ABSTRACT

Title of Dissertation: DYNAMICS AND DRIVERS OF FECAL INDICATOR BACTERIA AND ASSOCIATED BACTERIAL COMMUNITY MEMBERS IN ESTUARINE WATERS

Andrew Karl Leight
Doctor of Philosophy
2016

Dissertation directed by: Dr. Raleigh Hood
University of Maryland Center for Environmental Science

For over a century, specific types of bacteria have been monitored in natural water bodies as indicators of fecal pollution and increased risk of encountering human pathogens. One such type of bacteria is the fecal coliforms, a group of gram-negative, facultative anaerobes mostly from the Class Gammaproteobacteria and the Family Enterobacteriaceae, which are commonly found in the gut of warm-blooded animals. In the Chesapeake Bay, routine monitoring of coliform bacteria has been conducted since the 1920's to assess the likelihood of sewage pollution in shellfish harvest areas. The research for this dissertation examined the dynamics and drivers of fecal coliforms and potential pathogen groups in Maryland waters. First, the impacts of climate variability on densities of fecal coliforms in surface waters were examined, finding that annual precipitation and air temperature levels correlate well

to the proportion of stations with fecal coliforms in excess of the established regulatory criteria. A dominant climate pattern was identified for years with extreme precipitation and fecal coliform levels. Secondly, the validity of using precipitation totals as indicators of fecal coliform densities exceeding the regulatory criteria was examined. Precipitation levels over the previous two days were related to fecal coliforms in excess of the criteria for particular watersheds, depending on the percent of open water; non-tidal, forested wetlands; and soil types. The level of precipitation required to cause fecal coliform densities to exceed the FDA criterion varied between watersheds. Thirdly, high-throughput sequencing of 16S rRNA genes was used to study the community of bacteria at a long-term monitoring station in order to characterize community members over the course of 5 months. Water temperature and turbidity were found to be related to changes in community composition at the scale of Genera, while precipitation was a key driver for the presence of allochthonous bacteria such as fecal coliforms. The co-occurrence of some bacteria groups at the Class level of phylogeny was largely defined by the arrival of allochthonous groups into the autochthonous community. Further, a novel approach for estimating densities of bacteria from 16S rRNA amplicon pools was explored.

DYNAMICS AND DRIVERS OF FECAL INDICATOR BACTERIA AND ASSOCIATED BACTERIAL COMMUNITY MEMBERS IN MARYLAND WATERS

by

Andrew Karl Leight

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2016

Advisory Committee:
Professor Raleigh Hood, Chair
Dr. Byron Crump
Dr. Amy Sapkota
Dr. Mark Frana
Dr. Daniel Fisher

Dedication

This dissertation is dedicated to my wife, Susan, and our two daughters, Sophie and Anna Kate, who supplied unending encouragement and enthusiasm along the way.

Acknowledgements

Primary and heartfelt thanks go to the members of my research advisory committee, particularly to Byron Crump and Raleigh Hood who served as major advisor consecutively. All of my research committee members provided excellent guidance and taught me much. They were a diverse and impressive team. Much of the research for this dissertation leveraged monitoring data collected by the Maryland Department of the Environment. Kathy Brohawn, Heather Merritt, and Keeve Brine from MDE provided the data and answered numerous questions about how and why it was collected. Bill Beatty, Billy Evans, Rusty McKay, and Quentin Forrest were among the many individuals from MDE responsible for collection of samples. Tsvetan Bachvaroff, of the University of Maryland Center for Environmental Science Institute of Marine and Environmental Technology, provided access to and assistance with the Linux machine used for bioinformatics. Jason Steele and Juan Maldonado of the Center for Personalized Diagnostics, The Biodesign Institute, Arizona State University developed the DNA library and conducted the 16S sequencing. Lindy Fine provided instructions and advice on DNA extraction. Thomas Sharpton and Ryan Mueller of the Department of Microbiology, School of Agricultural Sciences, Oregon State University consulted on the use of an internal control as a means of estimating 16S rRNA gene densities. The entirety of this dissertation research occurred while I was employed by the National Oceanic and Atmospheric Administration at the Cooperative Oxford Lab. Many thanks go to the leaders of that lab, Bob Wood and Suzanne Skelley, respectively, and to all my coworkers, past and present.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

16S rRNA	16 Subunit Ribosomal Ribonucleic Acid; portion of the gene found in all prokaryotes responsible for coding part of the RNA making up the ribosome; differences in the DNA sequence of the gene are used to classify prokaryotes.
FC	Fecal Coliforms; a.k.a thermotolerant coliforms, presumptive gram–negative fecal bacteria that ferment lactose, producing acid and gas, in selective growth media incubated at 44.5°C.
m	Meters
mL	Milliliters
MPN	Most Probable Number; Use originates in early 1900's; Involves the practice of diluting 'portions' of water samples in selective media and counting containers that contain evidence of bacterial growth (e.g. gas and acid production) as an estimate of bacteria densities.
PDiff	Percent Difference; specific to Chapter 3 of this dissertation; the difference in percent of stations over the FDA criterion in dry periods compared to the percent over the criterion following an inch or more of rain.
TMDL	Total Maximum Daily Load; a calculated estimate of the total amount of a potential pollutant (e.g. excessive nutrients or fecal matter) that may be delivered to a body of water each day in order to bring the waterbody into compliance with water quality criteria.

.

Chapter 1: Introduction

The Chesapeake Bay, like many estuaries, suffers from excess nutrients, chemical contaminants, and human pathogens, which impact both the environmental quality and sustainable use of natural resources. In many of waterbodies, access to natural resources may be restricted due to poor water quality or the natural resource contamination caused by these pollutants. For example, restricted access to shellfish harvest areas and recreational waters occurs in many coastal areas due to high levels of fecal pollution in the surrounding waters or shellfish tissues. However, when making decisions about restricting resource use, managers of coastal environments often rely on arcane assessment and management tools that may not always be protective, or that are sometimes overprotective, of natural resource condition and human health. This is often the case for the management of waters relative to waterborne and foodborne pathogens.

Although our understanding of waterborne pathogens has increased over the last several centuries and reductions in pathogen presence in developed countries has lowered disease incidence, illnesses still occur and managers of environmental and public health require improved tools for pathogen monitoring and resource protection. These managers would benefit from having better understanding of what drives the presence of pathogens and their indicators over time and space, how these indicators relate to the pathogens they are intended to detect, and if there are better indicators of waterborne pathogen presence. They would also benefit from information about how to account for the influences that changes in land use and climate may have on the distribution and abundance of waterborne pathogens or their indicators for the protection of environmental and public health.

Worldwide, there are about one billion illnesses and 2.2 million deaths per year from waterborne pathogens, most of these cases occurring in developing countries (Montgomery

and Elimelech 2007). In the US for the years 2009 and 2010 combined there were 24 cases of disease outbreak related to recreational activities in natural water bodies, affecting 296 individuals and causing 22 hospitalizations (CDC 2014). However, these numbers likely underestimate the actual number of waterborne illnesses from recreational activities (Craun et al. 2006) as the estimates derive from voluntary reports to the Centers of Disease Control (CDC), they do not capture single case events, and do not account for cases where the ill person did not seek medical attention or did not attempt to discover where contact with the pathogen occurred. Significant numbers of illness have also been documented from shellfish consumption in the United States. For example, *Vibrio* bacteria accounted for thousands of cases of human illness from shellfish consumption over the period of 1996-2010, including hundreds of hospitalizations and about 70 deaths (Newton et al. 2012). The economic impact from these waterborne and seafood consumption-related illnesses is significant, with total annual costs for the US estimated at roughly $350 million (Ralston et al. 2011). These are relevant topics in the Chesapeake Bay where pathogens of human health concern, such as species of *Vibrio*, *Clostridium*, *Klebsiella*, and *Salmonella,* have been isolated from Bay waters and shellfish (Sayler et al. 1975, Morgan et al. 1976). Some of these pathogens tend to be related to fecal pollution while others, such as *Vibrio*, tend to be naturally occurring in waterbodies.

The ability to protect humans from waterborne and seafood related illness has long challenged science. It was not until the late 19th century that microbes were identified as the causative agents of illness from contact with water or consumption of seafood (Ashbolt et al. 2001). A primary concern and focus, both then and now, centers around enteric pathogens from sewage pollution impacting recreational beaches and shellfish growing areas. As more and more waterborne pathogens were discovered it became clear that direct detection of the suite of pathogens found in most aquatic areas would be impossible and detecting both

enteric pathogens resulting from sewage pollution or manures and endemic pathogens, separate from the vast majority of non-pathogenic microbes, would be challenging. There was a need for an easily measured indicator that would signal the presence of human pathogens from fecal pollution and therefore help assess the relative risk of infection. A relatively fast and simple method of detecting and quantifying coliform bacteria, thought to be associated with feces and good indicators of fecal pollution, was developed in 1905 (Macconkey 1905) and soon after became the standard for assessing the presence of fecal pollution in the aquatic environment (Ashbolt et al. 2001). Coliforms were defined as those bacteria resembling *Bacillus coli* (later renamed *Escherichia coli*) in that they ferment lactose, producing acid and gas. Following an outbreak of typhoid fever related shellfish consumption in 1924-1925, a US report recommended using *Bacillus coli* culture from surface waters as an indicator of fecal pollution (Hunt 1977). It was recognized by the early 20th century, however, that the coliform group included a range of bacteria, not exclusively of fecal origin (Winslow and Walker 1907). The edition of higher temperature incubation at 44°C was added to help select for thermostable coliform bacteria, which thrive in the elevated temperatures of warm-blooded animals but also include non-fecal members (Ashbolt et al. 2001). This group of bacteria was dubbed fecal coliforms.

The use of fecal coliform levels as indicators for monitoring shellfish meat and growing area waters also became codified in methods for assessing seafood safety and for general water quality. For example, the passing of the Federal Water Pollution Control Act (a.k.a Clean Water Act) in 1972 (FWPCA 1972) prompted the US Environmental Protection Agency (EPA) to establish water quality criteria for a number of pollutants, including fecal bacteria (EPA 1976). The EPA conducted studies examining rates of infection for people swimming in recreational waters and found that densities of fecal coliform bacteria provided a suitable indicator of relative risk of swimmers contracting gastrointestinal disease (EPA

1976). Though this effort lacked much epidemiological rigor, it led to the establishment of water quality criteria for indicator bacteria in recreational and shellfish harvesting waters. The EPA officially adopted the use of fecal coliform counts as an indicator of fecal pollution, due principally to a relationship between the occurrence of fecal coliforms and *Salmonella* (EPA 1976). In the 1970s and 1980s, advances in microbial culture techniques allowed the use of more specific groups of bacteria as indicators. Additional testing was conducted by the US EPA examining the concentrations of a handful of potential indicators (EPA 1986). The densities of the two bacterial groups *E. coli* and *Enterococcus spp.*, which are common in the feces of warm-blooded animals, were found to have better relationships to the incidence of gastrointestinal illness in swimmers than fecal coliforms. However, in 1973, the National Shellfish Sanitation Program (NSSP) reviewed shellfish criteria and decided that fecal coliforms were "scientifically and logistically superior" to the (total) coliform or fecal streptococci indicator groups as a microbiological indicator of fecal pollution in estuarine waters (Hunt 1977). Based on this decision, and the fact that many management agencies wanted to maintain data comparability over time and space, they choose to continue monitoring fecal coliforms at shellfish harvest areas. Despite concerns over the efficacy of using fecal coliforms as indicators of fecal pollution (EPA 1986, Wade et al. 2003, Schwab 2007), they remain the test group for assessment of human health risk from pathogens at shellfish harvest areas in many US states.

In the US, the NSSP establishes and publishes guidelines for the management of shellfish growing areas and for post-harvest handling of commercial shellfish catches. The NSSP is a cooperative partnership, started in 1925, between the US Federal Government and state agencies that is recognized by the US Food and Drug Administration and managed by the Interstate Shellfish Sanitation Conference. Their guidelines include criteria for acceptable levels of indicator bacteria, principally fecal coliforms, in surface waters of shellfish harvest

areas (FDA 2015). There are various criteria and the appropriate ones for a particular harvest area depend on the type of waters being monitored and their proximity to sources of human waste. Classification of shellfish beds into Approved, Restricted, and Conditionally-Approved categories is based on surface water measurements of indicator bacteria, shoreline surveys to identify potential contamination sources and, in the case of conditionally-approved areas, some environmental trigger, such as a rainfall event or the presence of seasonal boaters.

Before the advent of sewage treatment plants, minimally processed sewage was dumped directly into coastal waters, such as the Chesapeake Bay. Since that time there has been a gradual process of upgrading wastewater treatment plants to reduce solid matter, include disinfection of effluent and, more recently, the loadings of both bacteria and nutrients. With the improvements that have been made in waste treatment and septic systems, fecal coliform densities may be expected to decrease in the Bay as a whole, with significant improvements in areas where best management practices aimed at decreased wastewater introduction have been implemented. Although improvements to wastewater treatment systems have reduced the likelihood of raw sewage entering the Bay, breaches from large storms and system failures can still result in large introductions of fecal matter. Other sources include failing septic systems, wildlife, and livestock (MDE 2006), which require separate management strategies for decreasing loads to the Bay. One such management strategy has been increased requirements for water management on agricultural and urban lands in hopes of improving water quality by reducing the introduction of various pollutants from these lands. In contrast, the human population in the Chesapeake Bay watershed has increased dramatically and the amount of waste and the number of waste treatment facilitates have increased in response. Although efforts to decrease the presence of fecal bacteria have

shown success, most tributaries of the Chesapeake Bay have some upstream sections restricted to harvest due to excessive levels of fecal indicator bacteria (Figure 1.1).

In Maryland, the Department of Environment began sampling surface waters over shellfish beds for total coliforms in the 1920's and fecal coliform concentrations in the early 1950's. Following the NSSP recommendations (FDA 2015), the State has monitored fecal coliform bacteria densities at more than 1500 shellfish harvesting and sewage outfall areas. Surface water samples from these locations are tested for fecal coliform counts based on the multiple tube fermentation protocol, as described by the American Public Health Association (APHA 1998) and directed by the Food and Drug Administration (FDA 2015). The samples are typically collected monthly or twice a month, using a random sampling schedule, with the intent that sampling will, by chance, include the range of conditions typically found at these stations. MDE uses these fecal coliform data, along with qualitative information from sanitation surveys and other sources, to determine if shellfish harvest areas should be 'restricted' (closed for harvest), 'approved' (open for harvest), or conditionally approved. 'Conditionally approved' areas are closed to harvest for any day following one or more inches of rainfall. For restricted and approved areas, the State uses the median and 90^{th} percentile of counts from at least thirty samples to compare against the NSSP (FDA 2015) criteria to determine the closure status. In conditionally approved areas, the percent of the samples over the 90^{th} percentile FDA criterion must not exceed 10% when condition for closure (rainfall over an inch in this case) has not occurred in order for the area to remain in that classification.

Enormous effort and monies have been spent by Maryland, other coastal U.S. states, and local municipalities to monitor the densities of these indicator bacteria in recreational and shellfish harvest waters. In addition, beginning in the 1990's, the Total Maximum Daily Load (TMDL) process requires states to report waterbodies that violate water quality criteria,

and more significantly, to develop plans to remediate those waterbodies (EPA 1991). Among a collection of various pollutants, fecal bacterial indicators are required to be monitored, and assessments of indicator bacteria loading into various watersheds based on densities have been made (Figure 1.2). In order to identify and improve watersheds with high levels of fecal coliforms in the surface waters, inclusive but not limited to shellfish areas, Maryland state agencies have developed total maximum daily loads (TMDLs) to manage maximum allowable inputs of fecal bacteria from identified sources. Using reverse modeling (MDE 2006), predictive models of fecal coliform concentrations have been developed for specific watersheds. These models incorporate a broad suite of variables, including freshwater input, tidal exchange, wind, salinity and water temperature. These models tend to produce good predictions of surface water concentrations but have not been used to evaluate rainfall thresholds for fecal coliform introductions throughout the Bay. The TMDL process further emphasizes the need to better understand the ecological implications of indicator bacteria presence, to estimate the sources and sinks of indicator bacteria, and develop tools to more accurately and rapidly assess contamination.

The goal of improving and sustaining environmental and public health, however, involves not only the assessment of current conditions, but an understanding of how conditions change over time and what factors influence the presence of these indicators and associated pathogens in the water. Factors that commonly affect the amount and location of fecal indicator bacteria in the aquatic environment include source type (e.g. sewage, manure, endemic populations), water quality conditions (e.g. temperature, salinity, pH, suspended solids), light intensity, rainfall, sediment resuspension and water mixing (Campos et al. 2013). Additionally, the influence of these factors on pathogens of human health concern is likely to be different than for the indicator bacteria. A number of studies have examined the co-occurrence of pathogens and indicator bacteria in estuarine waters (Carney et al. 1975,

EPA 1976, 1986, Krometis et al. 2010). For example Hood et al. (1983) examined fecal coliforms, *E. coli*, and *Salmonella,* in shellfish from Florida and found that *E.coli* occurrence in the oyster tissue correlated very strongly with the occurrence of *Salmonella*. However, Carney et al. (1975) found that fecal coliforms and pathogens were not closely related. A more recent review of epidemiological studies for recreational waters (Wade et al. 2003) suggests that *E. coli* and *Enterococcus* are suitable indicators of risk of gastrointestinal illness, but fecal coliforms are not. How often and under what conditions indicators and pathogens co-occur remains uncertain, but the use of new molecular tools may be able to help shed light on the subject.

Bacterial communities in natural waters tend to be very diverse (Kan et al. 2006) and have complex and redundant functional groups (Comte and del Giorgio 2010). Until recently, studies of bacterial communities in shellfish harvest waters have relied on culture-based methods that exclude a large number of bacteria and require focused analyses of particular bacterial groups. However, the use of high-throughput DNA sequencing techniques, coupled with ever-improving databases of genetic sequence information for millions of both bacteria and archaea, have provided a means for better assessing the bacterial community in aquatic environments. For example, a study in the Santa Anna River, California using 16S rRNA gene amplicon pyrosequencing detected higher percentages of genera that contain human pathogens in areas of urban runoff and agriculture (Ibekwe et al. 2013). Likewise, a recent study of bacterial community variability at a shellfish harvest area in Spain noted seasonal shifts in the bacterial community and the presence of genera containing human pathogens (Pereira et al. 2015). Similar studies, focused on more highly resolved temporal changes in bacterial communities and potential pathogenic members at shellfish harvest areas would help inform relative human health risks and decisions about restricting shellfish harvest.

Indicator bacteria and related human pathogens in coastal waters tend to be from land-based sources and are discharged to receiving waters from point and nonpoint sources. Precipitation plays a key role in delivering these pollutants (Curriero et al. 2001, Lipp et al. 2001a, Chigbu et al. 2005). Therefore, the density and distribution of fecal indicator bacteria may be linked to climate in several ways. Curriero et al. (2001) found that most waterborne disease outbreaks in the US follow large precipitation events. Fecal coliform concentrations at beaches are also related to rainfall (Ackerman and Weisberg 2003, Kelsey et al. 2004). In a study of fecal coliform densities at shellfish areas in Galveston Bay, Texas, years with higher amounts of precipitation tended to have higher fecal coliform counts (Jensen and Su 1992). Moreover, survival of the fecal bacteria in the environment is related to water temperature (Howell et al. 1996) and salinity (Goyal et al. 1977), both of which are directly related to climate forcing. In addition, fecal bacteria have been found to survive in benthic sediments and have a positive relationship with total suspended solids in sediments, suggesting that wind-driven resuspension may influence their densities in shallow waters (Howell et al. 1996). Lipp et al. (2001a) found that fecal indicator organisms in a Florida estuary became more widespread during winter and fall when precipitation rates increased and noted qualitatively that this trend occurred during an El Niño event. These teleconnections (i.e. connections between climate and environmental phenomenon at large scales) are compelling as local and regional meteorological events, such as storm driven rains and winds, are reflective of large-scale climate conditions (Chapter 2). In the Chesapeake Bay, these climate patterns have been reduced to synoptic climate patterns which have been statistically linked to dynamics in nutrients and phytoplankton (Miller and Harding 2007), crustacean communities (Kimmel et al. 2006), and fish populations (Wood 2000).

Several previous studies of fecal indicator bacterial dynamics in Chesapeake Bay have been conducted, outside of State monitoring programs. Some early studies focused on

seasonal and spatial patterns of indicators and pathogens in discrete river systems (Carney et al. 1975) or in the mainstem of the Bay (Sayler et al. 1975). Erkenbrecher (1981) studied indicator bacteria in sediments and overlying waters in the Chesapeake Bay and found that they tended to be higher in headwaters than inlets, and greater in the sediment than above the bottom. Erkenbrecker also found that fecal indicator numbers were lower in May, July, and August than in spring and fall, despite warmer summer temperatures that would presumably promote fecal bacterial survival and growth. More recently, Huang (2010) examined the temporal and spatial distribution of indicator bacteria in Virginia's section of the Chesapeake Bay and found that a greater frequency of upstream stations had higher densities of fecal coliforms than downstream stations, that cold months (November to March) tended to have lower densities than warm months (April to October), and that tidal effects were less important than location in stream or season. Such a study had not been conducted for Maryland waters before the research described in this dissertation.

This dissertation examines drivers of fecal coliform bacteria variability in Maryland waters and the community of microbes that may co-occur with these indicator bacteria. The goals were to improve understanding of the factors that drive variability in fecal indicator bacteria and associated community members in shellfish harvest waters of the Chesapeake Bay, and provide information necessary to refine management of important and iconic Chesapeake Bay shellfish species. It was shown that levels of fecal coliform bacteria in Maryland shellfish harvest waters are related to inter-annual climate variability and large-scale climate patterns (Chapter 2), that the level of precipitation necessary to result in fecal coliform densities in excess of the FDA criterion is typically more than one inch and that it is not uniform between various small watersheds across Maryland's portion of the Chesapeake Bay (Chapter 3), and that the community composition of bacteria at a shellfish harvest area

varies seasonally, is related to environmental factors and contains members which co-occur with indicator bacteria and specific *Vibrio* species (Chapter 4).

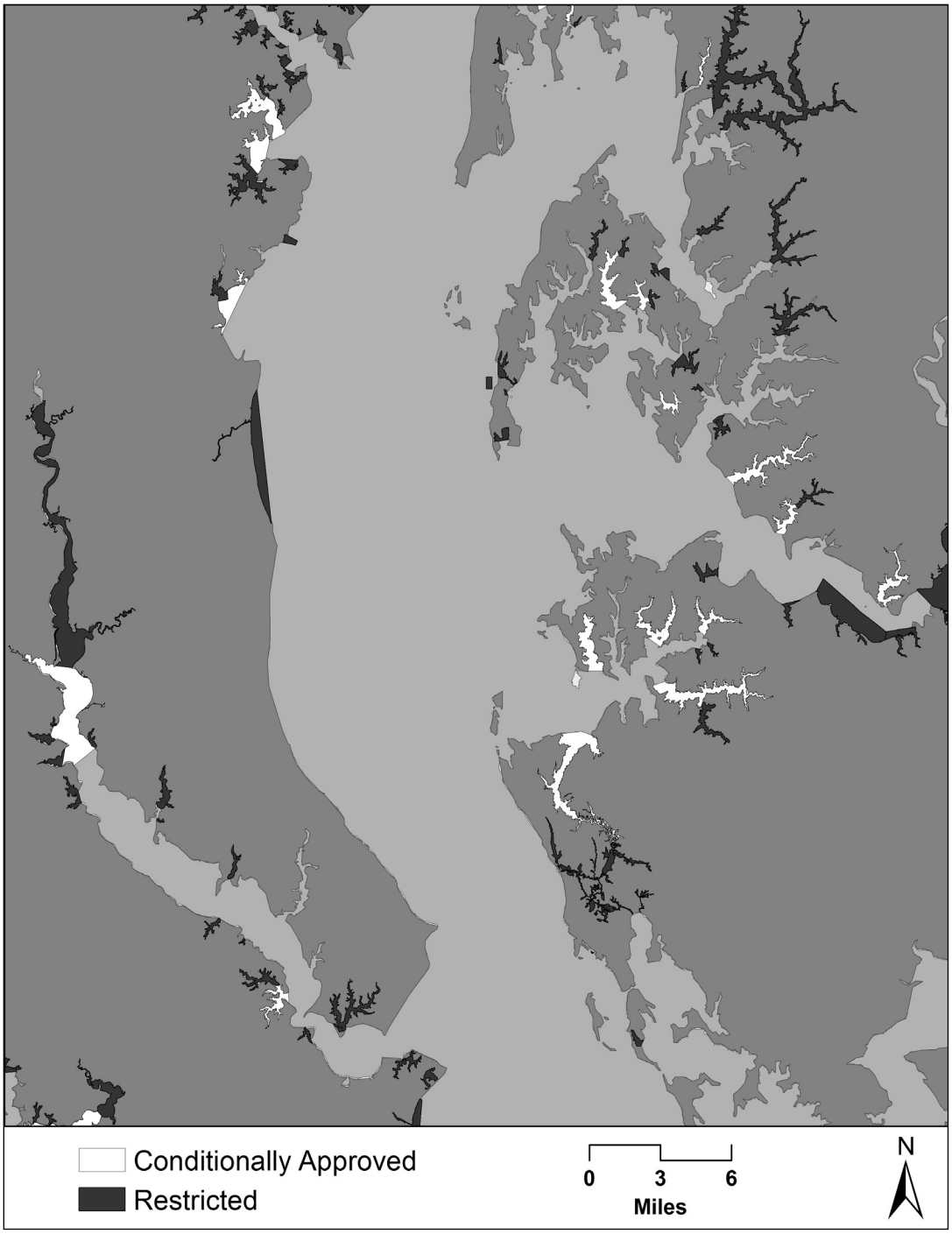

Figure 1.1. Maryland Department of Environment harvest classifications for shellfish harvest areas in the central portion of Marylands Chesapeake Bay, as of January 2016. Source data MDE (2014).

Figure 1.2. Watersheds in Maryland with approved Total Maximum Daily Load (TMDL) plans for reducing levels of fecal coliforms in the receiving waters. Data source MDE (2016).

Chapter 2: Climate Relationships to Fecal Bacterial Densities in Maryland Shellfish Harvest Waters

Abstract

Coastal states of the United States (US) routinely monitor shellfish harvest waters for types of bacteria that indicate the potential presence of fecal pollution. The densities of these indicator bacteria in natural waters may be related to climate in several ways, including through runoff from precipitation and survival related to water temperatures. The relationship between interannual precipitation and air temperature patterns and the densities of fecal indicator bacteria in shellfish harvest waters in Maryland's portion of the Chesapeake Bay was quantified using 34 years of data (1979-2013). Annual and seasonal precipitation totals had a strong positive relationship with average fecal coliform levels ($R^2 = 0.69$) and the proportion of samples with bacterial densities above the FDA regulatory criterion ($R^2 = 0.77$). Fecal coliform levels were also significantly and negatively related to average annual air temperature

($R^2 = -0.43$) and the average air temperature of the warmest month ($R^2 = -0.57$), while average seasonal air temperature was only significantly related to fecal coliform levels in the summer ($R^2 = -0.50$). River and regional fecal coliform levels displayed a wide range of relationships with precipitation and air temperature patterns, with stronger relationships in rural areas and mainstem Bay stations. Fecal coliform levels tended to be higher in years when the bulk of precipitation occurred throughout the summer and/or fall (August to September). Fecal coliform levels often peaked in late fall and winter, with precipitation peaking in summer and early fall. Continental-scale sea level pressure (SLP) analysis revealed an association between atmospheric patterns that influence both extratropical and

tropical storm tracks and very high fecal coliform years, while regional precipitation was found to be significantly correlated with the Atlantic Multidecadal Oscillation and the Pacific North American Pattern. Years with extremely high fecal coliform levels were also years with hurricanes and/or tropical storms impacting the US Atlantic seaboard. These findings indicate that management of shellfish harvest waters should account for changes in climate conditions and that SLP patterns may be particularly important for predicting years with extremely high levels of fecal coliforms.

Introduction

Shellfish harvest in many coastal areas may be restricted due to high levels of fecal bacteria in the surrounding waters or shellfish tissues. Although these fecal indicator bacteria may rarely be pathogenic themselves (Walters et al. 2007, Wilkes et al. 2011), their presence in shellfish waters acts as an indicator of fecal pollution and an increased risk of shellfish being contaminated with human pathogens (EPA 1986, Ashbolt et al. 2001, Wilkes et al. 2011), though the strength of this relationship has been questioned (Wade et al. 2003).

Bivalve shellfish (such as oysters and clams) contaminated with human pathogens represent a significant risk to humans. For example, the consumption of shellfish contaminated with pathogenic *Vibrio* bacteria has caused documented illnesses in Japan, Canada, Australia and the US, and likely many more undocumented cases worldwide (FAO 2011). Estimates for the annual number of *Vibrio parahaemolyticus* cases specifically from oyster consumption range from 66 in Japan to 186 in Canada (FAO 2011). In the US, *Vibrio* bacteria in shellfish accounted for thousands of human illnesses over the period of 1996-2010, including hundreds of hospitalizations and about 70 deaths (Newton et al. 2012). The economic impact from these waterborne and foodborne related illnesses is significant, with total annual costs for the US estimated at roughly $350 million (Ralston et al. 2011). These are also relevant concerns in the Chesapeake Bay where pathogens of human health concern,

such as species of *Vibrio*, *Clostridium*, *Klebsiella*, and *Salmonella* have been isolated from Bay waters and shellfish (Sayler et al. 1975, Morgan et al. 1976). *Vibrio* species found in the Bay are naturally occurring aquatic pathogens and generally are not considered aligned with land-based sources.

To prevent consumption of shellfish contaminated with human pathogens, US coastal states conduct sanitary surveys and monitor surface waters in shellfish harvest areas for bacterial indicators of fecal pollution, in accordance with the National Shellfish Sanitation Program Guide for the Control of Molluscan Shellfish, 2013 revision, the federal/state cooperative program recognized by the US Food and Drug Administration (FDA) (FDA 2015), and the Interstate Shellfish Sanitation Conference. With more than 1,000 historic oyster beds and numerous areas of clam harvest, Maryland officials routinely conduct shoreline surveys for roughly 5,650 kilometers of shoreline, looking for and assessing pollution sources, and monitor over 5,000 square kilometers of waters for indicator bacteria. In Maryland, the Department of Environment (MDE) measures fecal coliform bacteria densities at several hundred shellfish harvesting and sewage outfall areas, at least once per month. If the bacterial densities exceed the criteria, the State will close the shellfish beds for harvest and work to identify ways to mitigate fecal bacterial levels in those waters. Currently, most Maryland tributaries to the Chesapeake Bay have at least some portion of their shellfish grounds restricted to harvest (MDE 2014).

The density and distribution of fecal indicator bacteria and associated pathogens in natural waters, such as shellfish harvest waters, may be linked to climate through regional weather patterns, especially rainfall. The relationship between precipitation, indicator bacteria, and pathogens has been noted in a number of other studies (Mallin et al. 2001). Curriero et al. (2001), for example, found that most waterborne disease outbreaks in the US follow large precipitation events. In addition, fecal coliform concentrations at beaches are

often related to rainfall (Ackerman and Weisberg 2003, Kelsey et al. 2004). In a study of fecal coliform densities at shellfish areas in Galveston Bay, Texas, years with higher amounts of precipitation tend to have higher fecal coliform counts (Jensen and Su 1992). Because of this relationship between rainfall and fecal bacteria, in US coastal states, rainfall greater than a specified level (e.g. one inch) is a standard that may be used to trigger the closure of shellfish beds in waterbodies where the runoff of fecal bacteria appears to coincide with rain events (FDA 2015).

Other important factors that affect the presence and survival of indicator bacteria in coastal waters are water temperature (Howell et al. 1996) and salinity (Goyal et al. 1977), which both may be related to climate variability. For example, a small number of studies found water temperatures were negatively correlated with fecal bacterial levels (Burkhardt et al. 2000, Chigbu et al. 2005) and salinity (Chigbu et al. 2004). In addition, fecal bacteria have been found to survive in benthic sediments (Davies et al. 1995, Anderson et al. 2005) and have a positive relationship with total suspended solids, suggesting that winds can act to resuspend benthic sediments and increase fecal coliform densities in shallow waters (Howell et al. 1996).

These linkages of indicator bacteria levels to regional (i.e. synoptic (Barry 2005)) climate patterns may be described, in part, by larger scale climate phenomena. For example, Lipp et al. (2001b) found that fecal indicator organisms in a Florida estuary became more widespread during winter and fall when precipitation rates increased and noted that this trend occurred during an El Niño event. In the Chesapeake Bay, large-scale climate patterns have been reduced to synoptic climate patterns which have been statistically linked to dynamics in nutrients and phytoplankton (Miller and Harding 2007), zooplankton communities (Kimmel et al. 2006), and fish populations (Wood 2000). In addition to effects from climate variability, the potential for climate change to impact important ecological conditions in the

Chesapeake Bay appears likely (Najjar et al. 2010), including changes to factors that may influence the introduction of indicator bacteria, such as the magnitude and timing of precipitation, streamflow, and storm-related runoff.

The goal of the research described in this chapter is to better define the potential relationship of climate and fecal bacteria in Maryland waters of the Chesapeake Bay. The focus is not on the ability to predict individual observations of surface water fecal coliform concentrations (addressed in Chapter 3) or to define as many drivers of fecal pollution as possible, but to understand the more global impact of climate patterns on densities of these indicator bacteria. The specific objectives of this study are to: 1) assess the relationship of annual and seasonal precipitation and air temperature patterns to average annual and seasonal fecal coliform densities in shellfish harvest waters of Maryland's portion of the Chesapeake Bay, 2) examine climate relationships to indicator bacteria concentrations within individual rivers and bays, 3) identify within-year patterns of precipitation associated with years of high fecal coliform concentrations, and 4) identify dominant climate patterns corresponding to periods of high precipitation and fecal coliform concentrations.

Materials and Methods

Station and Time Period Selection

The Maryland Department of Environment (MDE) has collected surface water samples at shellfish harvest areas since the 1950s. However, monitoring stations for shellfish waters in Maryland have been added and removed over time. Therefore, a subset of monitoring stations that have been continuously sampled were selected for this analysis. A significant change in the number and locations of sampling stations occurred in 1978 and 1979, making that a necessary time period to begin our analysis. Although the Maryland

Department of Environment attempts to sample each station twice a month, year-round, weather and other logistics often result in missed sampling events. For the purposes of this study, we considered any station as continuously monitored if it had been sampled at least four months out of each year from 1979 to the 2013, with at least six samples for each year. Even within this subset, there were a number of stations missing several years of data, though these gaps did not tend to overlap for a large number of stations in any particular year. Therefore, stations were included that had no more than four non-contiguous years in which no samples were collected. These criteria resulted in the selection of 355 continuous sampling stations, which included most of the rivers and bays of Maryland Chesapeake Bay waters (Figure 2.1).

