Development and Analysis of a Water Quality Monitoring Program for the Pockwock Lake Watershed

by

Colin M. Ragush

Submitted in partial fulfilment of the requirements
for the degree of Master of Applied Science

at

Dalhousie University
Halifax, Nova Scotia
April 2011

DALHOUSIE UNIVERSITY

DEPARTMENT OF ENVIRONMENTAL ENGINEERING

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled “Development and Analysis of a Water Quality Monitoring Program for the Pockwock Lake Watershed” by Colin M. Ragush in partial fulfillment of the requirements for the degree of Master of Applied Science.

Dated: April 13 2011

Supervisor: ______________________________

Readers: ______________________________

____

____

DALHOUSIE UNIVERSITY

DATE: April 13, 2011

AUTHOR: Colin M. Ragush

TITLE: Development and Analysis of a Water Quality Monitoring Program for the Pockwock Lake Watershed

DEPARTMENT OR SCHOOL: Department of Environmental Engineering

DEGREE: MASc CONVOCATION: October YEAR: 2011

____

Signature of Author

Table of Contents

List of Tables

List of Figures

Abstract

Municipal Source Water Protection Plans (SWPPs) are instituted in efforts to maintain and protect water quantity and quality. An integral part of a SWPP is the Source Water Monitoring Plan (SWMP). Without well defined metrics it is impossible to determine if a SWPP is effective and meeting its general goals of maintaining water quantity and quality.

A 16 month intensive monitoring plan was implemented to examine how a SWMP should be structured and how acquired data needs to be analyzed in order to answer specific water quality questions that may be posed.

This thesis demonstrates the temporal and spatial variability of water quality data and discusses the utilization of common water quality metrics. The importance of developing goals for SWMP is stressed as, due to the range of information that can be acquired from different sampling strategies, SWMPs need to be tailored to meet the goals of a monitoring program.

List of Abbreviations and Symbols Used

AIC	Akaike Information Criterion
ARMA	Autoregressive Moving Average
ARIMA	Autoregressive Integrated Moving Average
BOD	Biological Oxygen Demand
C	Carbon
CCME	Canadian Council of Ministers of the Environment
CFU	Colony forming unit
DBP	Disinfectant byproduct
DIP	Dissolved Inorganic Phosphorous
DO	Dissolved Oxygen
GEV	Generalized Extreme Value
ISO	The International Standards Organization
MAC	Maximum Allowable Concentration
N	Nitrogen
N_2	Nitrogen gas
NH_3	Ammonia
NO_2^-	Nitrite
NO_3^-	Nitrate
NSEL	Nova Scotia Environment and Labour
NSE	Nova Scotia Environment
NTU	Nephelometric Turbidity Unit
P	Phosphorous
PO_4^-	Orthophosphate
SWMP	Source Water Monitoring Plan
SWPP	Source Water Protection Plan
TC	Time of Concentration
TDS	Total Dissolved Solids
THM	Trihalomethane
TOC	Total Organic Carbon
TP	Total Phosphorous
TSS	Total Suspended Solids
UMVU	Uniformly minimum variance unbiased estimator
WHO	World Health Organization

Acknowledgements

I would like to recognize my thesis supervisor Dr. Rob Jamieson for his mentorship, as he always exceeded his duties. Dr. Jamieson’s positive demeanor, patients, and encouragement created the foundations for an environment where I thrived both academically and personally.

I would like to thank Dr. Graham Gagnon and Dr. Serguei Iakovlev for their willingness to be part of my master thesis committee. I greatly appreciate the time taken to help develop my research and synthesize my work into the final thesis product.

I would also like to extend a thank you to Rick Scott for his aid in sample analysis as well as his willingness to share his knowledge and expertise. The experience of working with Rick Scott was rewarding both personally and professionally.

I also wish to recognize Dr. Jamieson’s research group from 2009-2011 as their consistent support was an integral component of my positive graduate school experience. I would like to more specifically thank the members of Dr. Jamieson’s research group who accompanied me to the Lake Pockwock watershed as well as served as sounding boards for ideas during the data analysis phase of my thesis. These members are, in no special order: Leah Boutlier, Katie Campbell, Nick Dorado, Tristan Goulden and Andrew Sinclair.

It is imperative that I recognize the financial contributions of Halifax Water and the Natural Science and Engineering Research Council of Canada as it was their financial support that made my research project a reality.

Finally, I would like to extend a special thank you to Dalhousie Faculty of Graduate Studies, and the Bruce and Dorthy Rossetti Scholarship Fund for their additional and unexpected funding. These two groups made my graduate life a little more comfortable, and for that I am very grateful.

Chapter 1 Introduction

Fresh water is an essential resource required to sustain life, and support industries like agriculture and manufacturing. Anthropogenic development is increasing pressure on water resources globally, and as a result there is increased interest in the protection of finite fresh water resources. Most critical to humans are water sources utilized for potable purposes, and as a result protection of our municipal water supplies is an important environmental and political issue. Municipal Source Water Protection Plans (SWPPs) are instituted in efforts to maintain and protect water quantity and quality. An integral part of a SWPP is the Source Water Monitoring Plan (SWMP). Without well defined metrics it is impossible to determine if a SWPP is effective and meeting its general goals of maintaining water quantity and quality. All public water utilities in Nova Scotia are required to develop a SWPP which is submitted for review to Nova Scotia Environment and Labour (NSEL) (Province of Nova Scotia, 2009).

Politically, there is pressure on service utilities to increase source water monitoring in an effort to guarantee water quantity and quality, but there has been little scientific evidence that this will benefit water quantity or quality now or into the future. It has long been established, in many industries, that more data does not always produce better information. Loftis et al. (1986) were the first to recognize that this problem was plaguing water management. Loftis et al. suggest that new technology will not alleviate the struggle, and as the problem continues to persist 25 years later it would appear that they are correct. With the development of “user friendly” monitoring equipment and continuous monitoring equipment the problem appears to have grown.

However, the technology is not the source of the problem. In general, SWMPs have not been adequately constructed, and there has not been sufficient consideration of three technical decisions that determine the information that will be obtained: (i) what parameters are measure, (ii) sampling frequency, and (iii) sampling technology to employ.

Improved technology, when appropriately applied, along with an effectively structured SWMP, will create a more cost effective program while not compromising the quality of information about the health of the source water. The overall goal of this research is to assess how to structure a SWMP to be more effective at providing information about water quality. This research uses the Pockwock Lake watershed as a case study, and assesses both discrete and continuous monitoring structures. Pockwock Lake is used by Halifax Water as the source water supply for over 300,000 people.

1.1 Research Objectives

Halifax Water currently conducts a monthly SWMP at 9 points within the Pockwock Lake Watershed. In this research project a 16 month continuous monitoring program was enacted on two of the primary tributaries that feed Pockwock Lake. The data set acquired will provide the ability to assess temporal and spatial variability of water quality within the watershed, as well as the quality of information generated from continuous data. Insight into how, and what, water quality parameters should be measured will also be gained. This study will aid in the future development of Halifax Water's and Pockwock Lake's SWPP and SWMP.

The three main objections of this research are:

1. Assess spatial and temporal variability in water quality parameters in the Pockwock Lake Watershed
2. Assess how high frequency water quality data needs to be processed and statistically analyzed.
3. Create recommendations for sampling method and frequency, and statistical processing for the Pockwock Lake SWMP

Chapter 2 Literature Review

2.1 Principles Of Source Water Management

Provincial governments, including Nova Scotia, are now requiring that some source water management techniques, such as protection plans, be constructed for all source watersheds in an effort to keep source waters sustainable. Municipal water utilities are responsible for managing all aspects of their potable water system, as stated in regulations put forth in a comprehensive drinking water strategy by Nova Scotia Environment (NSE) on October 17, 2002. The strategy aims to create a multi-barrier approach to achieve the goals of: keeping clean water clean, making water safe to drink, and proving water is safe to drink. A source water protection plan is focused on satisfying the first barrier of the strategy, which is keeping clean water clean.

The treatment of source water for human consumption is expensive, and the cost increases with decreasing water quality. The treatment plant for Pockwock Lake used over $2 M in chemicals and $ 1.5 M in energy in 2009 (Reid Campbell, Personal Communication, November 2010). From both the service provider's perspective and the consumers' it is beneficial to maintain the source water at the highest standard possible to be able to provide safe water at the lowest cost.

Historically source water sampling regimes have been focused on quantifiable biological and chemical indicators and solutions have been mechanically oriented (Pahl-Wostl, 2006). The goal has been to keep chemical constituents below guideline levels generally through technical solutions. This has been an adequate strategy for some issues but service providers are facing problems that do not currently have cost effective treatment solutions. Due to this current impasse, it is necessary to examine alternative solutions, such as source water management.

Source Water Protection Planning is an iterative process which requires a committee that is representative of the stakeholders in the watershed. The panel should include: government, industry, and community citizens (Technical Expert Committee, 2004). The Committee must: define the source protection area, identify potential risks, define goals, targets and objectives, map the area, rank threats and identify vulnerable areas, create and implement a source water protection plan, and establish regular monitoring and reporting practices (Pollution Probe, 2004).

The source water protection area is generally the watershed area. The watershed area includes all the land and water which directs precipitation, by means of groundwater flow or surface flow, to the point of use. This place of use may be a well, a source water treatment plant, or a river or lake (Pollution Probe, 2004). It is crucial that groundwater not be ignored, and that the assessment of zones of groundwater discharge and recharge are determined. The zones of discharge and recharge are especially critical in systems where wells are used to access the water resource, as they are generally more vulnerable to contamination than other areas within the watershed (Technical Expert Committee, 2004). Flow direction of both surface waters and groundwater must be documented as an understanding of the hydrological processes is essential when assessing the risks which could result in source water degradation (Technical Expert Committee, 2004)

Water quality can be influenced by various natural and anthropogenic factors before it arrives at the site of use, whether it is a farmer’s field, a geological deposit, an outflow from industry, or a residential septic system. All potential risks need to be documented which includes both point sources and non-point sources. A point source is a source such as an outflow from industry which is easily identified and analyzed for its effect on water quality, while non-point sources often are more difficult to quantify such as fertilization on an agriculture area (Chin, 2006). At this point of SWPP development all possible sources should be brought forward by all proponents for consideration. Although a potential source is considered, the source is not inherently of concern until further review is completed as it may not have any negative effects on water quality (Technical Expert Committee, 2004).

In conjunction with the identification of threats, the goals, objectives, and targets of the SWPP must be defined. It is a necessity that current goals are created, but also that future goals are also developed when designing the objectives and targets. A plan that fails to incorporate projections of changes in the land use or user base, or one that does not account for the need for water resources in the future, is a plan that will create management issues in the future (Pollution Probe, 2004).

Having identified the goals, objectives, targets and potential threats the committee now would assess and rank the risks (Nova Scotia Environment Water and Wastewater Branch, 2004). The assessment of risk is imperative to prioritize where to focus resources to most efficiently improve the safety of water resources. Risks need to be assessed based on the probability that the threat will impair the use of water resources and the severity of impairment if an event does occur. The process of risk assessment is one of judgment which requires input from experts such as scientists, farmers, engineers, and geologists. It is methodical and uncertainty should result in a bias on the side of safety.(Technical Expert Committee, 2004).

After prioritizing risks, the SWPP process involves implementing strategies to manage the risk. Strategies encompass three major groups of management techniques: (i) preventative (measures such as restricting access to waterways), (ii) mitigative (measures such as riparian buffer zone construction), or (iii) rehabilitative (measures such as stabilization of impacted stream banks or re-naturalization of stream channels). It should be recognized that most management techniques are policy based prevention and that a great deal of the effort by the committee will be focused on assuring that the policies are upheld (Pollution Probe, 2004). The public, government, and industry need to be informed of the implemented management strategies, especially ones involving policy, as the success is dependent upon accountability of the stakeholders for their watershed.

Having enacted management strategies, monitoring and reporting becomes imperative. Metrics must be instituted to measure the effectiveness of management strategies, and the results need to be reported to the entire community which has a stake in the water resource. The metrics must be chosen so the goals and objectives can be assessed to determine if they are being met. A SWPP is an iterative learning process and is never static, as over time the source water will change and the needs, objectives, and goals will also change. The continual scrutiny and revision of objectives and goals as well as management techniques will allow for adaptation to changes in the watershed (Pahl-Wostl, 2006).

2.2 Impacts Of Forestry And Rural Development On Water Quality

Forestry activities and rural development modify the local land cover which can drastically modify hydrology. Forestry and rural development create impervious surfaces, compacts soil, and removes vegetation from the land (Mallin, 2009). These actions decrease the infiltration rate through the soil's profile resulting in an increase in Horton overland flow (overland flow that occurs when the rainfall rate exceeds the infiltration capacity of the soil (Chin, 2006)), and a decrease in the natural evapotranspiration process that occurs in the vegetated areas (Mallin, 2009). Due to this increase in overland flow, storm waters reach the stream much more quickly and elevated peak flows and shorter duration storm flows occur (Booth, 2002). In addition to flow regime modification, the chemical and biological quality of the water can be significantly impacted because of the increased transport of soils and materials.

The storm runoff, being of greater quantity, contains more energy and is more able to carry particulates, and chemical and biological contaminants into receiving waters (Brezonik & Stadelmann, 2002). Storm flows contribute the majority of contaminants and suspended solids to water bodies, and since the removal of vegetation increases storm flow, the removal of vegetation can create a pollution problem (Mallin, 2009). Deforestation and farming on graded slopes also contributes to accelerated erosion (Lal, 1998). This is because the soil matrix is now particularly prone to erosion during severe storms because there is no longer material to cover and stabilize. This fine sediment is carried to the waterways as suspended solids.

Nutrient cycles within watershed ecosystems are also impacted by deforestation and rural development. Agricultural development results in the creation of both point and non-point nutrient sources within receiving water systems because of runoff from fertilized fields and the excrement of livestock (Chin, 2006). The addition of phosphorous (P) can lead to accelerated eutrophication in fresh water systems (Schoor, 1996). Organic carbon matter and nitrogen (N) export are also typically increased. Ahtianen (1992) found that there were significant increases in N, P and total suspended solid (TSS) exports in watersheds that had been clear-cut in Finland. Upon the completion of the three year monitoring plan for the watersheds, levels of nutrients and suspended solids still had not returned to baseline conditions. Ahtiainen (1992) suggests that his findings greatly depend upon soil type, as the soil's characteristics could greatly influence the availability of nutrients and severity of the impact that can be expected from deforestation. Macdonald et al. (2003) in British Columbia found an increase in freshet and storm water flows after harvesting occurred in small stream systems. These changes continued to persist for five years after the harvest. Although TSS concentrations initially increased, three years post harvesting the concentrations were back to or below baseline conditions. The harvesting practices included a 20 m riparian buffer zone, and it was determined that during some of the storms where elevated TSS levels were discovered the source was from stream crossing areas or log landing sites that had compromised the buffer. A study in Scotland by Neal et al. (2004) suggested that forestry management practices can have a strong impact on water quality and nutrient export to streams. They found that phased felling, instead of the conventional clear cutting, decreased nutrient export and the peak flow of stream hydrographs.

Though it is well documented that forestry activities will impact water quality, it has been expressed by Feller (2005) that there is a lack of knowledge and understanding of how forestry practices affect water quality. Study results have been highly variable because of differences in physical, biological and chemical characteristics of the study watersheds, as well as differences in forestry management practices. Compounding the problem, large temporal and spatial variations within watersheds themselves are often recorded making it challenging to definitively determine the effects of forestry activities. Feller identified eight knowledge gaps that need to be removed if there is to be confidence in predicting the impacts of forestry activities ion water quality. These eight areas are: (i) detailed chemical budgets of both impacted and unimpacted streams, (ii) understanding of longitudinal variation in stream chemistry, (iii) understanding of forest harvesting impacts on processes within streams, (iv) quantification of the behavior of trace metals in forested streams and the response of these trace metals to forest harvesting, (v) quantification of the influence forest harvesting has on geological weathering chemical release, (vi) quantification of forest harvesting’s influence on the snowmelt flux of chemical export to streams, (vii) quantification of the influence forest harvesting induced soil changes and stream temperatures has on stream water chemistry, and (viii) qualitative assessment of the relative importance of factors controlling stream water chemistry, particularly post forest harvesting.

2.3 Watershed Scale Water Quality Monitoring

2.3.1 Key Parameters

2.3.1.1 Microbiological Parameters

Microbiological pathogens that may be present in source water include bacteria, protozoans, and viruses. Despite the affluence of Europe and North America there continues to be cases of mass infection into the 21st century. The best known water contamination case in Canadian history was a bacterial contamination of drinking water in May 2000 in Walkerton, Ontario when a pathogenic strain of *Escherichia coli*. (0157:H7) resulted in 2,300 cases of gastroenteritis and 7 deaths in the town of 4,800 residents (Hrudey et al., 2003). Protozoan and viral contamination events have also been document with a Crytosporidia, an infectious protozoan, outbreak in North Battleford, Saskatchewan in 2001 resulting in over 1,900 cases of sickness (Stirling et al., 2001), and a source water contamination of virus aetiology resulting in over 1,500 cases of gastroenteritis in 1992 in Uggelose, Denmark (Laursen et al., 1994).

Coliforms are a group of bacteria widely used as indicator organisms of fecal contamination of a water supply. The coliform group of bacteria consists of organisms which reside in the intestinal system of both cold and warm-blooded animals. Coliforms are of interest because 93-98% of coliforms are fecal coliforms. Fecal coliforms are coliforms which are known to be present in the waste of warm blooded animals, and their presence suggests water contamination with animal excrement. Although most coliforms are not pathogenic, their detection serves as indication of the possible presence of more serious fecal pathogens such as *Streptococci*, *Vibrio cholera*, or *Campylobacter*. The species of coliform most often used as an indicator of fecal contamination is Escherichia coli. It is used because of its known presence in fecal material and well established culturing practices. (Droste, 1997)

Microbial pathogens represent a broad array of organisms, and therefore it is impractical to monitor source water for all of them. The current accepted practice is to monitor for the presence and concentrations of coliforms, specifically *E. coli*. Using coliforms as an indicator has been considered a conservative measurement as they tend to be able to survive the extra-intestinal environment better than other pathogens, and generally are in higher numbers (Edberg et al., 2000). The assumption is that if coliforms were removed to an acceptable level, so were all other pathogens. There is continued debate about the appropriateness of coliforms as an indicator, but coliforms continue to be the standard in water quality testing due to the cost effectiveness (Edberg et al., 2000).

There is currently no ability to continuously monitor for microbial contaminants. The primary methods of analysis are culture based, and culturing of microbes has an incubation period generally upwards of 24 hours, which results in a delay in information.

2.3.1.2 Temperature

Temperature is a key water quality parameter as it influences reaction rates, fluid dynamics, saturation levels of dissolved gases, organisms' metabolic rates and other biological, chemical, and physical processes (Tchobanoglous & Schroeder, 1985). Temperature also serves as a cue for natural events such as animal emergence and spawning. Alterations of stream temperature regimes can have an impact on the ecology of an aquatic system (Hauer & Lamberti, 2006). Temperature has an annual and diurnal cycle that mimics that of the air but is less extreme due to water's high heat capacitance. Fluctuations in water temperature can be a diagnostic tool in determining causation of water quality changes and, as result, solutions to water quality issues. It is generally not meaningful to take discrete measurements of temperature as temperature can vary as much as 30 ^{o}C annually and over 5^{o}C daily in the case of small poorly sheltered streams in temperate climates (Caissie, 2006). Temperature is one of the easiest and most cost effective parameters to monitor continuously with a high level of accuracy.

2.3.1.3 Conductivity

Conductivity is the measurement of how well electrical current travels through a solution. In a solution the current is able to traverse through dissolved ions (Sawyer et al., 2003). Therefore, conductivity is related to the total dissolved solids (TDS). TDS is composed of salts and other constituents that will not evaporate. Water becomes continually more laden with dissolved salts as it travels downstream through the stream system, and if the concentration of the solids reaches too high of a level the utilization of the water may be impaired. Water which has a high conductivity and thus a high TDS content can pose a problem to crop production, be more expensive to treat, and can cause environmental damage (Davis & Cornwell, 2008). TDS sources can be natural such a geological formations, or anthropogenic such as road salting (Allan, 1995).

Conductivity is much easier to measure on a continuous basis than TDS, and as a result is the parameter of choice. Conductivity provides insight into the local geochemistry and allows for monitoring of anthropogenic impacts in the watershed. Welch et al. (1977) recognized that forest clear-cutting and intensive potato farming resulted in increased conductivity in New Brunswick streams. Conductivity is a stable parameter and variations in it suggest perturbation of the local environment.

2.3.1.4 pH

pH is a measure of the concentration of hydrogen ions. Water with elevated levels of hydrogen ions will be acidic with a pH below 7, and water with elevated levels of hydroxyl ions will exhibit the property of being basic (pH > 7). These two ions are naturally formed by the disassociation of a water molecule. Local geology strongly influences the pH of water systems through geochemical reactions, as water travels through and over the soil. As the water incorporates minerals the pH will be modified. Unimpacted streams, river, and lakes have a range of pH of 4 to 10. Waters in Nova Scotia naturally are more acidic due to the local geology consisting mostly of non-weathering granite and metamorphic bedrock, which lacks buffering capacity. This lack of buffering capacity makes local waters susceptible to acidification (Ginn et al., 2007).

Acidification of streams and lakes can have severe impacts on the flora and fauna, as well as the chemistry of the system. Depressed pH can allow for the increased solubility of toxic metals and can have a negative impact on the utilization of the source water (Siegel, 2002). The ecology can also be compromised when species or life stages of species are under acid stress. The result of acid stress can ripple through the food web and have a profound effect on the ecological quality of a watershed (Okland & Okland, 1986). However, the pH of the water (provided it is in an acceptable range) may not be the main concern; the stability of the pH may be of more importance. Most aquatic plants and animals can tolerate a window of 3-4 pH units but are very sensitive to abrupt changes, and episodic acidification results in increased mortality (Baker et al., 1996).

pH is usually included in water quality monitoring programs because it is easy to measure, inexpensive , and can be monitored continuously. It can be an indicator of local disturbances in the watershed, seasonal variations or long term watershed changes.

2.3.1.5 Total Suspended Solids

Natural aquatic systems have particles suspended in the water column. These are often referred to as suspended solids, and can be composed of many different types of particles. These particles can be organic particles, both living and dead, such as algae and detritus, respectively. Or they can be inorganic silt, clay or even sand particles. In quiescent waters all but the living organic particles will settle out over time, but in flowing waters systems some particulates throughout the water column remain suspended. In general, increased river flow results in larger concentrations of suspended solids due to both the associated influx of material into the stream from storm events, and increased energy in the flow resulting in suspension of larger particles (Tchobanoglous & Schroeder, 1985). As a result, attention must be focused on storm events when analyzing the movement of sediment and suspended solids through a water system (Robertson & Roerish, 1999).

Suspended solids have a variety of impacts on water quality and ecosystem health. Excessive suspended sediment can destroy benthic habitats by changing the natural substrate, negatively impact the hunting ability of visual predators, and act as a transport mechanism for sorbed chemicals and nutrients through the ecosystem. The accumulation of sediment also can greatly change a habitat as the sediment will be removed from areas of the river experiencing stronger currents and deposited in areas of weak currents. Many pesticides, fertilizers, and nutrients are easily sorbed to sediment, and the transport of sediment is a major pathway for how these substances travel within a watershed. (Schoor, 1996)

The movement of sediments is a natural process, but when the rates are modified through human interaction with the environment, drastic changes in the environment and water chemistry may occur.

2.3.1.6 Turbidity

Turbidity is the measurement of the scattering of light caused by particles in a liquid (Tchobanoglous & Schroeder, 1985). In a natural environment these particles, or suspended solids, could be fine sediment or organics. Turbidity is measured in Nephleometric Turbidity Units (NTU) and is a relative measure. Turbidity, being a relative measurement, must be calibrated to water clarity or total suspended solids to have environmental meaning (Davies-

Colley & Smith, 2001). When assessing drinking water we want to assess both the optical quality and mass of sediments; the challenge is that the relationship between suspended solids concentration and turbidity is highly variable depending on the particles. The variability is described by Raleigh's law as both the size and the concentration of particles influence the measurement of turbidity (Droste, 1997).

Equation 2.1 Relationship between turbidity and particle characteristics.

$$Is \propto \left(\frac{V^2}{\lambda^2}\right)n$$

V = volume of particles
n = number of particles
λ = wavelength of light
I_s = intensity of scattered light

When turbidity is measured it is important to note the apparatus it was measured with, as turbidity is not an absolute measurement, and turbidity readings between equipment can vary substantially due to the design and construction of the apparatus (Davies-Colley & Smith, 2001). Turbidity is still a valuable parameter to measure because it is inexpensive, easy to measure, and there is the ability to measure it continuously. Through knowledge of site specific particle characteristics it may be possible to relate turbidity to TSS, a parameter of generally greater interest for potable water.

2.3.1.7 Nutrients

The three main nutrients which are measured in environmental systems are C, N, and P. In the environment nutrients, are present in different organic and inorganic forms. The balance and cycle of nutrients is integral to the health of an ecosystem and any disruption can result in ecosystem degradation (Allan, 1995).

Organic carbon is of particular concern in water treatment systems where residual disinfection with chlorine compounds is utilized. If organic C is present when chlorine is added to drinking water reactions can occur that result in the formation of disinfectant byproducts (DBPs). Some DBPs such as Trihalomethanes (THMs) are monitored because of their carcinogenic nature

(Droste, 1997). Total organic carbon concentrations vary greatly between watersheds as the parameter is greatly dependent upon land use and forest type. A group study conducted by Environment Canada and the Service Hydrolique National of Switzerland found that the range of average TOC concentrations for 100 streams was 2 mg/l – 30 mg/l with a mean of 10 mg/l (Meybeck, 1982).

N is present in the environment in many forms but there are five forms that are of importance to water quality; they are: organic nitrogen, ammonia (NH_3), nitrite (NO_2^-), nitrate (NO_3^-) and nitrogen gas (N_2). N_2 is fixated by bacteria and algae, which work in symbiosis with plants that do not have the capacity to fixate N themselves, into organic matter. This organic matter is transferred through the food chain (Tchobanoglous & Schroeder, 1985). Upon decay and decomposition of plant and animal waste bacterial decomposers break down organic matter containing N producing, what is a waste product to them, NH_3. Plants and algae preferentially uptake NH_3 over other N forms as it is more energy efficient (Schoor, 1996). Nitrification takes place in the natural environment by *nitrosomase* and *nitrobacter*, two genuses of bacteria, which respectively transform NH_3 into NO_2^- and from there into NO_3^- (Zhang et al., 2009). NO_3^- then undergoes denitrification which returns the nitrogen back to the atmosphere as N_2.

Nitrate$^-$ is generally the most prevalent species in aquatic environments because of its high mobility in soil, solubility, and stability (Health Canada, 2009). These characteristics make it the typical for of N included in SWMP.

Nitrogen is a nutrient of concern because of its potential to be a limiting nutrient in plant growth (Schoor, 1996). There are also documented health concerns related with elevated ingestion of NO_3^- and NO_2^- such as Methaemoglobinaemia (common cause of blue baby syndrome), carcinogenicity, congenital birth defects, and behavioral effects such as sluggishness (Health Canada, 2009).

Phosphorous is cycled in the environment between inorganic and organic forms. The phosphorous cycle is similar to the nitrogen cycle; however, the atmosphere fixation does not play a significant role. Orthophosphate (PO_4^-) is the only form of phosphorous that plants and microorganisms are readily capable of incorporating. Since P is often relatively scarce compared to other nutrients in terrestrial environments, soluble forms are quickly assimilated. Insoluble inorganic forms, such as calcium phosphate [$Ca_2(HPO_4)_2$], are converted into soluble forms (primarily PO_4^-) by microorganisms. Organic P from decaying plant and animal wastes are converted by bacteria into PO_4^-, which are then generally quickly assimilated by plants and microbes (Tchobanoglous & Schroeder, 1985).

Chemical processes also influence Dissolved Inorganic phosphorous (DIP) concentrations in the stream. DIP, which is often charged, is attracted and adsorbed to charged clays and organic particles when DIP levels are high. This effectively is a buffering system for DIP as the reaction favors desorption in the case of low DIP concentrations. Dissolved P under aerobic conditions may complex with metal oxides or hydroxides forming insoluble precipitates, but this P can be liberated during anaerobic conditions (Allan, 1995).

An excess of P, generally the limiting nutrient in freshwater systems results in accelerated eutrophication which is characterized by excessive plant growth, anoxia, and loss of species diversity (Schoor, 1996). Total phosphorous (TP) is typically analyzed in water samples because phosphorous is strongly bound to the sediment or particulates in the water. The result is that a significant portion of the P is transported in the movement of particulates and sediment (Brunet & Astin, 1998). Cooke and Prepas (1998) monitored two forested watersheds in the boreal plains of Alberta in 1994 and 1995 and calculated the loading intensity of TP to be an average of 11 and 14 $kg*km^{-2}*year^{-1}$ for that particular watershed. There currently is not technology available to economically measure phosphorous continuously in the environment.

