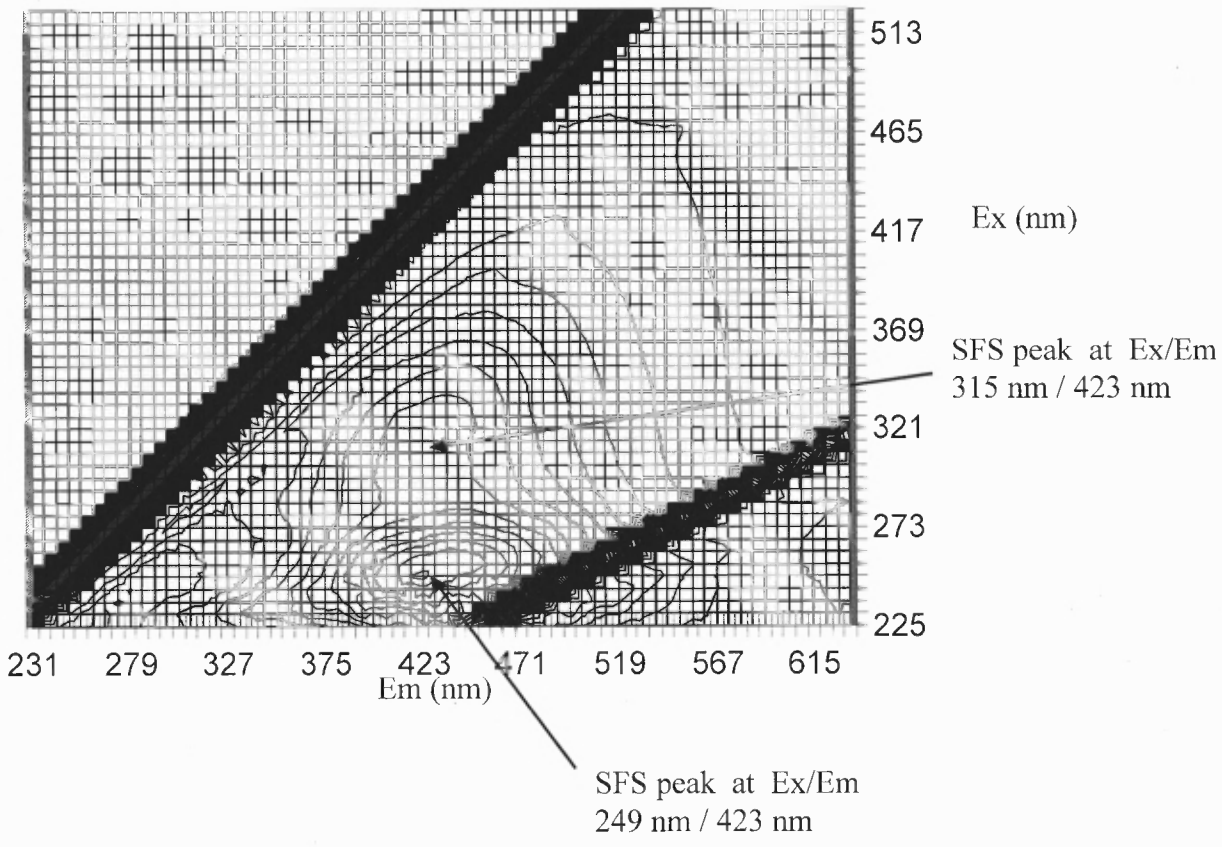


Figure 5.2(b) SFS contour of the water sample.

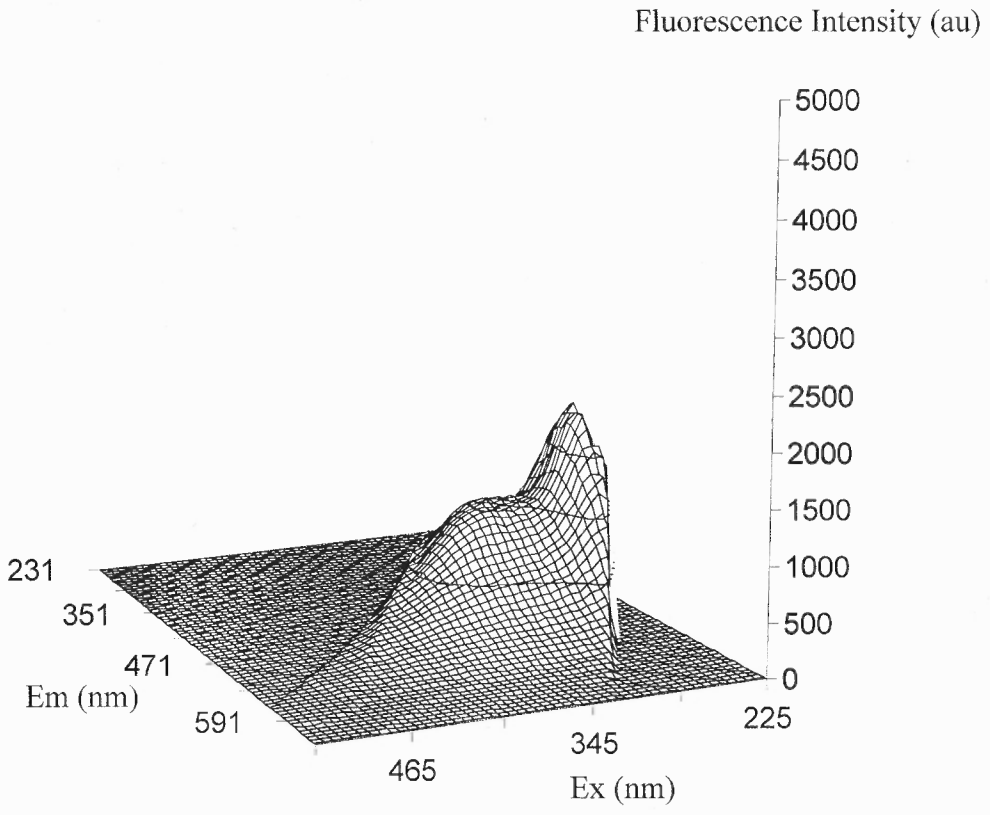


Figure 5.3(a) SFS of the water sample after the removal of Raman and Rayleigh scatters.

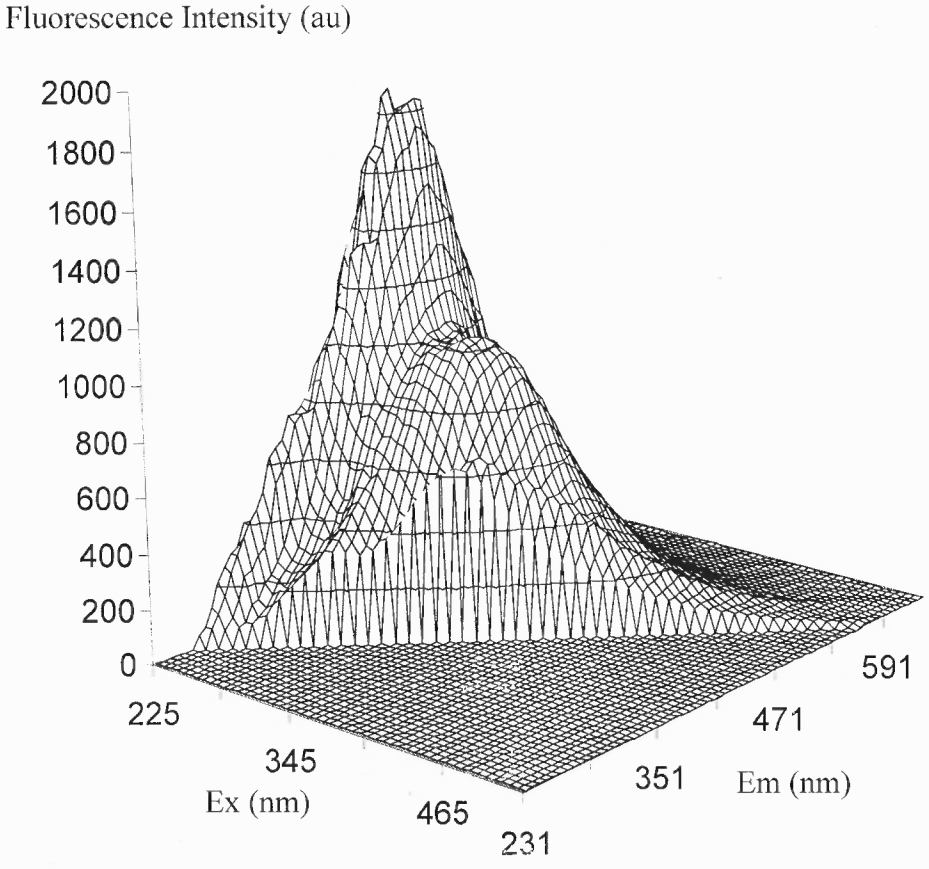


Figure 5.3(b) The same SFS from different angle.

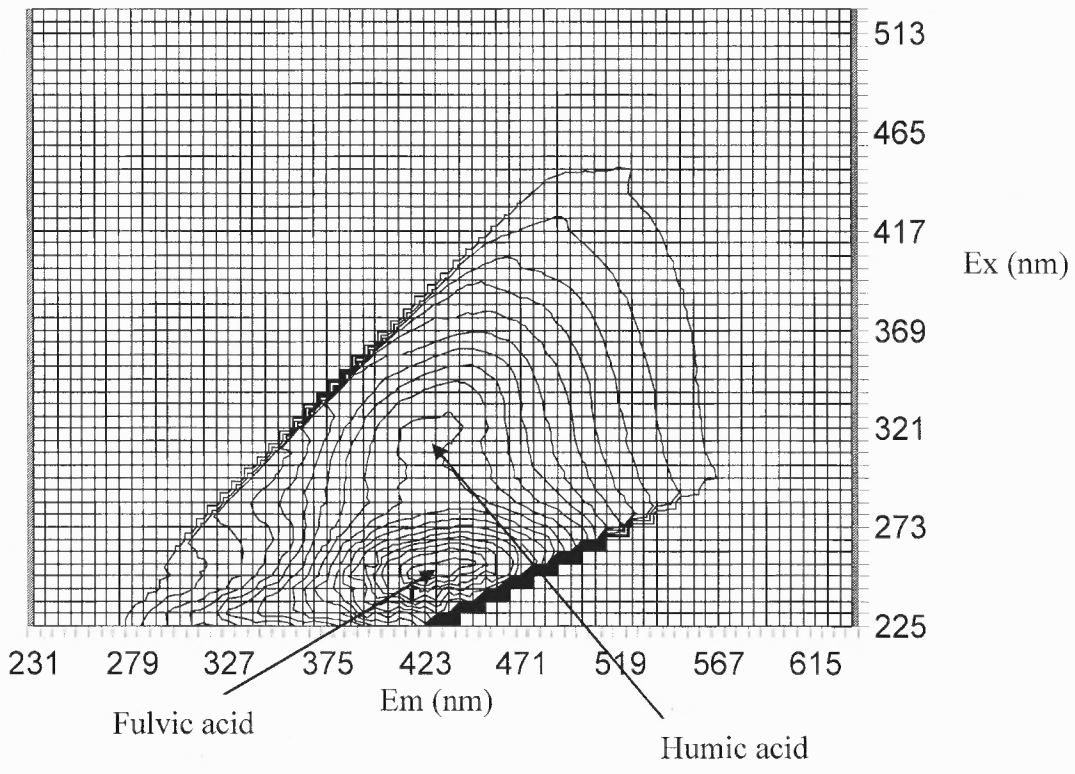


Figure 5.3(c) SFS contour after scatters removal.

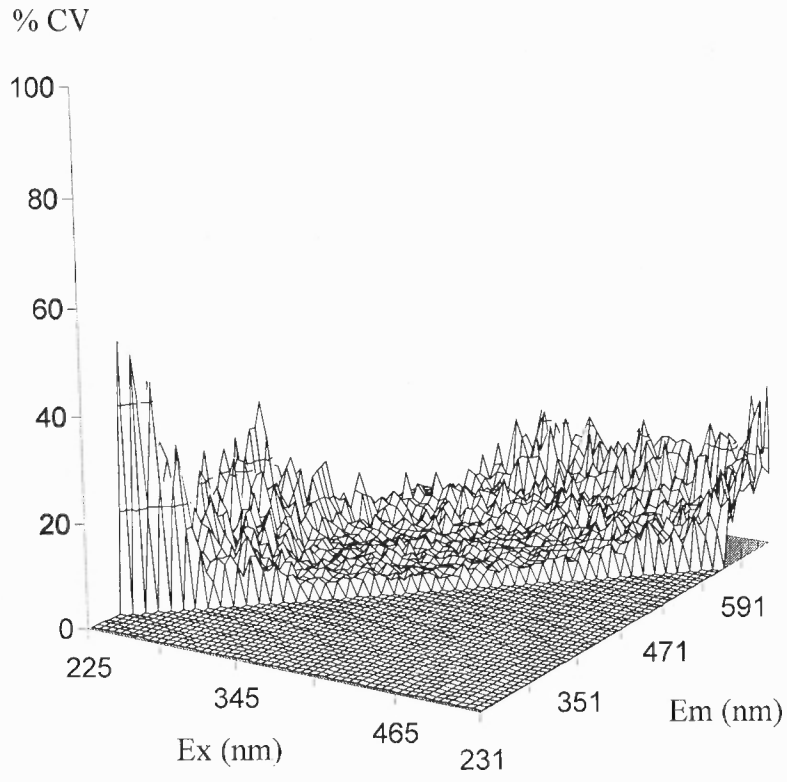


Figure 5.4(a) % CV of SFS intensity of the water sample in three dimensions.

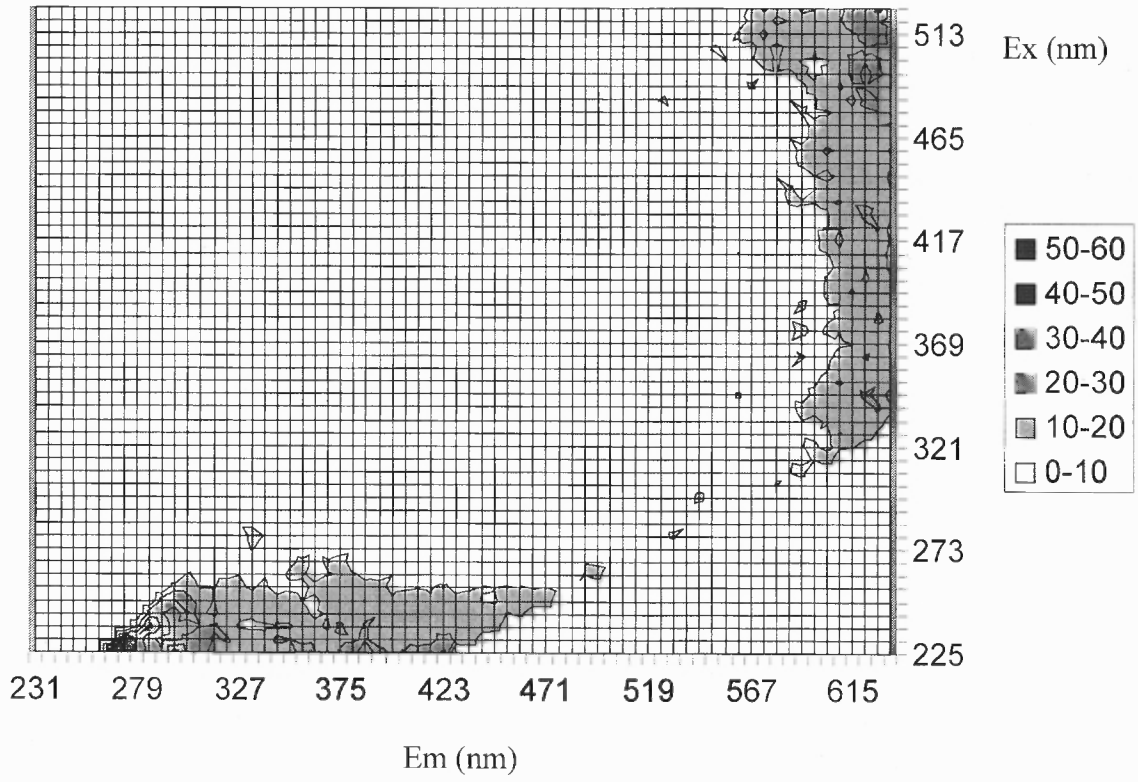


Figure 5.4(b) % CV contour of SFS intensity of the water sample.