Fecal Indicator Bacteria Data

Surface water estimates of fecal coliform bacteria densities from the selected stations were obtained from an existing MDE database. These data were generated in a consistent manner for the entire period of sample collection, using the three-tube broth dilution protocol as recommended by the American Public Health Association (APHA 1998) and FDA (FDA 2015). Data were excluded from the analysis if they were from targeted sampling around specific weather events (such as rain event sampling conducted in 1987) or samples collected from more than one meter below the surface of the water.

Climate Data

Precipitation and air temperature data were obtained from the National Climatic Data Center (NCDC) (www.ncdc.noaa.gov) for the same time period as the fecal coliform data. Monthly NCDC Climate Division data, aggregated from national weather service stations, were used because they capture weather conditions within geographic regions, called divisions. The six climate divisions that cover Maryland's portion of the Chesapeake Bay

were selected (Figure 2.1). NCDC reports climate division data as monthly total precipitation in inches and monthly average air temperature in degrees Celsius. In order to summarize conditions over the region, precipitation and air temperature values were first averaged between the five geographic divisions by month. For each year and season, the sum of monthly precipitation values, the average of monthly air temperatures, and the maximum of the monthly air temperatures were then calculated.

Sea Level Pressure Patterns and Teleconnections

Monthly sea level pressure (SLP) data were obtained from the Earth System Research Laboratory (ESRL) (www.esrl.noaa.gov) in the form of 'reanalysis' data, which has undergone a quality control process (Kalnay et al. 1996). Values for climate teleconnection indices, such as the North Atlantic Oscillation, were obtained from the National Centers for Environmental Prediction (www.ncep.noaa.gov). In order to identify differences in the SLP between years with very high fecal coliform counts and rainfall amounts and the long-term average for SLP, the three years with both the highest precipitation levels and highest fecal coliform concentrations were each compared to the average condition for all other years (1979-2013). These three years (1979, 1996, and 2003) were selected because the fecal coliform levels were notably higher than other years and therefore of the greatest relevance to natural resource managers. Average annual SLP data for each year of the three years of greatest rainfall were compared to the long-term average for the base period 1979-2013 to generate climate anomalies.

Analysis

Comparison of interannual variability in fecal coliforms and climate factors

In order to examine the relationship of fecal coliforms in surface waters to precipitation and air temperature patterns, three endpoints for evaluation of trends in fecal

coliform densities were selected. Densities were reported in Most Probable Numbers (MPN) which represents an estimate of the number of viable fecal coliforms in a particular volume of water based on replicated culture methods (APHA 1998). The average fecal coliform MPN/100mL of water, within the geographic region and time period of interest, the proportion of samples with densities exceeding the FDA criterion (FDA 2015) of 14 MPN/100mL, and the proportion exceeding the FDA criterion of 49 MPN/100mL were examined. The 14 and 49 MPN/100mL criteria were established to compare against the geometric mean and 90th percentile, respectively, of fecal coliform densities from individual monitoring stations over time (FDA 2015). Both FDA criteria were considered in order to compare the effect of climate variables on both high and low level exceedances. Because average air temperature has increased over this region during the time period examined, annual air temperature and fecal coliform densities were compared with and without their long term trends. Detrending was accomplished by regressing fecal coliforms and air temperature separately against year and then using the residuals associated with each for further comparison.

Seasonal comparison of fecal coliforms and climate variables

Because data from times of the year are considered when regulating shellfish harvest, we examined seasonal trends. Seasons were classified as follows: December through February as winter, March through May as spring, June through August as summer, and September through November as fall.

Comparison of fecal coliforms and climate variables within different rivers

When considering geographic trends, bacterial monitoring stations were grouped within geographic regions as defined by MDE, and regions that contained less than five sampling stations were excluded from our analysis in order to limit the potential influence of

single stations and the inherent spatial variability of environmental data. Each MDE region was compared to the nearest NCDC climate division, except for the stations in the mainstem of the Bay which were compared to aggregated climate data for all six divisions.

Statistical and geospatial methods

All tests for correlation were Pearson tests, as opposed to Spearman correlations, because plots of the raw data suggested that approximately linear relationships existed between fecal coliform levels and the climate variables. Linear regression models using various combinations of total annual precipitation, average annual air temperature, and air temperature of the hottest month against average annual fecal coliform densities, the proportion of samples over the 14 MPN/100mL criterion and the proportion of samples over the 49 MPN/100mL criterion were compared using adjusted R^2 values. Akaike's Information Criteria (AIC) and AIC weight were also calculated to inform model selection (Wagenmakers and Farrell 2004). A visual comparison of residuals versus predicted values, and a visual comparison of predicted and observed values were used to evaluate the linearity and heteroscedasticity of each model. Cluster analysis was conducted using Ward's minimum-variance method, which attempts to assign observations to clusters by minimizing the within-group variance (Ward 1963). All statistical test results were considered significant with a p-value of 0.05 or less.

ArcGIS (ESRI, Redlands, CA) was used to produce layered maps of monitoring stations and NCDC climate divisions. All statistical tests were done using SAS (SAS Institute, Inc., Cary, NC). Analyses of sea level pressure and wind vector data, and the generation of climate anomaly maps were conducted using NCAR (National Center for Atmospheric Research) Command Language (NCL) (www.ncl.ucar.edu).

Results

Annual Comparison of Indicator Bacteria to Climate Factors

Interannnual trends between fecal coliform densities (Figure 2.2, Panel A) and the proportion of samples exceeding the FDA criteria (Figure 2.2, Panel C and D) are well correlated ($p < 0.001$) to patterns of total annual precipitation (Table 2.1). The average annual air temperature and the average air temperature of the warmest month (Figure 2.2, Panel B) were inversely correlated (Table 2.1) to annual average fecal coliform densities, as seen both in the opposing peaks and valleys of the late 1980's and mid-1990's and in the long-term trend of decreasing fecal coliforms and increasing air temperatures. Annual total precipitation explained considerable amounts of variability for both the proportion of samples per year with fecal coliform densities above the 14 MPN/100mL and the proportion with densities over 49 MPN/100mL (Table 2.1). Long-term annual average air temperature and the average temperature of the hottest month followed a significantly positive trend, while fecal coliforms followed a significantly negative trend. However, detrending the data did not affect the strong negative correlation between air temperature and fecal coliform densities. Linear regression analysis showed that the combination of total annual precipitation and the average air temperature of the warmest month provided a significant linear model ($p < 0.001$) of average annual fecal coliform densities (Table 2.2; adjusted R^2=0.60). However, the model comparing the proportion of samples exceeding the 49 MPN/100mL criterion per year to both the total annual precipitation and average air temperature of the warmest month proved to be the strongest model out of those tested (Table 2.2; adjusted R^2=0.72) and was able to reasonably predict observed values (Figure 2.3).

Seasonal and Geographic Comparisons

Fecal coliform densities were highest during the fall (September through November), though there was considerable variability within season for the 34 years of data (Figure 2.4, Panels A, and B). In contrast, total precipitation was highest in spring and summer, slightly lower on average in the fall, and lowest in winter, showing a mismatch in the timing of peak precipitation and fecal coliform levels seasonally (Figure 2.4, Panel C). The comparison of fecal coliform densities to precipitation and air temperature between years and by season showed that for all seasons, total seasonal precipitation was significantly correlated to average seasonal fecal coliform densities and to the proportion of samples exceeding 49 MPN/100mL, with the highest correlation coefficients occurring in the summer and spring. (Table 2.3). Average and maximum monthly air temperature, although always negatively related to fecal coliform densities, was only significantly related in summer (Table 2.3).

Geographically, the relationship of annual precipitation and maximum monthly air temperature varied notably between different rivers (Figures 2.5, 2.6). Total annual precipitation within a river tended to explain more variability in the fecal bacteria values than did air temperature. The proportions of samples exceeding the 49 MPN/100mL criterion within several rivers on the eastern side of the Chesapeake Bay, as well as the two rivers on the western side of the Bay were moderately correlated to annual precipitation levels (correlation coefficients between 0.5 and 0.7) (Figure 2.5). However, fecal coliform exceedances in the remaining rivers and regions of the Bay were either not significant ($p \geq 0.05$) or had very low correlations to precipitation. Only a few rivers showed a low level of correlation between the proportion of samples exceeding the 49MPN/100mL standard and the average air temperature of the hottest month (Figure 2.6).

Characterization of Synoptic Precipitation Patterns

Using cluster analysis to examine the timing and magnitude of monthly precipitation totals, we were able to characterize several types of precipitation patterns (Figure 2.7). Monthly precipitation profiles for many of these clusters (Figure 2.8) showed distinct profiles that differed from 'normal' (average monthly values from 1979 to 2013) precipitation patterns. For example, Cluster A (Figure 2.8) tended to have above average precipitation in the late summer and fall, while Cluster C featured above normal precipitation throughout much of the year, particularly between May and September. Cluster B contained only one year, 2006, which had a very distinct below average pattern of precipitation for much of the year and large peaks in June and September. A large number of years were clustered together that generally contained low precipitation values throughout the year. Comparison of the annual proportion of samples exceeding the 49MPN/100mL criterion by cluster (Figure 2.9) shows that Clusters A and C contained years with very high exceedances relative to all other years. Despite having two months with precipitation totals well above average, fecal coliform levels (based on the 49MPN/100mL exceedance criterion) for 2006 were relatively low.

The precipitation-temperature regression model derived in this work accounted for a majority of the observed variance (72%) among annual fecal coliform exceedances of the 49mpn/100mL criteria from 1979-2013, and highlighted three years with extremely high precipitation and fecal coliform levels (1979, 1996, and 2003) relative to the other years (Figure 2.3). In order to better understand climate and weather conditions that may be common to years with high fecal coliform exceedances, a focused examination of the precipitation patterns and regional atmospheric circulation for these three years was conducted. The monthly precipitation and fecal coliform levels for these three years showed a delay in timing between some peaks of precipitation and fecal coliform exceedances (Figure

2.10). Also, there were both similarities and differences between precipitation patterns when comparing 1979, 1996, and 2003. For all three of these very high fecal coliform years, precipitation was generally high from May-October. However, in 1979 precipitation peaked in late summer through early fall (August – September), while 1996 featured very high spring precipitation (May-June) with an exceptionally high early summer peak (July). In 2003, precipitation was very high throughout most of the growing season (May-September).

Comparison of Climate Variables to Climate Indices and Synoptic Climatology

Precipitation patterns for the climate divisions included were significantly, though weakly correlated to annual average index values for the Atlantic Multidecadal Oscillation and the Pacific-North American Oscillation (Table 2.4). No discernible relationship emerged between precipitation patterns in these climate divisions and either the El Niño Southern Oscillation (ENSO) or North Atlantic Oscillation.

Comparison of sea-level pressure patterns for the years with extremely high precipitation and fecal coliform densities (1979,1996, 2003) to the average SLP conditions for the period 1979 to 2013, revealed climate patterns consistent with the precipitation patterns observed in these three years (Figure 2.11). Precipitation in May through June was very high in 2003, high in 1996, and moderately high in 1979. This pattern was inversely related to SLP over the mid-Atlantic, which was anomalously low in 2003, higher in 1996 as a result of intersecting high pressure areas centered to the Northwest and Southeast of the mid-Atlantic, and very high over the North American mid-latitudes in 1979. Anomalous SLP patterns for the three years with very high precipitation and fecal coliform levels were also consistent with observed precipitation patterns from July through October, when cold fronts from the west or northwest and tropical storms generally account for very high monthly precipitation levels. In 1979, when July-October precipitation was moderate to very high in all months, SLP was very high across the Atlantic basin, indicative of a stronger than normal Azores-Bermuda

high (ABH), while low pressure existed over the western Caribbean, much of the US and the Gulf of Mexico, and also over most of the area from eastern Canada to Iceland. A distinctive high SLP anomaly over the western Atlantic Ocean was present in 2003 coinciding with very high precipitation amounts. In 1996, only a weak and small high SLP anomaly existed off of the eastern US, corresponding to slightly less precipitation for August through October of that year, compared to 1979 and 2003.

Discussion

Precipitation and air temperature patterns explain a large percentage of the inter-annual variability in bacterial levels at shellfish harvest areas in Maryland waters. A small number of previous studies have demonstrated similar relationships, for other geographic areas and shorter time periods, between annual estimates of climate variables and levels of fecal indicator bacteria in estuarine and river waters (Lipp et al. 2001a, Chigbu et al. 2005, Huang 2010), though a positive relationship between fecal coliforms and air temperature was noted for Virginia waters (Huang 2010).

In Maryland shellfish waters, total annual precipitation was more strongly correlated to the proportions of samples over the 49 MPN/100mL than to average fecal coliform density. This finding may be related to the skewed distribution of both precipitation and fecal coliform data such that the 90^{th} percentile values are more representative of precipitation-related input events than the average MPN/100mL. This finding also supports the management approach of classifying shellfish beds as 'conditionally approved' based on the magnitude of precipitation events (FDA 2013). In Virginia waters, the density of fecal coliforms was strongly influenced by the amount of precipitation within a given rain event and the time period between the rain event and the sample collection (Huang 2010). Future work in Maryland should attempt to better define the relationship between climate patterns and the frequency and magnitude of individual precipitation events. Chapter 3 presents

findings from the evaluation of the relationship between individual precipitation events and the likelihood of fecal coliforms to exceed the FDA criteria.

The inverse relationship between air temperature and fecal coliform levels may be related to summer heat waves along the eastern US resulting in relatively low precipitation and fecal coliform levels. This interpretation is supported by the weak negative correlation between summer air temperatures and total summer precipitation amounts, and the more continental influence of air masses in summers with lower precipitation amounts. This finding also agrees with the relationship that other studies have noted for decreased indicator bacteria levels with increasing water temperatures (Burkhardt et al. 2000, Chigbu et al. 2005). Rhodes and Kator (1988), for example, showed that increasing water temperatures leads to increased predation on *Escherichia coli* (a type of fecal coliform bacteria) and may be correlated with increased stress from physical factors such as solar radiation. In contrast, the weak, positive relationship between fecal coliforms and air temperature found for Virginia rivers (Huang 2010) was hypothesized to result from reduced physical stress and lower mortality rates of fecal coliforms in warm waters.

In this study, seasonal differences were also detected in the relationships between fecal bacteria and climate variables, which may be important to managers since data from all seasons are typically treated the same when determining whether or not to close shellfish beds to harvest (FDA 2013). The relationship between climate and fecal coliforms was stronger from spring through fall than in winter, with the strongest relationship occurring in summer when both precipitation and air temperature patterns were significantly related to average annual fecal coliform densities and the proportion of samples above the 49 MPN/100mL criteria (FDA 2013). In contrast, Lipp et al. (2001b) found that, for an estuary in Florida, fecal coliform levels increased in fall and winter and hypothesized that the seasonality related to ENSO driven wintertime precipitation patterns.

The classification of within-year precipitation patterns, in our study, further defined the relationship between precipitation and fecal coliform presence in Maryland waters by identifying two distinct types of years that tend to lead to high fecal coliform levels, years with high precipitation in the late summer and fall and those years with above normal precipitation throughout much of May through September. The separation in time between periods of extremely high precipitation and elevated fecal coliform levels helps describe the linkages between climate, weather, and fecal coliform levels. The delayed response in fecal coliform levels that was observed in late fall of some years, suggests that periods of high precipitation for those years may create conditions that indirectly lead to elevated fecal coliform levels. Although fecal indicator bacteria levels in surface waters are often linked to runoff from individual precipitation events, elevated bacterial concentrations in months after the bulk of precipitation suggests that either conditions in the summer and early fall associated with highest rainfall levels are not favorable to the survival of fecal coliforms in surface waters or that periods of above normal precipitation set up conditions that support the subsequent introduction of fecal coliforms. It may be that solar radiation and stratification of waters are lower in fall and early winter than earlier in the year resulting in conditions more supportive of fecal coliform survival and distribution throughout the water column following rain events. Another related possibility is that intense rain events in the summer and early fall result in stormwater overflows and intense runoff of water from land, which loads benthic sediments with fecal coliforms. Subsequent wind events could pull these benthic sediments and associated fecal bacteria up into the water column.

Precipitation and air temperature patterns within a geographical region may be described and predicted by large scale climate patterns. Climate indices have been developed to assess the state of these climate patterns and may be used to explain annual and seasonal variability in fecal bacteria levels. For example, Lipp et al. (2001a)found significant seasonal

differences for levels of fecal indicators and pathogenic bacteria in waters and sediments of Charlotte Harbor, Southern Florida, and hypothesized that these differences were being driven by weather conditions resulting from changes in the ENSO. Lipp et al. (2001b) also found significant differences in fecal bacteria for tributaries of Tampa Bay, Florida, which correlated to patterns of the ENSO. Chigbu et al. (2004) found seasonal and annual patterns in fecal indicator densities in Mississippi waters over 11 years of field sampling that corresponded to changes in the ENSO.

In contrast to ENSO related effects along the Gulf Coast, regional climate conditions in the Chesapeake Bay have been linked to the North Atlantic Oscillation (NAO) (Cronin et al. 2003) and American Multidecadal Oscillation (AMO) (Nye et al. 2014), with winter climate conditions being related to biological endpoints, such as the abundances and predation patterns for jellyfish (Purcell and Decker 2005). In this study, the weak but significant correlation of fecal coliform levels above the 49 MPN/100mL to the AMO was surprising given that the multidecadal cycle of the AMO contrasted with the relatively short time frame and large interannual variability of the fecal coliform levels. As time goes on and more years of data are added, a more robust test of this relationship can occur. Some summer climate patterns for the Chesapeake Bay region have also been described. For example, Zhu and Liang (2013) defined an index of the Bermuda High for summer, where high SLP over the Gulf of Mexico and low SLP over the southern Great Plains of the US correlated with precipitation and air quality patterns over the central and eastern US. However, the interannual summertime anomalies of their Bermuda High Index (BHI) (Zhu and Liang 2013) do not appear to correspond with the patterns of total annual precipitation that we present in this study.

Analysis of SLP for the three years with very high precipitation and fecal coliform levels highlighted large scale climate patterns capable of affecting precipitation over the

Chesapeake Bay by influencing the generation of storms (cyclogenesis), storm tracks, storm durations, and storm intensity. That analysis suggested that high precipitation and high fecal coliforms result from storm generation and tracks that stretch from the western Gulf of Mexico, across the eastern US to the Canadian eastern coastline. That same analysis showed patterns of high pressure off the east coast of the US, sometimes characteristic of a stronger than normal Azores-Bermuda High, which can lead to tropical storm tracks that traverse the eastern US, bringing very high precipitation (Herbert 1980, Elsner 2003). In fact, the SLP patterns during the three years of extreme fecal coliform abundances corresponded to years with notable tropical storm seasons for Maryland. For example, tropical storms David and Frederic tracked across the Chesapeake Watershed in September of 1979. Additionally, the presence of high pressure over the western Atlantic in the summer and fall of 1996 occurred when Tropical Storms Bertha, Fran, and Josephine impacted the eastern US with heavy rainfall and storm surges. Similarly, the fall of 2003, when anomalously high pressure occurred over the northwestern Atlantic, was marked by Tropical Storms Bill, Grace, and Henri; with Hurricane Isabel causing historic flooding along the eastern US in September. Additional research into SLP anomalies for years with more moderate amounts of precipitation and fecal coliform levels and for the small set of years where the relationship of fecal coliforms, precipitation, and air temperatures was weaker (Figure 2.3) would help describe potential links between large-scale climate variability and the synoptic patterns of climate over the upper Chesapeake Bay.

The effect of climate on fecal coliform densities is also especially important for cases where a particular climate pattern occurs for consecutive years. This can be seen most distinctly in the late 1980's (Figures 2.2) when a drought over much of the United States resulted in sustained decreases in average annual fecal coliform densities and the proportion of samples exceeding both FDA criteria. Without taking climate into account, managers

might come to the conclusion that decreases in bacterial densities during that period was linked exclusively to other factors, such as best management actions implemented as a result of the Clean Water Act.

The goal of sustaining environmental and public health involves not only the assessment of current risks, but an understanding of how risks change over time and why. Another finding from this study is that the evaluation of long-term trends in fecal coliform densities, both retrospectively and in projections of future conditions, should both account for climate variability and climate change. In retrospective analysis, the influence of climate variables needs to be considered when examining long-term trends. For projections of future conditions, the linear relationship that we describe between fecal indicator patterns and climate variables may provide a predictive tool for forecasting broad-scale fecal coliform conditions in future years. In addition, the indirect link of fecal coliform levels to years with notable tropical storm activities calls for a robust analysis of future storm projections and the concurrence of fecal bacteria in shellfish harvest waters.

Our study suggests that a combination of climate variability and climate change forecasts with estimates of the effectiveness of practices and structures designed to reducing bacterial pollution could be used to help predict the success of management actions. For example, the State of Maryland continues to support and strengthen its activities in stormwater management (Md Code §4-201 (LexisNexis 2015)), with the goal of reducing the amount of various pollutants entering the Chesapeake Bay. Improvements in the management of stormwater, through the implementation of best management practices that are designed to reduce the direct runoff of rain water and stormwater overflows, should lead to lower inputs of fecal bacteria. Because these practices target precipitation-related runoff, the link between precipitation and fecal introduction may be lessened over time, as other drivers become more important, or may shift in nature, as low volume rain events become

less important and large scale rain events that overwhelm stormwater controls become more influential in indicator bacteria densities.

Our findings also indicate that the relationship between rainfall and fecal coliforms may be influenced by other environmental factors within each river. By focusing our analysis on annual averages and the proportions of samples exceeding FDA criteria, our goal was to describe the relationship between fecal bacteria in Maryland waters to climate on a large geographic and temporal scale. However, this approach limits the interpretation of our results to annual, seasonal, and river scale dynamics, and does not attempt to predict bacteria levels within individual samples or sample stations. Although the approach in this study of looking at annual data is in keeping with the practice by managers of aggregating data (typically 15 or more samples from each station are considered when making decisions about harvest closures; (FDA 2013), building models that can predict fecal coliform levels in individual samples might help explain drivers of fecal pollution at the local scale. Chapter 3 describes research focused at this more local geographic scale and in response to individual precipitation events.

In both Maryland (this study) and Virginia (Huang 2010) waters, the strength of the relationship between precipitation and fecal coliform levels weakened as the analysis focused on smaller time and spatial scales. This phenomenon likely results from the influence of factors other than rain and temperature when considering bacterial concentrations in surface waters at finer temporal and spatial scales, such as land-use and local hydrodynamics (Huang 2010). For example, several studies have found an increased level of fecal bacteria in areas of increased human development (Mallin et al. 2001, Interlandi and Crockett 2003) and impervious surface (Glasoe and Christy 2004, DiDonato et al. 2009), presumably from increased overland runoff of fecal matter related to precipitation events. In this study, the aggregate influence of precipitation on fecal coliform levels in various rivers appears to be

weaker for watersheds with high development, such as Baltimore, Washington D.C. and other highly developed areas, than for the agriculturally and more sparsely developed lands, primarily on the eastern side of the Chesapeake Bay. Further research into the sources of fecal coliforms found in Maryland shellfish waters is needed to better describe the relationship between source type and runoff potential. One approach would be to conduct a regression analysis using as many locally-relevant factors (e.g. land use status and change, hydrodynamics, fecal sources, etc.) to help describe the suite of factors that influence fecal coliform densities on smaller time and spatial scales.

Several conclusions can be drawn from the research presented in this Chapter. Firstly, climate variables, particularly annual rainfall, play a substantial role in large-scale fecal coliform trends for Maryland shellfish waters, and should be considered when evaluating conditions for management of shellfish beds. Secondly, the strength of these relationships is not uniform seasonally or geographically, suggesting that management criteria may be altered by season and location. The timing of rainfall within a year can influence fecal coliform levels for that year, with patterns of sustained spring through fall and intense early fall rains leading to the highest coliform levels. In several years, there was a delay in timing between peak rainfall and peak fecal coliform levels indicating that either the high rainfall sets up conditions for later growth of fecal coliforms or is predictive of some later climatological conditions. Large scale sea-level pressure patterns are predictive of years with very high fecal coliform levels and correspond to climate conditions that direct storms over the region or that produce storms.

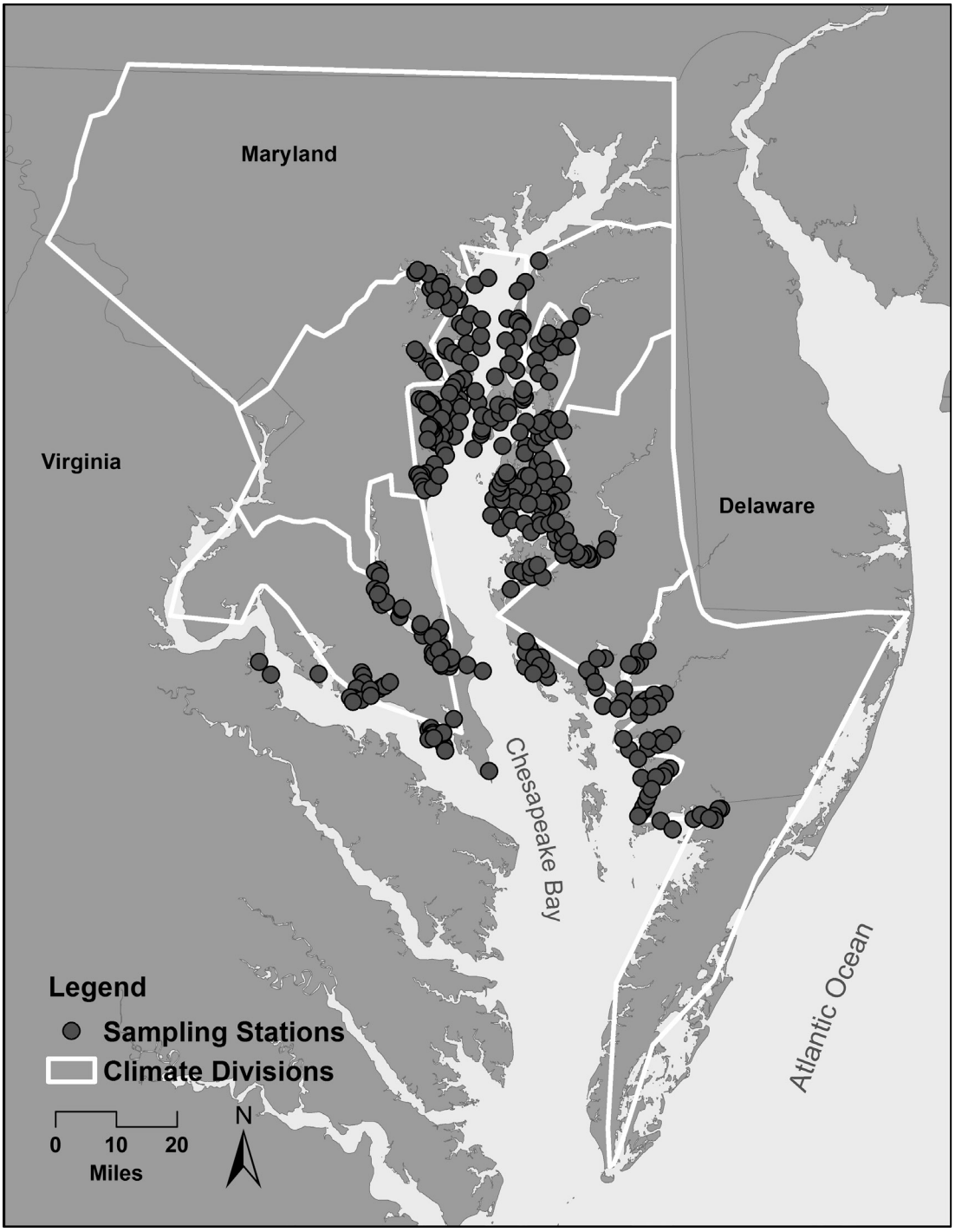

Figure 2.1. Sampling stations and climate divisions included in analysis.

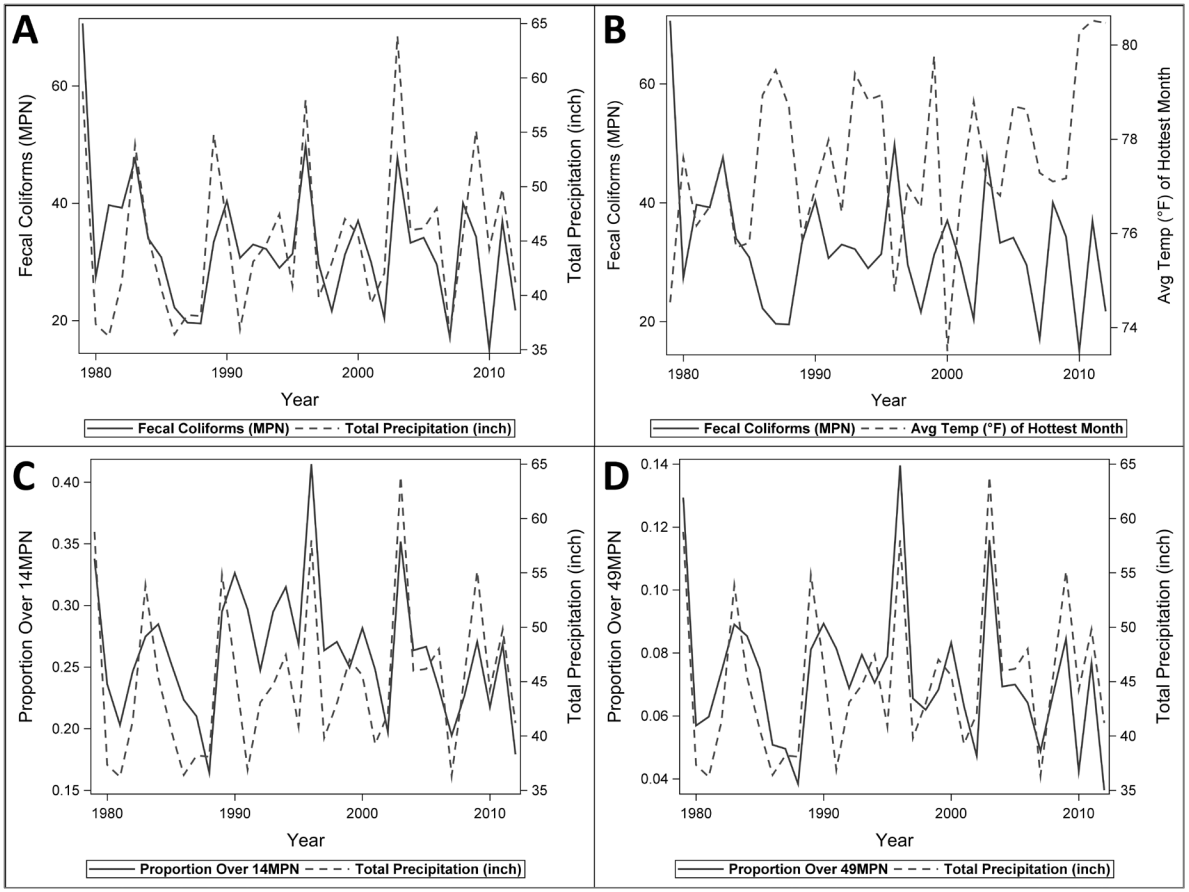

Figure 2.2. Plot of annual average fecal coliform densities (solid line) and total annual precipitation (dashed line) by year (Panel A), annual average fecal coliform levels and the average air temperature (dashed line) of the hottest month by year (Panel B), the proportion of samples with fecal coliform densities exceeding the 14MPN/100mL criterion and total annual precipitation by year(Panel C), the proportion of samples with fecal coliform densities exceeding the 49MPN/100mL criterion and total annual precipitation by year (Panel D).

Table 2.1. Correlation coefficients between fecal coliform densities (MPN) and climate variables. Numbers in parentheses are p-values. Numbers in bold are considered significant ($p<0.05$).

		Climate		
		Total Annual Precipitation	Average Monthly Air Temperature	Air Temperature of Hottest Month
Fecal Bacteria	Annual Average	**0.690 (<.001)**	**-0.469 (0.005)**	**-0.554 (<.001)**
	Proportion Above 14MPN	**0.711 (<.001)**	-0.302 (0.083)	**-0.459 (<.001)**
	Proportion Above 49MPN	**0.771 (<.001)**	**-0.430 (0.010)**	**-0.572 (<.001)**

Table 2.2. Fit statistics for regression models comparing the proportion of samples over 14MPN/100mL (P_{14}), the proportion over 49MPN/100mL (P_{49}) or the average annual fecal coliform density (X_{fc}) to total annual precipitation (Σprecip), average annual air temperature (X_{temp}), and average temperature of the hottest month (H_{temp}). AIC = Akaike's Information Criteria.

