# Toward understanding fecal coliform and *Escherichia coli* distributions in the Annapolis River

***Honours Thesis Research Project***
***By Caitlyn MacMaster***

Submitted Friday, 11 April, 2008
In fulfillment of the requirements for ENVS 4902
At Dalhousie University, Halifax, Nova Scotia

**ABSTRACT:**

The Clean Annapolis River Project (CARP) has identified fecal bacterial contamination as the single greatest threat to water quality in the Annapolis River, located in western Nova Scotia. The intense agriculture and use of private wells within the Annapolis River watershed make bacterial contamination an issue of particular concern. This *post hoc* exploratory study examines the relationships between concentrations of bacterial indicators and a variety of environmental factors. Concentrations of *Escherischia coli* and fecal coliform bacterial were transformed using the natural logarithm and were compared and contrasted based on differences in sampling location, month, year, rain rates, temperature, and pH, using box-and-whisker plots, analysis of variance (ANOVA), and Student's *t* tests. A step-wise multiple regression was then performed to determine the level of interaction between the predictor variables, which environmental factors had the strongest influence, and the total amount of variation that could be explained by the study. All predictor variables with the exception of pH demonstrated some statistically significant influence on logarithmically transformed bacterial concentrations. The multiple regression analysis indicated that rain a few days prior to sampling, sampling location, month, and water temperature were the strongest predictors of transformed concentrations, and collectively accounted for 21.3% of variability in the transformed concentrations.

**AUTHOR:**

Caitlyn A. MacMaster
Honours Environmental Science Candidate, 2008
Dalhousie University,
Halifax, Nova Scotia, Canada

E-mail: cxmacmaster@dal.ca

**SUPERVISOR:**

James C. Boxall, B.A., B.Ed., M.A., MLIS, FRGS
GIS Centre Director, Map Collection Curator
Adjunct Professor, School of Planning, School of Resource and Environmental Studies
Dalhousie University
Halifax, Nova Scotia, Canada

E-mail: James.Boxall@Dal.Ca

**INSTRUCTOR:**

Karen Harper, B.A., M.Sc., Ph.D.
Assistant Professor, Environmental Programmes Department
Honorary Adjunct Professor, School of Resource and Environmental Studies
Dalhousie University
Halifax, Nova Scotia, Canada

E-mail: Karen.Harper@Dal.Ca

# ACKNOWLEDGEMENTS

In addition to the assistance of James Boxall, my supervisor, and Karen Harper, the Honours Thesis Coordinator, I'd like to acknowledge the several other people for their contributions to the success of this project. Firstly, I'd like to thank Daniel Rainham, Associate Professor with Environmental Programmes at Dalhousie University, for his wonderful guidance and advice on this project. It truly has made a world of difference! I'd also like to thank Dr. W.C. Hart of the Water Resources Centre, Faculty of Engineering, Dalhousie University for reading my draft and providing comments. I'd also like to express my gratitude and appreciate for the assistance and cooperation of Andy Sharpe, Water Quality Researcher with the Clean Annapolis River Project (CARP), for providing the data record upon which this analysis is based. The final product just wouldn't be the same without you!

Thanks also to Jennifer Strang, GIS Analyst with the Maps and Geospatial Information Centre (MAGIC) at Dalhousie University, for assistance in acquiring data inputs for GIS mapping. Thanks to Gavin King, fellow Honours Environmental Science candidate, for GIS mentorship and to Vaneeta Kaur Grover, Statistical Consultant with the Department of Mathematics and Statistics, Dalhousie University, for tutelage in non-parametric statistical analysis (even though it turned out I could use parametric analyses!).

Last, but certainly not least, thanks to my family and friends – Mum, Dad, Kimbe, Matt, and so many more – for the support and encouragement and for acting as sounding boards for my ideas. Your inputs have made a great deal of difference to me and this project!

# TABLE OF CONTENTS

# TOPIC INTRODUCTION

The following report presents an exploratory study into water quality in Nova Scotia's Annapolis River. The study is aimed at addressing the role of various environmental factors in influencing the distribution of fecal coliform and *Escherichia coli* (*E. coli*) bacteria within this particular surface water body. A deeper understanding of these influences can then potentially be harnessed to improve the efficacy of best management practices and various projects aimed at improving water quality, simultaneously contributing to greater gains for lesser capital investments by targeting those projects appropriately.

The primary objectives of this research are to quantify the relationships between *E. coli* and/or fecal coliform concentrations and various environmental factors, including water temperature, pH, rain rate, and geographic location. Seasonal and long-term trends are also examined. The end goal of this research is to promote the local community's understanding of *E. coli* and/or fecal coliform distributions and to thereby support the continued management efforts of community groups attempting to protect the health of the region's populations – both human and non-human.

Water plays a significant role in human health and well-being. The first words of Maggie Black's book, *The No-Nonsense Guide to Water*, state: "Water is key to the survival and growth of all human, animal and plant life and all economics and environmental processes" (Black, 2004). Three days without hydration is a serious health risk to humans, but our need for water goes beyond direct consumption (USFNR, *n.d.*). Peter Gleick, co-founder of the Pacific Institute for Studies in Development, Environment and Security, estimates that the average person requires a minimum of 50 litres of water per day: 5L for direct consumption, 10L for food preparation, 15L for bathing, and 20L for other sanitation needs (Gleick, 1998). However,

domestic use only accounts for 8 to 11 percent of water usage in low to medium and high income countries, respectively (Black, 2004). Industry demands 10% of water consumed in low- and middle-income countries and 59% in high-income countries (Black, 2004). Agriculture is still the primary sink for freshwater resources, demanding 70% of the worldwide resources; this ranges from 30% in high-income countries to 82% in low- to medium-income countries (Black, 2004). During the growing season in the Annapolis Valley, agriculture is the single largest water use sector, accounting for 39% of water diverted from the Annapolis River between May and September (Timmer, 2003a). Based on 1999 data for a 150-day growing season, greenhouses consumed an average of 197 019 L of water per day, with orchards, livestock operations, and field crops using somewhat less, consuming, on average, 143 296 Lpd, 82,397 Lpd, and 78,263 Lpd, respectively (Timmer, 2003a).

Not only do humans require water for personal use, as Gleick suggested, but it is also necessary for the production of our food – both plant and animal matter. This is especially important when one considers that the explosion of the human population in recent decades has only been sustained through increased agricultural yields, which have been achieved primarily through irrigation. In 1983, 13.7% of the world's available cropland was irrigated and this proportion accounted for 41.3% of the food produced (Long, 1983). Alternatively, the 86.3% of the total cropland that was not irrigated contributed 58.7% of the world's food supply (Long, 1983). Irrigation can have considerable impacts on water quantity because demand for irrigation water peaks concurrently with the months of lowest streamflows (Timmer, 2003a). In fact, a rapid drainage regime paired with climatologically dry summers produced drought conditions in the Annapolis River in 4 of the six years between 1996 and 2002 (Timmer, 2003b). This increased the appeal of using groundwater for human consumption, as well as for

agricultural practices (Timmer, 2003b). Recent shortages of water during the summer months have provided additional incentive for further research – predominantly by the provincial government – into local management activities for water supplies (Timmer, 2003aa).

Given that the human population is expected to continue to increase rapidly, it is only logical to suppose that demand for food will increase and that this demand will be met with greater irrigation of cropland, to increase yields. This will undeniably result in an increased need for freshwater. However, because we have such a great dependence on freshwater, we are subject to great risk through its scarcity or contamination.

In 2003, the United Nations announced that eighty percent of all illnesses and deaths in the developing world were attributable to consumption of or exposure to unsafe water – killing a sum of 25 000 people per day (Annan, 2003). While this threat is not so dire in the developed world, thanks to better healthcare, water-borne diseases still pose a significant threat to human welfare. For example, the pathogens responsible for ear infections, dysentery, typhoid fever, viral and bacterial gastroenteritis, meningitis, and Hepatitis A are all waterborne and associated with the presence of fecal coliforms (Sullivan, 2004; Ferguson, de Roda Husman, Altavilla, Deere, and Ashbolt, 2003). Furthermore, *Salmonella, Campylobacter* and *Shigella* bacteria are associated with fecal coliforms, as are the protozoa *Giardia,* and *Cryptosporidia* (Sullivan, 2004; Ferguson, de Roda Husman, Altavilla, Deere, and Ashbolt, 2003). While it would be costly and time consuming to test for each of these possible pathogens, it is far simpler and more economical to test for the presence of fecal coliforms or *E. coli* directly (Ferguson *et al.*, 2003; USEPA, 2006). Fecal coliforms and *E. coli* are therefore referred to as bacterial indicator species, as their concentrations are used as indices of the potential for exposure to other pathogens. Ergo, the presence of high concentrations of fecal coliforms in surface waters is not

itself dangerous, but does imply an elevated risk of exposure to associated pathogens (Synova, 2006; USEPA, 2006). Accordingly, the occurrence of high levels of fecal coliform bacteria in water samples has led to the closure of beaches and shellfish harvesting areas, resulting in strong social and economic impacts (Sullivan, 2004). Because humans are not the only ones using these water resources, the impacts of high concentrations of fecal coliforms could also affect other species (Sullivan, 2004).

*E. coli* can also serve as an indicator for the presence of these harmful microbes. However, unlike fecal coliforms, *E. coli* may actually provide a threat in and of itself. There are hundreds of strains of E. coli, most of which are innocuous and habitually live in the intestines of healthy humans and animals (CCME, 2002). However, some strains – such as *E. coli* O157:H7, which was implicated in an outbreak in Walkerton, Ontario in 2000 – can have considerable deleterious effects on human and animal health (CCME, 2002). When humans (or other animals) ingest water or food contaminated with *E. coli*, they may experience watery or bloody diarrhea, fever, abdominal cramps, nausea, and vomiting (CDCP, *n.d.*). Symptoms tend to be more severe in young children, the elderly, and the chronically ill and may include kidney failure and death (CDCP, *n.d.*; CCME, 2002). Irrigation of cropland with contaminated water and improper management of manure fertilizer have been linked to *E. coli* infections in humans (UWOFMD, 2007). *E. coli* poisoning may also occur if humans ingest shellfish that inhabit contaminated water. Given these deleterious effects, contamination of water sources with *E. coli* should be of high concern and is therefore a common component of water quality monitoring programs.

Water quality in the Annapolis Valley can have far reaching effects, as the agricultural products of the valley are shipped throughout Canada and internationally. The Annapolis Valley

is one of Canada's three most fertile agricultural regions – along with the Holland Marsh in Ontario and British Columbia's Okanagan Valley (McMullen, 2007). The surface waters of the Annapolis River and its tributaries drain the south-western two thirds of the Annapolis Valley and are used for recreation, livestock watering, crop irrigation, hydroelectric power generation, and occasionally human consumption (Timmer, 2003a). Furthermore, many of the Valley's rural residents rely on private groundwater access for their domestic uses (Timmer, 2003a).

In addition to the communities of humans and domesticated animals that rely on the waters of the Annapolis Valley, natural communities also require a minimum quantity and quality of available ground- and surface-water. For example, thousands of migratory waterfowl populate bird sanctuaries throughout the valley (Timmer, 2003a). This may also influence water quality as, for example, the sanctuary above the Kentville Bridge is believed to contribute significant fecal coliform contamination to the Cornwallis River (which drains most of the eastern third of the Annapolis Valley) (Timmer, 2003a).

In summary, the Annapolis River provides a prime subject area for research into *E. coli* and fecal coliform research, because the agricultural activities both contribute to and are threatened by water contamination. Furthermore, dedicated community groups are already in place to support mitigation and management practices. Water quality is a key issue of concern and it is hoped that a better understanding of factors influencing *E. coli* and fecal coliform distribution will support continued efforts to minimize risks of exposure within the vibrant communities of Nova Scotia's Annapolis Valley.

# PROJECT SIGNIFICANCE

In a press release announcing 2003 as the International Water Year, the United Nations wrote that "[t]he availability of clean, fresh water is one of the most important issues facing humanity today – and will be increasingly critical for the future, as growing demands outstrip supplies and pollution continues to contaminate rivers, lakes and streams" (UNIS, 2002). Furthermore, the Clean Annapolis River Project (CARP) – a not-for-profit non-governmental organization that has been monitoring water quality within the Annapolis River watershed since 1992 – has acknowledged that the single greatest threat to water quality in the Annapolis River is contamination by fecal bacteria (Synova, 2006; CARP, 2007). Thus, this research project directly addresses the maintenance and protection of water quality in the Annapolis River by investigating influences on *E. coli* and fecal coliform distribution.

Within the Annapolis River watershed, CARP has been concerned with the condition of the river for the past two decades. Amongst its various projects, the organization records water quality data at several locations throughout the river's course through the Annapolis River Guardians (ARG) project. CARP has also implemented various projects aimed at restoring stream-banks and aquatic habitat, promoting ecologically sensitive farming, and ensuring ongoing stewardship of riparian land. The aim of these undertakings is to minimize the impact of riparian activities – including agriculture – on the surface waters of the river and to thereby protect water quality. It is hoped that the additional research presented herein will help those living and working on or near the Annapolis River to understand the distributions of *E. coli* and fecal coliform within the river and thereby contribute to improved efficacy of future management projects by identifying specific target areas and influences.

This research may also be of interest to various other groups that have also been active in addressing water quality and quantity issues within the Annapolis River watershed. The Nova Scotia Federation of Agriculture has partnered with departments of the provincial government to produce the Environmental Farm Plan Program, which is a voluntary program for farmers seeking assistance in assessing and reducing environmental risks on their property, including management of water quantity and quality (Timmer, 2003b). The Nova Scotia Agricultural College has produced a variety of fact sheets for farmers on soil and water conservation, in partnership with Horticulture Nova Scotia (Timmer, 2003b). The Growers Water Group has formed as a collection of representatives from agricultural production groups that are particularly interested in sources of irrigation waters for the Annapolis Valley, among other water issues associated with agriculture (Timmer, 2003b). Any of these groups may find that this Honours Thesis Project provides supportive supplementary material to be used in their various endeavours for environmental stewardship and management of agricultural pollution.

This research may also support the department of Agriculture and Agrifood Canada's new focus on best management practices (BMPs, also beneficial management practices). A better understanding of the relationship between the distribution of potentially noxious microbes and environmental conditions (including seasonal and geographic distributions, relationships with temperature, pH, etc.) can help to determine which actions may be taken in the future to minimize associated risks. This research is intended to support such projects and improve their effectiveness.

**Case Study: Walkerton, Ontario**

The following case study demonstrates the potential negative effects of an *E. coli* outbreak in drinking water in the town of Walkerton Ontario.

In May 2000, the rural Ontario village of Walkerton become host to Canada's worst-ever outbreak of *E. coli* contamination, which was primarily attributed to animal manure (Ferguson *et al.*, 2003; CBC News Online, 2004). In a town of only 5000 residents, half of those were affected, almost 900 were hospitalized, and nine deaths were attributed, either directly or indirectly, to *E. coli* in drinking water (Gloubeman, 2001). The ramifications of this event were not limited to this rural town, but were felt across the province of Ontario, as residents began to question their water quality. This caused an upset in the political arenas, cost the province $3.5 million in legal fees and $9 million to fix the town's water supply, resulted in an estimated $2.7 million of lost revenue, and caused a $1.1 million drop in total real estate values for the region (CBC News Online, 2004). From the Walkerton example it becomes apparent that *E. coli* contamination can have broad-reaching negative effects. The impacts described were strictly those felt by human populations affected by *E. coli*. Once the health of populations of other species is included, the effects are even greater.

The potential negative impacts associated with fecal contamination of water bodies serve as a strong motivation to implement mitigation measures to protect against E. coli and fecal coliform contaminations. Such measures are more likely to be effective if the relationship between these bacteria and various environmental variables are better understood. This is the motivation for the enclosed research.

# LITERATURE REVIEW

The Clean Annapolis River Project (CARP) has been monitoring water quality within the Annapolis River since 1992 as part of its Annapolis River Guardians (ARG) program – one of thirteen programs funded in part by the Atlantic Coastal Action Program (ACAP), established in 1991 by Environment Canada (Timmer, 2003c). CARP has also has partnered with local farm owners on stewardship projects, addressing aquatic habitat restoration and environmentally sensitive agricultural practices, among other things. The Annapolis River Guardians monitoring program has identified that the greatest environmental concern for the river is the abundance of fecal coliform bacteria, which indicate an increased risk of exposure to pathogenic bacteria (Synova, 2006; CARP, 2007). In order to target mitigation efforts to minimize this concern, it is important to understand the distribution and behaviour of the target organisms. Thus, the following sections will address the use of indicator organisms and the source, transport, and fate of microbes in ground and surface waters.