2.3.1.8 Dissolved Oxygen

Dissolved oxygen (DO) is an important chemical parameter in aquatic environments because it an indicator of ecosystem health. DO content of aquatic systems is regulated by five processes:

1. The release of oxygen into the water column during the day as a result of photosynthesis by both phytoplankton and benthic plants.
2. The uptake of oxygen from the water during respiration by organisms, and plants.
3. The consumption of oxygen by chemical oxidation processes to break down organic detritus.
4. The exchange of oxygen with the air; which is directly dependent upon the saturation gradient that exists between the air and the water.
5. Influx of oxygen with accrual of ground water into surface water. (Odum, 1956)

All living animals require oxygen to survive, including aerobic microorganisms who utilize oxygen during decomposition of organic matter (Schoor, 1996). This Biological Oxygen Demand (BOD) is a result of the requirement for oxygen to oxidatively break down this organic matter, and can cause a depression of the available DO in the environment. Natural detritus in the stream will always be present but the action of adding additional BOD to the environment through disposal of organic wastes can form an oxygen poor environment that is unsuitable to many organisms. Dissolved oxygen concentrations can also be depressed by an increase in water temperature, as a result of gas solubility's inverse relationship to temperature. As a result, temperature pollution or environmental modifications such as forestry can have a negative impact on ecosystems by decreasing DO levels (Binkley & Brown, 1993). Dissolved oxygen follows a diurnal cycle with the lowest concentrations typically occurring at night and the largest during the day. The diurnal cycle is attributed to photosynthesis of the plants resulting in oxygen production in the streams during the day, when there is sunlight, leading to a net production of oxygen. However, during the night, respiration results in a net deficit (Odum, 1956). Dissolved oxygen is one of the main parameters used to measure ecological productivity. Through the use of two stations on the stream an estimate of the productivity of the stream can be obtained (Odum, 1956). Dissolved oxygen can be easily measured through a Clark cell (Pettingill, 1933) and there is the ability to measure continuously.

2.3.2 Sampling Strategies

2.3.2.1 Discrete Methods

A Discrete sampling strategy involves collection of sampling intermittently over a sampling period. The samples can be taken manually, or a machine can be automated to do so. A sample only represents the value of the parameter in the environment at the particular point of sampling shortly before, during, and after the sample was taken. As the temporal or spatial distances increases between time, site of interest, and the measurement, there will be less correlation between the sample and environmental conditions. Little can be inferred about the parameter between two sampling occasions or locations, but approximations can be suggested. There are three main types of discrete samples: grab, integrated and composite. A grab sample is taken at a specific spatial point at a given time. If the parameter varies spatially over a water body the sample is not representative of the whole water body as it is only a representation at the specific location at which it was collected. Integrated samples are used to obtain a spatial average in a water body, and are especially useful in larger rivers. Composite samples are used to obtain a temporal average over a specific time period such as a day or week. (United States Department of Agriculture, 1996)

2.3.2.2 Continuous Methods

Continuous methods are applicable when electrometric means of measuring a parameter are available. This means that parameters like pH, temperature, conductivity, and dissolved oxygen can be adequately measured continuously (United States Department of Agriculture, 1996). The term “continuous” is a slight misnomer as samples are semi-continuously collected. It is common that electrometric readings are collected on fifteen minute to one hour intervals and it is inferred that any appreciable change between readings is linear or some other common function.

The array of available equipment in the area of continuous monitoring is growing and becoming more reliable and affordable, but there is still a relatively high capital cost. Although the reliability of equipment is improving, it is recognized that it is still difficult to obtain a reliable and representative continuous measurement series for some systems like rivers (Vandenberghe et al., 2005). The large amount of data generated and the complexity of the equipment also results in the requirement of trained personal if accurate results from the equipment and appropriate conclusions from the data are to be obtained.

Continuous monitoring continues to be used sparingly because of the cost and is still primarily deployed in research in an effort to extend our understanding of biological and chemical processes. Nonetheless, it is recognized that the utilization of continuous environmental monitoring equipment will increase because of the ability to detect sudden changes like accidental discharges, and the potential to construct complex models for decision making in water protection and water management from the high frequency data (Vandenberghe, et al., 2005).

2.4 Analysis And Interpretation Of Water Quality Data

2.4.1 Water Quality Guidelines And Assessment Tools

Drinking water guidelines have been developed by the World Health Organization (WHO) using risk assessment to predict maximum concentrations of chemical, microbial, and radiological contaminants in water that a person can ingest without any negative health impacts. The guidelines describe acceptable concentrations after the treatment process has been complete and should not be applied to raw water sources as technology today is capable of purifying even heavily contaminated water to a high quality final product (CCME, 1999). These guidelines are often referenced by nations as they develop their own risk management strategies to provide safe drinking water, often in the form of standards and regulations (World Health Organization, 2008). The provided maximum daily loads assume healthy adult individuals, and persons with compromised immune systems, or at different life stages, such as childhood, may require additional treatment of water to ensure that they will not be negatively affected throughout their life. Canada has national water quality drinking guidelines (Health Canada, 2009), but they are not law, as water resources fall under provincial jurisdiction. For example, in British Columbia only the standards for *E. coli* are legally binding in their Drinking Water Protection Regulation (Government of British Columbia, 2008), however, Nova Scotia has adopted the Canadian guidelines as a legally binding document for public drinking standards (Province of Nova Scotia, 2009). The guidelines are regularly modified as more in-depth research into epidemiology of contaminants is completed. The general result is that the maximum concentrations are decreased and regulations are becoming more stringent.

The Canadian Council of Ministers of the Environment (CCME) has developed guidelines for water quality with respect to different uses which include; agriculture, drinking, recreation, industry, and aquatic health. The guidelines have been created in an effort to provide guidance nationally to protect water resources for the designated use at a designated site. However, the local environment can have synergistic or inhibitive effects on the toxicity or availability of contaminants, resulting in a need for guidelines to be modified with site specific Maximum Allowable Concentrations (MACs). (CCME, 2003)

2.4.2 Statistical Analysis Of Water Quality Data

2.4.2.1 Applicable Probability Distributions

Fitting an appropriate probability distribution is the first step in the analysis of environmental data. The fitting of a probability distribution allows for inferences to be made about the data based upon known statistical probabilities of a distribution (Mcbean & Rovers, 1998). There are many distributions, and a number of these may be required to describe all aspects of water quality, but most parameters can adequately be described by the normal (Gaussian) or lognormal distribution (Mcbean & Rovers, 1998)

The normal distribution is the most permeating distribution throughout nature, as stated by the Central Limit Theorem. The Central Limit Theorem states "under very general conditions, as the number of variables in a sum becomes large, none of which is dominant, the distribution of the sum of random variables approaches the normal distribution, almost regardless of the distribution of the individual variables" (Mcbean & Rovers, 1998). This suggests that given enough data, if water quality is influenced by a number of random factors, the distribution of water quality samples should tend to be normal. In statistics it is convenient when the data is normally distributed, or can be numerically transformed to be normally distributed, because of the vast development and understanding of statistics related to the normal distribution. Many statistical tests and analysis tools have the assumption of normality built into them, and large deviations from normality makes these tools unusable and more complex methods of analysis necessary. However, normality has two assumptions that pose a problem for environmental data. It assumes that the range of possible values is from negative infinity to positive infinity, and that the data is symmetrical about the mean (Mcbean & Rovers, 1998). Many measurement systems of parameters do not go below zero and in nature we find that they may be significantly skewed; in these cases it is beneficial to utilize the lognormal distribution.

To arrive at the lognormal distribution the natural logs of measurements are taken and then treated in the same manner one would treat normally distributed values. It is important to recognize that the natural log of a mean is not the mean of the natural logs (Mcbean & Rovers, 1998). The mean of the samples and the natural log of the means are related by Equation 2.2. The same analysis techniques that are used for analysis of the normal distribution can be used on the lognormal transformation, with minor modifications, which makes the lognormal a powerful transformation for many environmental parameters. The benefits of the lognormal distribution are that it cannot contain values below zero, as the distribution has a lower bound.

Equation 2.2 Mean and Variance of lognormal distributions (Lawless, 1982).

$$u_y = E[y] = \exp\,(u_z + \frac{{\sigma_z}^2}{2})$$

$$\sigma_y{}^2 = Var[y] = u_y{}^2[\exp\left(\frac{{\sigma_z}^2}{2}\right) - 1]$$

$\mathrm{z} = \text{statistics in natural log transformation}$
$\mathrm{y} = \text{statistics in normal distribution}$

Other distributions may be useful in certain situations. These include but are not limited to the Binomial, Poisson and Extreme Value Distibuions. The binomial distribution is characterized by independent trials with two outcomes and a constant probability of each outcome between trials. The binomial distribution is useful for examining pass/fail or true/false phenomena, but is limited by the fact that the data must be binary. The Poisson distribution was derived for use with a large number of trials, where the probability of the outcome was small and constant. It is the limiting case of the binomial distribution when probability of occurrence is small and the number of trials is large. The Poisson distribution describes the number of events or occurrences that happen in a set number of trials, and in an environmental context may describe events such as the number of microorganisms of a certain type as quantified by a lab test. Extreme value distributions are a family of distributions that are used to characterize extreme events. They historically have been used to characterize storm events. The goal of extreme value distributions is not to describe the central tendency of populations, but to describe maximums or minimums. Normal and Lognormal distributions adequately describe the central tendencies of a population, but extreme value distributions are much more adequate at describing the tails of a distribution. (Mcbean & Rovers, 1998)

There are three main types of extreme value distributions which have been derived to represent the tails of different distributions. Type I (Gumbel) is an exponential decay and is representative of the tails of the normal distribution. Type II (Frechet) is a polynomial representation of the tails of distributions such as the lognormal distribution in the positive direction. Type III (Weibull) is representative of distributions that are bounded by a value in the direction of the extreme. The main problem experienced when choosing an extreme value distribution is that once a choice of extreme value distribution is made, subsequent model decisions assume that the initial model decision was correct and do not account for the uncertainty present in model selection. There is the additional option of using a generalized extreme value distribution (GEV) which does not hold the requirement that a specific extreme value distribution be selected, but it requires significantly more data (Coles, 2001).

2.4.2.2 Confidence Intervals

Confidence intervals are descriptions of the range a measured mean is expected to fall within with a particular level of confidence (Berthouex & Brown, 2002). The stated confidence suggests that the mean will fall within these bounds that percent of the time. An increase in confidence will also increase the confidence width, and as a result there is a balance between the two. The most commonly used confidence levels are the 90%, 95%, and 99% confidence levels (Devore & Farnum, 1999). These confidence intervals can be constructed for both normally and lognormally distributed data, but special considerations must be made when constructing them for lognormally distributed data.

Confidence intervals of lognormally distributed data are biased unless modifications are applied to correct them (Zhou & Sujuan, 1997). Ideally the exact method for the computation of the confidence interval would be used, but it is computationally tedious, and approximations in many applications are sufficient (Land, 1972). In many situations, researchers have applied what Land (1972) refers to as the "naïve" method for the computation of confidence intervals for lognormally distributed data.

Take the case when the outcome of X, a two-parameter log-normal distribution where all values of X are positive and non-zero so that Y = log(X) will be normally distributed with mean μ and variance σ^2 (Zhou & Sujuan, 1997). Equation 2.2 is a description of the resulting unbiased normalized mean. The construction of the confidence interval using the "naïve" method is described by Equation 2.3.

Equation 2.3 "Naïve" method of confidence interval construction for lognormally distributed data.

$$\exp\left(\bar{Y} \pm Z_{\left(1-\frac{\alpha}{2}\right)} \frac{s}{\sqrt{n}}\right)$$

The problem with the naïve method as expressed by Land (1972) is “the naïve method gives, in fact, exact confidence limits for e^{μ}, the median of X, and is tolerably accurate for the E(X) only when σ^2 is exceedingly small.” Both Land (1972) and Zhou and Sujuan (1997) state that not even a large sample size will reduce the bias present in the naïve method, and Zhou and Sujuan further determined through simulation that an increasing sample size will actually exacerbate the flaws of the naïve method. Consequently, it is suggested that the naïve method be avoided since reasonable alternatives are available.

The best approximation, suggested by Land (1972) and by Zhou and Sujuan (1997), for moderate to large sample sizes is the Cox transform. It benefits from both a relative simplicity when compared to other approximation methods and the exact method, and also has been shown to have good performance. Cox proposed that construction of the confidence interval be based on a simple modification that utilizes the uniformly minimum variance unbiased estimator (UMVU) of the mean and variance for log transformed data (Zhou & Sujuan, 1997). An UMVU is mathematically defined by Lehmann and Casella (1998) as unbiased if Equation 2.4 is satisfied. The unbiased estimate of the mean of log(X) is defined by equation 2.2 while the corresponding UMVU of the variance is $S^2/n + S^4/(2(n-1))$. Cox’s realization of the suggested UMVU confidence interval is provided in Equation 2.5.

Equation 2.4 Definition of an unbiased estimator.

$$E_{\theta}[\delta(X)] = g(\theta) \quad \text{for all } \theta \in \Omega$$

Equation 2.5 Cox's UMVU estimator method for construction of lognormally transformed data.

$$Y + \frac{S^2}{2} \pm Z_{(1-\alpha/2)}\sqrt{\frac{S^2}{n} + \frac{S^4}{2(n-1)}}$$

2.4.2.3 Statistical Hypothesis Testing

One key goal of water quality monitoring is to detect changes, differences or anomalies in a system. Hypothesis testing is used to statistically analyze temporal changes in parameters, spatial differences, or if a perceived anomaly is statistically significant. Hypothesis testing is a process where a hypothesis, for example that two populations are statistically different, is constructed and tested with a given statistical confidence (Berthouex & Brown, 2002).

These statistical procedures are designed to extract information from sample data and make inferences about the population. As a result, the tests are only as good as the data they are constructed from. The final statement is never a guarantee that our findings are true, as hypothesis testing is probability based and there is the chance of either type I error or type II error occurring. Type I error is the risk that the procedures will provide a false positive (the null hypothesis is rejected when it should be accepted), and type II error is a false negative (the hypothesis is accepted when it should be rejected). Type I and type II errors are not independent of each other, and as a result by attempting to minimize one type of error you will increase the other. In engineering practices we view type II as more serious, and as a result care should be exercised when designing tests with a high probability of type II error (United States Department of Agriculture, 1996). It also is not recommended to increase the confidence of a test from 95% to 99% because the probability of a type II error occurring increases from 0.245 to 0.5 (Mcbean & Rovers, 1998).

The student t-test is a commonly applied hypothesis test, and is used to compare a sample to the population mean. This is useful for quality control and assurance. It can also be utilized to compare central tendencies of data sets. There are implicit assumptions in the student-t test; the observations must be independent, normally distributed and both data sets must have equal variance. Independence between samples in environmental data is often assumed and it is a questionable assumption as both temporal samples and spatial samples are unlikely to be completely independent. There are modified t-tests that can be utilized to relax the implicit assumptions, but care must be exercised when utilizing them because they are approximate tests and as a result have their own limitations. The paired t-test is a notable modification as it allows for the determination if two populations are the same when it is known that they are not independent. This is very useful in analysis water quality when looking at temporal data for one location, or looking at spatial data along a stream as data independence cannot be assumed unless the distance between samples is large. (Mcbean & Rovers, 1998)

2.4.2.4 Outlier Detection

Outliers are data values that are extremely high or low, and diverge from the main body of a data set (Moore & McCabe, 2006). This means that statistically they do not fit the proposed distribution and as a result the outliers have a disproportional influence on statistical test outcomes. There are statistical tests that can be performed to analyze for outliers, but even if a test determines that statistically the value is an outlier it does not mean that the value is erroneous or invaluable. In water quality, and many measurement systems, the outliers could contain the most valuable information. Cases where this would be true are in extreme value distributions, such as those constructed for high flows and maximum concentrations. The inclusion or exclusion of an "outlier" from a data set is an experience and knowledge-based decision. Knowledge of how a sample was collected, tested, and recorded is essential in determining what data should be omitted in order not to bias the data set. Statistical tests should not be used to discard data, but to tag data that deserves further observation to determine its validity.

Methods to detect outliers include probability plotting, scattergrams, standard deviation tests, and standard normal tests. Probability plotting and scattergrams are limited statistically as they give you a visual understanding of the values, but do not provide statistical evidence by which to consider a value an outlier. As a result graphical methods are best used to detect possible outliers, which should be analyzed further. Standard deviation tests analyze the cumulative standard deviation and attempts to determine if a data point disproportionally influences the standard deviation. A more formalized, and the most common way to detect an outlier is to determine if a point falls outside three standard deviations from the mean. A value outside three standard deviations is considered very large or small and the value may be rejected based on the likelihood of this actually occurring (0.1%) (

Equation 2.6). (Mcbean & Rovers, 1998)

Equation 2.6 General Determination of Outliers.

$$Tn = \frac{largest\ value - mean}{standard\ deviation}$$

$$If\ Tn > 3\ Data\ point\ is\ an\ outlier$$

2.4.2.5 Time Series Analysis

2.4.2.5.1 Seasonal Trends

Various water quality parameters exhibit strong seasonal variability, which needs to be considered when analyzing water quality trends or statistics. The variability caused by the seasonal trends is due to hydrological, biological, and activity changes within the watershed (Hirsch et al, 1991). Commonly trigonometric deterministic models are applied to account for seasonal variability.

Equation 2.7 Deterministic seasonal model.

$$y_t = A(\cos\ wt + C)$$

Y_t = value of deterministic function
T = time step
W = 360 degrees / number of time intervals per cycle (Loftis & Ward, 1980)

If seasonal trends are not removed the data may not appear to meet all the requirements to be treated as normal or lognormal. Data that has not had seasonal trends removed will likely have correlated residuals related to when samples were taken, which violates independence assumptions.

2.4.2.5.2 Autocorrelation

One assumption of many statistical tests, and especially prevalent in parametric tests, is that measurements are independent. Environmental water quality data sets with relatively short interval (days or less) between samples exhibit a significant serial correlation. As a result, it is necessary to account for serial correlation to obtain meaningful statistical indicators such as means, variances, and confidence intervals (Loftis et al., 1991).

In environmental processes it is impossible to obtain information on all the processes that will influence data. It is typically necessary to model the unknowns as a random element. Models which assume a random element are called stochastic. Autoregressive Moving Average (ARMA), also known as the Box-Jenkins models (Equation 2.8), are stochastic models which assume that measurements which were taken closer in time will be more similar than measurements with a greater amount of time between them (Box et al., 2008). ARMA models consider both the autocorrelation in measurements and the autocorrelation in the random noise element.

Equation 2.8 ARMA model structure (Box et al., 2008).

$$Z_t = \phi_1 z_{t-1} + \cdots + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$$

Z_t = parameters present value in time series
Z_{t-1} = value of parameter at one time interval in the past
ϕ_p = autoregressive coefficient at time lag p
θ_q = moving average coefficient at time lag q
α_t = random noise term or shock at time t
α_{t-1} = random noise term at time t-1

The main assumption of an ARMA model is that it is stationary. Stochastic processes are deemed strictly stationary if the joint probability distributions stay constant in the instance of a time origin shift. A Gaussian process is stationary if it has a fixed mean for all values of n and a symmetrical autocovariance matrix (Equation 2.9). If this cannot be verified, the model must be extended to an Autoregressive Integrated Moving Average (ARIMA) model which allows for change in the mean with time. (Box et al., 2008)

Equation 2.9 Autocovariance matrix of a stationary process.

$$\Gamma_n = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{n-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \cdots & 1 \end{bmatrix}$$

2.4.3 Effect Of Autocorrelation On Statistical Measures

The effect of autocorrelation on statistical measures is not simply an improvement or depreciation of statistical confidence, as the influence of autocorrelation is dependent upon the statistical question that is posed. It can be shown that if the question is not bounded in time or space such as ‘what is the average annual TOC concentration’ autocorrelation increases the level of uncertainty and will result in a widening of confidence intervals (Loftis & Ward, 1980). This widening of the confidence interval is a result of the inability to treat the samples within the year as independent samples and as a result you statistically have a much smaller effective sample size from which to characterize the parameter's variance. It can be recognized from Equation 2.10 that the greater the correlation the smaller the effective sample size becomes, which results in an increase in variance and the confidence interval width.

Equation 2.10 Variance of dependent samples.

$$Var(\overline{X}) = \frac{\sigma^2}{{n_b}^*}$$

$$\frac{1}{nb^*} = \frac{1}{n} + \frac{2}{n^2}\sum_{j=1}^{n-1}(n-j)\rho(j)$$

$nb^* =$ number of independent samples
$n =$ number of samples
$t =$ time interval
$\rho(j) =$ Autocorrelation at lag j

Although the characterization of the annual mean of a parameter's concentration might be inhibited by autocorrelation, similarly worded questions experience very different outcomes when considering autocorrelation. The question "what is the mean parameter concentration for a given year," is a distinctively different question than the one previously asked and is impacted differently by autocorrelation than the previous question. This statistical problem raises the issues of "sampling of internally correlated lots" and was addressed in a two part paper by Muskens and Kateman (1978). Musken and Kateman devised equations for one-dimensional lots (either varying in time or space) for Gaussian stationary stochastic processes of the first order. Their work, as stated in part II of their paper is highly applicable to the sampling of water over a set duration in time. A lot is defined as the spatial or temporal group which is to be statistically analyzed. Generally, with respect to water quality, this would be a season or an annum. The variance of a water quality parameter with a given sampling frequency (assuming normality and first order process) is the realization of integrated variances present in Musken and Kateman's (1978) paper for sample sizes approaching 0, where sample size refers to the physical size of each sample (Equation 2.11). Autocorrelation, in this instance, is beneficial because highly correlated samples allow you to infer about the values between samples, and provides a greater confidence for estimates of the true mean.

Equation 2.11 Realization of variance for internally correlated lots with small sample size.

$$\begin{aligned}
\sigma^2{}_{est} = {} & \frac{\sigma_x^2}{n}\left[1+2\left\{\frac{\exp\left(-\frac{A}{T_x}\right)}{1-\exp\left(-\frac{A}{T_x}\right)} - \frac{\exp\left(-\frac{A}{T_x}\right)\left(1-\exp\left(-\frac{P}{T_x}\right)\right)}{n\left(1-\exp\left(-\frac{A}{T_x}\right)\right)^2}\right\}\right] \\
& - \frac{2\sigma_x^2 T_x}{nP}\left[2n - \frac{1-\exp\left(-\frac{p}{T_x}\right)}{1-\exp\left(-\frac{A}{T_x}\right)}\right. \\
& \left. + \frac{1-\exp\left(-\frac{P}{T_x}\right)}{1-\exp\left(\frac{A}{T_x}\right)}\right] \\
& + \frac{2\sigma_x^2 T_x^2}{\mathrm{P}^2}\left[\mathrm{P}/T_x - 1 + \exp(-\mathrm{P}/T_x)\right]
\end{aligned}$$

$\sigma^2{}_{est}$ = Estimated variance of the lot (year)
$\sigma^2{}_x$ = Computed variance of the year assuming independence
A = Distance between two adjacent samples (P/n) (1/365)
Tx = Time constant of the process (first order internal correlation parameter)
P = Lot size (1 year)
n = number of samples taken equally spaced from lot P (365 for daily samples)

Chapter 3 Methodology

Two subwatersheds, Lacey Mill River and Peggy’s Brook, in the Pockwock Watershed were selected based upon their land use, size, and hydrological characteristics for the study. A 16 month monitoring program was conducted where pH, temperature, conductivity, turbidity, DO, and flow were monitored continuously, and TP, TSS, TOC and *E. coli* were measured discretely on a weekly basis. The sampling period was from April 2009 until August 2010 inclusive. Data was then analyzed through a plethora of statistical methods to determine the most effective measurement practices including, choice of parameters, sampling interval, and sampling regime.

3.1 Site Description

3.1.1 Location

The Pockwock Lake watershed (Figure 3.1) Is the source water for Halifax Regional Municipality west of the harbor and is located in Nova Scotia at 44^{0} 48’ N 63^{0} 51’ W. The watershed is split by the Hants County - Halifax County line with the northwest portion of the lake and the encompassing watershed area belonging to Hants County and the southeast portion belonging to Halifax County. The Northern portion of the watershed is traversed by highway 101 as it passes south of the community of Mt. Uniacke. The entire watershed was delineated using a 10 m digital elevation model from 10 m contours and determined to have an area of 5502 hectares (55.02 km^2). Two subwatersheds were chosen for the study to represent different land use characteristics of the watershed and analyze how water quality parameters may vary throughout the watershed.

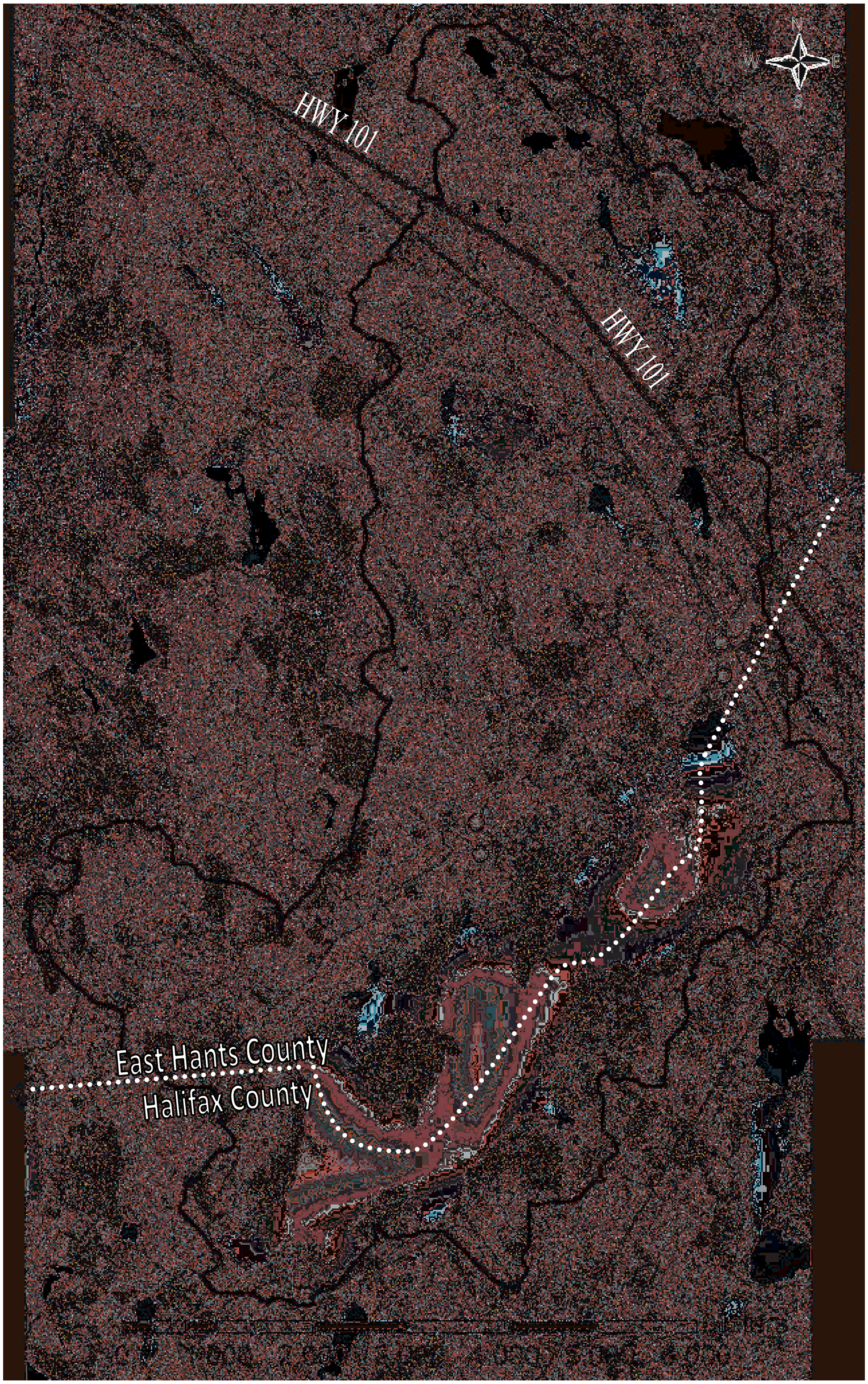

Figure 3.1 Watershed boundary of Pockwock Lake on an orthographic photo (Service Nova Scotia and Municiapl Relations, 2007).

Lacey Mill subwatershed (Figure 3.2) comprises the northern section of the watershed and contains a network of smaller lakes terminating in Little Indian Lake at the headwaters. The Lacey Mill subwatershed contains some notable development on West Lake, as part of Etter Settlement. The Lacey Mill subwatershed is of particular interest because of the development located within it and its large proportion of watershed area (one third of total watershed area). Forestry activity has been present inside the watershed and continued throughout the monitoring period of the current study. The Lacey Mill River subwatershed has an area of 1850 hectares (18.50 km^2), of which 1780 (17.8 km^2) were captured at the monitoring site located at 44^o 50.112’ N 63^o 48.648’ W.

Peggy’s Brook subwatershed (Figure 3.3 Peggy Brook subwatershed study catchment and location of monitoring site.) is located on the northern side of the lake has an area of 350 hectares (3.5 km^2) of which 320 hectares (3.20 km^2) is captured at the monitoring site located at 44^o 48.948 N 63^o 50.447 W. It is only accessible by logging roads and is completely forested. It has not been logged since the late 1980s.