Model		Adjusted R^2	R^2	AIC	AIC weight	Delta AIC
P_{49} =	Σprecip + H_{temp}	0.72	0.74	-297.5	0.99	0.0
	Σprecip + X_{temp}	0.64	0.66	-288.9	0.01	8.6
	Σprecip	0.60	0.60	-284.7	0.00	4.2
	H_{temp}	0.33	0.33	-267.4	0.00	17.3
	X_{temp}	0.19	0.19	-260.9	0.00	6.5
P_{14} =	Σprecip + H_{temp}	0.56	0.58	-226.5	0.79	0.00
	Σprecip	0.49	0.53	-222.7	0.12	3.79
	Σprecip + X_{temp}	0.49	0.51	-222.1	0.09	0.56
	H_{temp}	0.19	0.21	-206.8	0.00	15.30
	X_{temp}	0.06	0.09	-202.1	0.00	4.79
X_{fc} =	Σprecip + H_{temp}	0.60	0.62	134.2	0.86	0.0
	Σprecip + X_{temp}	0.55	0.58	138.0	0.13	3.8
	Σprecip	0.48	0.48	143.4	0.01	5.3
	H_{temp}	0.31	0.31	152.8	0.00	9.5
	X_{temp}	0.22	0.22	156.9	0.00	4.1

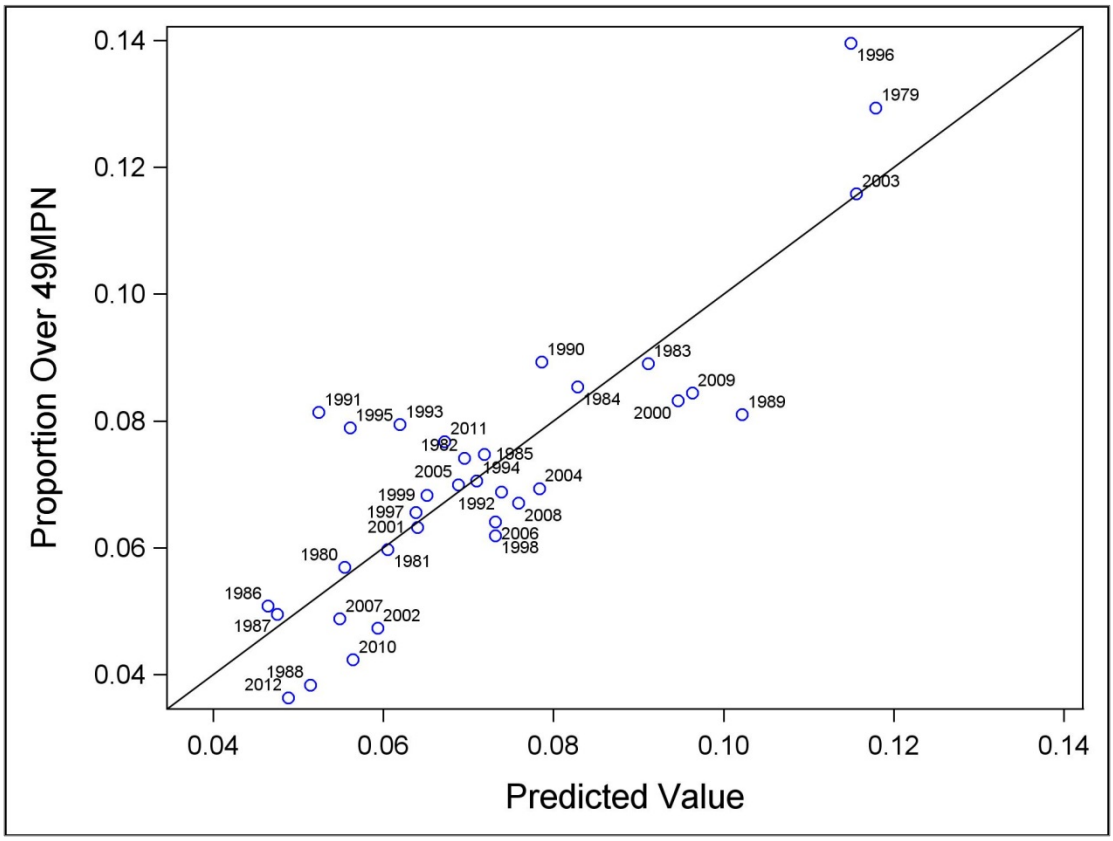

Figure 2.3. Plot of the proportion of samples with fecal coliform densities exceeding the 49MPN/100mL criterion (y axis) versus the predicted values (x axis). Predicted values are based on the regression model: predicted = 0.0022 * (total annual precip) - 0.0052 * (air temp of hottest month) + 0.38 (adjusted R^2=0.729, p<0.001).

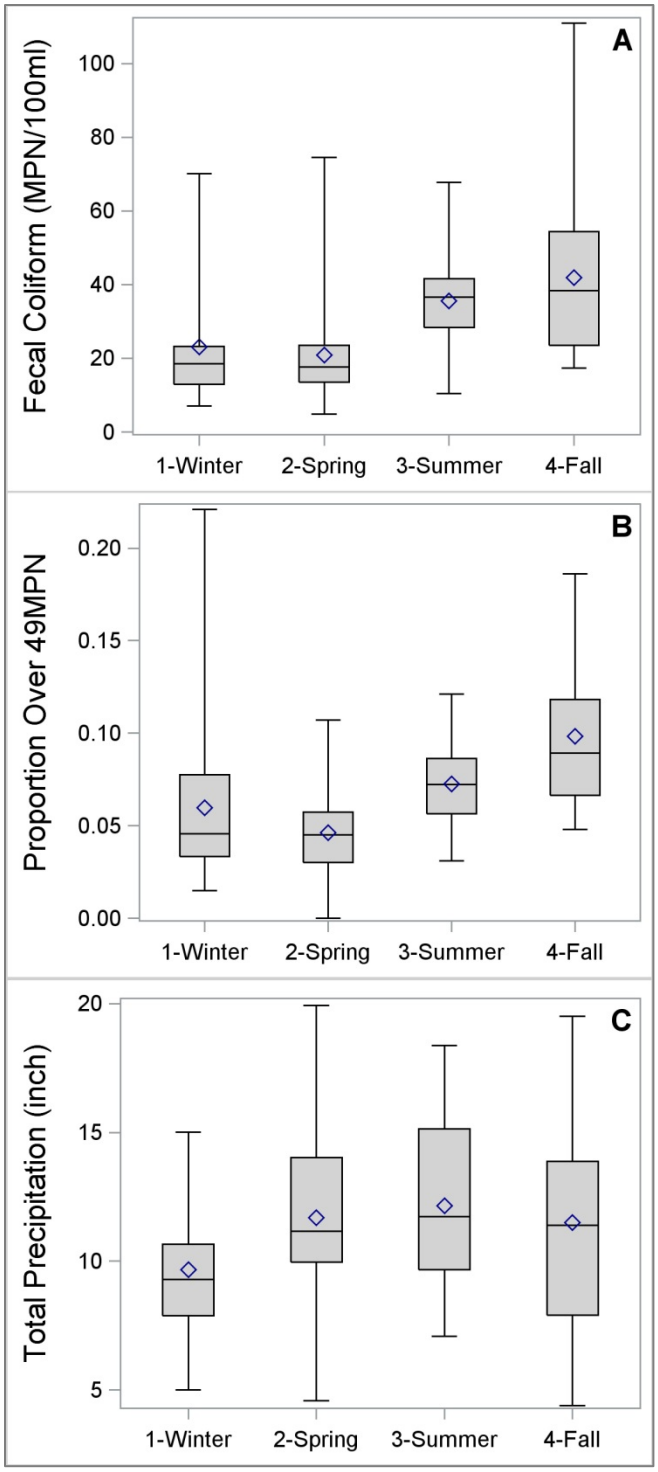

Figure 2.4. Boxplots of average fecal coliform densities (panel A), the proportions of samples with densities over 49MPN/100mL (panel B), and total precipitation amounts (panel C) by season. Boxes represent the range between the 25th and 75th percentiles of values. Whiskers extend to extreme values. Horizontal lines bisecting boxes indicate the median value and diamonds represent the mean value.

Table 2.3. Correlation coefficients for the proportion of measurements over 49MPN/100mL compared to total precipitation and the average air temperature of the hottest month by season. Bold text indicates a significant relationship ($p<0.05$).

	Total Precipitation	**Maximum Temperature**
Winter	**0.547** **(0.001)**	-0.034 (0.850)
Spring	**0.687** **(<.001)**	-0.138 (0.438)
Summer	**0.607** **(<.001)**	**-0.500** **(0.003)**
Fall	**0.520** **(0.002)**	-0.181 (0.307)

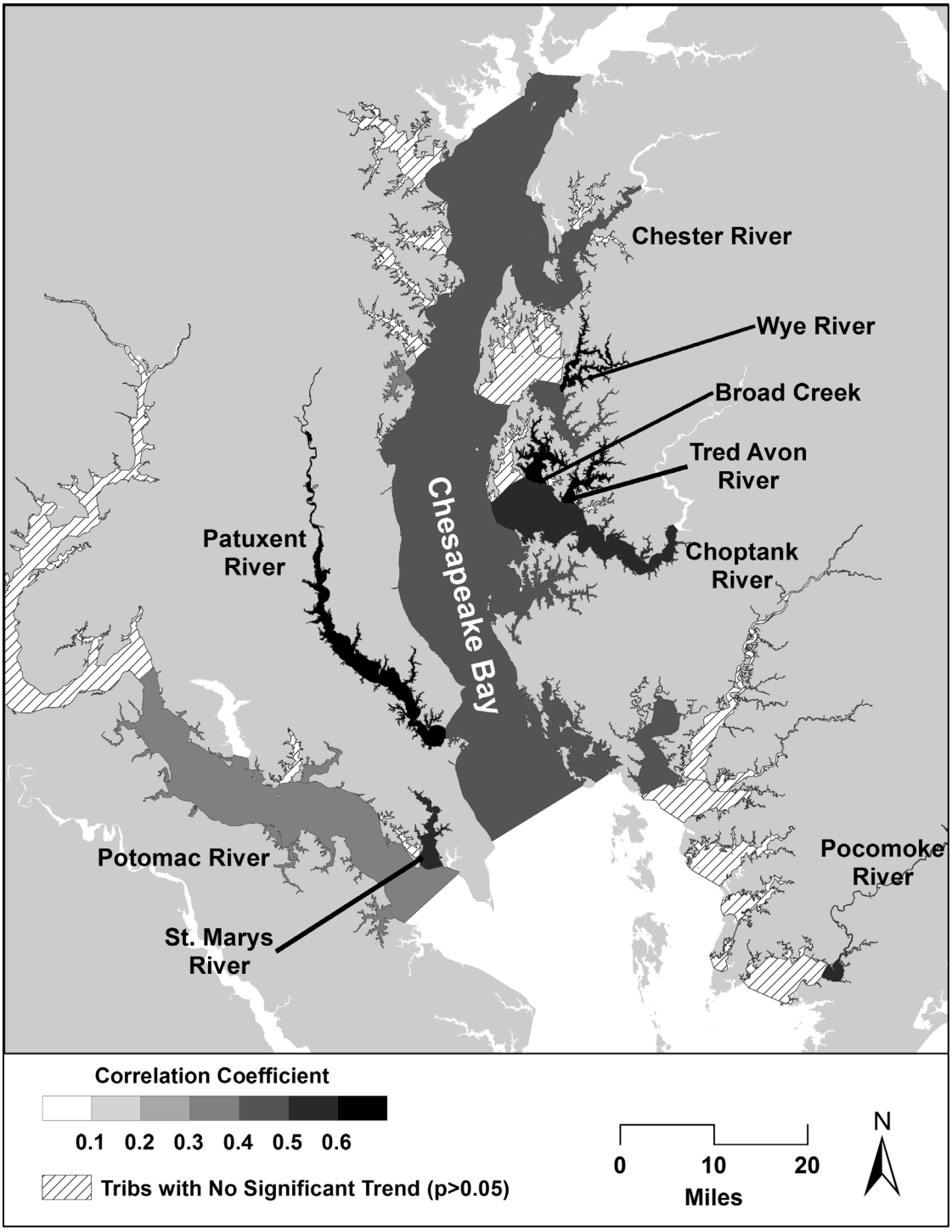

Figure 2.5. MDE monitoring regions, shaded by correlation coefficient for the proportion of samples with fecal coliform densities over the 49MPN/100mL criteria compared to total annual precipitation.

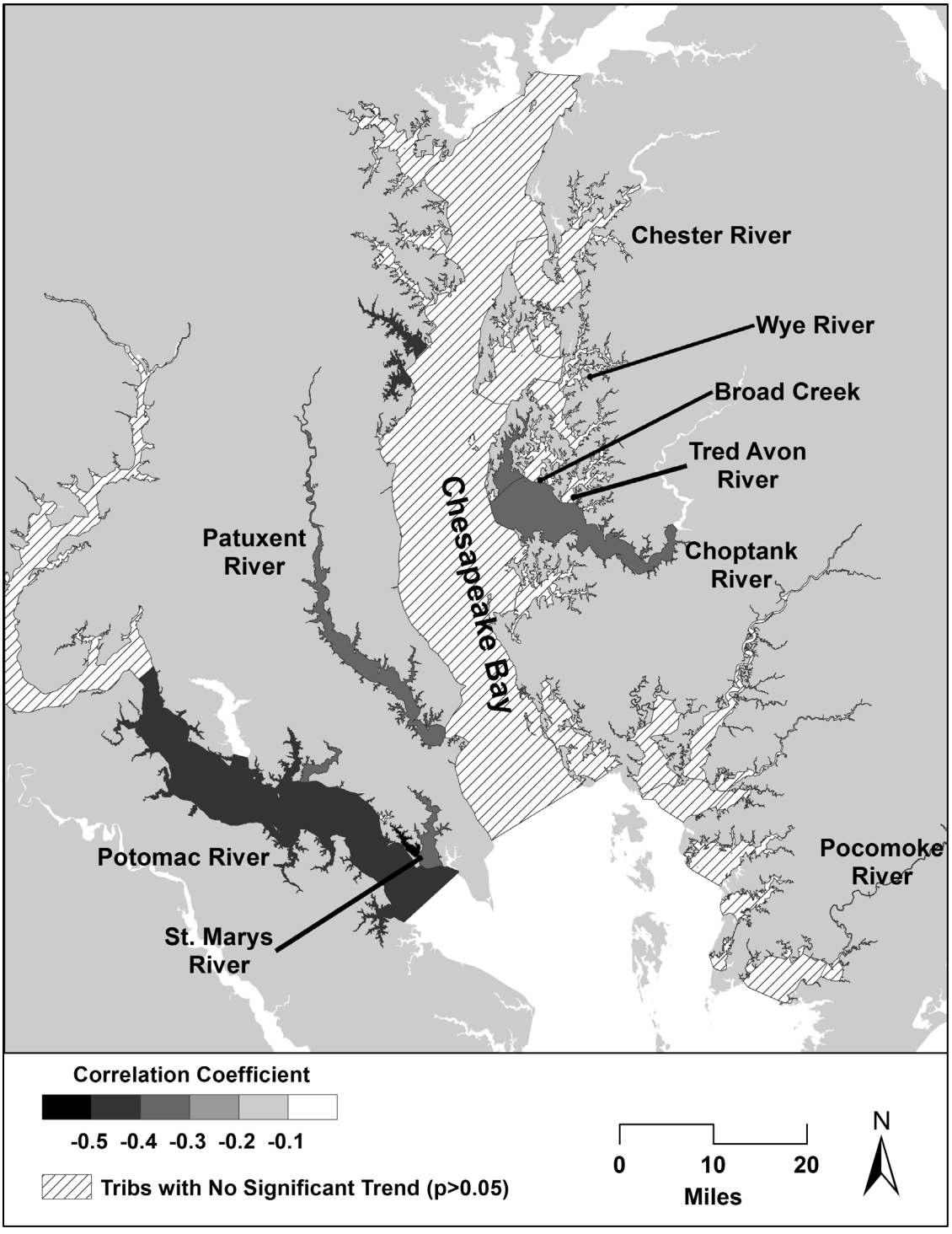

Figure 2.6. MDE monitoring regions, shaded by correlation coefficient for the proportion of samples with fecal coliform densities over the 49MPN/100mL criteria compared to temperature of the hottest month precipitation.

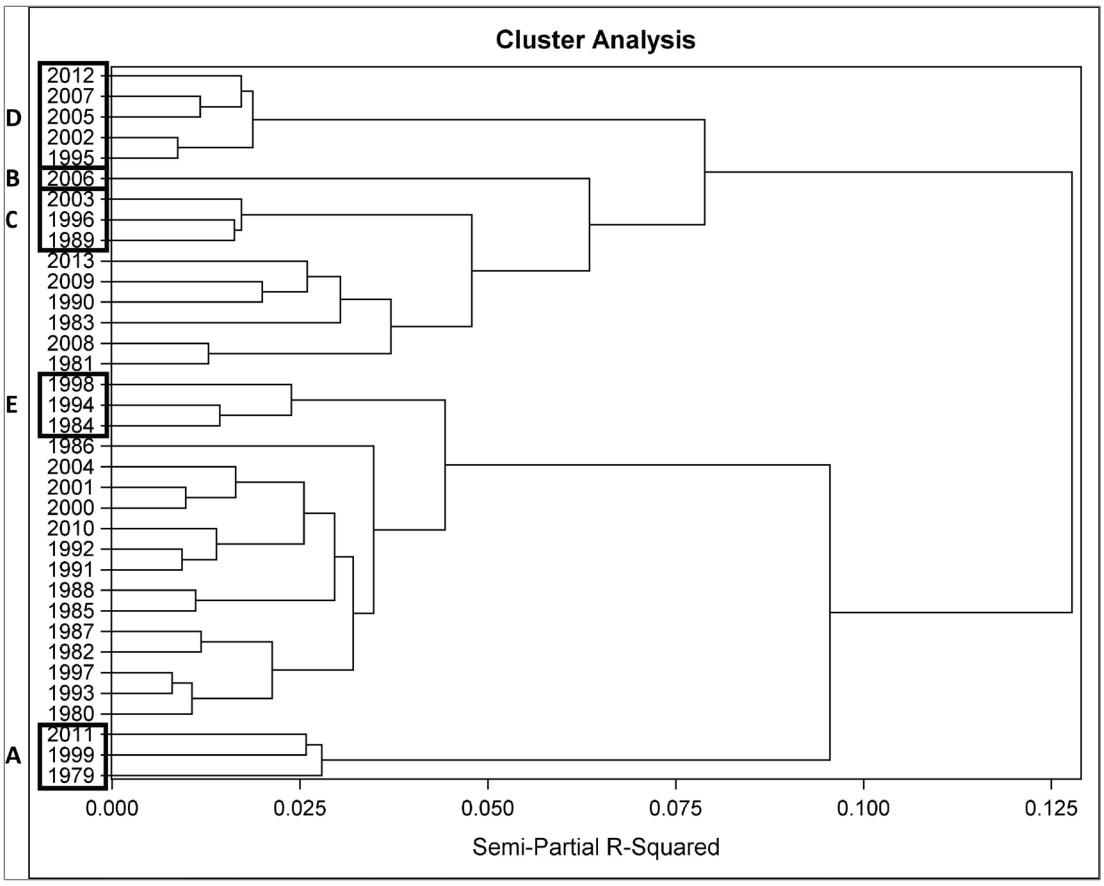

Figure 2.7. Dendrogram from cluster analysis showing how years group based on their pattern and magnitude of monthly precipitation levels. Clusters A-E were selected in order to define synoptic climatologies and for comparison of fecal coliform levels between clusters.

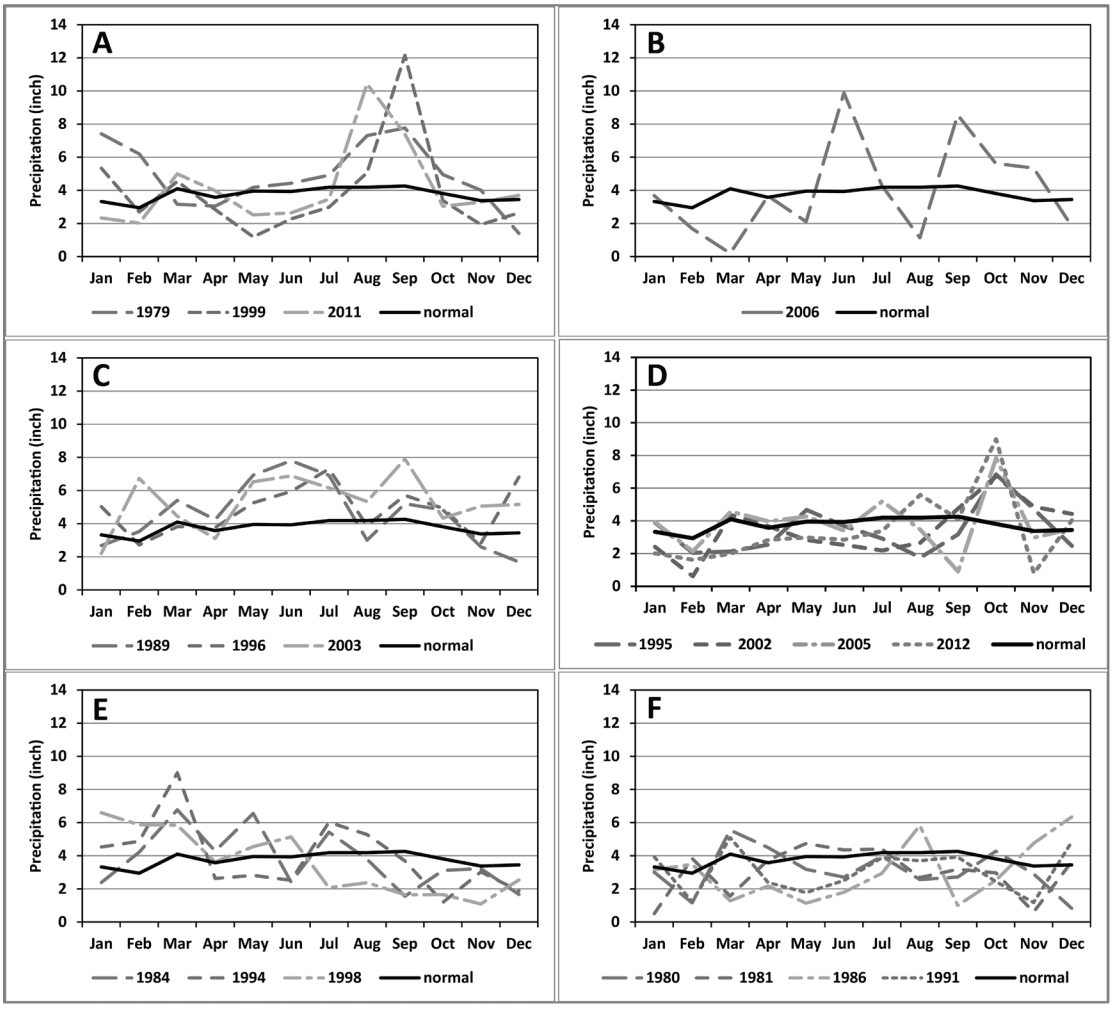

Figure 2.8. Monthly total precipitation values compared to the normal (the long term average for that month over the base period of 1979-2012). Panels A-E represent years from individual clusters corresponding to Figure 2.7. Cluster F includes a subset of years from those years not included in clusters A-E.

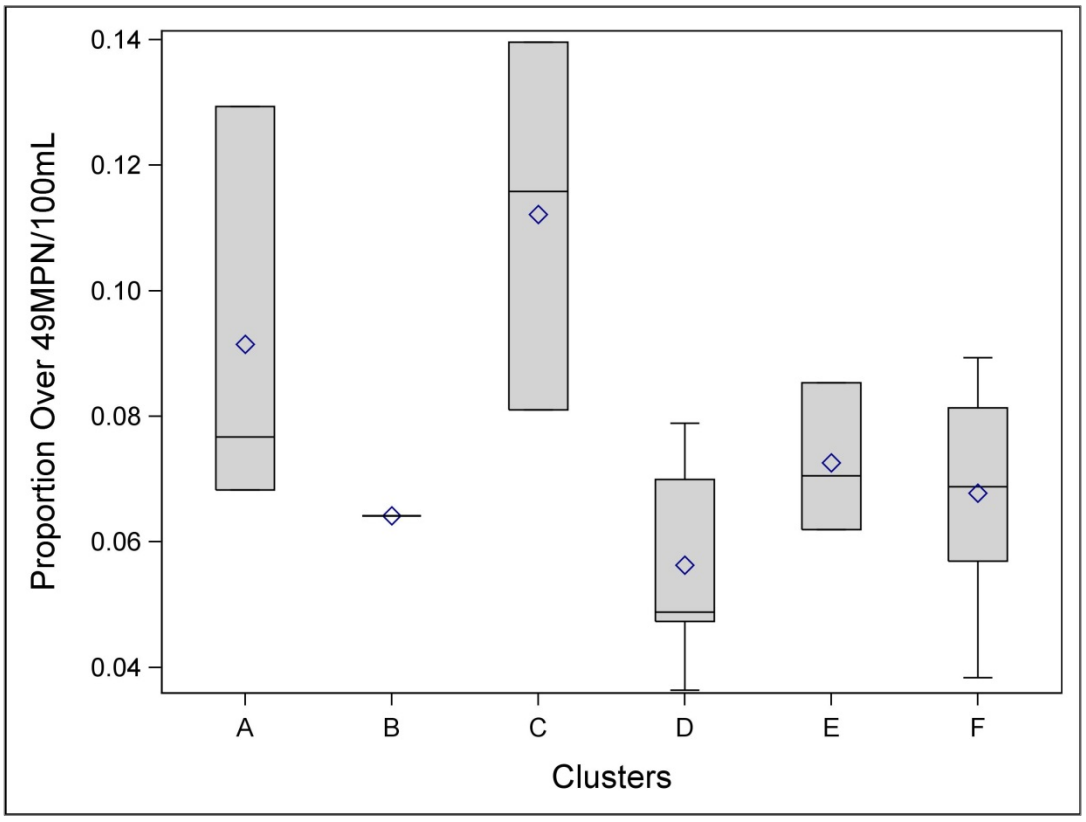

Figure 2.9. Comparison of the annual proportion of samples with fecal coliforms exceeding the 49MPN/100mL criterion (y axis) by clusters of years (x axis).

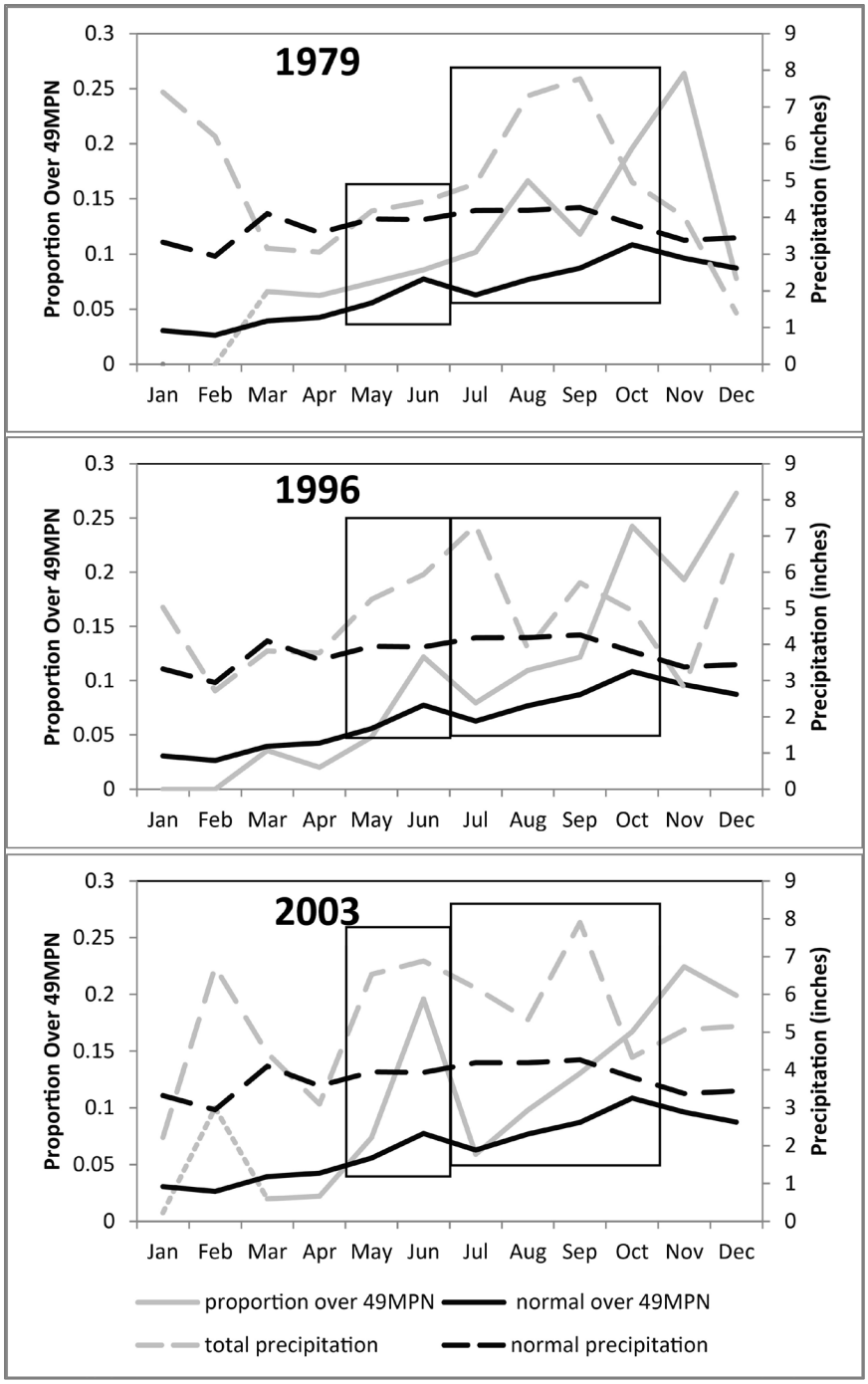

Figure 2.10. Monthly values for the proportion of samples above the 49MPN criteria (solid gray line) compared to the normal conditions (average monthly value for base period of 1979-2013; solid black line), and total precipitation levels (dashed gray line) compared to normal conditions (dashed black line), for 1979, 1996, and 2003. Dotted lines in January and February of 1979 and 2003 indicate low numbers of samples (< 100) for those months due to icy conditions. Black boxes indicate the time periods targeted for sea-level pressure analysis.

Table 2.4. Correlation coefficients, and p-values in parentheses, for the comparison of the proportion of water samples with fecal coliforms exceeding the criteria to existing climate teleconnection indices.

	North Atlantic Oscillation	**Atlantic Multidecadal Oscillation**	**Pacific North American Oscillation**	**Multivariate ENSO Index**
Total Precipitation	0.187 (0.283)	0.232 (0.210)	**-0.424** **(0.011)**	-0.137 (0.434)
Average Air Temperature	-0.116 (0.507)	**0.525** **(0.002)**	-0.254 (0.142)	-0.117 (0.505)
Maximum Air Temperature	-0.060 (0.731)	0.189 (0.309)	0.129 (0.459)	0.013 (0.940)

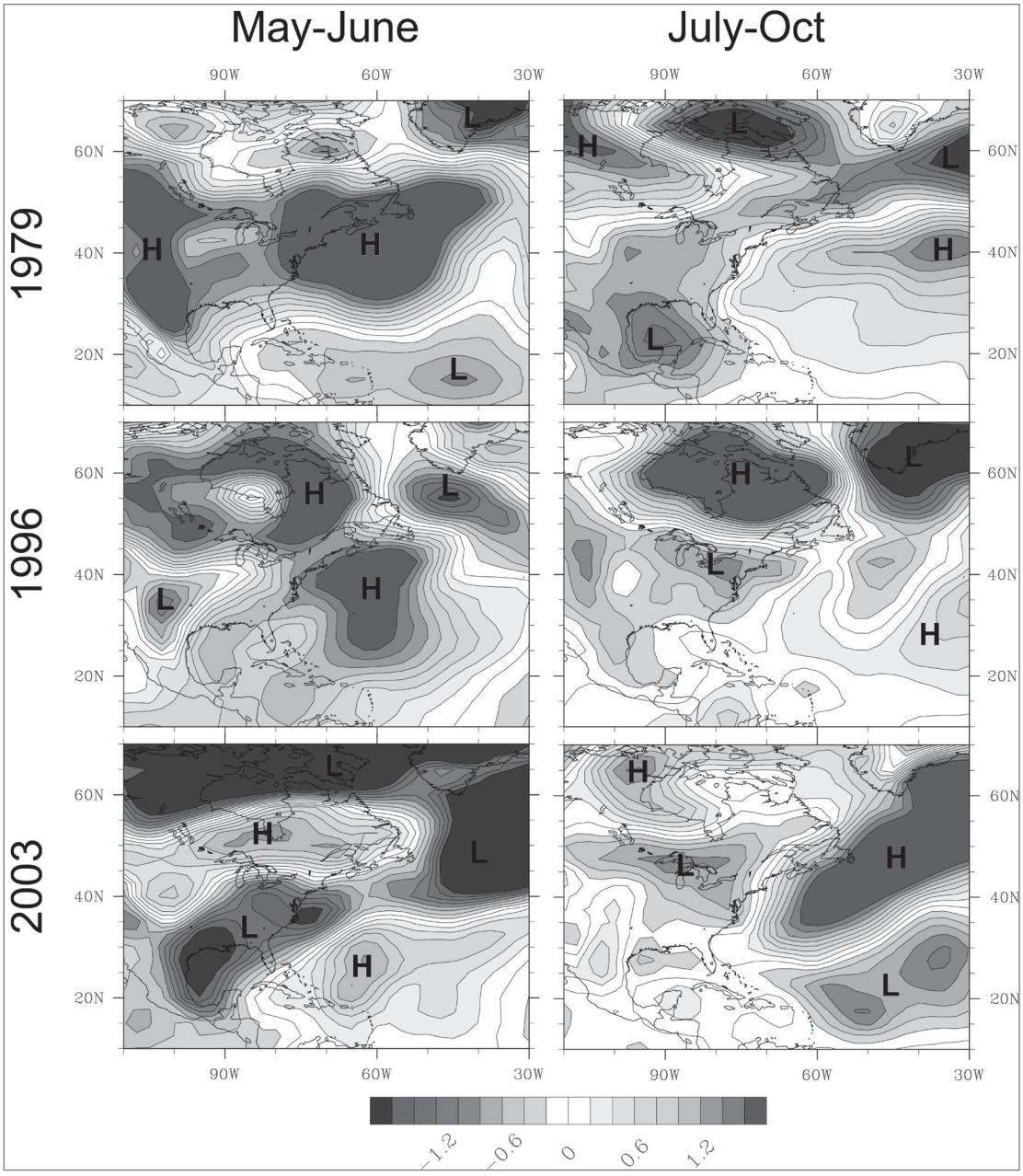

Figure 2.11. Maps of sea level pressure anomalies (in millibars; base period 1979-2013) for select months of the three years with very high precipitation and fecal coliform levels. Panels in the left column show pressure patterns for May through June, while the right-hand column shows patterns for July through October.

Chapter 3: Precipitation thresholds for fecal bacterial indicators in the Chesapeake Bay

Abstract

Many coastal states of the United States restrict harvest of shellfish from select areas based on some environmental trigger. In Maryland, the trigger is an inch or more of rainfall that has fallen in the last 24 hours. This study used 11 years of monitoring data to test the relationship between daily rainfall totals and the levels of fecal indicators in Maryland shellfish harvest areas. Precipitation data and fecal coliform monitoring data from 2004 to 2014 were matched by date and watershed. Measurements of antecedent rainfall conditions (i.e. rainfall in the preceding days or weeks), distance to land, and water column depth at each monitoring station were compared to the percent of samples exceeding the FDA criterion for managing shellfish harvest areas. Sample stations beyond 1000m from land had fecal coliform (FC) levels consistently below the FDA criterion and were excluded from further analysis. Rainfall events greater than an inch tended to result in elevated FC for the following two days, followed by lower levels thereafter. The total amount of rain in the last three weeks was positively related to the proportion of samples with FC greater than the FDA criterion. Bay-wide, the percent of samples exceeding the FDA criterion rose from seven percent for rainfall less than an inch to 37% following one or more inches of rain. Watersheds were classified based on the percent of FC densities over the criterion when rainfall was an inch or more, with 41 of 81 watersheds showing FC responses indicative of potential conditionally approved areas. These areas largely overlapped the current conditionally approved areas defined by Maryland. The percent of open water, wetlands, and poorly drained soils explained a significant amount of the variability in the difference in percent of samples exceeding the FDA criterion when rainfall was greater than an inch and

when it was less than an inch. Logistic regression analysis showed that the current trigger of one inch of rain in 24 hours is predictive of FC densities over the FDA criterion, though the appropriate threshold will most likely depend on how far the particular shellfish growing area is from land and antecedent rain conditions. In watersheds with relatively high percentages of open water to total watershed size, higher rainfall thresholds might be appropriate. The approach taken in this study could be applied to individual stations and sub-watersheds, likely allowing the reclassification of some shellfish harvest areas. The approach could also be easily applied to assessments in other States.