## Fecal Coliforms and *E. coli*

Fecal coliforms (sometimes written faecal coliforms) are a category of colony-forming bacteria that are typically associated with the intestinal tracts of humans and other warm-blooded animals, such as cattle and hogs (Synova, 2006). The genus *Escherichia*, including the species *Escherichia coli* (abbreviated *E. coli*), originates in fecal matter, but other fecal coliforms, such as the genera *Enterobacter, Klebsillia,* and *Citrobacter* may or may not be associated with fecal matter (Doyle and Erickson, 2006). Potential sources of fecal coliform in surface waters include wastewater treatment plants, malfunctioning on-site septic systems, domestic and wild animal manure, and storm runoff (CCME, 2002; Sullivan, 2004; USEPA, 2006). Submerged wood and

the effluent from pulp-and-paper or textile manufacturing facilities may also contribute coliforms of non-fecal origin – notably the genus *Klebsiella* (USEPA, 2006; Doyle and Erickson, 2006).

As previously mentioned, fecal coliform bacteria are not typically harmful in their own right, but rather are used as indicators for the possible presence of pathogenic bacteria, viruses, protozoa, and other parasites that also originate in the digestive tracts of host organisms (Synova, 2006; USEPA, 2006). It is much less time intensive and less costly to test for fecal coliforms than for each pathogen directly. *E. coli* can be tested for with similar ease and cost-effectiveness and also serves as an indicator for the potential presence of other pathogenic organisms originating from mammalian gastrointestinal systems. However, it is important to note that with both of these indicators, a considerable degree of uncertainty remains about the prevalence of the pathogens of concern. Multiple studies have shown that the concentrations of culturable indicators, such as fecal coliform, do not necessarily correlate well with the actual abundance of pathogenic microorganisms (Ashbolt, Grabow, and Snozzi, 2001; Grabow, 2001). This is primarily due to the sporadic incidence of infection in the human and animal populations that are the source of the fecal matter, compared to a relatively constant supply of fecal matter – i.e. fecal coliforms are contributed much more regularly than pathogenic microorganisms (Ferguson *et al.*, 2003). To some degree, the precautionary principle is applied here, whereby the elevated risk of exposure to associated pathogens (characterized by high concentrations of indicator organisms) is used in place of data pertaining directly to the concentration of target microbes

The Canadian Water Quality Standards (CWQS) identify four threshold concentrations for fecal coliforms, depending on the use of the water. Concentrations greater than 200 colony forming units per 100 ml water sample (cfu/100ml) are deemed “unacceptable for water-contact recreation” (Synova, 2006). Concentrations greater than 100 cfu/100ml are considered

"unacceptable for food crop irrigation," while concentrations in excess of 50 cfu/100ml are labeled "unacceptable for livestock watering" (Synova, 2006). The Canadian Water Quality Standards identify 0 to 50 cfu/100ml as "minimal bacteria detected," but Health Canada's Guidelines for Canadian Drinking Water specify that bacteria content should be zero for water consumed directly by humans (CCME, 2002; Synova, 2006).

However, fecal coliforms and *E. coli* – both of which are bacteria – have been shown to be unsuitable for characterizing the risk presented by protozoan pathogens and some enteric viruses, due to the differences in their biology (Ashbolt *et al.*, 2001). Thus, outbreaks of cryptosporidiosis have been documented even in areas where microbiological standards for water quality were met based on bacterial indicators (Ferguson *et al.*, 2003).

## Sources of Fecal Coliform in the Annapolis River

Origins of enteric pathogens in surface water bodies may be point or diffuse sources and are typically difficult to pinpoint, due to the transport over land and through water (Ferguson *et al.*, 2003). To better delimit source regions, the Clean Annapolis River Project coordinated a Sub-Watershed Investigative Monitoring (SWIM) program to research the contributions of fecal coliforms from various tributaries within the Annapolis River. It was hoped that this would help prioritize source regions and thus help target best management practices to where they would have the greatest impact. The program was successful in identifying that the mean fecal coliform contribution over the sampling period (July 6 to November 12, 2004) from the Fales River was roughly twice as great as the second highest contributor (of the six tributaries studied) (Sullivan and Sharpe, 2005). The results are summarized in **Table 1** (see **Appendix A**).

While the SWIM program managed to identify relative importance of source regions, it still did not identify the ultimate source of the fecal contamination. A developing approach

known as Microbial Source Tracking (MST) has been proposed to assist in the identification of contamination sources. There are a variety of techniques that fall under the MST umbrella, several of which can discriminate between human and non-human sources, and one technique is purported to be able to further delimit the source of fecal pollution between cattle and hogs (Sullivan, 2004). Though this technique has not yet been applied in the Annapolis River watershed, it is CARP's hope that the direct tracking of fecal pollution sources by MST would allow for better targeting of remediation efforts – with focus on the primary contributor species – thereby saving time and resources (Sullivan, 2004).

A direct relationship has been observed between the presence of cattle on agricultural land and increased concentration of fecal coliform concentrations in nearby surface waters (Tiedemann, Higgins, Quigley, Sanderson and Marx, 1987). It is suggested that animal access to streams might have been a greater contributor to fecal coliform concentrations than do stocking densities (Tiedemann *et al.*, 1987). The Clean Annapolis River Project has addressed the watershed access issue directly through its Aquatic Habitat Restoration projects of 2006 and 2007. As part of these works, CARP partnered with local land owners on eleven different agricultural properties to raise awareness and understanding of the impacts of current agricultural practices on water quality and aquatic habitat and to install a total of 5 295 m of fencing along the river and its tributaries to restrict cattle access to the waterway (Cliche and Gaul, 2007).

**Fate and transport of E. coli and fecal coliforms**

Event-related increases in concentrations of waterborne pathogens are related to surface run-off, but the process by which pathogens from fecal deposits excreted to land are dispersed, transported, attenuated, and inactivated is poorly defined (Ferguson *et al.*, 2003).

The use of indicator species, such as *E. coli*, as indices for the probable presence of other pathogens requires the survival of these organisms long enough to be measured, but demands that they do not grow and proliferate after deposition (egestion from the host's intestinal tract) (Ferguson *et al.*, 2003). The survival of both indicators and their associated pathogens in feces, soil, and ultimately water depends on a variety of environmental conditions, including the physical, chemical, and biological composition of each matrix (Ferguson *et al.*, 2003). These various influences are described in the following sections.

**Overland Transport**

Ferguson *et al.* (2003) wrote that "[w]ater flow is one of the most important parameters affecting transport of microbial contaminants." The occurrence of rainfall events has been correlated with increased microbial loading of surface waters, supporting the hypothesis that microorganisms are transported in surface run-off (Curriero, Patz, Rose, and Lele, 2001). Field experiments have shown peak *Cryptosporidium* oocyst release after 20 minutes of heavy rainfall, which could reasonably be expected to hold for other fecal microbes (Ferguson *et al.*, 2003). The same study found that the slope of the field plots had a significant effect on overland transport – i.e. increased elevation was associated with greater transport (Ferguson *et al.*, 2003).

Soil type has also been shown to strongly influence immobilization of bacteria from leaching and surface run-off (Ferguson *et al.*, 2003). Adding to that effect, shade-producing vegetation can extend fecal coliform survival by shielding from direct sunlight and lowering ambient temperatures (Ferguson *et al.*, 2003).

Ferguson *et al.* (2003) report that "there is practically no data on the efficacy of vegetated buffer strips for attenuating overland pathogen movement." However, filtration through wetlands was shown to decrease fecal coliforms in secondary sewage effluent by an average of

98.2% (Thurston, Gerba, Foster, and Karpiscak, 2001). Similarly, *Giardia* cyst and *Cryptosporidium* oocyst concentrations declined by 87.8% and 64.2% respectively in the same study (Thurston *et al.*, 2001). Thus, wetlands may play a vital role in the diminution of concentrations of harmful micro-organisms and their associated indicators in surface waters.

## The Role of Climate Change

Global climate change is associated with more erratic precipitation distribution, with a disproportionate increase in heavy rainfall events (Ferguson *et al.*, 2003). Ferguson also reports that the greatest exposures to fecal coliforms and *E. coli* "occur sporadically, particularly after large hydrologic events or during other exceptional circumstances" and that storm events have been linked to increases in pathogen loading by several orders of magnitude (Ferguson *et al.*, 2003). Studies have demonstrated that incidence of increased rainfall is correlated to waterborne disease outbreaks in the United States, between 1948 and 1994 (Curriero *et al.*, 2001). Thus, it can be anticipated that the issue of water quality and particularly indicator species for pathogens will become increasingly important as the effects of global climate change manifest.

## Factors Influencing Microbial Survival in Surface Waters

The survival of *E. coli*, fecal coliforms, and their related pathogens in surface water is most greatly impacted by temperature (Ferguson *et al.*, 2003; Crane and Moore, 1986). Elevated temperatures cause rapid inactivation and die-off of many microorganisms, especially in concert with dessication (Crane and Moore, 1986). It has been shown that coliform bacterial survival is inversely proportional to temperature in the range of 5 to 15 $^{o}$C (McFeters and Stuart, 1972). Furthermore, both visible and ultraviolet wavelengths are lethal to enteric pathogens and their indicators (including fecal coliform and *E. coli*) after sufficient exposure (Davies and Evison,

1991; Davies-Colley, Donnison, and Sped, 2000). However, some cells have UV repair mechanisms that can reduce impact and promote survival by some of the exposed population – a trait that is particularly prevalent among viruses (Davies-Colley *et al.*, 2000).

For *E. coli* organisms stressed by visible light, the typical survival strategy is to become progressively more dormant viable but non-culturable cells (VBNC) (Ferguson *et al.*, 2003). Ogden, Fenlon, Vinten, and Lewis (2001) report that the die-off rate of pathogenic *E. coli* strain O157 is approximately equivalent to that of other strains; therefore the evolution of the total *E. coli* concentration is a suitable indicator for changes in pathogenic concentrations. Similar to the indicator species *E. coli*, the inactivation rates of *Cryptosporidium* oocysts, *Giardia* cysts, and *Salmonella* – which are contributed through livestock feces – have been shown to be more rapid in marine waters under conditions of exposure to natural sunlight (Ferguson *et al.*, 2003; Johnson, Entriquez, Pepper, Davis, Gerba, and Rose, 1997). Other fecal coliforms are less stressed by visible light when in the presence of humic substances, but may be more sensitive to lowered pH (Ferguson *et al.*, 2003). However, most fecal coliforms and their potentially associated pathogens are generally pH-tolerant, as demonstrated by their survival and proliferation in gastrointestinal systems of host organisms (Ferguson *et al.*, 2003).

It has been shown that the addition of an exogenous nitrogen source increased survival of *E. coli* bacteria in aquatic environments, even in the presence of competing microorganisms (Ferguson *et al.*, 2003). Thus, sewage outflow and surface run-off not only serve as sources of pathogens and their indicators, but also of nutrients that promote the survival of those microbes.

**Interactions with Soil and Sediment**

Microorganisms typically survive longer in soils and sediments than in surface waters, as they are protected from threats such as sunlight. Thus, soils and sediments may act as reservoirs

for microorganisms (Grimes, 1975; Davies and Evison, 1991; Ferguson *et al.*, 2003). Field studies in coastal areas of Texas and Florida found that sediments contained ten to ten thousand times greater concentrations of enteroviruses (on a volume basis) than in overlying seawater (Gerba, Smith, and Melnick, 1977; Labelle and Gerba, 1980). This ratio also applies to virus concentrations in sediment near outfalls of treated sewage (Ferguson *et al.*, 2003). As long as the microorganisms remain adsorbed to the sediments, they do not pose a significant public health threat (Ferguson *et al.*, 2003).

Based on the preferential affinity of microorganisms for sediment particles as compared to water, percolation of microbiologically contaminated water through the soil matrix may be an effective way to remove pathogens. As water travels through interstitial pores in the soil structure, sorption/desorption, inactivation, and predation can reduce the concentration of bacteria, protozoa, and viruses (Ferguson *et al.*, 2003). Laboratory experiments have demonstrated that increased hydraulic flow (reducing contact and retention time within the soil matrix) corresponds to increased concentrations of *E. coli* in the effluent, and *vice versa* (Stevik, Ausland, Hanssen, and Jenssen, 1999). Requisite subsurface travel distances depend on the size, adsorption characteristics, degree of inactivation of the targeted pathogen, and the velocity of ambient groundwater (Ferguson et al., 2003). Conditions of lower pH, higher ionic strength, and greater organic carbon content in the soil tend to enhance viral attachment to soil particles and thus slow their transport, based on laboratory experiments (Ferguson *et al.*, 2003). For bacteria, size, hydrophobicity, and electrostatic charge of the cell surface influence transport and are all subject to change based on nutrient availability to the cell (Ferguson *et al.*, 2003).

Soil percolation is deemed an excellent barrier against bacteria and protozoa, but is slightly less effective in protecting against contamination by viruses, which are typically two to

three orders of magnitude smaller (Ferguson et al., 2003). Germany and the Netherlands have typically identified minimum subsurface travel times on the order of 50 to 60 days, which was thought to be sufficient to inactivate pathogenic microorganisms (Ferguson et al., 2003). However, recent insight into the persistence of *Cryptosporidium*, *Giardia*, and some enteric viruses suggests that this travel time might be insufficient (Ferguson et al., 2003).

It has also been reported that enteric bacteria and viruses survive longer in groundwater than in surface water, due to lower temperatures, protection from sunlight, and limited microbial antagonism (Ferguson *et al.*, 2003). They have been shown to travel 1000m to 1600m in channeled limestones and 250 to 408 m in glacial silt-sand aquifers (Ferguson *et al.*, 2003). The USEPA standard set-back distance is 30.5m, but this may be ineffective for certain hydrogeological settings (Ferguson *et al.*, 2003). In this way, groundwater may become contaminated with fecal microbes, just as surface water is susceptible.

Pathogens are removed from groundwater flows by binding to soil particles, but they may similarly bind to fluvial sediments, where they are believed to be capable of multiplication (Ferguson *et al.*, 2003; Sullivan, 2004). Studies have demonstrated that viruses bound to sediments may provide a source of infectious viruses that are capable of re-entering the water system for up to 230 days after their initial exposure (Ferguson *et al.*, 2003). Furthermore, sediment resuspension, triggered by motorboat activity, swimming, water currents, or wildlife activity, can increase microorganism concentrations in the overlying water through desorption (Ferguson *et al.*, 2003). This occurs more easily in shallow waters, such as those of the upper reaches of the Annapolis River, than in deeper freshwater lakes or open marine environments (Ferguson *et al.*, 2003). Resuspension may also be associated with meteorologic events of high rainfall, amplifying the effects of increased surface run-off. Desorption rates may also be

influenced by changes in water chemistry, such as pH, as viruses have been shown to aggregate at acidic pH but not at alkaline pH values (Ferguson *et al.*, 2003). Thus, anthropogenic activities influencing river characteristics can indirectly affect concentrations of water-borne pathogens.

Relative microorganism prevalence in soil or sediment versus pore or overlying water is influenced by various physico-chemical properties of soil. For example, grain size can affect the amount of surface area available for adsorption sites, with more surface area corresponding to greater adsorption of *E. coli* and thus, reduced bacterial transport (Stevik *et al.*, 1999). Such physical characteristics can help explain why topsoil provides a more favourable environment for fecal coliform persistence (surviving 7 to 20 days), as compared to pasture or subsoil (2 to 6 days) (Thelin and Gifford, 1983).

Another example of a relevant soil characteristic is its composition – the relative abundance of sand, silt, clay, and organic matter. Clays have negatively charged surfaces, which increase their microorganisms adsorption efficacy, as well as their ability to retain organic matter (Ferguson *et al.*, 2003). The greater concentration of nutrients and organic matter in this soil type promotes *E. coli* and *Salmonella* survival directly, while the clay soils protect microorganisms from environmental stressors, such as solar radiation, pH extremes, dessication, antibiotics, and predation (Ferguson *et al.*, 2003). Sands and organic soil materials are comparatively poor adsorbents, especially in the presence of high pH, high concentrations of dissolved solids, or divalent cations (Ferguson *et al.*, 2003). Thus, soil type is an important determinant in adsorptive capacity.