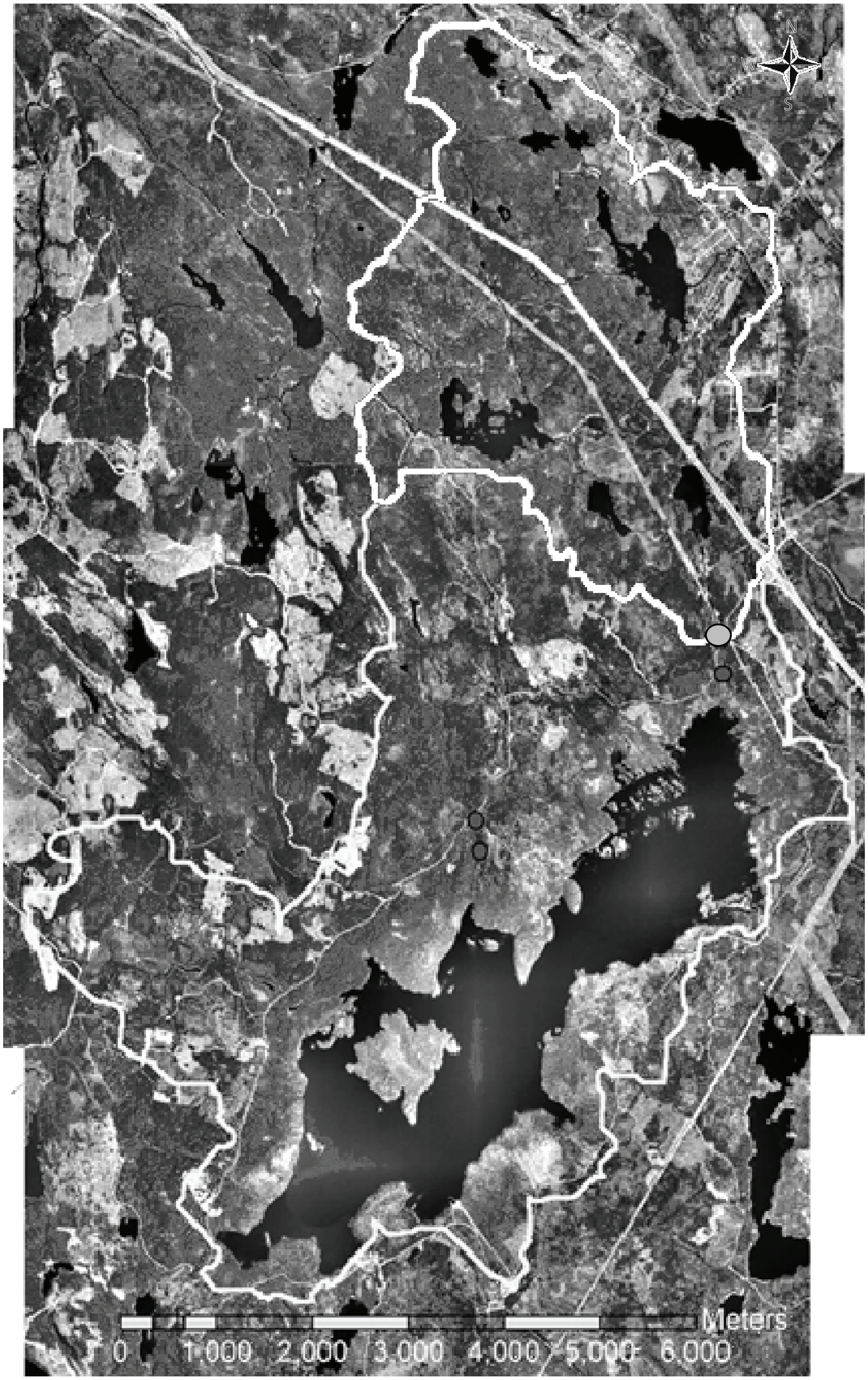

Figure 3.2 Lacey Mill subwatershed study catchment and location of monitoring site (Service Nova Scotia and Municiapl Relations, 2007).

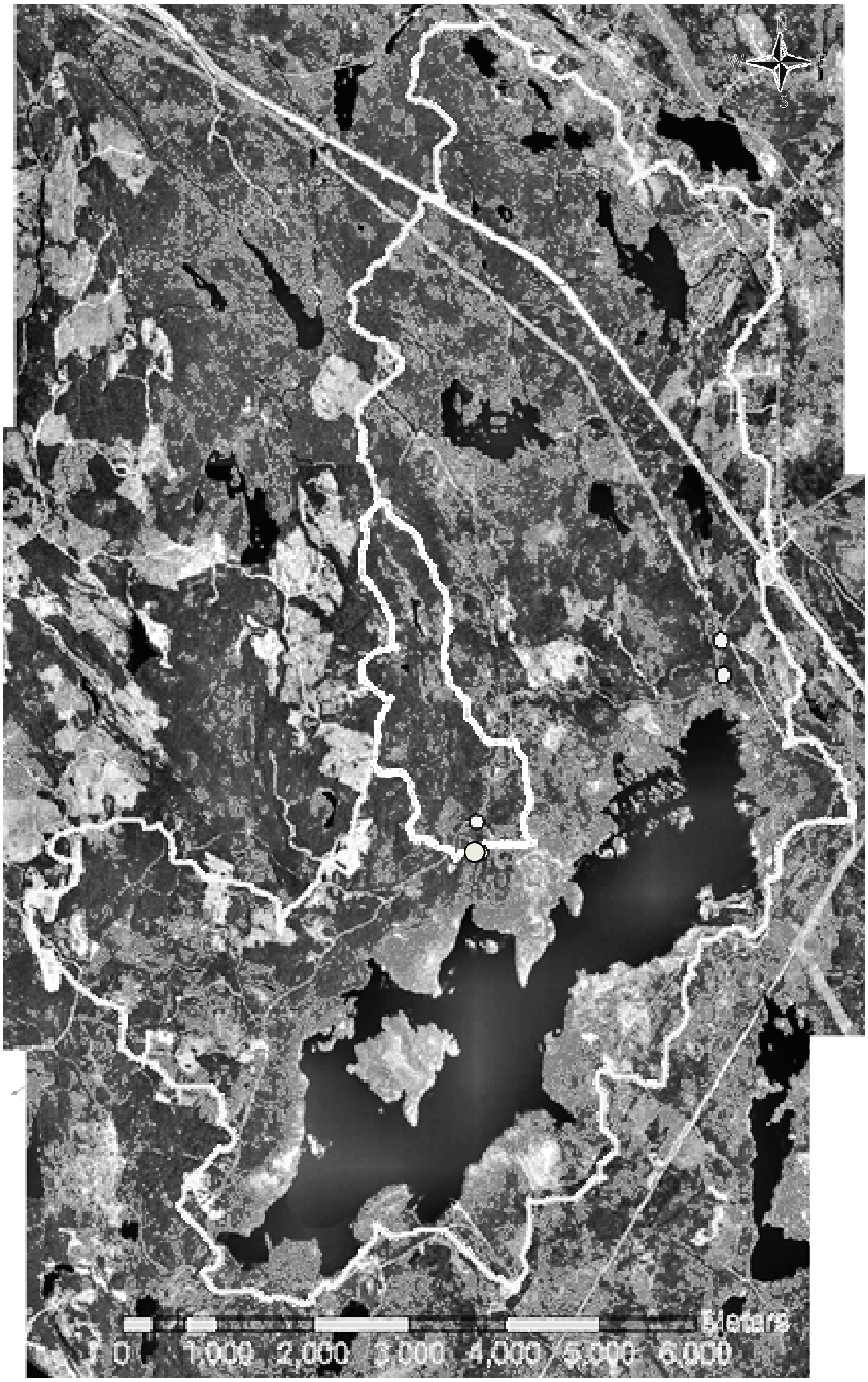

Figure 3.3 Peggy Brook subwatershed study catchment and location of monitoring site (Service Nova Scotia and Municiapl Relations, 2007).

3.1.2 Land Use

The watershed south of highway 101 is crown land and is a designated watershed under the Nova Scotia Environment Act. This designation provides the source water protection under provincial law and institutes restriction on activities such as fishing and motorized vehicle use.

Development has been limited and greatly suppressed by Halifax Water with all land south of highway 101 being restricted to development because of its status as designated watershed. However, Halifax Water only has direct control of land north of highway 101 that they have acquired. This has led to an institutional mandate to purchase land within the watershed north of 101 as it becomes available, as well as work with community members to best protect the northern portion of the watershed that does not fall within the designated watershed.

There is small scale logging in the watershed focused on maintaining tree stand health. The main forestry techniques used are variable retention and selection management (Table 3.1). Clear cutting is still an optional technique but not has been utilized since 2006 (Barry Geddes, Personal Communication, November 2010). The watershed south of highway 101 has minimal gravel roads that are only maintained for logging access and are gated to the general population.

Table 3.1 Breakdown of used forestry management technique (Barry Geddes, Personal Communication, November 2010).

Year	Variable Retention (% of total)	Selection Management (% of total)
2007	100%	0%
2008	87%	13%
2009	61%	39%
2010	44%	56%

As most of the land is crown land the public is allowed to use the watershed for recreational purposes such as hiking, cross country skiing and hunting, but is not to interact with the lake or streams through fishing, boating or swimming. Use of motorized vehicles is also prohibited which includes but is not limited to skidoos, all terrain vehicles, and cars (Province of Nova Scotia, 2009).

3.1.3 Climate

Environment Canada's climate normals for 1971-2000 recorded a daily average temperature of 6.3 ^{o}C and an average precipitation of 1452 mm/yr. In 2009 the average daily temperature was considerably higher at 11.1 ^{o}C, but the precipitation was normal when compared to the climate normals, with 1472 mm/yr. Though normal through the year, August was wetter than usual with 179.6 mm compared to the normal of 92.7 mm. A drier November was also experienced with 95.1 mm in comparison to 146 mm. The eight months of sampling in 2010 had an average temperature similar to the normals but the precipitation for the first eight months of 2010 was considerably lower than the climate normals with only 635.5 mm compared to the expected 919.3 mm. The spring was unusually dry with only 88 mm of rain from April to May compared to the expected 225 mm in accordance with the climate normals. With the exception of June and July the precipitation totals for any month were less than 70% of the expected and in April was only 30% that of the expected rainfall. The data suggest a fairly normal climate with respect to temperature but a deviation from normal climate when considering precipitation. Climate variability is an important factor to consider when making inferences from relatively short term watershed monitoring studies, as drawn conclusion may not be valid for years when the climate is very different than the period that was studied.

3.2 Monitoring Methods

The initial monitoring plan involved deployment of a YSI 6600 mutiparameter sonde and a Hobo water level logger continuously throughout the sixteen months of data collection. The continuous monitoring was to be supplemented by weekly discrete monitoring for water quality indicators that could not be sampled continuously, and for quality control and drift correction of parameters that were monitored continuously. Due to lack of accessibility caused by heavy snowfalls and logging roads that are not maintained in the winter, continuous monitoring equipment was removed from Peggy's Brook and monitoring was interrupted from January till March, and discrete sampling of Lacey Mill River monitoring was also scaled back to biweekly. There also are data gaps, of about a day's length, every four to five weeks caused by the necessity to remove monitoring equipment for maintenance and calibration.

3.2.1 Continuous

Continuous monitoring equipment was deployed to take measurements of several water quality parameters at fifteen minute intervals. A fifteen minute sampling interval was chosen to be the minimum reasonable interval of data collection. This was based upon the amount of data that would be collected and the length of maximum deployment. Shorter sampling intervals will have a heavier drain on the power supply, and as batteries were the only power source, shorter intervals increased risk of data loss due to power failure. It was also assumed that any meaningful water quality event would last for fifteen minutes and be captured by the sonde.

At the monitoring sites of both Lacey Mill River and Peggy's Brook a YSI 6600 multiparameter sonde was deployed. These are self contained data loggers which can be fitted with sensors to measure an array of biological and chemical water quality parameters. The sondes were inspected weekly and the data downloaded at least on a bi-weekly basis. The sondes were removed monthly to be cleaned, inspected, and recalibrated. Hobo water level loggers were also placed at each site to record continuous water levels. These sensors, provided they did not fail over the period of the study, were left in-site and checked weekly for functionality.

3.2.1.1 Water Quality

The sondes were fitted with sensors to measure DO, conductivity, temperature, pH, and turbidity. This provided a broad selection of biological and chemical indicators that could be measured by continuous monitoring equipment currently available on the market.

3.2.1.1.1 Dissolved Oxygen

YSI 6600 model sondes use a patented *Rapid Pulse* system for measurements of dissolved oxygen. The benefit over historically standard electrochemical detectors for DO is that the Rapid Pulse is not significantly affected by flow. The *Rapid Pulse* system utilizes a Clark-type sensor, which is common in many membrane-covered DO probes. The membrane encloses a small layer of electrolyte which is required for the reduction of the silver electrode while isolating the electrodes from the external environment. Clark cells measure the current associated with the reduction of oxygen which diffuses through the Teflon membrane. The current is proportional to the partial pressure of the oxygen in the solution (Eutech Instruments, Singapore). Due to biofouling, electrode oxidation, and electrolyte depletion, drift can be expected in long term deployments.

During the monthly maintenance retrievals the membrane was removed and the electrodes were lightly sanded to remove oxidation products that had formed. New electrolyte was placed onto the electrodes while a new membrane was carefully installed making sure not to have any air bubbles present between the membrane and the electrode. The sensor was then recalibrated and tested for functionality which included response time, stability, accuracy, and precision.

3.2.1.1.2 Conductivity

The sonde utilizes a cell with four pure nickel electrodes to measure conductance in an aqueous environment. Two of the electrodes are driven by the power supply while the other two are used to measure voltage drop. The voltage drop is then related conductance.

Instead of measuring conductivity specific conductance was measured. Conductivity is dependent on temperature while specific conductance is a generalized conversion to incorporate the changes that are expected due to temperature changes in the solution.

Equation 3.1 Specific conductivity relation to conductivity.

$$specific\ conductance(25\ {}^{0}\mathrm{C}) = \frac{conductivity}{1 + TCx(T - 25)}$$

$\mathrm{T} = \text{temperature in Celsius}$
$\mathrm{TC} = \text{temperature coefficient } 1.91\%/{}^{0}\mathrm{C}\ (\mathrm{TC}=0.0191)$

The sonde, as used in this study, provided measurements of specific conductance.
The conductivity cell was inspected for debris during weekly visits and was cleaned upon monthly maintenance retrievals. The cell was calibrated during the monthly maintenance retrievals and diagnostics were run to test for functionality.

3.2.1.1.3 Temperature

The temperature sensor on the sonde is a thermistor composed of sintered metallic oxides. Thermistors predictably change resistance with change in temperature. The measurable change in resistance can be related to temperature and is constant so there is typically no requirement for calibration or maintenance of the temperature sensor. However, temperatures were compared between the sondes, a thermometer, and hobo sensor loggers at launch, and were checked throughout the study for proper functionality.

3.2.1.1.4 pH

The pH probe consists of a combination electrode consisting of a proton selective glass bulb containing pH 7 buffer solution and a Ag/AgCl reference electrode using gelled electrolyte. Protons (hydrogen ions) interact with each side of the glass bulb, creating a potential gradient. Since the concentration of protons on the interior bulb is constant, the pH of the external solution can be determined from the gradient. The pH probe is exhaustible as the gelled electrolyte degrades, and this results in drift. Under good operating procedures with deployments less than a couple of months the pH probe is stable (drift of less than 0.1 ph units), but as the probe approaches exhaustion more care must be taken to ensure accurate measurements. Probes need to be replaced at least every two years, but depending on application and length of deployment might need to be replaced as often as every six months.

The probes for pH were replaced once during the study at a point about three months into the study. The pH probes were recalibrated during the monthly maintenance retrieval with a two point calibration with buffers 4 and 7. The buffers were chosen because measured pH in the streams was found to be generally between 4 and 5.5. Diagnostics were also run on the probe to analyze health of the probe and the response time of the probe.

3.2.1.1.5 Turbidity

The turbidity sensor consists of a light emitting diode radiation source and a photodiode detector with high sensitivity. The International Standards Organization (ISO) recommends the use of a light source with a wavelength between 830 and 890 nm and a 90 degree angle between emitter and detector. YSI has conformed to these recommendations. YSI has two models of turbidity probe, and their latest model (6136) was utilized in this study. YSI claims its readings are more comparable to large cell volume laboratory turbidimeters like the Hach 2100AN readings, which have generally been accepted as industry standard. (Onset, Pocasset, MA)

The turbidity sensor is not prone to drifting, but can be prone to fouling or physical damage. Two point calibrations at 0 NTU and 100 NTU were performed. The turbidity sensor has a linear response, and as a result using a high turbidity standard is more practical than attempting to calibrate with standards closer to expected values, which will be more sensitive during the calibration process. Initially calibrations were performed on a monthly basis during the monthly maintenance retrieval; this was changed to every second month as the sensors did not drift. The sensors were inspected for damage and tested in standards if any problems in functionality were experienced.

3.2.1.2 Flow

For water level logging a Hoboware u20 9-meter depth water level loggers were deployed at each site. A stage discharge curve, which is a relationship between water level and stream flow, was constructed to compute continuous flow measurement. One water level logger was deployed in a stationary position submerged in each stream, and one was deployed in the air which allowed for barometric pressure corrections. The sensors are a ceramic pressure sensor with the capability to records pressures up to 207 kpa, which corresponds to a maximum depth of 9 m (Onset, Pocasset, MA). Knowing the difference between barometric pressure as measured by the sensor in the air and the pressure recorded by the logger in the stream, and utilizing the density of water, the water level can be determined. The sensors cannot be recalibrated by the user, and if recalibration is required they must be sent back to Onset, the manufacturer. Provided there was no physical damage, the sensors were precise. The data was regularly retrieved, at least on a biweekly basis, and analyzed for quality. Sensors failures are easy to identify by unusual fluctuations in pressure and were replaced with a backup upon discovery.

3.2.2 Discrete Sampling Program

Discrete samples for TOC, TP and bacteria were taken by grab sample during weekly site visits. The technology has not advanced to a point that equipment is readily available that can continuously monitor these parameters. Discrete measurements for all continuously monitored parameters were also taken for quality control and quality assurance. Lastly, regular stream flow measurements were taken in order to create a stage discharge curve for each stream.

3.2.2.1 Continuous Monitoring Confirmation

A handheld YSI multiparameter sonde, the 600R, was used during weekly site sampling to monitor for parameter drift and to assure accuracy of the data. The parameters measured by the handheld sonde are: temperature, dissolved oxygen, conductivity, and pH. The 600R used has identical sensor technology to the 6600 which was deployed in-situ.

3.2.2.2 Turbidity

Samples were collected and brought back to the lab weekly and turbidity was measured on a bench top turbidity meter. A Lamotte 2020e turbidity meter was used for the majority of the study, but it failed on August 4th 2010 and a Hach 2100AN was used for the final month of sampling.

3.2.2.3 Total Organic Carbon

TOC samples were collected and acidified immediately in order to preserve the samples. Prior to analysis, the samples were stored in a fridge at 4 ^{0}C. The combustion-infrared method, method 5310 B in Standard Methods for Examination of Water and Wastewater (1999), was used to analyze the TOC concentration, and the analyzer used was a Shimadzu TOC-V CSH.

3.2.2.4 Phosphorous

Phosphorous samples were collected in clean plastic bottles, frozen immediately and then analyzed when there were a sufficient number of samples. The storage period was as long as four months. They were analyzed using the ascorbic acid method in accordance to Standard Method 4500-P (1999 version) with detection performed by an LKB Ultrospec 4051 spectrophotometer (American Public Health Association, 1998).

3.2.2.5 Flow

Stream flows were measured at least bi-weekly during the study using the velocity-area method in accordance with the Canadian General Standard Boards recommendations developed in 1991. Flow was measured using a Model 622 Price meter made by Gurley Precision Instruments.

3.2.2.6 Bacteria

The presence and concentration of *E. coli* and Total Coliforms were enumerated using the media filtration method. 100 ml of water was filtered and plated on the selective media m-coliBlue24 Broth produced by Hach. The plates were then incubated for 24 hours at 37.5 ^{o}C, and following the incubation the colony forming units (CFUs) were counted. During a shortage of m-coli Blue24 Broth, which occurred from June - August 2009, modified mTEC Agar was used as a selective media. The procedures remained the same, but the plates were incubated at 44.5 ^{o}C and only *E. coli* was enumerable with the mTEC media.

3.2.3 Quality Assurance and Quality Control

At least twice during the sampling period triplicates of all the parameters that were measured discretely were taken to examine both the variability in the stream, and the quality of lab procedures. To assess the lab procedure one large sample was drawn from the sampling site and three aliquots were made from the sample and then analyzed. To measure stream variability three separate samples were drawn from the stream in quick succession and then analyzed. The variability across the streams was not analyzed because Lacey Mill River was five meters wide at its largest point near the sampling site and Peggy's Brook was only three meters. Due to the physical nature of these streams at the sampling site they could be considered well mixed.

3.3 Data Analysis

3.3.1 Data Correction

The challenges of continuously deploying monitoring equipment are: (i) that equipment is consistently in an active state, (ii) the environment potentially causes biofouling, and (iii) the measurements are strongly affected by intermittent water quality conditions. Sensors that consist of one or more expendable components are prone to calibration drift as the functionality of the sensor will change as the expendable portions are consumed. Equipment located in productive water environments will grow biofilms, this natural process is referred to as biofouling and can modify functionality of a sensor by impeding chemical or physical processes that the sensor's performance is dependent on. Intermittent water quality conditions, although important, can make data analysis challenging because it can be difficult to determine cause and effect.

In this study sensors predicted to be particularly susceptible to drifting were pH and dissolved oxygen. It was found that these two parameters were the most susceptible to drift but on a monthly deployment still functioned adequately. Turbidity was the parameter most influenced by biofouling, and at times biofouling resulted in unusable turbidity data. No other sensors were affected by biofouling to an extent that functionality of the sensor was compromised.

Data correction is both a quantitative and qualitative process. It was assumed that the hand-held sonde, that was used weekly, would provide a more accurate measurement of a parameter than the in-situ sonde, and the handheld sonde measurement served as the "true" value. Offset corrections of the measurements would be applied to the in-situ sonde data to be in accordance with weekly data from the handheld sonde. Linear drift corrections of the continuous data were applied between weekly site visit data points. Qualitative assessment of how measurements from the handheld sonde related to the in-situ measurement were made and the handheld readings were, on occasion, rejected when suspected of providing a poor reading. This was often a result of poor acclimation, especially in colder water conditions.

Outliers and noise were removed only in cases when they were known not to be caused by extreme events. Periods when outliers and noise were removed include durations it could be ascertained that: (i) sensors were removed from the water for inspection and downloading, (ii) a significant level of biofouling had occurred, (iii) a sensor had failed, or (iv) unrepresentative intermittent measurements were recorded. Significant biofouling and sensor failure were determined by physical inspection of sensors and inspection of data. Data suggested biofouling when successive readings were outside the range expected in the natural environment and no recent precipitation had occurred. Sensor failure was determined by data analysis and verification of the sensor failure through diagnostics in the lab. These diagnostics included both software diagnostics and comparison to standards. Most data segments could be further verified for removal as grab samples were collected during the same time period and produced different results.

With the exception of turbidity, intermittent anomalies were generally accepted based on the premise that there was no physical data to reject that they occurred. Turbidity sensors are affected by debris interacting with the sensor, and events that lead to turbidity shifts of an order of magnitude or greater were removed as their likely causation was debris. A large number of data points will ensure that sporadic outliers will not have a significant weighting on water quality statistics and as a result removal of unexplainable outliers was considered acceptable

3.3.2 Gap Filling

Missing sections of continuous data were filled depending upon: (i) the length of data gap, (ii) the current and antecedent conditions in the watershed (recent rainfall events), and (iii) the parameter being filled. If the conditions of the watershed were stable (no rainfall event and a stable trend in stream flow) and the data gap was small (less than six hours) a linear interpolation was applied to fill data gaps for all types of data. Linear interpolations for missing flow records were reliable as long as no recent rainfall events occurred shortly before the rainfall event or during the rainfall event. Even if these interpolations are not ideally accurate, because of the numbers of samples these gap filling techniques were applied to less than 1% of the data. For longer data gaps or gaps occurring when there are storm flows, statistical correlations were applied.

3.3.3 Trend Removal

3.3.3.1 Seasonal

Deterministic trigonometric functions were fit to describe seasonal trends. For a more in-depth discussion see Appendix F.

3.3.3.2 Autocorrelation

ARMA models were developed and assessed using a number of criteria which included: (i) t-testing for parameter significance, (ii) graphical examination, and (iii) Akaike Information Criterion (AIC). AIC is a metric of most likely outcome, and is used to determine the autoregressive and moving average orders which most effectively represent the "true" orders of the ARMA model, considering accuracy versus complexity. To determine autocorrelation the data must be gapless, and in order to meet this criterion analysis was performed on the five longest gapless periods and the results were averaged.

3.4 Loading Estimates

Loading estimates are the product of the concentration of the pollutant or nutrient being analyzed and the flow in the stream. As a result it is imperative that both an accurate estimate of both flow and concentration be obtained. The loading estimate can only be as accurate as the parameter with the least confidence; concentration measurements typically limit the estimate (Littlewood, 1995).

3.4.1 Flow Data Processing

Stage discharge relationships were constructed to provide continual flow measurements from pressure transducer readings from within the streams. Two issues resulted in the need to fill-in sections of the temporal flow record: (i) sensor failure and (ii) inability to characterize high flows with the stage discharge relationship. In order to fill data which was not collected by the sensor due to failure a multiple regression equation was constructed between the two streams in the watershed with reasonable results (r-squared values of 0.73 and 0.80 respectively for predicting Lacey Mill River and Peggy's Brook). If the stage discharge relationship did not cover a high flow event the high storm flow for the section of data was extrapolated from the data before the event, and after, and shaped as a typical hydrograph (Appendix C: Filling Extreme Flow Events). This is the best way to estimate flows until a more complete storm discharge relationship, one that captures higher flows, is created.

It is common practice to create a flow relationship between two streams that are in close proximity and deemed to have similar characteristics (geology, geography etc.) to predict one stream's flows continuously using continuous flows of the other. During examination of paired flows between the steams it was apparent that there was hysteresis in the relationship (figure 3.4), and as a result it was also important to examine whether the watershed was in a wetting or drying trend. It is beneficial to use a multiple regression that takes into consideration of how the flow is changing, as quality of prediction improves.

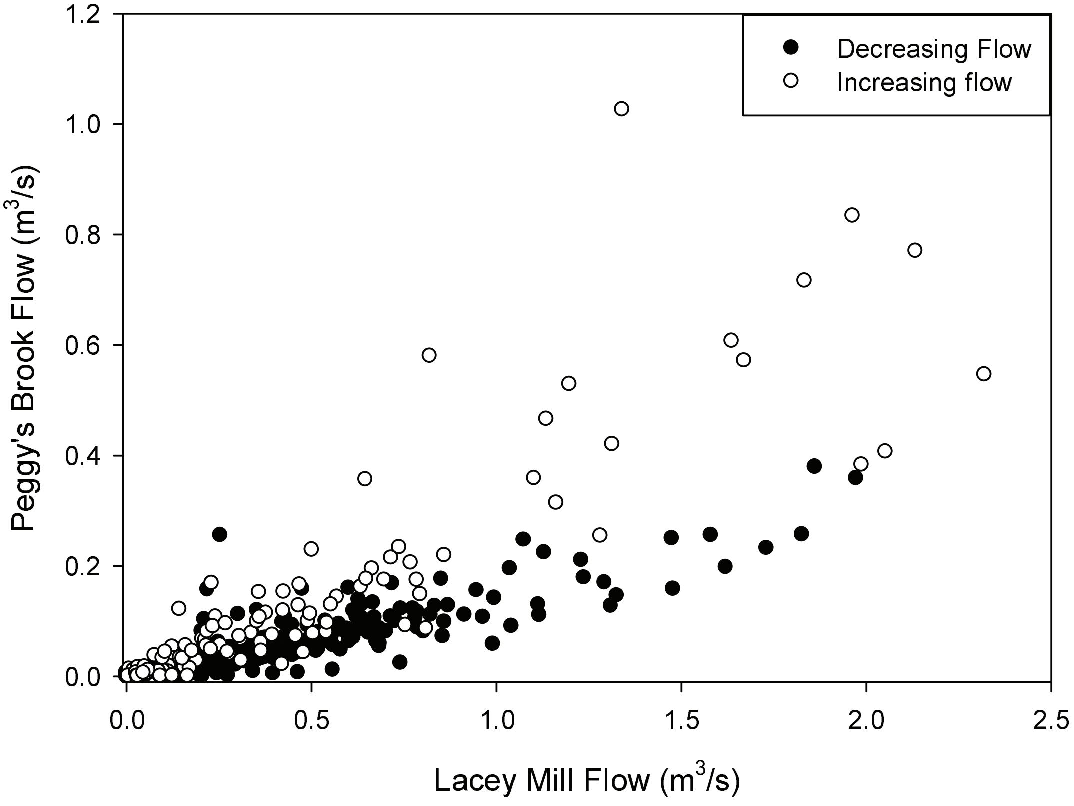

Figure 3.4 Hysteresis in stream flow correlations on a daily time step.

For construction of the regression model a one hour smoothing window was applied to the flows to remove some of the noise present in the signal. The predictors in the model were the flow in the measured stream, and how the flow in the measure stream had changed over some period in time (equation 3.1 and 3.2). Numerous time periods were examined for the lag, and in order to keep the analysis parsimonious 12 hours was selected. This time period was optimal for predicting Peggy's Brook but was suboptimal for predicting Lacey Mill River. This value was chosen because it was close to the determined time of concentrations for both Lacey Mill River and Peggy's Brook which had times of concentration 12 and 10.5 hours respectively (Appendix A). Lacey Mill River's actual time of concentration is likely significantly higher than the calculated 12 hours as the equation used ignores the effect of the headwater lakes. Also when examining time lags for predicting Lacey Mill River flows there was considerable improvement with lags up to 24 hours (r-squared 78%). The calculation of time of concentration and further description can be found in Appendix A.

Equation 3.2 Prediction of Lacey Mill River flows.

$$\begin{aligned} Lacey\ Flow(t) = {} & 0.1648 + 3.2633 * Peggy\ Flow(t) \\ & - 2.0249 * \big(Peggy\ Flow\ (t) - \ Peggy\ Flow\ (t-48)\big) \end{aligned}$$

$$\text{t} = \text{current 15 minute time interval}$$

Equation 3.3 Prediction of Peggy's Brook flows.

$$if\ Peggy\ Flow\ (t) > 0$$

$$\begin{aligned} Peggy\ Flow\ (t) = {} & -0.013841 + 0.21565 * Lacey\ Flow(t) \\ & + 0.492844 * \big(Lacey\ Flow(t) - Lacey\ Flow(t-48)\big) \end{aligned}$$

$$else\ Peggy\ Flow\ (t)\ = 0$$

$$\text{t} = \text{current 15 minute time interval}$$

It was observed in the regression analysis that the nature of the storm (intensity and duration) and the watershed conditions previous to the storm greatly affect the hydrological response. Predictions for a storm have highly correlated errors as it can visibly be discerned in storm events illustrated on Figure 3.5. However the watershed's responses after the rainfall are not as discernable as the decline in flow trend is consistent between storms. The regression also appears to not capture transition at the hydrograph peak, likely a result of instituting a twelve hour lag in the regression. Flows around the peak are first under predicted and then over predicted. The prediction of average storm flows for the four storms in Lacey Mill River were 14.4% below the values measured using the stage discharge curves (Table 3.1). This suggests that storms flows are generally underestimated by this model. Predictions for lower flows and smaller storms are significantly better than major storm event predictions (Table 3.3) Although this is expected, it is of concern as storm events are likely to contribute the vast majority of pollutants and nutrients to the loading estimates. High flow events are of great importance to loading because storm flows in Pockwock can be an order of magnitude larger than average flow conditions.