Introduction

Fecal indicator bacteria in estuarine waters are used as indicators of fecal pollution and an increased risk of encountering human pathogens in nearby shellfish (Ashbolt et al. 2001, FDA 2015). In areas where non-point sources are present or stormwater overflows occur, fecal pollution in estuaries is often related to rainfall (Kelsey et al. 2004, FDA 2015). Thus, in the United States (US), under the Food and Drug Administration (FDA) guidance for management of shellfish harvest (FDA 2015) and Environmental Protection Agency (EPA) guidance for recreational water use (EPA 1986), access to these natural resources may be restricted after a specified level of precipitation. For shellfish, such areas are classified as 'conditionally approved' meaning that harvest is conditional upon some set of environmental conditions. Shellfish harvest in many US coastal states is prohibited from conditionally approved beds whenever a specified amount of rain has fallen in the last 24 hours.

In Maryland, the classification of conditionally approved shellfish beds is based on field studies that identified conditions generally meeting the FDA criteria except after large rainfall events. The bulk of these studies occurred in the 1980's. The threshold for closure, one inch of rain in the last 24 hours, was established by quantifying fecal coliform densities in surface waters during and after rain events in select areas of the Bay in 1987. The level of

fecal coliforms in surface waters was measured at several times over several days following rain events (Kathy Brohawn, MDE, personal communication). Although this process provided evidence of the link between rainfall and FC densities in water and the extent of time that elevated FC densities occurred, it was limited in the number of samples, watersheds, and rain events tested. Therefore, the amount of rain necessary to produce significant runoff may not have been thoroughly investigated, nor were antecedent rainfall conditions (i.e. rainfall that occurred in the days or weeks preceding the current rain event) always taken into account. Additionally, field studies designed to assess the impact of rain events did not quantify related factors, such as land use or soil types, which may affect levels of fecal pollution.

Studies in other aquatic systems have assessed the relationship of fecal indicator bacteria, such as FC, in surface waters to a large number of environmental variables and, in some cases, have been able to develop models with moderate capability to predict fecal indicator bacteria densities (Mallin et al. 2001, Maimone et al. 2007, EPA 2010b, Kelsey et al. 2010, Gonzalez et al. 2012). Predictive models for fecal indicator bacteria in recreational waters at beaches has been of particular focus (EPA 2010b). Variables found to be predictive in previous studies included rainfall, wind velocity, turbidity, water temperature, and riverflow. These studies primarily focused on using linear regression models (Kelsey et al. 2010) or decision trees (Maimone et al. 2007) to provide guidance for risk of fecal bacterial densities exceeding established criteria, though some use of logistic regressions has attempted to predict probability of occurrence (Eleria and Vogel 2005). The predictive power of rainfall and/or riverflow (typically a function of rainfall) for fecal indicator densities relates to the land-based source of most fecal bacteria (Kelsey et al. 2004). In general, the ability to predict concentrations of fecal indicator bacteria in natural water bodies with low uncertainty has proved challenging (Novotny and Olem 1994, EPA 2010b). For convenience, the use of a

rainfall threshold that is predictive of excessive fecal bacteria may serve as a tool for shellfish managers to make decisions about shellfish area closures based on the relative risk of having fecal pollution in the growing waters. An assessment of fecal coliform densities at California beaches following large storm events underscored the utility of precipitation thresholds (Ackerman and Weisberg 2003).

In this study, empirical data was used to examine the relationship between fecal coliform densities and precipitation in Maryland's estuarine waters. The null hypotheses were that the level of precipitation necessary to result in fecal coliform densities in excess of the FDA criterion is at least an inch and is uniform between various small watersheds across Maryland's portion of the Chesapeake Bay. Supporting hypotheses were that the relationship between precipitation and FC levels was not influenced by the distance of the monitoring stations in each watershed to land, antecedent rain conditions, wind speed, and air temperature (as a proxy for seasonal patterns). The response of FC densities to rainfall was further compared to characteristics of the watershed, such as the percent of open water, impervious surface and soil types. Logistic regressions were used to assess the amount of rainfall resulting in a significant probability of fecal densities exceeding the FDA criterion for management of shellfish harvest areas.

Materials and Methods

Meteorological Data

Several sources of precipitation data were considered, including daily estimates from National Weather Service weather stations as well as estimated rainfall based on Doppler radar images. Ultimately, rainfall estimates produced by the Middle Atlantic River Forecast Center (MARFC) called Multi-Sensor Precipitation Estimates (MPE)

(http://www.weather.gov/marfc/Multisensor_Precipitation) were chosen, primarily due to their use by MDE to regulate closures of their conditionally approved shellfish beds and the relatively fine spatial coverage (grid size is approximately 16km^2). Some MPE data was excluded from our analysis on the advice of the MARFC (Jason Nolan, MARFC, personal communication). MPE data prior to 2004 was excluded based on a lower level of confidence in the estimates, and data for the months January through March were excluded because the radar precipitation estimates are not as accurate for frozen precipitation and suffer from 'bright banding' - where melting snow registers as large raindrops. The archived MPE data represents 24 hour estimates of total precipitation in inches from 8:00pm to 8:00pm, the same time span used by MDE for conditionally approved shellfish area closures. Antecedent rainfall amounts were calculated by summing previous rainfall amounts for each day up to a week and then by week up to a month prior to the target date.

Wind speed and air temperature data was gathered from the Global Historical Climatology Network-Daily (GHCN-D) database through the National Centers for Environmental Information (NCEI) (www.ncei.noaa.gov/, accessed 4/7/2016). Mean daily air temperature and wind speed data were chosen from NOAA weather stations based on their completeness of record and geographic locations. Air temperature data came from weather stations at the Conowingo Dam (USC00182060) and Royal Oak (USC00187806) while wind data came from Baltimore/Washington National Airport (USW00093721) and Salisbury/Wicomico Regional Airport (USW00093720). Temperature and wind data were averaged by day between stations and, to account for sharp changes in air temperature between consecutive days that might not have been reflected in water temperatures, an average of the air temperature for the day of sampling and the previous day was used.

Station and Watershed Data

Water quality stations monitored by the Maryland Department of Environment (MDE) were selected for this study based on the frequency and duration of sampling. All monitoring stations in Maryland's portion of the Chesapeake Bay sampled consistently from 2004 to 2014 were considered, in order to match the time period of the precipitation data. This provided data from 509 stations, with an average sample count of 107 per station and a total of 54,580 observations. MDE typically collects data from these stations twice a month, though gaps exist, due primarily to winter ice conditions and extreme weather events.

Distance from each MDE station to the closest point of land was calculated by plotting the stations in ArcMap (version 10.2.2; ESRI, Inc.) and using the Distance tool and a map of the Chesapeake Bay waterline made by the Chesapeake Bay Program (CBP 2003). Average water depth from mean low tide at each station was also identified from a map of topography created by The Nature Conservancy (TNC 2013).

United States Geological Survey (USGS) Hydrodynamic Unit Code 12-Digit watersheds (HUC12) were used to compare precipitation and fecal coliform levels (http://water.usgs.gov/GIS/huc). These watersheds are the smallest of the watershed delineations included in the USGS Watershed Boundary Dataset (http://nhd.usgs.gov/wbd.html) and are the closest in size to the current conditionally approved shellfish harvest areas in Maryland (average size approximately 3km^2). MDE monitoring stations were associated with the USGS HUC12 watershed in which they fell (ESRI ArcMAP; spatial join tool). These stations were located within 81 HUC12 watersheds, with an average of 6 stations per watershed. MPE precipitation grids were associated with USGS HUC12 watersheds in which any part of the grid fell. The Chesapeake Bay mainstem (020600010000), Eastern Bay (020600020609), Lower Choptank River mainstem (020600050508), and the Lower Potomac River (020700111001) HUC12 watersheds were

excluded because their watersheds consisted of greater than 99% open water and they likely receive riverine and tidal inputs from larger areas than the other watersheds.

Land use characteristics and soils data for each watershed were collected from the National Oceanic and Atmospheric Administration (NOAA) Coastal Change Analysis Program (C-CAP) (https://coast.noaa.gov/digitalcoast/tools/lca) and the United States Department of Agriculture (USDA) Soil Survey (http://websoilsurvey.nrcs.usda.gov), respectively.

Analysis

Precipitation and FC data were matched by HUC12 watershed and date (see Figure 3.1 for example). Median daily precipitation levels were calculated for each watershed and day. The US FDA criterion of 49MPN/100mL for fecal coliforms was used to assess response to rainfall, where MPN stands for an estimate of bacterial concentration called the most probable number (FDA 2015). In the US, shellfish harvest area closures are based on comparison of the levels of fecal indicator bacteria in ≥30 samples for areas of non-point sources and ≥15 samples for areas of point sources (e.g. wastewater treatment plants) to multiple FDA criteria (FDA 2015). Maryland uses the three-tube dilution method for estimating FC densities, which involves inoculation of selective liquid culture media (A1) at three different volumes of sample water (0.1, 1.0, and 10.0 mL) followed by incubation in waterbaths at 44.5°C (APHA 1998). When using this three-tube dilution method, the median FC density estimate must be below 14MPN/100mL, while the calculated 90th percentile estimate must fall below 49MPN/100mL for the area to remain open to harvest.

Management of existing conditionally approved areas is slightly different than other areas in that the classification is based on the most recent 15 samples collected during the open status (without rainfall) in comparison to the 49MPN/100mL criterion. The 49

MPN/100 mL criterion is used because it is more indicative of conditions resulting from an introduction event, such as rainfall or stormwater overflow (FDA 2015), whereas the median value is more likely to exceed the 14 MPN/100mL criterion in the absence of an event. The percent of samples for each monitoring station that exceed 49MPN/100mL is used to confirm response of the area to rain events when assessing the area for proper classification. If less than 10% of samples collected when rainfall was not present are below 49MPN/100mL, the area remains classified as conditionally approved. Because the focus of this study was to assess responses at the watershed level, the percent of samples over 49MPN/100mL for a watershed was used instead of the station-based percentile.

Frequency Analysis

The relationship of rain and fecal coliforms was first examined by comparing the frequency of samples above the 49MPN/100mL FDA criterion with the number of days since the last daily rainfall amount over an inch. The potential effect of antecedent rainfall on the levels of FC resulting from a rain event greater than an inch was tested by comparing the amount of rainfall in the previous one, two, three and four weeks, exclusive of rainfall over the previous two days, to FC levels after a daily rainfall greater than an inch. Based on these analyses, the greater of the rainfall amounts for the previous two days (Rain2Day) was used for all subsequent comparisons to FC densities and any Rain2Day amount greater than one inch was considered a RainEvent. This criterion was chosen over a cumulative amount of rain for the previous two days because it is in keeping with the current management strategy of closing conditionally approved beds after a 24hr total for rainfall.

In order to assess the use of rainfall as an appropriate variable for conditional closure of Maryland shellfish beds, the percent of FC densities over 49MPN/100mL following a RainEvent, both at the bay-wide scale and by HUC12 watershed were compared. A set of decision points was used to classify HUC12 watersheds based on these frequencies (Figure

3.2). The first decision point was whether or not there were at least three RainEvents sampled. For those watersheds with three or more RainEvents, the next decision point was whether or not the percentage of samples over 49MPN/100mL when a RainEvent had not occurred was less than 10%, a threshold similar to the 90th percentile management rule used to classify shellfish growing areas (FDA 2015). For those watersheds with less than 10% FC exceedance of the FDA criteria without RainEvents, those with greater than 10% exceedances after RainEvents were considered responsive watersheds, while those that had less than 10% exceedances after RainEvents were considered not responsive to rainfall, a less stringent standard than would likely be required to classify an entire watershed as conditionally-approved.

Comparison of Frequencies to Watershed Characteristics

For those watersheds determined to be responsive to rainfall, the difference in percentage of samples over the 49MPN/100mL criterion when a RainEvent occurred versus the percentage when rainfall was less than an inch (hereafter called PDiff) was compared by Pearson correlation tests to the percent of various land use classifications and the percentage of soil groups within each watershed. Classification and regression tree (CART) analysis was also conducted to identify the most important land use and soil-type variables and corresponding thresholds that best categorize the range of PDiff values. A correlation matrix revealed strong linear relationships (r >0.70) between types of developed land in the watersheds (i.e. high intensity, medium intensity, etc) and between cropland and pasture, so these were combined into total development and total cropland percentages, respectively, before CART analysis. The appropriate number of branches for the tree (pruning) was determined using cross-validation (K=10) and by choosing the tree that had the minimum sums of square error (Lemon et al. 2003). The results from the frequency analysis were also used to select watersheds for inclusion in the logistic regression analysis described below.

Logistic Regression Analysis

Logistic regressions were conducted to assess the amount of rainfall (Rain2Day) leading to exceedance of the FDA criterion. Specifically, the 50% probability of a surface water sample to contain FC exceeding 49MPN/100mL was calculated as a function of the greater amount of rainfall for the previous two days (Rain2Day), the total amount of rainfall in the previous three weeks (excluding the previous two days), the distance of each sample from the nearest shoreline, air temperature, and the average wind speed for the previous day. For this analysis, the FC density for each sample was converted to a binary response variable as either above or below the 49MPN/100mL criterion. An analysis of all bay-wide data from within 1000m of land was first conducted. Models were compared using Akaike's Information Criteria (AIC) and the area under the Receiver Operating Curve (ROC) (Hastie et al. 2009). The AIC is a measure of model parsimony, which is the amount of variance explained relative to the number of variables included in the model. The area under the ROC is a measure of correct classification of observations. Smaller AIC and larger ROC values are indicators of more parsimonious and accurate models, respectively. Parameter estimates were assessed using bootstrapping techniques with 1000 iterations of model run (Kelsey et al. 2010). In order to further evaluate the ability of the chosen model to correctly predict probability of exceeding the FDA criteria, data were binned by model predicted probability to the nearest 0.5 percent and compared to the observed frequency of occurrence for the raw data in each bin (Jacobs et al. 2014). The predicted probability and percent occurrence values were then compared using linear regression. Logistic regression analyses using the same predictor variables from the selected bay-wide model were then conducted for each of the watersheds that were considered to be responsive to rainfall based on the decision points used in the frequency analysis. Logistic regression results were reported for watersheds with significant relationships between rainfall and the proportion of samples over 49MPN/100mL

($p<0.05$), with a percent concordance greater than 75%, and sufficient data to calculate a rainfall threshold with 95% confidence limits.

The logistic regression approach was further tested by application to an existing conditionally approved area as characterized by MDE. Although many of the existing conditionally approved areas were established decades ago and included targeted sampling efforts around precipitation events and sampling stations that no longer exist, sufficient data existed to test the logistic regression approach for several of them. One of these is the St. Mary's River. There were only two stations located in or beside the conditionally approved area at the head of the east branch of the St. Mary's River that were consistently monitored from 2004 to 2014 (Figure 3.3). One hundred fifty six samples, including those from 6 RainEvents, were collected over that time span. Because logistic regression results may be biased by such small sample numbers, the Firth bias correction method (Firth 1993) was used to estimate model parameters.

Analysis Software

Datasets were combined and queried using SAS (SAS Institute, Inc., Cary, NC). The frequency analysis was also conducted in SAS. Logistic regressions, including Firth bias correction, and bootstrapped parameter estimates were performed using R (The R Foundation for Statistical Computing, version 3.3.1). Classification and Regression Tree Analysis (CART) were also conducted using R.

Results

A decrease in FC levels was noted with increasing distance from land, with all stations beyond 1000m having less than 10% of their samples above 49MPN/100mL, the FDA criterion for restricting harvest (Figure 3.4, panel A). Plotting river depth for each

station, as a proxy for distance to land and from headwaters, versus the 90th percentile FC densities revealed a similar trend in fecal levels, with monitoring stations having a water column depth of 5m or greater primarily having fecal coliform densities below 49MPN/100mL (Figure 3.4, panel B), though with several outliers. Distance to land was used instead of depth for subsequent analyses, due to the fewer number of outliers. Based on these findings, stations greater than 1000m from land were excluded from further analysis.

A distinct relationship between FC and rainfall was seen when comparing the number of days between a RainEvent and the date of surface water sampling (Figure 3.5, Panel A). Greater than 30% of all samples within 1000m of land contained FC above the 49MPN/100mL criterion when a RainEvent occurred in the previous two days. However, this percentage decreased sharply after two days to less than 20 percent and continued to decrease until it reached a level below 10 percent after five days. Antecedent rainfall, the total amount of rain over the previous three weeks exclusive of the last two days, was also significantly and positively related to FC densities that were detected following a RainEvent (Table 3.1), though the R^2 values indicate antecedent rainfall explained very little of the variation in the FC data. The amount of rainfall within the previous three weeks (Figure 3.5, Panel B) had the highest R^2 value (0.16) of the time periods considered. The larger percentage of violations that were found when total 3-week rainfall was 9 inches resulted from extremely high FC densities in samples collected directly after Super Storm Sandy in 2011. Wind was associated with increased proportions of samples over the FDA criterion for several speeds over 10mph (Figure 3.5, Panel C). Average air temperature for the day samples were collected and the previous day had a generally negative relationship with the proportion of samples containing FC over the FDA criterion (Figure 3.5, Panel D). However, a significant drop in the proportion of samples over the criterion occurred in the 35°F bin relative to the bins on either side of it. This large discrepancy results from late winter samples

(November and December) with wet and dry winters generally falling into different temperature bins.

Frequency Analysis

Seven percent of water samples, collected within 1000m of land, contained FC exceeding the 49MPN/100mL criterion when rainfall for each of the previous two days was less than one inch. In contrast, 37% of the samples, nearly a six-fold increase, exceeded the criterion after a RainEvent. When analyzed at the watershed (HUC12) scale, 41 watersheds contained less than 10% exceedances of FC over the criteria when rainfall was less than an inch and greater than 10% after a rainfall over an inch. These watersheds were classified as being responsive to rainfall (Figure 3.6). The difference in percent exceedance between samples above and below one inch of rainfall (PDiff) ranged from 3% to 68% for watersheds classified as responsive (Figure 3.7). Twenty-seven of the 34 conditionally approved shellfish harvest areas, as designated by MDE, fall within these responsive watersheds (Table 3.2). Five of the remaining areas classified by MDE as conditionally approved were located in watersheds where there were insufficient numbers of RainEvents sampled (<3) to make a reasonable assessment. One of the MDE conditionally approved areas, Bolingbroke Creek, resides within watersheds having more than 10% of the samples above the FDA criterion of 49MPN/100mL.

Comparison of Frequencies to Watershed Characteristics

For each watershed, the difference in the percent of samples with FC densities over the 49MPN/100mL criterion when rainfall was less than an inch versus the percent when rainfall was greater than an inch (RainEvent) was not significantly related to the number of RainEvents, the average amount of rainfall for all Rainevents, or the maximum rainfall amount (p value > 0.05). However, PDiff was significantly correlated to several land use and

soil type characteristics (Table 3.3). The strongest relationship was a negative correlation to the percent of open water in the watershed, followed by a negative relationship to the percent of unconsolidated shoreline, and a positive relationship to the percent of cultivated crops. Unconsolidated shores are those lacking permanent vegetation and composed of sediment which is prone to inundation and redistribution by waves (https://coast.noaa.gov/digitalcoast/tools/lca). Classification and regression tree analysis also showed that the percent of open water is the strongest indicator of how FC densities respond to rainfall and that a percentage of open water for a watershed above 25.8% tends to result in lower levels of FC following RainEvents (Figure 3.8). For those watersheds with less than 25.8% open water, the next most important variable was a significant negative relationship to the percent of palustrine forested wetland in the watershed. Based on the wetlands classification scheme used by NOAA's C-CAP (NOAA 2016), palustrine forested wetlands are low salinity, non-tidal wetland areas primarily consisting of woody vegetation that is at least 5m in height and total vegetation coverage greater than 20 percent. The next most important variable for classification was the percentage of C/D soils in the watershed. These are soils that have slow to very low infiltration rates and relatively high runoff potentials (see Appendix D for description of soil types). Altogether, the percentages of open water, palustrine forested wetland, and C/D soils account for 72% of the variability in the PDiff for samples exceeding the 49MPN/100mL threshold relative to RainEvents. The average distance of the stations in each watershed to land was not correlated to PDiff, suggesting that distance to land is not important in how many samples exceed the criteria, but may be important in the amount of rainfall required to result in FC densities above the FDA criterion.

Logistic Regression Analysis

Logistic regression analysis of pooled data from all of the responsive watersheds revealed a predictable increase in the probability of FC exceeding the FDA criterion with

increasing amounts of rainfall, increasing total amount of rainfall over the last three weeks, decreasing air temperature, and decreasing distance from land. Analysis of data from all responsive watersheds combined produced a logistic model with a concordance of 76.5% when including all of these independent variables, and a concordance of 60.2% when only including recent rainfall (Rain2Day) in the model (Table 3.4). Tests for variance inflation indicated that no significant collinearity existed between the predictor variables or between the predictor variables and the intercept (α 0.05, p-value > 0.05). The value for Rain2Day corresponding to a 50% chance of exceeding the 49MPN/100mL criterion, under average antecedent rainfall and distance from land, was between 2 and 2.2 for all models tested. Although the inclusion of air temperature lowered the AIC, it did not significantly raise the area under receiver response curve. Wind speed was not a significant variable (p-value>0.05) in the model. Therefore, the model using only rainfall (Rain2Day and 3wkRainfall), distance from land, and antecedent rainfall was chosen for further testing. For the 50% probability threshold, the percent of false negatives (observations above the FDA criterion predicted to be below the criterion) was low (<10%) while the percent of false positives (observations below the FDA criterion predicted to be above the criterion) was higher (> 10%). However, a plot of the observed proportion of samples over the 49MPN/100mL criterion within each of the predicted probability bins (Figure 3.9) shows a significant, positive relationship, with some notable, nonlinear scatter for bins close to 50% probability. The negative relationship between probability of exceedance and air temperature revealed a significant seasonal pattern. The selected model estimated a Rain2Day value of 2.16 inches for a 50% probability of exceeding the FDA criterion and an odds ratio indicating that there is 4.59 times greater odds of exceeding the criterion for every one inch increase in rain (Table 3.5). Bootstrapping showed that there was relatively small error associated with the parameter estimates for this model (Table 3.6).

Logistic regression analysis, conducted for each of the responsive watersheds revealed statistically significant models with concordance above 75% and sufficient data to determine 95% confidence limits for 10 of the 41 watersheds (Table 3.7). The percent of false negatives remained very low, while the percent of false positives ranged from 15% to 52%. The amount of rainfall (Rain2Day) corresponding to a 50% chance of FC exceeding the FDA criterion, based on average watershed values for all other significant variables, ranged from 1.2 to 2.4 inches (Table 3.7). These rainfall probability threshold values had a significant, positive relationship to the percentage of open water in the watershed and to highly impervious soils (type D) (Figure 3.10), but were not related to the time of year that the exceedances occurred ($p > 0.05$).

Application of the logistic approach to an existing MDE conditionally approved area in the St. Mary's River produced superior results to the watershed-scale analysis. Testing for the probability of FC exceeding the 49MPN/100mL criterion confirms the findings for this watershed and provides strong evidence for the dominant influence of rainfall on FC densities in this conditionally approved area (Table 3.8). With an odds ratio of 50.9, and no significant variable other than Rain2Day remaining in the model, the relationship between FC densities and rainfall was highlighted.

Discussion and Conclusions

The findings from this study underscore the importance of rainfall in effecting FC densities in estuarine surface waters and provide evidence that the effect of rainfall on these densities varies between different watersheds. The amount of rainfall in the last 2 days (Rain2Day) provided a convenient and useful metric for separating watersheds based on the frequency of encountering high densities of fecal coliforms and was consistently more

important (highest odds ratio) than antecedent rainfall, wind speed, and air temperature in determining the probability of high FC densities. These findings both confirm the importance of rainfall in surface water conditions and supports the use of rainfall as a conditional closure tool in select areas. The drop in FC densities after two days post RainEvent also suggests the current management strategy of closing conditionally approved areas for a defined number of days is protective from a human health risk perspective. However, the proportion of samples below the FDA criterion did not fall below 10% for five days after a RainEvent, similar to results found for California beaches (Ackerman and Weisberg 2003). If all of the stations less than 1000m from land were part of conditionally approved areas then there would likely be some times and/or places when the 72 hr rule was not fully protective. Any establishment of a new conditionally approved area would need to measure the rate of fecal coliform dilution and decay.

The frequency analysis also provided key insights into the influence of land-based conditions on FC densities in shellfish harvest waters. Other studies (Mallin et al. 2001, Glasoe and Christy 2004, DiDonato et al. 2009) have found statistically significant relationships between impervious surfaces and/or urban development on indicator bacteria densities in coastal systems. These land use types may help explain the excessive number of samples over the FDA criteria regardless of rain in the heavily urbanized watershed draining Baltimore City and two agricultural watersheds on the Eastern Shore. However, this study found that the frequency of FC densities exceeding the FDA criterion was highly variable within region and had more connection to characteristics of the water body (percent open water) and land margin (percentages of palustrine forested wetland) than to dominant upland characteristics. The percent of soils with moderate to high runoff potential (C/D soil) and the percent of cultivated crops were also significantly positively related to the frequency of exceedances (PDiff), but explained less variability than the percentage of open water. The

strong relationship to open water, despite the exclusion of stations greater than 1000m from land, suggests FC-laden runoff may be diluted in tidal systems with relatively large proportions of water to land. This finding is supported by the Total Maximum Daily Load (TMDL) modeling efforts of MDE which show that inputs of fecal coliforms from land are often diluted to very low levels within several kilometers of where they are introduced (MDE 2006). More importantly, the significant relationship to open water implies that watersheds with small water-to-land ratios are more prone to rainfall-driven FC densities in exceedance of the FDA criterion. The positive correlation to percentage of C/D soils in the watershed suggests that soils with decreased infiltration rates and higher runoff potential may increase runoff introduction of fecal coliforms following rain events. The significant negative relationship of FC to certain types of wetlands might be expected as both natural and constructed wetlands have been shown to reduce or retain fecal bacteria and pathogens (Green et al. 1997, Knox et al. 2008). However, others (Grant et al. 2001, Huang 2010) have detected a significant positive relationship between wetlands and numbers of indicator bacteria in estuarine waters. Huang (2010) hypothesizes this positive relationship results from the use of wetlands by warm-blooded animals, such as geese and deer. One of the watersheds in this study, the Monie Bay watershed on the lower eastern shore, had high FC densities regardless of rainfall in extensive estuarine wetlands (18% estuarine emergent wetlands, (https://coast.noaa.gov/digitalcoast/tools/lca)) indicating that types of wetlands, based on salinity levels and the kind of vegetation, might be important in their relationship to fecal bacteria in adjacent waters.

The results of the logistic regression analysis for individual watersheds lead to several important conclusions. Foremost, FC densities in exceedance of the FDA criterion in these watersheds are strongly related to rainfall. Secondly, the level of rainfall required to produce FC densities in exceedance of the FDA criterion may vary between watersheds. In

part, this variability is related to the average distance of the sampling stations to land. However, the differences between odds ratios (the odds that an increase in 2DayRain will result in FC concentration of 49MPN/100mL or greater) for Rain2Day between the watersheds suggest different degrees of influence of rainfall on FC densities in these systems. The findings of this study also indicate that the level of rainfall necessary to produce FC densities exceeding the FDA criteria is dependent on a handful of variables, including the distance of the sample stations from land and the proportion of open water that makes up the watershed. These findings agree with those from the frequency analysis, with percentages of open water and soil type influencing not only the frequency of encountering high levels of fecal coliforms but also the amount of rainfall needed to produce those high levels.

Logistic models proved capable of estimating the relative importance of environmental variables on FC densities in surface waters and in estimating the level of rainfall associated with a 50% chance or greater of exceeding the FDA criteria. Although recent rainfall and distance to land had the strongest relationship to probability of exceedance, antecedent rainfall and air temperature also showed significant influence. The negative relationship to air temperature showed a significant seasonal influence with higher probabilities of exceeding the FDA criteria in cooler times of the year. A previous study of FC data in Maryland found an increase in FDA exceedances late in the year (October through December), with a lag in time between peak precipitation and FC levels (Leight et al. 2016). The findings from this study showed that, for a constant distance from land and antecedent rainfall level, the amount of recent rainfall (Rain2Day) necessary to result in a 50% chance of exceeding the FDA criterion varies seasonally. One potential reason for elevated numbers of exceedances in cool time periods is an increase in the amount of rainfall that leads directly to runoff when trees lose their leaves and ground cover has died back. This phenomenon is a known factor in Chesapeake Bay runoff, such that it is included in the Chesapeake Bay

Watershed model used by EPA and relevant states to set criteria for pollution (EPA 2010a). Other possibilities include an increase in larger rain events and therefore greater soil saturation from late year tropical storms leading to increased runoff of FC bacteria, a decrease in deactivation of bacterial cells by ultraviolet radiation, and an increase in the presence of migratory waterfowl, particularly Canada geese (*Branta canadensis*).

By selecting those watersheds that were responsive to rainfall, based on the frequency analysis, the logistic regression was targeted at watersheds where exceedances related to rainfall would tip the balance towards violation of the FDA criterion. Both the predictive capability of the logistic model seen in the observed and predicted probabilities, and the odds ratios for Rain2Day support the importance of rainfall driven violations in the watersheds selected. Application of the logistic approach to a defined growing area in the St Mary's River produced low percentages of false negatives and false positives, a very high odds ratio, and very high concordance values, despite the modest number of samples included in the analysis.

In development of our logistic regression models, one important consideration was the FC density used for assessing violations. FC densities in surface waters after rain events may peak at levels well above 49MPN/100mL. Increasing the FC concentration considered excessive to some value above 49MPN/100mL would increase the concordance and lower the rate of false positives in the logistic regression. However, the 90th percentile criterion used here was chosen based on the existing management structure and understanding of risk for illnesses. Additionally, picking a higher probability of exceedance would lower the percent of false positives, but would increase the percent of false negatives and, therefore, run the risk of not being as protective of human health.

The co-location of most areas currently classified as conditionally approved by MDE with the watersheds defined in this study as being responsive to rainfall supports the classification of these areas and the general use of the conditionally approved classification. The frequency analysis also suggests that FC densities in many of the watersheds in Maryland's portion of the Chesapeake Bay are responsive to rainfall such that they do not exceed the 90th percentile over 49MPN/100mL FDA criterion without significant rain events. The decision in this study to consider fecal coliform levels at the watershed scale to be responsive to rainfall based on 10% or more of the samples exceeding 49MPN/100mL when rainfall was greater than an inch represents a less stringent standard than what managers would need for classifying an entire watershed as conditionally approved. However, the response of fecal coliform densities to precipitation in half of all watersheds and the level of response in some watersheds such that at least 50% of samples after an inch of rain exceed the 49MPN/100mL criterion suggest that targeted sampling would likely uncover new areas suitable for being classified as conditionally approved. Focused analysis, such as that done for St. Mary's River in this study, may be used to target assessments for classification. Any conversion of restricted areas to conditionally approved would have important economic and sociopolitical ramifications for both shellfish harvesters and shellfish aquaculturists. Further, any increases in the precipitation threshold would decrease the number of days that shellfish beds in conditionally approved areas are closed.

Employing logistic regressions, such as those developed in this study, could provide important information when considering such management decisions. While distance to land has been shown previously to be an important driver of FC densities (MDE 2006), the inclusion of this variable in a regression model might provide an informative way to make decisions about conditional closure areas using empirical data. Results from the logistic regression analysis underscore the importance of how far the sampling location is from land,

with decreasing influence of rainfall on FC densities as distance to land increases. By using the average conditions for distance for each watershed and antecedent rainfall the probability values calculated by logistic regression support the use of a one inch rainfall threshold, at least at the watershed scale. Models including not only measurements of rainfall in the last 24 hours, but also air temperature and antecedent rainfall could be used to better estimate the current probability of exceeding the threshold for specific growing areas.

The chief limitation in conducting this study was the relatively modest number of samples collected soon after rain events. Several watersheds with areas currently classified as conditionally approved lacked sufficient rain event data to include in the analyses. With increasing amounts of data, as the MDE monitoring program continues, additional data from samples collected within 48 hours of rain events should improve estimates of rainfall triggers and the appropriate areas for their application. Further, our analysis excluded information from January through March based on the advice of the Middle Atlantic River Forecast office, which produces the rainfall product. Although exclusion of winter precipitation coincides with a large window of oyster harvest from natural beds in Maryland, earlier research (Chapter 2) showed that both precipitation amounts and FC densities in January through March tend to be lower than the rest of the year. One management strategy would be to exclude any frozen precipitation as a trigger for closing conditionally approved shellfish areas. Another important consideration for maintaining conditionally approved areas and understanding rainfall triggers is that shellfish collections other than oysters from natural bars (principally clams) and oysters grown by the burgeoning aquaculture industry in Maryland are permitted year round. This importance is underscored by the current MDE strategy which uses numbers of FC from year-round sampling to make management decisions about shellfish harvest closures.