Ferguson *et al.* (2003) suggest that that microorganism survival can be expected to be hampered near the soil-air interface, as groundwater is scarcer and drying adversely affects their survival. However, despite the threat of dessication, the drying of soils between rainfall events

does not significantly reduce leachate trends for *E. coli* (Gagliardi and Karns, 2000). In fact, it has been found that the intermittent wetting from rainfall events was sufficient to sustain *E. coli* within surface soils, and that rapid growth of these populations could actually occur when soils were saturated during precipitation events (Gagliardi and Karns, 2000).

**Indicator Regrowth**

As previously mentioned, the utility of indicator species requires their persistence in the environment, but necessitates that their populations do not continue to grow after release (Ferguson *et al.*, 2003). This presumes that the same population growth behaviour is typical of the pathogens with which the indicator species are associated. It has been reported that viruses may become bound to sediment and those individuals subsequently act as a source of infectious organisms that may continue for up to 230 days (Ferguson *et al.*, 2003). This behaviour may or may not be similar in their bacterial indicator species (fecal coliforms and *E. coli*), as adsorption and desorption characteristics vary from one microorganism to the next based on morphological characteristics (Ferguson *et al.*, 2003).

**Agricultural Waste Management Practices**

Livestock excreta are identified as a primary source of pathogens on agricultural land (Khaleel, Reddy, and Overcash, 1980). Domestic cattle and sheep can be major sources of *Campylobacter*, *Salmonella*, and enterohemorrhagic *E. coli* to a watershed (Ferguson *et al.*, 2003; Jones, 1999). Even some clinically healthy cattle can contribute these pathogens (Ashbolt *et al.*, 2001).

More intensive farming practices have led to changes in disposal methods for animal wastes. Whereas waste material was traditionally collected as solid waste on straw bedding, it is

now more common for waste to be collected as semi-solid slurry (Ferguson *et al.*, 2003). Furthermore, this slurry is no longer composted, whereby heat pasteurization and the biotic effects of aerobic activity can contribute to declines in pathogenic populations, but is stored under conditions more favourable to the persistence of enteric pathogens (Ferguson *et al.*, 2003).

The application of manure as a fertilizer and the storage of animal wastes on-site can elevate the risk of fecal contaminants leaching into groundwater (Ferguson *et al.*, 2003). A farm groundwater survey conducted in Ontario indicated that contamination of well-water is inversely proportional to distance from feedlots or exercise yards and that higher risk wells were located near pastures that spread manure (Conboy and Goss, 2000). This is pertinent in the Annapolis Valley, where many residents rely on private wells for drinking water (Timmer, 2003a). Also, the adoption of no-till planting practices, intended to reduce soil erosion and increase organic matter in surface soil, results in more consolidated soil and the establishment of preferential flow paths, which allow contaminants to travel to depth more rapidly than typical filtration through the soil matrix (Conboy and Goss, 2000). Groundwater contamination threatens both private wells and surface water bodies that recharge from groundwater outflow.

# RESEARCH METHODS

## Data Acquisition

The data used for this research project were retrieved from the archives of the Annapolis River Guardians project (ARG). The ARG project has involved a total of almost 100 volunteers since its inception in 1992 to monitor water quality within the Annapolis River and its tributaries. Water sampling was performed fortnightly (when possible) between late March or early April and late November or early December. No samples were retrieved when the river was frozen. Samples were taken at 1m depth from the upstream side of bridges at various locations along the river (See **Figure 1** for sampling locations). (Beveridge, Sharpe, and Sullivan, 2006)

Among the parameters recorded are fecal coliform bacteria, dissolved oxygen, weather conditions, and air and water temperatures. Water samples collected for fecal coliform or *E. coli* testing were refrigerated until delivery to the Synova Diagnostics laboratory in Lawrencetown, Nova Scotia where bacterial content was measured. The samples were typically delivered within 24 hours of collection (Beveridge *et al.*, 2006). Field and travel blanks were also tested as part of the quality assurance / quality control procedures.

From 1992 to 1997 and from 2003 to 2004, the samples were tested for fecal coliforms using the membrane filtration method. This technique involves the filtration of a measured volume of sample through a filter pad, the contents of which are then transferred to an absorbent pad containing growth media to promote formation of bacteria colonies. Samples were then incubated for 24 ± 2 hours at 44.5$^{o}$C before counting the number of visible colonies, which appeared as navy blue spots (Beveridge *et al.*, 2006). Results were expressed as the number of colony forming units (cfu) per 100ml of sample. Synova's internal quality control procedures indicate that fecal coliform data are precise to within 11 percent (Beveridge *et al.*, 2006).

In 2005, a switch was made to the IDEXX Colilert procedure, to yield a Most Probable Number (MPN) of *E. coli* bacteria present in the sample (Beveridge *et al.*, 2006). The Colilert procedure can be used to detect total coliforms and *E. coli* simultaneously (USEPA, 2007). The technique is based upon the production of specific enzymes by the target bacteria, which then react with test substrates. The enzyme B-galactosidase, produced by coliforms, cleaves the media substrate O-nitrophenyl-B-d-galactopyranoside (ONPG), which results in a yellow colour when O-nitrophenyl is released (USEPA, 2007). Similarly, a fluorescent substance is produced when the enzyme B-glucuronidase, produced by *E. coli*, hydrolyses 4-methylumbelliferyl-B-dglucuronide (USEPA, 2007). Data from 1997 to 2002 (inclusive) were also reported as MPN of *E. coli* per 100ml of sample (Beveridge *et al.*, 2006). Precision levels for MPN of *E. coli* per 100ml are reported by Synova laboratories to be 10.4 percent (Beveridge *et al.*, 2006).

Bacterial analyses were duplicated for an interim period of two months (4 samples at each of eight locations), in which data were produced using both membrane filtration and most probable number methods, to compare the two results (Beveridge *et al.*, 2006). The difference in concentration values returned by each method was not shown to be statistically significant over the time period of duplication, thus *E. coli* and fecal coliform measurements may be and are used interchangeably for the Annapolis River system (Beveridge *et al*., 2006).

All the relevant data reporting *E. coli* and fecal coliform concentrations were compiled into an electronic spreadsheet, with sampling location across the top (from highest upstream to farthest downstream location) and sampling date down the side (from earliest to most recent). This spreadsheet was then used as the basis for further analysis, as outlined below.

In addition to the data retrieved by CARP's River Guardian program, this research makes use of geographic data from the archives of Access Nova Scotia. Data were retrieved with the assistance of Jennifer Strang and the Dalhousie Maps and Geospatial Imaging Centre (MAGIC).

**Data Analysis**

First, the software program ArcGIS was used to produce a map of the sample area including topographic contours, surface water bodies, and sample locations (**Figure 1**). Eight of the forty-nine ARG sampling locations were chosen for more detailed study, based on the completeness of their data records. These eight locations were plotted on the map and are outlined in **Table 2**. Additionally, the six tributaries studied as part of the Clean Annapolis River Project (CARP) Sub-Watershed Investigative Monitoring (SWIM) project were included on the map for purposes of relative positioning. Data on roads and buildings were retrieved from Access Nova Scotia archives, but were omitted from the map creation for clarity of the final image.

Next, the data were compiled into spreadsheets in Microsoft Excel and the observed fecal coliform and *E. coli* concentrations were amalgamated to produce a continuous record. In instances where both measurements were available for common samples, the arithmetic mean was calculated and used. This amalgamation was completed based on the statistical tests previously performed by CARP, which indicated that there was no statistically significant difference between the values of *E. coli* and fecal coliforms in the Annapolis River.

Microsoft Excel's statistical package was used to plot a histogram showing the frequency distribution for the combined *E. coli* - fecal coliform record (hereafter referred to as "bacterial concentration" or "indicator concentration"). Based on the observed distribution, the data were transformed using a natural logarithm and another histogram was plotted. To do so, all null observations (0 cfu / 100ml for fecal coliform analyses or MPN / 100ml for *E. coli* tests) were eliminated. This is not expected to diminish the strength or accuracy of any further analyses, as zero values are more likely to be indicative of lack of source than of any influences of

environmental factors. Microsoft Excel was also used to produce a table of summary statistics outlining the characteristics of the data for each of the eight primary ARG sampling sites selected previously and for the dataset as a whole.

The compiled and transformed data were then exported from Microsoft Excel to MiniTab for additional statistical testing. The MiniTab statistical software package was then used to quantitatively test for normality and to plot box-and-whisker plots of bacterial concentrations at each of the eight primary ARG sampling sites. Next, an analysis of variance test (ANOVA) was performed to test for statistically significant differences among the eight primary ARG sample sites. Based on the results of the ANOVA, two-sample Student's t-tests were performed (under assumptions of unequal variances) between all possible pairings of sample locations to determine which locations had statistically significant differences. The results were summarized in a table. A 95% confidence level ($\alpha = 0.10$) was utilized for these statistical tests, based on the recommended values within the statistical software.

Similar methods were used in MiniTab to test for temporal variability in bacterial contamination throughout the entire river. All available transformed data were used (from all 49 sampling locations) and grouped together by month, with no distinctions made for sampling location. Box-and-whisker plots were produced comparing concentration distributions by month and an ANOVA was performed, using a 95% confidence level, to test the null hypothesis that there were no differences in mean transformed concentrations among months. The results of the ANOVA then dictated the necessity of Student's *t* test against null hypotheses of no differences in mean concentrations for each pair of months. The p-values were again reported in a table. These procedures were then repeated for interannual variability, replacing "month" with "year" in the previous description.

Box-and-whisker plots, ANOVA tests, and Student's *t* tests were also used to test the response of transformed bacterial concentrations to rain rate the day of sampling and also to the rain rate three days before sampling. The ANOVA was used to test the null hypothesis that there were no differences in mean transformed concentrations for different rain rates the day of sampling and for three days before the sampling date. Based on these results, one-tailed Student's *t* tests were used where appropriate, to test the null hypothesis that one sample's mean was larger than that of the other sample.

To test the influence of temperature and acidity (pH) on *E. coli* and fecal coliform abundance, the logarithmically transformed concentrations were plotted against both temperature and pH and the plots were fitted with trendlines. Both MiniTab and Excel were used to create scatter plots and perform regression analyses for these two variables. Linear and quadratic fits were tried for both temperature and pH as the predictor variable.

To test for the combined influences of multiple environmental variables, a step-wise multiple regression was performed manually in MiniTab. To facilitate this test, nominal variables, such as rain rate, were converted to index values, such that no rain was represented as 0, drizzle as 1, moderate rain as 2, and heavy as 3. Predictor variables were then added by forward selection, with the strongest predictor (greatest R-squared value) added first and subsequent independent variables added in order of decreasing predictive power (i.e. decreasing R-squared values). Once all desired variables were added and the changes to the total predictive power of the model were minimal for each addition (changes to R-squared values of 0.1%), reverse elimination was used to remove one, two, or three variables at a time to find the model that best suited the data. Due to limitations in sample size, pH values could not be included in this regression analysis. This was the final analytical technique used for this research project.

# DATA PRESENTATION & ANALYSIS

## Data Distribution and Summary Statistics

The first component of data analysis was the evaluation of data distribution and tests for normality. Based on the histogram presented in **Figure 2** (**Appendix B**), the data were strongly skewed to the right and visibly violated the normality assumptions, as they were not distributed symmetrically around a mean. In actuality, 1048 of the 2047 data points – equivalent to 51.20% of the data – were included in the first bin (concentration values ranging from 0 to 100). Another 18.61% (381 observations) were included in the second bin, including values from 101 to 200. Thus, the first two bins combined account for 69.81% of all data points, as demonstrated by the cumulative percentage line in the graph (**Figure 2**). Subsequent bins contained considerably fewer measurements. The final bin – labeled "more" and including all observations with concentration values greater than 1000 – is noteworthy in that it contains 7.38% of all observations (151 individual entries), which makes it the fourth most populated bin. This contributes considerably to the strong right skew of the data and makes parametric statistical analysis undesirable, as these high values (reaching a maximum of 23 400 cfu / 100ml) would bear disproportionate influence on calculated statistics.

This initial examination of the data was supported by a table of summary statistics produced from the Excel spreadsheets and presented in **Table 3**. Each of the eight primary ARG sites selected for more detailed analysis are shown in this table, as are the summary statistics for the dataset as a whole (all 49 sampling locations). The differences between the various measures of central tendency (mean, median, and mode) give further evidence of the non-normality of the distribution and support the use of non-parametric statistics. The table also demonstrates the

potential for high values to have excessive influence on statistical tests, as the maximum values are far further from the means than are the minimum values (providing additional evidence for the non-symmetric distribution of data).

Furthermore, the summary statistics for the complete dataset help to quantify the distribution of the 2047 data points. Again, the maximum (23400) is several times further from the mean (429) than the minimum (0) is. The same is true when considering the median for the entire dataset (100). This distribution would yield unwarranted weight to the higher values if parametric statistics were used. It is interesting to note, however, that the most frequent concentration recorded amongst the more than 2000 observations was zero. This is an encouraging result for those who rely on water from the Annapolis River. However, this news is tempered by the demonstration that roughly half (49%) of the data exceed the 100 cfu or MPN per 100ml guideline for use of water for crop irrigation and that 30% of all measurements exceed even the 200 cfu or MPN per 100 ml guideline that is the maximum allowable concentration for contact recreation activities. This implies that nearly one third of the measurements of bacterial concentration within the Annapolis River violate even the most lax guidelines, emphasizing that this is an issue of concern for Annapolis River communities. These results – 65% of measurements over the past 15 years exceed guidelines for livestock watering and 49% exceed guidelines for crop irrigation – have considerable implications for agriculturalists in the area.

Logarithmic transformation of the amalgamated concentration record produces a distribution that is considerably more normal, as per the histogram displayed in **Figure 3**. A normality test was performed using MiniTab that reported a p-value of less than 0.005, indicating that this distribution is effectively normal. Therefore, parametric statistical methods are appropriate when using the transformed data. Given that information is lost in the conversion of

data to ranks for non-parametric statistical analyses, it is preferable to use a transformation to achieve a normal distribution whenever possible (McDonald, 2008). Thus, for increased descriptive power, the transformed data were used whenever possible in all subsequent analyses.

The following analyses will probe deeper into factors that might influence the distribution of *E. coli* and fecal coliform measurements, examining which locations are subject to lesser or greater contamination, what trends are present over various time periods, and the role of various environmental conditions on the concentrations of these indicator organisms.

## Geospatial Distribution

Of the forty-nine sites that the Annapolis River Guardians sampled, only eight were sampled with reliable frequency or for a considerable duration. These eight sites are outlined in **Table 2**. The summary statistics for these eight sites (presented in **Table 3**) indicate a sample size of 185 ± 17 for seven of the sites; Aylesford Rd. (Site AY40) is an exception, with only 44 entries. This comparative lack of data is attributable to its more recent inclusion in the sampling program. The earliest testing at Aylesford Road commenced in 2005, whereas the records for the other seven sites extend back to October of 1992. However, the sampling frequency is quite good for AY40 and a sample size of 44 is generally considered large enough to produce viable statistical results, so it was included for further testing. Among the other 41 sites, sample sizes ranged from 1 to 51 samples, with a mean sample size of 18.65. The lack of data at these other locations is one limitation to the full understanding of geospatial distribution of concentrations of indicator bacteria within the Annapolis River.

Further examination of the summary statistics presented in **Table 3** provides some insight into the differences amongst sample sites. All three measures of central tendency indicate that that the Aylesford site (00) is considerably more polluted than the others, with the Aylesford

Road site (AY40) coming in second, based on the means and medians. The modes also suggest that Wilmot (18) and Middleton (25) experience frequent high concentrations. These tendencies are also supported by the lowest four rows in the table, which delimit the percent of entries at each location that exceed the various guidelines for water quality, as previously outlined. The Aylesford site (00) has the greatest proportion of observations in excess of 50 cfu / 100ml (or MPN / 100ml), the greatest proportion of observations above 100 cfu / 100ml, and the highest proportion of observations greater than 200 cfu / 100ml, making it the most polluted of the eight sampling sites overall. The Aylesford Road site (AY40) had the second highest proportions above each of those threshold values. Both sites have considerably greater proportions in excess of each guideline level than did the river as a whole (as represented by the summary statistics for the complete dataset, presented in the right-most column of **Table 3**). However, the fact that the statistics for the eight sites are scattered around those values for the river as a whole suggests that collectively, the eight sites may be reasonably representative of the river system.