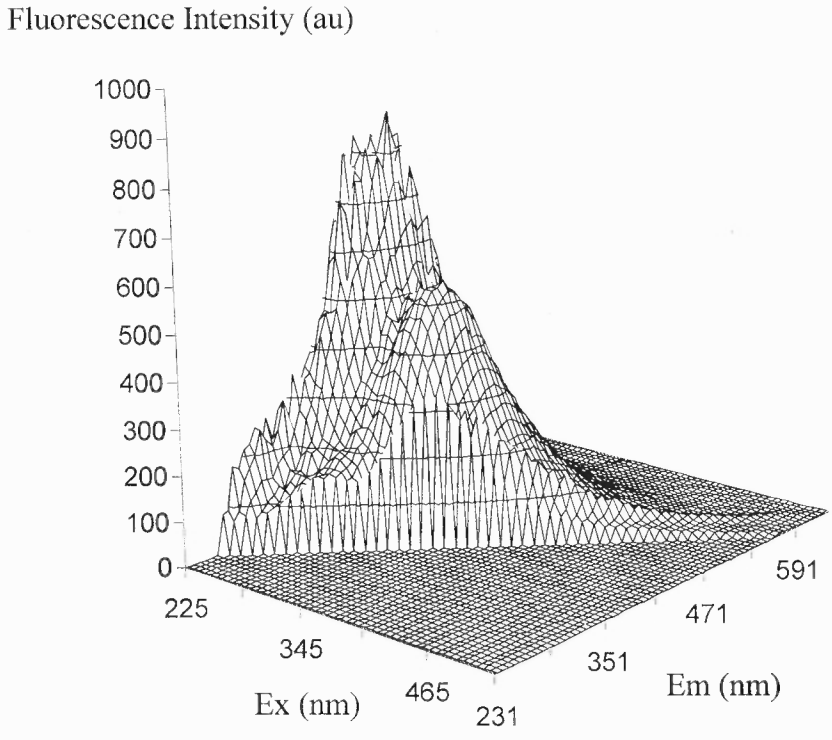


Figure 5.5(a) Three dimensional SFS of 1 mg/L HPOA.

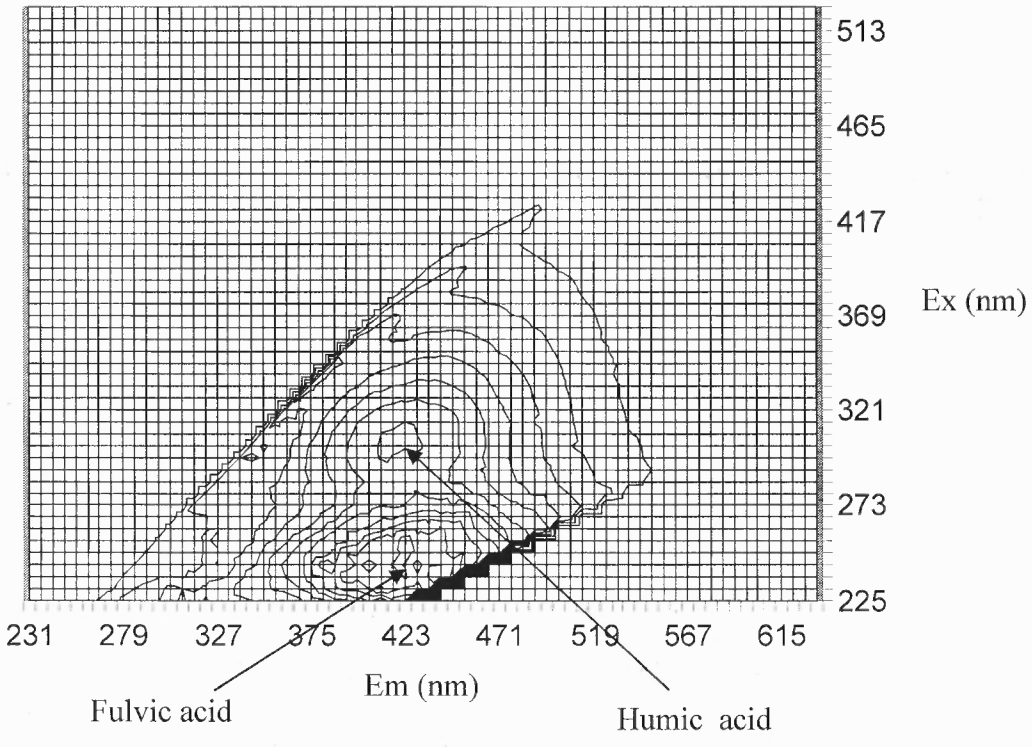


Figure 5.5(b) SFS contour of 1 mg/L HPOA.

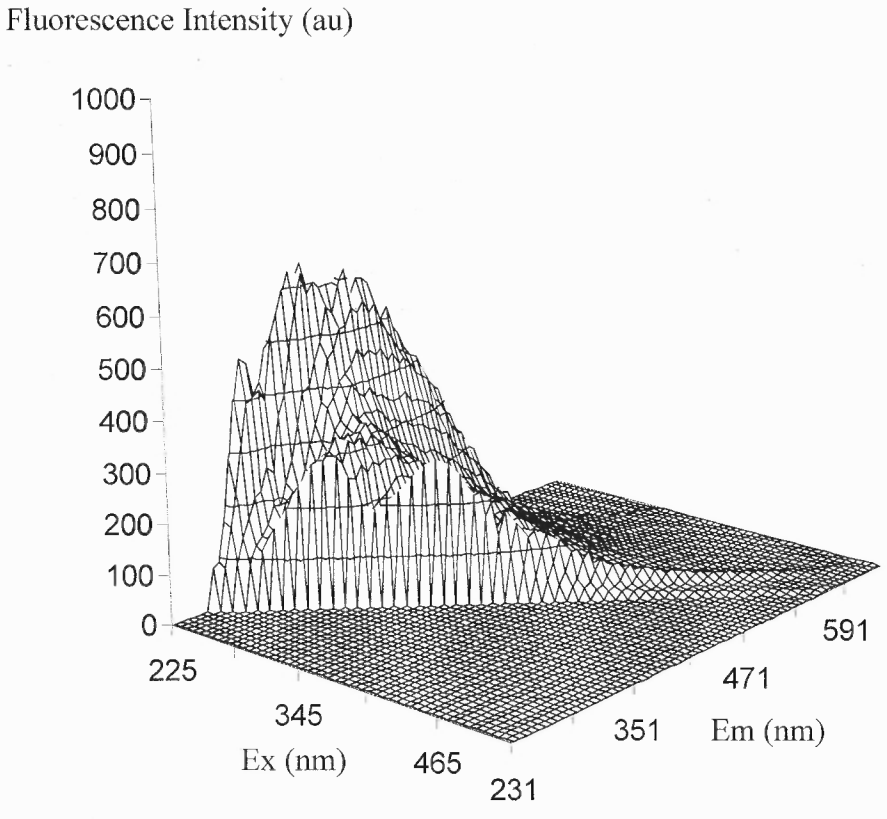


Figure 5.6(a) Three dimensional SFS of 1 mg/L HPON.

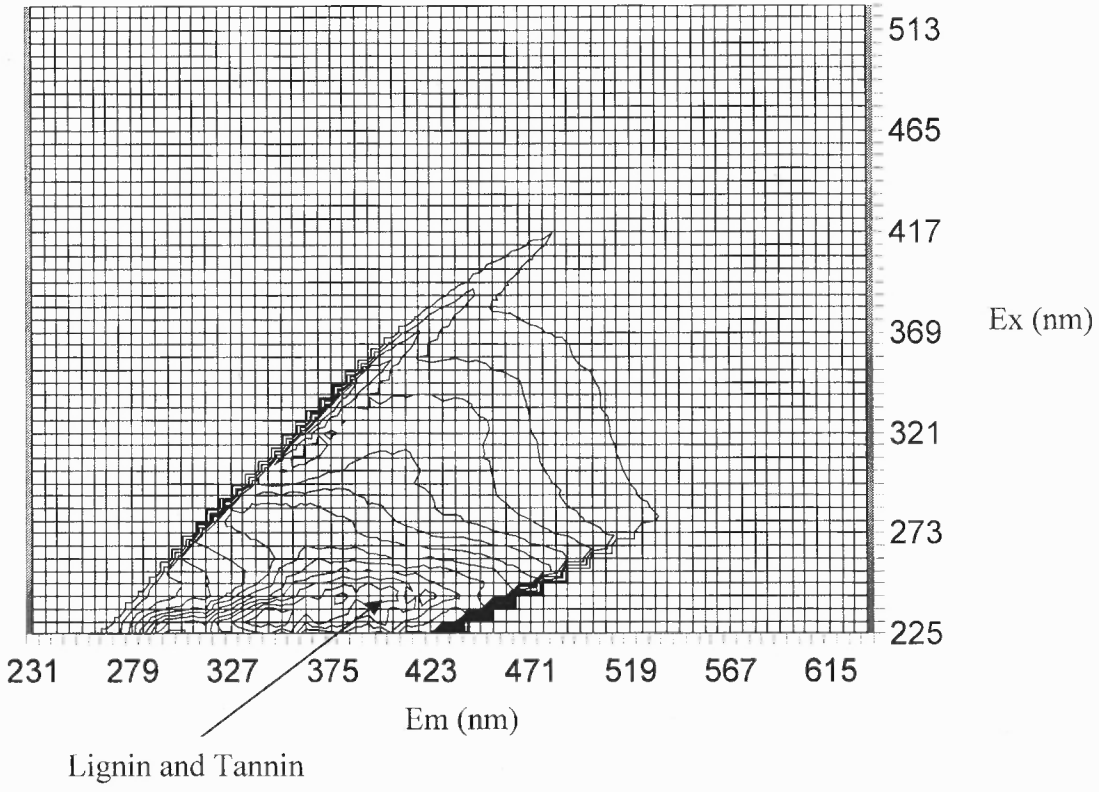


Figure 5.6(b) SFS contour of 1 mg/L HPON.

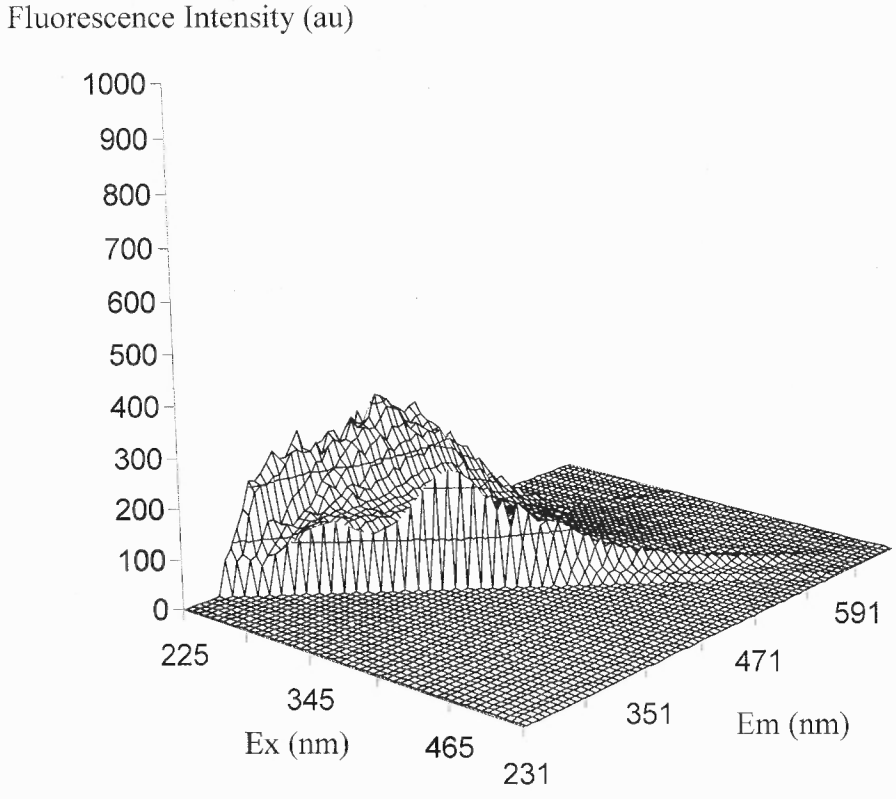


Figure 5.7(a) Three dimensional SFS of 1 mg/L HPOB.

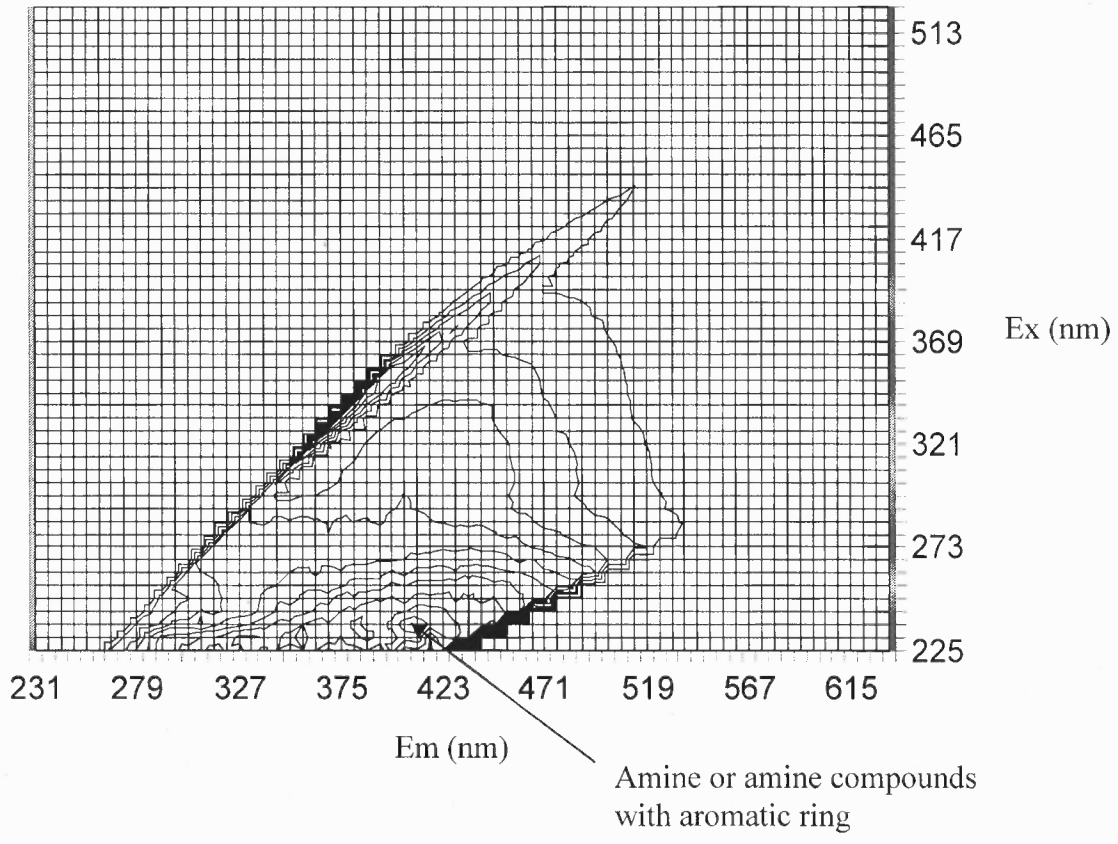


Figure 5.7(b) SFS contour of 1 mg/L HPOB.