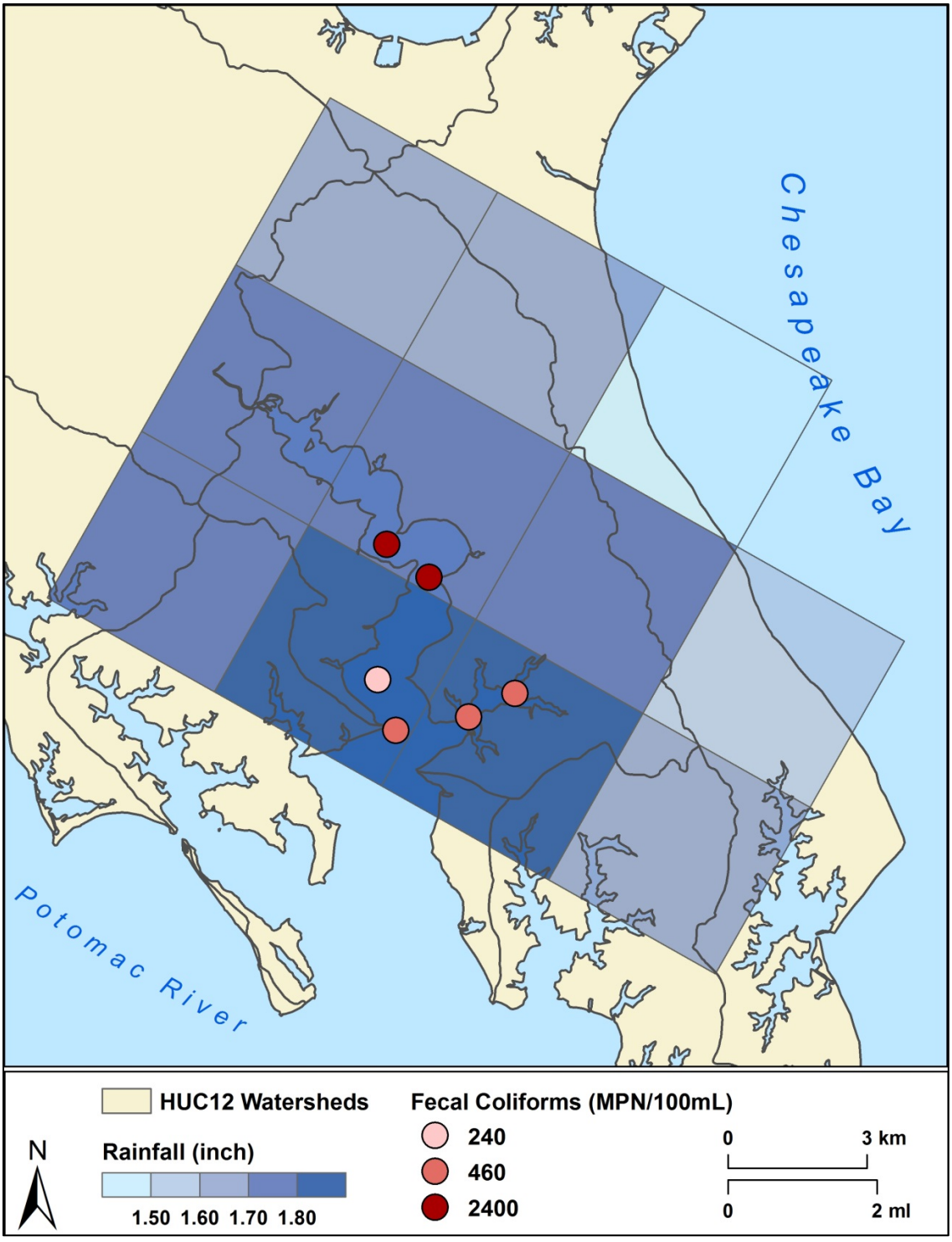

Figure 3.1. Example of a RainEvent that coincided with MDE sampling.

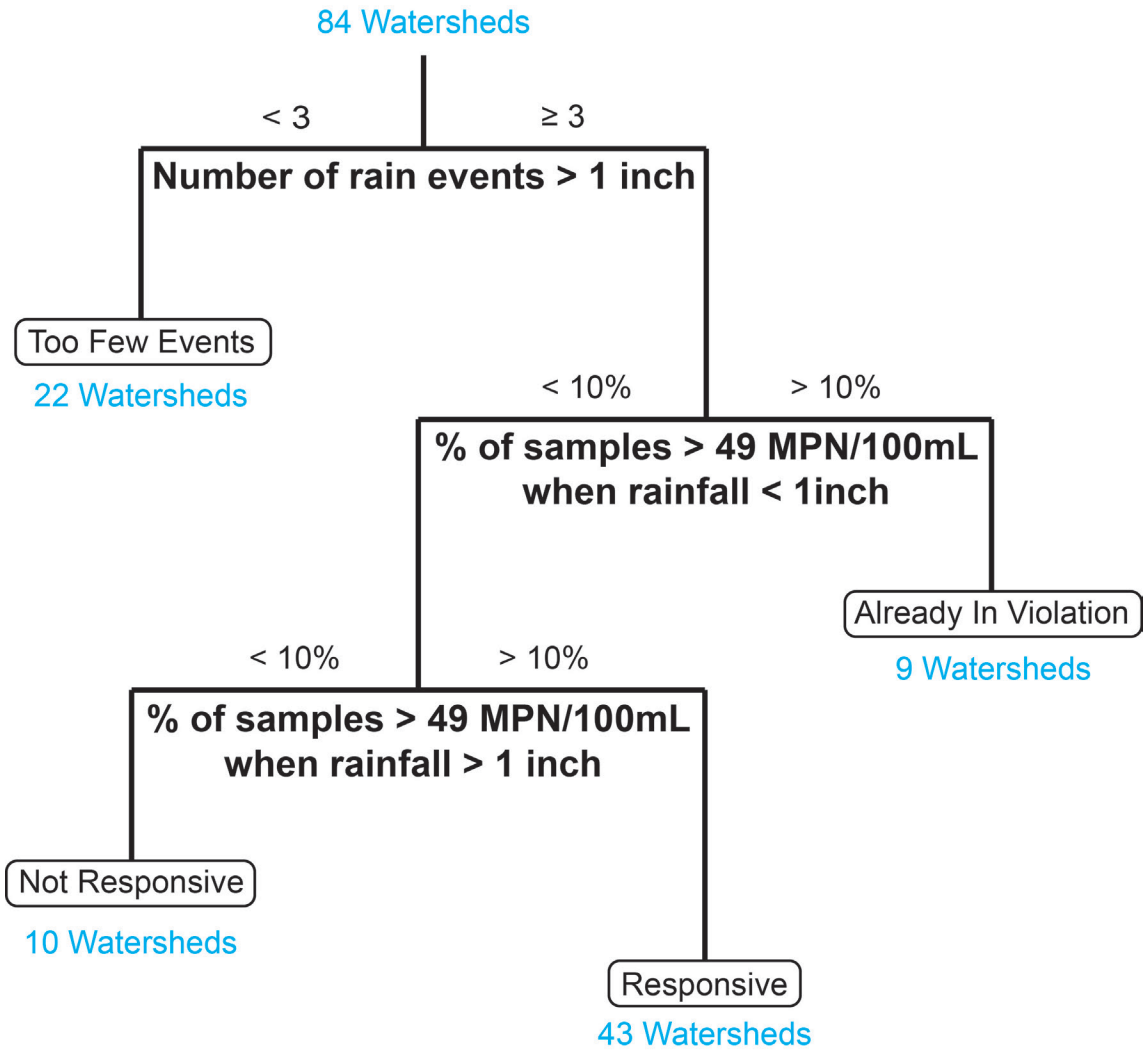

Figure 3.2. Decision points used to classify watersheds relative to their responsiveness to rainfall events of one or more inches of rain. Responsive watersheds had less than 10% of samples with fecal coliforms below the 49MPN/100mL criterion when there was less than an inch of rain and an increase of at least 10% of samples above the FDA criterion when rainfall exceeded an inch. Watersheds with less than 3 rain events were classified as having too few events.

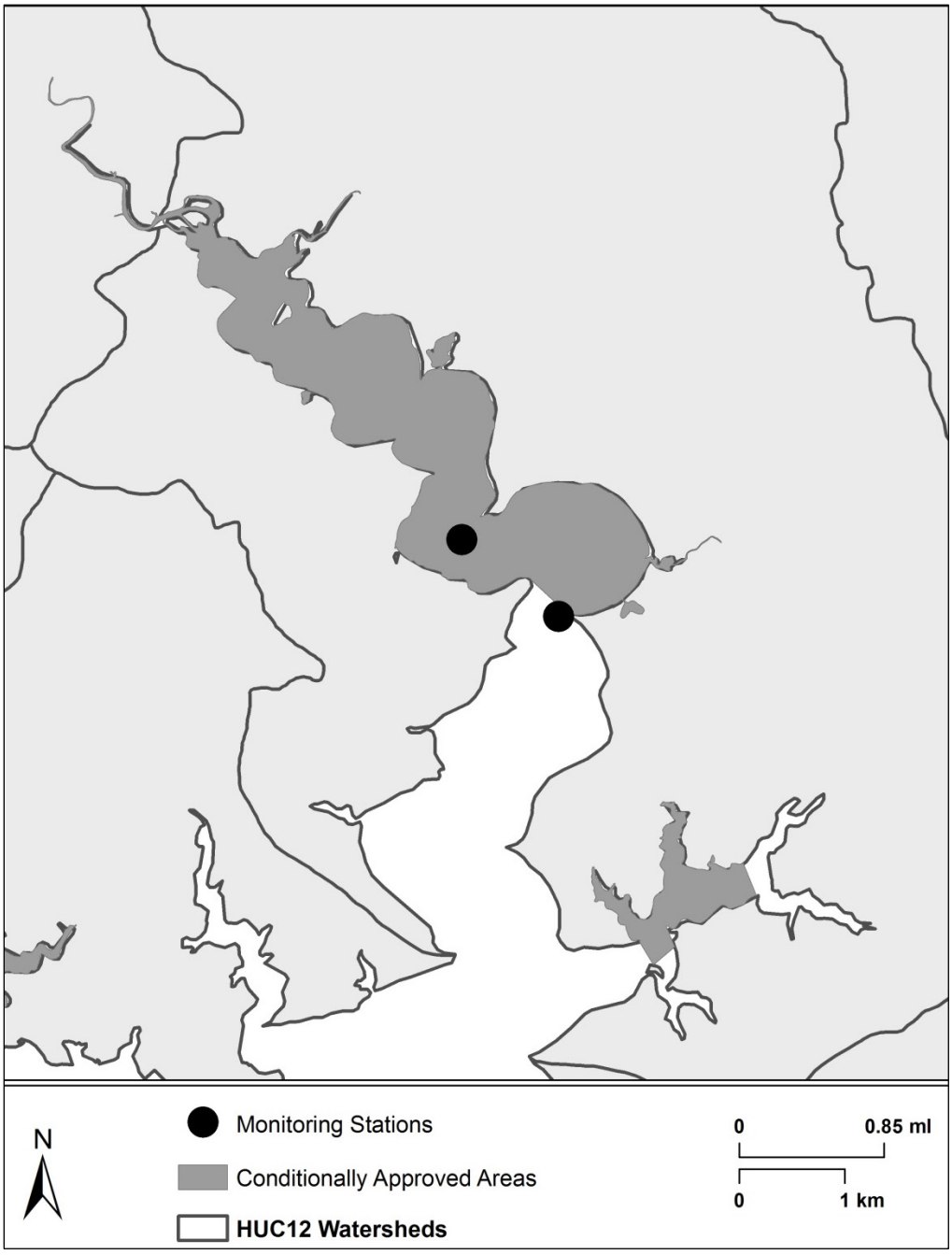

Figure 3.3. MDE monitoring stations that fall within or directly beside the current conditionally-approved shellfish harvest area in the St. Mary's River, Maryland.

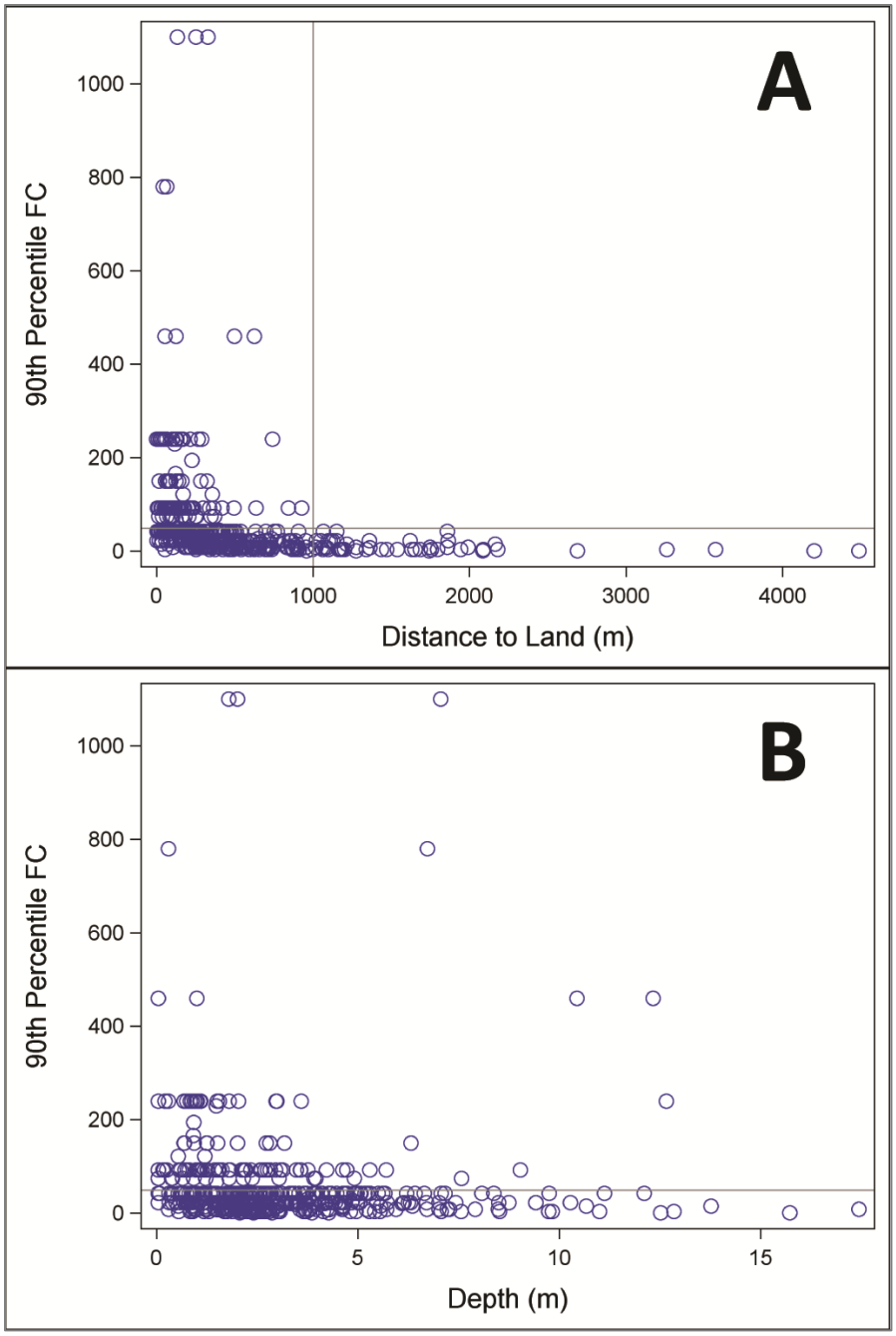

Figure 3.4. Comparison of the 90th percentile of fecal coliform values (MPN/100mL) for each monitoring station with distance (panel A) of that station to the nearest point of land and total depth (panel B).The horizontal reference line indicates 49MPN/100mL.

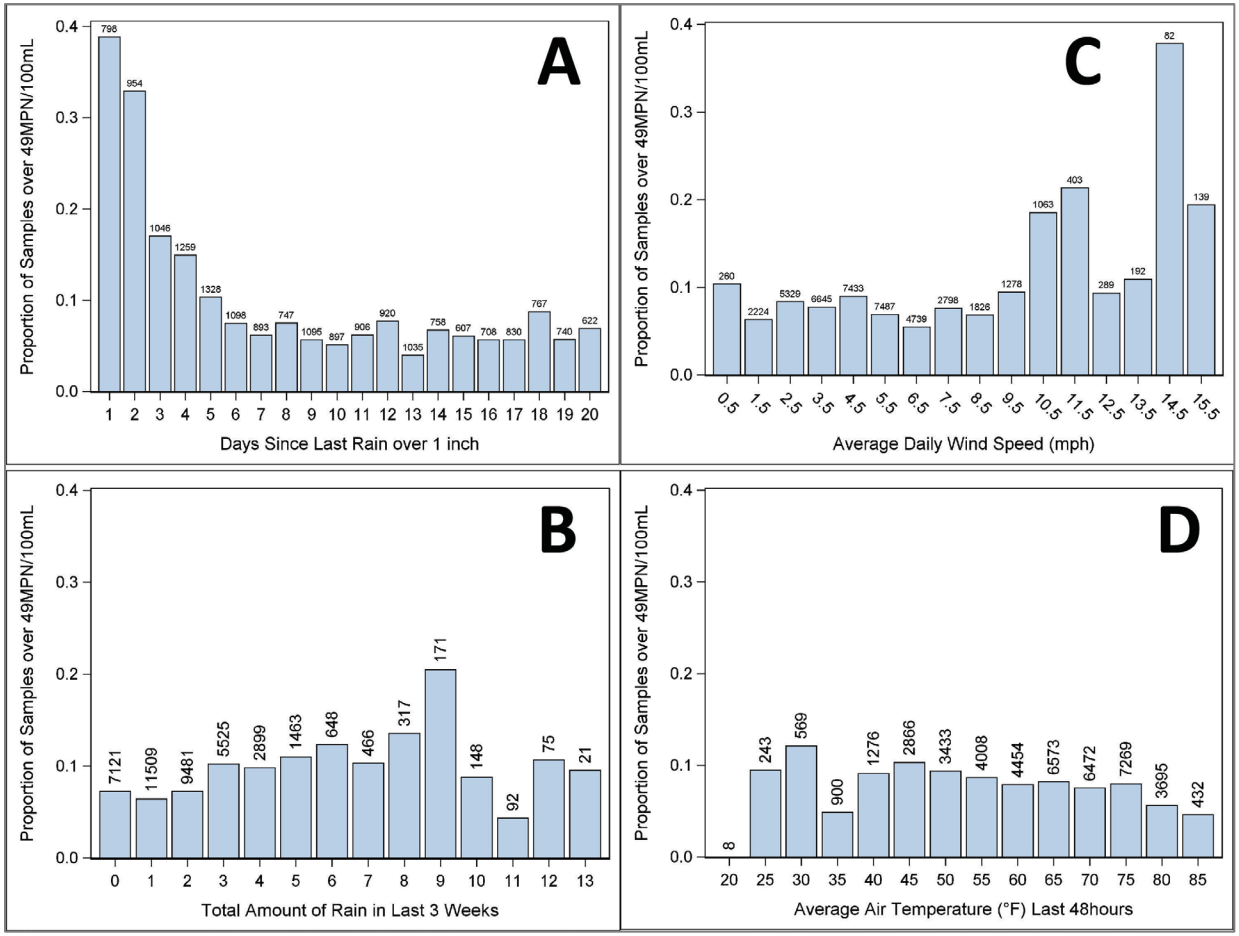

Figure 3.5. Comparison of the proportion of samples with FC over 49MPN/100mL to the number of days since last rain over an inch (Panel A), the total amount of rainfall in the previous month (Panel B), average daily wind speed for the preceding day (Panel C), and average air temperature over the last 48hours (Panel D). Numbers above bars indicate how many samples were present in each bin.

Table 3.1. Results of Pearson correlation analysis of antecedent rainfall and fecal coliform densities. Asterisks indicate total rainfall for the stated period not including the previous two days.

Rainfall total for:	**R^2**	**P-value**
greater of last 2 days	0.334	<.001
previous week*	0.124	<.001
previous 2 weeks*	0.128	<.001
previous 3 weeks*	0.165	<.001
previous 4 weeks*	0.150	<.001

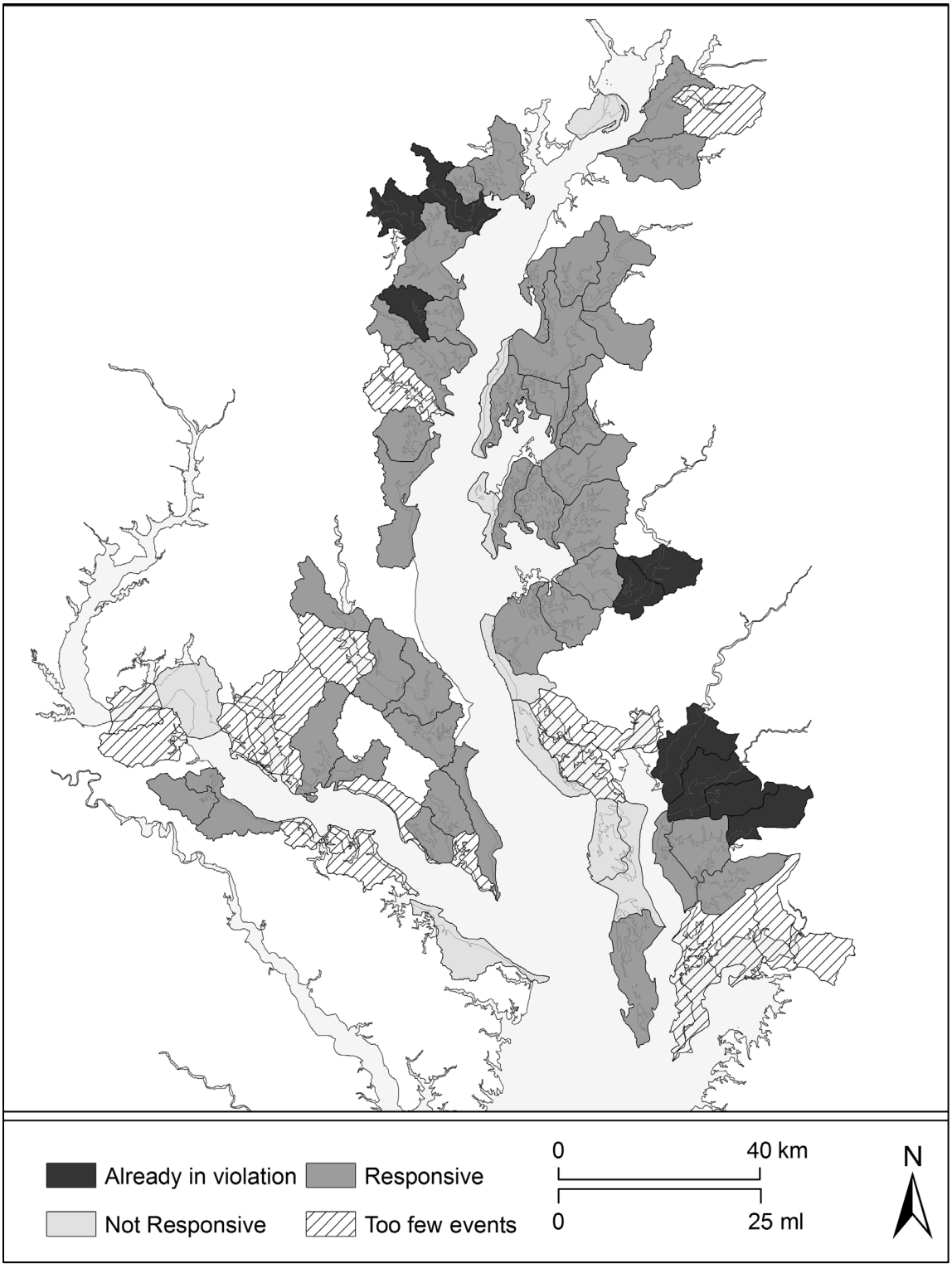

Figure 3.6. Map showing classification of watersheds based on the change in fecal coliform densities relative to rain events of 1+ inches of rain (Rain2Day).

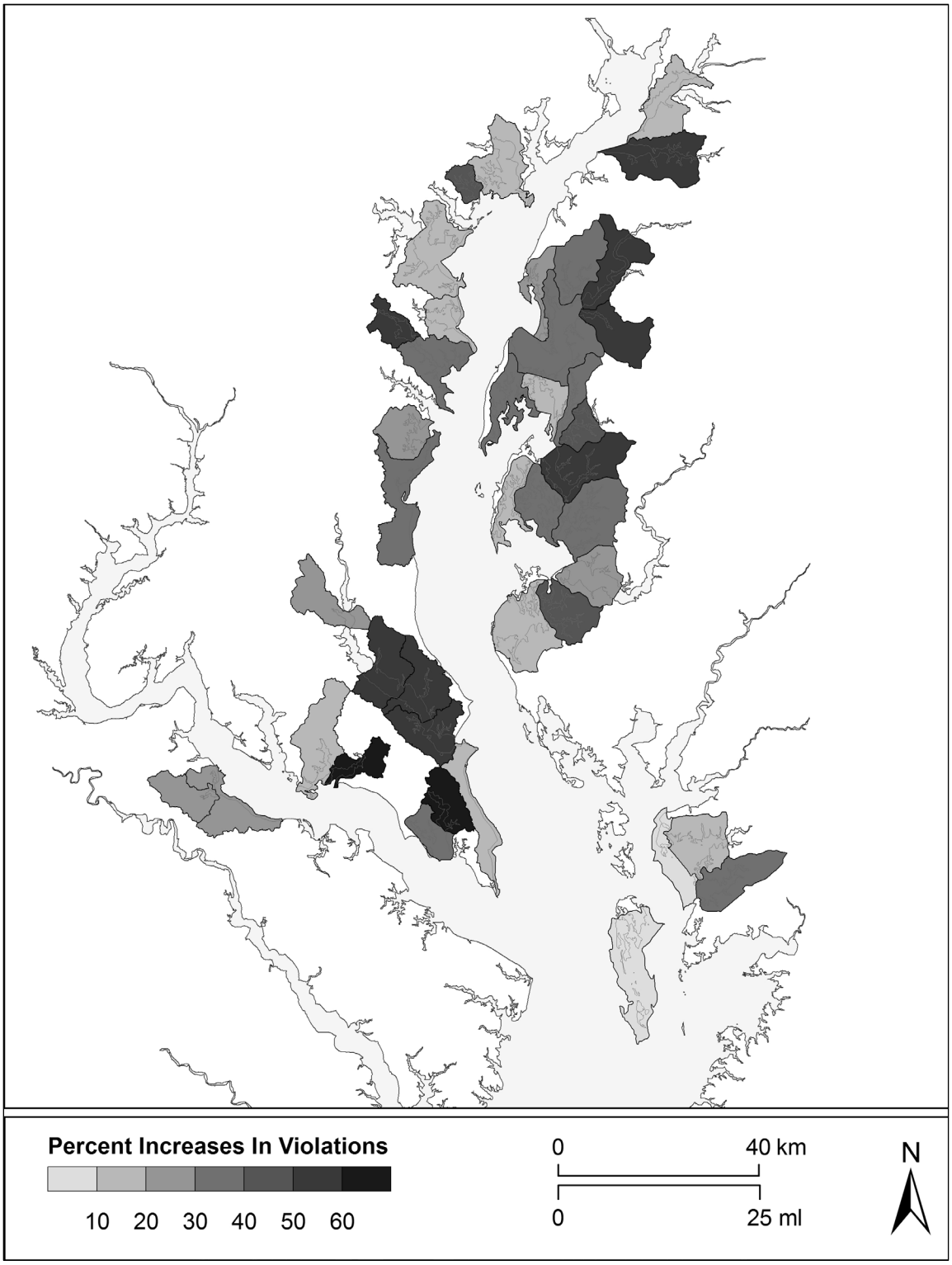

Figure 3.7. Map of those watersheds classified as responsive to rain (based on decision points in Figure 2), showing the percent increase in fecal coliform densities exceeding the FDA 49MPN/100mL criterion after a rain event of ≥1 inches of rain (Rain2Day) from the percent occurring when rainfall was < 1 inch.

Table 3.2. Conditionally approved areas designated by MDE as of June 2016 and the corresponding assessment by the frequency analysis in this study.

Conditionally Approved Area Name	Larger Waterbody	USGS HUC12 Watershed Number	HUC12 Name	Frequency Analysis Assessment
Bolingbroke Creek	Choptank River	020600050502	Bolingbroke Creek-Choptank River	Already in violation
Chester River	Chester River	020600020408	Middle Chester River	Responsive
Langford Creek	Chester River	020600020410	Langford Creek	Responsive
Greenwood Creek	Eastern Bay	020600020606	Prospect Bay-Eastern Bay	Responsive
Cox and Thompson Creeks	Eastern Bay	020600020607	Cox Creek-Eastern Bay	Responsive
Crab Alley Bay and Crab Alley Creek	Eastern Bay	020600020607	Cox Creek-Eastern Bay	Responsive
Rhode River	Rhode River	020600040401	Rhode River-West River	Responsive
West River	West River	020600040401	Rhode River-West River	Responsive
Herring Bay	Herring Bay	020600040402	Tracys Creek-Herring Bay	Responsive
Tanner Creek	Tanner Creek	020600040404	Saint Jerome Creek-Chesapeake Bay	Responsive
Little Choptank	Little Choptank River	020600050401	Fishing Creek-Little Choptank River	Responsive
Phillips and Beckwith Creeks	Little Choptank River	020600050401	Fishing Creek-Little Choptank River	Responsive
Fishing Creek	Little Choptank River	020600050401	Fishing Creek-Little Choptank River	Responsive
Brooks Creek	Little Choptank River	020600050402	Slaughter Creek-Little Choptank River	Responsive
Slaughter Creek	Little Choptank River	020600050402	Slaughter Creek-Little Choptank River	Responsive
La Trappe Creek	Choptank River	020600050503	LaTrappe Creek-Choptank River	Responsive
Island Creek	Choptank River	020600050504	Tred Avon River-Choptank River	Responsive
Broad Creek	Broad Creek	020600050505	Broad Creek-Choptank River	Responsive
San Domingo Creek	Broad Creek	020600050505	Broad Creek-Choptank River	Responsive
Irish Creek	Broad Creek	020600050505	Broad Creek-Choptank River	Responsive
Upper Patuxent River	Upper Patuxent River	020600060507	Swanson Creek-Upper Patuxent	Responsive
Cuckold, Nats and Mill Creeks	Patuxent River	020600060604	Mill Creek-Patuxent River	Responsive
Breton Bay	Potomac River	020700110703	Glebe Run-Breton Bay	Responsive
St. Mary's River A	St. Mary's River	020700110902	Eastern Branch-Saint Marys River	Responsive
St. Mary's River B	St. Mary's River	020700110902	Eastern Branch-Saint Marys River	Responsive
Schoolhouse Brook	St. Mary's River	020700110903	Saint George Creek-Saint Marys River	Responsive
St. George Creek	St. Mary's River	020700110903	Saint George Creek-Saint Marys River	Responsive
Big Annemessex River	Big Annemessex River	020801100405	Big Annemessex River	Responsive
Cuckhold Creek	Potomac River	020700110306	Piccowaxen Creek	Too few events
Herring Creek	Potomac River	020700110704	Poplar Hill Creek	Too few events
Piney Point Creek	Potomac River	020700110704	Poplar Hill Creek	Too few events
Jutland Creek	Smith Creek	020700110904	Jutland Creek	Too few events
Marumsco Creek	Pocomoke Sound	020801110501	Marumsco Creek-Pocomoke Sound	Too few events
Ape Hole Creek	Pocomoke Sound	020801110502	East Creek-Pocomoke Sound	Too few events

Table 3.3. Pearson correlation coefficients and p-values associated with comparison PDiff to watershed characteristics. Significant correlations ($p < 0.05$) are in bold-face type.

Variable	R^2	p-value
Cultivated Crops	**0.370**	**0.017**
Pasture Hay	**0.339**	**0.030**
Deciduous Forest	**0.316**	**0.044**
Palustrine Aquatic Bed	0.281	0.075
Developed - Open Space	0.234	0.140
Scrub Shrub	0.218	0.171
Type C Soils	0.208	0.192
Mixed Forest	0.199	0.213
Developed - High Intensity	0.190	0.234
Type B Soils	0.185	0.248
Developed - Medium Intensity	0.171	0.286
Type A Soils	0.166	0.298
Developed - Low Intensity	0.145	0.367
Developed - total	0.144	0.368
Type B/D Soils	0.096	0.552
Grassland Herbaceous	0.070	0.665
Evergreen Forest	0.060	0.708
Palustrine Scrub Shrub Wetland	0.053	0.743
Impervious Surfaces	0.005	0.979
Palustrine Forested Wetland	-0.012	0.941
Total Land Area	-0.087	0.589
Bare Land	-0.087	0.589
Palustrine Emergent Wetland	-0.096	0.550
Type C/D Soils	-0.113	0.483
Type D Soils	-0.116	0.470
Estuarine Scrub Shrub Wetland	-0.167	0.298
Type A/D Soils	-0.193	0.227
Estuarine Forested Wetland	-0.295	0.061
Estuarine Emergent Wetland	**-0.369**	**0.018**
Unconsolidated Shore	**-0.447**	**0.003**
Open Water	**-0.512**	**0.001**

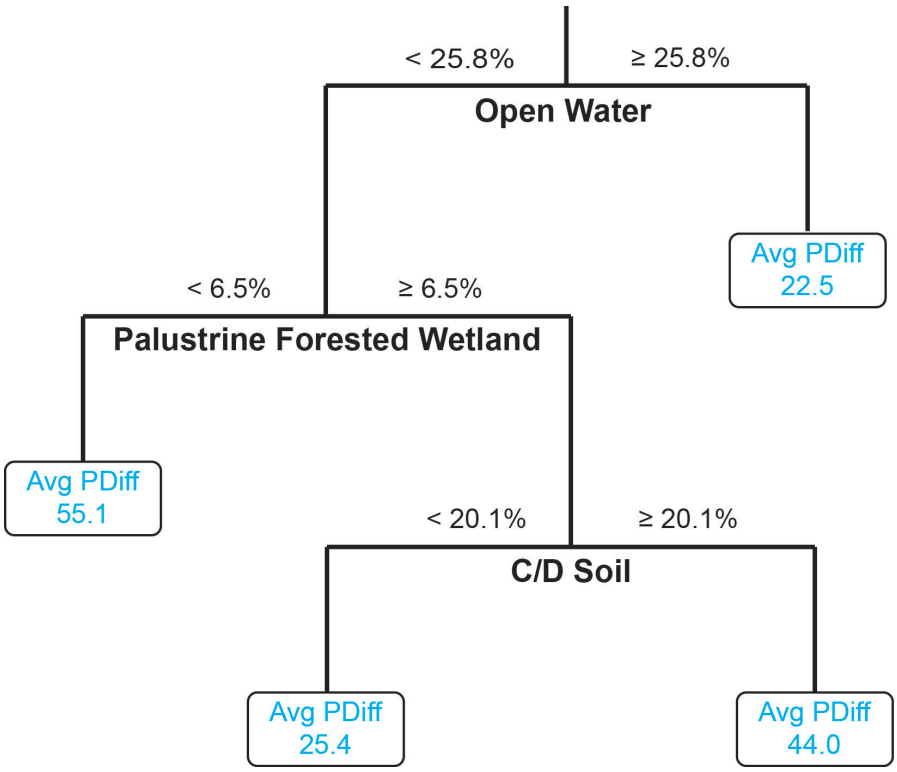

Figure 3.8. Classification and regression tree results for the comparison of PDiff to watershed characteristics. PDiff is the difference in percent of samples exceeding the 49MPN/100mL criterion (FDA 2015) when < 1 inch of rain has fallen and when ≥ 1 inch of rain has fallen.

Table 3.4. Bay-wide logistic regression results for four different models. The first three models include only measurements of rainfall (Rain2Day and 3wkRainfall) and distance from nearest shoreline (Distance). The fourth model also includes average air temperature (Air Temp).