The relative contamination of the eight sites is further evident in the box-and-whisker plots shown in **Figure 4**. The box-and-whisker plot displays the distribution of the logarithmically transformed measurements of bacterial concentrations at each of the eight sites, for all available dates. This plot supports the observation based on **Table 3** that the Aylesford site (00) is the most contaminated and the Aylesford Road site (AY40) is the second most contaminated (based on a comparison of their respective medians and third quartile ranges against the other sites). However, the methods so far used to assess the relative contamination bear little statistical power, as they are predominantly qualitative (in the case of the box-and-whisker plots) or rely on scant indicative numbers (in the case of the table of summary statistics). Thus, the ANOVA procedure was used to test for the existence of any statistically significant

differences amongst the eight sites (i.e. testing a null hypothesis of no differences in sample means amongst the eight locations). The MiniTab output for this test reported an F-statistic of 10.28 and a p-value of 0.000 – both of which support the existence of differences among the sampling locations that are significant at the 95% confidence level. The output also included an adjusted R-squared 4.91 percent. The MiniTab output also included an approximate visual depiction of the 95% confidence intervals for each of the sample locations.

The existence of statistically significant differences among the sample locations prompted the use of Student's t-tests (also at the 95% confidence level) to determine which locations were significantly different from which others (against a null hypothesis of zero difference between each set of two sites). The results of these tests are summarized in **Table 4**. Based on a 95% confidence level ($\alpha$=0.05), there are statistically significant differences between Site 00 (the Aylesford site) and all locations downstream of it (sites 13 to 49); p-values of 0.000 were determined for all comparisons between Site 00 and other sites (excepting Site AY40). Site AY40 was found to be significantly different from sites 13, 35, 40, and 49. Statistically significant differences were also found between sites 13 and 18, 18 and 49, and 25 and 49.

In summary, the ANOVA and Student's *t* tests demonstrate that there are statistically significant differences in the concentrations measured among the eight sites. When combined with the qualitative analysis based on the box-and-whisker plot, this suggests that at the 95% confidence level, the concentrations measured at Site 00 are significantly greater than all locations downstream of it. Similarly, the concentrations measured at Site AY40 are significantly higher than four of the other seven sites.

Thus, it appears that – as a generalization – *E. coli* / fecal coliform concentrations tend to be greater in the headwaters and lower throughout the rest of the body of the river. Based on the

understanding of microbial fate and transport presented in the literature section of this paper, it might therefore be posited that the contamination in these upstream locations may flow downstream and contribute to measured concentrations downstream. The higher values in the headwaters may be attributable to sediment-bound organisms persisting in the riverbed or soils in the drainage basin and a greater likelihood of resuspension, if the headwater sites are shallower than the sites downstream. The differences may also be attributable to activity types or density in that vicinity that differ from locations further downstream.

These results may be applied in targeting riparian management projects to achieve the greatest possible efficacy – as determined by larger reductions of contaminant concentrations achieved with fewer inputs, whether financial, time, or material. If projects can be implemented to mitigate further *E. coli* / fecal coliform loading of these headwaters, the benefits may be felt downstream, if indeed these locations act as sources for microbial concentration in downstream sample sites.

This spatial distribution of bacterial contamination can also be compared with the results from the CARP Sub-Watershed Investigative Monitoring (SWIM) project, presented in **Table 1**, using the map in **Figure 1** that displays positions of the sampling sites. The results summarized in **Table 1** indicate that the Fales River (FR20) contributed the most to bacterial contamination of any of the six tributaries tested as part of the SWIM project, with a flux of $6.042 \times 10^6$ fecal coliform units per second. The Nictaux River (NX20) was the second greatest contributor, adding $3.398 \times 10^6$ fecal coliform units per second – slightly more than half as much as the Fales River. The other four tributaries contributed considerably fewer fecal coliforms to the main river body. It is interesting to note, however, that both of these high-contribution sites are downstream of the Aylesford sites (AY40 and 00), as shown on the map in **Figure 1**. Thus, their

contributions do not seem to reflect the general trends in the concentrations at the ARG sites (high concentrations upstream and lower concentrations downstream). However, both of these sites are relatively close to the Wilmot sampling location, so might contribute to its elevated concentration values.

**Temporal Distribution – Seasonality**

The time series plot displayed in **Figure 5** relates the logarithmically transformed concentrations to their sampling dates and is fitted with a moving average trend line, based on a period of fifty data points. The trend line accentuates a general tendency for concentrations to decrease from January to July and to increase again from July to January. The winter plateaus may be indicative of the paucity of measurements taken during the winter months, when the water was inaccessible due to ice cover. Visual inspection of this plot suggests that there may be some seasonal effects on the concentrations of these indicator bacteria. However, the convexity of the plot produced after eliminating the years from the date fields (see **Figure 6**) appears to contradict the trends identified on the full time series. Based on **Figure 6**, it appears that bacterial concentrations have a tendency to increase over the course of the spring, peak during the summer months and decline slightly during the fall. The box-and-whisker plot (presented in **Figure 7**) compares the distribution of logarithmically transformed concentrations observed in each month, for all locations over the entire sampling program to date (1992 to 2007). Visual inspection of this box plot supports the generalized seasonal trend from **Figure 6**. February is a visible exception, based on the box plot, as it has anomalously high concentrations, compared to other months. This may be a valid result or it may be attributable to the small number of samples collected during February. The Annapolis River Guardians did not collect samples when the river surface was frozen, resulting in a very small sample size for the months of January,

February, and March. There are nine, eleven, and twenty-four samples in each month respectively over the duration of the sampling program (1992 to 2007) and for all sampling locations. The paucity of data collected during winter months may limit the accuracy of these data in describing typical winter conditions.

The examination of seasonality of bacterial concentrations may also be limited by the consistency of low concentrations throughout all months. The observance of low concentrations may be associated with environmental conditions that promote inactivation and die-off of fecal bacteria or it may simply be indicative of a lack of source fecal material in the particular time and vicinity of a given sample. Thus, the issue of source material is a confounding factor that may affect the strength of any conclusions drawn about seasonality of fecal bacterial concentrations.

Statistical tests were also performed to support the graphical analyses already presented on seasonality of bacterial concentrations. A one-way ANOVA analysis was performed on the complete ARG dataset, grouped by month, which revealed the existence of statistically significant differences amongst the months at the 95% confidence level. The analysis tested the null hypothesis that there was no difference in mean concentrations among the 12 months. The test returned an F-statistic of 20.48, a p-value reported by MiniTab to be 0.000, and an adjusted R-squared value of 10.02 percent. This latter value makes sampling month one of the strongest predictors of transformed bacterial concentrations tested in this study. These outputs support the use of Student's *t* tests to determine which months are different from which others.

The results of the month-wise Student's *t* tests are summarized in **Table 5**. An α-value of 0.05 was used for these tests, which tested a two-tailed alternative hypothesis against the null hypothesis that each pairing had no difference between its means. The results, shown in **Table**

**5**, indicate that the distribution of bacterial concentrations in April was statistically significant from that in all other months, with p-values ranging from 0.000 to 0.032. Combined with qualitative analysis based on the box plots and scatter plots, this confirms that, within the Annapolis River, April has the lowest concentrations of *E. coli* and fecal coliforms than any other month, for the period of 1992 to 2007. May was also found to have statistically significant differences from the distributions in most other months, excepting January, February, and March. Considering the graphs again, this suggests that May also typically has lower bacterial concentrations than later months. July and September were found to be significantly different from all months but February and each other, based on the p-values in **Table 5**. When paired with visual inspection of the box plot in **Figure 7**, this would suggest that these two months have some of the highest concentrations each year. These tests seem to support the trend for lower concentrations in the spring and higher concentrations during the summer, as depicted in the scatter plots and box plots in **Figures 5** and **6** respectively.

These seasonal differences may be related to water fluxes to and in the river, which tend to change with the seasons. The months of July, August, and September are typically warm and dry, and water shortages have been identified as a concern for the Annapolis River (Timmer, 2003a). If these conditions contribute to enhanced evaporative losses and reduced groundwater recharging of the Annapolis River, concentration values may rise due to depletion of solvent (water) quantities, rather than changes in bacterial loading. Similarly, high rainfall rates and spring melts may contribute to diluted samples in the spring months and/or may encourage flushing of the river, washing potential contaminants out to the river mouth. While these mechanisms may influence or help to explain the observed trends, the reality remains that regardless of the reason, elevated concentrations of fecal bacteria pose a health risk for

recreationists, agriculturalists, or other users of the Annapolis River waters when they are contaminated. Thus, these trends may be useful in planning activities that require water inputs.

**Temporal Distribution – Interannual Variability**

The Clean Annapolis River Project's Annual Water Quality Reports typically indicate the trend in concentrations relative to the previous year. As part of this research project, differences amongst all sampling years were investigated. The box-and-whisker plot in **Figure 8** portrays the distributions of values for each of the years sampled. Based on this plot, it appears that 1994, 1996, 2004, and 2006 were years of anomalously high concentrations, whereas 1995 appears to be the year with the lowest typical concentrations. It should be noted, however, that sampling only began in October of 1992, so this sample year is biased toward the values measured in October, November, and December of that year. The spring lows and summer highs identified in the previous section would not be included in the distribution for the year 1992.

Quantitative statistical analysis revealed significantly different concentrations amongst the 16 sampled years. The MiniTab output from a one-way ANOVA testing a null hypothesis of no differences in mean transformed concentrations among the various years included an adjusted R-squared value of 9.38 percent, an F-statistic of 14.29, and a p-value of 0.000 at the 95 percent confidence level. Further analyses using Student's *t* tests indicated which years differed from which others at the 95 percent confidence level (see **Table 6**). The year 1994 was demonstrably the year with the highest concentrations, based on p-values indicating statistically significant differences between it and every other year (except 2006) and combined with visual inspection of **Figure 8** and the confidence intervals returned in MiniTab's ANOVA output. Similarly, 2006 was found to be significantly different from all years except 1993, 1994, 1996, and 2004. When this is considered alongside the graphical displays, it suggests that 2006 was also a year of high

concentrations. P-values less than α (0.05) indicate that the mean logarithmically transformed concentration for 1996 was significantly different from 1994, 1995, 1997 to 2003, and 2005. The same is true for 2004. Based on the 95 percent confidence intervals for the means, available in the MiniTab ANOVA output, it appears that 1996 and 2004 had lower mean logarithmically transformed concentrations than 1994, but higher such concentrations than the other years listed. Conversely, 2002 was found to have statistically significant differences in mean logarithmically transformed concentrations from all years except 1992, 1995, and 1999 to 2001. Again, considering this alongside the qualitative outputs from the box plots and 95 percent confidence intervals, this ranks 2002 among the least contaminated years in the dataset.

This examination of interannual trends may be particularly useful if considered in light of changes in activity on or near the river. It may be helpful in determining the success of reclamation projects, the impact of new developments or changes in practices on riparian properties, or the changes in response to changes in nearby livestock populations. However, in the absence of this information, the temporal record does not provide much transferable information. There is no generally apparent trend in bacterial concentrations – increasing or decreasing – based on these observations. Further research comparing differences in annual concentration distributions to changes in activities or populations on nearby land is recommended.

**Influences of Rain Rate**

Rain rates were recorded subjectively for the sampling date and the day three days prior to sampling. Each of these records included the values "None," "Drizzle," "Moderate," and "Heavy," as perceived by the volunteer who was retrieving that particular sample. The box-and-whisker plot in **Figure 9** displays the distributions of logarithmically transformed concentrations

recorded for each nominal value in both records (i.e. the left four boxes represent None, Drizzle, Moderate, and Heavy rain rates the day of sampling and the right four boxes show the same four descriptors for the rain rates three days before samples were retrieved). Qualitative analysis of this figure suggests a weak but relatively consistent increase in concentrations with rain rate the day of sampling and a stronger increase associated with rain rates three days before sampling. A one-way ANOVA performed at the 95 % confidence level on natural logarithms of bacterial concentrations versus rain rates the day of sampling produced an F-statistic of 17.26, a p-value of 0.000, and an adjusted R-squared value of 3.39 percent. The p-value indicates the existence of at least one statistically significant difference amongst the four groups at the 95 % confidence level. The same test performed on rain rates three days prior to sampling yielded an F-statistic of 44.16, a p-value of 0.000, and an adjusted R-squared value of 9.27 percent. This makes rain rate three days prior to sampling one of the strongest predictors of concentrations tested to this point.

The significance of both ANOVAs justified the use of Student's *t* tests. The null hypothesis for these tests was one of no difference between each set of two rain rates. Based on the box plot in **Figure 9**, the one-tailed alternative hypothesis was that the greater rain rates yield greater mean transformed concentrations. Based on a 95% confidence level, the tests indicated that moderate and heavy rainfall the day of sampling produced significantly higher mean transformed concentrations than did no rain, but no significant increase in drizzle over no rain. While moderate rainfall was also found to produce significantly higher transformed bacterial concentrations than drizzle, heavy rains were not found to be significantly different from either drizzle or moderate rains at the 95% confidence level on the day of sampling.

These results suggest that rainfall tends to increase the concentration of *E. coli* and/or fecal coliform above levels present in the absence of rainfall. They further suggest that moderate

rains elevate concentrations above levels observed during light rains (drizzle), but that heavy rains do not significantly increase concentrations any more than drizzle or moderate rains during the rainfall event. This result may be related to rates of river flushing, in that heavy rains are likely to cause swelling in the river that might carry contaminants downstream. They are also likely to be influenced by rates of stream-bed sediment resuspension, in response to turbulent mixing of the water column. Further research in this area might compare concentrations with observed rain rates at sites further upriver to test this hypothesis and should include information about the duration of rainfall, in addition to intensity.

When the previously outlined tests were repeated for rain rates three days before sampling, each nominal rain rate was found to have a significantly higher mean transformed concentration than its predecessors. That is to say that heavy rains produced significantly higher concentrations than moderate ($p = 0.000$), drizzle ($p=0.000$), or none ($p=0.000$); moderate rains produced significantly higher concentrations than drizzle ($p=0.002$) or none (0.000); and drizzle yielded significantly higher concentrations than no rain ($p=0.041$).

Based on these statistical analyses, it is apparent that rain rates three days prior to sampling have a strong influence on bacterial concentrations. This is likely related to overland flushing of fecal matter and turbulent mixing suspending organisms previously adsorbed to the riverbed sediments. Given that the only temporal set-back reported was three days and that sampling resolution was fortnightly at best for any given location, it is not possible to describe the evolution of surface water bacterial concentrations following hydrologic events (i.e. predict the number of days after rainfall when concentrations peak). However, the results do indicate that, for some days following rain events, water users are more likely to be exposed to elevated bacterial contaminant concentrations, which mimic the risk of exposure to microbial pathogens.

It is also possible that there is some interaction between rain rates on the sampling day and three days prior to sampling, which may be appropriate for further investigation in subsequent studies.

**Temperature and pH**

The responses of logarithmically transformed bacterial concentrations to changes in temperature and acidity (pH) are shown in the scatter plots in **Figures 10** and **11**. Both linear and quadratic fits were attempted for each of these predictor variables, but neither modeled the data particularly well. The quadratic fits generally had greater R-squared values and so were selected for display in the appendix. However, with R-squared values of roughly 1% for temperature and 3% for pH, it seems that these variables are among the weakest predictors of concentrations tested within this study (i.e. bacterial concentrations are relatively impervious to changes in temperature or acidity). MiniTab's regression analysis provided an adjusted R-squared value of 0.0% for the response of logarithmically transformed *E. coli* concentrations to variations in pH and the square of pH, and yielded p-values of 0.381 and 0.392 for the two predictors, respectively. This is insufficient to reject the null hypothesis of no relationship between pH and ln(*E. coli*). However, when the same analysis was performed for water temperature and its square, the regression output included p-values of 0.003 for H2OTemp and 0.038 for H2OTemp^2. These p-values are sufficient to reject the null hypothesis of no relationship between water temperature and logarithmically transformed concentrations. Thus, it can be concluded that pH does not influence *E. coli* abundance within the range of pH values measured by the Annapolis River Guardians, but that temperature is a weak, but statistically significant predictor of logarithmically transformed fecal bacterial concentrations, explaining only 1.3% of the changes in the response variable.

It should be noted that tests for response to pH are limited by the small sample size. Only samples retrieved from site AY40 had pH records, and even then, not all samples from AY40 had an associated pH recording. Thus, the total sample size for this test was only 20 data points, which serves as a limitation to the test's accuracy. Furthermore, the Annapolis River has fairly stable, circumneutral pH – apparent in the pH records for site AY40 (with the 20 available data points ranging from 6.02 to 7.59) and reported in annual Annapolis River Water Quality Reports published by CARP. These values are unlikely to stress organisms greatly. It is possible that more pronounced trends would be apparent over a wider range of pH values. This might be appropriate for a laboratory experiment.