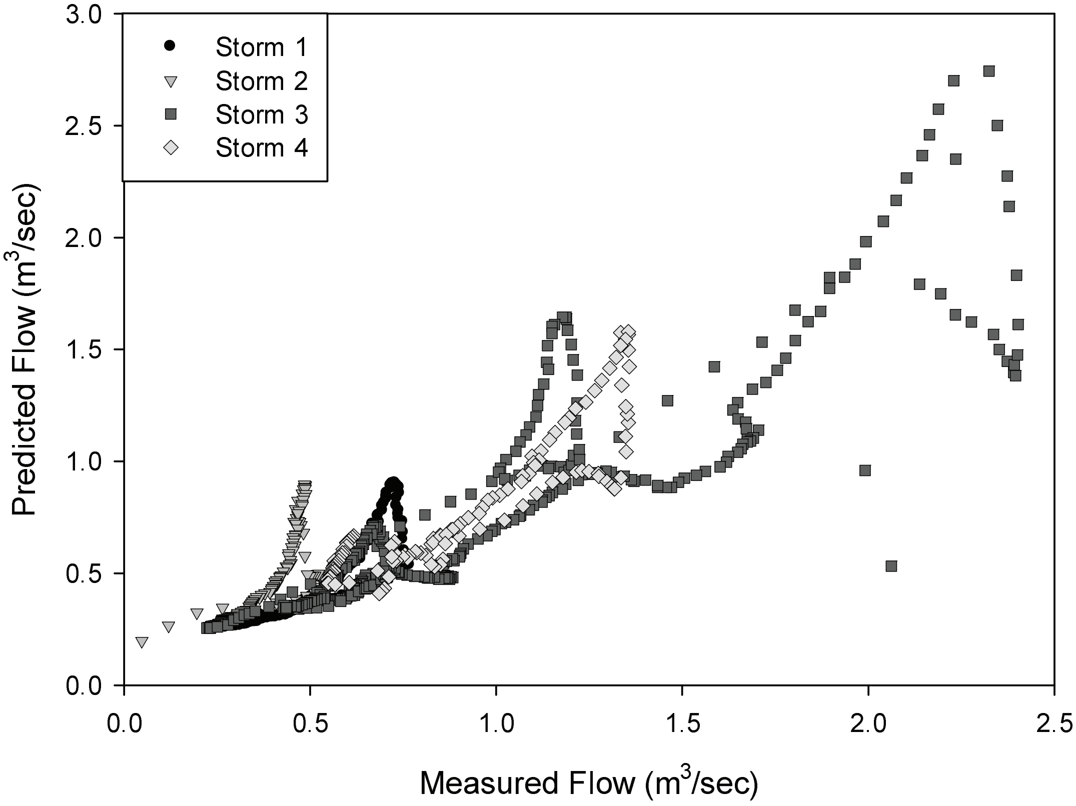

Figure 3.5 Prediction of Lacey Mill River flows for four storm events.

Table 3.2 The accuracy of stream regression model's prediction of storm flow events for Lacey Mill River.

	Average storm flow(m^3/s)		
Storm	Actual	Predicted	% Difference
1	0.538	0.469	-12.8%
2	0.825	0.702	-14.9%
3	0.917	0.798	-13.0%
4	0.717	0.596	-17.0%
			Average Difference: -14.4%

Table 3.3 The accuracy of stream regression model's prediction of stable and low intensity rain events for Lacey Mill River.

	Average storm flow(m^3/s)		
Event	Actual	Predicted	% Difference
1	0.356	0.348	-2.0%
2	0.252	0.236	-6.2%
3	0.236	0.235	-0.5%
4	0.259	0.258	-0.2%
			Average Difference: -2.2%

3.4.2 Estimation Of Concentration

Measurements of TOC and TP, the only two parameters of interest from a loading perspective, were taken on a weekly basis. This presents a challenge as there is a limited ability to estimate concentrations between individual discrete measurements. The general accepted method is linear interpolation, and so linear interpolation between measurements was applied.

Chapter 4 Results And Discussion

4.1 Determination Of Probability Distributions

The decision to use a normal distribution or a lognormal distribution to describe environmental data is crucial, as the choice of distribution will give rise to biases in the statistics performed, which are based on the chosen distribution. It is rare that any specific probability distribution is a perfect description of the data set, and as a result a great deal of personal discretion is exercised when choosing a distribution. Some parameter datasets like turbidity (Figure 4.1) and *E. coli* (figure 4.2) are described better by lognormal distributions, and as a result lognormal distributions were used. However, for other parameters, such as conductivity (Figure 4.3) and pH (Figure 4.4), the normal and lognormal distributions appear to provide similar fits. The lognormal distribution does have the inherent benefits of being non-negative, and the shape of the tails, which are of importance to extreme value distributions, appear to conform to the data better. However, the normal distribution is much simpler to use, as most statistical procedures were initially designed for the normal distribution, and when operating in a different distribution additional considerations must be made. The decision was made to operate in the lognormal distribution for all parameters in order to keep the analysis parsimonious.

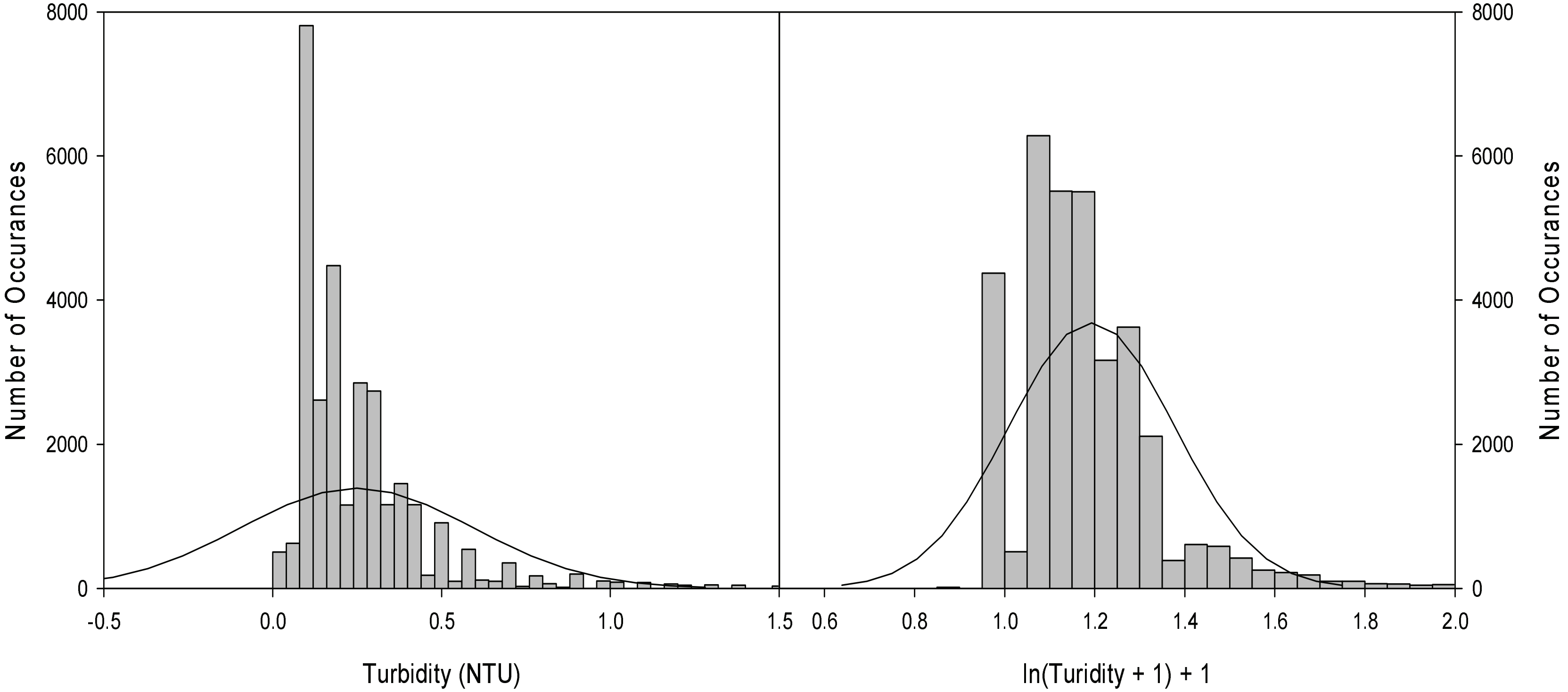

Figure 4.1 Comparison of using the normal distribution (left) and modified lognormal distribution (right) for turbidity using data from Peggy's Brook.

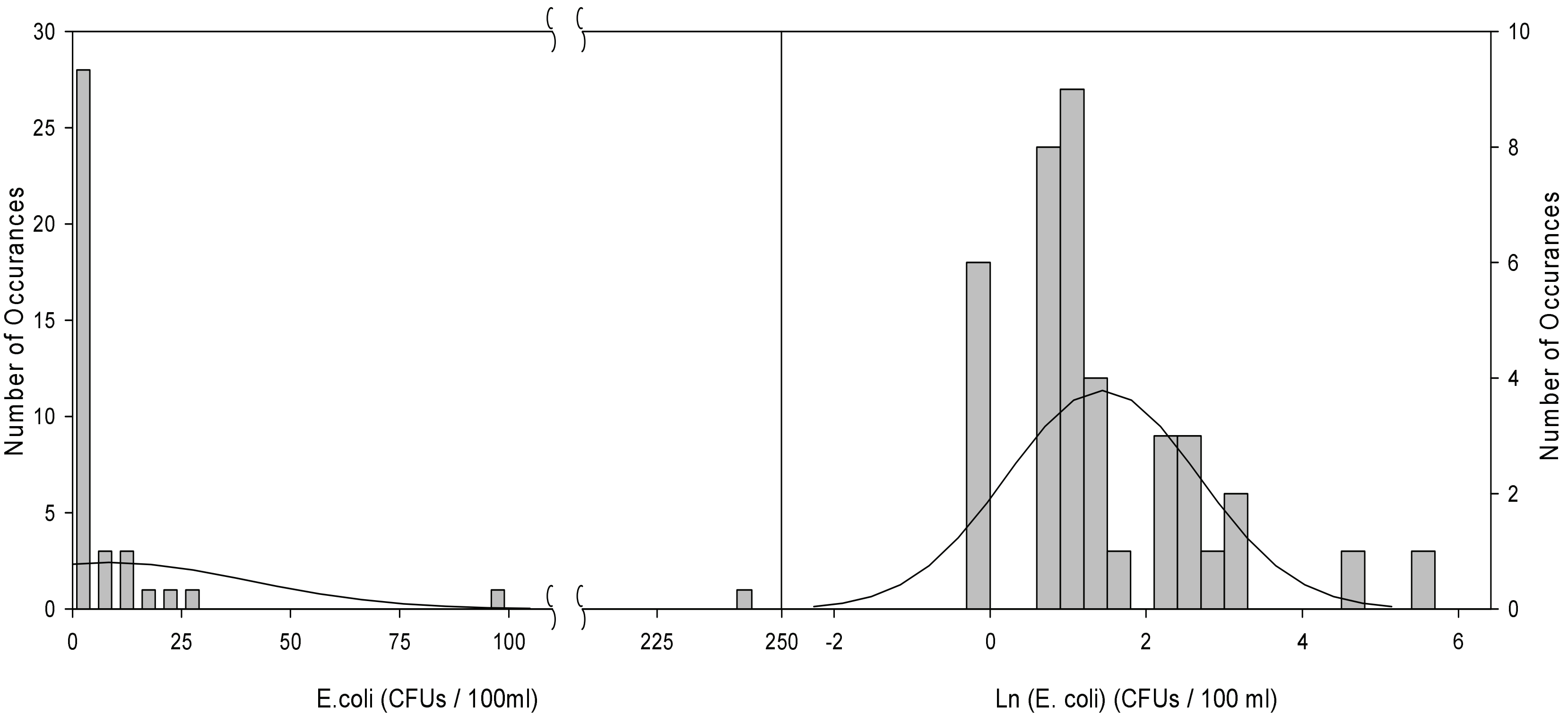

Figure 4.2 Comparison of using the normal distribution (left) and lognormal distribution (right) for *E. coli* using data from Lacey Mill River.

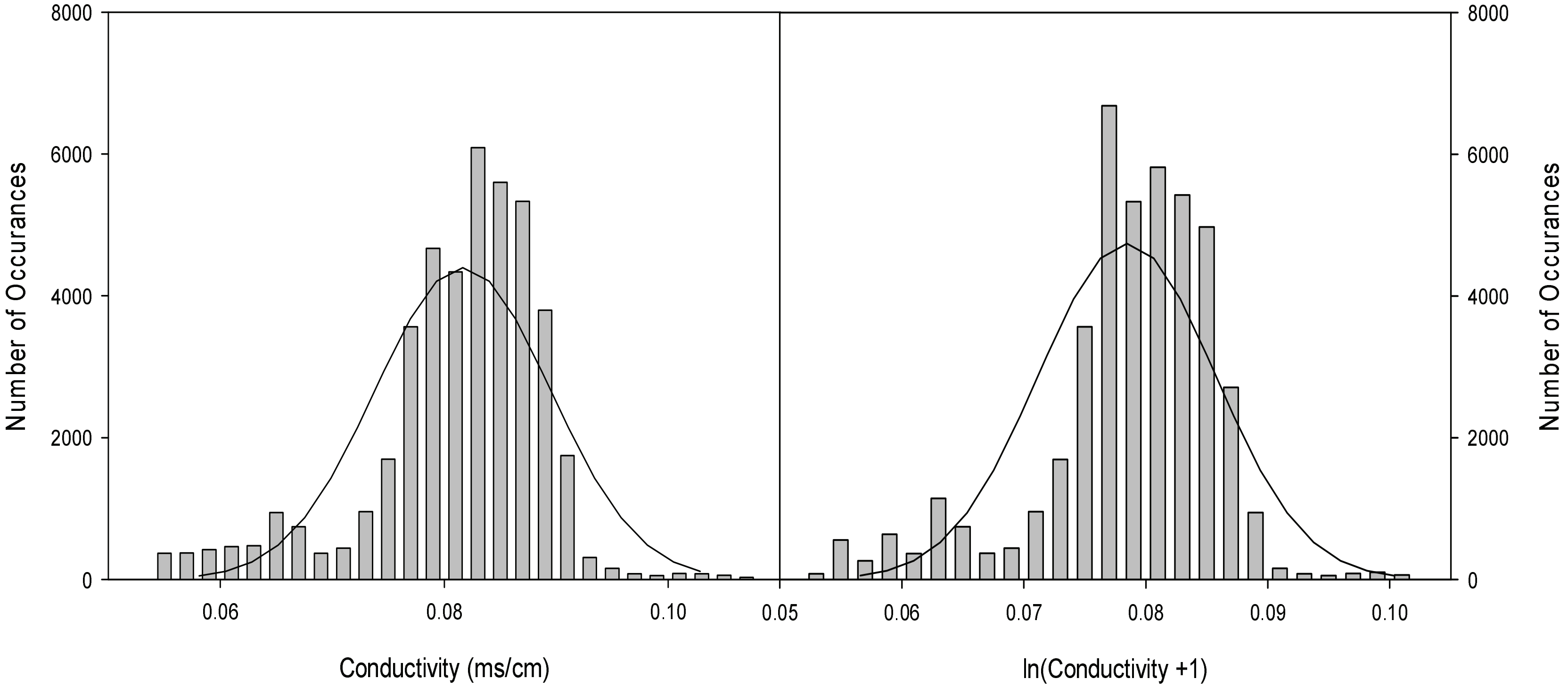

Figure 4.3 Comparison of using the normal distribution (left) and lognormal distribution (right) for conductivity using data from Lacey Mill River.

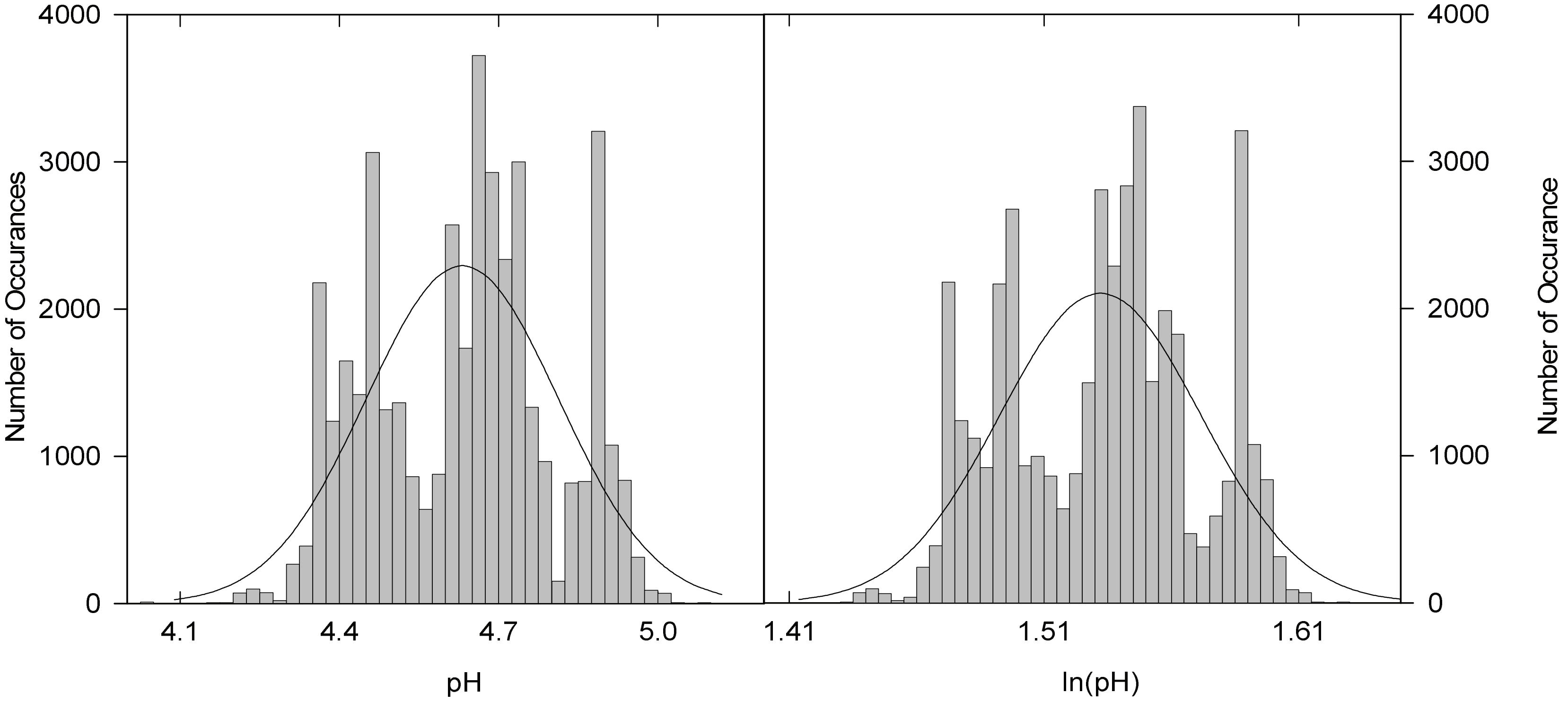

Figure 4.4 Comparison of using the normal distribution (left) and modified lognormal distribution (right) for pH using data from Lacey Mill River.

4.2 Determination Of Autocorrelation Structure

It was determined that water quality data, after seasonal trends were removed, could be analyzed as having a stationary mean on an annual basis, as the assumption of non-stationary means did not improve ARMA models performances. All parameters were analyzed and it was determined that a first order autoregressive process was an acceptable model based on graphical methods, significance testing and the AIC. Autocorrelation graphs for the analyzed parameters displayed an exponential decay, which is characteristic of a positive first order process for short time lags of the autocorrelation structure, but the trend begins to deviate for higher time lags, even dipping into negative autocorrelation (Figures 4.5 and 4.6). The initial lags, which have the most influence, are the most important and the general structure has an acceptable fit for a positive first order autoregressive process. The short time lags were found to be highly significant through t-tests ($p < 0.05$). The partial autocorrelation graphs verified that a first order autoregressive model is an acceptable model as the partial autocorrelations were only significant for the smallest time lag (Figure 4.7). Using the AIC there were very modest improvements in the model by increasing the autoregressive order or adding in a moving average. These modest improvements did not suggest that the first order autoregressive model was poor compared to other models. Though the criteria did not appear to suggest that the first order model was unacceptable, it is possible that the first order autoregressive process is not the most appropriate model, but using a higher order model would greatly complicate the analysis and is outside the scope of this research. Assuming a first order model is at the very least a great improvement over assuming independence. All statistics will be carried out with the assumption that the model of a first order regressive process is acceptable.

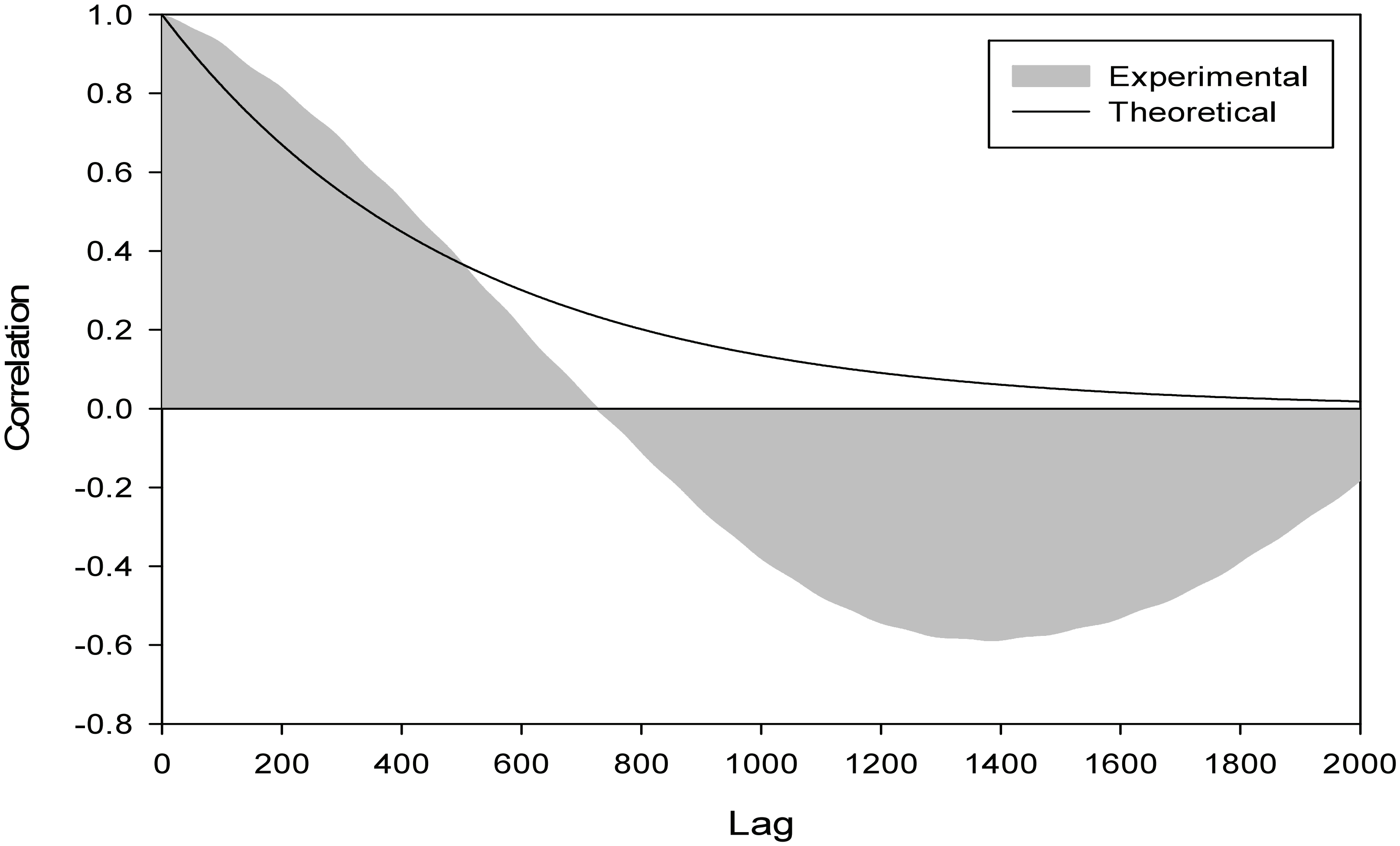

Figure 4.5 Autocorrelation structure of pH in Lacey Mill River with 15 minute time lags (theoretical first order p=0.998).

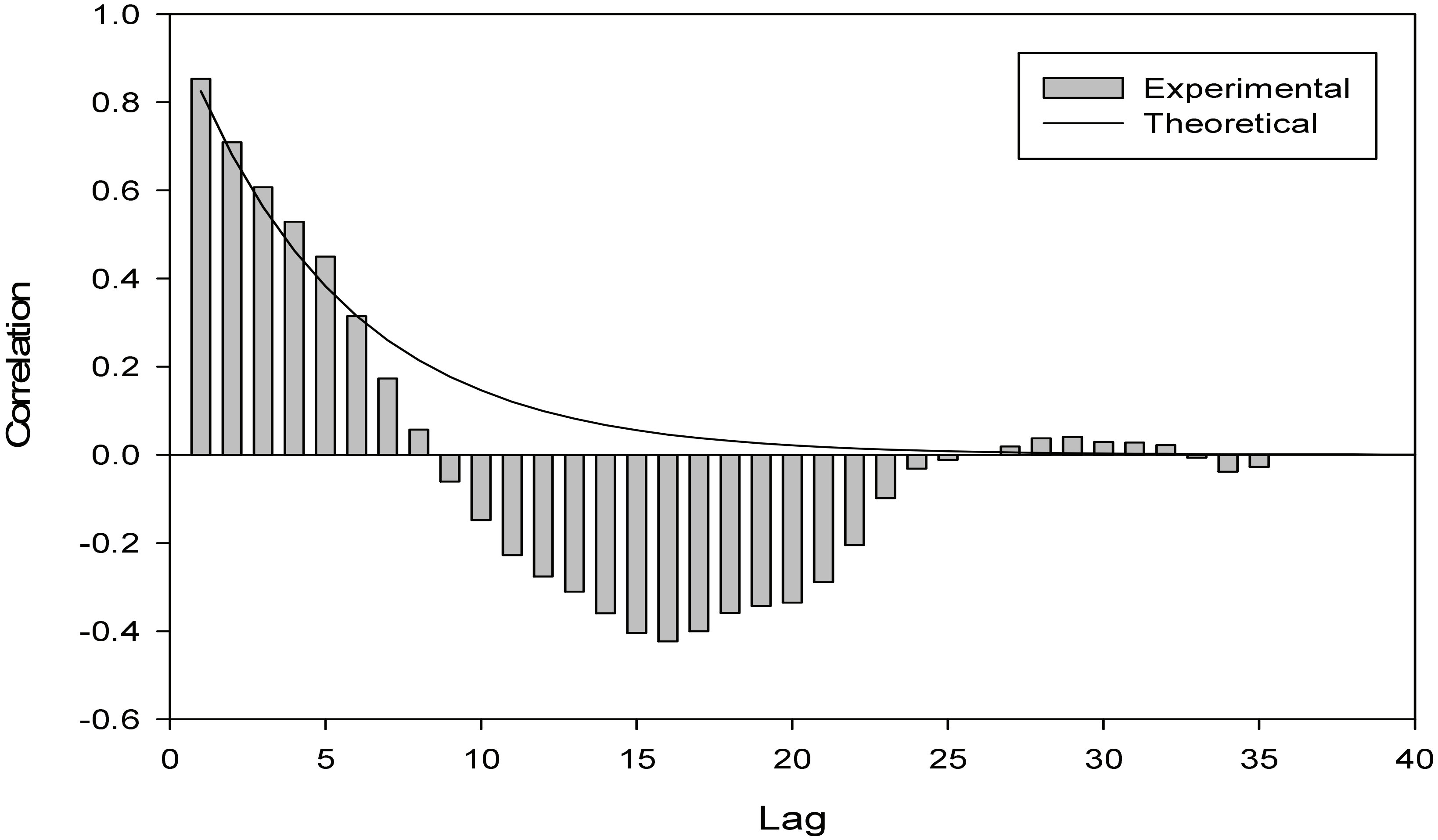

Figure 4.6 Autocorrelation structure of pH in Lacey Mill River with one day time lags (theoretical first order p = 0.825).