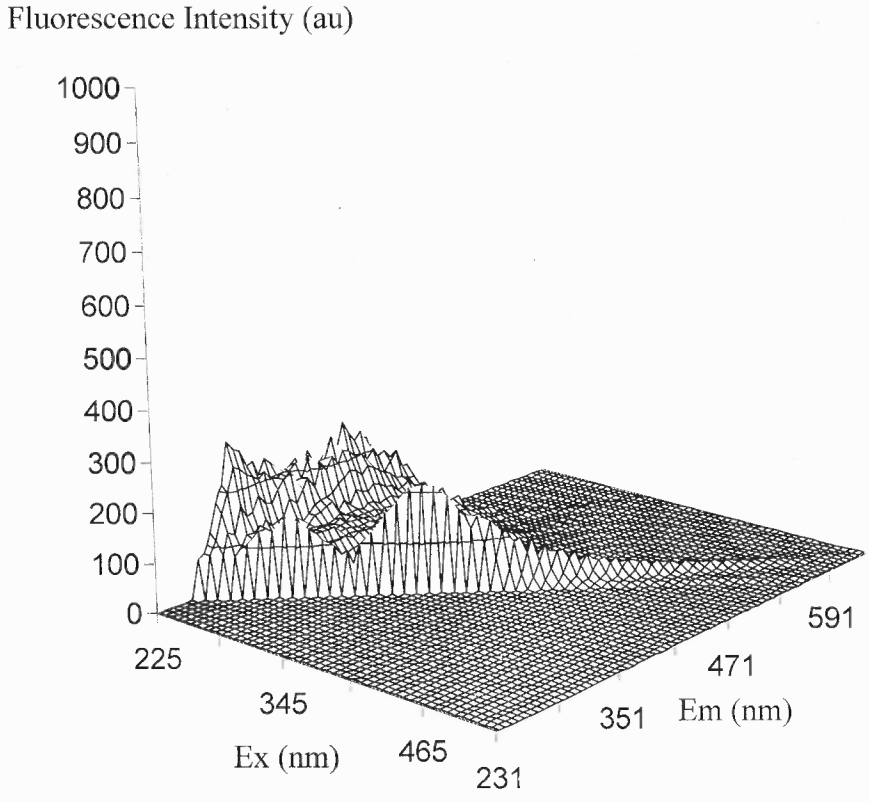


Figure 5.8(a) Three dimensional SFS of 1 mg/L HPIA.

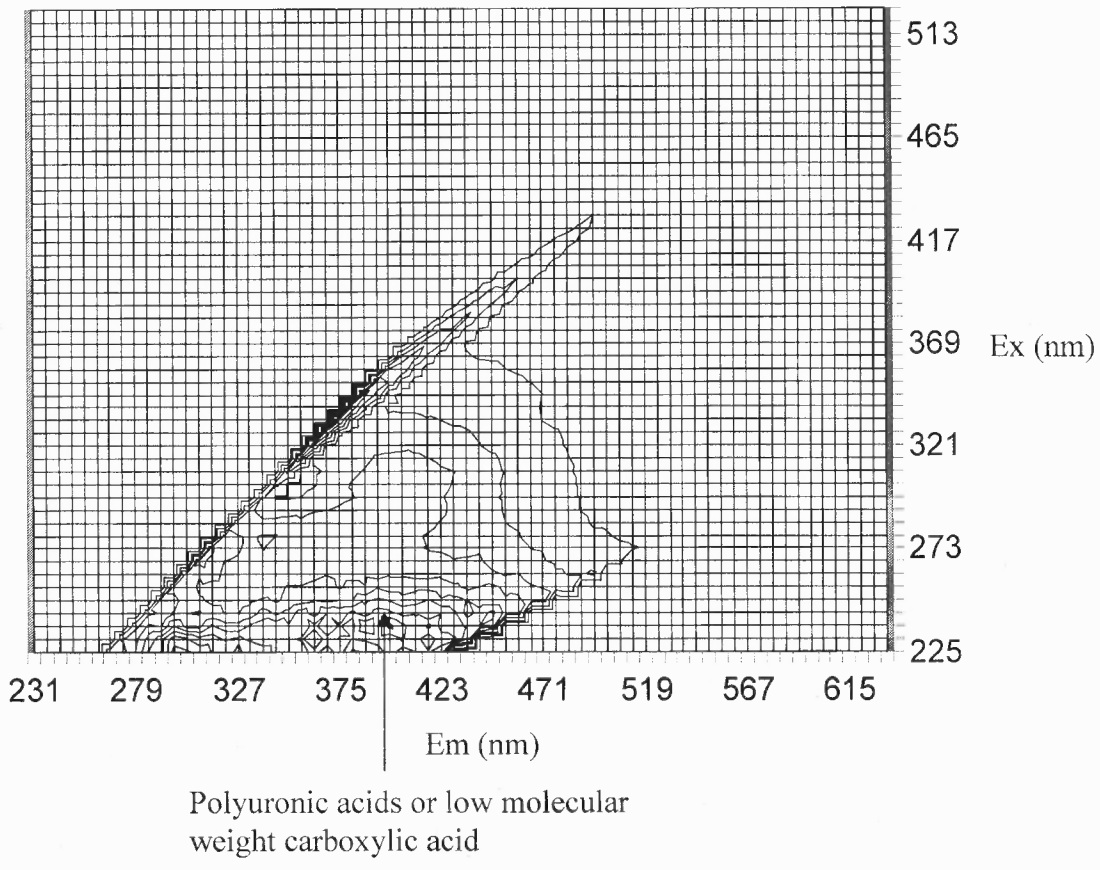


Figure 5.8(b) SFS contour of 1 mg/L HPIA.

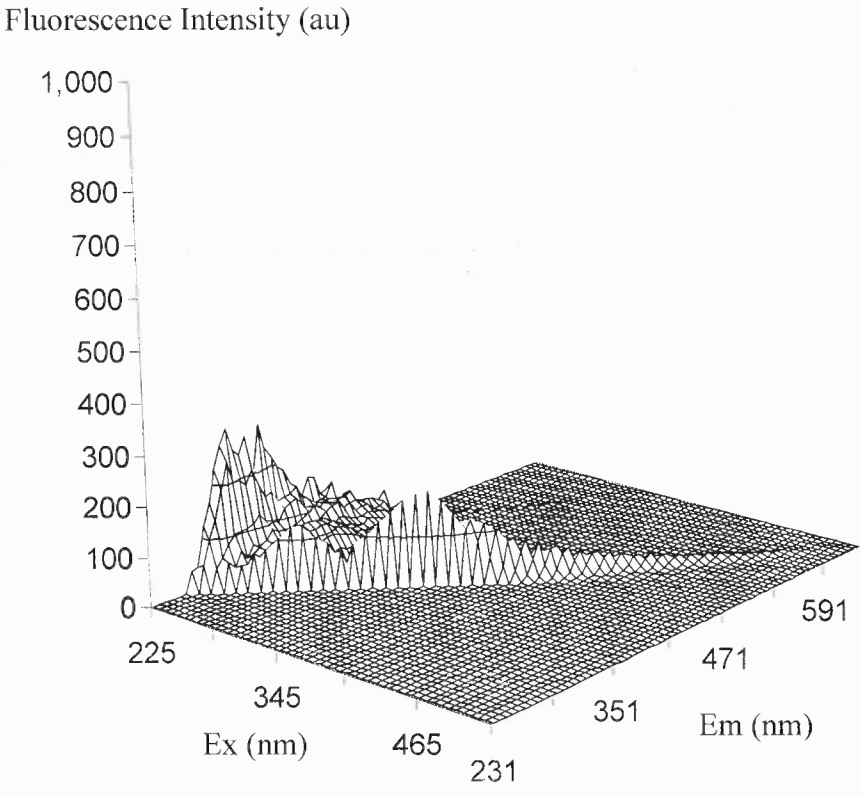


Figure 5.9(a) Three dimensional SFS of 1 mg/L HPIN.

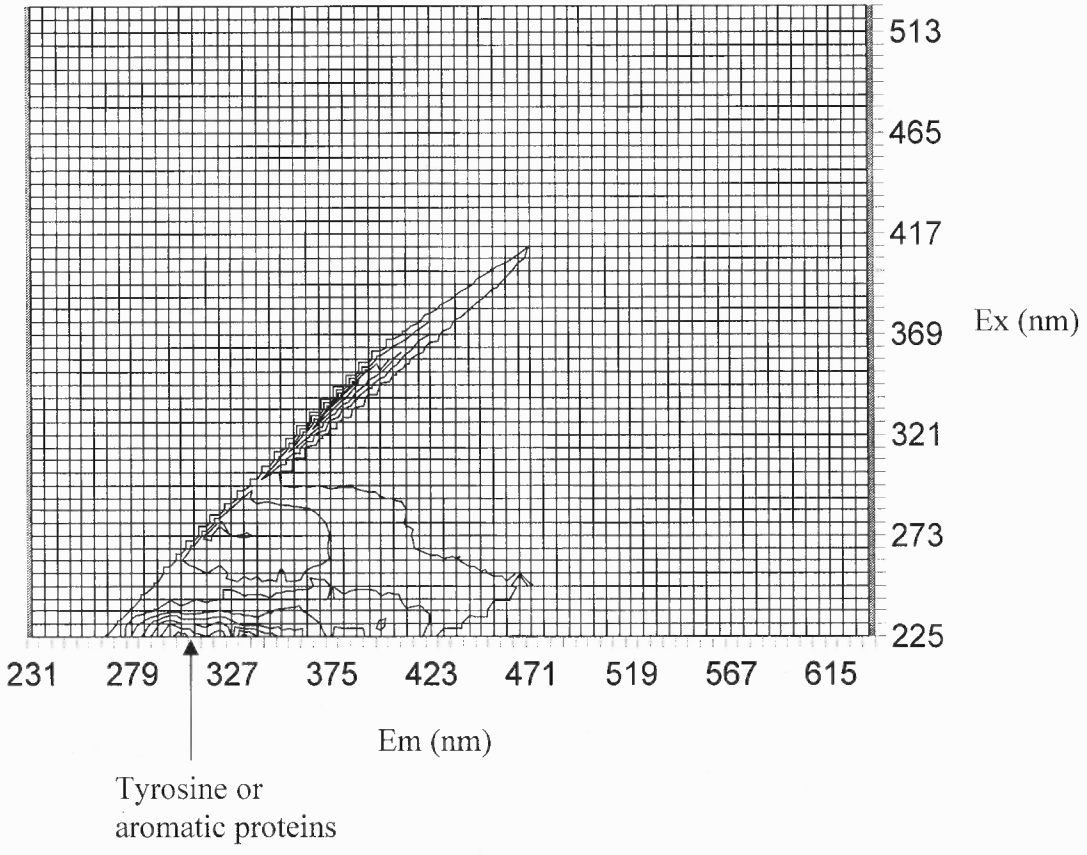


Figure 5.9(b) SFS contour of 1 mg/L HPIN.

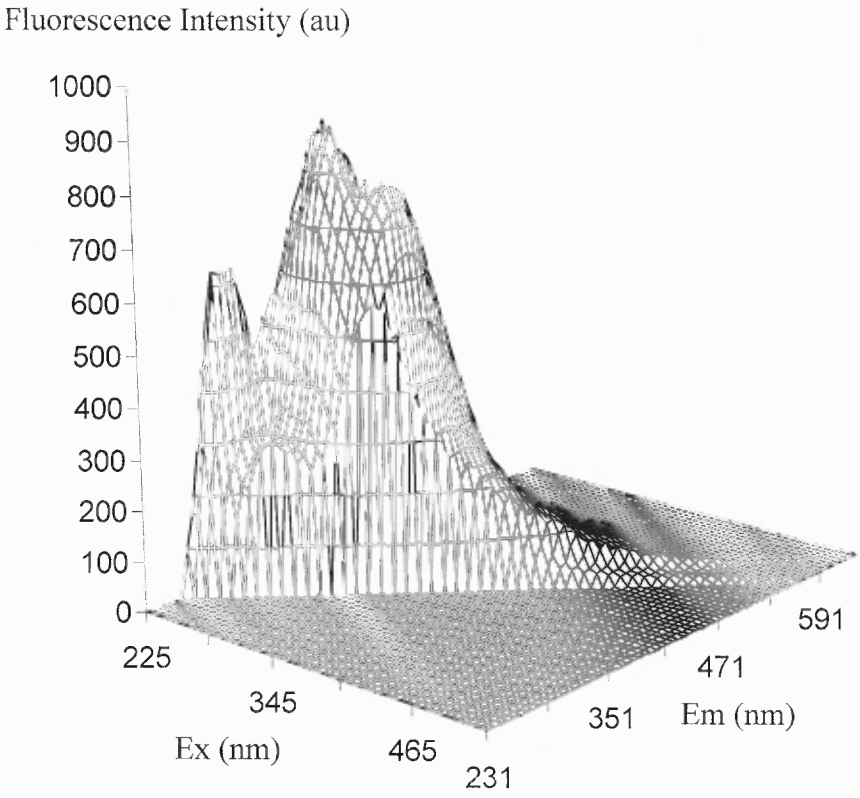


Figure 5.10(a) Three dimensional SFS of 1 mg/L HPIB.

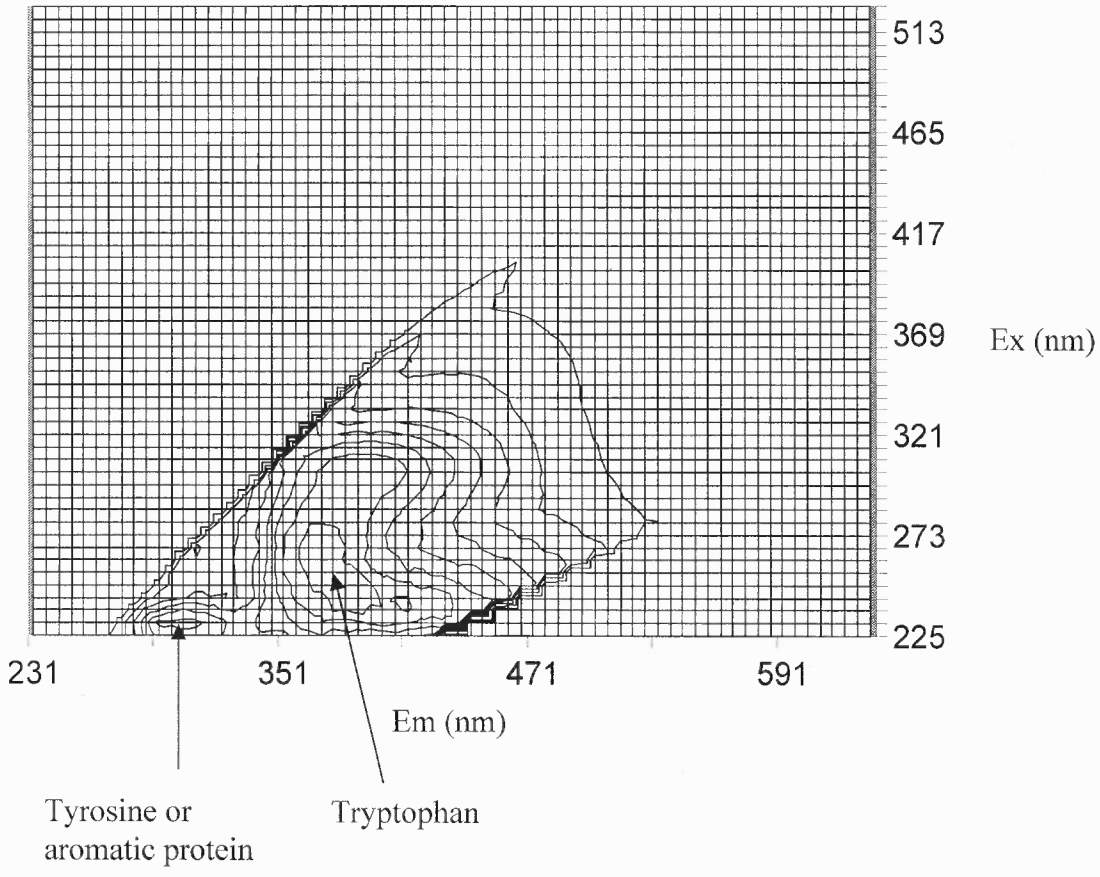


Figure 5.10(b) SFS contour of 1 mg/L HPIB.