**Multiple Regression**

Based on the analyses done to this point, there is no single variable that stands out above all others as the key predictor of bacterial concentrations. It is expected that a multitude of environmental factors and measured variables influence concentrations and that there may be interactions amongst these predictors. Thus, as outlined in the methods section, a step-wise multiple regression was performed to seek a model with the greatest possible predictive power based on the fewest variables. New predictor variables were added by forward selection to determine the added predictive power of each variable. Subsequently, one, two, or three of the independent variables were removed to establish whether there were interactions between the variables (e.g. if the addition of one variable contributes greatly to the R-squared value, but the addition of the next contributes only slightly, the former might be removed and the R-squared value might only decrease slightly, indicating that these two variables are closely related and each alone helps to predict the response of the dependent variable, but their combination does not contribute greatly to predictive power). The null hypothesis for the multiple regression was that

there did not exist a relationship between the independent variables and the dependent variable (logarithmically transformed concentrations). The alternative hypothesis was one of a relationship amongst the dependent and independent variables. P-values associated with these tests are used to reject the null hypothesis, where appropriate, given a 95 % confidence level.

It should be noted that pH would ideally be included in this multiple regression analysis, however the paucity of data combined with the lack of measurements for any of the other independent variables – except water temperature – for those particular samples with associated pH measurements necessitated the exclusion of pH as an independent variable in this regression analysis.

Throughout the forward selection process, the R-squared value increased with each added variable, but variables added early (based on higher R-squared values in single-predictor tests) had stronger influences on the total R-squared value. The initial model included only R3DBIndex and resulted in an adjusted R-squared value of 8.1 percent. The addition of Month as a predictor increased the model's predictive power to 11.2 percent. The fit increased to 14.9% when H2OTemp was added and again to 21.3% with the addition of Site ID. Adding Rain Index as an independent variable raised the adjusted R-squared value to 22.5 percent. Including year in the model brought the adjusted R-squared value to 22.8 percent and the addition of the square of water temperature (H2OTemp^2) actually brought the adjusted R-squared value down to 22.7 percent. In this most complete inclusion of all available independent variables, each variable was associated with a p-value less than 0.003, indicating that the null hypothesis can be rejected at the 95% confidence level for all the variables tested (i.e. all the variables tested are related to the dependent variable). Based on the forward selection process, it seems that the ideal multiple regression model to describe variations in logarithmically transformed *E. coli* and fecal coliform

concentrations within the Annapolis River includes the first four variables: R3DBIndex, Month, H2OTemp, and SiteID. This was further tested by removing independent variables by reverse elimination.

Through reverse elimination, each variable was removed individually in turn and also as sets of two, three, or four variables removed. The removal of Year yielded an adjusted R-squared value of 22.4 percent, whereas the removal of squared water temperature resulted in an adjusted R-squared value of 22.8 percent. The elimination of both of these variables together produced an R-squared value of 22.5 percent. It is thereby appears that the square of water temperature contributes least to the model's predictive power. Further testing removing RainIndex alone (adjusted R-squared = 21.5%), RainIndex and H2OTemp^2 (adjusted R-squared = 21.6%), and Rain Index, H2OTemp^2, and year (adjusted R-squared = 21.3%) support this supposition. This latter model (adjusted R-squared = 21.3%) matched the optimal model based on forward selection (i.e. included R3DBIndex, Month, SiteID, and H2OTemp). Each of the four remaining variables was then removed one at a time resulting in adjusted R-squared values of 19.3% (for the removal of water temperature alone), 15.7% (for the removal of month only), 14.9% (for the removal of Site ID alone), and 9.3% (for the removal of only the rain rate three days prior to sampling). Thus, based on both forward selection and reverse elimination, it is apparent that the optimal model includes the variables R3DBIndex, SiteID, Month, and H2OTemp, and accounts for 21.3 % of variations in logarithmically transformed bacterial concentrations.

Given that the adjusted R-squared value did not increase by the R-squared value reported for each variable individually, it is apparent that there are some interactions between these variables (for example, month and water temperature would be expected to be correlated). It

also summarizes that the research completed as part of this project can, at best, explain 22.8% of variations in logarithmically transformed fecal bacterial concentrations. This implies that either other factors exist that influence concentrations of fecal bacteria in the Annapolis River's surface waters or that the concentration distribution is inherently chaotic. It is also possible that both of these options are true – the concentrations may be affected by environmental factors not examined here but that still do not account for all of the variability without accounting for some natural chaotic behaviour.

# CONCLUSIONS AND RECOMMENDATIONS

This research project was performed with the intent to improve community understanding of the patterns in concentrations of bacterial indicator species. A better understanding of what environmental factors influence *E. coli* and fecal coliform abundance can be employed to target reclamation and riparian management projects by local community groups, such as the Clean Annapolis River Project (CARP). The results of this project may also be of interest to other community groups, such as the Nova Scotia Federation of Agriculture, the Nova Scotia Agricultural College, and the Growers Water Group. This report may also be helpful to families or farms assessing the risk of using surface water at a given location, during a given month, or under other such environmental conditions.

The Annapolis River was a prime study area, as the Clean Annapolis River Project has been recording water quality data since 1992, when they received partial funding for the Annapolis River Guardians (ARG) project from the Atlantic Coastal Action Program (through Environment Canada). Furthermore, the intense agriculture within the Annapolis Valley not only plays a role in contributing fecal contamination to the river, but also stands to be negatively impacted by such contamination. As the human population continues to grow, so too will the demand for increased agricultural production, which will exacerbate the demand for water for irrigation in this area. Thus, water quality within the Annapolis River is a prime concern for the local community.

*E. coli* and fecal coliform are used as indicators for the probability of exposure to other microbial pathogens, as they are inexpensive and easy to test for. Based on the literature reviewed, it is apparent that the use of *E. coli* and fecal coliforms to indicate the possible presence of other pathogens of fecal origin is imperfect, as viruses and protozoa may move

differently within the environment than bacteria do, based on their sizes, and may be subject to different stresses that influence their abundance. Furthermore, fecal coliforms may be present in the absence of other pathogens, based on whether the source fecal matter was infected. However, *E. coli* and fecal coliform testing is still widely used, as it is relatively easy and inexpensive, has been relatively effective in the past at providing a ball-park estimate for the contamination risk in a given water body, and is an application of the precautionary principle.

The importance of understanding the fate and transport mechanisms for these bacterial indicators, as well as their sources within a given watershed, are exemplified in the case of Walkerton, Ontario, which experienced an *E. coli* outbreak in the summer of 2000. The ramifications to human health, as well as the local and provincial economies and real estate markets, were considerable. The health of other populations may also be affected by *E. coli* or fecal coliform contamination, which could result in financial losses if commercially productive fisheries or shellfish breeding grounds are affected. Consumption of contaminated animal products, produce, or water can have significant deleterious health effects, including kidney failure and even death, especially in young children, the elderly, and those with pre-existing health conditions. These factors were key motivators in the pursuit of the research presented in this report.

Statistical analysis of the ARG data records for water quality in the Annapolis River yielded many interesting and significant results. Summary statistics indicated that 65% of data points pertaining to the Annapolis River between 1992 and 2007 (inclusive) exceeded the recommended maximum concentration in water for livestock, 49% of data points exceeded the recommended maximum concentration for water used for irrigation, and 30% of data points exceeded the recommended maximum concentration for contact recreation. Box-and-whisker

plots, analysis of variance tests, and Student's *t* tests were all used to test for differences in distributions of logarithmically transformed concentrations by location, month, year, rain rate, temperature and pH. The results indicated that, with the exception of pH, significant relationships existed at the 95% confidence interval for each of these variables with logarithmically transformed concentrations. Concentrations were found to be significantly higher in the headwaters than downstream. Late summer was shown to have the highest concentrations (with July and September significantly higher than any other month), whereas spring was shown to have the lowest concentrations (notably the distribution of logarithmically transformed concentrations was significantly lower in April than any other month). With respect to interannual variability, the years 1994, 1996, 2004, and 2006 had significantly greater concentrations than other years, but no general trend was apparent over the course of the data record, to indicate that concentrations were either increasing or decreasing overall. Furthermore, rain rates three days before sampling demonstrate a clear positive correlation (i.e. stronger rains elevated the distribution of logarithmically transformed concentrations), whereas rain the day of sampling did not exhibit such a clear trend. Temperature was only found to have a very weak relationship with bacterial concentrations, yielding an adjusted R-squared value of 1.1 % for a quadratic fit, whereas pH was not found to have any statistically significant relation to logarithmically transformed concentrations for the 20 available data points. The pH analysis may have been limited by the circumneutral pH values apparent within the Annapolis River, which are unlikely to cause stress. When the influences of these predictor variables were combined, through the application of a step-wise multiple regression, it was found that the four most influential predictors of logarithmically transformed bacterial concentrations were rain three days prior to sampling, sample location, month, and water temperature. The regression

model incorporating these four variables accounted for 21.3% of the variation in transformed *E. coli* and fecal coliform concentrations. These results can be better understood when considered within the framework of the literary research also presented within this report.

The literature implicated livestock populations and the application of manure for fertilizer as the primary sources of fecal bacteria to watersheds. Tiedemann *et al.* (1987) observed a direct relationship between the presence of cattle on agricultural land and increased fecal coliform concentrations in nearby surface waters. Data on the spatial and temporal variations in livestock population densities within the Annapolis Valley could not be retrieved in time to include such analyses in this project, but may help to explain some of the interannual and/or geospatial variability in bacterial concentrations. Thus, further research into the correlation between *E. coli* and/or fecal coliform concentrations and livestock population densities is recommended. It is also recommended that Microbial Source Tracking (MST) research be pursued for the Annapolis River, as this would allow for direct identification of sources of fecal contamination in each sampled area (i.e. whether sources are cattle, pigs, sheep, other livestock, wild animals, or faulty septic or municipal sewage systems). Such research would allow for better targeting of projects aimed at reducing fecal contamination of the Annapolis River.

Tiedemann *et al.* (1987) also found that direct livestock access to streams was a primary contributor to elevated bacterial concentrations – an issue that CARP tackles directly in their community-based riparian management activities. Furthermore, wetlands have been shown to be capable of diminishing fecal coliform concentrations by 98% (Thurston *et al.*, 2001). These two factors may have influenced the results of CARP's Sub-Watershed Investigative Monitoring (SWIM) project, which pinpointed regions of greatest loading, among the six tributaries monitored in 2004. It is interesting to note that while location within the river was shown

through this research project to be one of the strongest predictors (among the variables studied) for *E. coli* and fecal coliform concentrations, the distributions of logarithmically transformed concentrations at each of the eight primary ARG sampling sites did not seem to be well explained by the relative loading from the six tributaries studied in the SWIM program. Thus, it is recommended that additional tributaries of the Annapolis River be monitored, if better spatial modeling of fecal contamination sources is desired.

The existing literature suggested that hydrologic events increasing surface run-off rates strongly influence on bacterial loading of a surface water body (Ferguson *et al.*, 2003). The statistical analyses performed as part of this research project are in keeping with these observations, as the index for rain three days prior to sampling was one of the strongest predictors of measured bacterial concentrations. However, rain rates were recorded subjectively – volunteers collecting samples reported rain rates as 'none,' 'drizzle,' 'moderate,' or 'heavy' – and were only recorded on the day of sampling and three days prior to sampling. Furthermore, there was no indication of the duration of rain, which could play an important role in flushing. To more accurately investigate the response of *E. coli* and fecal coliform concentrations to hydrologic events, more quantitative rain rates would be beneficial, as well as more frequent observations (e.g. rates for the date of sampling, as well as every day or every other day for the five to seven days prior to sampling). Even if this could be recorded at only one site, it would allow for more detailed investigation of the relationship between rain rates and the subsequent evolution of bacterial concentrations. Such information could be gathered using a rain gauge, which could indicate the depth of precipitation received over a specific time period. A makeshift gauge could be easily and inexpensively constructed using a cylindrical bucket or a short length of PVC pipe closed off at one end and with a ruler inserted into the column or depth markings

indicated along the cylinder sides. This tube could then be situated vertically, so that precipitation would fall into the open end and accumulate to an appropriate depth. A wider opening would improve accuracy where strong winds cause rain to fall at an angle from the vertical, however it would increase evaporative rates during periods without precipitation, which may decrease the accuracy of the recording. In short, such an instrument would be an inexpensive way to better quantify the relationship between bacterial concentrations and rain rates. This information could be used to determine how long after a rainfall concentrations peaked and whether this was dependent upon how heavy the rain was.

Additional hydrologic information that may be pertinent would relate to temporal records of stream flow rates. Sampling month was found to one of the strongest predictors of logarithmically transformed bacterial concentrations (among the variables tested), with an adjusted R-squared value of 10.02 percent, but may have been confounded by stream flow rates. Since *E. coli* and fecal coliform concentrations are recorded as a function of volume, decreased stream flow during summer and fall months may contribute to the significantly higher concentrations observed in July and September than in other months, and to the significantly depressed concentrations in April. Practically, the concentration value – independent of stream volume – is the number of greatest interest to those using the water, but from scientific and best-management perspective, it would be very interesting to determine how much of the seasonal fluctuation in concentration is due to variation in water volume. Further research into volumetric flow rates by season could potentially identify how much of the seasonal variation in bacterial concentrations is based in water volume, with the presumed implication that the majority of seasonal variation that is not explained by water volume changes is related to source loading rates. This information would then be useful in the targeted application of riparian management

techniques. Additionally, in order to calculate volumetric flow rates, a benthic profile would be required for each of the sampling sites. This would provide information about the depth of the river at each site, which could then be used to help interpret the differences in concentration distributions among the eight sites.

Literary research into the fate and transport of bacterial indicator species within surface water bodies supported the apparent conclusion that elevated concentrations in Annapolis River headwaters may be contributing to observed concentrations downstream. The two furthest upstream sampling sites were found to have significantly higher concentrations at the 95% confidence level than the sites further downstream. The literature reviewed suggests that as long as water temperatures remain low and the surface waters are shaded from direct insolation, bacterial populations could survive long enough to contaminate the other sampling locations. The implications of result for riparian management techniques are considerable. If the proposed mechanism is true, then focusing initial implementation of best management practices on the region surrounding the Aylesford sampling sites could potentially contribute to lower concentrations throughout the run of the river. Thus, the greatest efficacy would be achieved by projects implemented in the headwaters, whereas riparian management projects implemented downstream might not be as effective as desired, if contamination is contributed by natural stream flow.

This issue could be further exacerbated if a source population exists in the fluvial sediments or riparian soils at these upstream locations. Based on the literature, *E. coli* and fecal coliform populations, as well as those of their associated pathogens, can persist in greater concentrations and for greater durations in soil and sediment matrices than in surface water, and are also believed to be capable of multiplication and growth under these conditions (Grimes,

1975; Gerba, Smith, and Melnick, 1977; Labelle and Gerba, 1980; Davies and Evison, 1991; Ferguson *et al.*, 2003). Groundwater flow through such soils or turbulent suspension of fluvial sediments could mobilize these organisms from these populations into the surface waters. Thus, it is possible that contaminated soils or sediments in the headwaters are contributing contamination to those sites and others downstream. This might therefore support investigations into bacterial concentrations within the soil or sediment, particularly in the regions surrounding sites AY40 and 00 – the two most contaminated sites among the eight studied.

In summary, the seasonal and geographic distributions of concentrations outlined in this report can help to focus mitigation measures where and when they will have the greatest impact. Furthermore, users of the Annapolis River's waters should be aware of the impact rain has in elevating microbial contamination for several days following rain events. However, this research only explains about 20% of the variability of bacterial indicator concentrations within the Annapolis River, leaving many questions yet unanswered.

# RESOURCES

Annan, K. (2003). " 'Water-related diseases responsible for 80 per cent of all illnesses, deaths in developing world', says secretary-general in Environment Day message." *United Nations Press Release SG/SM/8707 OBV/348*. Retrieved 27 October, 2007, from http://www.un.org/news/Press/docs/2003/sgsm8707.doc.htm.

Ashbolt, N.J., Grabow, N.O.W., and Snozzi, M. (2001). "Indicators of microbial water quality." *Water Quality: Guidelines, Standards, and Health*. Risk Assessment and Management for Water-Related Infectious Disease, Chapter 13. Ed. Fewtrell, L. and Bartram, J. p. 289

Beveridge, M., A. Sharpe, and D. Sullivan. (2006). "Annapolis River 2005 Annual Water Quality Monitoring Report." *Clean Annapolis River Project*. Annapolis Royal, Nova Scotia.