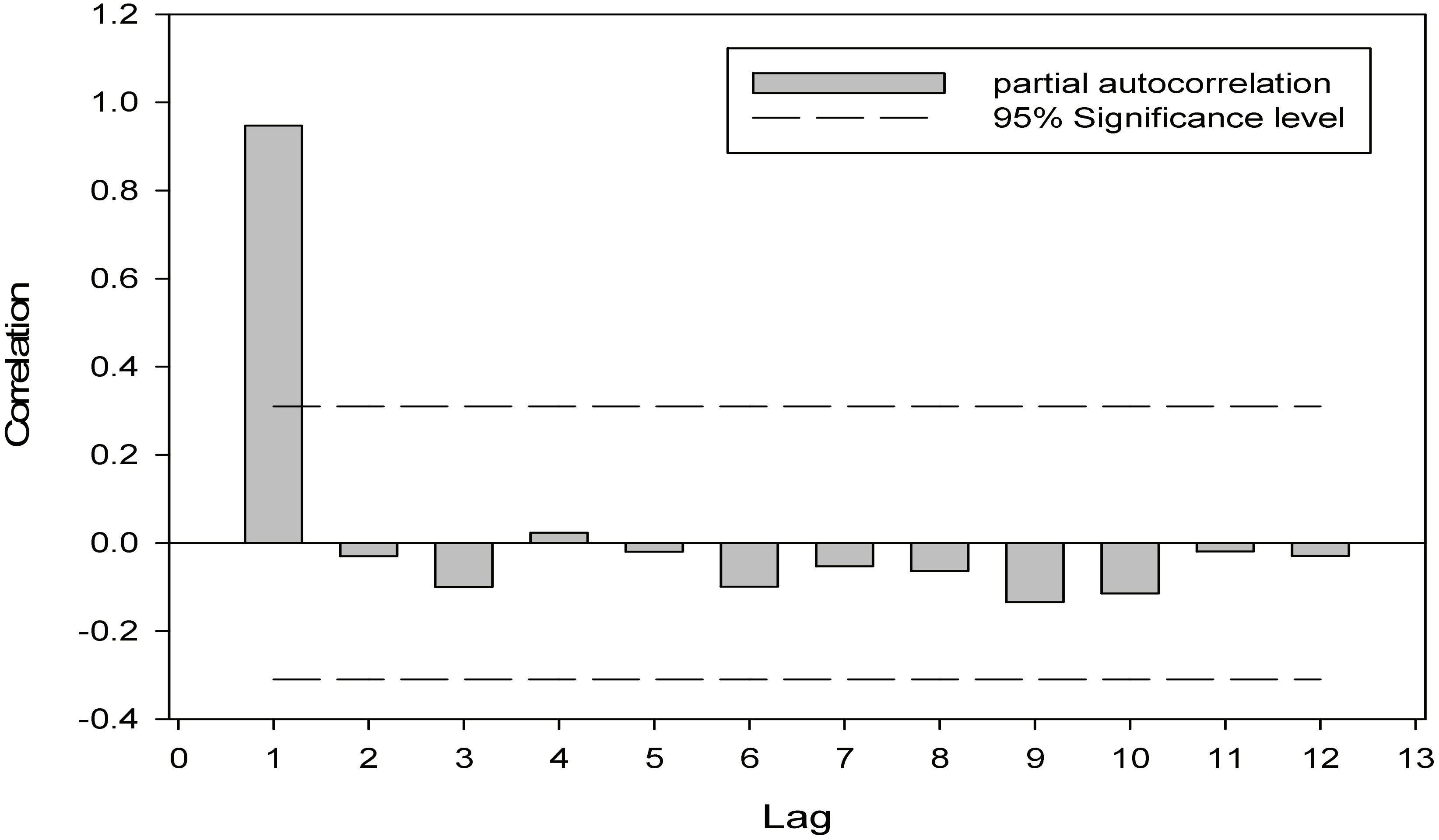

Figure 4.7 Partial autocorrelation structure of pH in Lacey Mill River with one day time lags.

4.3 Nutrient Loading Estimates

Obtaining accurate nutrient loading estimates present challenges due to the impact of storm events on nutrient transport. Historically, linear interpolation is typically used to estimate concentration during time periods in-between discrete sampling intervals. Unless the sampling frequency is high this approach is projected to provide poor estimates of total loading. In this study the top 10% of flow events account for more than a third of C and P loading (Table 4.1 and 4.2). Failure to accurately characterize the concentration during these events can bias loading estimates.

The average TOC concentrations for Lacey Mill River and Peggy's Brook, for the sampling periods that were used to generate loading estimates were 8.6 and 13.6 mg/l respectively. It should be recognized that there is some bias in the estimate of Peggy's Brook loading as the winter is omitted, however these fall within the range of TOC concentrations for temperate forested watersheds presented by Meybeck (range of 2 to 30 mg/l) (1982). The Calculated annual loading for Lacey Mill Brook and Peggy's Brook were 8.8 and 16.8 kg km^{-2} yr^{-1} TOC with the Peggy's Brook estimate missing January through March 2010. These are on the same order of magnitude as a study on boreal forest by Cooke and Prepas (1998), but Peggy's Brook does appear to have a relatively high loading. The study by Cooke and Prepas showed a significant difference between years because of watershed conditions (5 and 22 kg km^{-2} yr^{-1} TOC), and as a result a one year loading study cannot accurately quantify annual loading estimates.

Phosphorous loading rates of 51.1 and 160.4 g hectare^{-1} yr^{-1} were respectively calculated for Lacey Mill River and Peggy's Brook for the period of June 2009 until July 2010. This is a large range, but both calculations are within the range suggested for P export coefficients by Brylinski (2004) for forested watersheds in Nova Scotia (50 – 600 g hectare^{-1} yr^{-1}). This range of suggested P export coefficients is a result of differences in geology and watershed conditions which currently are not extensively outlined in the guide. It is suggested in Brylinski's P prediction manual that unmanaged forests generally have the lowest P export; however, this study did not confirm that to be true. Pockwock may be a unique case where the unmanaged forest is exporting more phosphorous, or the one year annual loading estimate may not accurately portray the average P export regimes of the stream.

Table 4.1 Total organic carbon loading on a monthly basis for Lacey Mill River. Annual loading estimate is for time period of July, 2009 – July, 2010. Loading was computed using weekly, bi-weekly, and monthly sampling frequencies for concentrations.

Month	Weekly (kg)	Bi-weekly (kg)	Monthly (kg)	Contribution of top 10% of flows
Jun-09	2200	2200	2200	26%
Jul-09	2800	2800	2700	23%
Aug-09	4300	4200	3700	42%
Sep-09	2200	2300	1500	47%
Oct-09	10400	12800	9700	21%
Nov-09	6700	7100	7000	14%
Dec-09	15400	13700	16300	20%
Jan-10	6500	6400	7400	25%
Feb-10	4200	3900	4200	48%
Mar-10	8000	8100	8900	26%
Apr-10	3600	3400	3500	23%
May-10	1300	700	700	14%
Jun-10	2100	2100	2300	29%
Jul-10	45400	22400	14500	67%
Aug-10	21000	36800	3400	47%
Annual Sum (kg)	***67500***	***67500***	***67900***	***34%***

Table 4.2 Total phosphorous loading on a monthly basis for Lacey Mill River. Annual loading estimate is for time period of July, 2009 – July, 2010. Loading was computed using weekly, bi-weekly, and monthly sampling frequencies for concentrations.

Month	Weekly (kg)	Bi-weekly (kg)	Monthly (kg)	Contribution of top 10% of flows
Jun-09	9.2	13.2	16.4	21%
Jul-09	7.5	6.5	10.2	12%
Aug-09	20.5	24.2	32.6	27%
Sep-09	2.1	2.4	2.4	40%
Oct-09	21.1	10.4	8.5	11%
Nov-09	5.7	7.5	4.6	12%
Dec-09	15.0	16.9	9.5	34%
Jan-10	19.6	7.7	6.5	49%
Feb-10	8.5	16.8	6.6	67%
Mar-10	39.1	57.9	27.9	42%
Apr-10	4.2	6.8	9.3	24%
May-10	3.5	0.7	0.7	22%
Jun-10	2.1	2.2	1.9	29%
Jul-10	5.7	5.5	4.2	53%
Aug-10	2.0	2.0	1.8	27%
Annual Sum (kg)	**149.0**	**160.2**	**120.7**	**43%**

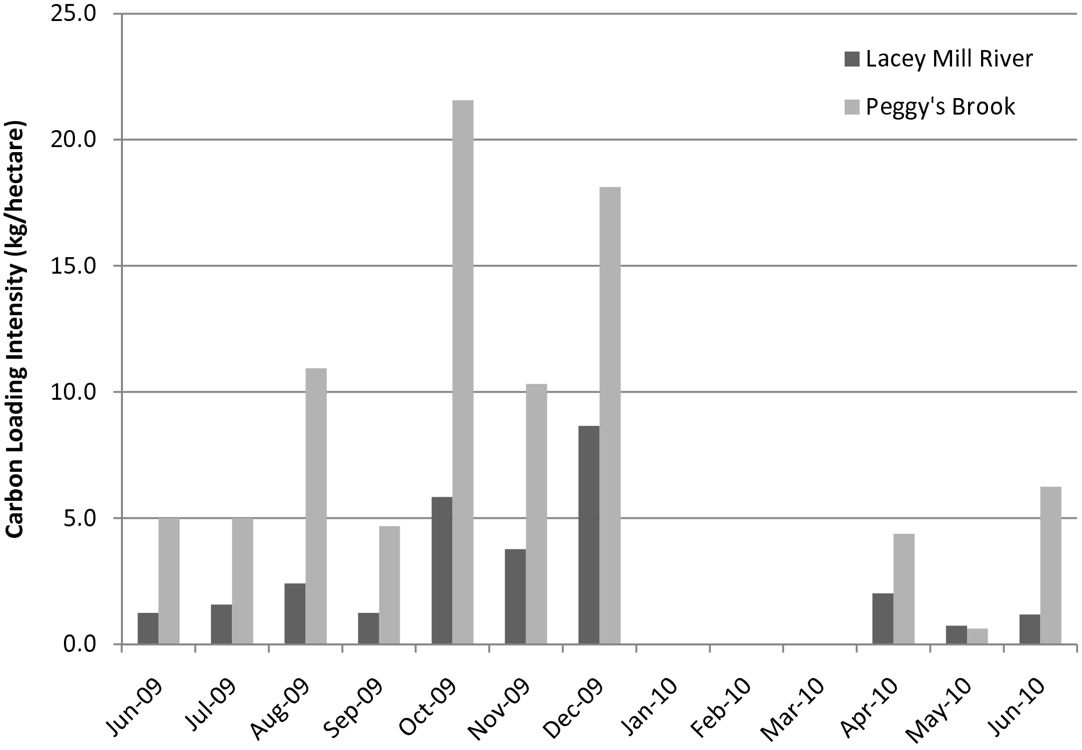

Figure 4.8 Comparison of TOC loading intensity estimates for Lacey Mill River and Peggy's Brook (June 2009 - June 2010).

There are seasonal trends in flows (Figure 4.8), TOC (Figure 4.10), and TP concentrations which result in periods that will predictably contribute a majority of the loading on an annual basis. Flows are lowest during the summer with a small number of rainfall events contributing to most of the rivers total seasonal flow (Figure 4.8). Concentrations of TOC and TP are lowest in the spring from May-July, during the growing season, as nutrients are readily assimilated by growing vegetation (Figure 4.8). This results in relatively low loading rates during the late spring and the summer. In the fall however, large flushes of nutrients are observed due to the increased nutrient availability caused by forest litter fall and decomposition (Table 3.2 and 4.2). The fall is also generally wetter, and larger average flows are experienced leading to a disproportionate amount of loading to occur in the fall (48% of Lacey Mill River’s estimated carbon loading).

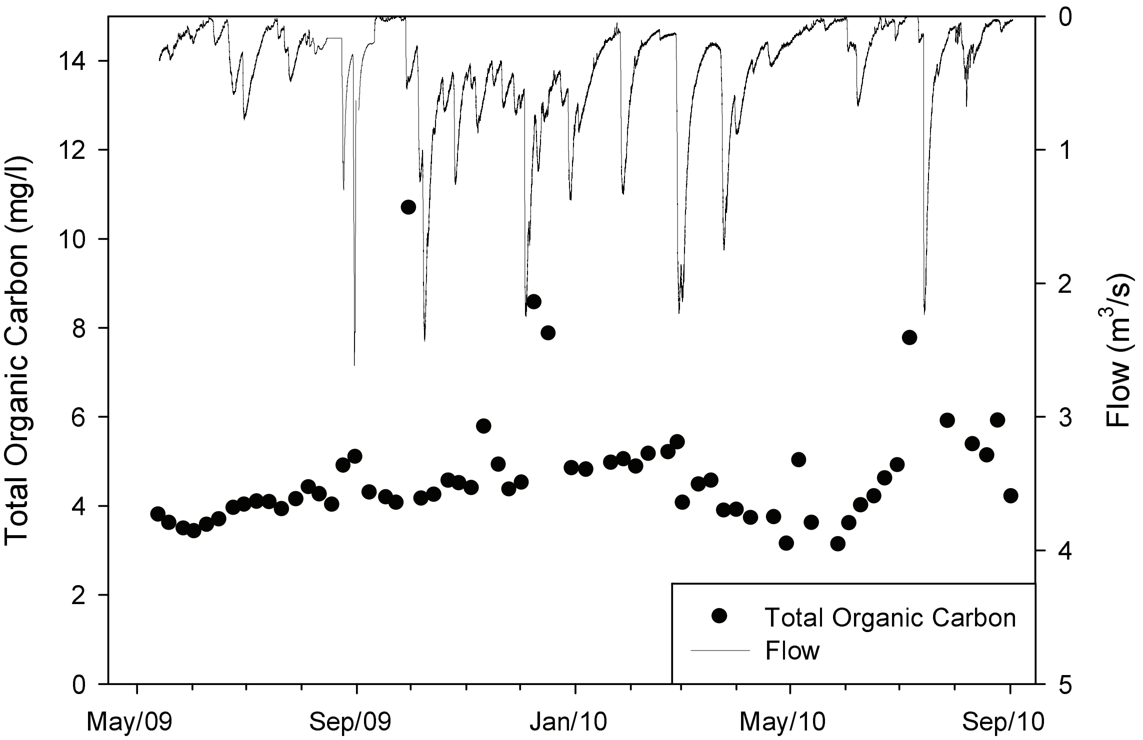

Figure 4.9 Flows and TOC concentrations during the monitoring period for Lacey Mill River.

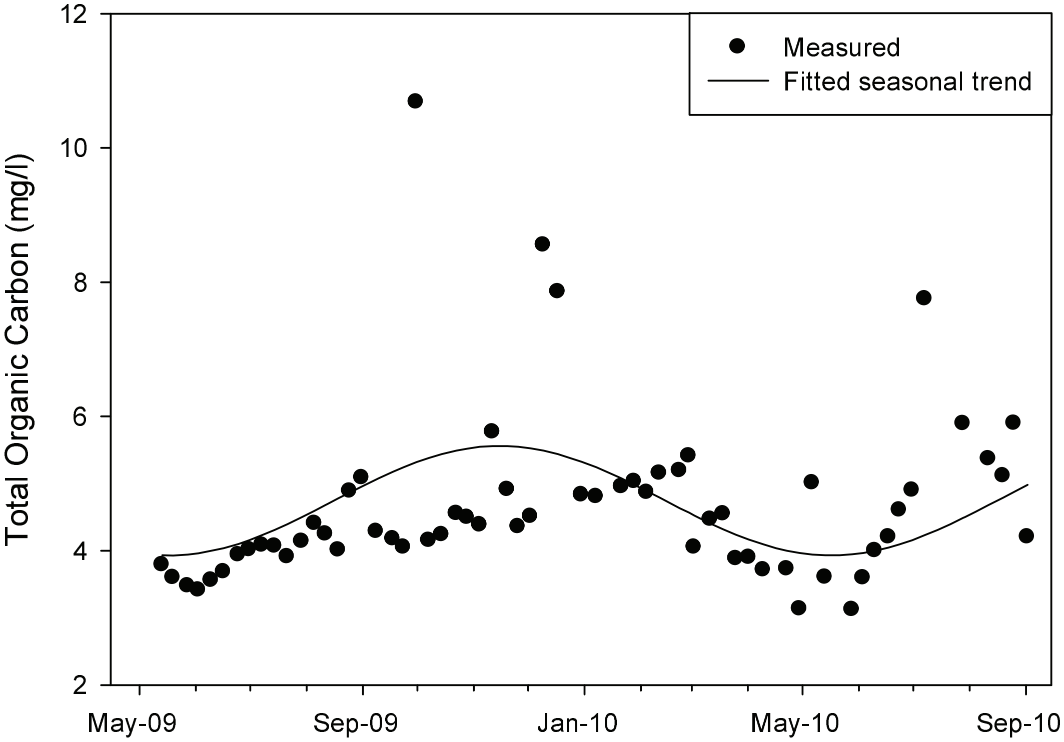

Figure 4.10 Seasonal trends in total organic carbon concentrations in Lacey Mill River.

Nutrient loading estimates can be easily biased if linear interpolation is used. This is highly visible for carbon loading in Lacey Mill River in July and August of 2010, where estimated carbon loadings were computed to be 45,400 and 21,000 kg respectively (Table 4.1). Contributing to these large loading values were measured concentrations of 80 mg/l and 169 mg/l respectively. These samples were collected at the start of two large storms, which possessed peak flows of 2.20 m^3/s and 0.61 m^3/s. These are very significant storms when compared to the monthly average flows of 0.35 m^3/s and 0.15 m^3/s for July and August, respectively. These concentrations are unlikely to be representative of the entire storm and are expected to be high, when compared to the storm average, because pollutant concentrations are generally higher during the rising limb of storm hydrographs (Inamdar et al., 2004). Tables 4.1 and 4.2 depict that there is little difference in loading estimates as a function of sampling frequency. This was surprising as it would be expected that a higher sampling frequency would provide better loading estimates. However, the concentrations of many water quality parameters that are of interest from a loading perspective are transient and are susceptible to fluctuation during storm events. This means that unless the sampling interval is shorter than the duration of intermittent events potentially large bias will be present in loading estimates that use linear interpolation for concentrations. Interval sampling will not be able to capture the transient states that occur between samples, and consequently event sampling (storm chasing) is necessary to improve estimates. The minimum sampling interval for TP and TOC in this study was weekly, and it is unlikely that a weekly sampling frequency is short enough to adequately capture storm event loading phenomena.

A common ideology persists that a healthy watershed, from the perspective of providing sustainable potable water, would have limited human intervention, and that forestry or other development results in adverse affects to water quality. This may be true in certain situation, but appropriate forestry management could be a beneficial practice to improve and maintain water quality from a drinking water perspective. Tables 4.3 and 4.4 suggest there is much greater nutrient loading intensity in the Peggy's Brook subwatershed as compared to the Lacey Mill River subwatershed. This is contrary to expected results as Lacey Mill River has development in the upper subwatershed and is therefore considered to be more impacted. The consistent result may be explained by Peggy Brook's subwatershed being comprised of old growth forest, which may have limited ability to assimilate nutrients, and this lack of assimilation capacity possibly has resulted in Peggy's Brook subwatershed becoming an exporter of nutrients. If this is the case, forestry management could potentially be utilized to reinvigorate the forest in an attempt to improve nutrient assimilation with the goal of decreasing nutrient loading into Pockwock Lake.

Table 4.3 Carbon loading intensity (kg/hectare) comparison of Lacey Mill River and Peggy's Brook subwatersheds.

Month	Loading Intensity (kg/hectare)	
	Lacey Mill River	Peggy's Brook
Jun-09	1.2	5.0
Jul-09	1.6	5.0
Aug-09	2.4	10.9
Sep-09	1.2	4.7
Oct-09	5.8	21.6
Nov-09	3.8	10.3
Dec-09	8.7	18.1
Apr-10	2.0	4.4
May-10	0.7	0.6
Jun-10	1.2	6.3
Jul-10	25.5	24.7
Aug-10	11.8	2.5

Table 4.4 Total phosphorus loading intensity (g/hectare) comparison of Lacey Mill RIver and Peggy's Brook subwatersheds.

Month	Loading Intensity (g/hectare)	
	Lacey Mill River	Peggy's Brook
Jun-09	5.17	7.61
Jul-09	4.23	4.64
Aug-09	11.53	9.00
Sep-09	1.15	7.29
Oct-09	11.87	45.93
Nov-09	3.20	51.00
Dec-09	8.44	25.85
Apr-10	2.34	4.27
May-10	1.97	0.96
Jun-10	1.20	3.88
Jul-10	3.19	7.40

4.3.1 Disinfectant Byproduct Formation Potential

The general concern with loading of organic carbon into source water is the potential for the interaction of organic compounds and chlorine compounds to form toxic and or carcinogenic compounds during the drinking water treatment process. Peggy’s Brook subwatershed has an elevated intensity in carbon loading when compared to Lacey Mill River subwatershed, but it cannot be precluded that Peggy’s Brook will have a higher intensity of DBPs. A small study was performed on October 28, 2010 to examine the potential for formation of DBPs from water samples collected in each stream. This study used one sample from each site, where the samples were analyzed, after reaction with chlorine, for Trihalomethanes (THMs) and Haloacetic Acids (HAAs). Table 4.5 is a summary of the results from this study. Peggy’s Brook had a higher concentration of TOC and an increase in concentration of three types of DBPs: chloroform, chloroacetic, and dichloroaecetic acid. This small study suggests that elevated levels of TOC may have a strong relationship to elevated levels of DPBs from treated water being produced from the Pockwock Watershed. However, due to the small sample size caution should be exercised when using these findings.

Table 4.5 Disinfectant byproduct potential for the associated Total Organic Carbon samples collected on October 28, 2010.

	Lacey Mill River	Peggy's Brook
Total Organic Carbon (mg/l)	3.78	16.4
DBPs (ug/l)		
Chloroform	251.5	1580.3
Dichlorobromomethane	16.2	14.1
Dibromochloromethane	9.1	0.0
Chloroacetic Acid	1.8	20.8
Dichloroacetic Acid	132.2	496.4
Trichloroacetic Acid	125.1	160.3
Bromochloroacetic Acid	2.6	1.6
Bromodichloroacetic Acid	6.2	5.61

4.4 Dissolved Oxygen

With respect to water quality monitoring, the key characteristics of DO are the minimum and maximum concentrations, as they are indicators of ecosystem health. The minimum and maximum are dependent on temperature and sunlight which can be highly variable on daily to seasonal time scales. Because of the daily and annual variability in temperature and light, there are diurnal and seasonal trends. Super saturation of oxygen within the water can result from two issues: (i) the result of a dam that has not been properly designed, or (ii) excessive algal growth. The development of anoxic conditions in the water column is often a result of biological oxygen demand caused by excessive respiration or decomposition occurring in the stream. With respect to the assessment of aquatic health, it is of most interest to be able to characterize DO minima.

If the goal of a water quality monitoring program is to characterize the DO minima the use of a discrete sampling program would be inadequate. Figure 4.11 and 4.11 illustrate that the time of day when the DO minima occurs is highly variable. Peggy's brook commonly experiences DO minima in the middle of the night; however Lacey Mill River's occurs in the early morning. This suggests that an in-situ monitoring probe would need to be installed to capture the minima, as most monitoring programs do not incorporate night sampling, and because the timing of the minima is site dependent. Diurnal oxygen trends are controlled by the biological and hydrological characteristics of the stream or river, and as a result two streams can be very different (Figure 4.13). Rivers which are productive and quiescent will experience very different diurnal oxygen trends than the alternative. Lacey Mill River and Peggy's Brook show the contrast very well as can be seen in Figure 4.13. Lacey Mill River characteristically has DO minimums that occur shortly before sunrise. This can be attributed to the river's typical low water velocity and high productivity, as they contribute to continual decline of DO concentration until sunrise, because the respiration rate exceeds the reaeration rate. However, Peggy's Brook's respiration rate and reaeration rate reach a point of equilibrium during the night. This can be discerned from the relatively constant oxygen saturation in Peggy's Brook throughout the night (Figure 4.13). The factors causing the Peggy Brook DO trend are contrary to that of the Lacey Mill River trend. Peggy's Brook's stream morphology results in more turbulent flows, and the stream is less productive because it is sunlight limited due to a narrow channel width (generally less than 2 m) which results in greater shading of the stream.

There is a discernable seasonal trend (figure 4.13) in DO concentration and saturation, which is influenced by changes in temperature and biological productivity. Over the winter DO concentrations are much higher and less variable. This is a result of lower temperatures and lack of biological production. Due to the consistent temperature of the water, within a couple degrees of freezing, DO content of the water does not change significantly from saturation. The timing of DO measurements are less important during winter months as the daily range of DO is small and independent of time of day. During the spring and summer photosynthesis and respiration have a major impact on DO in streams as these biological processes respectively greatly increase and decrease the oxygen concentration in the water. Summer provides the lowest minima, highest maxima and the largest daily range of oxygen saturation and concentration (Figure 4.13, Table 4.6, and Table 4.7). It is most likely that concerns related to DO in the watershed will occur between May and October.

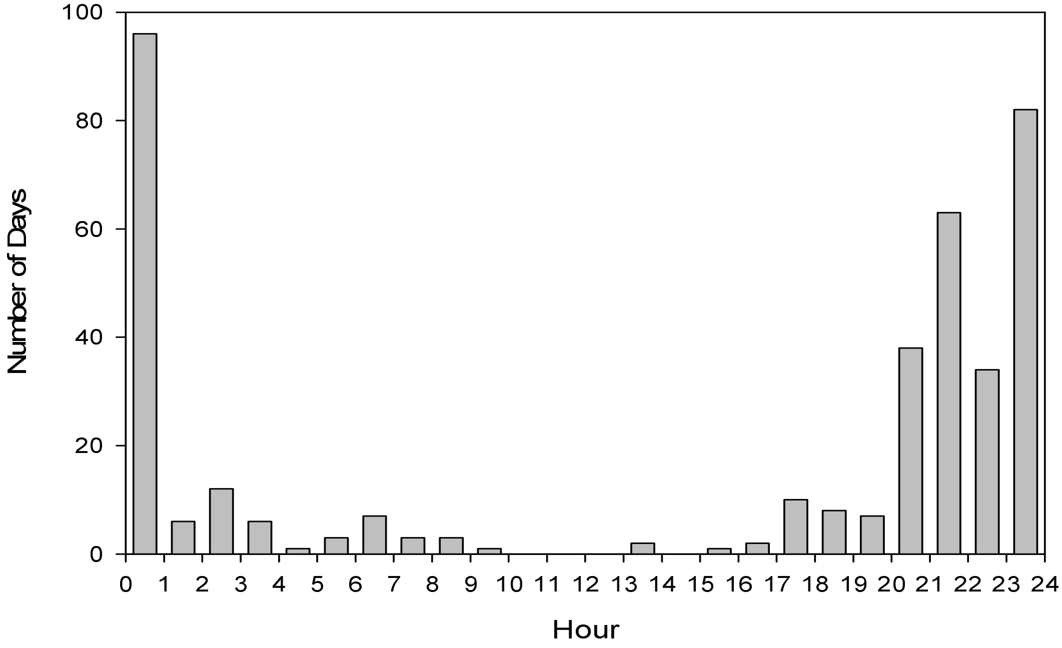

Figure 4.11 Depiction of when oxygen minima occurred (hour of the day) during monitoring period in Peggy's Brook.

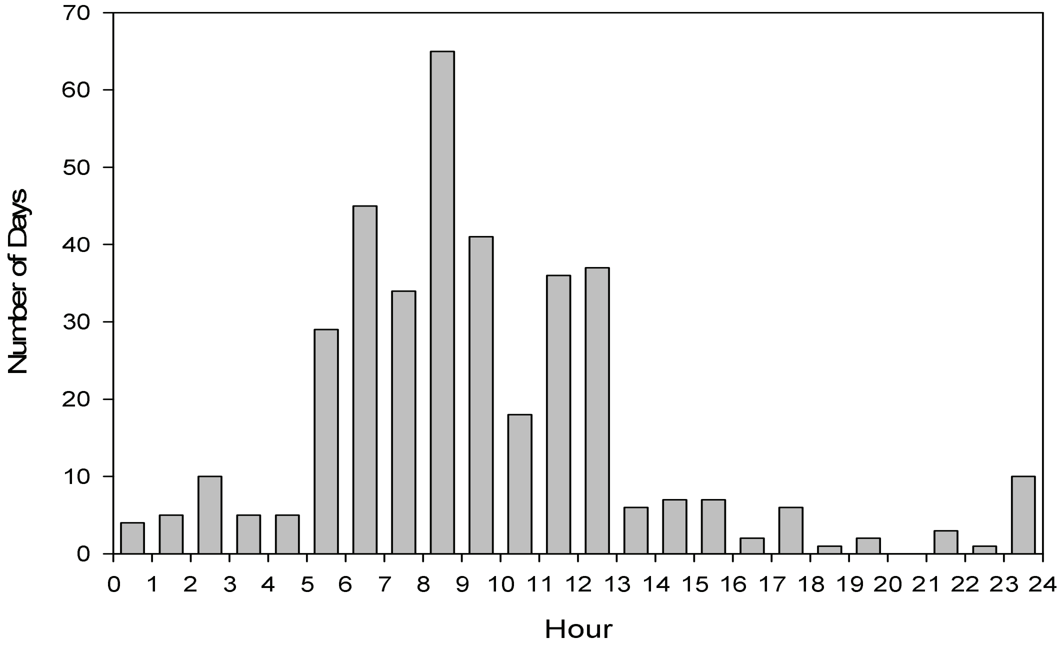

Figure 4.12 Depiction of when oxygen minima (hour of the day) occurred during monitoring period in Lacey Mill River.