Model Parameters	% Concordant	c[1]	AIC[2]
Intercept 2DayRain	60.2	0.673	12710
Intercept 2DayRain Distance	75.6	0.756	12180
Intercept 2DayRain Distance 3wkRainfall	76.1	0.761	12110
Intercept 2DayRain Distance 3wkRainfall Average Daily Temp	76.5	0.76	12100

[1]-Area under the Receiver Operator Curve
[2]-Akaike Information Criterion

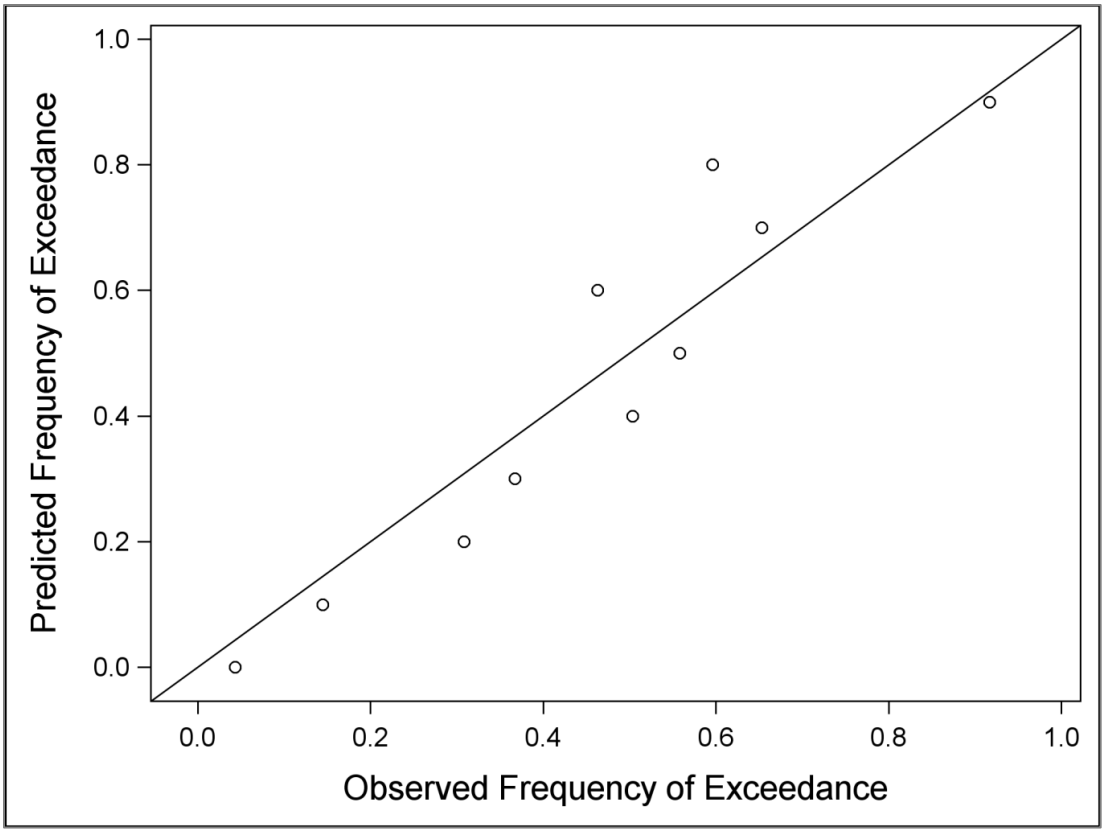

Figure 3.9. Comparison of predicted probability of exceedance (binned) to the proportion of observations in each bin that exceeded the criteria. Adjusted R^2 for the linear relationship was 0.88. Predicted probabilities are for all responsive watersheds based on logisitc model with RainEvent, antecedent rainfall, and the average distance of stations to shore.

Table 3.5. Bay-wide model parameter estimates and results. The estimated rainfall threshold value is the amount of rainfall (Rain2day) in inches where there is a 50% probability of exceeding the 49MPN/100mL criterion. It uses the average conditions that were observed for all other variables besides Rain2Day.

Variables	β[1]	Odds Ratios	False Positives	False Negatives	RainFall Threshold[2]	
					Threshold	CL
Intercept	-2.587		38.0	6.3	2.16	(2.07, 2.25)
2DayRain	1.525	4.59				
Distance	-0.004	1.00				
3wkRainfall	0.105	1.11				

[1]-Parameter Estimates; [2]- The estimated rainfall threshold value and 95% confidence limits (CL) are based on the amount of rainfall (Rain2day) in inches where there is a 50% probability of exceeding the 49MPN/100mL criterion, using the average conditions that were observed for all other variables besides Rain2Day.

Table 3.6. Parameter bias from bootstrap analysis, using 1000 iterations.

Variable	Original β	Bias	Standard Error
Intercept	-2.5793	-0.0032	0.0569
Rain2Day	1.5370	0.0024	0.0477
Distance	-0.0037	0.0000	0.0002
3wkRain	0.1042	0.0001	0.0111

Table 3.7. Logistic regression results for responsive watersheds with significant relationships to Rain2Day, concordance above 75%, and rainfall estimates having 95% confidence limits.

Watershed	# Samples	# RainEvents	Average Rainfall Threshold[1]	Odds Ratio Rain2Day	Concordant (%)	False Positive (%)	False Negative (%)	Distance p-value	Antecedent Rainfall p-value
Lower Wye East River	691	5	1.8 (1.4, 2.5)	5.9	76.1	40	6.5	> 0.050	< 0.001
Round Bay-Severn River	463	8	1.8 (1.4, 2.4)	6.2	84.9	17.6	6.7	< 0.001	0.049
Whitehall Creek-Severn River-Chesapeake Bay	1447	9	2.3 (2.0, 2.7)	5.6	84.8	31	4.5	< 0.001	0.002
Battle Creek-Patuxent River	507	3	1.6 (1.3, 2.0)	16.9	82.5	37.5	3	0.056	0.004
Saint Leonard Creek-Patuxent River	1149	10	1.7 (1.4, 2.0)	5.2	80.2	36.4	8.8	< 0.001	> 0.050
Mill Creek-Patuxent River	944	7	1.7 (1.4, 2.2)	6.2	77.7	31.8	8.4	< 0.001	> 0.050
Glebe Run-Breton Bay	719	9	1.2 (1.0, 1.4)	16.4	84.7	15	8.1	< 0.001	< 0.001
Eastern Branch-Saint Marys River	483	6	1.4 (1.2, 1.7)	16.8	87.1	29.4	3.3	> 0.050	> 0.050
Saint George Creek-Saint Marys River	1182	6	2.4 (2.1, 3.0)	4.9	86.0	52	5.3	< 0.001	> 0.050
Big Annemessex River	736	14	1.7 (1.5, 2.1)	10.5	85.7	39.1	4.5	0.003	> 0.050

[1]The estimated rainfall threshold value is the amount of rainfall (Rain2day) in inches where there is a 50% probability of exceeding the 49MPN/100mL criterion, using the average conditions that were observed for all other variables besides Rain2Day

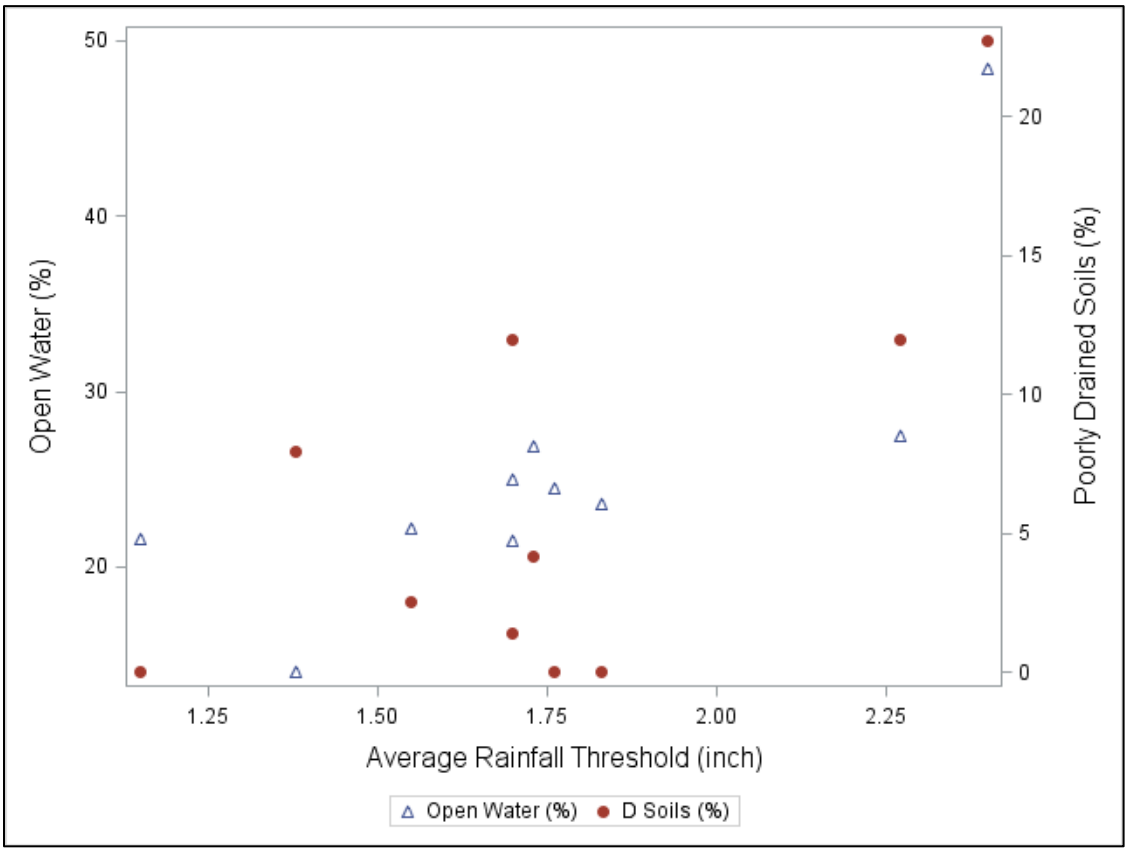

Figure 3.10. Comparison by watershed of the percent of open water (triangles) and soils with high runoff potential (circles; soil type D) compared to the average rainfall threshold.

Table 3.8. Comparison of watershed-scale and conditionally approved area specific logistic regression results for St. Mary's River.

Stations	# Samples	# RainEvents	Rainfall Threshold[1]	Odds Ratio Rain2Day	Concordant (%)	False Positive (%)	False Negative (%)	Distance p-value	3wkRainfall p-value
All Stations in Watershed	483	6	1.4 (1.2, 1.7)	16.6	82.4	29.4	3.3	>0.05	>0.05
2 Stations in Conditionally Approved Area	153	6	1.3 (1.0, 1.7)	50.9	99.1	18.2	1.4	>0.05	>0.05

[1]The estimated rainfall threshold value is the amount of rainfall (Rain2day) in inches where there is a 50% probability of exceeding the 49MPN/100mL criterion, using the average conditions that were observed for all other variables besides Rain2Day

Chapter 4: The Co-occurrence of Fecal Indicator Bacteria and Potential Pathogen Groups at a Shellfish Harvest Area

Abstract

Prevention of human illness from bacterial pathogens in shellfish harvesting areas requires better understanding of pathogen occurrence. This study investigated the co-occurrence of pathogen-containing groups and fecal indicators in 40 water samples from a shellfish harvest area in the Chesapeake Bay, USA, over 6 months in 2014. Sample treatment involved sequencing of 16S rRNA gene amplicons, culture-based methods for fecal indicators and quantitative PCR for select *Vibrio* species. Addition of *Thermus thermophilus* genomes to the DNA samples provided an experimental method of estimating bacterial 16S gene densities. Sequencing of 16S rRNA amplicons detected fecal coliforms and potential human pathogens in most samples and *Vibrio* in all samples. Calculated 16S rRNA gene densities ranged from $5.42x10^5$ to $5.08x10^6$ per sample, but this method overestimated Enterobacteriaceae and underestimated *Vibrio* densities when compared to culture based and qPCR techniques, respectively. A distinct and temporary change in bacterial community composition and fecal coliform and pathogen densities occurred after a 1.3 inch rain event, but not after smaller rain events. Enterobacteriaceae and cultured fecal coliforms co-occurred, due primarily to spikes in densities following this rain event. Bacteriodales and Lachnospiracaea, potential fecal indicators, also co-occurred with this spike in culture-based fecal coliforms.

Introduction

Monitoring for indicator microorganisms, such as fecal coliforms, *Escherischia coli*, and *Enterococcus* spp., at shellfish beds and swimming beaches has been conducted since the early 1900's to assess the likelihood of encountering human pathogens (EPA 1986, Ashbolt

et al. 2001). However, the presence and density of these bacteria are not always correlated with the presence of human pathogens (Noble and Fuhrman 2001) and their strength in risk assessments depends on cell density (EPA 1986, FDA 2013) and the types of pathogens present (Wade et al. 2003). More reliable indicators of fecal pollution and human pathogens are being sought and several other bacterial groups have been proposed, including several members of the Order Bacteriodales (Bernhard and Field 2000, Walters et al. 2007) and the Family Lachnospiracaea (Newton et al. 2011).

Several studies have explored the relative abundance of pathogens and their co-occurrence with indicator bacteria in bivalves or their surrounding waters. For example, Walters et al. (2007) found that *Salmonella*, *Campylobacter*, and *E. coli* O157:H7 did not commonly co-occur in water samples, suggesting different sources or life-histories for these pathogens. While Hood et al. (1983) found that high levels of fecal coliforms and *E. coli* were about equally predictive of the presence of *Salmonella* in oyster tissue, a single indicator for the diverse types of pathogens found at shellfish beds or swimming beaches may not be feasible (Wade et al. 2003).

In the Chesapeake Bay, monitoring for indicator bacteria over natural shellfish beds occurs year-round. While the harvest of oysters from these natural beds is currently prohibited from late spring to early fall, harvest from aquaculture is permissible year-round, with additional post-harvest handling requirements during summer months (MDDNR 2015). In Maryland, monitoring for fecal indicators in waters over natural shellfish beds involves measurement of fecal coliform densities twice a month. Similar routine monitoring is conducted during the summer by local health agencies in recreational areas of the Chesapeake Bay used for boating and swimming.

Despite these monitoring efforts, little is known about the bacterial communities that co-occur with fecal indicator bacteria. Bacterial communities in natural waters tend to be very diverse (Kan, 2006) and have complex and redundant functional groups (Comte and del Giorgio 2010). Until recently, studies of bacterial communities in shellfish harvest waters have relied on culture-based methods that exclude a large number of bacteria and require focused analyses of particular bacterial groups. A study in the Santa Anna River, California using 16S rRNA gene amplicon pyrosequencing detected higher percentages of genera that contain human pathogens in areas of urban runoff and agriculture (Ibekwe et al. 2013). Likewise, a recent study of bacterial community variability at a shellfish harvest area in Spain, using denaturing gradient gel electrophoresis (DGGE), noted seasonal shifts in the bacterial community and the presence of genera containing human pathogens (Pereira et al. 2015). Similar studies, focused on more highly resolved temporal changes in bacterial communities and potential pathogenic members at shellfish harvest areas would help inform relative human health risks and decisions about restricting shellfish harvest.

The objectives of this study were to improve understanding of bacterial pathogen and fecal indicator bacteria occurrence in shellfish harvest waters by characterizing the bacterial community composition over time at a monitoring station in the Chesapeake Bay, and to identify members of the bacterial community that co-occur with indicator bacteria and *Vibrio* species. The specific goals were to 1) assess the occurrence of fecal indicator bacteria and *Vibrio* bacteria at the monitoring station over time; 2) assess the composition of the bacterial community at the station over time and at varying depths; 3) identify members of the bacterial community that fall into taxonomic groups known to contain human pathogens; and 4) assess the relationships of the indicator bacteria, *Vibrio* bacteria and bacterial community members with environmental variables such as precipitation, water temperature, and salinity.

Materials and Methods

Water Collection

The sampling location was a long-term Maryland Department of the Environment shellfish harvest monitoring station near the mouth of Town Creek (Figure 4.1), surrounded by the Town of Oxford (Maryland, USA) to the west and agricultural fields and low density residential areas to the east. The Town of Oxford manages a secondary wastewater treatment facility that releases approximately 125,000 gal/day of effluent (Scott Delude, Public Works Director, personal communication). Because of proximity to the treatment plant, shellfish harvest and aquaculture are currently forbidden in Town Creek, though routine testing of treatment plant effluent shows fecal bacteria levels below the detection limit (<1.8 MPN/100mL (Most Probable Number), (standard method 9221E; (APHA 1998)). Total depth at the sampling location is approximately 3.0 m, tidal amplitude at this location averages about 0.6 m (NOAA 2015), and salinity averages 10 ppt (current study).

Forty samples were collected on 14 days between April 14 and September 3, 2014, from both surface and bottom waters, including replicate samples on most sampling days. Water was collected in sterile polypropylene bottles (500 mL) by hand from surface waters and by Van Dorn sampler (Wildco, Inc) from bottom waters (0.25 m above bottom), placed on ice, and processed within 30 minutes of collection. Water physicochemical measurements were collected with a YSI 6600 datasonde (YSI, Inc., Yellow Springs, OH). Precipitation was measured with a tipping bucket rain gauge (Onset, Inc.) stationed approximately 1.6 km from the water collection site (Figure 4.1).

Recovery of DNA by water filtration was performed using established methods (Fortunato and Crump 2011). Between 250 and 300 mL of water was filtered through a 0.2 µm Sterivex-GP filter (EMD Millipore, Darmstadt, Germany), the volume filtered was

recorded, and 1mL of filter-sterilized DNA extraction buffer (DEB; 0.1M Tris-HCL (pH 8), 0.1M Na-EDTA (pH 8), 1.5M NaCl, 5% Cetyltrimethyl ammonium bromide) was added to each filter assembly. The filters were then stored at -80°C, and later extracted according to published methods (Crump et al. 2003), including several freeze/thaw cycles and final DNA isolation by isopropanol precipitation. During extraction, approximately 60 million copies of the *Thermo thermophilus* genome (i.e., 250 ng *T. thermophilus* DNA) were added to each sample, in order to approximate a final ratio of 1.00 to 0.04 bacterial cells to *Thermus thermophilus* genomes, assuming $5x10^6$ bacterial cells/mL in each water sample.

DNA sequencing

DNA library preparation and sequencing followed Fadrosh et al. (2014). A segment of the 16S gene including the V3 and V4 hypervariable regions was amplified using PCR. Amplification was confirmed by running several samples on an electrophoresis gel and checking for bands with an approximate length of 469 base pairs. Universal 16S primers (319F and 806R) were complemented with a heterogeneity spacer sequence of variable length and an Illumina (Illumina, Inc., San Diego, CA) index sequence. Sequencing was performed on the MiSeq platform (Illumina, Inc., San Diego, California, USA), which produced an average of 131,000 reads per sample. Processing of reads was conducted using Mothur (Schloss et al. 2009, Mothur 2015) and following Kozich et al. (2013). All reads were archived in the Sequence Read Archive (accession #SRP073436).

A phylotypic approach (Mothur command 'phylotype') was used for clustering sequences into operational taxonomic units (OTUs), with a cutpoint of 97% similarity for classification (Mothur 2015). OTU representative sequences were classified with the Silva database, version 1.19 (Silva 2015), and OTUs not classified as Eubacteria (i.e. Eukaryota, mitochondria, chloroplasts, or unclassified) were deleted. The per liter abundance of 16S copies per OTU in the original sample population was calculated from the number of reads of

the internal control, *T. thermophilus*. A correction factor was applied to the number of reads in each sample based on the ratio of 16S genomes added to original samples compared to the number of reads for *T. thermophilus* after sequence processing (Eq. 1). Although this approach does not account for gene copy number per cell or differential amplification success of different bacteria, it provided a rough estimate of the number of 16S genes present in the samples. For the taxa Enterobacteriaceae and *Vibrio*, where multiple isolates have been evaluated for 16S rRNA gene copy (Stoddard et al. 2015), the number of cells per milliliter was estimated using an approximation of 7 and 10 copies of the 16S rRNA gene per cell, respectively. These estimates for gene copy number per cell were based on examination of all entries for both phylogenetic groups in the Ribosomal RNA Database (https://rrndb.umms.med.umich.edu/) (Stoddard et al. 2015).

(Eq. 1) ((*Tt* genomes added) / (2 * *Tt* seqs)) * (non- *Tt* seqs) * (1 / volume filtered) = 16S genes / mL

Where *Tt* = *Thermus thermophilus* and seqs = sequences

Fecal coliform and Vibrio density estimations

Fecal coliform densities were estimated using the standard three-tube culture method (APHA 1998). Test tubes containing inverted gas-tubes and various concentrations of A1 broth were inoculated with aliquots of sample water, incubated in a 35°C water bath for 3 hr (±30 min), and then in a 44.5°C water bath for 21 hr (±2 hr). Tubes showing turbidity and gas production were counted as positive for growth of fecal coliforms. Fecal coliforms were measured for all samples, except those from August 22 when an equipment failure

compromised the cultures. An estimate of the number of fecal coliform bacteria in 100mL of water was made by dividing the Most Probable Number (MPN) by 100.

Estimates of *V. vulnificus* (*Vv*) and *V. parahaemolyticus* (*Vp*) concentrations, independent of the 16S community analysis, were conducted using quantitative PCR (qPCR) (Jacobs et al. 2014), with primers and probes adapted from Panicker and Bej (2005) and Nordstrom et al. (2007), respectively (Table 4.1). A unique internal control was incorporated simultaneously into the qPCR assay to test for the presence and influence of inhibitors (Nordstrom et al. 2007). Samples of known *Vibrio* density, developed previously, were used for the standard curve and were based on a slightly different DNA extraction protocol (Jacobs et al. 2014). A comparison of 21 environmental samples indicated the extraction protocol used in this study produced 2.1 times the DNA yield as the protocol used for the *Vibrio* standard curve (data not shown), and estimates of *Vibrio* density by qPCR were adjusted accordingly.

Detection of pathogen groups and other specific genera

Bacterial genera that largely contain human pathogens or members endemic to the human gut were identified. These taxonomic groups included *Aeromonas*, *Arcobacter*, *Campylobacter*, *Legionella*, *Helicobacter*, *Toxoplasma,* and members of the Enterobacteriaceae (e.g. *Yersinia*, *Klebsiella*, *Escherichia*, *Shigella*). In addition, detection of genera that contain common foodborne (Scallan et al. 2011) and waterborne (Pond 2005) human pathogens, but also include non-pathogenic members, was assessed. These genera included *Clostridium, Leptospira, Listeria, Mycobacterium, Staphylococcus*, and *Vibrio*.

Data Analyses

Prior to diversity calculations, the number of sequences per sample was rarefied to 17,500. Alpha diversity measurements (Shannon Diversity and the Chao1 estimate of taxon richness) were calculated for each sample using R (The R Foundation for Statistical Computing, version 3.1.2). Beta diversity was compared between samples using Bray-Curtis Similarity matrices and plotted with non-metric multi-dimensional scaling (nMDS) using R and by cluster analysis using PRIMER for Windows (PRIMER-E Ltd, version 6). Comparison of community composition to environmental variables was conducted via distance-based linear modeling (DistLM) using PRIMER. All other comparisons involved estimated 16S gene densities. For rare taxa, with 16S genes detected in only one or two samples, presence versus absence were compared between 16S phylotypes, fecal coliforms (culture method), *V. vulnificus* and *V. parahaemolyticus* (qPCR) estimates. For more commonly occurring potential pathogen groups (detected in 3 or more samples), estimated abundances of bacterial genera (16S sequences), fecal coliforms (culture method), *V. vulnificus* and *V. parahaemolyticus* (qPCR), and environmental factors (e.g. water clarity, rainfall) were compared using Spearman correlation analysis in SAS (SAS Institute, Inc; version 9.4). All statistical tests were considered significant at $p<0.05$.

Results

Environmental conditions

Conditions were similar between surface and bottom waters, with a small decrease in dissolved oxygen in bottom waters during June. Changes in water conditions generally followed seasonal trends (Table 4.2). Mean water temperatures increased from 13°C in April to 27°C in June, and remained about the same level into September. Salinity at the sampling

station ranged from 9.0 to 11.9 ppt, with the lowest values in mid-summer. Water clarity, as measured by Secchi depth, decreased in June and remained low thereafter. Approximately 15 precipitation events occurred during the study, ranging from a very small amount of rain to 1.3 inches of rain, with several sampling events taking place during or within 24 hours following rainfall (Figure 4.2). May 1 was the only sampling date with greater than 1 inch of rain occurring in the previous 24 hours, though three other sampling dates had 0.5 to 0.85 inches of rain fall within the preceding 24 hour.

Internal control recovery and estimation of 16S gene densities

After deleting unclassified sequences and those classified as Eukaryota, mitochondria, and chloroplasts, the 16S dataset contained 5,190,588 reads classified to 1,429 OTUs. Despite notable variability in *T. thermophilus* sequence numbers between samples, all but two samples were estimated as having between $5.4x10^5$ and $5.1x10^6$ 16S genes/mL (Figure 4.3). These two samples, Jun13S and Jul09B2, contained very few *T. thermophilus* sequences, indicating very poor recovery of the internal control and resulting in very high estimated 16S gene densities, and thus were eliminated from further analyses. One other sample, Aug22B2, contained a large number of *T. thermophilus* gene sequences, though the overall number of sequences was relatively low, leading to an estimated 16S gene density within the range of the other samples.

Bacterial community composition

Using the phylotypic approach to sequence classification (Mothur 2015), the samples contained an average of 541 operational taxonomic units (OTUs). The bacterial community at this station was dominated by Actinobacteria, Cyanobacteria and Proteobacteria (Figure 4.4). Temporal trends were detected for several bacteria phyla and classes, highlighted by increases in the percent of Cyanobacteria from late May through August and decreases in the

percent of Actinobacteria for the same period. Replicate samples were generally similar in their composition of major taxonomic groups, except for surface water samples collected July 9, with Jul09S1 having a notably higher percentage of Cyanobacteria than Jul09S2. Shannon diversity ranged from 3.4 to 4.2 with a trend towards increased diversity over time, except for the Jul09S1 sample (Figure 4.5, panel A). The Chao1 estimate of taxonomic richness was fairly stable through time with a notable increase for the May01S sample and slightly higher numbers of OTUs in bottom waters than in surface waters for August samples (Figure 4.5, Panel B).

Comparison of bacterial community composition between samples showed a separation of samples in April and early May from all other samples, with May 1 being the most different (Figure 4.6). Trends in water temperature, turbidity, and salinity were significantly related to changes in bacterial community composition between samples, though they only explained a very small amount (between 29 and 14% respectively) of the overall variability in bacterial community composition (Table 4.3). The amount of rain that fell in the last 24 hours was significantly related to community composition for the overall study but also explained very little of the variability (Table 4.3). Other variables, including wind speed, wind direction, and tidal stage, were not significantly related to differences in bacterial community composition between samples.

Detection of select genera

A number of bacterial genera that include human pathogens were detected in the 16S sequences. For example, *Arcobacter* (Epsilonproteobacteria) and *Legionella* (Gammaproteobacteria) were detected in all samples, while *Aeromona* (Gammaproteobacteria) was detected in most samples. Bacteria were also detected from 11 genera of the family Enterobacteriaceae, which includes most genera defined as fecal coliform bacteria. *Enterococcus* spp., a common gut bacteria and existing fecal indicator

bacteria, were detected in two samples. Among the bacteria genera selected due to the pathogenicity of a few members, all 38 samples contained members of the *Clostridium* (Firmicutes), and *Vibrio*, while 37 samples contained *Mycobacterium* (Table 4.4). The 16S gene densities for a number of individual Enterobacteriaceae groups as well as *Clostridium*, *Aeromonas*, and *Arcobacter*, were notably elevated or were exclusively detected on May 1. Several bacterial genera of commonly of concern in shellfish harvest waters not detected in the 38 samples analyzed included *Escherichia, Salmonella*, *Listeria*, and *Helicobacter*.

Most members of the potential pathogen groups that occurred in 3 or more samples were not correlated (p>0.05) to general trends in environmental factors (Table 4.5). Except for *Aeromonas*, these more commonly detected genera were not well correlated to rainfall, despite the 16S gene being detected at elevated concentrations for several genera following the 1.3 inch rainfall event. Several genera (*Clostridium, Arcobacter, and Legionella*) were weakly correlated to water temperature, though some trends were negative and others positive.

Comparisons of 16S sequencing to fecal coliform bacteria and specific Vibrio species

Fecal coliforms were detected at low densities in most samples using standard culture methods (Figure 4.7, Panel A). Fecal coliforms occurred on 11 of the sampling dates and in both surface and bottom waters, with all but one sample containing less than 0.5 cells per mL (estimated from the MPN). A spike in fecal coliform density of 4.5 cells/mL occurred on May 1, concurrent with the largest rain event of 1.3 inches in the previous 24 hours. In contrast, samples collected on May 28, June 13 and July 24 contained relatively low fecal coliform densities, despite occurring after slightly smaller rain events. Fecal coliforms were only correlated to *Staphylococcus*, among all potential pathogen groups detected in 3 or more samples.

Vibrio vulnificus was detected by qPCR assay in most samples after May 13 (Figure 4.7, Panel B), with greater than 400 cells per mL in the May 28 samples, which exceeded the estimated density of total *Vibrio* (after accounting for an average of ten 16S gene copies per *Vibrio* genome (rrnDB 2015) for 24 of the 38 samples used for density analysis. *Vibrio vulnificus* densities declined in mid-summer, with an increase in bottom water samples on August 6. *Vibrio parahaemolyticus* occurred much less often with only 4 samples containing detectable levels and a maximum density of 4.7 cells/mL in one of the August 6 samples. *Vibrio vulnificus* and *V. parahaemolyticus* densities did not correlate to fecal coliforms or potential pathogen groups, including total *Vibrio* (Table 4.5).

Detection of Potential and Existing Indicators

Several bacterial genera from families that have been proposed as indicators of fecal pollution had a significant spike in density on May 1 (Table 4.6), concurrent with a number of potential pathogen groups. None of these indicators, including the existing indicator genus *Enterococcus* (EPA 1986), were correlated to the potential pathogen genera that were detected in 3 or more samples or had elevated densities on May 1.

Discussion

The diversity and abundance of 16S rRNA genes at the sampling location provided indications of bacterial diversity and the relative occurrence of both endemic and transitory pathogen groups. Sequencing also detected and characterized an important shift in the bacterial community following a relatively large rain event, and informed the relative utility of indicator bacteria in serving as proxies for human pathogens. Although this study was unable to determine the 'species' or virulence factors of the bacteria classified, it provided a much more detailed characterization of the bacterial community and the occurrence of

bacterial genera containing pathogens than could have been determined by conventional culture-based methods or PCR assays alone.

The bacterial community at the shellfish monitoring station examined in this study was dense and diverse, and showed changes in composition over time that were related to environmental factors. Seasonal differences between community composition on the scale of phyla and classes, such as those detected in this study, have been noted in other studies of estuarine bacterial communities (Crump et al. 2003, Fortunato and Crump 2011, Pereira et al. 2015), although spatial variability (Fortunato et al. 2012) and changes in organic matter composition (Crump et al. 2003) have also been linked to changes in bacterial community composition. Kan et al. (2006) found that temporal changes in bacterial communities in the Chesapeake Bay were greater than spatial patterns, with seasonal differences in environmental variables such as temperature playing a large role. The only test of spatial variability in this study, water column depth, did show slightly higher OTU richness in bottom waters.

The greatest change in community composition at the monitoring station occurred after a rainfall of approximately 1.3 inches in the preceding 24 hours. The number of different bacterial groups also was highest on this day, suggesting the introduction of allochthonous (i.e. non-endemic) bacteria to the community. This rain event also suggests that runoff from land can play a very important role in influencing the relative chance of encountering some pathogen containing groups in nearshore water, but only for relatively short periods of time. It also suggests a nonlinear relationship between precipitation and the bacterial community, with some threshold below 1.3 inches of rain capable of causing a distinct shift in the community.

Trends in the presence of bacterial genera of common concern to shellfish harvest managers also showed signs of both indigenous and non-indigenous sources. The findings suggest that a balance between allochthonous (land-based or benthic) and autochthonous (endemic) bacteria occurs in these waters. In keeping with their association to runoff and land-based sources (Ferguson et al. 1996, Ackerman and Weisberg 2003), several of the pathogen groups, such as members of the Enterobacteriaceae and *Aeromonas*, had notable spikes in 16S rRNA gene abundance or were exclusively detected following the large rain event on May 1. In contrast, the detection of *Vibrio vulnificus* by qPCR in most of the samples is in keeping with their ubiquitous presence in aquatic environments (Wright et al. 1996, Jacobs et al. 2014), particularly in summer months. Several other pathogen genera, such as *Campylobacter* and *Escherichia*, were never detected, despite previous evidence of the presence of some of these groups in Chesapeake Bay waters (Sayler et al. 1975) and shellfish meat (Rawles et al. 1995).

The occurrence of both allochthonous and autochthonous pathogen-containing genera underscores the need for a management strategy capable of assessing potential pathogens in shellfish and recreational waters from multiple sources. For example, the co-occurrence of several Enterobacteriaceae and *Aeromonas* genera with a spike in culture-based fecal coliform densities supports the use of these bacteria as indicators (FDA 2013). Also, the co-occurrence of several genera of Bacteriodales and Lachnospiracaea with several pathogen-containing groups on May 1 encourages further study into the use of members of these families as alternative indicators, and the possibility that they may be particularly indicative of human waste (Bernhard and Field 2000, Newton et al. 2011). However, a different strategy is necessary for those pathogen groups such as *Vibrio vulnificus*. The strategy for assessing the presence and/or abundance of endemic pathogens will continue to rely upon

qPCR detection in shellfish tissues (Brasher et al. 1998) or the use of spatially-explicit predictive models (Jacobs et al. 2014).

Despite challenges in relating individual pathogen species to genus-level data, the approach of using 16S rRNA gene classification to characterize the bacterial community at this station provided a better understanding of these communities than culture-based or targeted qPCR approaches could provide individually. However, for most of the bacterial groups of interest to shellfish managers, species-level and virulence information would be preferred when making management decisions, as many of these bacterial groups contain nonpathogenic members and many pathogenic species have virulent and nonvirulent strains. A reasonable next step might be to use metagenomics to identify virulence and functional genes in such communities. Although virulence factors are complex and subject to evolution and transfer between microbes (Hacker et al. 1997), a study of their occurrence in shellfish waters using metagenomic analysis might improve our assessment of relative risk to human health.

Using an internal control to normalize sequences across samples and to estimate 16S gene densities provided a novel approach allowing for comparison of bacterial densities between the 16S sequencing study and culture-based estimates. Estimates of total bacterial 16S rRNA genes in our samples were within the range of values for density of total bacterial cells in estuarine waters (Ducklow 1982). However, the number of 16S rRNA genes per bacteria varies between 1 and 15 (Kembel et al. 2012), with greater numbers typical of copiotrophic bacteria that maintain large amounts of genetic material in order to quickly take advantage of nutrient introductions (Klappenbach et al. 2000). Although methods for estimating bacterial cell abundance based on 16S rRNA gene copies have been developed (Kembel et al. 2012), there are still large groups of uncultured bacteria for which the number of 16S genes per cell are unknown. This approach may be strengthened in future studies by

making additional estimates of cell densities from staining and direct counting of total bacterial cells, from fluorescence in-situ hybridization using probes for specific groups, or using qPCR methods.