CHAPTER 6

PREDICTION OF DISSOLVED ORGANIC CARBON AND TRIHALOMETHANES FORMATION POTENTIAL

6.1 Introduction

Disinfection of drinking water by chlorine produces DBPs such as THMs and HAAs. These are undesirable compounds due to their potential toxicity to human health. The DBPs formation is from the reaction of chlorine with DOM in source water (Rook, 1974).

Fluorescence spectroscopy is a sensitive technique for characterization of THMs precursors in water samples as discussed in Chapter 5. Consequently, multivariate statistical analysis of fluorescence spectra for prediction of DOC and THMFP was explored and demonstrated this chapter. Moreover, DOC and THMFP prediction models were also developed by MLR using UV as predictor variables. Then, models' performances were compared. For consistency and comparison, all the models were developed from the same 40 calibration samples.

6.2 Data Preparation and Model Calibration

Water samples were characterized for DOC, THMFP, and fluorescence spectra. For models developed using SFS, regions of scatter and undesirable regions of SFS intensity more than 10% CV were removed as demonstrated in Chapter 5. HPOA was the most reactive fraction to THMs formation, and also highly sensitive to fluorescence spectroscopy. Hence, a particular SFS region on Ex/Em matrix covering two major peaks of the HPOA fraction was selected for quantitative analysis. This region includes Ex wavelength from 243 nm to 327 nm and Em wavelength from 375 to 459 nm. Prediction

models for DOC and THMFP were developed by multivariate statistical analysis (i.e., PCA & MLR) of fluorescence spectra. Moreover, THMFP prediction model was developed using UV and DOC by MLR. Similarly, DOC model was also developed using UV. All statistical analysis and model development were performed using Minitab statistical software version 15.

Model calibration is the process of constructing a mathematical model to relate the response or an output of an instrument (such as DOC or THMFP) to properties of samples or an input (e.g., fluorescence spectra or UV absorption). Prediction is the process of using the mathematical relationship to calculate the response from given sample properties (Babee et al. 1998).

Fluorescence intensity or UV absorption can be mathematically related to the concentration of organic compounds in the sample (Edwald et al. 1985; Robinson, 1995; Skoog et al. 1998). To construct the models, 40 samples with a wide range of UV, fluorescence intensity, DOC, and THMFP were used for calibration. A mathematical relationship was developed to relate the SFS or UV to predict DOC & THMFP of the samples. Unknown samples were characterized for SFS and UV and then used as the input to the model. Finally, DOC and THMFP of the samples were predicted.

6.3 Models Development by PCA and MLR Using SFS as Predictors

Fluorescence intensities on selected SFS region are highly correlated, so general linear regression cannot be performed directly on fluorescence spectra. For this reason, fluorescence intensities on the selected area were analyzed by PCA to transform a large number of fluorescence intensities into a few, un-correlated PCs (Johnson and Wichern,

2007). The number of PCs for further analysis can be determined by their Eigen values. The PCs having Eigen values larger than one are kept for further analysis, and the rest are discarded (Field, 2005).

6.3.1 Principle Component Analysis

PCA was performed on 225 highly correlated fluorescence intensities on selected SFS region and transforms the intensities into a few PCs. As shown in Figure 6.1, only the first three PCs had Eigen values larger than one those were 218.70, 1.94, 1.74, respectively. Additionally, the new data of three PCs can explain 98.8% of total variation of the original data. Each PC is calculated from 225 fluorescence intensities using the coefficients shown in Table F.1, and three PCs of the 40 calibration samples is in Table F.2. Now higher correlated 225 fluorescence intensities were represented by only three independent PCs. Then, MLR on PCs was applied for DOC and THMFP prediction.

6.3.2 Parameters Selection for Regression Analysis

"Best Subsets," an algorithm in Minitab15, is a technique used for selection of predictor variables into the regression models based on the following criteria: R-square (R^2), adjusted R-square (Adj. R^2), Mallows' Cp, and Root mean square error (RMSE). R^2 shows how well the model fit the data. Use R^2 for a regression model with single predictor variable while Adj. R^2 for multiple predictor variables. Mallows' Cp is an indicator to show if the model is over-fitting. RMSE is square root of mean of error square of prediction values from the model. The best possible models should have (1) high R^2 and Adj. R^2, (2) Mallows' Cp closest to number of predictor variables and constant in the models, (3) low RMSE (Anderson et al. 2002; Moore and McCabe, 2003; Kutner et al. 2004; Minitab 15, 2007). Then, MLR can be performed to obtain the best

regression model according to "Best Subsets" results, based on the three above criteria. Using the same process, DOC and THMFP prediction models were developed from regression analysis of PCs as shown in Figure 6.2 and Figure 6.3, respectively.

The results of "Best Subsets" in Figure 6.2(a) show the possible seven alternative regression models for DOC prediction. The best DOC model should include all three PCs, because it has highest R^2 and Adj. R^2, its Mallows' Cp equal to four, and lowest RMSE. Then regression analysis for DOC prediction was performed using all three PCs as predictor variables. The DOC regression model is shown below:

$$DOC = 5.75 + 0.199\ PC1 + 0.152\ PC2 - 0.940\ PC3 \quad (6.1)$$

As shown in the Figure 6.2(b), the DOC model passes F-test as its P-values less than 0.05 (n=40). This means that at the regression model is statistically significant to DOC (Bee et al. 1998; Kutner et al. 2004).

The Durbin-Watson statistic is a statistical value used to detect the presence of autocorrelation in the residuals from a regression analysis. Residuals are defined as actual values minus predicted values. The Durbin-Watson statistic is a number that always lies between zero and four. A value of two indicates no autocorrelation. As a rough rule of thumb, if the Durbin-Watson statistic is less than one, autocorrelation of the residuals could be the problem (Minitab 15, 2007). Because Durbin-Watson statistic of DOC model is 1.4, this model shows no problems of autocorrelation or it means that the model likely includes all the predictor variables in the regression model.

Variance Inflation Factor (VIF) indicates the extent to which multicollinearity (correlation among predictors) is present in a regression analysis. Multicollinearity is problematic because it can increase the variance of the regression coefficients, making

them unstable and difficult to interpret (Kutner et al. 2004). If VIF = 1 (the minimum value), predictor variables are not correlated. If 1 < VIF < 5, predictor variables are moderately correlated. If 5 < VIF < 10, predictor variables are highly correlated. If VIF > 10, multicollinearity is a serious problem in the model. The DOC model has VIF = 1.40, which indicates that multicollinearity is not problematic. In addition, Adj. R^2 shows that the equation can explain 95.7% of the total variation of raw data. This means the model fit the data well.

Figure 6.3(a) shows "Best Subsets" results for the THMFP model. Among seven alternatives, the best THMFP model should include all three PCs, because it has highest R^2 and Adj. R^2, Mallows' Cp equal to four, and lowest RMSE. If regression analysis was performed using three PCs, the THMFP model is shown in Figure 6.3(b). However, after regression analysis was performed, PC2 didn't pass the t-test at 95% confidence, as its P-value was much larger than 0.05. This implied that PC2 was not statistically significant for THMFP prediction. Hence, only PC1 and PC3 were included in the THMFP model. Regression analysis was performed again including only PC1 and PC3 as shown in Figure 6.3(c). The THMFP regression model is shown below:

$$\text{THMFP} = 423 + 19.8\ \text{PC1} - 83.6\ \text{PC3} \qquad (6.2)$$

The model passed all the criteria, as shown in the ANOVA tables. The model passed F-test as P-value less than 0.05. The Durbin-Watson statistic showed no autocorrelation between the residuals. The VIF was equal to one meaning no multicollinearity among predictor variables. Adj. R^2 was also good at 85.81%.

Prediction Sum of Square (PRESS) is used as criterion to justify a good prediction performance. PRESS is obtained by deleting one data point, estimating the regression

model from the remaining data sets, and then using the regression equation to predict the response of the deleted data. PRESS is then the summation of squared prediction error over all the number of data sets. Thus, a model with small PRESS value fits well as it has small prediction error. As illustrated in Figure 6.3(b) and Figure 6.3(c), THMFP model including PC1 and PC3 had higher PRESS, or it had better prediction performance. Moreover, the "Best Subsets" results in Figure 6.3(a) also showed good values of Adj. R^2, Mallows' Cp, RMSE for THMFP model using PC1 and PC3.

The DOC model's prediction performance was better than THMFP model's as it had higher Adj. R^2, lower PRESS, and lower RMSE. A larger variation of THMFP could possibly be due to reasons such as halide ions interferences (Richardson et al. 2003; Chow et al. 2007) or suspended colloidal particles (particles with diameter of 5-200 nanometers) in the samples after filtration (Chow et al. 2005).

Model validation is a measure of the prediction performance of a model. The validation process is an essential step to make sure that the model works quantitatively for similar data sets. The validation data set should be treated similarly with respect to sampling techniques, sample preparation, and water quality measurement methods. Full cross-validation was done by removing one data point (predictor variables: X, response: Y) from the calibration data set, then developing the model from the remaining calibration data set, and estimating the prediction error. The same process was repeated for the all the data points in the calibration data set (Esbensen, 2004). The full cross-validation of DOC and THMFP models using SFS are demonstrated in Figure 6.4(a) and 6.4(b), respectively. The diamonds represent predicted results, while the stars are laboratory results.

6.4 Models Development by GLR Using DOC or UV as Predictors

MLR was used to develop DOC prediction model from UV absorption. Figure 6.5 presents a strong correlation between DOC and UV absorption with Adj. R^2 of 92.5% and the regression equation is shown below.

DOC = 0.759 + 25.1 UV, where UV ranges from 0.0748 – 0.6194 1/cm (6.3)

As shown in the above mathematical relationship, when the sample had no UV absorption, its DOC was 0.759 mg/L. This implied that some organic compounds may not absorb UV.

Not only DOC but THMFP also had high correlation with UV absorption as illustrated in Figure 6.6 with Adj. R^2 of 85.1%. In contrast, THMFP also has strong correlation with DOC as shown in Figure 6.7 with Adj. R^2 of 94.5%, respectively. The mathematical relationships of both correlations are shown below.

THMFP = -65.2+ 2459 UV, where UV ranges from 0.0748 – 0.6194 1/cm (6.4)

THMFP = - 146 + 99.0 DOC, where DOC ranges from 2.3093 to 15.3625 mg/L (6.5)

The intercepts of both equations are negative. This means that the samples could have a very low level of either DOC or UV absorption while their THMFP values are zero. As seen in equation 6.4, when THMFP is zero, UV is 65.2 divided by 2459 equal to 0.0265. Or, samples with UV absorptions less than 0.0265 cm^{-1} may not produce THMs. Similarly, if THMFP of a sample is zero in equation 6.5, then DOC is 146 divided by 99.0 and become 1.47 mg/L. This means that from this calibration sample data set, in average, samples having DOC equal or less than 1.47 mg/L may not have THMs formation.

One of the assumptions for general linear regression is that predictor variables are independent of one another or they are not correlated (Kutner et al. 2004). Therefore,

THMFP model can not include both UV and DOC in the same equation as UV and DOC are highly correlated as shown in equation 6.3.

6.5 Comparison of Models Performances

The DOC prediction models from SFS and UV absorption are compared in Table 6.1. It shows that the DOC model from SFS (equation 6.1) has higher Adj. R^2, lower RMSE, and lower PRESS or it has better prediction performance. This means SFS is a better DOC prediction than UV absorption.

THMFP prediction models using SFS, UV, and DOC are compared in Table 6.2. THMFP model from DOC (equation 6.5) has better Adj. R^2, lower RMSE, and lower PRESS than the other two models, or DOC could be a better predictor variable for THMFP than SFS or UV absorption based on this calibration data sets. However, equation 6.5 could give a large error if DOM in the sample is not reactive to THMs formation, while equation 6.2 targets chemically discrete component to THMs formation regardless of the total DOM.

Comparing equations 6.2 and 6.4, their Adj. R^2 and RMSE are quite close but PRESS of equation 6.2 is much lower (612431 vs. 734615). This means that equation 6.2 has better prediction capability or SFS is a better predictor variable than UV absorption for THMFP estimation.

6.6 Limitations of the Models

The DOC models from SFS use the principle components calculated from selected 225 fluorescence intensities of DOM in the samples as predictor variables. However, it should be noted that not all types of organic compounds are equally sensitive to fluorescence spectroscopy. In other words, the models could be used for similar types of organic compounds having about the same fluorescence sensitivity. Then, the fluorescence intensities can be related to their concentrations accurately. Similarly, DOC prediction using UV absorption would be accurate when DOM in the sample absorb UV.

In addition, not all parameters affecting THMs formation are included in the THMFP model. Environmental conditions and incubation conditions such as temperature, pH, chlorine dosage and contact time, trace ions also greatly impact THMs formation (Amy et al. 1998; Sohn et al. 2004; Uyak et al. 2005). Salinity and pH of the samples also increase the yield of fluorescence spectra (Zepp et al. 2004). However, these variables are not included in this model.

The presence of bromide ion (Br^-) and iodine ion (I^-) further enhances the formation of the bromated and iodated DBP species and thereby increasing total THMs formation (Richardson et al. 2003; Chow et al. 2007). However, these ions were not measured in this study nor incorporated in the models.

All samples were incubated at controlled conditions (25 ^{o}C, pH = 7, 3-5 mg/L free chlorine residual, in dark for seven days) according to standard method 5710 B. Therefore, this model should be used to predict THMs formation at standard incubation conditions or to give an idea of the impact from DOM to THMs formation. These

incubation conditions for THMFP usually give the highest yield of THMs formation and may not reflect the actual conditions in the distribution system.

At very low DOC or low UV absorption (i.e., DOC < 1.47 mg/L or UV < 0.0265 cm^{-1}), THMFP models cannot be applied for THMs formation prediction. However, the THMFP model from SFS may be applied for all DOC and UV ranges investigated.

6.7 Chapter Summary

DOC and THMFP models give a good prediction, as they passed many statistical criteria (e.g. t-test, F-test, VIF, Durbin-Watson statistic test) used in model development. In addition, THMFP model does not account for the effects of interferences, environmental conditions, or incubation conditions, as discussed previously. The unknown samples should be collected using the same techniques and treated in the same way with the calibration samples to ensure good prediction performance.

Table 6.1 Comparison of DOC Models

DOC Models	Equation	Adj. R^2	RMSE	PRESS
DOC = 5.75 + 0.199 PC1 + 0.152 PC2 - 0.940 PC3	(6.1)	95.7%	0.680	30.4
DOC = 0.759 + 25.1 UV	(6.3)	92.5%	0.897	37.8

Table 6.2 Comparison of THMFP Models

THMFP Models	Equation	Adj. R^2	RMSE	PRESS
THMFP = 423 + 19.8 PC1 - 83.6 PC3	(6.2)	87.6%	117.3	612431
THMFP = -65.2 + 2459 UV	(6.4)	85.1%	128.3	734615
THMFP = -146 + 99.0 DOC	(6.5)	94.5%	78.4	255549

Eigenvalue	218.70	1.94	1.74	0.34	0.28	0.26	0.18	0.16	0.15
Proportion	0.972	0.009	0.008	0.002	0.001	0.001	0.001	0.001	0.001
Cumulative	0.972	0.981	0.988	0.990	0.991	0.992	0.993	0.994	0.994

Figure 6.1 Principle component analysis of fluorescence intensities.

Best Subsets Regression: DOC versus PC1, PC2, PC3

Response is DOC

Vars	R-Sq	R-Sq(adj)	Mallows Cp	S	PC1	PC2	PC3
1	81.2	80.7	133.3	1.4362	X		
1	14.4	12.1	734.1	3.0632			X
1	0.4	0.0	859.9	3.3038		X	
2	95.6	95.3	5.8	0.70555	X		X
2	81.6	80.6	131.5	1.4391	X	X	
2	14.8	10.2	732.3	3.0967		X	X
3	96.0	95.7	4.0	0.68043	X	X	X

Figure 6.2(a) Best subsets results for DOC model.