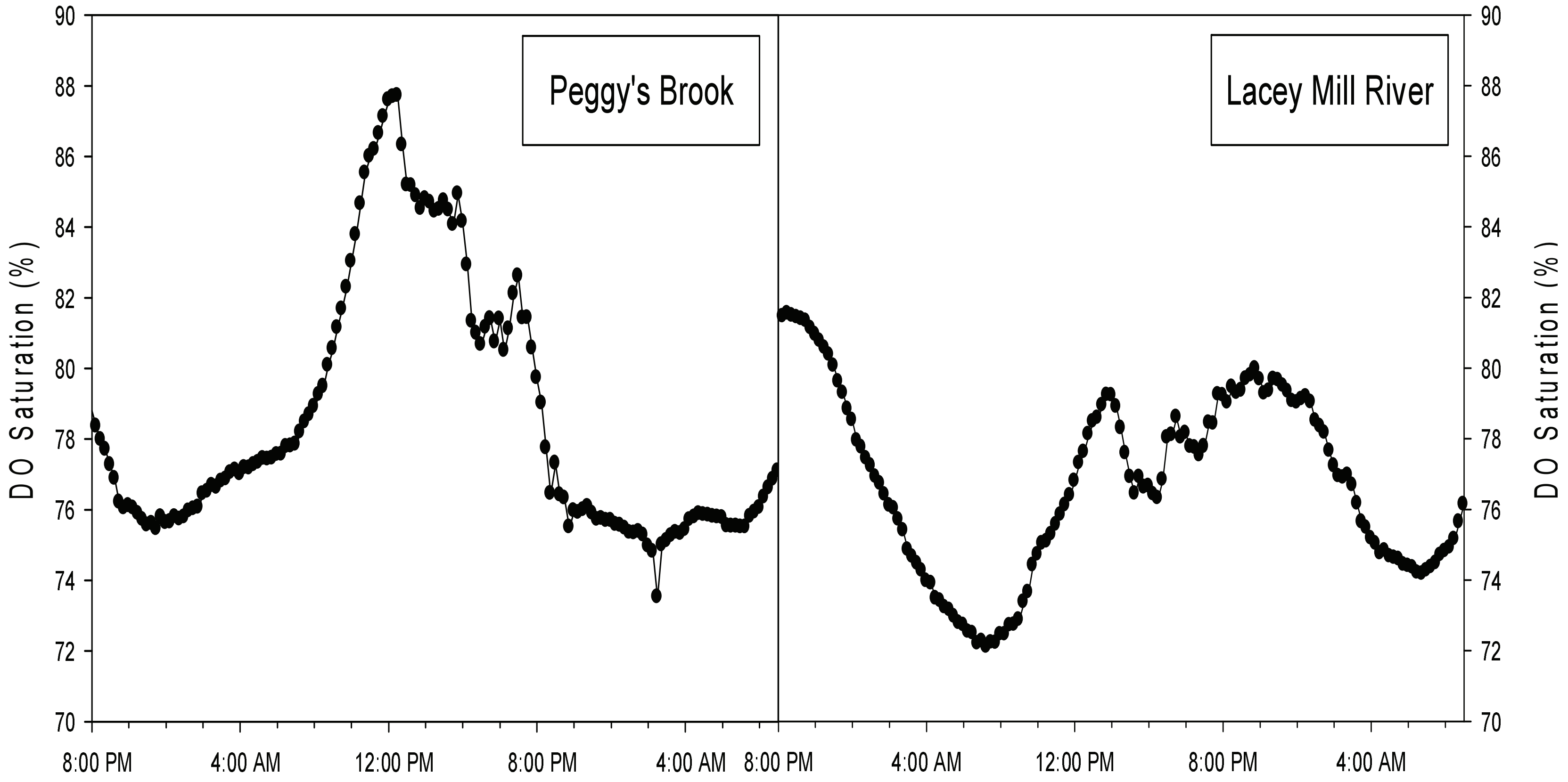

Figure 4.13 Dissolved oxygen saturation levels diurnal trend (20/08/2009 8:00 pm - 22/09/2009 10:00 am

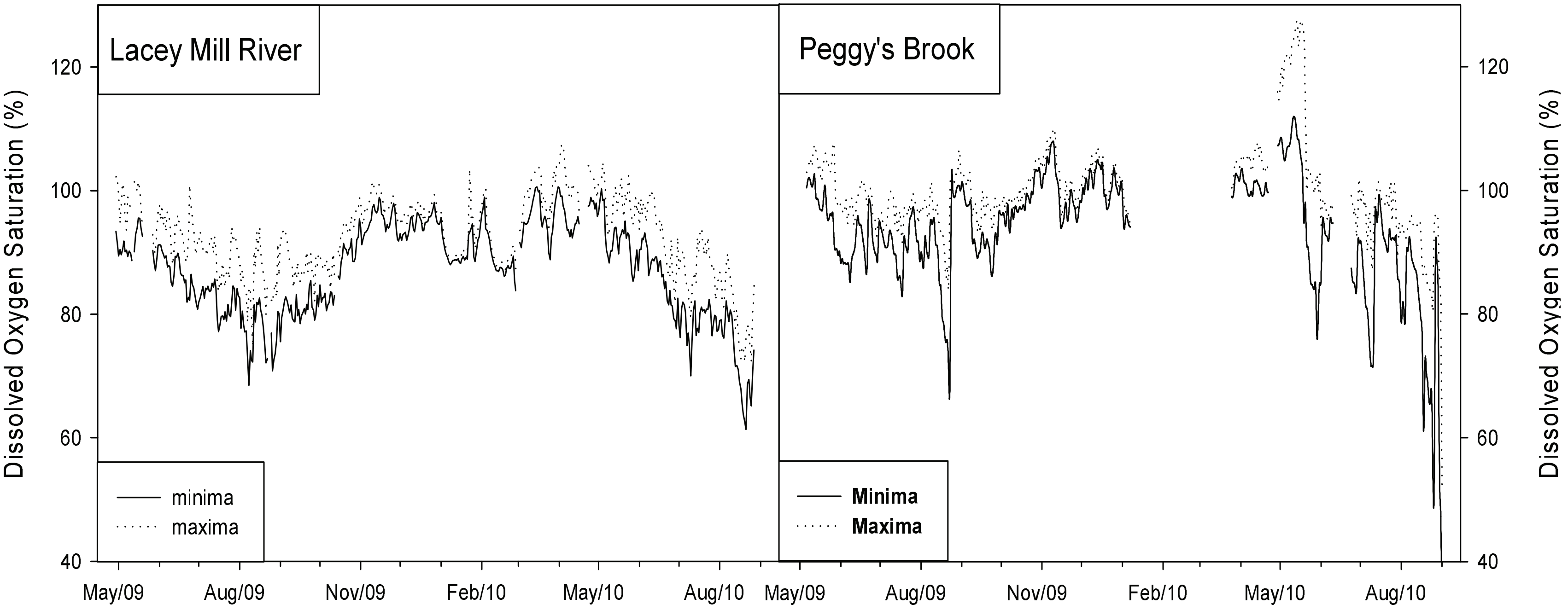

Figure 4.14 Daily dissolved oxygen saturation maxima and minima in Lacey Mill River and Peggy's Brook.

Table 4.6 Summary statistics for monthly dissolved oxygen in Peggy's Brook.

Month	DO saturation (%)				DO concentration (mg/l)			
	Avg	Min	Max	Max Range	Avg	Min	Max	Max Range
May-09	100.01	89.47	107.39	15.41	11.46	9.82	12.82	1.96
Jun-09	93.97	85.12	101.35	12.30	9.89	8.63	10.95	1.27
Jul-09	93.37	83.75	99.82	10.67	9.32	8.33	10.32	1.09
Aug-09	92.49	69.22	106.63	30.62	8.89	6.35	10.95	2.98
Sep-09	94.89	86.13	102.76	9.48	10.07	8.94	11.23	0.98
Oct-09	99.13	93.64	105.28	4.68	11.64	9.87	13.56	1.16
Nov-09	101.29	94.17	109.61	5.57	12.72	11.17	14.44	1.20
Dec-09	102.04	97.09	107.00	3.62	14.26	12.14	15.44	1.02
Jan-10	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Feb-10	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Mar-10	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Apr-10	102.85	99.05	115.90	8.68	12.83	11.79	14.19	1.05
May-10	105.69	76.02	127.49	30.40	12.14	8.28	14.99	3.40
Jun-10	93.07	84.07	101.87	12.80	9.93	8.63	11.18	1.14
Jul-10	92.02	71.57	101.21	21.28	8.90	6.67	10.18	2.20
Aug-10	81.68	47.25	95.93	41.69	8.17	4.52	9.42	3.90

Table 4.7 Summary statistics for monthly dissolved oxygen in Lacey Mill River.

Month	DO saturation (%)				DO concentration (mg/l)			
	Avg	Min	max	Max Range	Avg	Min	max	Max Range
May-09	93.67	87.02	101.50	11.33	9.94	8.92	10.72	0.73
Jun-09	89.76	80.77	100.92	18.73	8.53	7.21	9.59	1.09
Jul-09	85.62	77.16	94.09	11.95	7.67	6.94	8.44	0.79
Aug-09	81.15	68.47	94.11	15.88	6.90	5.92	7.65	1.23
Sep-09	84.47	75.55	93.90	16.88	8.08	7.04	8.85	1.30
Oct-09	89.22	80.56	98.79	8.36	9.80	8.23	11.32	0.74
Nov-09	96.78	91.21	101.53	4.68	12.05	10.89	12.68	0.68
Dec-09	95.48	91.87	99.34	3.49	13.10	11.25	14.12	0.65
Jan-10	90.99	88.00	103.50	10.26	13.15	12.64	14.92	1.49
Feb-10	90.28	83.80	100.29	6.49	13.07	12.04	14.56	0.99
Mar-10	97.66	88.81	104.14	9.00	12.91	11.72	14.00	0.85
Apr-10	98.84	92.34	107.54	9.76	11.33	10.21	13.40	0.53
May-10	95.18	85.30	104.58	11.21	9.81	8.09	11.11	0.84
Jun-10	89.76	77.63	100.03	13.06	8.60	7.26	9.68	0.75
Jul-10	83.09	70.02	93.93	14.73	7.16	5.89	7.81	0.91
Aug-10	76.33	61.32	89.82	14.99	6.71	5.34	7.64	1.09

The results of this study have illustrated that it would be difficult to characterize DO minima and maxima without continuous monitoring, as the timing of these incidents vary on a daily basis depending on watershed conditions. However, it may not be necessary to have continual deployment of the sensor to meet monitoring objectives that require DO data. If a continuous program for DO is established, monitoring equipment may be removed during the winter in an effort to prolong the life of the equipment, as DO may not provide insight into the water quality or ecosystem health throughout the winter. In monitoring a watershed for DO, efforts should generally be placed on the growing season and continue through the fall months prior to ice cover.

4.5 Turbidity

Turbidity is a water quality parameter that is vary transient, varying by orders of magnitude both within, and between storm flow events. (

Figure 4.15). Turbidity data is both difficult to collect and interpret because the durations of extreme events, when we typically see increases in turbidity, are short. If the sampling frequency is low, or the sensor is not appropriately placed, turbidity events will be missed. As well, even if the event is captured it can be difficult to assess the validity of the data as turbidity sensors can be influenced by debris, and sporadic measurements can be outliers.

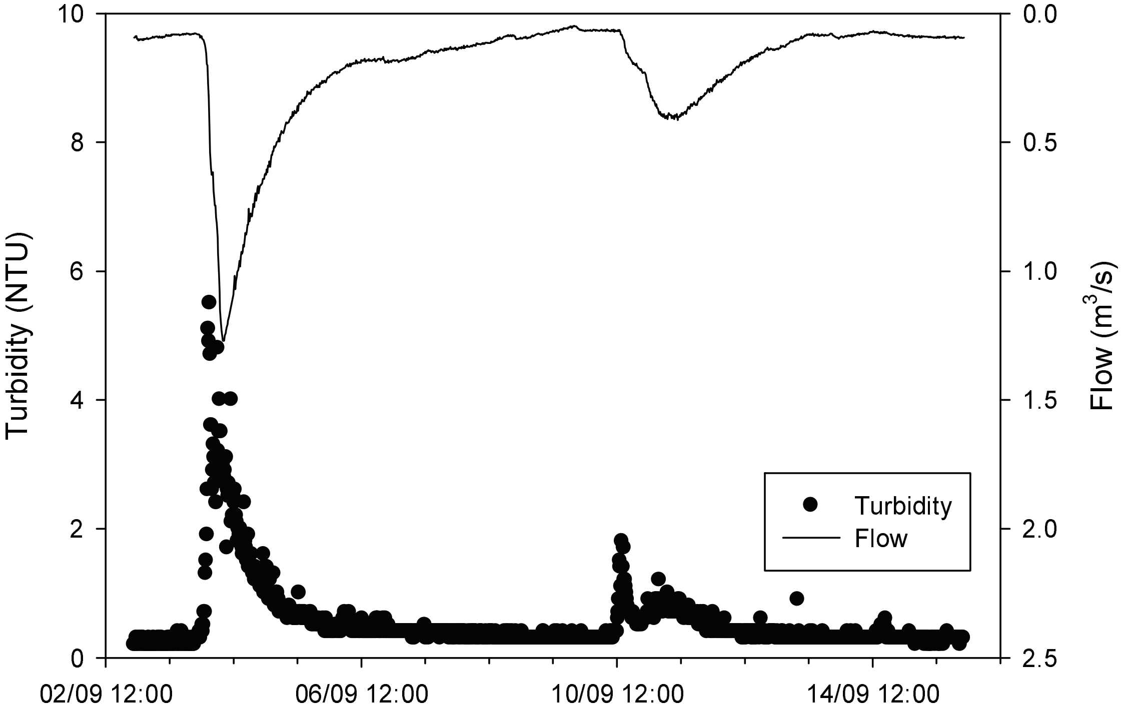

Figure 4.15 Variation of turbidity during a storm flow event in Peggy's Brook.

There is no significant seasonal or diurnal trend in turbidity, but measurements are strongly autocorrelated (ρ = 0.98 for fifteen minute intervals). Turbidity, in Peggy’s Brook and in Lacey Mill River, possesses average baseflow levels of 0.21 and 0.46 NTU, respectively. This is representative of conditions when there is no overland flow and stream flows are not at their extremes. The sample size from low flow conditions is much larger than that from storm flow events, and the result is that the mean is heavily weighted to the general background conditions, although extremes might fall well outside the expected values according to the sample distribution (Peggy's Brook 95% confidence interval upper limit is 1.2 NTU). This suggests that the quantification of the mean of turbidity is not a useful statistic for water quality monitoring, as storm events can cause dramatic increases in turbidity.

Table 4.8 Turbidity statistics for Peggy's Brook (June 20/09 – September 02/10; winter 2009-2010 omitted).

		95% Confidence interval (+/- NTU)		
Interval	**Mean (NTU)**	**No autocorrelation**	**Autocorrelation (0.98)**	**Maximum (NTU)**
15 minutes	0.215	0.002	0.023	8.6
1 hour	0.215	0.004	0.024	8.6
6 hours	0.215	0.010	0.023	6.7
1 day	0.207	0.026	0.027	6.7

Table 4.9 Turbidity statistics for Lacey Mill River (June 20/09 - February 27/10).

		95% Confidence interval (+/- NTU)		
Interval	**Mean (NTU)**	**No autocorrelation**	**Autocorrelation(0.98)**	**Maximum (NTU)**
15 minutes	0.459	0.002	0.022	12.1
1 hour	0.455	0.004	0.021	9.1
6 hours	0.451	0.010	0.024	7.6
1 day	0.435	0.021	0.023	7.6
2 days	0.479	0.036	0.036	7.6

An event was recorded on August 4th 2010 around noon during a site visit. During this site visit a torrential downpour was experienced for thirty minutes. The precipitation for the day was 28.7 mm, most of which fell within this short period. The antecedent conditions were dry with no appreciable precipitation since the 22nd of July. This extreme rain resulted in opaque water downstream of a bridge that spans Lacey Mill River, and the river bottom was no longer discernable at less than a half meter depth. Turbidity samples were collected upstream and downstream which respectively measured 1.06 and 28.2 NTU. The opaque nature of the water lasted only thirty minutes as sediment was washed into the stream. The source of this sediment was runoff from the logging road; this runoff entered the stream unimpeded through a ditch. This event was not captured by the sonde because the sonde's placement is upstream. This recorded event shows the sensitivity of both sampling duration and equipment placement when attempting to characterize turbidity. Since the equipment was upstream of the source, the equipment did not capture the event.

Turbidity, as a proxy parameter for suspended sediment, can be used to detect large flushes of sediment into receiving streams. Peggy's Brook, with a less impacted watershed as compared to Lacey Mill River, still possesses large variability in turbidity levels (maximum of 8.6 NTU over the sampling period). Peggy's Brook would be considered to possess a background, or natural, sediment transport regime. Therefore, it would be expected that turbidity levels in tributaries in the Pockwock watershed could increase by orders of magnitude during extreme storm events under natural conditions. It would be expected that greater increases might occur in areas where there has been significant disruption to land cover; this would be useful to characterize from a water quality management perspective. However, the challenge in using turbidity as a water quality metric is establishing background levels. This would involve deployment and collection of continuous data for several years, and in several different watersheds. Turbidity sensors may have more value in strategic deployment in areas expected to have impacted sediment loading due to riparian zone disruption, such as river crossings, sections of streams in close proximity to roads, or downstream of logging activities. As well, the use of turbidity data should be more qualitative than quantitative, as NTUs are a relative measurement scale, and setting arbitrary limits and guidelines upon it is not a reasonable practice (Davies-Colley & Smith, 2001).

4.6 General Characterization of Water Temperature, Conductivity, and pH

Water temperatures are strongly correlated with air temperatures, with stream waters reaching peak temperatures of 27^{o}C in Lacey Mill River and 20^{o}C in Peggy's Brook in the August. The difference in maximum temperatures observed in the two streams is likely largely attributed to less stream shading in the Lacey Mill River subwatershed, as well as more developed land and upstream ponding. Both streams flowed continuously during the winter months, with temperatures in both streams through January and February hovering around 0 ^{o}C. The daily temperature range shows that the temperatures in both Peggy's Brook and Lacey Mill River can change at a rate greater than 5^{o}C/day. From a source water quality perspective there were no concerns associated with temperature, however the local fish population, which is comprised of mostly trout and salmonid species, is expected to be negatively impacted by the high maximum annual temperatures experienced in Lacey Mill River (Wehrly & Wang, 2007). If the health of aquatic species in the river is of concern mitigative actions should be taken to reduce the water temperatures in Lacey Mill River.

The streams in the Pockwock watershed characteristically have low conductivity and low pH. This is a result of the local geology, which has a limited ability to contribute dissolved solids to the stream network, and also possesses limited buffering capacity (Ginn et al., 2007). Peggy's Brook has a much lower conductivity than Lacey Mill River which may naturally be caused by differences in geology and the shorter residence time of Peggy's Brook subwatershed, or a result of human activities such as development and road salting within the Lacey Mill River subwatershed. Peggy's Brook has a more stable conductivity, experiencing a daily range of 0.015 ms/cm whereas Lacey Mill River's conductivity could change as much as 0.044 ms/cm/day. The large daily ranges generally occur during storm events, but storms can both increase or decrease conductivity measurements, which is likely a function of watershed conditions. Storms in the winter months greatly increased conductivity in Lacey Mill River, however during non-winter months a decrease in conductivity was observed during storm events. The increase in conductivity during winter runoff events is likely attributable to road salting activity on the highway and other roadways within the Lacey Mill River subwatershed. Peggy's Brook did not have sufficient data to examine the effect of winter storms on conductivity, but for the rest of the months conductivity decreased during storm events.

Table 4.10 Water temperature characterization and comparison of Lacey Mill River and Peggy's Brook on a monthly basis (^{o}C).

Month	Average		Minimum		Maximum		Maximum Daily Range	
	Lacey	*Peggy's*	*Lacey*	*Peggy's*	*Lacey*	*Peggy's*	*Lacey*	*Peggy's*
May-09	13.34	9.24	10.86	6.67	18.30	14.93	4.30	4.16
Jun-09	18.71	12.98	15.27	9.90	24.23	17.60	4.17	3.48
Jul-09	20.85	15.35	17.26	11.93	25.64	19.37	3.54	3.16
Aug-09	23.06	17.37	17.51	13.37	27.03	20.22	4.70	6.10
Sep-09	16.82	12.61	12.89	8.32	21.63	16.49	4.34	3.86
Oct-09	10.15	8.57	5.68	4.24	16.74	14.17	3.18	3.78
Nov-09	6.04	5.66	3.86	3.31	9.08	8.52	2.52	2.72
Dec-09	1.72	1.71	0.04	0.01	6.43	6.51	2.53	2.95
Jan-10	0.38		0.04		0.86		0.47	
Feb-10	0.44		0.10		0.99		0.49	
Mar-10	3.57		0.01		7.38		3.09	
Apr-10	10.37	5.89	6.00	3.04	13.51	8.76	3.95	2.71
May-10	14.74	9.14	9.95	5.82	22.13	15.61	5.22	5.31
Jun-10	18.48	12.37	13.25	8.67	24.31	16.36	4.94	3.51
Jul-10	22.84	16.76	18.48	12.25	26.41	19.49	3.39	3.73
Aug-10	21.85	15.99	19.39	12.85	24.78	18.45	3.95	3.87
Annual	Sep-09 – Aug-10		0.01	0.01	26.41	19.49	5.22	5.31

Table 4.11 Conductivity characterization and comparison of Lacey Mill River and Peggy's Brook on a monthly basis (ms/cm).

Month	**Average**		**Minimum**		**Maximum**		**Maximum, Daily Range**	
	Lacey	*Peggy's*	*Lacey*	*Peggy's*	*Lacey*	*Peggy's*	*Lacey*	*Peggy's*
May-09	0.079	0.027	0.042	0.024	0.088	0.030	0.044	0.003
Jun-09	0.087	0.028	0.064	0.023	0.091	0.034	0.023	0.007
Jul-09	0.087	0.029	0.084	0.025	0.109	0.035	0.025	0.005
Aug-09	0.089	0.029	0.076	0.021	0.097	0.044	0.015	0.017
Sep-09	0.082	0.028	0.074	0.025	0.092	0.040	0.008	0.015
Oct-09	0.079	0.035	0.074	0.030	0.083	0.041	0.006	0.009
Nov-09	0.078	0.034	0.072	0.031	0.082	0.036	0.005	0.004
Dec-09	0.079	0.033	0.072	0.029	0.093	0.034	0.010	0.004
Jan-10	0.075		0.058		0.860		0.029	
Feb-10	0.068		0.057		0.096		0.036	
Mar-10	0.076		0.055		0.107		0.017	
Apr-10	0.082	0.027	0.074	0.024	0.086	0.028	0.012	0.002
May-10	0.084	0.024	0.081	0.023	0.086	0.026	0.004	0.003
Jun-10	0.088	0.027	0.081	0.023	0.100	0.030	0.012	0.006
Jul-10	0.087	0.028	0.076	0.025	0.109	0.035	0.021	0.009
Aug-10	0.084	0.027	0.078	0.024	0.087	0.033	0.004	0.006
Annual	Sep-09 - Aug-10		0.055	0.023	0.109	0.041	0.036	0.015

Values of pH ranged by over 1 unit in both streams, and a half unit change typically occurred during storms. Storms depress pH in both Peggy's Brook and Lacey Mill River due to acid deposition and the lack of buffering capacity in the soils and geological formations within the watershed area. The maximums in pH observed do not pose any major concerns, however, the minimums experienced are ecologically a concern as some fish and aquatic fauna are expected to experience negative impacts, with particular difficulties in reproduction (Baker et al., 1996).

Table 4.12 pH characterization and comparison of Lacey Mill River and Peggy's Brook on a monthly basis.

	Average		**Minimum**		**Maximum**		**Daily Range**	
Month	*Lacey*	*Peggy's*	*Lacey*	*Peggy's*	*Lacey*	*Peggy's*	*Lacey*	*Peggy's*
May-09	4.86	4.58	4.81	4.43	4.91	4.72	0.08	0.17
Jun-09	4.87	4.55	4.72	4.29	4.97	4.75	0.15	0.29
Jul-09	4.87	4.54	4.72	4.31	5.02	4.82	0.08	0.25
Aug-09	4.86	4.47	4.35	4.06	4.98	4.76	0.56	0.63
Sep-09	4.61	4.53	4.35	4.19	4.81	4.85	0.13	0.57
Oct-09	4.43	4.24	3.47	4.02	5.16	4.49	1.65	0.32
Nov-09	4.45	4.30	4.37	4.21	4.50	4.42	0.06	0.18
Dec-09	4.37	4.25	4.30	4.13	4.45	4.38	0.09	0.24
Jan-10	4.45		4.32		4.54		0.09	
Feb-10	4.51		4.36		5.16		0.18	
Mar-10	4.62		4.45		4.79		0.13	
Apr-10	4.66	4.35	4.59	4.19	4.74	4.61	0.05	0.15
May-10	4.72	4.65	4.61	4.37	4.81	5.00	0.06	0.24
Jun-10	4.66	4.22	4.58	4.08	4.71	4.65	0.09	0.48
Jul-10	4.64	4.11	4.48	3.70	4.93	4.33	0.27	0.35
Aug-10	4.70	4.28	4.63	3.96	4.81	4.57	0.07	0.32
Annual	Sep-09 - Aug-10		3.47	3.70	5.16	5.00	1.65	0.57

4.7 Confidence Intervals for Average Annual Concentrations

Confidence intervals of annual means, which describes the 95% confidence interval that the annual mean of any given year is expected to fall within, were assessed for pH and conductivity in Lacey Mill River. These two parameters were selected because they were monitored on a continuous basis and had a relatively complete data set for an entire year (33300 samples of 35040). Annual means were constructed considering seasonal trends and autocorrelation, as well as not considering these trends. Removal of seasonal trends will always improve your confidence interval, while autocorrelation will always widen your confidence interval. The seasonal trend is fit based on minimizing the residuals (Appendix F: Fitting of a Deterministic

Seasonal Model), which will always result in a lower variance and as a result a smaller confidence interval. Autocorrelation however, will always result in an increase in the confidence interval width; the greater the level of autocorrelation the more substantial the effect will be. This result is expected as the effect of autocorrelation on annual means estimates is a decrease in the sample size, which results in an increase of the variance (Equation 2.7).

Seasonal trends have a more pronounced effect on decreasing the confidence interval width as the sampling interval increases. At very small intervals the seasonal trend can effectively be ignored because fitting a deterministic model will not appreciably change the confidence interval width; the confidence interval width, when not considering the seasonal, is already very narrow because of the large number of samples. However, at large sampling intervals the seasonal trend has a notable effect as the extremes of your measurements are decreased and some of the residual variance is removed by seasonal considerations. It is noteworthy that the fitted trend only encompassed 16 month duration. Ideally, a dataset would be sufficiently long so that it would fully describe annual variability. However, this is impractical as, depending on the site, the required dataset length might be extremely large.

At high sampling frequencies autocorrelation is very strong and has a very pronounced effect on confidence intervals widths of annual means. Values of pH possessed a first order autocorrelation factor of 0.997 and conductivity had a first order autocorrelation factor of 0.991. Because the autocorrelation factors are very strong, it is erroneous to consider samples as being independent at short time lags. Using the analysis of scale of fluctuation, Equation 4.1, which is a theoretical measure of distance in time or space between measurements, before they can largely be assumed independent of each other (Fenton, 1999), times of 6.9 days and 2.3 days are acquired for pH and conductivity, respectively. This functional independence is perceptible from Figure 4.16 and Figure 4.17 as the distinction between the trends considering autocorrelation and not considering autocorrelation becomes less perceptible as the time sampling interval increases, and is relatively irrelevant at a sampling interval greater than the scale of fluctuation.

One issue with longer sampling periods is the omission of intermittent events that have a large impact on annual estimates. This can be seen in Figure 4.16 as at longer intervals (10, 20, and 30 days) the trends begin to deviate. This deviation is caused by the increased chance that sampling at these time steps would miss a number of extreme events, and thus resulted in an under prediction of confidence interval widths. This is the peril of not sampling enough, as your population might not be representative. Realistically, as the number of sampled years increased, this effect would decrease, and the general trend would be realized.

There is a practical aspect associated with the creation of these control charts, in that a water quality manger can modify their statistical post processing depending upon the sampling interval. If the sampling interval is shorter than the scale of fluctuation, both autocorrelation and seasonal trends should be considered. Sampling intervals that are larger than the scale of fluctuation, and up to twice your scale of fluctuation, can be analyzed without the removal of trends, and minimal statistical impact will be present. And lastly, datasets generated using sampling intervals greater than twice your scale of fluctuation should only require corrections for seasonal trends, as autocorrelation at this length of interval will not significantly bias the results. It is further noteworthy that ignoring seasonal trends will provide a conservative estimate which, though not statistically optimal, may be practical in many situations.

Equation 4.1 Scale of fluctuation (correlation length) of a first order autoregressive model (Fenton, 1999).

$$\theta = -\frac{2|\tau|}{\ln\left(p(\tau)\right)}$$

$\theta = \text{scale of fluctuation (\# of time intervals)}$
$\rho(\tau) = \text{first order correlation at distance of } \tau$
$\tau = \text{length of time between two sample (\# of time intervals)}$

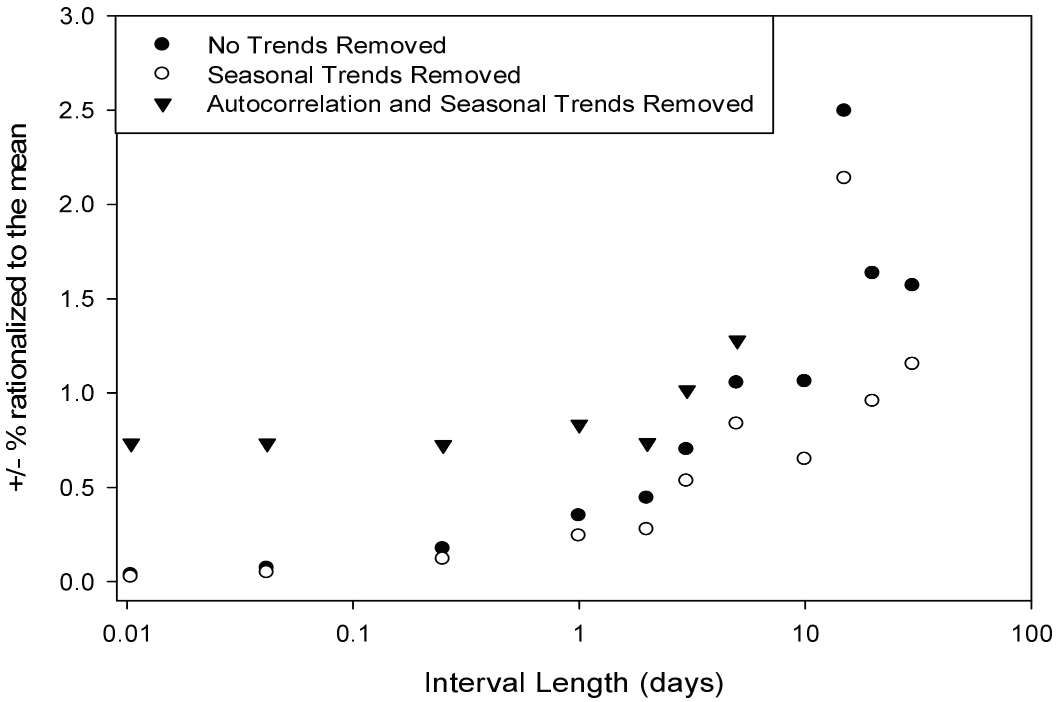

Figure 4.16 The confidence interval of the annual mean of pH in Lacey Mill River as affected by sampling frequencies and statistical post-processing.