Black, M. (2004). *The No-Nonsense Guide to Water*. Toronto, Canada: New Internationalist Publications.

Canadian Council of Ministers of the Environment (CCME). (2002). "E. coli." Retrieved 16 February, 2008, from http://www.ccme.ca/sourcetotap/ecoli.html.

CBC News Online. (2004). "Inside Walkerton: Canada's worst-ever E. coli contamination." *CBC News In Depth*. Retrieved 27 October, 2007, from http://www.cbc.ca/news/background/walkerton/ .

Center for Disease Control and Prevention (CDCP). (n.d.) "Escherichia coli Infection and Farm Animals." Retrieved 27 October, 2007, from http://www.cdc.gov/healthypets/diseases/ecoli.htm.

Clean Annapolis River Project (CARP). (2007). "Annapolis River Guardians." Retrieved 16 March, 2008, from http://www.annapolisriver.ca/guardians.htm.

Cliche, L. and Gaul, J. (2007). "Annapolis Watershed Aquatic Habitat Restoration Project: Final Project Report." *Clean Annapolis River Project.* Retrieved 5 March, 2008, from http://www.annapolisriver.ca/downloads/aquatic%20_habitat_restoration_2007.pdf.

Conboy, M.J., and Goss, M.J. (2000). "Natural protection of groundwater against bacteria of fecal origin." *Journal of Contamination Hydrology*, 43(1), 1.

Crane, S.R., and Moore, J.A. (1986). "Modeling enteric bacterial die-off: a review." *Water, Air, and Soil Pollution*, 27, 411.

Curriero, F.C., Patz, J.A., Rose, J.B., and Lele, S. (2001). "The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948-1994." *American Journal of Public Health*, 92(8), 1194.

Davies, C.M. and Evison, L.M. (1991). "Sunlight and the survival of enteric bacteria in natural waters." *Journal of Applied Bacteriology*, 70(3), 265.

Davies-Colley, R.J., Donnison, A.M., and Sped, D.J. (2000). "Towards a mechanistic understanding of pond disinfection." *Water Science Technology*, 42(10-11), 149.

Doyle, M.P., and Erickson, M.C. (2006). "Closing the door on the fecal coliform assay." *Microbe*, 1, 162.

Gagliardi, J.V., and Karns, J.S. (2000). "Leaching of Escherichia coli O157:H7 in diverse soils under various agricultural management practices." *Applied Environmental Microbiology*, 66(3), 877.

Gerba, C.P., Smith, E.M., and Melnick, J.L. (1977). "Development of a quantitative method for detecting enteroviruses in estuarine sediments." *Applied Environmental Microbiology*, 34(2), 158.

Gleick, P.H. (1998). *The World's Water 1998-1999: The Biennial Report on Freshwater Resources*. Washington D.C., U.S.A.: Island Press.

Gloubeman, S. (2001). "Walkerton Water and Complex Adaptive Systems." *Hospital Quarterly*, 4(4), 28.

Grabow, W.O.K. (2001). "Bacteriophages: update on application as models for viruses in water." *Water South Africa*, 27(2), 251.

Grimes, D.J. (1975). "Release of sediment-bound fecal coliforms by dredging." *Applied Microbiology*, 29(1), 109.

Jones, D.L. (1999). "Potential health risks associated with the persistence of Escherichia coli O157 in agricultural environments." *Soil Use Management*, 15(2), 76.

Khaleel, R., Reddy, K.R., and Overcash, M.R. (1980). "Transport of potential pollutants in runoff water from land areas receiving animal wastes: a review." *Water Resources*, 14, 421.

Labelle, R.L., and Gerba, C.P. (1980). "Influence of estuarine sediment on virus survival under field conditions." *Applied Environmental Microbiology*, 39(4), 749.

Long, R.B. (1983). "Measuring the contribution of land and water to agriculture." *Water Resources Bulletin: American Water Resources Association*, 19(4), 613.

McDonald, J.H. (2008). "Kruskal-Wallis test and Mann-Whitney U test." *Handbook of Biological Statistics*. Retrieved 25 March, 2008, from http://udel.edu/~mcdonald/statkruskalwallis.html.

McMullen, T. (2007). "Various Topics." *Environmental Field School, Dalhousie University*. Nova Scotia.

Ogden, I.D., Fenlon, D.R., Vinten, A.J.A., and Lewis, D. (2001). "The fate of Escherichia coli O157 in soil and its potential to contaminate drinking water." *International Journal of Food Microbiology*, 66, 111.

Sullivan, D. (2004). "Microbial Source Tracking (MST): Towards Effective Identification of Fecal Pollution Sources." *Clean Annapolis River Project MST Applications Workshop Final Report*. Retrieved 4 March, 2008, from http://www.annapolisriver.ca/downloads/mst_report.pdf.

Sullivan, D. and Sharpe, A. (2005). "Annapolis River Guardians Volunteer Water Quality Monitoring Program." *Clean Annapolis River Project*. Retrieved 5 March, 2008, from http://www.annapolisriver.ca/downloads/RG2004reportMarch10.pdf.

Stevik, T.K., Ausland, G., Hanssen, J.F., and Jenssen, P.D. (1999). "The influence of physical and chemical factors on the transport of E. coli through biological filters for wastewater purification." *Water Resources*, 33(18), 3701.

Synova BioScience Corporation. (2006). "FundyBay.Com – More About Fecal Coliform Bacteria." *Annapolis River Guardians Program*. Retrieved 16 February, 2008, from http://www.fundybay.com/datahut/datahut.htm.

Tiedemann, D.A., Higgins, D.A., Quigley, T.M., Sanderson, H.R., and Marx, D.B. (1987). "Responses of fecal coliform in streamwater to four grazing strategies." *Journal of Range Management*, 40(4), 322.

Timmer, D. (2003a). "Annapolis Valley Report." *Guelph Water Management Group*. Retrieved 5 March, 2008, from http://www.uoguelph.ca/gwmg/wcp_home/Pages/A_wi.htm.

Timmer, D. (2003b). "Local water quantity management in the Annapolis Valley." *Guelph Water Management Group*. Retrieved 5 April, 2008, from http://www.uoguelph.ca/gwmg/wcp_home/Pages/A_ia_lqn.htm

Timmer, D. (2003c). "Watershed stakeholders – Annapolis Valley." *Guelph Water Management Group*. Retireved 5 April, 2008, from http://www.uoguelph.ca/gwmg/wcp_home/Pages/A_ws.htm

Thelin, R., and Gifford, G.F. (1983). "Fecal coliform release patterns from fecal material of cattle." *Journal of Environmental Quality*, 12, 57.

Thurston, J.A., Gerba, C.P., Foster, K.E., and Karpiscak, M.M. (2001). "Fate of indicator microorganisms, Giardia and Cryptosporidia in subsurface flow constructed wetlands." *Water Resources*, 35(6), 1547.

United Nations Information Service (UNIS). (2002). "WaterYear2003: International year aims to galvanize action on critical water problems." *Press Release*. Note No. 226. Vienna, Austria.

United States Environmental Protection Agency (USEPA). (2006). "Fecal Bacteria." *Monitoring and Assessing Water Quality*. Retrieved 16 February, 2008, from http://www.epa.gov/volunteer/stream/vms511.html.

United States Environmental Protection Agency (USEPA). (2007). "Colilert, Colilert-18 and Colisure coliform and E. coli water analysis." *Region 9 Laboratory*. Retrieved 4 March, 2008, from http://www.epa.gov/region09/lab/sops/sop1103.html.

United States Forestry and Natural Resources (USFNR). (*n.d.*). "Lesson 3: Survival Skills." *Unit 18: Outdoor Safety and Survival Skills*.

University of Western Ontario, Faculty of Medicine and Dentistry (UWOFMD). (2007). "Case study for the 4th year course in ecosystem health: Walkerton E. coli outbreak." *Ecosystem Health*. Retrieved 27 October, 2007, from http://www.med.uwo.ca/ecosystemhealth/education/casestudies/walkertonenviro.htm.

# APPENDIX A – Tables

**Table 1.** Relative Contributions of Fecal Coliforms (FC) for Selected Tributaries to the Annapolis River (SWIM Sites)

| Site | FC Geomean (cfu/100ml) | Mean Discharge Rate ($m^3/s$) | FC Contribution (1000 cfu/s) |
|---|---|---|---|
| SA30 – South Annapolis River, upstream of Victoria Rd. | 125 | 0.94 | 1175 |
| FR20 – Fales River, downstream of Highway 201 bridge | 959 | 0.63 | 6042 |
| BR19 – Black River, downstream of Hwy 201 bridge | 246 | 0.47 | 1156 |
| NX20 – Nictaux River, downstream of Hwy 201 bridge | 141 | 2.41** | 3398 |
| RH20 – Round Hill River, upstream of Hwy 201 bridge | 76 | 0.76 | 578 |
| MS20 – Moose River, upstream of Clementsport Dam | 3* | N/A | N/A |

* All fecal coliform results recorded at the CARP laboratory for the Moose River site fell below the acceptable range of colonies per fecal coliform plate (20 to 200 colony forming units per membrane). At this site, every water sample filtered yielded counts of 4 cfu/100ml or lower. Since there were no other results available within the acceptable range, the geomean was calculated using these data. All other sites use data only within the acceptable fecal coliform range.

** Discharge data was obtained from Nova Scotia Power (NSP), which operates a power generation station that regulates flow on the river. Using the times and flows provided by NSP, the total cubic meters of water flowing per sampling day was calculated. The mean discharge was then estimated by taking the average of all sampling days (10 days total) and converting to cubic meters per second. A base flow of $0.17 m^3/s$ was used in the calculations when the station was "off".

**SOURCE:** Modified from Sullivan and Sharpe, 2005.

**Table 2.** Sample Site Identities and Locations for Eight Primary ARG Sample Sites

| Site ID | Site Name | Site Description | UTM Easting | UTM Northing |
|---|---|---|---|---|
| AY40 | Aylesford Road | Bridge at Aylesford Road | 357328.48 | 4987755.13 |
| 00 | Aylesford | Bridge at Victoria Road | 353313.34 | 4985418.70 |
| 13 | Kingston | Bridget at Bridge Street | 346748.46 | 4982480.39 |
| 18 | Wilmot | Bridge at Old Mill Road | 342100.00 | 4979500.00 |
| 25 | Middleton | Bridge at Highway 10 | 336981.58 | 4978044.59 |
| 35 | Lawrencetown | Bridge at Lawrencetown Lane | 329581.15 | 4971984.70 |
| 40 | Paradise | Bridge at Paradise Lane | 325738.51 | 4970620.51 |
| 49 | Bridgetown | Bridge at Queen Street | 318900.00 | 4967621.30 |

**SOURCE:** Modified from Annapolis River Guardians (ARG) archival data.

**Table 3.** Summary Statistics for ARG Sample Sites

| | *AY40* | *00* | *13* | *18* | *25* | *35* | *40* | *49* | *Complete Dataset* |
|---|---|---|---|---|---|---|---|---|---|
| **No. of entries** | 44 | 176 | 192 | 172 | 168 | 185 | 167 | 202 | **2047** |
| **Min.** | 2 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | **0** |
| **Max.** | 2420 | 23400 | 3000 | 15000 | 15000 | 3750 | 15000 | 7000 | **23400** |
| **Mean** | 509 | 757 | 185 | 301 | 330 | 260 | 329 | 205 | **429** |
| **Median** | 142 | 200 | 91 | 109 | 78 | 78 | 82 | 70 | **100** |
| **Mode** | 166 | 2420 | 0 | 300 | 200 | 70 | 100 | 70 | **0** |
| **% > 0** | 100% | 98% | 96% | 100% | 99% | 98% | 98% | 98% | **94%** |
| **% > 50** | 75% | 79% | 68% | 74% | 65% | 63% | 60% | 60% | **65%** |
| **% > 100** | 61% | 69% | 46% | 53% | 43% | 44% | 41% | 38% | **49%** |
| **% > 200** | 41% | 49% | 23% | 26% | 22% | 24% | 24% | 19% | **30%** |

**Table 4.** P-values for Two-Sample Student's *t*-Tests Comparing Differences in Means for Pairs Amongst the Eight Primary ARG Sampling Sites

| | **AY40** | **00** | **13** | **18** | **25** | **35** | **40** | **49** |
|---|---|---|---|---|---|---|---|---|
| **AY40** | x | 0.444 | 0.015 | 0.139 | 0.053 | 0.031 | 0.026 | 0.003 |
| **00** | 0.444 | x | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| **13** | | | x | 0.046 | 0.333 | 0.581 | 0.746 | 0.264 |
| **18** | | | | x | 0.333 | 0.158 | 0.121 | 0.002 |
| **25** | | | | | x | 0.676 | 0.549 | 0.040 |
| **35** | | | | | | x | 0.840 | 0.098 |
| **40** | | | | | | | x | 0.168 |
| **49** | | | | | | | | x |

**Table 5.** P-values for Student's *t* Tests between Pairs of Months

| | **Jan** | **Feb** | **Mar** | **Apr** | **May** | **Jun** | **Jul** | **Aug** | **Sep** | **Oct** | **Nov** | **Dec** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Jan** | x | 0.238 | 0.832 | 0.032 | 0.297 | 0.087 | 0.004 | 0.093 | 0.010 | 0.151 | 0.135 | 0.429 |
| **Feb** | | x | 0.296 | 0.029 | 0.091 | 0.694 | 0.546 | 0.673 | 0.811 | 0.574 | 0.598 | 0.397 |
| **Mar** | | | x | 0.015 | 0.204 | 0.141 | 0.002 | 0.158 | 0.012 | 0.246 | 0.225 | 0.611 |
| **Apr** | | | | x | 0.004 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| **May** | | | | | x | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.003 |
| **Jun** | | | | | | x | 0.000 | 0.896 | 0.009 | 0.434 | 0.550 | 0.132 |
| **Jul** | | | | | | | x | 0.000 | 0.127 | 0.000 | 0.000 | 0.000 |
| **Aug** | | | | | | | | x | 0.007 | 0.526 | 0.646 | 0.160 |
| **Sep** | | | | | | | | | x | 0.001 | 0.002 | 0.001 |
| **Oct** | | | | | | | | | | x | 0.880 | 0.323 |
| **Nov** | | | | | | | | | | | x | 0.284 |
| **Dec** | | | | | | | | | | | | x |

**Table 6.** P-values for significance in Student's *t* tests between pairs of years

| | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1992** | x | 0.181 | 0.001 | 0.109 | 0.086 | 0.619 | 0.438 | 0.956 | 0.217 | 0.234 | 0.066 | 0.920 | 0.061 | 0.813 | 0.020 | 0.535 |
| **1993** | | x | 0.000 | 0.000 | 0.404 | 0.057 | 0.000 | 0.036 | 0.000 | 0.000 | 0.000 | 0.005 | 0.251 | 0.001 | 0.054 | 0.281 |
| **1994** | | | x | 0.000 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.006 | 0.000 | 0.203 | 0.000 |
| **1995** | | | | x | 0.000 | 0.000 | 0.187 | 0.032 | 0.460 | 0.555 | 0.788 | 0.013 | 0.000 | 0.019 | 0.000 | 0.001 |
| **1996** | | | | | x | 0.009 | 0.000 | 0.009 | 0.000 | 0.000 | 0.000 | 0.001 | 0.749 | 0.000 | 0.217 | 0.094 |
| **1997** | | | | | | x | 0.020 | 0.398 | 0.000 | 0.005 | 0.000 | 0.234 | 0.004 | 0.124 | 0.001 | 0.769 |
| **1998** | | | | | | | x | 0.311 | 0.462 | 0.502 | 0.081 | 0.245 | 0.000 | 0.332 | 0.000 | 0.037 |
| **1999** | | | | | | | | x | 0.085 | 0.120 | 0.010 | 0.953 | 0.005 | 0.794 | 0.001 | 0.352 |
| **2000** | | | | | | | | | x | 0.959 | 0.249 | 0.037 | 0.000 | 0.055 | 0.000 | 0.004 |
| **2001** | | | | | | | | | | x | 0.364 | 0.076 | 0.000 | 0.107 | 0.000 | 0.010 |
| **2002** | | | | | | | | | | | x | 0.002 | 0.000 | 0.003 | 0.000 | 0.000 |
| **2003** | | | | | | | | | | | | x | 0.000 | 0.799 | 0.000 | 0.245 |
| **2004** | | | | | | | | | | | | | x | 0.000 | 0.338 | 0.057 |
| **2005** | | | | | | | | | | | | | | x | 0.000 | 0.160 |
| **2006** | | | | | | | | | | | | | | | x | 0.013 |
| **2007** | | | | | | | | | | | | | | | | x |