The differences in density estimates between our 16S sequencing and the culture and qPCR-based methods suggest that there are some unquantified biases in these techniques. Enterobacteriaceae (16S sequencing) and fecal coliform (culture) bacteria co-occurred, due primarily to the peak in density on May 1. However, Enterobacteriaceae densities were an order of magnitude higher than the fecal coliform densities, despite adjusting the Enterobacteriaceae values to reflect an approximate gene copy number of 7 16S rRNA genes per cell. The primary cause for this discrepancy is likely that the fecal coliform culture media is only selective for the thermo-tolerant subset of the Enterobacteriaceae. Another potential cause is that Enterobacteriaceae densities based on 16S rRNA data may include bacteria that were non-culturable and therefore not represented in the fecal coliform counts, which is possible given the well-known difficulties in quantitatively culturing bacteria from the environment (Staley and Konopka 1985).

It is more difficult to explain the higher densities of *Vibrio vulnificus* estimated by qPCR than the 16S method found for all *Vibrio* spp. The qPCR method quantifies *V. vulnificus* by comparing sample data to a standard curve prepared with known densities of *V. vulnificus* cells. This method could overestimate *V. vulnificus* if the DNA extraction efficiency for our samples was much higher than that used to prepare standard curves. In fact, the *V. vulnificus* densities were unusually high in May, compared with previous studies in similar waters (Heidelberg et al. 2002, Jacobs et al. 2014), but the internal controls and assay diagnostics for these PCR runs were within normal range (Jacobs et al. 2014) suggesting that the qPCR estimates were accurate. The 16S method could underestimate *Vibrio* density if the actual 16S gene copy number in *Vibrio* cells was much lower than our

estimated 10 per cell. It is also possible that PCR amplification of 16S rRNA genes was not representative of natural abundances. The PCR primers used for 16S rRNA amplification exactly match *T. thermophilus* and all Enterobacteriaceae and *Vibrio* taxa in the Silva database (v1.19), suggesting that primer bias is not an issue. Suzuki et al. (1998) showed that over-amplification of genes in a mixed bacterial community can cause over-representation of rare taxa (e.g. Enterobacteriaceae/fecal coliform) and underrepresentation of more common taxa (e.g., *Vibrio* sp.). With 30 cycles of PCR to prepare 16S rRNA amplicons for DNA sequencing, differential amplification could partially explain the differences in microorganism density between methods. An important next step in using *Thermus thermophilus* or similar internal control approaches for estimating cell abundances would be the development of a quantitative PCR assay to assess the recovery of the internal control after sample extraction.

Several conclusions may be drawn from this study. Significant changes in bacterial community composition, including fecal bacteria and fecal indicators may occur from single rain events with a threshold of precipitation occurring at this sampling location. Water temperature may also play a role with regard to bacterial community composition over time. Also, PCR quantification of *V. vulnificus* and *V. parahaemolyticus* showed that only a subset of samples contained these pathogens and that their densities did not correlate with the ubiquitous *Vibrio* 16S rRNA gene densities. As sequencing methods produce higher numbers of reads per sample, and as read classification improves with the addition of greater amounts of data, assessments of bacterial communities and associated pathogens in shellfish harvest and recreational-use areas will improve our understanding of the ecology of these bacterial groups and how to best manage human health risks.

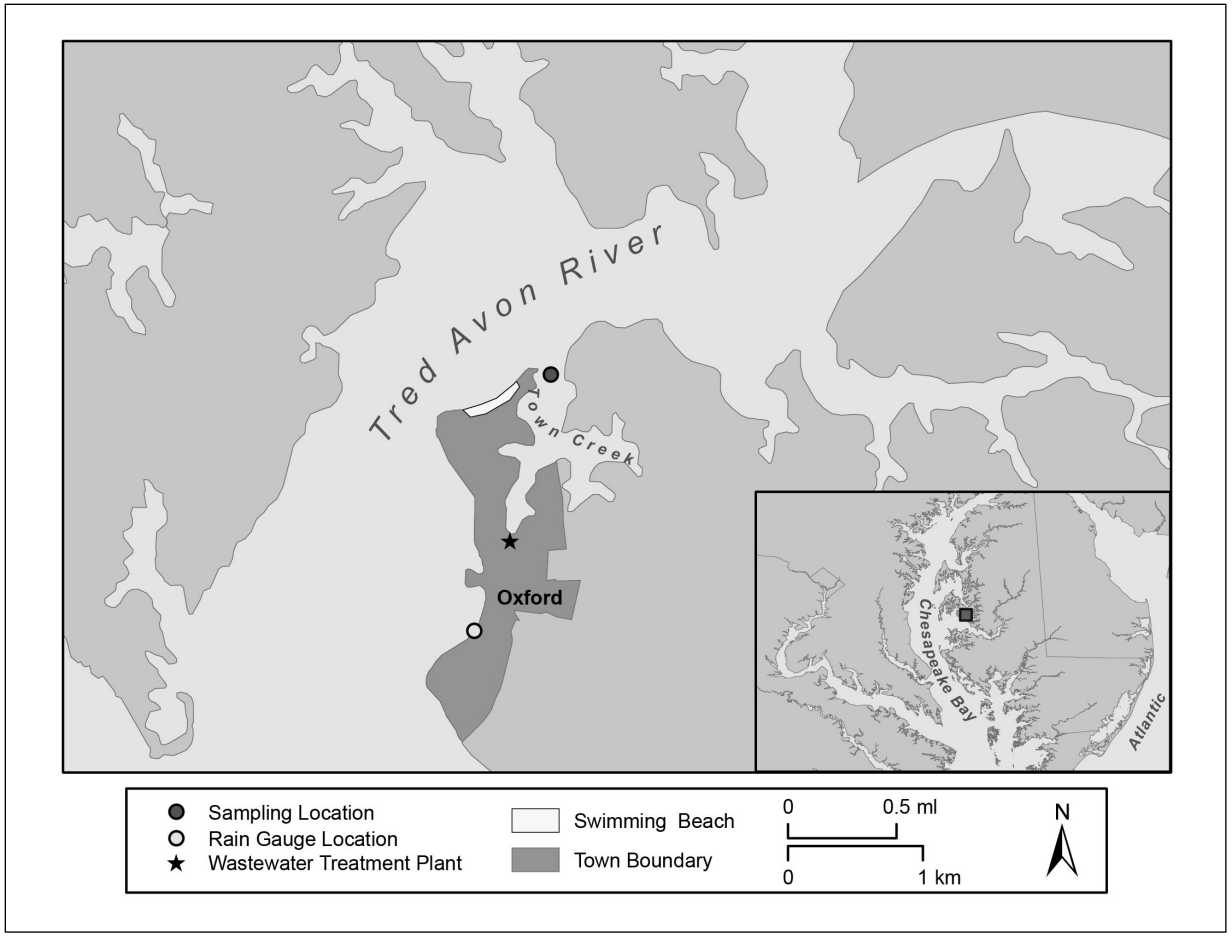

Figure 4.1. Location of sampling station (black dot), rain gauge (open circle), and wastewater treatment plant.

Table 4.1. Real-time PCR primers and probes used to detect and quantify select *Vibrio* species.

Bacteria	**References**	**Gene Target**	**Primer or Probe**	**Sequence (5'-3')**
V. vulnificus	Panicker and Bej (2005), Jacobs et al. (2014)	hemolysin a	*Vvh* forward	TTCCAACTTCA AACCGAACTATGA
			Vvh reverse	TTCCAGTCGATGCGAATACGTTG
			*Vvh*874 probe	56-FAM/AACTATCGTGCACGCTTTGGTACCGT/3IABkFQ/
V. parahaemolyticus	Nordstrom et al. (2007)	thermo-labile hemolysin	*tlh* forward	ACTCAACACAAGAAGAGATCGACAA
			tlh reverse	GATGAGCGGTTGATGTCCAA
			tlh probe	56-FAM/CGCTCGCGTTCACGAAACCGT/BHQ2

Table 4.2. Physical and chemical characteristics of surface and bottom waters at the sampling location.

Month	Depth	Number of observations	Water Temp (°C)		Salinity (ppt)		Dissolved Oxygen (%)		Secchi Depth (m)	
			Mean	Stderr	Mean	Stderr	Mean	Stderr	Mean	Stderr
April	Surface	11	13.10	0.58	11.15	0.12	108.2	2.0	0.94	0.06
	Bottom	10	12.84	0.60	11.19	0.12	106.4	2.2	0.94	0.06
May	Surface	17	20.38	0.68	9.75	0.11	89.8	2.7	0.95	0.04
	Bottom	17	20.01	0.68	9.77	0.12	83.5	2.8	0.95	0.04
June	Surface	15	25.39	0.42	9.16	0.02	83.9	1.5	0.70	0.03
	Bottom	15	25.08	0.43	9.16	0.02	76.4	2.2	0.70	0.03
July	Surface	20	27.32	0.19	9.30	0.02	91.5	1.4	0.58	0.02
	Bottom	20	27.14	0.19	9.30	0.02	86.8	1.3	0.58	0.02
August	Surface	14	26.10	0.15	9.85	0.10	90.6	1.7	0.68	0.03
	Bottom	14	25.97	0.13	9.90	0.13	83.5	1.9	0.68	0.03
September	Surface	1	27.30	.	10.87	.	93.2	.	0.70	.
	Bottom	1	27.22	.	10.87	.	92.8	.	0.70	.

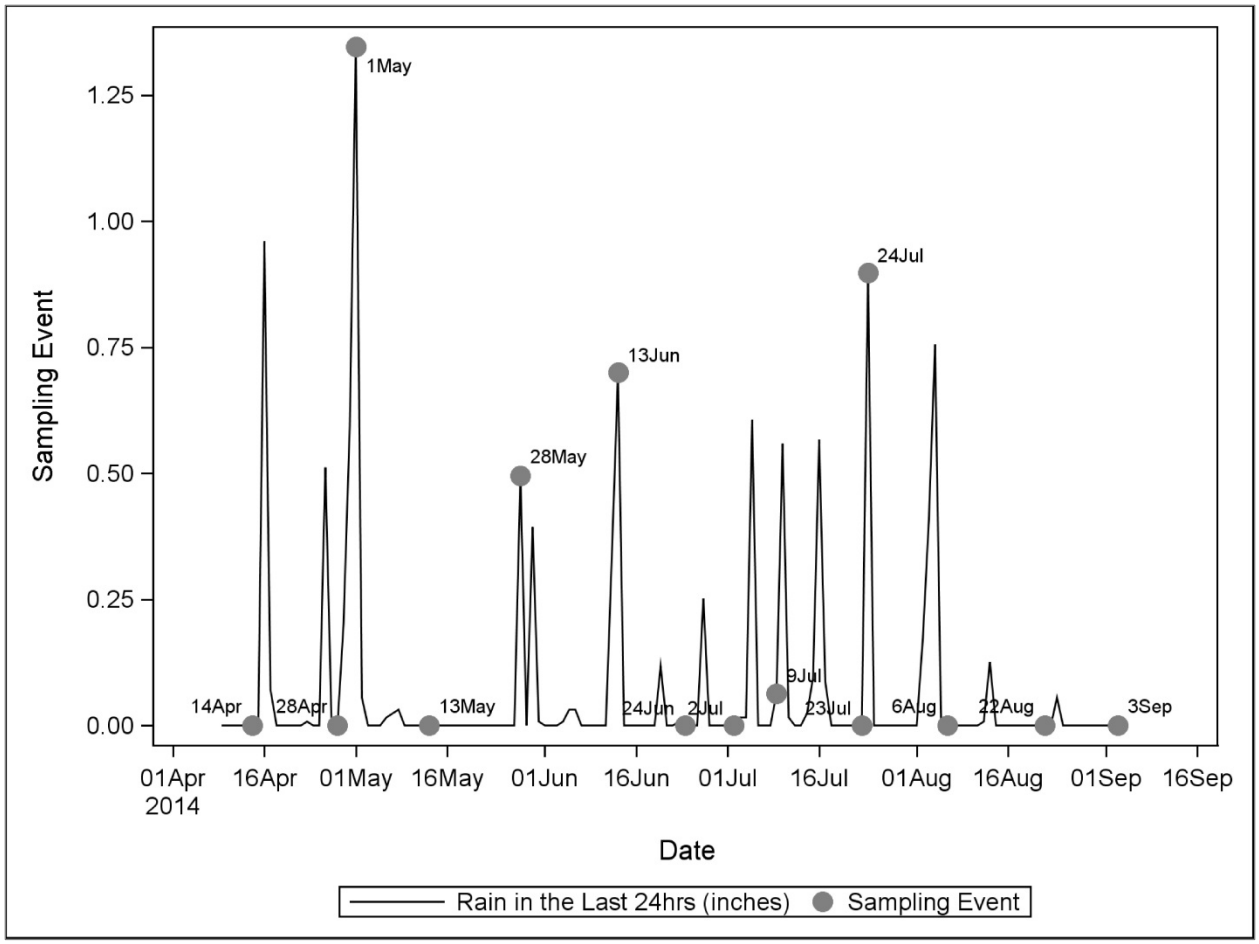

Figure 4.2. Total rain (inches) in the 24 hours preceding sampling for each sample collection date (gray dots). May 1 was the only sampling date with greater than 1 inch of rain, with three other sampling dates having 0.5 to 0.85 inches of rain preceding sampling.

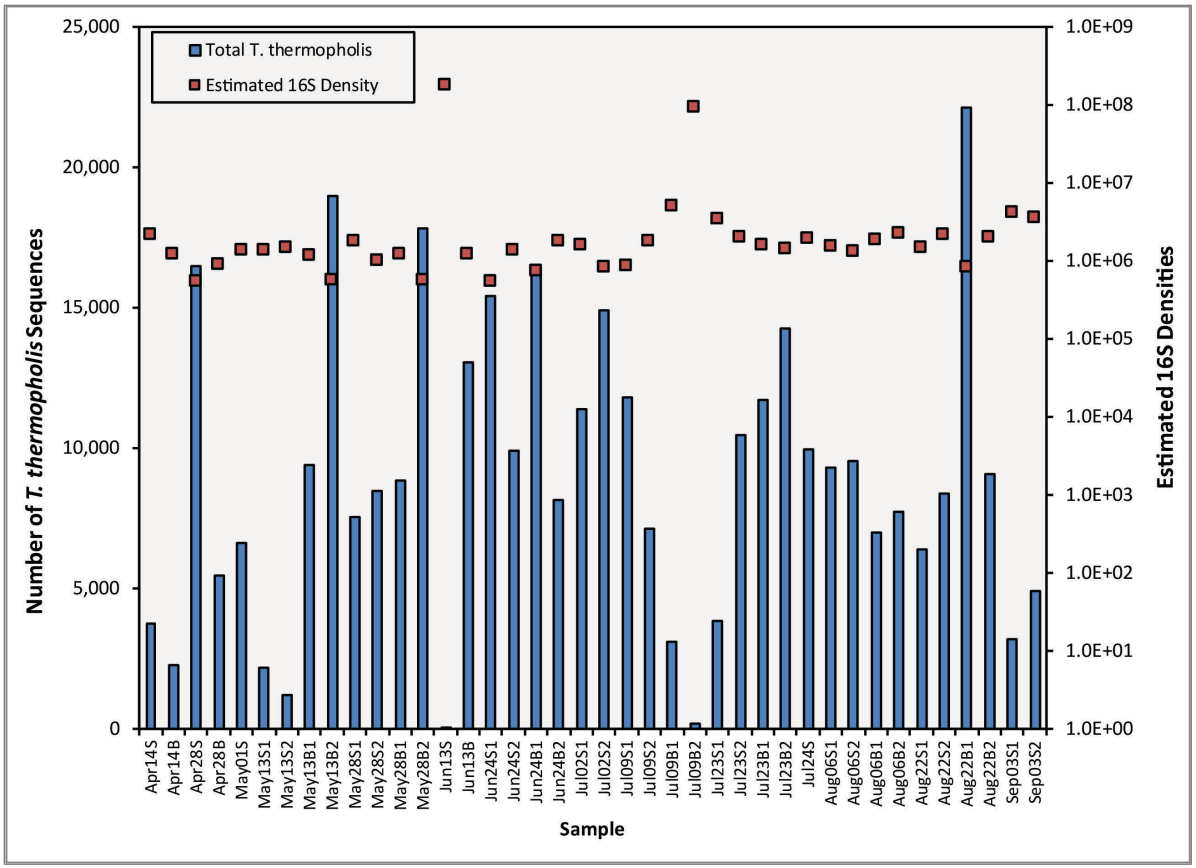

Figure 4.3. Comparison of the number of *T. thermophilus* (*Tt*) 16S rRNA reads (blue bars) and the estimated total 16S rRNA gene densities (#/mL) (red squares) for the 40 samples in this study. Sample names include month and day followed by S (surface) or B (bottom) and replicate number. Jun13S and Jul09B2 had unusually low numbers of *Tt* reads, indicating an unusually high density of bacteria or some error in the addition of *Tt* genes to those samples.

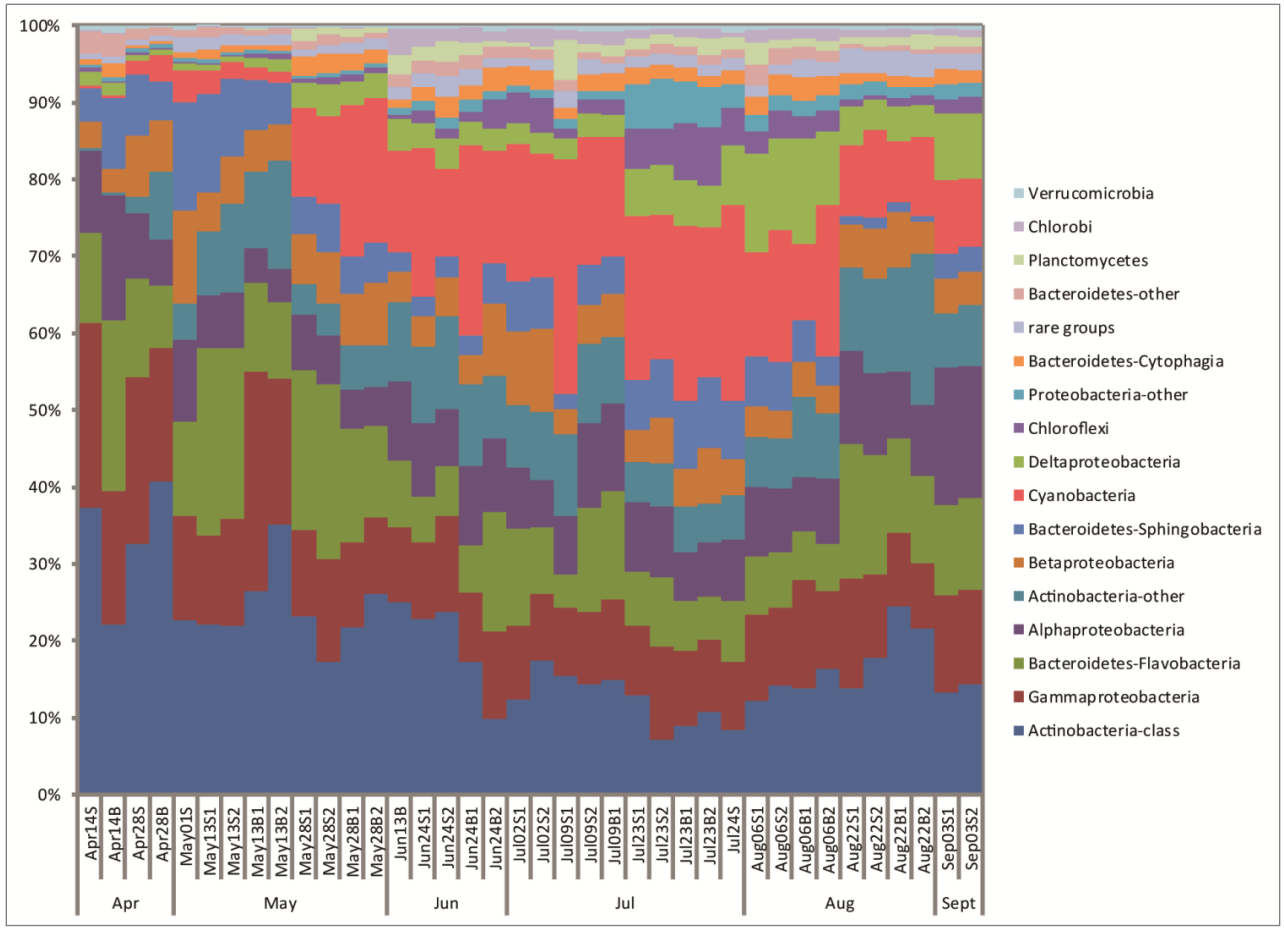

Figure 4.4. Percent 16S rRNA copy relative abundance for the 17 most common taxonomic groups by sample. Proteobacteria separated into Alpha-, Beta-, Gamma-, Delta-, and 'other' Classes. Bacteroidetes separated into Cytophagia, Sphingobacteria, Flavobacteria, and 'other' Classes. Actinobacteria separated into 'class' (for the Class Actinobacteria) and 'other' (for all other classes).

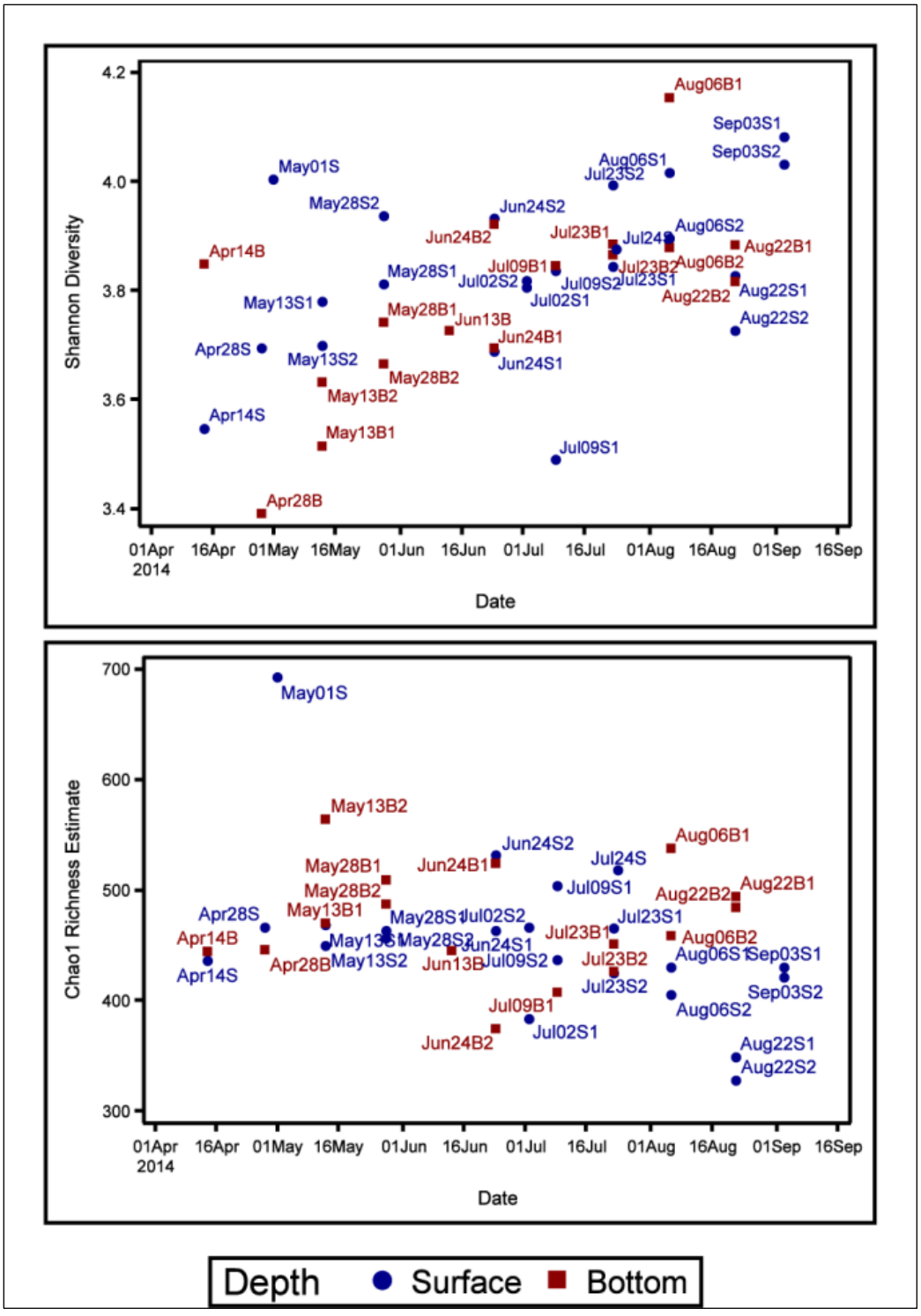

Figure 4.5. Measurements of alpha diversity by sample date and sample depth (surface and bottom). Panel A shows Shannon diversity values and panel B shows Chao1 richness estimates.

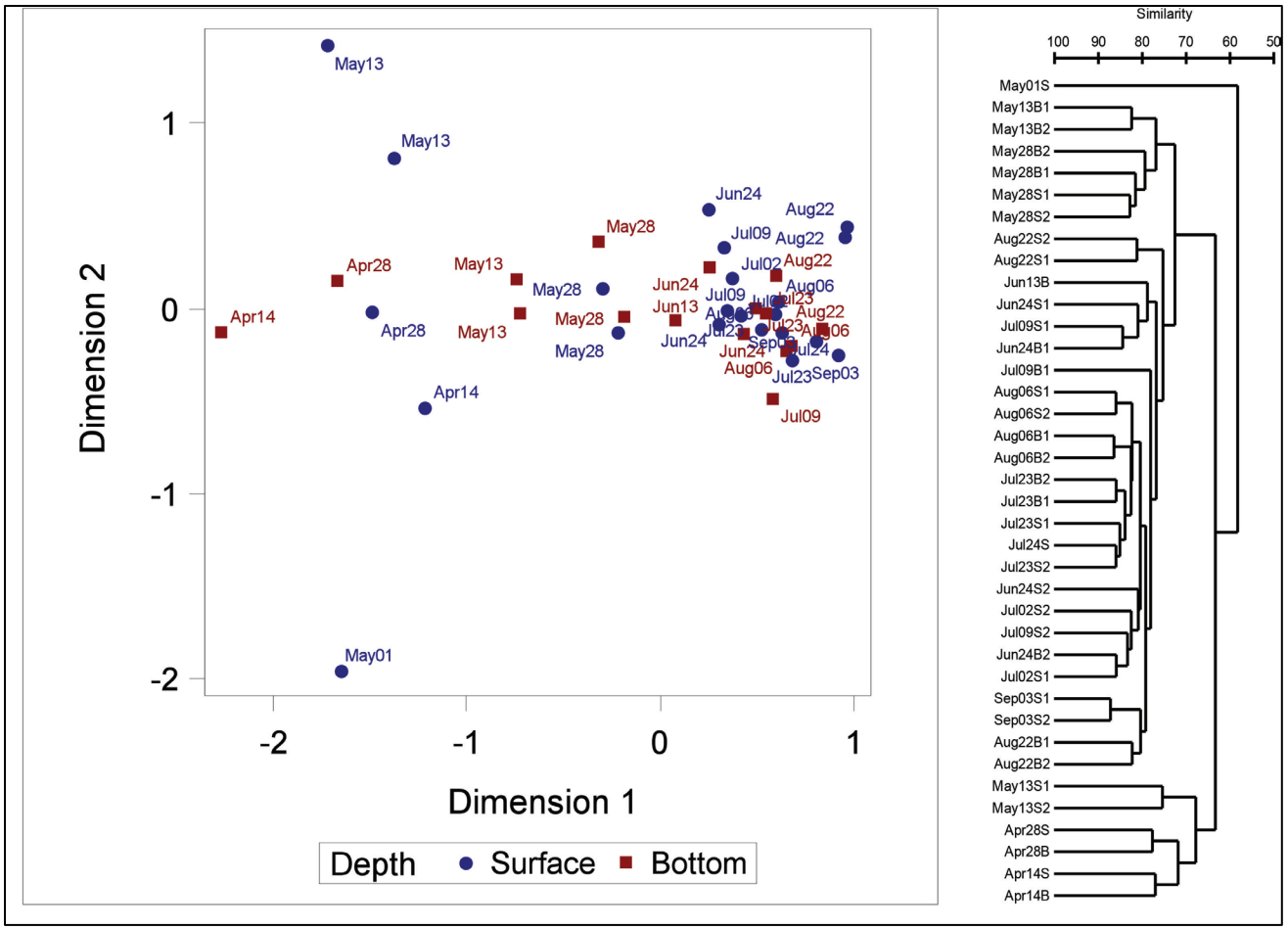

Figure 4.6. MDS plot and dendrogram of sample communities using Bray-Curtis Similarity comparisons of log transformed 16S rRNA copy abundance. MDS plot based on 25 iterations with final stress level of 0.01.

Table 4.3. Associations between environmental variables and patterns of community composition, as determined by distance-based linear modeling (DistLM). Statistically significant p-values are in bold.

Variable	Pseudo-F	P value	Proport.
Water Temp	14.80	**0.001**	0.291
Secchi Depth	8.19	**0.001**	0.185
Salinity	5.89	**0.003**	0.141
Dissolved Oxygen	3.78	**0.002**	0.095
Wind Direction	3.59	**0.006**	0.091
Rain in Last 24hrs	2.16	**0.038**	0.057
Tide State	1.62	0.096	0.043
Wind Speed	1.61	0.095	0.043
Cloud Cover	1.11	0.257	0.030

Table 4.4. Occurrence and estimated 16S rRNA gene density of select bacterial groups. The occurrence of genera on May 1 is highlighted by noting those that occurred on that date (Occurred), those that were elevated at least 2 times the standard deviation above their mean (Elevated), and those that were only detected on May 1 (Exclusive).

Family	Genus	Number Samples Detected	Total 16S rRNA Genes	Average 16S rRNA Gene/mL	1-May		
					Occurred	Elevated	Exclusive
Aeromonadaceae	*Aeromonas*	31	1460	38.4	✓	✓	
Campylobacteraceae	*Arcobacter*	38	14409	379.2	✓	✓	
Campylobacteraceae	*Campylobacter*	0					
Clostridiaceae_1	*Clostridium_sensu_stricto_1*	38	24327	640.2	✓		
	Clostridium_sensu_stricto_10	3	38	1.0	✓	✓	
	Clostridium_sensu_stricto_11	3	71	1.9			
	Clostridium_sensu_stricto_12	9	191	5.0	✓	✓	
	Clostridium_sensu_stricto_13	16	408	10.7	✓		
	Clostridium_sensu_stricto_14	2	17	0.4			
	Clostridium_sensu_stricto_17	17	486	12.8			
	Clostridium_sensu_stricto_18	7	99	2.6			
	Clostridium_sensu_stricto_19	1	10	0.3			
	Clostridium_sensu_stricto_2	6	182	4.8	✓	✓	
	Clostridium_sensu_stricto_3	1	15	0.4	✓	✓	✓
	Clostridium_sensu_stricto_5	3	114	3.0	✓	✓	
	Clostridium_sensu_stricto_7	6	88	2.3			
	Clostridium_sensu_stricto_8	2	18	0.5			
Clostridiaceae_4	*Clostridium_sensu_stricto*	4	66	1.7			
Enterobacteriaceae	*Brenneria*	1	15	0.4	✓	✓	✓
	Citrobacter	3	44	1.2			
	Enterobacter	1	6	0.1			
	Erwinia	2	229	6.0	✓	✓	
	Escherichia-Shigella	0	0	0.0			
	Klebsiella	1	15	0.4	✓	✓	✓
	Morganella	2	23	0.6			
	Plesiomonas	3	148	3.9			
	Providencia	1	11	0.3			
	Rahnella	2	147	3.9	✓	✓	
	Serratia	2	24	0.6	✓	✓	
	Yersinia	1	91	2.4	✓	✓	✓
Enterococcaceae	*Enterococcus*	2	29	0.8			
Helicobacteraceae	*Helicobacter*	0					
Legionellaceae	*Legionella*	38	26155	688.3	✓		
Leptospiraceae	*Leptospira*	11	244	6.4	✓		
Listeriaceae	*Listeria*	0					
Moraxellaceae	*Acinetobacter*	14	1502	39.5	✓		
Mycobacteriaceae	*Mycobacterium*	37	9848	259.1	✓		
Pseudomonadaceae	*Pseudomonas*	34	13777	362.6	✓		
Staphylococcaceae	*Staphylococcus*	9	375	9.9	✓		
Vibrionaceae	*Vibrio*	38	27545	724.9	✓		

Table 4.5. Spearman correlation coefficients for the comparison of selected genera (16S rRNA genes) detected in 3 or more samples to fecal coliform densities (estimated by A1 tube culture method), *Vibrio parahaemolyticus* and *Vibrio vulnificus* (estimated using qPCR), and environmental variables. Significant correlations ($p<0.05$) are in bold font.