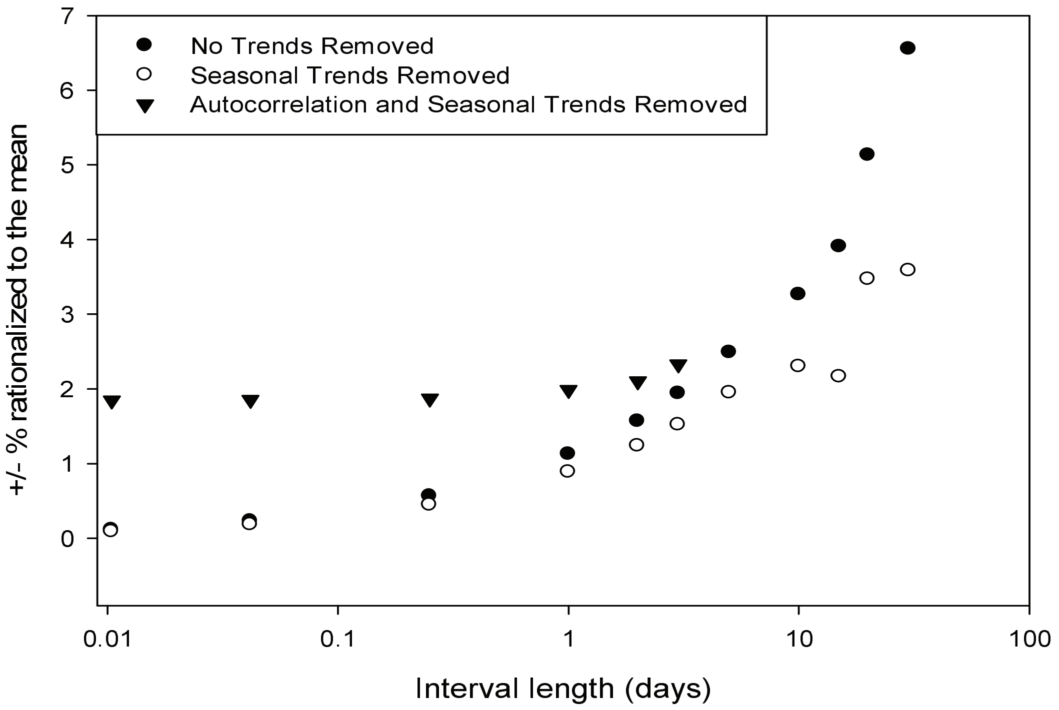

Figure 4.17 The confidence Interval of the annual mean of conductivity in Lacey Mill River as affected by sampling frequencies and statistical post-processing.

4.8 Confidence Intervals Of A Given year

The phrasing of statistical questions is extremely important as a slight change in the question being asked can have an impact on the statistics that need to be used, and the results which are produced. The questions of "what is the mean of a given year" and "what is the annual mean" may seem minimally different, but they differ greatly in the statistics that are applied, and more importantly how autocorrelation affects the estimate.

As was discussed earlier, the question of "what is the annual mean" is negatively impacted by autocorrelation as there is greater uncertainty in estimating the mean of any year when the measurements of a given year are correlated because the data is being used to extrapolate. However, if the goal is to know what the mean is for a specified measurement period, the estimate benefits from autocorrelation because it is an additional source of information that adds greater confidence to estimating values between the sampling intervals. Autocorrelation improves estimates of annual means to the extent that even if the measurement period is every five days for pH your confidence interval is smaller than that of your measurement error, or measurement resolution. The implication of this result is that a monitoring program can reduce the number of samples while still maintaining the necessary level of confidence in the estimate of the annual mean of a given year.

Table 4.13 Confidence intervals of the annual mean of pH for Lacey Mill Brook from August 28 2009 - August 28 2010.

interval	CIW	Mean	% +/- Rationalized to the mean
15 minutes	8.62E-06	4.57	0.00%
1 hour	7.53E-05	4.57	0.00%
6 hours	2.69E-04	4.57	0.00%
1 day	9.49E-04	4.57	0.01%
2 days	1.50E-03	4.57	0.02%
3 days	3.94E-03	4.56	0.04%
5 days	8.74E-03	4.56	0.10%

Table 4.14 Confidence Intervals of the annual mean of conductivity for Lacey Mill Brook from August 28 2009 - August 28 2010.

interval	CIW	Mean	% +/- rationalized to the mean
15 minutes	8.62E-06	0.0803	0.01%
1 hour	1.75E-05	0.0803	0.01%
6 hours	4.49E-05	0.0804	0.03%
1 day	1.05E-04	0.0803	0.07%
2 days	1.78E-04	0.0803	0.11%
3 days	2.57E-04	0.0805	0.16%

4.9 Escherichia Coli

Escherichia coli is currently the accepted indicator of fecal contamination for microbial water quality monitoring program. However, it must be recognized that it is conservative measurement of microbial contamination, and it is normal to detect low levels of *E. coli* in environments not impacted by fecal contamination. **Error! Reference source not found.** provides an illustration of the variability in *E. coli* concentrations in the Pockwock watershed, which is largely forested. Lacey Mill River and Peggy's Brook have comparable levels of bacteria and a similar seasonal trend.

The seasonal trend shows elevated concentrations during the summer with peaks from mid July to late August, and very low concentrations in the winter, with many results coming back as non-detect. The seasonal trend can be attributed to the environmental conditions, and more specifically the temperature. *E. coli*, though resilient in the natural environment, requires a sufficient temperature to multiply and this requirement will not be met during the winter (Ratkowsky et al., 1982). The peaks in the seasonal trend of measured *E. coli* concentrations correspond with the warmest water temperature (20 – 27 ^{0}C), which is approaching the ideal temperature condition for the replication of most *E. coli* strains (37 ^{0}C) (Cooper, Bennett, & Lenski, 2001).

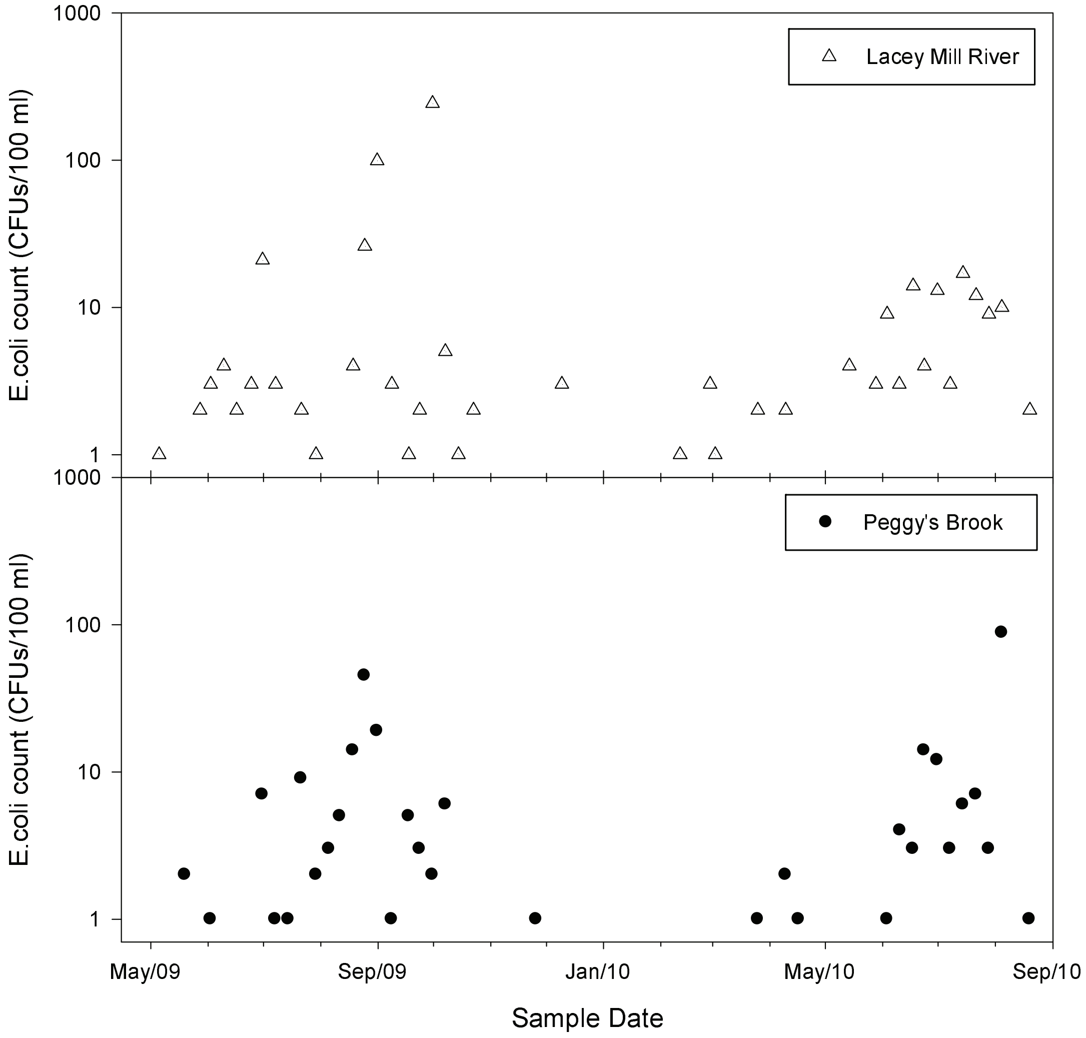

Figure 4.18 Measurements of E. coli throughout the monitoring period (zeros are omitted due to log scale).

Although temperature is essential for the growth of *E. coil*, it can also be recognized that there is a strong relationship between storm events and increased concentrations on *E. coli* in water samples. The relationship is not as clear when comparing storm intensity to *E. coli* concentration changes. This is expected as the antecedent moisture conditions in the watershed will also greatly affect the hydrologic response of the watershed, and transport of bacteria concentrations. Strong relationships between flow and *E. coli* concentration, and temperature and *E. coli* concentration, have been observed by other researchers (Edwards et al., 1997; Boyer and Kuczynska, 2003).

An extreme value distribution, type II (Frechet), was constructed for *E. coli* concentrations in Lacey Mill River, to develop a more complete expectation of extremes with different return periods. Based on this distribution, the concentration of *E. coli* associated with a 16 month return period would be 64 CFUs/ 100 ml, and the 20 year return period concentration would be 200 CFUs/100 ml. These predicted extremes suggest that high concentrations, when put into context with current watershed conditions, are expected. Pertaining to Pockwock Lake, this may not be of great concern, as bacteria populations will be diluted and attenuated in the lake before reaching the treatment facility. However, for source waters where this attenuation between source and treatment will not be significant, as in a river system, there is cause for concern with extreme events producing elevated bacteria concentrations at drinking water intakes.

Table 4.15 Expected measured extremes of *E. coli* concentrations In Lacey Mill River.

Time of Return	Expected Extreme (CFUs/ 100 ml)
16 Months	64
3 Year	80
5 Year	105
10 Year	147
20 Year	200

4.10 Quality Assurance and Quality Control

4.10.1 Sampling Variability

Triplicates were taken in quick succession for parameters measured by grab sample to assess the variability present in sampling. Triplicate sampling of this nature was conducted twice during the sampling period for important parameters, and the mean and 95% confidence intervals were computed for the triplicates. The results are summarized in Table 4.16.

Table 4.16 Sampling QA/QC for grab sample parameters.

Parameter	Site	Date sampled	sample 1	sample 2	Sample 3	Mean	95% confidence Interval (mean +/-)
E. Coli (Cfu/100 ml)	Lacey	18-Aug-09	3	4	4	4	1
		09-Apr-10	2	3	2	2	1
	Peggy's	18-Aug-09	11	15	16	14	3
		09-Apr-10	4	1	0	2	2
Total Coliform (Cfu/100 ml)	Lacey	09-Apr-10	6	9	6	7	2
	Peggy's	09-Apr-10	3	7	3	4	3
TOC (mg/l)	Lacey	18-Mar-10	4.28	4.65	4.70	4.54	0.26
		24-Aug-09	5.03	4.88	4.75	4.89	0.16
		28-Jul-10	5.70	13.01	5.94	8.21	4.70
	Peggy's	24-Aug-09	21.33	20.61	21.42	21.12	0.50
		18-Mar-10	8.04	8.09	8.27	8.14	0.14
		28-Jul-10	15.98	15.37	15.76	15.70	0.35
TP (ug/l)	Lacey	24-Aug-09	62.5	6.0	5.6	24.7	37.0
		25-Mar-10	159.0	4.4	4.8	56.1	100.9
		28-Jul-10	6.0	5.9	5.1	5.7	0.6
	Peggy's	24-Aug-09	15.7	16.3	16.1	16.0	0.4
		25-Mar-10	10.4	10.6	18.2	13.1	5.0

Parameter	Site	Date sampled	sample 1	sample 2	Sample 3	Mean	95% confidence Interval (mean +/-)
		28-Jul-10	15.9	17.0	21.9	18.3	3.6
TSS (mg/l)	Lacey	18-Aug-09	1	1	0	1	1
		16-Apr-10	0	1	0	0	1
	Peggy's	18-Aug-09	0	0	0	0	
		16-Sep-10	0	0	0	0	
Turbidity (NTU)	Lacey	09-Apr-10	0.45	0.51	0.65	0.54	0.12
		28-Jul-10	0.78	0.60	0.57	0.65	0.13
	Peggy's	09-Apr-10	0.17	0.28	0.25	0.23	0.06
		28-Jul-10	0.00	0.00	0.10	0.03	0.07

The QA/QC results for the sampling procedure generally show positive results with small confidence intervals. However, TOC and TP both had incidents where the QA/QC results were unacceptable. The triplicate sampling of TOC in Lacey Mill River on July 28 produced a high level of variability. This level of variability would naturally exist within a flowing water system as an excess of detritus or sediment may have been captured in the sample. When examining the results for TP, however, there were unacceptable magnitudes of variability on two occasions. The source of this error was never determined. This brings into question the elevated values that were used in the loading estimates of TP from the two streams. Two possible causations of the discrepancies are excessive sediment captured in one sample, and/or samples not being analyzed on a timely manner.

4.10.2 Analytical Variability

To quantify the variability in the analytical procedure one sample was taken from each stream, shaken vigorously in an effort to homogenize the sample, and then split into three bottles. Table 4.17 contains a summary of the analytical QA/QC results for grab sampled parameters.

Table 4.17 QA/QC of analytical procedures for grab sample parameters

Parameter	Site	Date	sample 1	sample 2	sample 3	mean	95% confidence Interval (mean +/-)
E. coli (Cfu/ 100 ml)	Lacey	23-Sep-09	1	2	2	2	1
		29-Apr-10	1	0	0	0	1
	Peggy's	23-Sep-09	3	4	2	3	1
		23-Apr-10	0	0	0	0	
Total Coliform (Cfu/ 100 ml)	Lacey	23-Apr-10	3	10	3	5	4.57
	Peggy's	23-Apr-10	0	0	0	0	
TOC (mg/l)	Lacey	17-Sep-09	4.03	4.06	4.43	4.17	0.25
		22-Apr-10	3.61	3.45	4.11	3.73	0.39
	Peggy's	17-Sep-09	12.07	12.57	12.72	12.45	0.39
		22-Apr-10	11.33	10.3	10.62	10.75	0.60
TP (ug/l)	Lacey	23-Sep-09	9.8	6.6	7.6	8.0	1.9
		14-Jul-10	6.8	8.1	6.4	7.1	1.0
		21-Jul-10	4.1	4.4	5.2	4.6	0.6
	Peggy's	23-Sep-09	16.6	17.8	16.9	17.1	0.7
		14-Jul-10	13.1	13.1	13.3	13.2	0.1
		21-Jul-10	62.5	13.1	12.9	29.5	32.3
Turbidity (NTU)	Lacey	30-Sep-09	0.24	0.33	0.24	0.27	0.06
		04-Aug-10	28.4	28.9	27.2	28.2	1.0
	Peggy's	30-Sep-09	0.30	0.22	0.42	0.31	0.11

In general, the QA/QC results suggest that the analytical procedures were performed adequately, with triplicates producing small confidence intervals for most parameters. However, TP proved to be an issue again, with a large confidence interval produced for the July 21, 2010 sampling event in Peggy's Brook. It is possible that the same errors apply as in the QA/QC for sampling, or the samples were not homogenized well enough, with one sample receiving a disproportionate amount of sediment that was high in TP.

Chapter 5 Practical Considerations for Designing Monitoring Programs

The use of continuous monitoring equipment to monitor environmental impacts and environment quality is increasing. Although these new programs are well intentioned, they often lack direction, and as a result effectiveness. The lack of direction is caused by organizations perceiving environmental monitoring as a cost of operation, and not a task that could potentially improve operations and reduce costs. An organization must invest both the monetary and human resources for continuous monitoring to become beneficial, but the investment alone will not reap benefits as the corporate culture must be willing to incorporate information from their monitoring program into management decisions. In this chapter, practical, general considerations are provided. As well, specific recommendations for water monitoring within the Pockwock Watershed are presented.

5.1 Design Of Monitoring Program

- Before any equipment is even purchased a full scale risk assessment should be carried out. In this assessment you should attempt to incorporate all available data in an effort to identify parts of the watershed that are most likely to impact water quality.
- Determine the questions that you want to answer and then determine, to the best of current knowledge, what will be required to answer this question. Decisions should be made for monitoring sites, monitoring strategy and equipment.
- Do not over extend the network. It is better to concentrate resources on fewer sites and obtain high quality data than to obtain low quality data on many sites.
- Design and implement a plan for data storage and management before implementing a large scale monitoring program as retroactive implementation will be much more difficult.
- It is necessary to not only consider the capital cost of field monitoring equipment but to consider the capital cost of software. It might be necessary to obtain software in these categories: database, statistical, and programming.

5.2 Use Of Continuous Monitoring Equipment

Multiparameter sondes are currently the most common continuous water quality monitoring probes available. Some practical suggestions for using this equipment are:

- The sondes have onboard computers and power sources, but also have the ability to connect to an external power source and datalogger. It is suggested that the extra money be invested for external power sources and dataloggers. The benefit of having a secondary backup when collecting continuous data is invaluable, as both power sources and the sonde’s onboard computer can fail. The secondary benefit is that with an external datalogger the sonde can be connected to upload real time data to a website.
- The ability to see real time data, especially if the site is accessible, would be a great asset. If monitored, this real time data would be a useful tool to mobilize water sampling teams to capture unique events to better understand the watershed. If you do not have timely data access, events are recorded but the context of the event is lost along with the knowledge and understanding that may come from the event. Also, the continual monitoring allows you to detect equipment failure and problems early minimizing the number of missed samples.
- For each sensor/data acquisition system, it is recommended to have at least one backup equivalent piece of equipment. This would greatly streamline monthly calibrations, as one piece of equipment may be removed and simply replaced with a freshly calibrated piece of equipment.
- There needs to be a balance between security and accessibility. If the sonde will not be accessible for part of the desired monitoring season another site should be considered for the equipment. However, due to the capital invested in this equipment, it should be as inconspicuous as possible and secured.
- A strategically placed video camera could prove very useful in both security and providing context for water quality events. There are times where intermittent conditions occur, and without a physical presence it is difficult to determine if the event was an electronic anomaly or caused by a physical event. Video data could provide insight in many of these cases, and the equipment is relatively inexpensive.

- When placing and securing equipment you need to plan for the highest flow conditions. Lacey Mill Brook at high flows overflowed the rivers banks and river stage was estimated to have increase by over 1.5 m. If sensitive equipment, like dataloggers, were placed on the stream bank they would have sustained damage.
- Logs of equipment maintenance must be kept, especially for consumable probes. Their performance needs to be monitored and probes need to be replaced on a timely manner.

5.3 Data Analysis

- Data needs to be stored in an accessible and organized manner with sufficient backup.
- Data correction and verification has to take place on a timely manner. To do this efficiently, software may need to be purchased or designed.
- Excel, although a powerful program for reporting and some analysis, will not be adequate for all analyses.
- It cannot be expected that purchasable software will meet all needs with respect to data analyses, and as a result some in-house programming may be necessary.
- Data analysis on a timely manner is imperative because data analysis provides the ability to determine if your monitoring program can suitably answer the question, as well as provide insight into the optimization and improvement of the monitoring program.

5.4 Pockwock Water Sampling Parameter Recommendations

The general water quality in Lacey Mill River and Peggy's Brook was good; however, the short duration of this study results in high levels of uncertainty. This section will suggest sampling frequency and seasons for parameters that were measured in this study. The main assumption in the recommendations is that forestry activity and developments stay the same as they were through the monitoring period of April 2008 – August 2010. If the activity changes then it may also be necessary to change the monitoring strategy. It further must be recognized that the most important part of a SWMP is to be able to assess that the organization is meeting the SWPP goals, and in doing so the monitoring strategy should be modified as needed to gain all necessary information.

Tables 5.1 and 5.2 summarize the suggested measurement season and frequency for the parameters measured in the study, but the choice of the parameters will depend on the goals of the monitoring program. The next four paragraphs discuss the parameters, and their associated sampling seasons and frequency, with specific reference to the following sampling goals for the Pockwock watershed..

(i) *The goal is to track general changes in ecosystem health and functionality*.

The parameters that are required to be measured are temperature, DO, pH, and conductivity. DO and temperature only need to be monitored from spring into fall, while pH and conductivity require measurement year round to adequately monitor ecosystem health. This program would have to be instituted on a long term basis. To develop baseline dataset for the watershed, at least 3 consecutive years of data should be collected. The program could be designed to have semi-continuous periods (such as monitoring every other year). The risk of utilizing a semi-continuous monitoring strategy is a delayed response to changes in water quality.

(ii) *The goal is to track changes in landscape processes and assess the effects of development activities.*

The parameters that require measurement are pH, *E. coli* and conductivity. pH and conductivity need to be measured year round while *E. coli* only requires measurement from spring - fall. The monitoring program requires long term monitoring with cataloguing of landscape changes and development activities.

(iii) *The goals are to predict and manage TOC and TP loading into Pockwock Lake.* An intensive storm chasing sampling program would need to be instituted for TOC and TP. A stream stage discharge curve would need to be developed to provide flow data on 15 minute intervals in order to compute loading. To meet the goals, catchment scale models would need to be developed for each subwatershed to be assessed. These models would be necessary to predict the impacts of management decisions. To develop the model, detailed spatial information such as forest composition, land slope, land use, and soil types would be required.

(iv) *The goal is to monitor impacts associated with a particular activity inside the watershed.* Activities in the watershed would include tree falling, bridge repair, and road construction. The parameters to be measured would be case dependent, but in general, pH, conductivity, temperature, DO, *E. coli,* flow, TP, TOC, turbidity, and TSS should all be monitored with sampling season and frequency to reflect tables 5.1, which is representative of an impacted stream.

It is suggested that current sampling efforts in the Pockwock Lake watershed be focused on Lacey Mill River because it is larger, there is more uncertainty associated with measurements, events that suggest there are ecological issues with pH and conductivity were recorded, and the presence of development within the subwatershed.

For most parameters it was not beneficial to sample at a rate faster than hourly. Exceptions are flow and turbidity as flow is easy to measure, and the increased resolution will improve loading estimates, and turbidity is event driven and hourly measurements may not capture extreme events. Sampling on hourly time intervals should be adequate to allow for the statistical description of all other parameters that can be measured continuously.

Parameters which cannot be sampled by continuous monitoring equipment (microbial contaminants, P and TOC) should have a sampling strategy other than monthly grab samples.

Microbial contaminants should be sampled as often as possible in Lacey Mill River because of the development in the upper watershed. However, streams which are not impacted by forestry activities or development, such as Peggy's Brook at the current time, are not expected to require as frequent sampling, and therefore the current strategy of monthly sampling is adequate. It should be recognized, because of the attenuation in the lake, that microbial contamination at the plant intake is unlikely, but it is potentially a good parameter to ascertain information about the effectiveness of the SWPP. It would be beneficial to have an intensive nutrient loading monitoring program, which utilizes storm chasing, for both Peggy's Brook and Lacey Mill River as there are potentially watershed management strategies that could help alleviate nutrient related problems, such as DBP formation.

Measuring some parameters all year round may not be the most effective strategy. Temperature, dissolved oxygen and microbial contamination, which are strongly correlated with temperature, may not require measurement from late fall into early spring. Through the winter, water temperatures will consistently be around freezing, dissolved oxygen will be high, and microbial counts will be very low or non-detect. Turbidity is a unique parameter and it is only beneficial at specific sites that may be impacted by intermittent activities. If turbidity sensors are to be deployed, they should be deployed before the activity is to commence and remain until activity completion.

Table 5.1 Recommendations for sampling season and frequency for Lacey Mill River.

Parameter	Measurement Season	Sampling Frequency
Temperature	Late Spring - Early Fall	Hourly
pH	Year Round	Hourly
Conductivity	Year Round	Hourly
Dissolved Oxygen	Late Spring – Early Fall	Hourly
Microbial	Late Spring - Early Fall	Weekly
Phosphorous	Year Round	Storm Chasing
Organic Carbon	Year Round	Storm Chasing
Flow	Year Round	15 minute
Turbidity	Activity Dependent	15 minute
TSS	Do Not Monitor	

Table 5.2 Recommendations for sampling season and frequency for Peggy's Brook.

Parameter	Measurement Season	Sampling Frequency
Temperature	Late Spring - Early Fall	Monthly
pH	Year Round	Monthly
Conductivity	Year Round	Monthly
Dissolved Oxygen	Late Spring – Early Fall	Monthly
Microbial	Late Spring - Early Fall	Monthly
Phosphorous	Year Round	Storm Chasing
Organic Carbon	Year Round	Storm Chasing
Flow	Year Round	15 minute
Turbidity	Activity Dependent	15 minute
TSS	Do Not Monitor	

Chapter 6 Conclusions

This research has highlighted the importance of defining the water quality goals and questions during the creation of a SWPP. The data suggests that SWMP should be custom tailored to meet specific goals and answer particular questions associated with source water quality.

Continuous monitoring will provide useful information for particular SWMPs. However, it may not be appropriate for all SWMPs because of the high capital cost, high maintenance cost, labor intensive nature of analysis, and the still limited array of parameters that can be measured. Continuous monitoring is a useful tool for assessing water quality, but is unable to provide answers to water quality questions.

Data cannot simply be collected and stored with future aspirations to gain information from it. Careful considerations must be applied, in advance, to how data is to be stored, processed, and analyzed in order to obtain information that has value.

Monitoring strategies have been recommended for the Pockwock Lake watershed, and the strategies include both discrete and continuous monitoring elements. It is recommended that Halifax Water reexamines the goals of their SWPP and SWMP for Pockwock Lake watershed in order to determine how to most effectively allocate available funds to the programs.

The data recorded over the 16 month monitoring period suggests that watershed management strategies could potentially aid Halifax Water in meeting their water treatment goals for DBPs by reducing the concentration of TOC at the intake of the plant. However, an intensive monitoring program would have to be implemented to assess the spatial and temporal variability of TOC in the watershed in order to assess potential strategies.

6.1 Recommendations for Future Research

1. Extend the monitoring period of this research from 16 months to over 3 years.

2. Design and implement a 3 year pilot study to assess TOC and TP loadings into Pockwock Lake. This plan would include both sampling and modeling components in an effort to assess watershed management strategies.

3. Implement an intensive pilot scale SWMP study, similar to the one constructed in this study, for a river source water to assess monitoring strategies for flowing source water systems.

References

Ahtiainen, M. (1992). The Effects of Forest Clear-Cutting and Scarification on the Water Quality of Small Brooks. *Hydrobiolgia* , 465-473.

Allan, D. (1995). *Stream Ecology: Structure and function of running water.* Dordecht: Kluwer.

American Public Health Association; American Water Works Association; Water Environment Federation. (1995). *Standard Methods for Examination of Water and Wastewater, 19th ed.* Washington, DC: American Public Health Association.

American Public Health Association; American Water Works Association; Water Pollution Control Federation; Water Environtment Federation. (1998). *Standards Methods for Examination of Water and Wastewater, 20th ed.* Washington, DC: American Public Health Association.

Baker, J., Sickle, J. V., Gagen, C., DeWalle, D., Sharpe, W., Carline, R., et al. (1996). Episodic Acidification of Small Streams in the Northeastern United Sates: Effects on Fish Populations. *Ecological Applications* , 422-437.

Berthouex, P., & Brown, L. (2002). *Statistics for Environmental Engineers 2nd. Ed.* Boca Raton: CRC Press.

Binkley, D., & Brown, T. (1993). Forest Practices as Nonpoint Sources of Pollution in North America. *Journal of the American Water REsources Association* , 729-740.

Booth, D., Harley, D., & Jackson, R. (2002). Forest Cover, Impervious-Surface Area, and the Mitigation of Stormwater Impacts. *Journal of AMerican Water Resources Association* , 835-845.

Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2008). *Time Series Analysis: Forcasting and Contriol.* Hoboken: John Wiley & Sons.

Boyer, D. G., & Kuczynska, E. (2003). Storm and Seasonal Distributions of fecal coliforms and cryptosporidium in a spring. *Journal of the American Water Resources Association* , 1449-1456.

Brezonik, P., & Stadelmann, T. (2002). Analysis and prective models of stormwater runoff volumes, load, and pollutant concentrations from watershed in Twin Cities Metropolitan area, Minnesota, USA. *Water Research* , 1743-1757.

Brunet, R., & Astin, K. (1998). Variation in phosphorous flux during a hydrological season: the River Adour. *Water Research* , 547-558.

Brylinsky, M. (2004). *User's Manual for Prediction of Phosphorous Concentrations in Nova Scotia Lakes: A Tool for Decision Making Version 1.0.* Nova Scotia Enviornment and Labour.

Caissie, D. (2006). The thermal regime of rivers: a review. *Freshwater Biology* , 1389-1406.

Canadian General Standards Board. (1991, June 01). Liquid Flow Measurements in Open Channels - Velocity-Area Methods .

CCME. (2003). *Canadian Water Quality Guidelines for the Protection of Aquatic Life.*

CCME. (1999). *Community Water Supplies.* Winnipeg.

Chin, D. A. (2006). *Water-Resources Engineering.* Upper Saddle River: Pearson Prentice Hall.

Chin, D. (2006). *Water-Quality Engineering in Natural Systems.* Hoboken: John Wiley & Sons.

Coles, S. (2001). *An Introduction to Statisticfal Modeling of Extremes.* London: Springer-Verlag.

Cooke, S. E., & Prepas, E. E. (1998). Stream phosphrous and nitrogen export from agricultural and forested watershed on the Boreal PLain. *Canadian journal of fisheries and aquatic sciences* , 2292-2299.

Cooper, V. S., Bennett, A. F., & Lenski, R. E. (2001). Evolution of Thermal Dependence of Growth Rate of Escherichia Coli opulations During 20,000 Generations in a Constant Environment. *Evolution* , 889-896.

Davies-Colley, R., & Smith, D. (2001). Turbidity, Suspended Sediment, And Water Clarity: A Review. *Journal of the American Water Resources Association* , 1085-1101.

Davis, M. L., & Cornwell, D. A. (2008). *Introduction to Environmental Engineering: Fourth Edition.* New York: McGraw Hill.

Devore, J., & Farnum, N. (1999). *Applied Statistics for Engineers and Scientists.* Pacific Grove: Brooks/Cole Publishing Company.

Droste, R. L. (1997). *Theory and Practice of Water and Wastewater Treatment.* Hoboken: John Wiley & Sons.

Edberg, S., Rice, E., Karlin, R., & Allen, M. (2000). Escherichia xcoli:the best biological drinking water indicator for public health protection. *Journal of Applied Microbiology* , 106-116.

Edwards, D., Coyne, M., Daniel, T., Vendrell, P., Murdoch, J., & Moore, P. (1997). Indicator Bacteria Concentrations of Two Northwest Arkansas Streams in Relation to Flow and Season. *American Society of Agricultural Engineers* , 103-109.

Eutech Instruments. (n.d.). *Introduction to DO.* Retrieved January 15, 2011, from Eutech Instruments Water Analysis Solutions: http://www.eutechinst.com/tips/do/01.pdf

Federal-Provincial-Territorial Committee on Drinking Water. (2008). *Guidelines For Canadian Water Quality Summary Table.* Health Canada.

Feller, M. C. (2005). Forest harvesting and Streamwater Inorganic Chemistry in Western North America: A Review. *Journal of the American Water Resources Association* , 786-811.

Fenton, G. (1999). Estimation for Stochastic Soil Models. *Journal of Geotechnical and Geoenvironmental Engineering* , 470-485.

Ginn, B. k., Cumming, B. F., & Smol, J. P. (2007). Assessing pH changes since pre-industrial times in 51 low-alkalinity lakes in Nova Scotia, Canada. *Natural Research Council Canada* .

Government of British Columbia. (2008). Drinking Water Protection Act. Victoria: Queen's Printer.

Hauer, F. R., & Lamberti, G. A. (2006). *Methods In Stream Ecology.* London: Elsevier.

Health Canada. (2009, March 09). *Environmental and Workplace Health*. Retrieved January 23, 2011, from Nitrate/Nitrite: http://www.hc-sc.gc.ca/ewh-semt/pubs/water-eau/nitrate_nitrite/index-eng.php

Hirsch, R., Alexander, R., & Smith, R. (1991). Selection of Methods for the Detection and Estimation of Trends in Water Quality. *Water Resources Research* , 803-813.

Hrudey, S., Payment, P., Huck, P., Gillham, R., & Hrudey, E. (2003). A fatal waterborne disease epidemic in Walkerton, Ontario: comparison with other waterborne outbreaks in the developed world. *Water Science and Technology* , 7-14.

Inamdar, S. P., Christopher, S. F., & Mitchell, M. J. (2004). Export mechanisms for dissolved organic carbon and nitrate during summer storm events in a glaciated forested catchment in New Your, USA. *Hydrological Process* , 2651-2661.

Kateman, G., & Muskens, P. (1978). Sampling of Internally Correlated Lots. The Reproducibility of Gross Samples As A Function of Sample Size, Lot Size and Numer of Samples Part II. Implication for Practical Sampling and Analysis. *Analytica Chimica Acta, Computer Techniques and Optimization* , 11-20.

Lal, R. (1998). Deforestation and land-use effects on soil degradation and rehabilitation in western Nigeria. III. Runoff, soil erosion and nutrient los. *Land Degradation & Development* , 99-119.

Land, C. (1972). An evaluation of Approximate Confidence INterval Estimation Methods for Lognormal Means. *Technometrics* , 145-158.

Laursen, E., Mygrind, O., Rasmussen, B., & Ronne, T. (1994). Gastroenteritis: a woterborne outbreak affecting 1,600 people in a small Danish Town. *Journal of epidemiology* , 453-458.

Lawless, J. (1982). *Statistical Models and Methods for Lifetime Data.* New York: John Wiley & Sons.

Lehmann, E., & Casella, G. (1998). *Theory of Point Estimation.* New York: Springer.

Littlewood, I. (1995). Hydrological Regimes, Sampling Strategies, and Assessment of Errors in Mass Load Estimates for United Kingdom Rivers. *Environment International* , 211-220.

Loftis, J. C., McBride, G. B., & Ellis, J. C. (1991). Considerations of Scale in Water Quality Monitoring and Data Analysis. *Journal of the American Water Resources Associatoin* , 255-264.

Loftis, J., & Ward, R. (1980). Sampling Frequency selection for regulatory Water Quality Monitoring. *Journail of the American Water Resources Association* , 501-507.

Loftis, J., Ward, R., & McBride, G. (1986). The "Data-rich but Information-poor" Syndrom in Water Quality Monitoring. *Environmental Management* , 291-297.

Macdonald, J., Beaudry, P., MacIsaac, E., & Herunter, H. (2003, July 4). The effects of forest harvesting and best management practices on streamflow and suspended sediment concentrations during snowmelt in headwater streams in sub-boreal forests of British Columbia, Canada. British Columbia, Canada: NRC Research Press.

Mackenzie, D., & Cornwell, D. (2008). *Introduction to Environmental Engineering.* New York: McGraw Hill.

Mallin, M. A. (2009). Effect of Human Land Development on Water Quality. In M. A. Mallin, *Handbook of Water Purity and Quality* (pp. 67-94).

Mcbean, E. A., & Rovers, F. A. (1998). *Statistical Procedures for Analysis of Environmental Monitoring Data & Risk Assessment.* Saddle River: Prentice Hall PTR.

Meybeck, M. (1982). Carbon, Nitrogen, and Phosphorous Transport by World Rivers. *American Journal of Science* , 401-450.

Moore, D., & McCabe, G. (2006). *Introduction to the Practices of Statistitics.* New York: W.H. Freeman and Company.

Muskens, P., & Kateman, G. (1978). Sampling of Internally correlated Lots. The Reproducibility of Gross Samples as a Function of Sample Size, Lot Size and Number of Samples, Part I. Theory. *Analytica Chimica Acta, Computer Techniques and Optimization* , 1-9.

Neal, C., Reynods, B., Neal, M., H., W., Hill, L., & Williams, B. (2004). The impact of conifer harvesting on stream water quality: the Afon Hafren, mid-Wales. *Hydrology and Earth System Sciences* , 503-520.

Nova Scotia Environment Water and Wastewater Branch. (2004). *Developing a Municipal Source Water Protection Plan: A guid for Water Utilities and Municipalities.* Retrieved January 17, 2011, from Developing a Municipal Source Water Protection Plan: http://www.gov.ns.ca/nse/water/sourcewater.asp

Odum, H. T. (1956). Primary Production in Flowing Waters. *American Society of Limnology* , 102-117.

Okland, J., & Okland, K. (1986). THe effects of acid deposition on benthic animals in lakes and streams. *Cellular and Molecular Life Sciences* , 471-486.

Onset. (n.d.). *HOBO U20 Water Level Data Logger 30-Foot Depth*. Retrieved October 19, 2010, from Onset Computer Corporation: http://www.onsetcomp.com/products/data-loggers/u20-001-01

Pahl-Wostl, C. (2006). Transitions towards adaptive management of water facing climate and global change. *Water Resources Management* , 49-62.

Pettingill, C. (1933). *Patent No. 1,919,858.* United States of America.

Pollution Probe. (2004). *The Source Water Primer.*

Province of Nova Scotia. (2009, September 08). *Drinking Water Quality and Treatment*. Retrieved March 22, 2010, from Nova Scotia Environment: http://www.gov.ns.ca/nse/water/waterquality.asp

Province of Nova Scotia. (2009, September 8). *Government of Nova Scotia*. Retrieved January 17, 2011, from Developing a Municipal Source Water Protection Plan: www.gov.ns.ca/nse/water/sourcewater.asp

Province of Nova Scotia. (2009, March 20). *Pockwock Lake Watershed Protected Water Area Designation and Regulations.* Retrieved January 15, 2011, from http://www.gov.ns.ca/just/regulations/regs/envpwpoc.htm

Ratkowsky, D., Olley, J., McMeekin, T., & Ball, A. (1982). Relationship Between Temperature and Growth Rate of Bacterial Cultures. *Journal of Bacteriology* , 1-5.

Relations, S. N. (2007). Orthographic images frames 1044800063900, 1044850063800, 104476063800, 1044800063700, 1044800063800 1:10,000.

Robertson, D. M., & Roerish, E. D. (1999). Influence of various water quality sampling strategies on load estimates for small streams. *Water Resrouces Research* , 3747-3759.

Sawyer, C., McCarty, P., & Parkin, G. (2003). *Chemistry for Environmental Engineering and Science.* New York: McGraw-Hill.

Schoor, J. L. (1996). *Environmental Modeling: Fate and Transport of Pollutants in Water, Air, and Soil.* New York: John Wiley & Sons.

Service Nova Scotia and Municiapl Relations. (2007). Orthographic images frames 1044750063800, 1044800063700, 1044800063800, 1044800063900, 1044850063800.

Siegel, F. (2002). *Environmental Geochemistry of Potentially Toxic Metals.* Berlin: Springer-Verlag.

Stirling, R., Aramini, J., Ellis, A., Lim, G., Meyers, R., Fleury, M., et al. (2001). Waterborne cryptosporidiosis outbreak, North Battleford, Saskatchwan. *Can. Comm. Dis. Rep.* , 27.

Tchobanoglous, G., & Schroeder, E. (1985). *Water Quality.* Reading: Addison-Wesley.

Technical Expert Committee. (2004). *Watershed-Based Source Protection Planning.* Toronto: Government of Ontario.

United States Department of Agriculture. (1996). *National Handbook of Water Quality Monitoring.* Washington: Department of Agriculture.

Vandenberghe, V., Goethals, P., Griensven, A. V., Meirlaen, J., De Pauw, N., Vanrolleghem, P., et al. (2005). Applications of Automated Measurement Stations for Continuous Water Quality Monitoring ofthe Dender River in Flanders, Belgium. *Environmental Monitoring and Assessment* , 85-98.

Vanmarcke, E. (1983). *Random fields, analysis and synthesis.* Cambridge: MIT Press.

Wanielista, M., Kersten, R., & Eaglin, R. (1997). *Hydrology: Water Quantity and Quality Control 2nd ed.* New York: John Wiley & Sons.

Wehrly, K. E., & Wang, L. (2007). Field-Based Estimates of Thermal Tolerance Limits for Trout: Incorporating Exposure Time and Temperature Fluctuation. *Transactions of the American Fisheries Society* , 365-374.

Welch, H., Symons, P., & Narver, D. (1977). Some effects of potato farming and forest clear-cutting on small New Brunswick streams. *Fisheries and Marine Service Technical Report* , 745.

World Health Organization. (2008). *Guidelins for Drinking-water Quality.* Switzerland: World Health Organization.

Zhang, Y., Love, N., & Edward, M. (2009). Nitrification in Drinking Water Systems. *Critical Reviews in Environmental Science and Technology* , 153-208.

Zhou, X., & Sujuan, G. (1997). Confidence intervals for the log-normal mean. *Statistics in Medicine* , 783-790.

Appendix A: Time Of Concentration Calculations

Time of concentration is defined by Wanielista et al. (1997) as "the time of equilibrium of a catchment under a steady rainfall excess, or as the longest travel time it takes surface runoff to reach the discharge point of the catchment."

The process below is the NRCS method as designed by Soil Conservation Services in 1986

Lengths and slopes of overland, and river paths were estimated using GIS software.

The time is calculated in three parts; sheet flow, shallow concentrated flow and river flow using the equations below

P_2 was available from Environment Canada, but was also calculated and verified using 30 years of data.

$$tf = \frac{0.0288(nL)^{0.8}}{{P_2}^{0.5}{S_0}^{0.4}}$$

tf = sheet flow time (hours)
P_2 = two-year 24-hour rainfall (cm)
L = flow length (max of 100 meters)
So = slope of path
N = manning's coefficient (0.4 for dense shrubbery and forest litter (Chin 2006))

$$t_{sc} = \frac{L_{sc}}{V_{sc}}$$

$$V_s = k{S_o}^{1/2}$$

tsc = shallow concentrated flow
Lsc= Length of path (m)
Vsc = Velocity (m/s)
K= intercept coefficient (m/s) (0.76 Forest with heavy litter(Chin 2006))

	Lacey Mill River	**Peggy's brook**
For Tf		
n mannings	0.4	0.4
P2 (cm/24hr)	6.7	6.7
Slope	0.017	0.008
length (meters)	100	100
tf (hours)	**1.09**	**1.46**
for tsc		
K	0.76	0.76
Length	330	840
Slope	0.044	0.019
tsc(hours)	**0.58**	**2.22**
For T channel		
L	8380	3200
V	0.228	0.129
Tch	**10.21**	**6.89**
Ttotal	**11.87**	**10.58**

Appendix B: Flow Regression Statistics

Lacey Mill River P*rediction*

The regression equation is

lacey flow moving average = 0.168 + 3.26 peggy flow moving average
- 2.02 12 hour change in flow

39709 cases used, 6073 cases contain missing values

Predictor	**Coefficient**	**SE coefficient**	**T**	**P**
Constant	0.168	0.00124	135.82	0.000
Peggy flow moving average	3.263	0.01013	322.04	0.000
12 hour change in flow	-2.02488	0.01527	132.61	0.000

S = 0.206330   R-Sq = 72.7%   R-Sq(adj) = 72.7%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	4509.3	2254.6	52960.78	0.000
Residual Error	39706	1690.4	0.0		
Total	39708	6199.7			

Source	DF	Seq SS
peggy flow moving average	1	3760.7
change in flow 12 hr	1	748.6

Peggy's Brook Prediction

The regression equation is

peggy flow moving average = - 0.0138 + 0.216 lacey flow moving average
+ 0.493 12 hour change in flow

39884 cases used, 5898 cases contain missing values

Predictor	**Coefficient**	**SE Coefficient**	**T**	**P**
Constant	-0.01384	0.003513	-39.40	0.000
Lacey flow moving average	0.2156	0.000620	347.58	0.000
12 hour change in flow	0.4928	0.002560	192.49	0.000

S = 0.0501977 R-Sq = 80.3% R-Sq(adj) = 80.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	408.61	204.31	81079.69	0.000
Residual Error	39881	100.49	0.00		
Total	39883	509.10			

Appendix C: Filling Extreme Flow Events

There were two events that occurred during the monitoring period that the stage discharge curve was unable to predict flows because the flows were out of the range that the stage discharge curve predicted. This is the process that was followed to obtain the best possible predictions for flows.

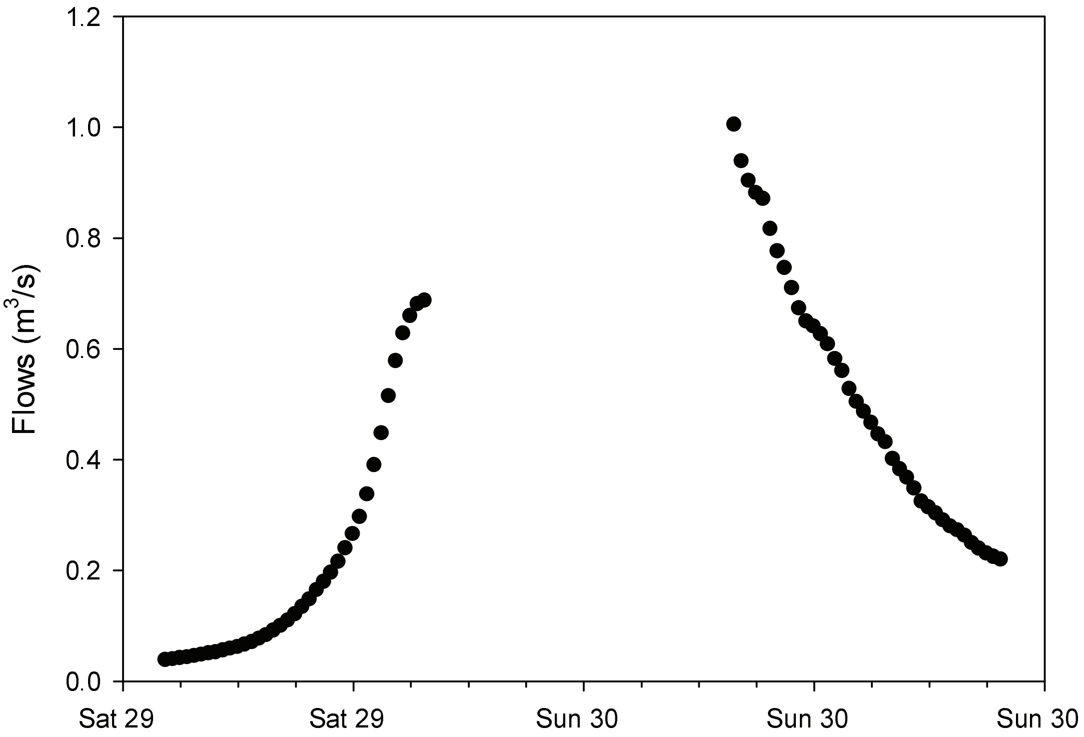

Figure C.1 Initial data of storm event

The rising and falling limb of the storm flow were extrapolated towards the peak of the storm to try and determine where the peak occurred and the level of the peak. The peak was taken to occur when the two extrapolated lines intersect.

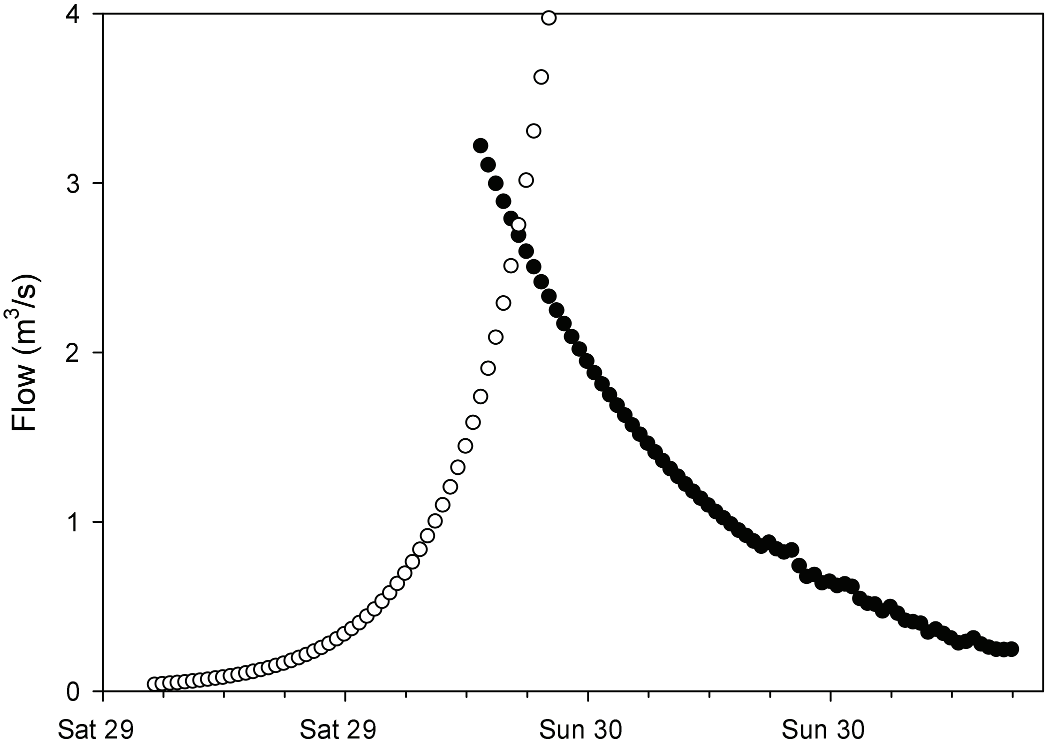

Figure C.2 Extrapolation of curves

The Peak was then shaped to provide a more natural storm flow curve. This was performed graphically using experience from the shape of other storm events.

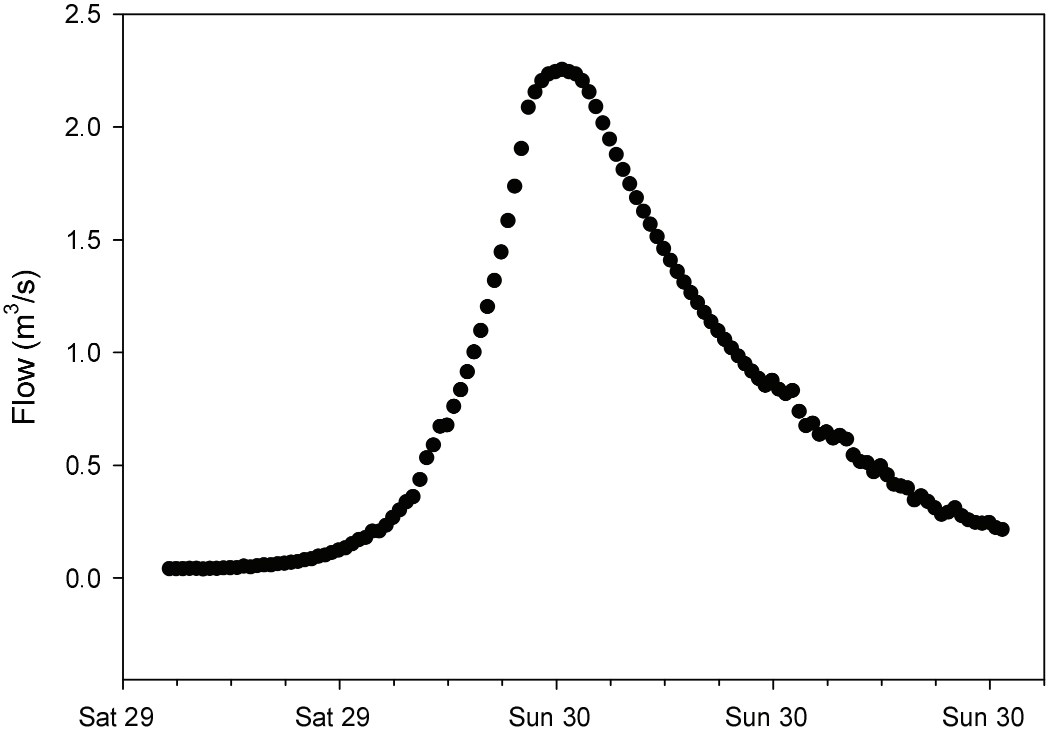

Figure C.3 Final estimated storm flows

Appendix D: Nutrient Loading For Peggy's Brook

Table D.1 Total organic carbon loading on a monthly basis for Peggy's Brook

Month	**Monthly (kg)**	**Bi-weekly (kg)**	**Weekly (kg)**	**Contribution of top 10% of flows**
Jun-09	1600	1500	1300	46%
Jul-09	1600	1500	1400	36%
Aug-09	3500	3200	2500	72%
Sep-09	1500	1500	1100	60%
Oct-09	6900	7000	6100	38%
Nov-09	3300	3200	3500	21%
Dec-09	5800	5800	6100	35%
Jan-10	N/A	N/A	N/A	N/A
Feb-10	N/A	N/A	N/A	N/A
Mar-10	N/A	N/A	N/A	N/A
Apr-10	1400	1400	1300	29%
May-10	200	200	200	36%
Jun-10	2000	1800	1900	53%
Jul-10	7900	3800	2900	90%
Aug-10	800	700	700	58%
Annual Sum (kg)	***26200***	***25600***	***24100***	***48%***

Table D.2 Total phosphorous Loading on a monthly basis for Peggy's Brook

Month	Monthly (kg)	Bi-weekly (kg)	Weekly (kg)	Contribution of top 10% of flows
Jun-09	2.4	2.7	1.2	40%
Jul-09	1.5	1.5	1.7	34%
Aug-09	2.9	2.7	8.4	61%
Sep-09	2.3	3.3	5.4	56%
Oct-09	14.7	11.4	13.2	37%
Nov-09	16.3	5.9	7.3	16%
Dec-09	8.3	8.3	10.0	33%
Jan-10	N/A	N/A	N/A	N/A
Feb-10	N/A	N/A	N/A	N/A
Mar-10	N/A	N/A	N/A	N/A
Apr-10	1.4	1.3	1.8	24%
May-10	0.3	0.4	0.4	36%
Jun-10	1.2	1.1	1.2	50%
Jul-10	2.4	3.2	2.2	86%
Aug-10	1.7	2.5	1.7	61%
Annual Sum (kg)	***47.7***	***34.9***	***48.2***	***48%***

Appendix E: Extreme Value Distribution

The general theory of how to construct an extreme value distribution is transferable but the representative equations for the chosen distribution are different and should be obtained from an appropriate source.

Equation E.1 Shape factor for lognormal distribution extreme values

$$\alpha = \frac{\sqrt{2\ln n}}{\sigma_{lnx}}$$

Equation E.2 Ln of 'characteristic' largest value

$$u_n' = \sigma_{lnx}\sqrt{2\ln n} - \frac{\sigma_{lnx}(\ln(\ln n) + \ln(4\pi))}{2\sqrt{2\ln n}} + u_{lnx}$$

Equation E.3 'Characteristic' Largest Value

$$u_n = \exp\{u_n'\}$$

Equation E.4 Cumulative Extreme Value Distribution

$$F_{Y_n}(y) = \exp\{-\left(\frac{u_n}{y}\right)^{\alpha}\}$$

Calculation of Extreme value distribution for E.coli in Lacey Mill RIver

Assume mean and variance from the 16 month monitoring period are representative.

$\mu_{lnx} = 1.44 \quad \sigma_{lnx} = 1.23$

$$\alpha = \frac{\sqrt{2\ln(65)}}{1.23} = 2.34$$

$$u_n' = 1.23\sqrt{2\ln 65} - \frac{1.23(\ln(\ln(65)) + \ln(4\pi))}{2\sqrt{2\ln(65)n}} + 1.44 = 4.16$$

$$u_n = \exp\{4.16\} = 64$$

This μ_n represents the expected maxima for a 16 month monitoring period (assuming 2 summer in measurement). It was constructed using 16 months of data instead of one year to moderately improve the statistics. This will institute a bias towards representing the summer, but this is when extremes are expected anyways. This distribution is F_{y65} as this is the extreme distribution of the 65 samples collected in the 16 month period.

To construct 5 year extreme value:

$$\mathrm{N} = 5 \text{ years} * \left(12\frac{\text{months}}{\text{year}}\right) * \frac{1}{16 \text{ months}} = 3.75$$

$$\mathrm{P}[\mathrm{y}_{65} > y] = \frac{1}{\mathrm{n}} = \frac{1}{3.75}$$

So

$$1 - \mathrm{F}_{\mathrm{y}_{65}}(\mathrm{y}) = \frac{1}{3.75}$$

$$1 - \exp\left\{-\left(\frac{u_n}{y}\right)^{\alpha}\right\} = \frac{1}{3.75}$$

$$1 - \frac{1}{3.75} = \exp\left\{-\left(\frac{64}{y}\right)^{2.34}\right\}$$

Y = 105 CFUs / 100 ml

Appendix F: Fitting of a Deterministic Seasonal Model

The seasonal deterministic model was fit by minimizing the residual associated with the model. In order to recreate the results a program capable of optimizing nonlinear problems is required. Matlab and Excel's Solver both have this capability.

$$A * \sin\left(2 * \frac{\pi}{n} * (x - b)\right) + c$$

$\mathrm{N} =$ time steps in the year (35040 for 15 minute time steps)
$\mathrm{C} =$ vertical offset (mean)
$\mathrm{B} =$ the shift of the wave from the first measurement to the time interval which is representative of the mean in # of intervals
$\mathrm{A} =$ Amplitude of wave

The equation is setup to fit A, B and C by minimizing the square of the residuals, where the residuals are: the measured data – the deterministic model.