# APPENDIX B – Figures

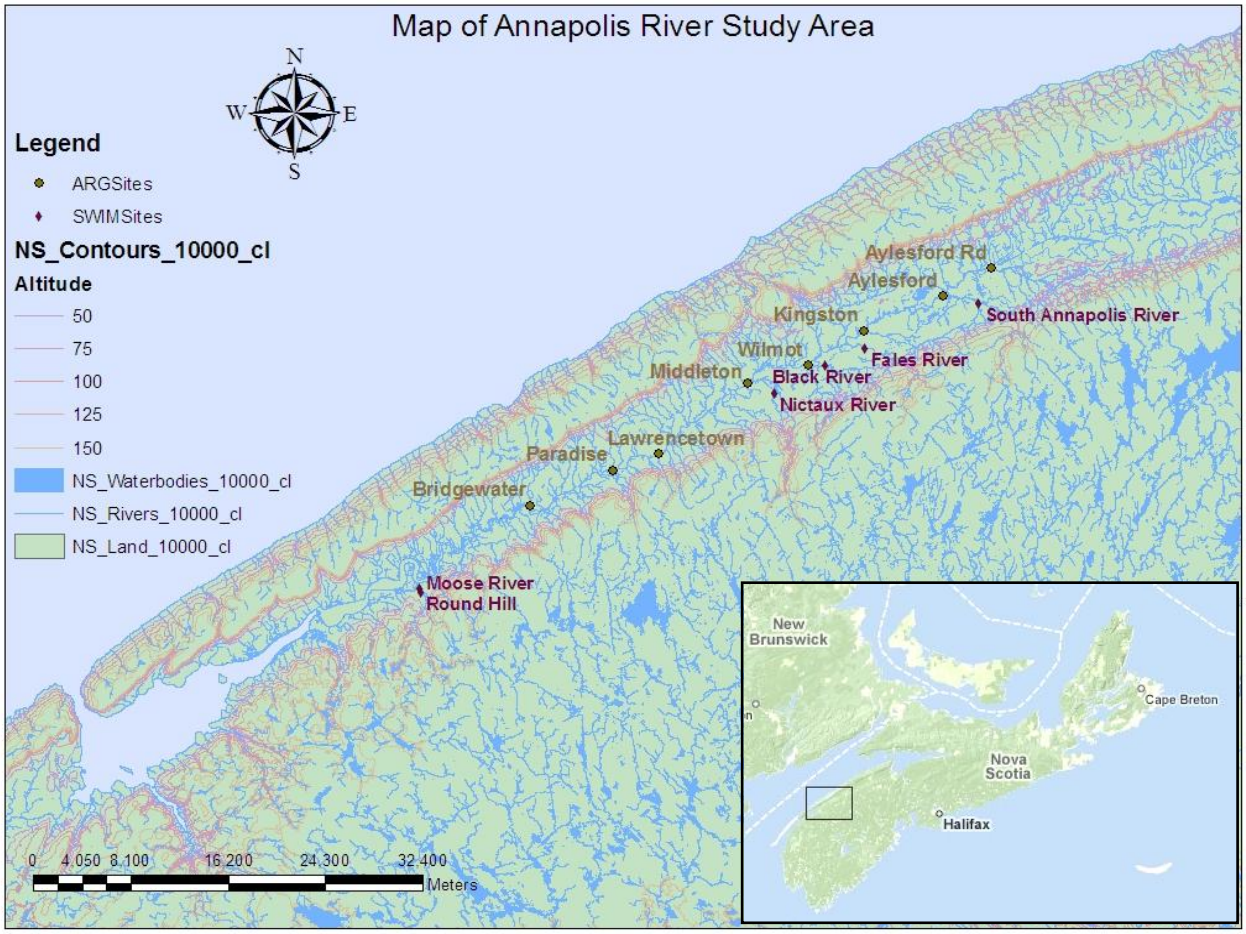


**Figure 1.** Map of Annapolis River study area. Brown dots represent eight primary Annapolis River Guardian (ARG) sample sites. Smaller purple diamonds represent six Sub-Watershed Investigative Monitoring (SWIM) sample sites. Elevation contours, surface water bodies, and rivers also included. Inset at bottom right shows location of map with respect to Nova Scotia. Base map data courtesy of Access Nova Scotia and Dalhousie University's Maps and Geospatial Information Centre (MAGIC).

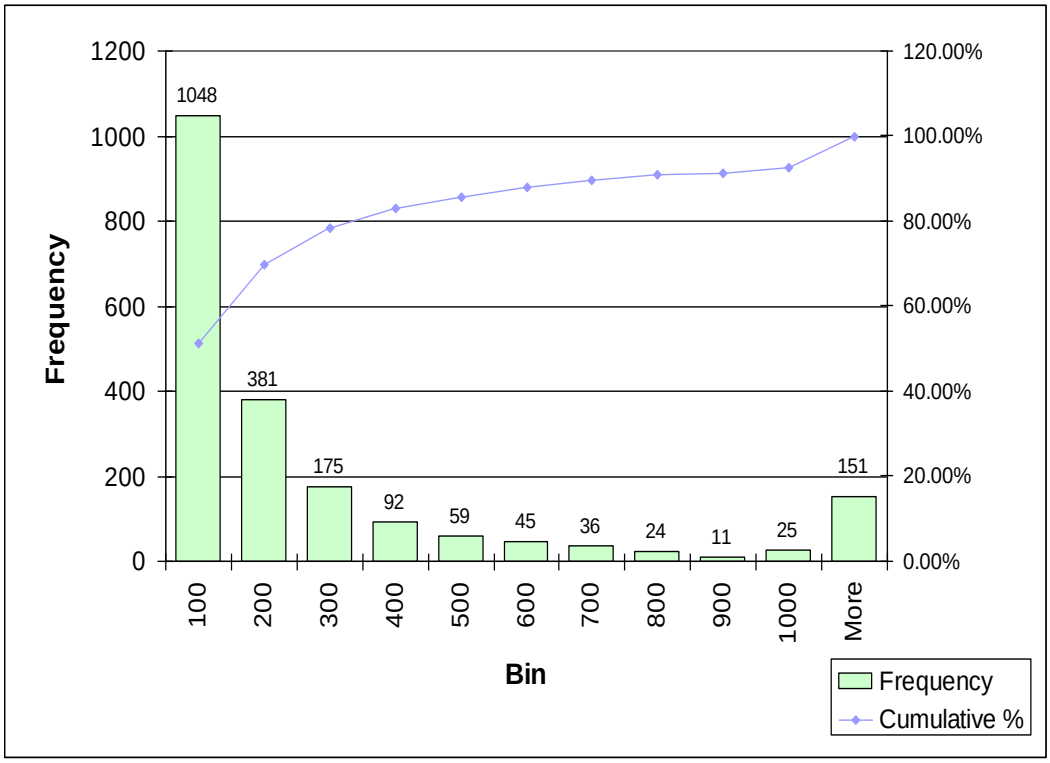


**Figure 2.** Histogram showing the frequency distribution for measurements of bacterial concentrations (combined *E. coli* and fecal coliform records).

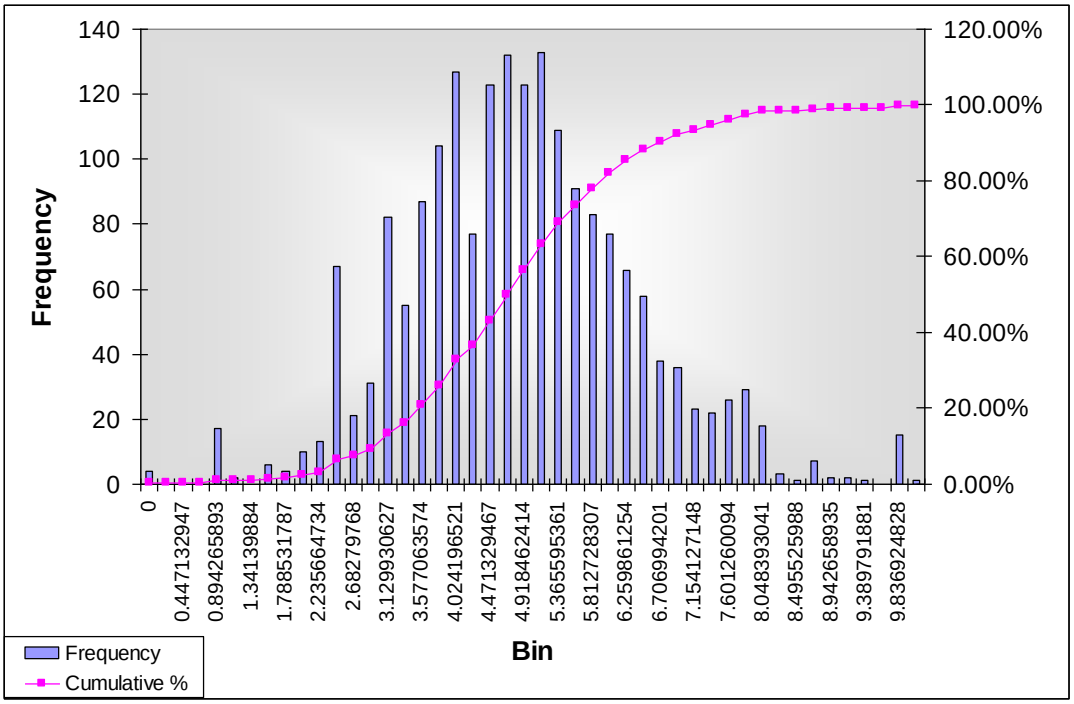


**Figure 3.** Histogram showing the frequency distribution for measurements of bacterial concentrations (combined *E. coli* and fecal coliform records) transformed using natural logarithm.

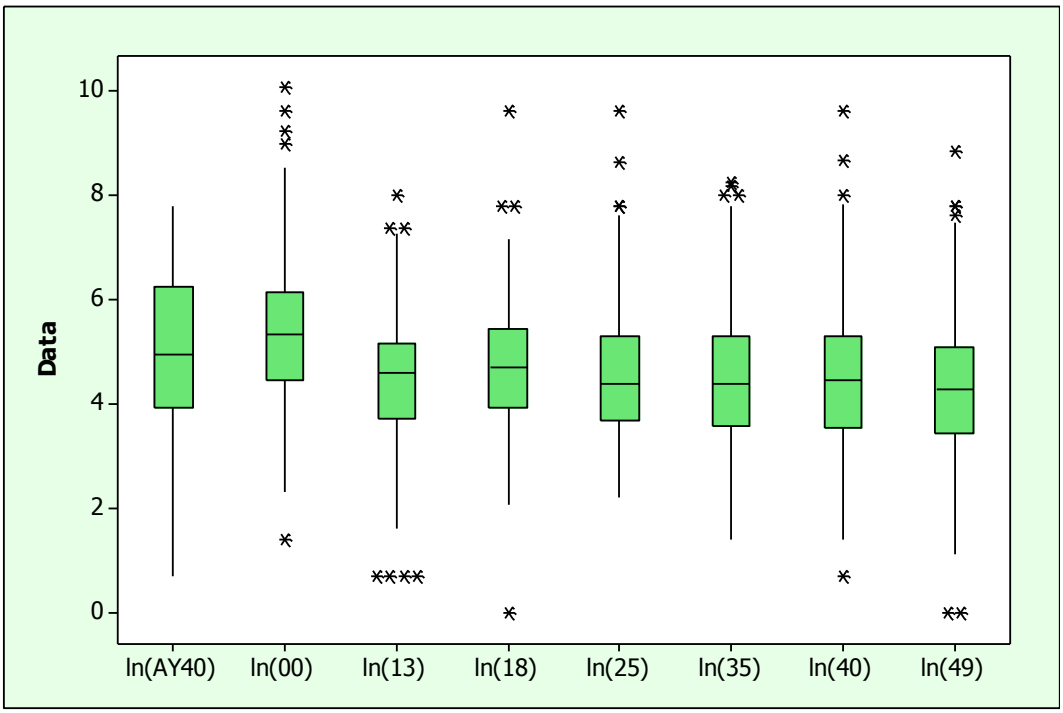


**Figure 4.** Box-and-whisker plot comparing logarithmically transformed measurements of bacterial indicator concentrations (amalgamated *E. coli* and fecal coliform record) among eight primary ARG sampling locations in the Annapolis River. Outliers (more than three standard deviations from the mean) are denoted by *.

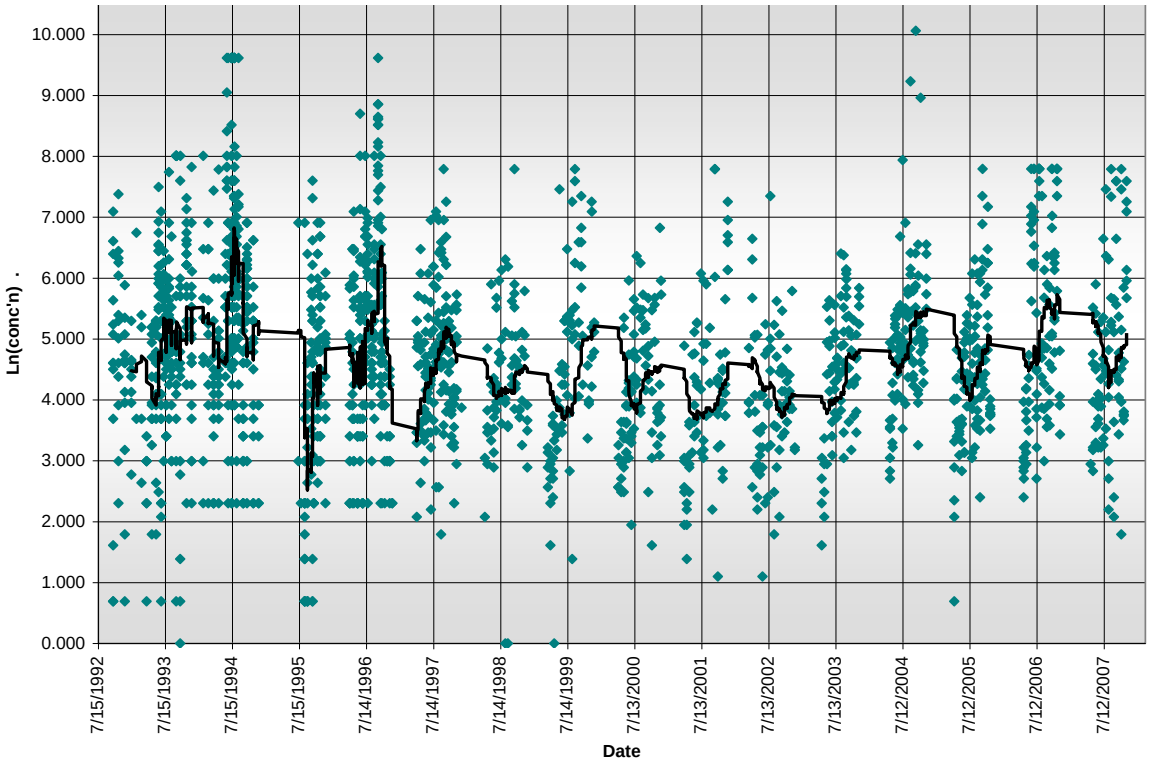


**Figure 5.** Time series plot relating logarithmically transformed concentrations of bacterial indicators (combined *E. coli* and fecal coliform record) versus sampling date for all available samples within the Annapolis River Guardians data archives (1992-2007, all sample sites).