Regression Analysis: DOC versus PC1, PC2, PC3

The regression equation is
DOC = 5.75 + 0.199 PC1 + 0.152 PC2 - 0.940 PC3

Predictor	Coef	SE Coef	T	P	VIF
Constant	5.7545	0.1076	53.49	0.000	
C239	0.199108	0.007368	27.02	0.000	1.000
C240	0.15224	0.07828	1.94	0.060	1.000
C241	-0.93956	0.08256	-11.38	0.000	1.000

S = 0.680429 R-Sq = 96.0% R-Sq(adj) = 95.7%

PRESS = 30.4372 R-Sq(pred) = 92.69%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	399.85	133.28	287.88	0.000
Residual Error	36	16.67	0.46		
Total	39	416.52			

Durbin-Watson statistic = 1.40000

Figure 6.2(b) Regression analysis for DOC model.

Best Subsets Regression: THMFP versus PC1, PC2, PC3

Response is THMFP

Vars	R-Sq	R-Sq(adj)	Mallows Cp	S	PC1	PC2	PC3
1	77.2	76.6	37.2	160.87	X		
1	11.0	8.6	250.1	317.96			X
1	0.6	0.0	283.5	336.01		X	
2	88.2	87.6	3.9	117.31	X		X
2	77.8	76.6	37.3	160.88	X	X	
2	11.6	6.8	250.2	321.15		X	X
3	88.8	87.9	4.0	115.88	X	X	X

Figure 6.3(a) Best subsets results for THMFP model.

Regression Analysis: THMFP versus PC1, PC2, PC3

The regression equation is
THMFP = 423 + 19.8 PC1 + 18.5 PC2 - 83.6 PC3

Predictor	Coef	SE Coef	T	P	VIF
Constant	423.36	18.32	23.11	0.000	
C239	19.767	1.255	15.75	0.000	1.000
C240	18.47	13.33	1.39	0.174	1.000
C241	-83.56	14.06	-5.94	0.000	1.000

S = 115.877 R-Sq = 88.8% R-Sq(adj) = 87.9%

PRESS = 670118 R-Sq(pred) = 84.47%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	3832733	1277578	95.15	0.000
Residual Error	36	483391	13428		
Total	39	4316124			

Durbin-Watson statistic = 1.08178

Figure 6.3 (b) Regression analysis for THMFP model using all three PCs as predictor variables.

Regression Analysis: THMFP versus PC1, PC3

```
The regression equation is
THMFP = 423 + 19.8 PC1 - 83.6 PC3


Predictor     Coef  SE Coef      T      P    VIF
Constant   423.36    18.55  22.82  0.000
C239       19.767    1.270  15.56  0.000  1.000
C241       -83.56    14.23  -5.87  0.000  1.000


S = 117.308   R-Sq = 88.2%   R-Sq(adj) = 87.6%

PRESS = 612431   R-Sq(pred) = 85.81%


Analysis of Variance

Source          DF       SS       MS       F      P
Regression       2  3806964  1903482  138.32  0.000
Residual Error  37   509159    13761
Total           39  4316124


Durbin-Watson statistic = 1.30692
```

Figure 6.3(c) Regression analysis for THMFP model using PC1 and PC3 as predictor variables.

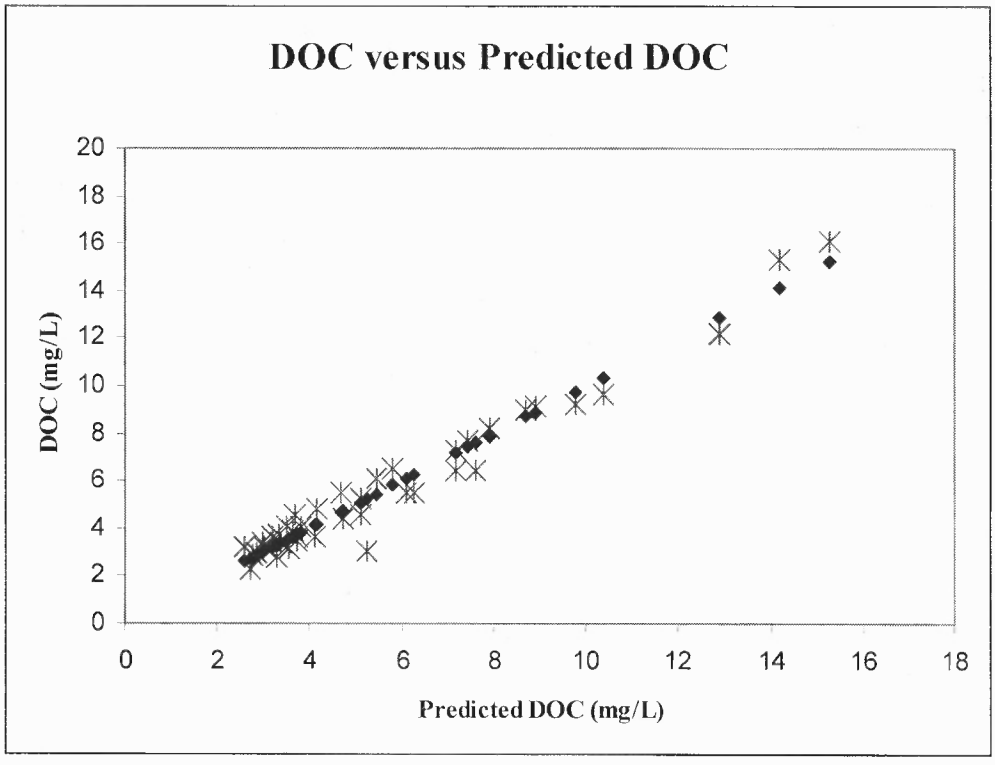


Figure 6.4 (a) Comparison of laboratory DOC and predicted DOC.

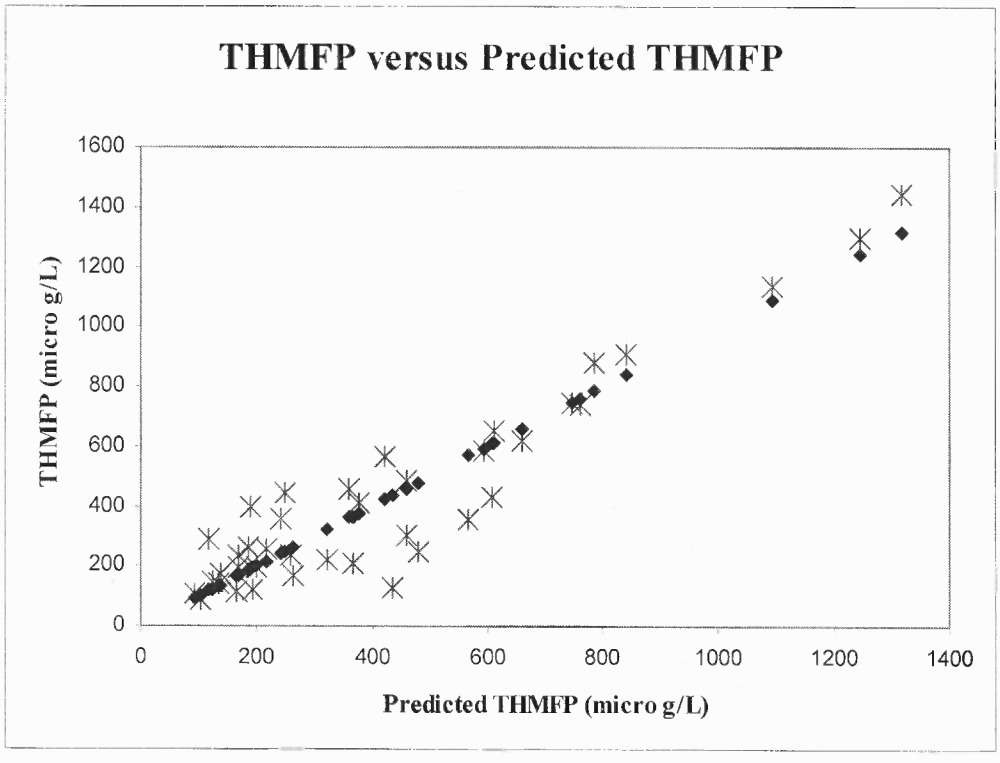
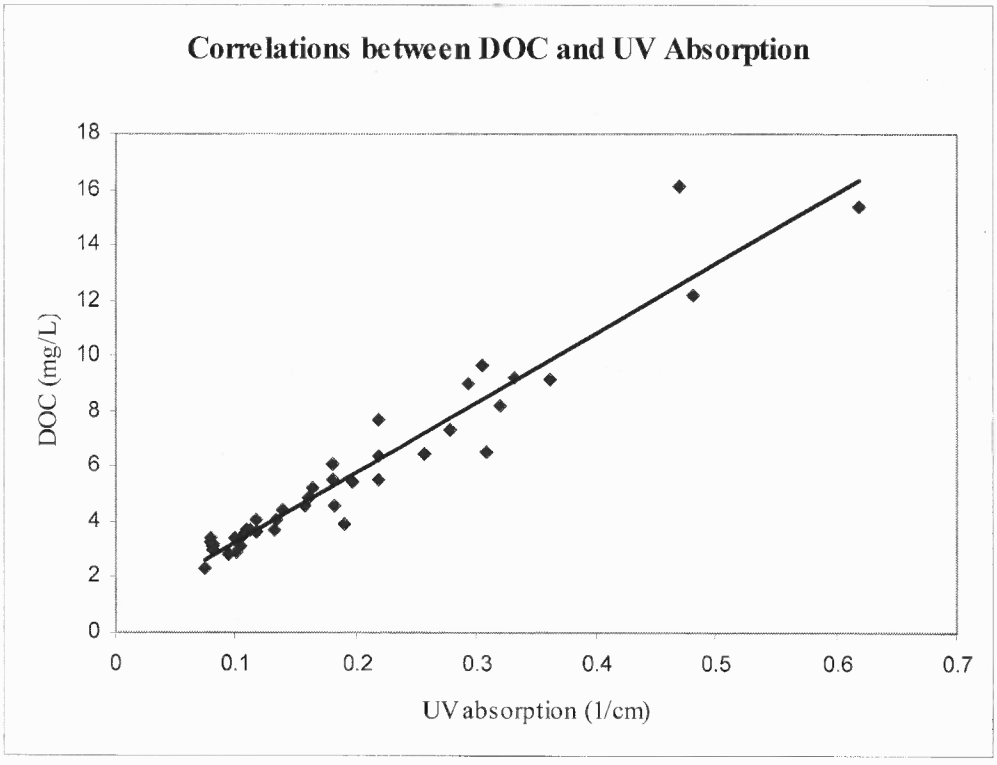


Figure 6.4(b) Comparison of laboratory THMFP and predicted THMFP.



Regression Analysis: DOC versus UV

The regression equation is
DOC = 0.759 + 25.1 UV

Predictor	Coef	SE Coef	T	P	VIF
Constant	0.7589	0.2688	2.82	0.008	
UV	25.143	1.149	21.89	0.000	1.000

S = 0.897497 R-Sq = 92.7% R-Sq(adj) = 92.5%

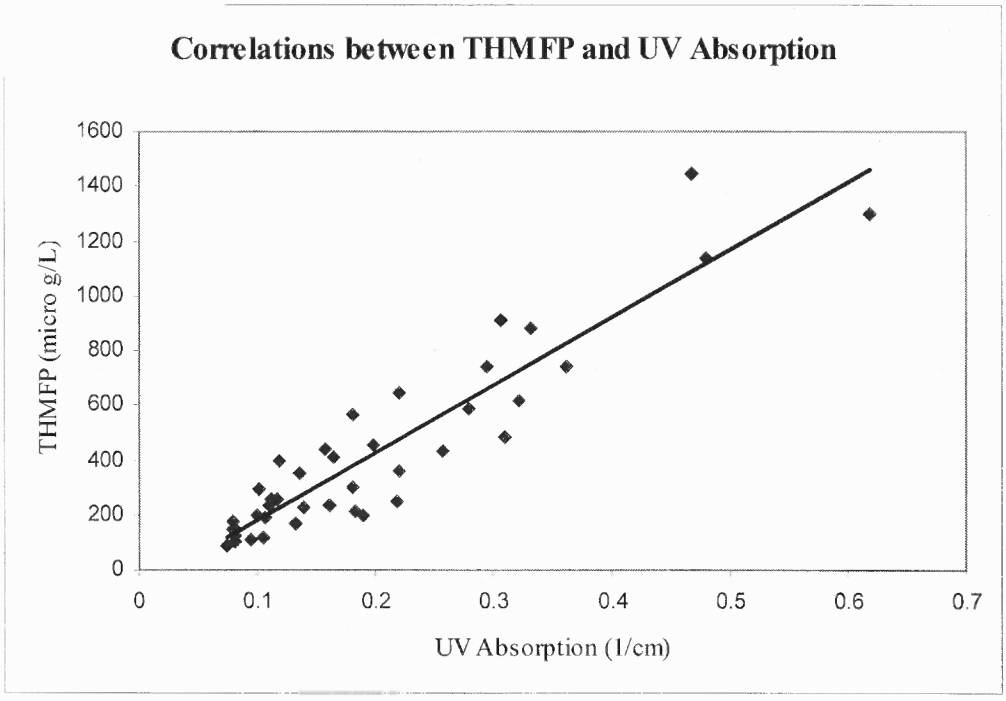
PRESS = 37.8492 R-Sq(pred) = 90.91%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	385.91	385.91	479.09	0.000
Residual Error	38	30.61	0.81		
Total	39	416.52			

Durbin-Watson statistic = 1.75102

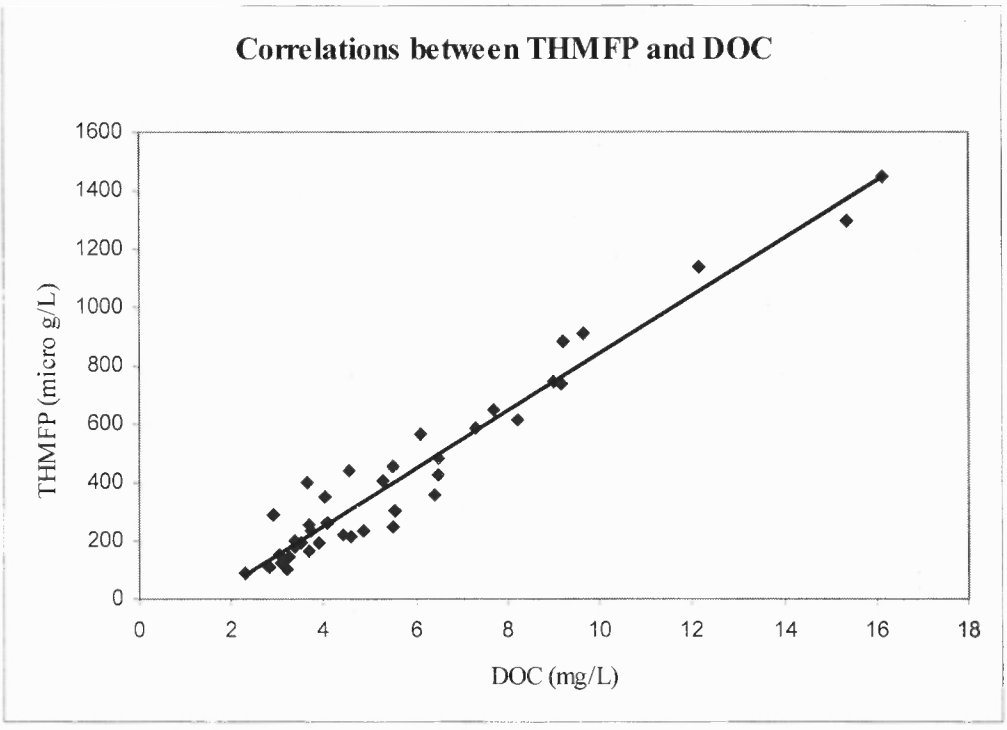
Figure 6.5 Correlations between DOC and UV absorption.



Regression Analysis: THMFP versus UV absorption

```
The regression equation is
THMFP = - 65.2 + 2459 UV


Predictor     Coef  SE Coef      T      P    VIF
Constant   -65.18    38.42  -1.70  0.098
UV         2458.8    164.2  14.97  0.000  1.000


S = 128.294   R-Sq = 85.5%   R-Sq(adj) = 85.1%

PRESS = 734615   R-Sq(pred) = 82.98%


Analysis of Variance

Source           DF       SS       MS       F      P
Regression        1  3690666  3690666  224.23  0.000
Residual Error   38   625458    16459
Total            39  4316124

Durbin-Watson statistic = 1.69795
```

Figure 6.6 Correlations between THMFP and UV absorption.



Regression Analysis: THMFP versus DOC

The regression equation is
THMFP = - 146 + 99.0 DOC

Predictor	Coef	SE Coef	T	P	VIF
Constant	-146.37	25.33	-5.78	0.000	
DOC	99.006	3.840	25.78	0.000	1.000

S = 78.3641 R-Sq = 94.6% R-Sq(adj) = 94.5%

PRESS = 255549 R-Sq(pred) = 94.08%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	4082768	4082768	664.85	0.000
Residual Error	38	233355	6141		
Total	39	4316124			

Durbin-Watson statistic = 1.23505

Figure 6.7 Correlations between THMFP and DOC.

CHAPTER 7

CONCLUSIONS AND RECOMMENDATIONS

7.1 Conclusions

DOM is a complex mixture of aromatic and aliphatic hydrocarbon structures that have various attached functional groups. However, not all organic compounds are equally reactive to THMs formation. Resin adsorption was used to isolate DOM in source water into six fractions. The three major fractions in most of samples were HPIN, HPON, and HPOA. HPOA fraction showed the highest reactivity to THMs formation, followed by HPIN and HPON. The other three fractions were usually found at lower concentrations and were not as reactive to THMs formation.

SFS was applied for the characterization of water samples and was found very sensitive to the HPOA fraction (i.e., humic and fulvic acids). Not only was HPOA the most reactive to THMs formation but also one of the most abundant fractions. Following the removal of scattering and interferences, two hundred and twenty five (225) fluorescence intensities covering peaks regions of HPOA fraction were selected for the prediction of THMs formation. PCA was used to reduce a number of fluorescence intensities into a few PCs. Then, the Best Subset Algorithm was performed to select the most significant predictor variables into regression models by the following criteria: Adj. R^2, Mallows' Cp, and RMSE. Regression analysis was performed according to the Best Subset results for DOC and THMFP prediction. The DOC prediction models using SFS showed better performance over the models that used UV absorption. For THMFP prediction, either DOC, SFS, or UV absorption could be used as predictor variables. The THMFP prediction model from DOC showed better prediction performance than the

model from SFS and UV absorption for the source water investigated. Nonetheless, the model from DOC cannot be used when DOC < 1.47 mg/L, while the model from SFS may be applied at all DOC ranges investigated.

SFS is a fast, inexpensive, reagent-free characterization technique compared to DOC measurement. Moreover, SFS characterized the most reactive DOM to THMs formation, while DOC measures the total DOM concentration. In cases where DOM is not very reactive to THMs formation, SFS could give a better THMFP prediction. There are limitations to the models. Some organic compounds may not be as sensitive to fluorescence spectroscopy as others. The DOC model could give an accurate prediction if DOM in the samples is sensitive to fluorescence spectroscopy. The THMFP model predicted the THMs formation potential based on seven-day incubation at standard conditions. These conditions may not reflect the real conditions in the distribution system and may predict the worst case scenarios. However, the models could be used for qualitative prediction of THMs formation.

7.2 Recommendations

SFS is a rapid, inexpensive, reagents-free technique for characterization of organic compounds in source water. Thus, this technique could be applied at water treatment plants and source water location for spatial and temporal characterization. A sample set can be collected and analyzed for DOC, fluorescence spectroscopy, and THMFP. These data sets can then be used for calibration and model development. Unknown samples are characterized by PCA coupled with MLR. DOC and THMFP values are predicted promptly from the SFS of the water samples. The rapid response from the models would

determine the organic characteristics and THMFP of the source water, allowing ample time for the water utilities to optimize their treatment processes to cope with the dynamic changes in water quality. In addition, the technique can be applied by watershed managers to better control point and non-point sources of problematic organics that impact water quality.

Future research could be the development of fluorescence spectroscopy for real-time measurement of labor-intensive tests of water quality parameters (e.g., COD, BOD, organic character, and THMFP)

APPENDIX A

STANDARD OPERATING PROCEDURE FOR TOTAL/DISSOLVED ORGANIC CARBON (TOC/DOC)

This follows the protocols outlined in Standard Methods 5310 B.

Sample preparation:

- Pre-treat filter and filter assembly by submerging overnight in 10% HNO_3 and 5% HCl solution and washing with tap water and organic-free water.
- Dry pre-treated filter and filter assembly in an oven at 550°C overnight.
- Wash the cooled filter and filter assembly by pass 100-ml organic-free water three times.
- Filter enough water samples to remove colloidal particles.
- Pre-treat sample test vials same as for filter and filter assembly.

Measurement procedure:

- Instrument operation- Set up instrument according to manufacture's instruction. Specially monitor flow rate to insure stability.
- Sample treatment- Adjust sample pH between 2 and 10 with H_3PO_4/NaOH if necessary.
- Standard Curve preparation- Prepare DOC standards by diluting stock solution to cover the expected range in samples. Run standards and draw standard curve.
- Sample test

Quality assurance for TOC:

- Sample test: triple runs for each.
- Standard runs: after every five samples and three replicates for each standard
- RSD: ±4% for each sample or standard.

- Performance verification: ±4% of the "real" concentration of standard.
- R^2 of the standard curve: > 99.99%.
- Organic free water.

APPENDIX B

STANDARD OPERATING PROCEDURE FOR UV ABSORBANCE

UV absorbing constituents measured following the protocols outlined in Standard Methods 5910 B.

Sample preparation:

- Dry pre-cleaned 1000 ml filter and filter assembly in an oven at 550°C overnight.
- Wash the cooled filter and filter assembly by passing 100-ml organic-free water three times.
- Filter enough water samples to remove colloidal particles.

Measurement procedure:

- Equilibrate the UV instrument according to manufacture's instruction.
- Prepare two 1-cm-quartz-cell organic-free water blanks.
- Check UV lamp response by scanning the water blanks from 225 nm to 900 nm.
- Set wavelength at 254 nm.
- Zero UV instrument's background adsorption.
- Measure sample UV adsorption.

Quality assurance for UV:

- After every five-sample run, check background adsorption and zero it.
- Analyze every tenth sample in duplicate to assess instrument precision.
- Verify UV instrument performance according to the standard method 5910B. Wash sample cell between runs.

APPENDIX C

STANDARD OPERATING PROCEDURE FOR SPECTRAL FLUORESCENCE SIGNATURES (SFS)

This follows the procedures developed for the Hitachi F-4500 instrument.

Sample preparation:

- Dry pre-cleaned 1000 ml Filter and filter assembly in an oven at 550°C overnight.
- Wash the cooled filter and filter assembly by pass 100-ml organic-free water three times.
- Filter enough water samples to remove colloidal particles.
- Adjust sample pH around 7.

Sample Measurement:

- Turn on the instrument.
- Warm the instrument for 30 minutes.
- Perform sensitivity and wavelength accuracy tests following the manufacture instruction.
- Select three-dimension mode on water sample.

Set up sample measurement parameters:

Parameters	Excitation (nm)	Emission (nm)
Starting	225	249
Ending	525	633
Sampling interval	6	6
Slit	5.0	10.0

Response: Auto

PMT Voltage: 950 v

Scan speed: 30000 nm/min.

Conduct sample tests.

Quality assurance for fluorescence:

$$\frac{\text{Signal}}{\text{Noise}} \geq 100\% \text{ and Drift} \leq 2\%$$

- Perform sensitivity test.
- Raman peak wavelength: 398±1 nm
- Perform wavelength Accuracy test.
- Signature peak wavelength: 450±2 nm
- Acid solution (10% HNO_3 and 5% HCl) soaks sample cell overnight, and then organic-free water rinses cell.

APPENDIX D

STANDARD OPERATING PROCEDURE FOR CHLORINE INCUBATION

Trihalomethane Formation Potential is carried out following the protocols outlined in Standard Methods 5710 B. Water samples are buffered at pH 7 $\pm$ 0.2, chlorinated with excess free chlorine, and stored at 25 $\pm$ 2 ^{o}C for 7 days. The free chlorine residual should be in between 3-5 mg/L at the end of reaction time. Nitrogenous species may interfere in the determination of free residual chlorine. Add enough free chlorine to oxidize chlorine demand substances and leave free chlorine residual at least 3 mg/L.

The free chlorine residual is measured using standard method 4500 F: DPD Ferrous Tritimetric Method. In this method N,N-diethyl-p-phenylenediamine (DPD) is used as indicator in the titrimetric procedure with ferrous ammonium sulfate (FAS). The procedures are outlined below

- Place five ml of each Phospate buffer solution (pH 6.2-6.5) and DPD indicator solution with 100 ml sample and mix well. Now the solution turns red. If the sample is added before buffer, the test does not work.
- Titrate rapidly with standard FAS titrant until the red color discharged.
- One ml of FAS is equivalent to 0.100 mg free chlorine.

The incubation procedures are out lined as follows.

- Adjust the pH of the samples to 7 $\pm$ 0.2 using H_2S0_4 and NaOH.
- Add 1 ml phosphate buffer solution for each 60 ml of incubation bottle.
- Add appropriate amount of standard sodium hypochlorite solution obtained previously from chlorine demand test.
- Fill up the incubation bottle completely with the sample.
- Incubate the samples at 25 $\pm$ 2 °C in dark for seven days

- After seven days, the samples were dechlorinated using sodium sulfite (Na_2SO_3), and their pHs were adjusted to pH around 4.5 to 5.5 by the addition buffer salt.

- Buffer salt for pH 4.5 to 5.5 was prepared by a mixture of 2 g Na_2HPO_4, 198 g KH_2PO_4, and 1.2 Na_2SO_3. Then buffer salt is washed by methanol 4 times, acetone 2 times, and MTBE 2 times to remove the impurity.

- About ten ml of sample was removed, and 3 ml of MTBE, 20 g of Na_2SO_4 was added. Shake vigorously for four minutes. THMs were extracted from the sample into MTBE phase and were measured by Varian GC-ECD 3400.

APPENDIX E

STANDARD OPERATING PRODCEDURE FOR DECTION OF THMFP BY VARIAN GC-ECD 3400

Determination of the THMs and HANs by Method 551.1 is Liquid-Liquid Extraction Gas Chromatography (LLE GC). This method is applicable to the determination of halogenated acetic acids in all drinking, ground, or raw water. Briefly, a 40 mL volume of sample is adjusted to pH < 0.5 and extracted with 4-mL of MTBE. The haloacetic acids that have been partitioned into the organic phase are then converted to their methyl esters by the addition of the acidic methanol followed by slight heating. The acidic extract is neutralized by a back-extraction with a saturated solution of sodium bicarbonate and the desired analytes are identified and measured by capillary column gas chromatography using an electron capture detector (GC/ECD).

1. Scope and Application

- This method is used to determine the concentrations of four trihalomethanes in extracts from aqueous samples.
- The four trihalomethanes listed below have been tested by this method: Chloroform, Bromodichloromethane, Bromoform, Dibromochloromethane.
- The THMs are separated with two parallel columns sharing a single injection port and detected with two Electron Capture Detectors (ECDs).

2. Summary of Method

- A 50 ml of sample is extracted with 3 ml MTBE.
- After extraction, 0.25 ml of extract phase is transferred into a 0.25 ml insert out-packed with two ml GC vials.
- Extracts are analyzed by injecting one µl aliquot together with 1µl internal standard into the gas chromatograph equipped with two parallel columns each

followed with its own ECD.

- The total run is 45 minutes.

3. Interferences

- Sources of interference in this method can be grouped into two categories: Contaminated solvents, reagents, and sample processing hardware; Contaminated GC carrier gas, parts, column, and detectors.

- The extracting solvent MTBE has been found containing residual chloroform. Distillation of MTBE between 50 and 60 °C can decrease but not eliminate background level of chloroform. Solvent background is examined for subtraction from sample response.

- Samples, standards, and reagents are stored in sealed amber vials with Teflon caps to prevent any contamination in the laboratory.

4. Sample Handling and Preservation

- Incubated samples for THMFP are in 68 ml vials capped with Teflon lined septa.

- Extracted samples are in 0.25 ml clear GC inserts placed into two ml amber GC vials to protect from light.

- Extracted samples and standards are stored under 4 °C in a refrigerator.

5. Laboratory Apparatus

- Varian 3300 with CAC A200S Auto sampler and injector

- Primary column 30 m × 0.25 mm ID × 1.0 μm (RTX-1301, catalog # 10153)

- Confirmative column 30 m × 0.25 mm ID × 1.0 μm (RTX-1, catalog # 16053)

- EZ CHROM Elite software (Version 2.61)

- Carrier gas: ultra high purity grade He

- Make up gas: Nitrogen

6. Regents and Standards

- MTBE: HPLC grade (Fisher catalog # E127-4)

- Acetone: HPLC grade (Fisher catalog # A949-4)
- Stock standards

Name	Compound	Conc. (mg/L)	Catalog #	Company
THMs	Chloroform Bromodichloromethane Chlorodibromomethane Bromoform	2,000	3021	Restek
LPC	Bromodichloromethane Hexachlorocyclopentadiene G-BHC Trichloroethene	30 20 0.2 30	M-551.1-MLPC-PAK	AccuStandard
IS		10,000	M-551.1-IS-100X	AccuStandard
Surrogate	Decafluorbiphenyl	1,000	M-551.1-SS-100X	AccuStandard

- Primary and working standard

LPC	Stock	Primary	Working
Solvent	MTBE	MTBE	MTBE
Concentration	Varied	Varied	Varied
Preparation	Commercial	1 ml of stock into 10 ml MTBE	100 µl of primary into 10 ml MTBE

Surrogate	Stock	Primary	Working
Solvent	Acetone	Acetone	Buffer/dechlorating solution
Concentration	1000 mg/L	10 mg/L	10 µg/L
Preparation	Commercial	500 µl of stock into 50 ml Acetone	50 µl of primary into 50 ml buffer/dechlorating solution

IS	Stock	Primary	Working
Solvent	Acetone	Acetone	MTBE
Concentration	10,000 mg/L	100 mg/L	1 mg/L
Preparation	Commercial	500 µl of stock into 50 ml Acetone	1 µl of primary directly into GC injector

7. Procedure

7.1. Laboratory Performance Check

- Directly inject one µl LPC working solution into GC for five repetitions under the following conditions:

Initial column temperature-	35 °C
Initial column hold time -	22 min
Program 1 final column temperature -	145 °C
Program #1 column temperature rate -	10 °C/min
Program #1 final temperature holding time-	2 min
Program #2 final column temperature -	225 °C
Program #2 column temperature rate -	20 °C/min
Program #2 final temperature holding time-	1 min
Program #3 final column temperature -	250 °C
Program #3 column temperature rate -	25 °C/min
Program #3 final temperature holding time-	3 min
Injector temperature -	150 °C
Detector temperature -	290 °C
Detector A initial range -	1
Detector B initial range -	1

- Check if LPC performance meets criteria listed in table below.

Parameter	Analyte	Conc. (µg/ml)	Criteria
Sensitivity	G-BHC	0.00020	S/N > 3
Chromatographic performance	Hexachlorocyclopentadiene	0.020	0.8 < PGF < 1.15
Column performance	Bromodichloromethane Trichlorethylene	0.030 0.030	Resolution > 0.50

PGF (Peak Gaussian Factor) = $\frac{1.83 \times W\,(1/2)}{W\,(1/10)}$

Where W (1/2) is the peak width at half height, W (1/10) is the peak width at tenth peak height

Resolution = $\frac{Rt2 - Rt1}{(W1 + W2)/2}$, Where $Rt2 - Rt1$ is the retention time difference between the two designated compounds and $(W1 + W2)/2$ is the average peak width between them

S/N is defined as the average response of the five runs of LPC working standards divided by the standard deviation.