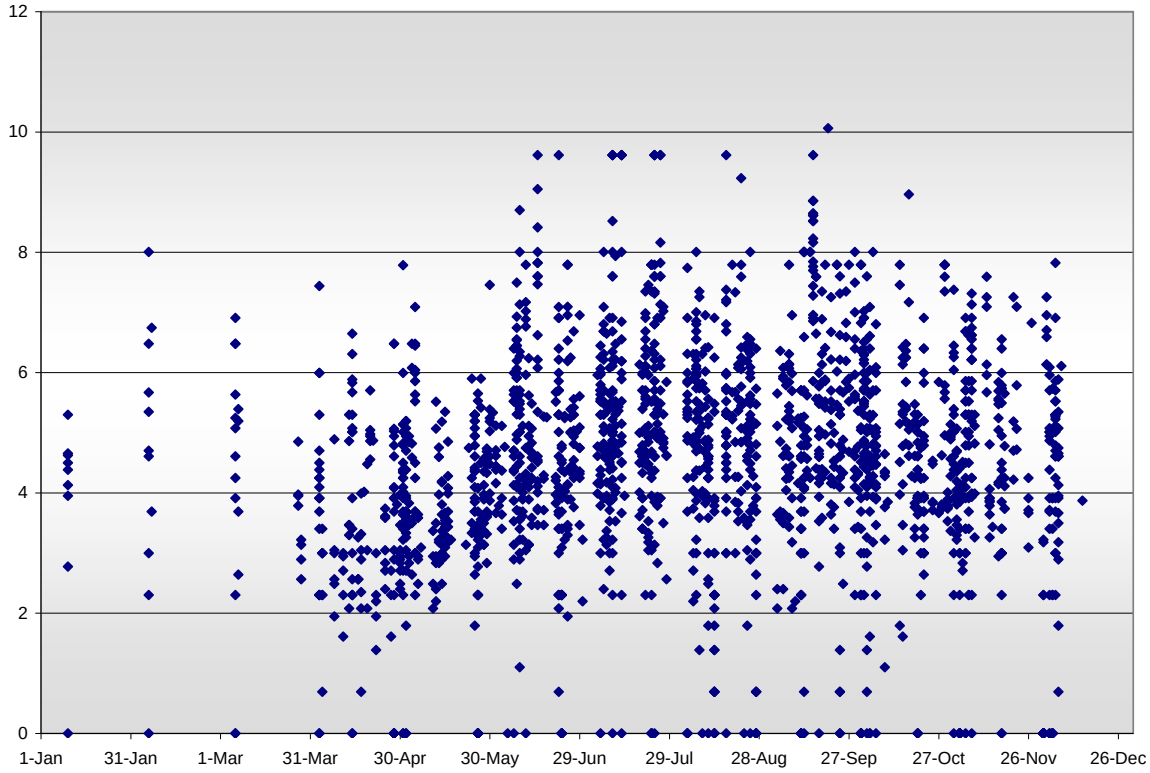


**Figure 6.** Scatter plot relating logarithmically transformed bacterial concentrations (*E. coli* and fecal coliforms) to sampling date (year omitted) to portray seasonality of concentrations.

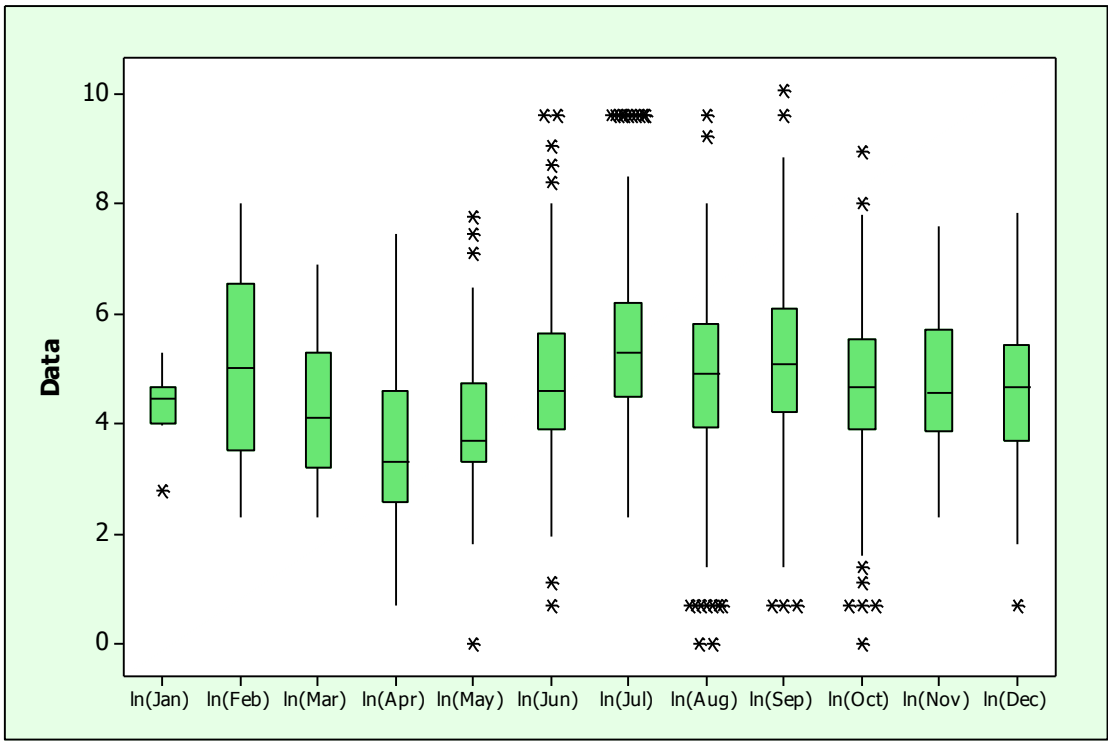


**Figure 7.** Box-and-whisker plot showing distribution of logarithmically transformed fecal bacterial concentrations from all available Annapolis River samples in the ARG data archives (1992-2007, all sampling sites), grouped by sample month (independent of year). Outliers (more than three standard deviations from the mean) are shown as *.

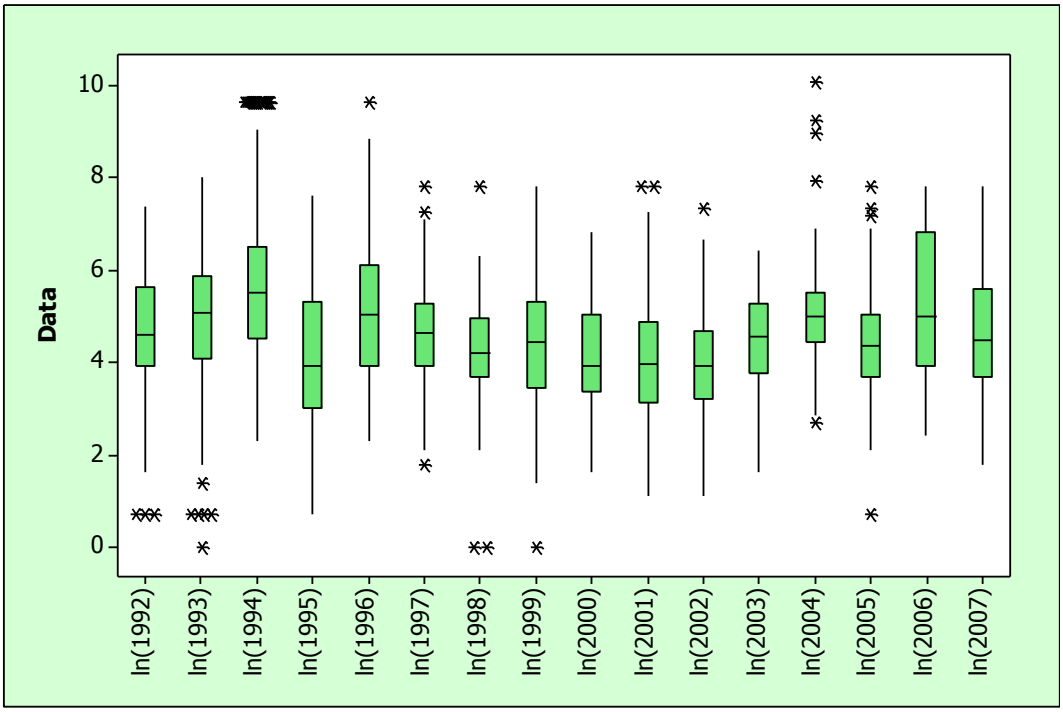


**Figure 8.** Box-and-whisker plot comparing logarithmically transformed measurements of bacterial indicator concentrations in the Annapolis River (combined *E. coli* and fecal coliform records for complete dataset) in each year sampled as part of the ARG program (1992 to 2007). Outliers (more than three standard deviations from the mean) are denoted by *.

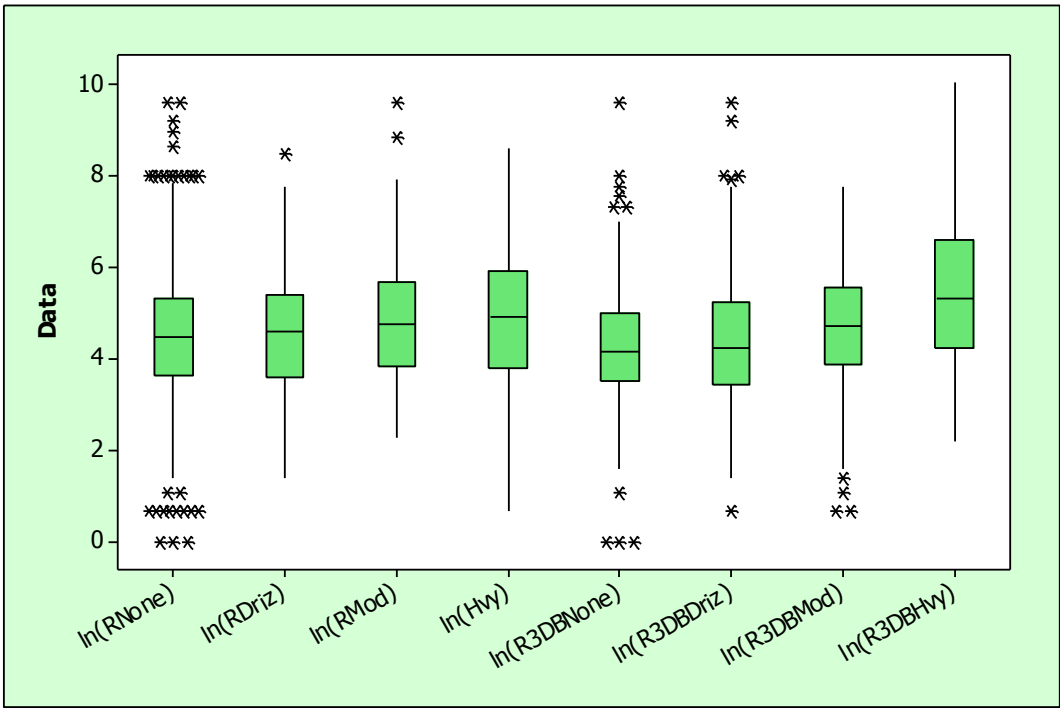


**Figure 9.** Box-and-whisker plot showing distributions of logarithmically transformed measurements of bacterial indicator concentrations in the Annapolis River for all 49 sampled sites through all sampled years, grouped by rain rate the day of sampling (left four boxes) and by rain rate three days before sampling (right four boxes). Outliers (more than three standard deviations from the mean) are denoted by *.

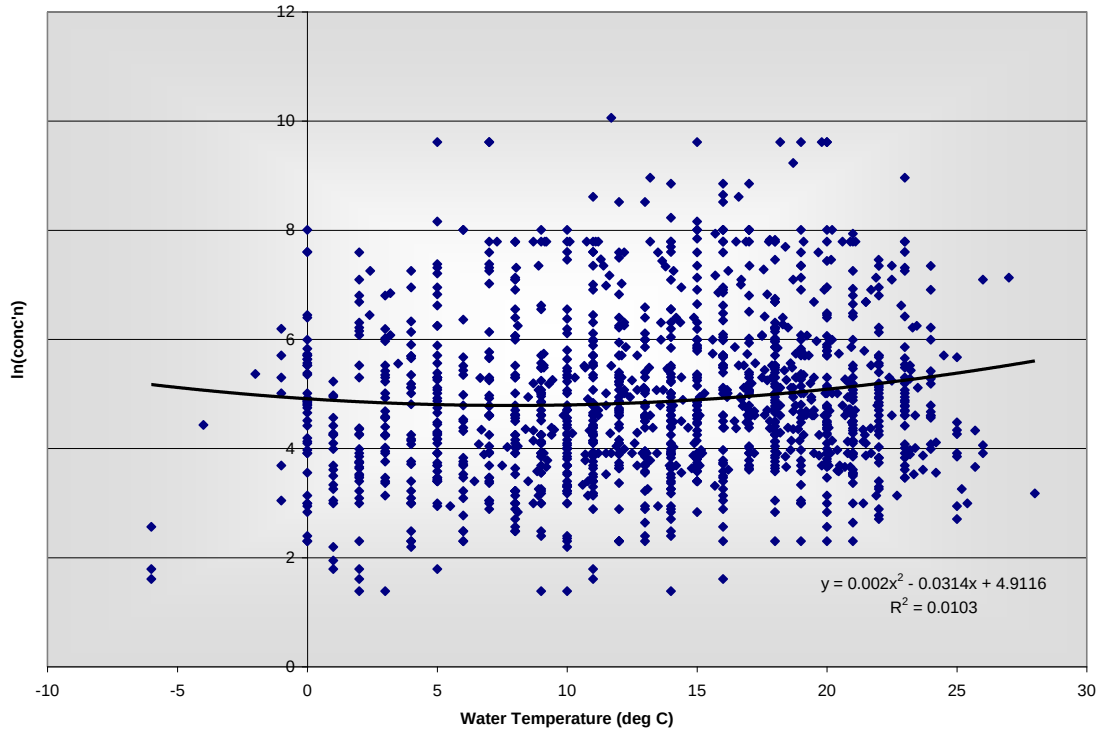


**Figure 10.** Scatterplot and quadratic best fit trendline for logarithmically transformed bacterial concentrations versus water temperature at the time of sampling for the complete Annapolis River dataset.

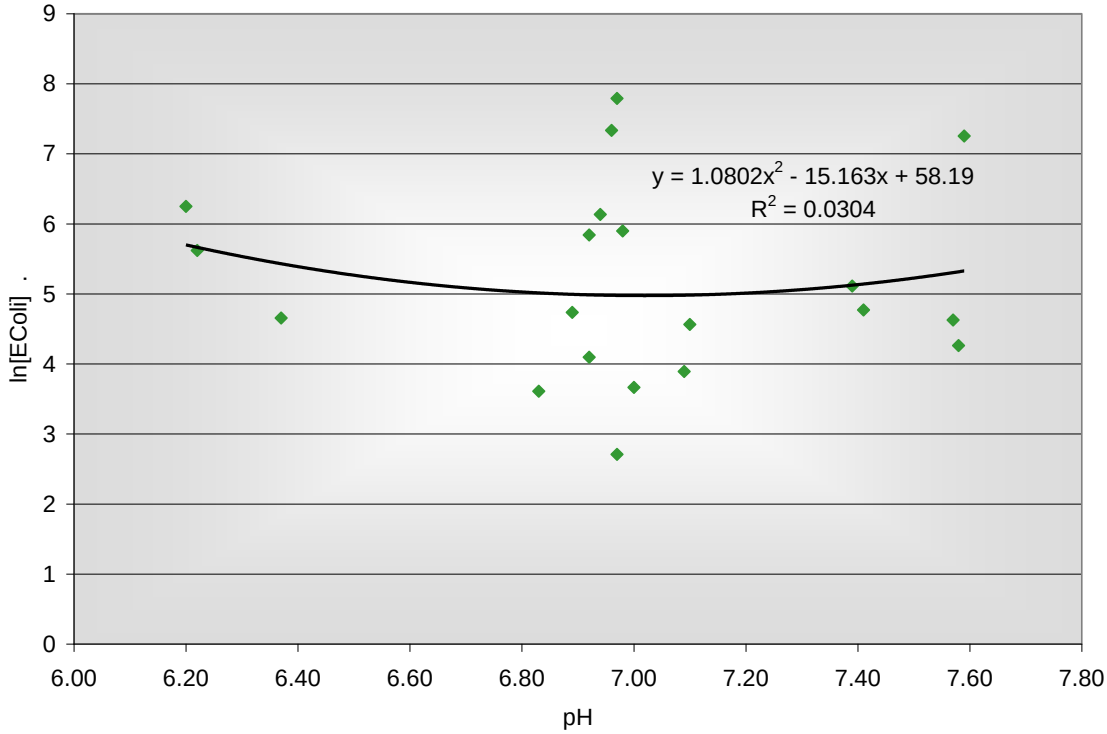


**Figure 11.** Scatter plot and quadratic best fit line comparing logarithmically transformed *E. coli* concentrations to pH at the time of sample retrieval for site AY40 in the Annapolis River.