7.2. Initial Calibration

- Initial calibration is to determine target average relative response factor (RRF) and retention time (R.T.) as followings:

$$RRF = \frac{[\text{Response/conc.}]_{\text{target}}}{[\text{Response/conc.}]_{\text{IS}}}$$

RT window = average retention time ± 3 standard deviation.

- Five calibration standards, containing each individual component of THM, are prepared at 0, 5, 20, 50, and 100 μg/L.
- Inject each standard with three repetitions.
- Determine average RRF and RT windows for each compound.
- The initial calibration must confirm a less than 15% relative standard deviation (%RSD) for each RRF and surrogate response. For internal standard, the % RSD of response must be less than 10%. Otherwise, investigate possible causes of exceeding these criteria.

7.3. Method Blank

In this SOP, the method blank is the same as the standard 0 μg/L

7.4. Calibration Verification

- The calibration verification or so-called continuous calibration is to check the RRF and RT window determined during initial calibration is still applicable.
- Inject a mixture standard of 50 μg/L three times (standard solution from providers other than the one of initial calibration is preferred)
- Calculate RRF and R.T. for each compound.
- Calibration verification should verify that RRF of each target is within ± 20% of that determined in initial calibration and R.T. falls into RT window determined in initial calibration. Response of internal and surrogate standards should be less than ± 20% of those determined in initial calibration.

7.5. Blank and Blank Spike

- The blank sample is different from method blank in that blank sample is subject to seven days incubation.

- Blank is to examine possible contamination during incubation.
- Blank spike is prepared at the same time as Blank by adding 34 μl of primary standard into Blank to make a final concentration of 50 μg/L for Blank spike.
- There should have no or insignificant amount of target compound able to be detected in Blank. The relative recovery of each target compound in Blank Spike should be within ± 10%.

7.6. Matrix and Matrix Spike

Except for original samples, all fraction samples are prepared with MQ water with some NaOH and H_2SO_4 (at most 0.1 N NaOH or H_2SO_4). It is presumed that no matrix interference presents during incubation and GC identification. It is expected that matrix of original sample is not harsh enough to result significant interference. This SOP suggests an optional run for matrix or matrix spike.

7.7. Sample

- When all above runs are confirmed to meet corresponding criteria, runs of samples can be initiated.
- For every run of 5 samples, a calibration check follows to examine if instrument is in acceptable running condition.
- An end check must be arranged at the end of each batch.
- Criteria for calibration check and end check is the same as Continuous Calibration.

APPENDIX F

PRINCIPLE COMPONENT ANALYSIS RESULTS

Table F.1 Principle Component Analysis Coefficients

Intensity number	PC1	PC2	PC3
1	0.056676	-0.270667	0.001247
2	0.064129	-0.134695	-0.097435
3	0.063303	-0.126198	-0.059284
4	0.064028	-0.090132	-0.022823
5	0.065784	-0.102148	0.045148
6	0.065591	-0.076907	0.040649
7	0.065973	-0.069237	0.045950
8	0.065528	-0.003459	0.077615
9	0.065504	-0.052426	0.110965
10	0.065610	-0.020987	0.090055
11	0.064775	-0.039105	0.126649
12	0.065557	-0.027271	0.145496
13	0.064724	-0.037191	0.125319
14	0.064146	-0.049417	0.146383
15	0.064218	0.010600	0.133077
16	0.061238	-0.210704	-0.130559
17	0.063721	-0.142367	-0.095087
18	0.065469	-0.118898	0.001354
19	0.065257	-0.136546	0.000886
20	0.066409	-0.077981	0.011202
21	0.065232	-0.117348	0.064899
22	0.065245	-0.085703	0.128851
23	0.065559	-0.090935	0.066976
24	0.065974	-0.076127	0.093176
25	0.065844	-0.066402	0.079588
26	0.066416	-0.031598	0.078089
27	0.065865	-0.076153	0.087001
28	0.065414	-0.053553	0.142743
29	0.065123	-0.018160	0.159196
30	0.065196	-0.025996	0.148000
31	0.063894	-0.192994	-0.042734
32	0.065148	-0.142413	0.009629
33	0.065563	-0.126667	-0.003701
34	0.066211	-0.109581	0.027905
35	0.066142	-0.088117	0.075260
36	0.066511	-0.079440	0.043520
37	0.066200	-0.069091	0.075421
38	0.066705	-0.070140	0.053698
39	0.066368	-0.063486	0.075059
40	0.066565	-0.036014	0.083316

Intensity number	PC1	PC2	PC3
41	0.066370	-0.027921	0.108687
42	0.066442	-0.031125	0.102062
43	0.066137	-0.020678	0.123532
44	0.066625	-0.034119	0.106622
45	0.066725	-0.014501	0.087276
46	0.064400	-0.171164	-0.054166
47	0.065405	-0.151306	-0.027162
48	0.066062	-0.127048	0.018771
49	0.066394	-0.096079	0.027595
50	0.066616	-0.079084	0.057925
51	0.066911	-0.039098	0.049696
52	0.066732	-0.047733	0.055032
53	0.066794	-0.043951	0.075748
54	0.066723	-0.046144	0.076804
55	0.067124	0.004894	0.066584
56	0.067160	0.003532	0.048989
57	0.067108	0.014323	0.066369
58	0.066520	-0.006725	0.095180
59	0.066550	0.014083	0.107635
60	0.066813	0.015905	0.089783
61	0.065510	-0.141384	-0.042821
62	0.066283	-0.111816	-0.017474
63	0.066525	-0.091010	-0.010994
64	0.066833	-0.060634	0.003087
65	0.067135	-0.020028	0.006557
66	0.067052	-0.029367	0.032515
67	0.067178	0.002829	0.047934
68	0.066957	-0.020843	0.063531
69	0.067004	-0.007326	0.082156
70	0.066890	0.031263	0.078184
71	0.067146	0.011548	0.052637
72	0.067048	0.026739	0.079993
73	0.067218	0.014767	0.061082
74	0.066584	0.036686	0.092745
75	0.067077	0.016320	0.071690
76	0.065335	-0.135772	-0.084025
77	0.066331	-0.106803	-0.046767
78	0.067020	-0.049532	-0.037192
79	0.067086	-0.047715	0.011114
80	0.066917	-0.028435	0.050607
81	0.067237	-0.026973	0.045896
82	0.067199	-0.018612	0.045550
83	0.067022	0.010876	0.074017
84	0.067095	0.035183	0.068893
85	0.066992	0.029393	0.073181
86	0.067090	0.049872	0.051508
87	0.067265	0.026178	0.052268
88	0.066875	0.047265	0.086841

Intensity number	PC1	PC2	PC3
89	0.066676	0.065564	0.081128
90	0.066639	0.064357	0.083440
91	0.066079	-0.119871	-0.066343
92	0.066559	-0.093527	-0.057026
93	0.067101	-0.057458	-0.024122
94	0.067246	-0.048768	0.002706
95	0.067434	-0.000016	-0.004048
96	0.067334	0.010467	-0.013669
97	0.067485	0.005364	-0.002267
98	0.067467	0.019167	0.011795
99	0.067257	0.044570	0.030301
100	0.067156	0.053225	0.032171
101	0.066935	0.069570	0.045896
102	0.066940	0.070701	0.050419
103	0.067030	0.057336	0.064907
104	0.066770	0.072925	0.061162
105	0.066218	0.077400	0.098688
106	0.065722	-0.103476	-0.122609
107	0.066443	-0.079319	-0.096819
108	0.067088	-0.046900	-0.058078
109	0.067150	-0.032892	-0.053780
110	0.067324	-0.014828	-0.039398
111	0.067431	0.001424	-0.037623
112	0.067426	0.025709	0.007906
113	0.067409	0.024834	-0.001511
114	0.067294	0.054172	0.010645
115	0.067170	0.059688	0.044575
116	0.067197	0.055948	0.030425
117	0.067076	0.067964	0.042222
118	0.066843	0.077045	0.059815
119	0.066841	0.073255	0.062418
120	0.066632	0.088915	0.064331
121	0.065559	-0.085322	-0.144161
122	0.066518	-0.061021	-0.095695
123	0.066590	-0.041259	-0.106626
124	0.067132	-0.013279	-0.073895
125	0.067217	0.013130	-0.065334
126	0.067323	0.025308	-0.056018
127	0.067276	0.025019	-0.047604
128	0.067325	0.047720	-0.030610
129	0.067372	0.040932	-0.000397
130	0.067256	0.056893	-0.004932
131	0.067184	0.066042	0.018070
132	0.067077	0.074170	0.030023
133	0.067029	0.078099	0.026094
134	0.066951	0.078959	0.039298
135	0.066831	0.087432	0.041097
136	0.065907	-0.065730	-0.136151

Intensity number	PC1	PC2	PC3
137	0.066398	-0.033562	-0.119574
138	0.066908	-0.029799	-0.096515
139	0.066906	-0.005934	-0.097417
140	0.067255	0.010510	-0.067468
141	0.067271	0.007632	-0.065794
142	0.067329	0.020839	-0.054769
143	0.067308	0.036560	-0.049855
144	0.067306	0.049314	-0.035720
145	0.067429	0.037964	-0.022988
146	0.067421	0.048153	-0.004091
147	0.067287	0.062315	-0.012055
148	0.067216	0.062106	0.004276
149	0.067039	0.079661	0.026367
150	0.066601	0.096722	0.041901
151	0.066386	-0.005990	-0.131885
152	0.066627	-0.020676	-0.120671
153	0.066910	0.000185	-0.103685
154	0.066948	0.003417	-0.094922
155	0.067183	0.013855	-0.075111
156	0.067182	0.022033	-0.075536
157	0.067347	0.039098	-0.047780
158	0.067280	0.041533	-0.052413
159	0.067326	0.041275	-0.040119
160	0.067309	0.056621	-0.029752
161	0.067349	0.058211	-0.012166
162	0.067319	0.057723	-0.018787
163	0.067303	0.060444	0.000163
164	0.067255	0.065789	0.002820
165	0.067000	0.077505	0.024678
166	0.066836	0.001220	-0.102873
167	0.067011	0.019641	-0.090182
168	0.067087	0.004125	-0.085967
169	0.067017	0.023429	-0.090975
170	0.067228	0.021111	-0.065577
171	0.067250	0.024035	-0.061446
172	0.067323	0.035994	-0.051032
173	0.067286	0.037647	-0.053986
174	0.067333	0.040058	-0.049197
175	0.067314	0.049815	-0.041064
176	0.067366	0.050008	-0.025897
177	0.067363	0.055515	-0.017215
178	0.067285	0.064970	-0.006890
179	0.067298	0.063997	-0.004017
180	0.067136	0.076165	0.010074
181	0.066725	0.018259	-0.114009
182	0.066999	0.028423	-0.085573
183	0.067151	0.025302	-0.076882
184	0.067212	0.025274	-0.070338

Intensity number	PC1	PC2	PC3
185	0.067293	0.023691	-0.059068
186	0.067252	0.031510	-0.058821
187	0.067269	0.045555	-0.052468
188	0.067270	0.042939	-0.047688
189	0.067349	0.042380	-0.039528
190	0.067274	0.056467	-0.035432
191	0.067293	0.056277	-0.034750
192	0.067308	0.054358	-0.032933
193	0.067347	0.059811	-0.016134
194	0.067246	0.070722	-0.009222
195	0.067210	0.074134	-0.007429
196	0.067275	0.030537	-0.062990
197	0.067141	0.034305	-0.078408
198	0.067197	0.042092	-0.063107
199	0.067219	0.037385	-0.061611
200	0.067392	0.030287	-0.045271
201	0.067300	0.041124	-0.048904
202	0.067280	0.044046	-0.043875
203	0.067249	0.053490	-0.048633
204	0.067310	0.049027	-0.043120
205	0.067258	0.060414	-0.031198
206	0.067294	0.053370	-0.036013
207	0.067277	0.064755	-0.022850
208	0.067187	0.068776	-0.030797
209	0.067225	0.071513	-0.010795
210	0.067141	0.079661	-0.006618
211	0.067220	0.033142	-0.061029
212	0.067181	0.049990	-0.056404
213	0.067140	0.046966	-0.064043
214	0.067273	0.040734	-0.053709
215	0.067253	0.040215	-0.054702
216	0.067250	0.050662	-0.052230
217	0.067236	0.058681	-0.041933
218	0.067211	0.056501	-0.045137
219	0.067253	0.055681	-0.041346
220	0.067199	0.061781	-0.040649
221	0.067246	0.063228	-0.030589
222	0.067254	0.064985	-0.025859
223	0.067160	0.070950	-0.034437
224	0.067193	0.076233	-0.021062
225	0.067135	0.079169	-0.022522

Table F.2 Principle Component Scores Values

Sample Number	PC1	PC2	PC3
1	-15.6580	0.3800	-0.0647
2	33.0483	2.6897	-0.1614
3	29.7845	-0.5631	-2.7583
4	-7.0900	-0.6900	1.1040
5	-0.1875	-0.8182	2.0418
6	6.4547	-1.3448	2.2729
7	17.1456	-0.5557	2.0229
8	21.8164	0.0338	2.3351
9	25.9243	-0.0456	2.1033
10	-7.9344	-0.6161	0.2677
11	-2.0793	-1.1659	0.0577
12	3.9165	-1.8325	0.4722
13	13.9472	-2.3664	-0.5880
14	-17.1077	1.0664	-0.2137
15	-12.7646	0.9812	0.0552
16	-5.4258	0.2960	0.6294
17	1.5948	0.2319	1.0615
18	7.7945	0.4785	1.1652
19	12.8520	0.4496	0.8157
20	-12.8569	0.5971	-0.2814
21	-5.7361	-0.2815	-0.1514
22	-1.0105	-0.4945	-0.6616
23	4.2631	-0.5427	-0.7311
24	1.8636	-1.5525	0.4731
25	7.7812	-1.6477	-0.4303
26	-16.3585	0.6406	-0.2969
27	-15.1998	0.5409	-0.1841
28	-11.9789	-0.0298	-0.3791
29	-11.6155	1.0303	0.0956
30	-14.2217	0.7725	-0.3205
31	-13.6573	0.9199	-0.2181
32	19.8982	2.7510	-0.2688
33	-13.0116	0.9841	-0.4234
34	16.7091	2.3425	-0.3512
35	-6.0985	-5.0983	-1.5911
36	23.7002	0.2204	-5.0820
37	-10.4172	0.0367	-0.4953
38	-13.4326	0.4024	-0.5134
39	-16.8886	0.7261	-0.5542
40	-17.7633	1.0740	-0.2535

REFERENCES

Abdi, H. (2003) Encyclopedia of Social Science Research method, Sage Publication, 1528p.

Ahmad, S.R., and Reynolds, D.M. (1995) Synchronous fluorescence spectroscopy of wastewater and some potential constituents. Water Research, 29(6), pp. 599-1602.

Amy, G., Siddiqui, M., Ozekin, K., Zhu, H.W., and Wang, C. (1998) Empirically based models for predicting chlorination and ozonation by-products: haloacetic acids, chloral hydrate, and bromate. USEPA Report CX 819579.

Anderson, D.R., Sweeny, D.J., and Williams, T.A. (2002) Statistics for business and Economics (8^{th} ed.) South-Western publishing, 884p.

APHA, AWWA, WEF (1995) Standard Methods for the Examination of Water and Wastewater, (19^{th}, ed.), Washington DC, 1995p.

Ates, N., Kitis, M., and Yetis, U. (2007) Formation of chlorination by-products in waters with low SUVA—correlations with SUVA and differential UV spectroscopy. Water Research, 41(18), pp. 4139-4148.

Baker, A. (2001) Fluorescence excitation-emission matrix characterization of some sewage-impacted rivers. Environmental Science and Technology, 35(5), 948-953.

Baker, A. (2002a) Fluorescence properties of some farm wastes: implications for water quality monitoring. Water Research, 36(1), pp. 189-195.

Baker, A. (2002b) Fluorescence excitation-emission matrix characterization of river waters impacted by a tissue mill effluent. Environmental Science and Technology, 36(7), 1377-1382.

Baker, A., Inverarity, R, Charlton, M., and Richmond, S. (2003) Detecting river pollution using fluorescence spectrophotometry: case studies from the Ouseburn, NE England. Environmental Pollution, 124(1), pp. 57-70.

Baker, A., and Inverarity, R. (2004) Protein-like fluorescence intensity as a possible tool for determining river water quality. Hydrological Processes, 18(15), pp. 2927-2945.

Bee. K.R., Pell, R.J., and Seasholtz, M.B. (1998) Chemometrics: A Practical Guide, John Wiley and Sons Inc., 348p.

Bengraïne, K. and Marhaba, T.F. (2003a) Comparison of spectral fluorescent signatures-based models to characterize DOM in treated water samples. Journal of Hazardous Materials, 100(1), pp. 117-130.

Bengraïne, K., and Marhaba, T.F. (2003b) Using principal component analysis to monitor spatial and temporal changes in water quality. Journal of Hazardous Materials, 100(1), pp. 179-195.

Bengraïne, K., Marhaba, T.F., Pu, Y., and Aragó, J. (2003) Spectral fluorescence and partial least squares regression: model to predict dissolved organic carbon in water, Journal of Hazardous Materials, 97(1), pp. 83-97.

Bengraïne, K., and Marhaba, T.F. (2004) Predicting organic loading in natural water using spectral fluorescent signatures. Journal of Hazardous Materials, 1089(3), pp. 207-211.

Brereton, R.G. (2003) Chemometrics: Data Analysis for the Laboratory and Chemical Plant, John Wiley and Sons, 489p.

Bourne, D.W.A. (2008) Boomer Manual. Retrieved on September 2008 from the World Wide Web: http://www.boomer.org/c/p3/c03/c0303.html

Bull, R.J. (1982) Health effects of drinking water disinfectants and disinfectant by-products. Environmental Science and Teachnology, 16(10), 554a–559a

Chen, W., Westerhoff, P., Leenheer, J.A., and Booksh, K. (2003) Fluorescence excitation-emission matrix regional integration to quantify spectra for dissolved organic matter. Environmental Science and Technology, 37(24), pp. 5701-5710.

Chow, A.T., Guo, F., Gao, S., Breuer, R., and Dahlgren, R.A. (2005) Filter pore size selection for characterizing dissolved organic carbon and Trihalomethane precursors from soils. Water Research, 39(7), pp. 1255-1264.

Chow, A.T. (2006) Disinfection byproduct reactivity of aquatic humic substances derived from soils. Water Research, 40(7), pp. 1426-1430.

Chow, A.T., Guo, F., Gao, S., and Breuer, R.S. (2006) Trihalomethane reactivity of water- and sodium hydroxide-extractable organic carbon fractions from peat soil. Journal of Environmental Quality, 35(1), pp. 114-121.

Chow, A.T., Dahlgren, R.A., and Harrison, J.A. (2007) Watershed sources of disinfection byproduct precursors in the Sacramento and San Joaquin rivers, California Environmental Science and Technology, 41(22), pp. 7645-7652.

Coble, P.G. (1996) Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy. Marine Chemistry, 51(4), pp. 325-346.

Croué, J.P., Debroux, J.-F., Amy, G. L., Aiken, G. R., and Leenheer, J. A. (1999) Natural organic matter: structural characteristics and reactive properties in formation and control of disinfection by-products in drinking water, American Water Works Association, 91, pp. 65–93.

Croué J.-P., Korshin, G. V., and Benjamin, M. (2000) Characterization of natural organic matter in drinking water, American Water Works Association Research Foundation.

Delaware & Raritan Canal State Park (2008) Park maps. Retrieved November 2008 from the World Wide Web: http://www.dandrcanal.com/maps.html

Edzwald, J.K., Becker, W.C., and Wattier, K.L. (1985) Surrogate parameters for monitoring organic matter and THM precursors. Journal of American Water Works Association, 77(4), pp. 122-131.

Esbensen, K.H. (2004) Multivariate Data Analysis in Practice (5th ed.), CAMO Process AS, 598p.

Field, A. (2005) Discovering Statistics Using SPSS (2nd ed.), SAGE Publishing, 779p.

Gang, D., Clevenger, T.E., and Banerji, S.K. (2003) Relationship of chlorine decay and THMs formation to NOM size. Journal of Hazard Material. Hazard, 96(1), pp. 1-12

Geladi, P., and Kowalski, B.R. (1986) Partial least-square regression: A tutorial. Analytica Chimica Acta, 185, pp. 1-17.

Goslan, E.H., Voros, S., Banks, J., Wilson, D., Hillis, P., Campbell, A.T., and Parsons, S.A. (2004) A Model for predicting dissolved organic carbon distribution in a reservoir water using fluorescence spectroscopy. Water Research, 38(3), pp. 783-791.

Johnson, R.A., and Wichern, D.W. (2007) Applied Multivariate Statistical Analysis (6th ed.), Prentice Hall, 773p.

Kanokkantapong, V.. Marhaba, T.F.. Panyapinyopol, B., and Pavasant, P. (2006) FTIR evaluation of functional groups involved in the formation of haloacetic acids during the chlorination of raw water. Journal of Hazardous Materials, 136(2), pp. 188-196.

Kebbekus, B., and Mitra, S. (1998) Environmental Chemical Analysis, ACRC Press. 330p.

Kim, H.C., Yu, M.J. (2005) Characterization of natural organic matter in conventional water treatment processes for selection of treatment processes focused on DBPs control. Water Research, 39(19), pp. 4779-4789.

Knocke, W.R., West, S., and Hoehn, R.C. (1986) Effect of low temperature on the removal of Trihalomethanes precursors by coagulation. Journal of American Water Work Association, 78, pp. 189-196.

Kutner, M.H., Nachtsheim, C.J., and Neter, J. (2004) Applied Linear Regression Models (4th edition), McGraw Hill Press, 701p.

Korshin, G.V., Li, C.W., and Benjamin, M.M. (1997) Monitoring the properties of natural organic through UV spectroscopy: A consistent theory, Water Research, 31(7), pp. 1787–1795.

Krasner, S.W., Weinberg, H.S., Richardson, S.D., Pastor, S.J., Chinn, R., Sclimenti, M.J., Onstad, G.D., Thruston Jr., A.D. (2006) Occurrence of a new generation of disinfection byproducts. Environmental Science and Technology, 40(23), pp. 7175-7185.

Larsson, T, Wedborg, M., and Turner, D. (2007) Correction of inner-filter effect in fluorescence excitation-emission matrix spectrometry using Raman scatter. Analytica Chimica Acta, 583, pp. 357-363.

Lee, S., and Ahn, K.H. (2004) Monitoring of COD as an organic indicator in waste water and treated effluent by fluorescence excitation-emission (FEEM) matrix characterization. Water Science and Technology, 50(8), pp. 57-63.

Leenheer, J.A., and Booksh, K. (2003) Fluorescence excitation-emission matrix regional integration to quantify spectra for dissolved organic matter. Environmental Science and Technology, 37, pp. 5701-5710.

Leenheer, J.A., and Huffman, E.W.D. (1976) Classification of organic solutes in water by using macro-reticular resins. Journal Research US Geological Survey, 4(6), pp. 737-751.

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), pp. 578-587.

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

Leenheer, J.A. (2004) Comprehensive assessment of precursors, digenesis, and reactivity to water treatment of dissolved and colloidal organic matter. Water Supply, 4(4), pp. 1-9.

Ma, H., Allen, H.E., and Yin, Y. (2001) Characterization of isolated fractions of dissolved Organic matter from natural waters and a wastewater effluent. Water Research, 35(4), pp. 985-996.

Marhaba, T.F., Lippincott, R.L., and Van, D. (2000) Characterizing dissolved organic matter fractions using spectral fluorescent signatures and post-processing by principal component analysis, Fresenius' Journal of Analytical Chemistry, 366(1), pp. 22- 25.

Marhaba, T.F. (2000) Fluorescence technique for rapid identification of DOM fractions. Journal of Environmental Engineering, 126(2), pp. 145-152.

Marhaba, T.F., and Lee Lippincott, R. (2000) Application of fluorescence technique for rapid identification of DOM fractions in source waters. Journal of Environmental Engineering, 126(11), pp. 1039-1044.

Marhaba, T.F., and Pu, Y. (2000) Rapid delineation of humic and non-humic organic matter fraction in water. Journal of Hazardous Materials, 73(3), pp. 221-234.

Marhaba, T.F., Pu, Y., and Bengraine, K. (2003) Modified dissolved organic matter fractionation technique for natural water. Journal of Hazardous Materials, 101(1), pp. 43-53.

Marhaba, T.F., and Van, D. (2000) The variation of mass and disinfection by-product formation potential of dissolved organic matter fractions along a conventional surface water treatment plant. Journal of Hazardous Materials, 74(3), pp. 133-147.

Milot, J., Rodriguez, M.J., and Serodes, J.B. (2000) Modeling the susceptibility of drinking water utilities to form high concentrations of Trihalomethanes. Journal of Environmental Management, 60(2), pp. 155–171.

Minitab 15 (2007) Minitab Stat Guide: On-Line Software Manual Minitab Inc.

Mobed, J.J., Hemmingsen, S.L., Autry, J.L., and Mcgown, L.B. (1996) Fluorescence characterization of IHSS humic substances: Total luminescence spectra with absorbance correction. Environmental Science and Technology, 30(10), pp. 3061-3065.

Moore, D.S., and McCabe, G.P. (2003) Introduction to the Practice of Statistic (4^{th} ed.) W.H. Freeman and Company. 827p

Munch, D.J., and Hautman, D.P. (1995) Method 551.1: Determination of chlorination disinfection by-products, chlorited solvent, and halogenated pesticides/ herbicides in drinking water by liquid-liquid extraction and gas chromatography with electron-capture detector. National exposure laboratory, Office of research and development, US EPA, Cincinnati, Ohio.

Musikavong, C., Wattanachira, S., Marhaba, T.F., and Pavasant, P. (2005) Reduction of organic matter and Trihalomethane formation potential in reclaimed water from treated industrial estate wastewater by coagulation. Journal of Hazardous Materials, 127(1), pp. 58-67.

Norwood, D.L., Christman, R.F., and Hatcher, P.G. (1987) Structural characterization of aquatic humic material: Phenolic content and its relationship to chlorination mechanism in an isolated aquatic fulvic acid. Environmental Science and Technology, 21(8), pp. 791-798.

Panyapinyopol, B., Marhaba, T.F., Kanokkantapong, V., and Pavasant, P. (2005a) Characterization of precursors to Trihalomethanes formation in Bangkok source water. Journal of Hazardous Materials, 120(1), pp. 229-236.

Panyapinyopol, B., Kanokkantapong, V., Marhaba, T.F., Wattanachira, S., and Pavasant, P. (2005b) Kinetics of Trihalomethane formation from organic contaminants in raw water from the Bangkhen water treatment plant. Journal of Environmental Science and Health - Part A Toxic/Hazardous Substances and Environmental Engineering, 40(8), pp. 1543-1555.

Pokorna, L., Gajdosova, D., Mikeska, S., Homolac, P., and Havel, J. (2001) The stability of humic acids in alkaline media. Humic Substances: Structures, Models and Functions, 12, pp. 133-152.

Pu, Yong (2005) Development and application of fractionation procedure for drinking water organic matter. Ph.D. dissertation, 138p.

Reynolds, D.M., and Ahmad, S.R. (1997) Rapid and direct determination of wastewater BOD values using a fluorescence technique. Water Research, 31(8), pp. 2012-2018.

Reynolds, D.M. (2002) The differentiation of biodegradable and non-biodegradable dissolved organic matter in wastewaters using fluorescence spectroscopy. Journal of Chemical Technology and Biotechnology, 77(8), pp. 965-972.

Richardson, S. D., Thruston, A.D., Rav-Acha, C., Groisman, L., Popilevsky, I., Juraev, O., Glezer, V., McKague, A.B., Plewa, M.J., and Wagner, E.D. (2003) Tribromopyrrole, brominated acids, and other disinfection byproducts produced by disinfection of drinking water rich in bromide. Environmental Science and Teachnology, 37, pp. 3782-3793.

Robinson, J.W. (1995) Undergraduate Instrumental Analysis (5^{th} ed.), Marcel Dekker, Inc. 858p.

Rodriguez, M.J., Sérodes, J.B., and Levallois, P. (2004) Behavior of Trihalomethanes and haloacetic acids in a drinking water distribution system. Water Research, 38(20), pp. 4367-4382.

Rook, J.J. (1977) Chlorination reaction of fulvic acids in natural waters. Environmental Science and Technology. 11(5), pp. 478-482.

Sadiq, R., and Rodriguez, M.J. (2004) Disinfection by-products (DBPs) in drinking water and predictive models for their occurrence: a review. Science of the Total Environment, 321, pp. 21-46.

Sarai, D.S. (2006) Water Treatment Made Simple: For Operators, Wiley Publication, 272p.

Singer, P.C. (1999) Formation and control of disinfection by-products in drinking water, American Water Work Association, 424p.

Singer, P.C. (2004) Disinfection by-products in US drinking waters: past, present and future. Water Science and Technology: Water Supply, 4(1), pp. 151-157.

Sharp, E.L., Parsons, S.A., and Jefferson, B. (2006) Seasonal variation in natural organic matter and its impact on coagulation in water treatment. Science of the Total Environment, 363: pp. 183-194.

Skoog, D.A., Holler, E.J., and Nieman, T.A. (1998) Principle of Instrumental Analysis, Thomson Learning, Inc.,849p.

Sohn, J., Amy, G., Cho, J., Lee, Y., Yoon, Y. (2004) Disinfectant decay and disinfection by-products formation model development: Chlorination and ozonation by-products. Water Research, 38(10), pp. 2461-2478.

Thermo Electron Corporation (2006) Algorithms - Principal Component Regression. Retrieved on April 2006 from the World Wide Web: http://www.thermo.com/com/cda/resources/resources_detail/1,2166,13414,00.html?fromPage=search&keyword=Principal+Component+Regression+%28PCR%29

Thurman, E. M. (1985) Organic Geochemistry of Natural Water, M. Nijhoff and W. Junk Publishers.

Toroz, I., and Uyak, V. (2005) Seasonal variations of Trihalomethanes in water distribution networks of Istanbul City. Desalination, 176, pp. 127-141.

Tranter, R.L. (2000) Design and Analysis in Chemical Research, CRC Press.

Uyak, V, Toroz, I., and Meriç, S. (2005) Monitoring and modeling of Trihalomethanes (THMs) for a water treatment plant in Istanbul. Desalination, 176, pp. 91-101.

USEPA, U.S. Environmental Protection Agency (1998a) National Primary Drinking Water Regulations: Interim Enhanced Surface Water Treatment Final Rule., 40 CFR Part 9, 141 and 142.

USEPA, U.S. Environmental Protection Agency (1998b) National Primary Drinking Water Regulations: Stage 1 Disinfectants and Disinfection By-products Rule., Federal Register: volume 63, number 241.

USEPA (1999) Enhanced Coagulation and Enhanced Precipitative Softening Guidance Manual. 237p.

USEPA, U.S. Environmental Protection Agency (2006a) National Primary Drinking Water Regulations: Stage 2 Disinfectants and Disinfection Byproducts Rule, Federal Register: volume 71, number 2.

USEPA, U.S. Environmental Protection Agency (2006b) Long Term 2 Enhanced Surface Water Treatment Rule, Federal Register: volume 71, number 24.

Weishaar, J.L., Aiken, G.R., Bergamaschi, B.A., Fram, M.S., Fujii, R., and 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), pp. 4702-4708.

Xie, Y.F. (2007) Disinfection Byproducts in Drinking Water: Formation, Analysis, and Control, Taylor & Francis Press. 176p.

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, pp. 15-36.