Family	Genus	Fecal Coliform MPN	*Vv* (cells/mL)	*Vp* (cells/mL)	Water Temperature	Salinity	Secchi Depth	Dissolved Oxygen	Rain in Last 24 hrs
Aeromonadaceae	*Aeromonas*	0.08	-0.09	0.07	-0.01	-0.28	-0.10	0.15	**0.37**
Campylobacteraceae	*Arcobacter*	-0.26	**-0.34**	-0.31	**-0.56**	**0.54**	**0.32**	0.28	0.06
Clostridiaceae_1	*Clostridium_sensu_stricto_1*	-0.15	-0.04	0.01	**0.32**	0.30	**-0.33**	-0.22	-0.30
	Clostridium_sensu_stricto_10	0.23	-0.10	-0.10	-0.27	0.18	0.21	0.00	0.10
	Clostridium_sensu_stricto_11	-0.13	-0.10	0.06	-0.17	0.02	-0.05	-0.01	-0.17
	Clostridium_sensu_stricto_12	0.19	-0.02	-0.28	**-0.34**	0.29	**0.38**	0.08	-0.03
	Clostridium_sensu_stricto_13	-0.08	-0.28	-0.08	-0.14	0.09	0.00	0.27	-0.11
	Clostridium_sensu_stricto_17	0.13	-0.29	-0.01	-0.16	0.03	0.25	-0.27	-0.06
	Clostridium_sensu_stricto_18	0.21	-0.16	0.11	0.08	-0.26	0.00	-0.18	-0.13
	Clostridium_sensu_stricto_2	-0.02	-0.15	-0.15	**-0.45**	0.27	0.28	0.13	0.11
	Clostridium_sensu_stricto_5	0.12	-0.10	-0.19	**-0.36**	0.10	0.12	0.15	0.11
	Clostridium_sensu_stricto_7	-0.07	0.09	-0.27	-0.15	0.03	0.19	0.00	-0.25
Clostridiaceae_4	*Clostridium_sensu_stricto*	0.09	0.16	0.15	-0.18	-0.09	0.18	-0.30	0.01
Enterobacteriaceae	*Citrobacter*	0.06	0.18	**0.32**	-0.14	-0.17	0.16	-0.26	0.28
	Plesiomonas	0.06	-0.10	-0.05	0.00	0.06	0.06	-0.12	0.03
Legionellaceae	*Legionella*	-0.04	0.11	-0.09	**0.52**	0.09	**-0.45**	0.00	-0.31
Leptospiraceae	*Leptospira*	0.03	0.27	0.23	0.19	0.05	-0.22	0.13	-0.08
Moraxellaceae	*Acinetobacter*	-0.04	-0.12	0.12	-0.28	-0.08	0.10	0.09	0.44
Mycobacteriaceae	*Mycobacterium*	-0.02	-0.10	0.18	0.29	0.01	-0.21	**-0.47**	-0.07
Pseudomonadaceae	*Pseudomonas*	0.04	**-0.40**	0.13	-0.29	0.11	0.07	-0.10	0.13
Staphylococcaceae	*Staphylococcus*	**0.48**	0.15	-0.01	-0.01	**-0.32**	0.17	-0.04	0.11
Vibrionaceae	*Vibrio*	0.06	-0.20	0.02	-0.13	0.21	0.21	0.00	0.08

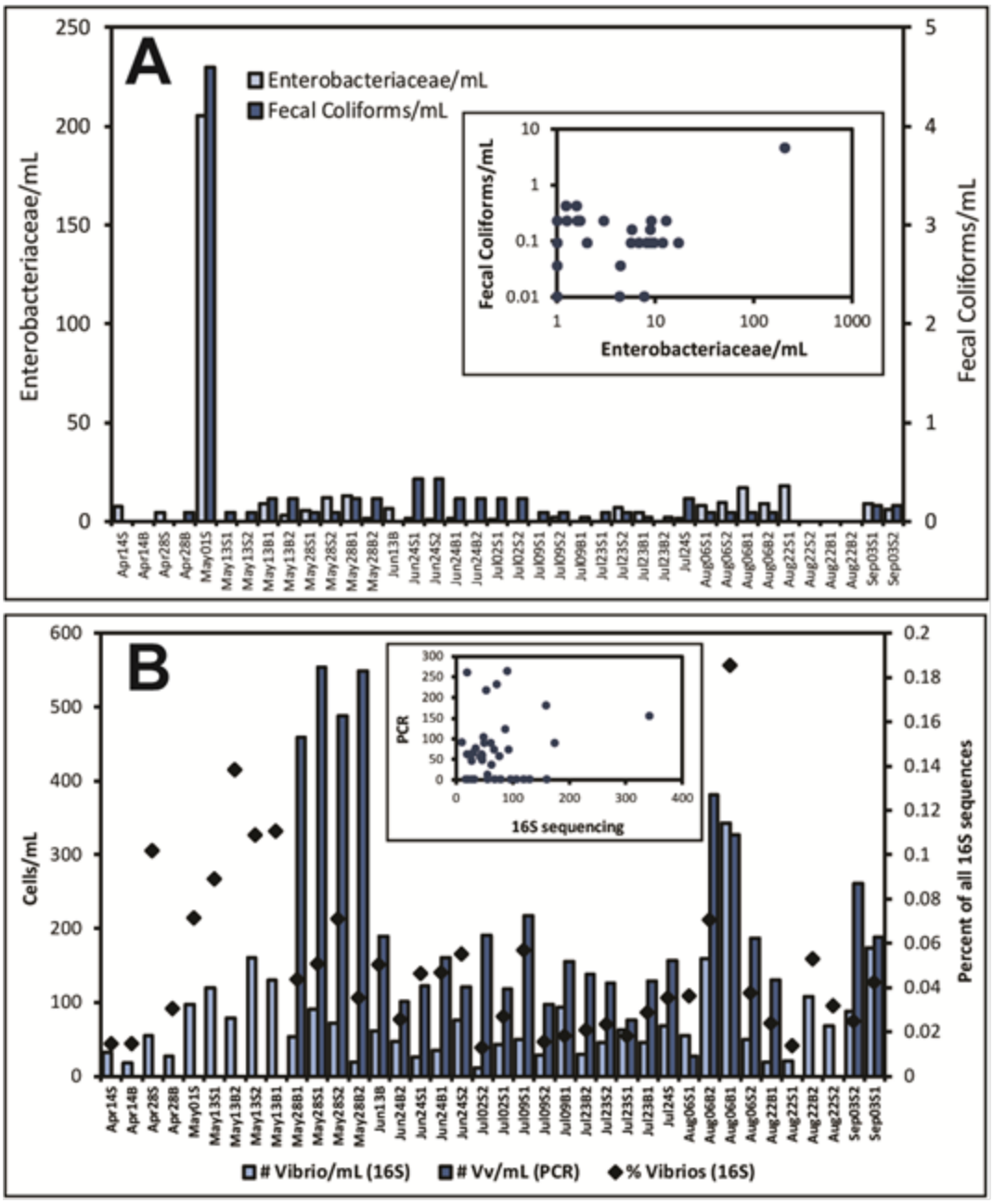

Figure 4.7. Comparison of Enterobacteriaceae (light bars; left axis) and fecal coliforms (dark bars; right axis) in Panel A, and *Vibrio vulnificus* (dark bars) with 16S rRNA sequencing estimates for total *Vibrio* (light bars), and the percent of *Vibrio* 16S rRNA genes in samples (dark diamonds) in Panel B. Enterobacteriaceae estimates based on 16S rRNA gene density/7 gene copies per cell. Fecal coliform estimates based on A1 tube method. Inset chart shows fecal coliform densities (y axis) compared to Enterobacteriaceae densities (x axis), with minimum values set to detection limits of 0.01 and 1.0 cells/mL, respectively. Panel B inset shows *Vibrio vulnificus* densities by PCR (x axis) compared to 16S rRNA sequencing densities/10 gene copies per cell (y axis).

Table 4.6. Potential and existing fecal bacteria indicator groups that were detected on May 1 (Occurred), those that were elevated on May 1 at least 2 times the standard deviation above their mean (Elevated), and those that were only detected on May 1 (Exclusive).

	Order	Family	Genus	Number of Samples	May 1		
					Occurred	Elevated	Exclusive
Potential Indicators	Bacteroidales	Bacteroidaceae	*Bacteroides*	33	✓	✓	
		Marinilabiaceae	*Mangroviflexus*	2	✓	✓	
			Marinifilum	14	✓		
		Porphyromonadaceae	*Dysgonomonas*	1	✓	✓	✓
			Macellibacteroides	2	✓	✓	
			Paludibacter	16	✓	✓	
			Parabacteroides	2	✓	✓	
		Prevotellaceae	*Prevotella*	6	✓	✓	
		S24-7	unclassified	4	✓	✓	
		unclassified	unclassified	19	✓	✓	
	Clostridiales	Lachnospiraceae	*Anaerosporobacter*	3	✓	✓	
Existing Indicator	Lactobacillales	Enterococcaceae	*Enterococcus*	2			

Chapter 5: Discussion

Estuaries have long served as receiving waters for unwanted or excess substances produced by humans. Some of these substances, such as nitrogen and phosphorus, occur naturally and are essential in modest concentrations, but become pollutants at high concentrations. Others, such as sewage and manure, are not desirable at any concentration and are considered especially problematic above some relatively low concentrations. Advances in science and technology have provided structures and methods for decreasing the amount of sewage and manure entering waterways and for mitigating the risk of illnesses from waterborne and seafood-borne pathogens, especially in developed countries. Yet, fecal bacteria and associated pathogens are still introduced to waterbodies and illnesses related to these pathogens still occur. In addition to pathogens, the introduction of fecal pollution brings antibiotic resistance and virulence genes into the aquatic environment, where they may be shared across bacterial groups (Baquero et al. 2008, Klein et al. 2014).

In order to protect public health and preserve environmental quality, therefore, it is imperative to understand the drivers and dynamics of fecal bacteria and the pathogens with which they may occur. In many cases, however, the tools for assessing the presence of fecal pollution and human pathogens in natural waters are arcane, imprecise and time consuming. The large temporal and spatial variability in allochthonous pathogen groups, those being released from point sources or washing off of land, coupled with the logistical challenges of monitoring extensive shellfish growing and recreational areas, make the protection of human health from these pathogens a challenging endeavor. Complicating matters, some human pathogens

may be present as part of the natural microbial community, may not be accurately detected or quantified using traditional indicators, and may require a variety of management actions to mitigate. Thus, numerous types of pathogens may be present in an area, with the timing and magnitude of their presence likely driven by a suite of environmental factors, with differential influence based on the characteristics of the particular pathogen.

These are important concerns in the Chesapeake Bay. As the nation's largest estuary the Chesapeake Bay represents an iconic and economically important waterway which, due to its very high land to water ratio, is also impacted by the pressures of an ever increasing human population. While the Bay has provided a bounty of natural resources, overfishing and disease have diminished fisheries, including shellfish stocks. Water quality also suffers, in places, from inputs of excess nutrients, contaminants, and fecal bacteria, resulting in restrictions on harvest of fish and shellfish. Decades of monitoring fecal indicator bacteria and conducting surveys of potential point sources (i.e. sanitary surveys) have resulted in few documented cases of human illness resulting from the consumption of shellfish or from recreational activities in Maryland. However, violations of the water quality standards for fecal indicator bacteria still occur in some sections of Maryland waters and both consumption related illnesses and severe wound infections continue to occur. A recent resurgence in aquaculture in Maryland (Green and Tracy 2013) only increases the need to accurately assess the condition of Bay waters relative to environmental condition and human health and place increased demands on managers

to improve water quality in areas that have otherwise been restricted due to the presence of fecal indicators.

The research described in this dissertation helps to better define the relationship of some environmental factors with the presence of fecal indicator bacteria in shellfish harvest areas of Maryland, USA, and provides managers of these waters information necessary to refine their management approaches. Firstly, the role of climate variability on fecal coliform densities was evaluated. The influence of annual precipitation amounts provided the most important predictor of the percent of fecal coliform densities exceeding the management criterion set forth by the National Shellfish Sanitation Program and promulgated by the FDA (FDA 2015). This finding agrees with those for the lower portion of Chesapeake Bay (Huang 2010), as well as estuaries along the Gulf of Mexico (Lipp et al. 2001a, Chigbu et al. 2004). Air temperature provided further predictive power and, as found in previous studies (Burkhardt et al. 2000, Chigbu et al. 2005), was negatively related to FC densities. These two factors, annual precipitation totals and maximum summer air temperature, accurately explained about 77% of the inter-annual variability in the percent of fecal coliform monitoring stations exceeding the FDA criteria. Seasonality also proved to be very important in describing fecal coliform densities, with highest densities occurring in fall and early winter.

An important separation in time between peak precipitation and FC densities was also discovered. The precipitation threshold analysis (Chapter 3) and the bacterial community assessment (Chapter 4), as well as other studies (MDE 2006), show that high levels of FC resulting from most rain events tends to be short lived, on

the order of days. Therefore the separation in time between months with high precipitation amounts and months with average fecal coliform numbers exceeding the FDA criteria suggests an additional component to the relationship. In Chapter 3, antecedent rainfall was shown to be an important factor controlling the magnitude of fecal coliform densities resulting from individual events, which might help explain this separation in time. Another possibility is that the runoff of fecal coliforms becomes more pronounced as the amount of runoff relative to the amount of rain increases in late fall when leaves fall from trees and colder air temperatures result in less absorbent soils. A delay between peak rainfall and peak fecal coliform levels may also represent an accumulation of fecal bacteria within the aquatic environment as winter storm winds resuspend bottom sediment associated fecal coliforms. Survival of fecal indicator bacteria in aquatic sediments has been detected elsewhere (Davies et al. 1995, Anderson et al. 2005). A positive relationship between the proportion of samples exceeding the FDA criterion with increasing wind speed was detected (Chapter 3) with some relatively high percentages of samples containing fecal coliforms above the FDA criteria when winds were greater than 10mph. However, recent rainfall explained a greater amount of the variability in the percentage of violations than high winds, without showing signs of collinearity to wind. Management strategies, mostly static in their current application, should reflect these seasonal changes in fecal coliform dynamics. Further, the random-stratified sampling design mandated by FDA and applied by coastal states attempts to account for the natural variability of fecal coliforms in shellfish waters. The research in this dissertation shows that there are a variety of forces at multiple scales, from regional

climate patterns to individual rain events, which influence fecal coliform concentrations in Maryland waters.

A dominant climate pattern is also described, associated with years of very high precipitation and correspondingly high FC densities for the Chesapeake Bay region. Sea-level air pressure anomalies in the summer and/or fall of these years reveal a sub-hemispheric climate pattern with higher than normal pressure values for areas over the central US and western Atlantic Ocean and a trough of lower than normal air pressure over the Atlantic seaboard. This air pressure system directs tropical and extratropical storms over the eastern seaboard of the US and establishes conditions conducive to storm generation, i.e. cyclogenesis. An index defined by the relative difference in sea-level air pressure between the western Atlantic and Chesapeake region showed moderate success in predicting years of high precipitation and preliminary success at identifying years with excessive FC densities. This index represents a true climate teleconnection in that it provides a measure of the water quality in Maryland waters resulting from large-scale climate patterns. It would also likely indicate other Bay dynamics, such as runoff of phosphate and sediments, which are also closely tied to rainfall.

In addition to climate variability, as described here, a common and important concern among scientists and policy makers is the impact of long-term climate change on conditions within the Chesapeake Bay. Predictions of the impacts of climate change on precipitation and air temperature patterns vary for the Chesapeake Bay. Average annual precipitation levels are projected to increase modestly over the next 50 to 100 years (Ning et al. 2012), though rainfall intensity will likely increase

more significantly in winter and spring (Najjar et al. 2010). Recent research has projected an increase in the number of extreme rain events and the duration of summertime heat spells (St.Laurent et al. 2016). Any substantial increase in the overall precipitation levels or the magnitude of individual events will likely increase the introduction of land-based fecal pathogens and indicators. Extreme precipitation events have already been linked to increases in the waterborne pathogen outbreaks for the US (Curriero et al. 2001). An increase in summertime high temperatures, however, may serve to lower concentrations of fecal bacteria, and associated pathogens, coming from nonpoint sources. A more important result will likely be that increasing air temperatures will change the location and extend the intra-annual time period that autochothonous pathogens, such as *Vibrios*, and algal toxins, such as paralytical shellfish disease, will be found in surface waters and shellfish (Trtanj et al. 2016).

While the climate variability and synoptic climatology work detailed in Chapter 2 provided important insights into large scale influences on fecal indicator bacteria and the importance of antecedent conditions in Maryland's shellfish harvest waters, there is a more direct need to understand the impact of individual rain events. In Chapter 3, watersheds in Maryland were identified where rainfall may introduce fecal bacteria, either from point or non-point sources, such that levels in the receiving waters exceed established criteria for healthy shellfish consumption. Current management strategies include the classification of particular shellfish harvest areas as conditionally approved, where such relationships to precipitation have been established through intensive field collections. About half of the watersheds tested

showed a relationship between fecal bacterial levels and rainfall, most of these watersheds having a portion already classified by Maryland natural resource managers as conditionally approved areas, thus supporting the historic field collections that were conducted to establish such areas. The level of rainfall necessary to introduce fecal coliforms at levels exceeding the FDA criterion was always above one inch and varied between watersheds, suggesting that the current threshold of an inch is conservative at the watershed level, especially in areas with large water to land ratios, and that for some areas this threshold may be overly restrictive.

The analysis in Chapter 3 also revealed the role of several environmental drivers of fecal coliform levels other than recent rainfall. For example, distance from land and antecedent conditions were shown to be important in the response of fecal coliform densities to rainfall. Land use and watershed characteristics also proved important. Both the frequency of exceedances and the rainfall threshold correlated to the percent of open water in the watershed and the types of soils found in the watershed, but not to the amount of development or impervious surfaces as reported elsewhere. Other studies have noted a positive relationship between fecal indicator bacteria and land use (DiDonato et al. 2009, Gonzalez et al. 2012). In this study, a variety of land uses were important for fecal coliform levels in Maryland waters, such that both the heavily developed Patapsco River and the relatively forested and marsh-lined Nanticoke and Wicomico Rivers were classified by the frequency analysis (Chapter 3) as already exceeding the criteria regardless of rain. The impacts of urban development, with notoriously dated and leaky infrastructure, are likely the culprit in

the Patapsco while relatively long water residence time and agricultural sources likely explain the elevated fecal coliform levels in the Nanticoke and Wicomico Rivers. Additionally, long residence time and extensive wetlands may play a role in the relatively narrow and poorly flushed Wicomico River. The role of wetlands in elevated fecal pollution has been implicated before (Huang 2010), likely results from warm-blooded animals using the wetland habitat, and should be further investigated.

All of the fecal coliform data utilized in Chapters 2 and 3 were collected by the Maryland Department of Environment. The availability of this data to conduct multi-decadal retrospective analyses highlights the importance of maintaining a long-term monitoring program with consistent methods and endpoints. While other states, such as Virginia, have opted to switch the method of fecal coliform measurement to a direct count used for assessing shellfish harvest areas, Maryland continues to use the standard dilution-tube method. Given the challenges with interpretation of MPN values, Maryland should also consider changing their method. However, any conversion to other data endpoints will need to ensure data comparability over time by including a rigorous methods comparison, including samples taken over a range of conditions for key drivers, such as rainfall, antecedent rainfall, wind, and temperatures. The investigation of rainfall thresholds was underpinned by the development of spatially resolved rainfall estimates based on multiple sensors. As the number of years of such data increase, more refined analyses of rainfall effects on water quality will be possible.

One of the limitations in the fecal coliform data collected by Maryland, using the sampling method prescribed by FDA, is the periodicity of collections. In order to

account for the low frequency of collection, the risk of encountering sewage pollution at a shellfish harvest area is assessed by comparing both the median and 90th percentile at least thirty samples to the FDA criteria. However, in order to understand the true distribution of fecal coliforms, at any location where their density is variable, more frequent sampling is required. For this doctoral research, an attempt was made to understand the impact of sampling frequency on the estimation of this true distribution. Daily samples were collected and measured for fecal coliform densities at the same station monitored for bacterial community dynamics in Chapter 4. Unfortunately, this sampling frequency also proved too coarse, as the distribution of fecal coliforms contained sharp peaks and valleys defined by single observations. Future studies of sampling frequency bias will likely require hourly or even sub-hourly collection schedules.

The study of dynamics within the bacterial community, presented in Chapter 4, shows that the presence of fecal coliforms, as well as other allochthonous bacteria, can also be event-driven on the local scale. This 16S rRNA bacterial diversity study represents an area of rapid and exciting advances in environmental microbiology. The ability to place the presence and numbers of select bacterial types in context with the larger bacterial community offers insights not possible with traditional microbiological methods. The co-occurrence of a number of bacterial groups known to contain pathogens with fecal coliforms and several other proposed fecal indicator bacterial groups provides a more holistic investigation of indicators than could be achieved with direct culture. This 16S diversity study also underscores the diversity and complexity of the bacterial community in which fecal indicator bacteria and

pathogens reside. Many of these community members are uncultured and undescribed. Understanding both the co-occurrence and interactions among these community members will provide a much better understanding of microbial aquatic ecology and human health risk assessment. As the efficacy of using fecal indicator bacteria to manage risk of infection has been questioned (Wade et al. 2003, Schwab 2007), and a number of pathogens have been isolated from Maryland waters, the future of pathogen assessments may reside in the use of such cutting-edge technologies as genomics and metagenomics. The management of natural waterbodies for public health protection is likely to undergo a paradigm shift as these technologies allow greater understanding of pathogen occurrence as well as bacterial community composition.

Some temporal dynamics in the bacterial community were uncovered by the 16S rRNA study. For example, seasonal changes in air temperature, turbidity, and salinity proved important for bacterial community composition, though they only explained a small amount of variation in the data. Previous studies have indicated that spatial differences in salinity also have large effects on the composition of bacterial communities (Kan et al. 2006, Fortunato et al. 2012). Although rainfall was not related to general trends in the bacterial community described in Chapter 4, distinct and temporary changes were detected following a rainfall of approximately 1.3 inches. With a particularly long residence time, waters in the Chesapeake Bay do not turn over rapidly, highlighting the event-driven nature of these allochthonous bacteria and their poor growth in the water column. However, it is possible that some members of the allochthonous bacterial community persisted in the underlying

sediments, as observed for some fecal indicator bacteria (Davies et al. 1995, Anderson et al. 2005). Any persistence of these bacteria in the environment increases the opportunities for allochthonous and autochthonous bacteria to share genes, raising concerns over the spread of genes of human health concern, such as those responsible for antibiotic resistance and virulence.

The initial development of a method of estimating bacterial densities using an internal control presents an important advancement in the characterization of microbial communities. Combined with more information about the number of 16S rRNA gene copies per cell for rare bacteria, this method would allow a much better understanding of bacterial communities by not only assessing the types of bacteria present but their relative densities. Currently, the influence of cell densities on bacterial community dynamics, such as quorum sensing and transformation, relies on studies with identification of select species and molecular targets.

Though high-throughput instrumentation was used for the research presented in Chapter 4, the number of samples included in the community analysis conducted for this dissertation was purposely kept to a low level in order to increase the number of reads per sample. The use of an internal control, while important for estimation of total bacteria cells in each sample and as a quality control for accurate processing of DNA samples, decreased the total number of reads from each sample. Further improvements in the science of genomics and metagenomics will continue to lower the cost and analytical demand on the study of bacterial communities and virulence markers, allowing more robust study designs and improving classification databases for aquatic microbes.

The research described in this dissertation focused primarily on understanding common drivers of fecal coliform concentrations in surface waters. Although rainfall amounts were integral at all three scales of research (regional, event, and local bacterial community), the impact of extreme events was not directly addressed. Recent efforts to characterize and quantify the occurrence of extreme climate conditions, such as consecutive days of extreme heat or cold, have been conducted for the Chesapeake Bay (St.Laurent et al. 2016). Comparison of the fecal coliform data used in Chapter 3 to these climate indicators might provide further insight and refine our understanding of the relationship between climate and fecal coliform levels in the Maryland portion of the Bay.

Appendix A: Development of a climate index for years with extremely high precipitation and fecal coliform densities in the upper Chesapeake Bay

Climate indices, such as the Northern Atlantic Oscillation and El Niño Southern Oscillation, have been developed to describe climate patterns with connections (so called teleconnections) to environmental conditions. Research presented in Chapter 2 defined a climate pattern for the northern Chesapeake Bay indicative of years with high precipitation amounts. These years correlated well with years of high proportions of samples from shellfish harvest areas containing fecal coliforms above the FDA criteria. Sea-level pressure anomalies for the three years with extremely high fecal coliform levels (1979, 1996, 2003) where selected and compared to normal sea-level pressures for 1979 to 2013 to define a synoptic-scale climate index for the Chesapeake.

Development of the climate index was based on comparing sea-level pressure anomalies between the western Atlantic Ocean and the Chesapeake Bay region for the months May to October. Monthly sea level pressure (SLP) data were obtained from the Earth System Research Laboratory (ESRL) (www.esrl.noaa.gov) in the form of 'reanalysis' data (Kalnay et al. 1996), as described in Chapter 2. Based on examination of sea-level pressure anomalies for the three extreme years, two regions were selected for development of the Chesapeake Bay Index. Average sea-level air pressure anomalies over the western Atlantic were calculated for the region 30°N to 45°N latitude and 50°W to 70°W longitude and compared to average sea-level air pressure anomalies over the Chesapeake region as defined by the area 37°N to 42°N

latitude and 74°W to 84°W longitude (Figure A.1). Annual values for the index were calculated as in Equation A1.

Eq A1. If anomWA > 0 then,

CB_Index = anomWA –anomCB

Else

CB_index = -1

Where anomWA = average sea-level pressure for western Atlantic for May to October of each year compared to the average sea-level pressure (1979-2013); anomCB = average sea-level pressure for Chesapeake Bay region for May to October of each year compared to the average sea-level pressure (1979-2013)

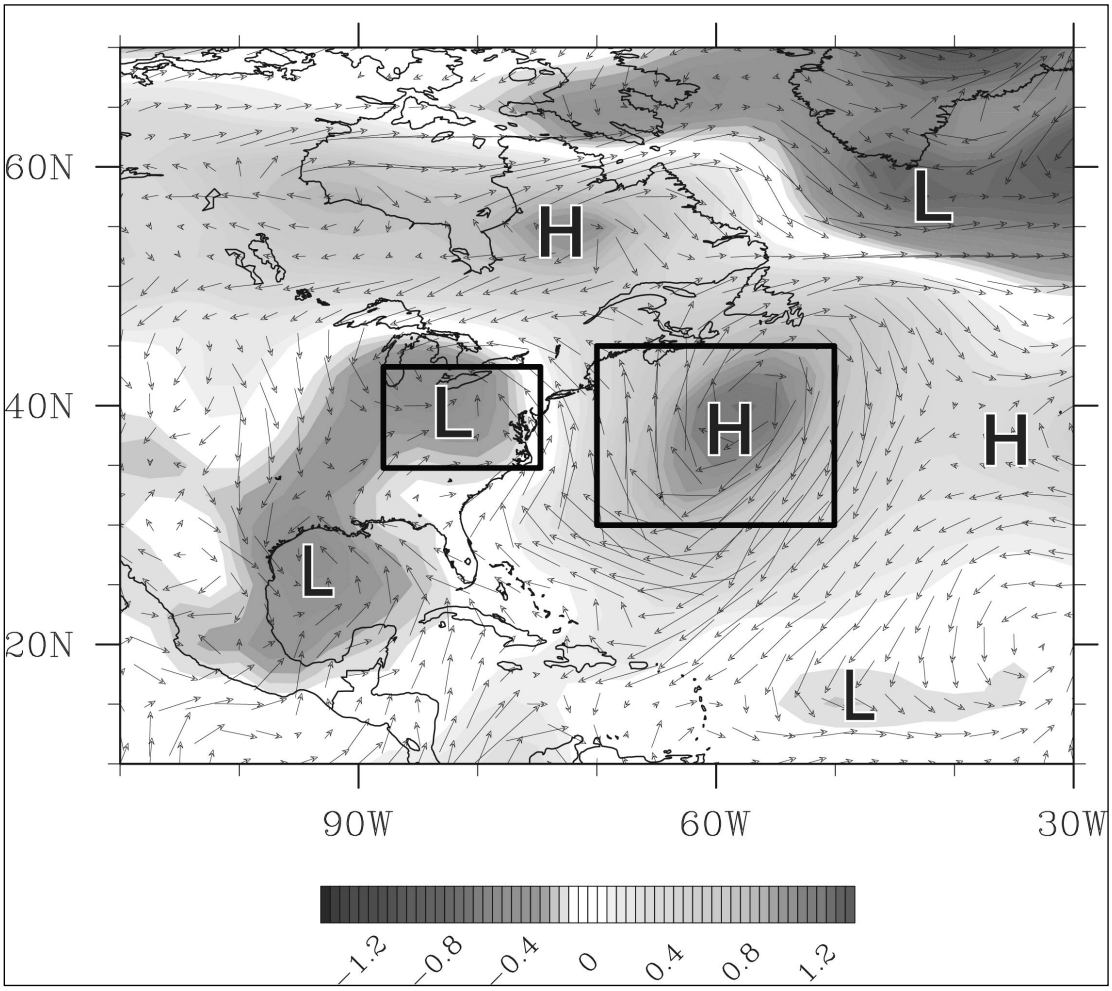

Figure A.1. Sea-level pressure anomalies for the three years with extreme FC exceedances (1979,1996, and 2003) compared to the long-term average of for the period 1979-2014, for the months May to October.

A clear, positive relationship was seen when comparing annual CB Index values to annual average precipitation levels (as defined in Chapter 2) with the three highest years of precipitation (1996, 1979, and 2003) having index values over 0.5 and four of the six highest precipitation years having index values over 0.5 (Figure A.2). No years with greater than 50 inches of rain had an index value over 0.5. Similarly, index values over 0.5 selected two of the three years (1979 and 1996) with the lowest maximum summer air temperatures, but also included two years (2003 and 2009) with normal air temperatures (Figure A.3). Years where a very high proportion of samples containing fecal coliform densities over the 49MPN/100mL criterion tended to also have CB Index values over 0.5 (Figure A.4), based largely on the relationship between rainfall and high fecal coliform levels described in Chapter 2.

Further refinement of this index would be useful for describing climate forcing of conditions in the Chesapeake Bay as they relate to precipitation, fecal coliform levels, as well as other ecological conditions.

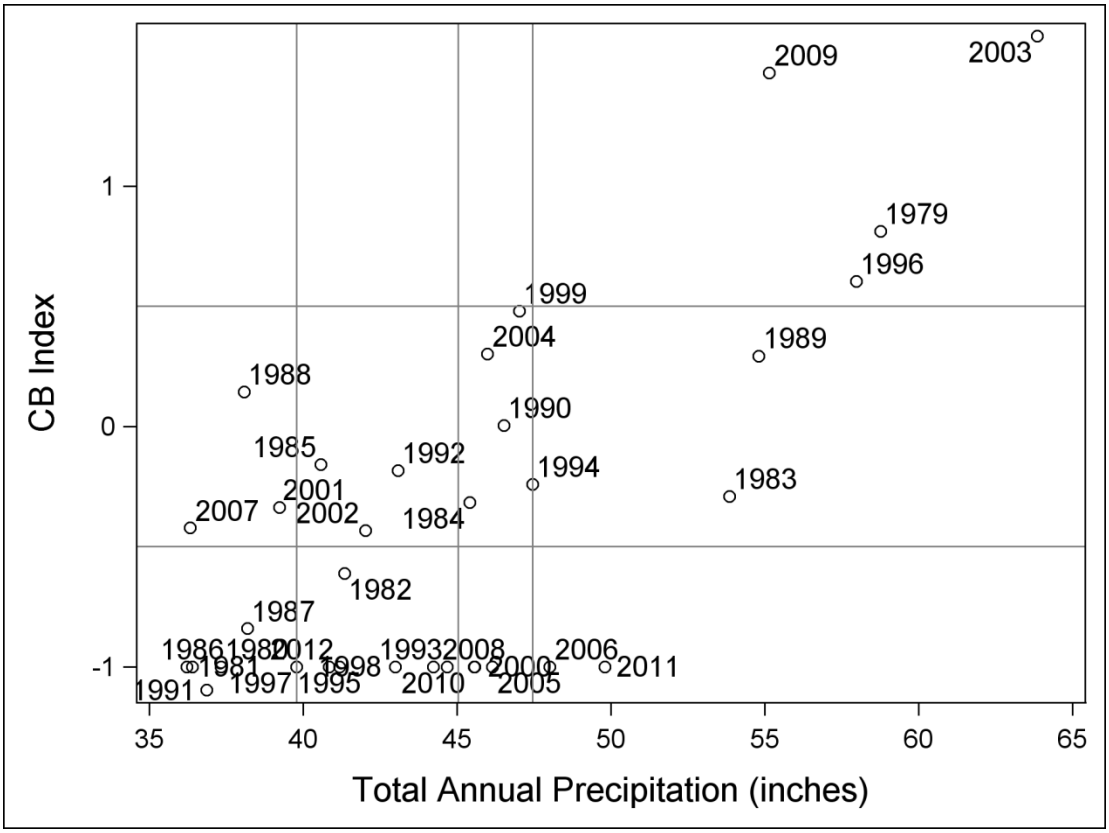

Figure A.2. The CB Index compared to annual average precipitation levels for the Northern Chesapeake Bay.

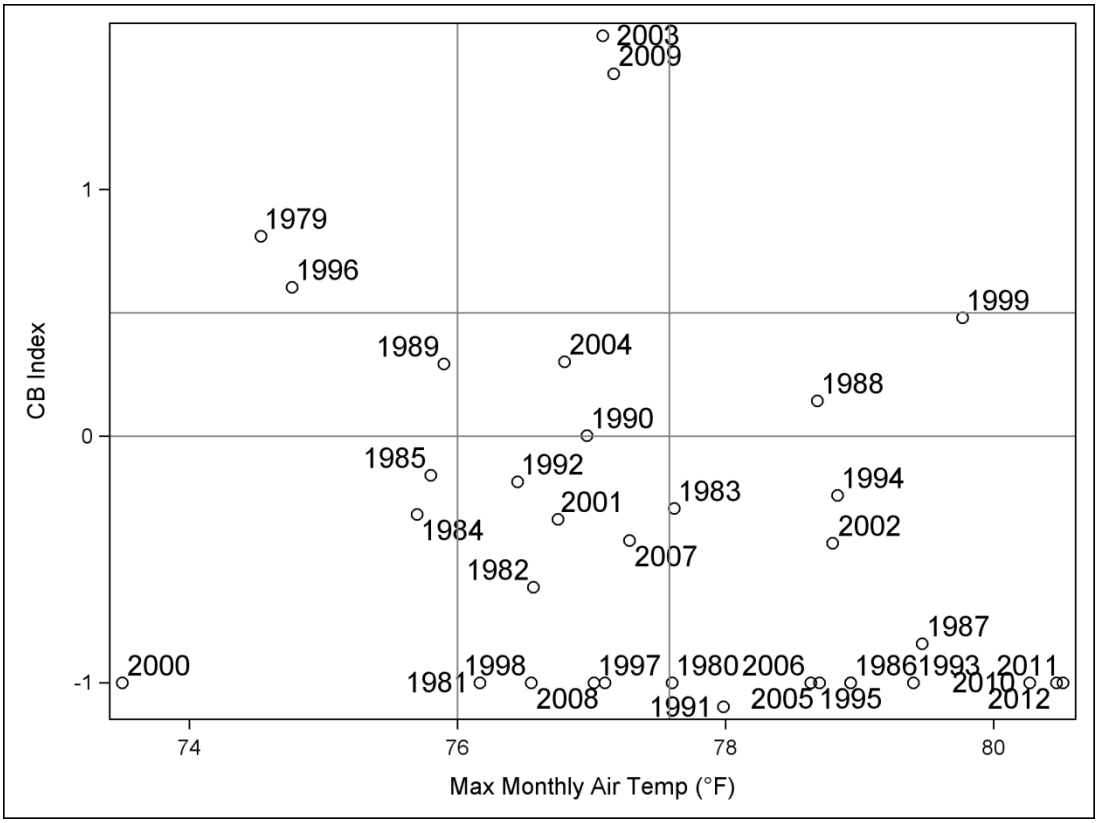

Figure A.3. The CB Index compared to maximum monthly air temperature for the Northern Chesapeake Bay.

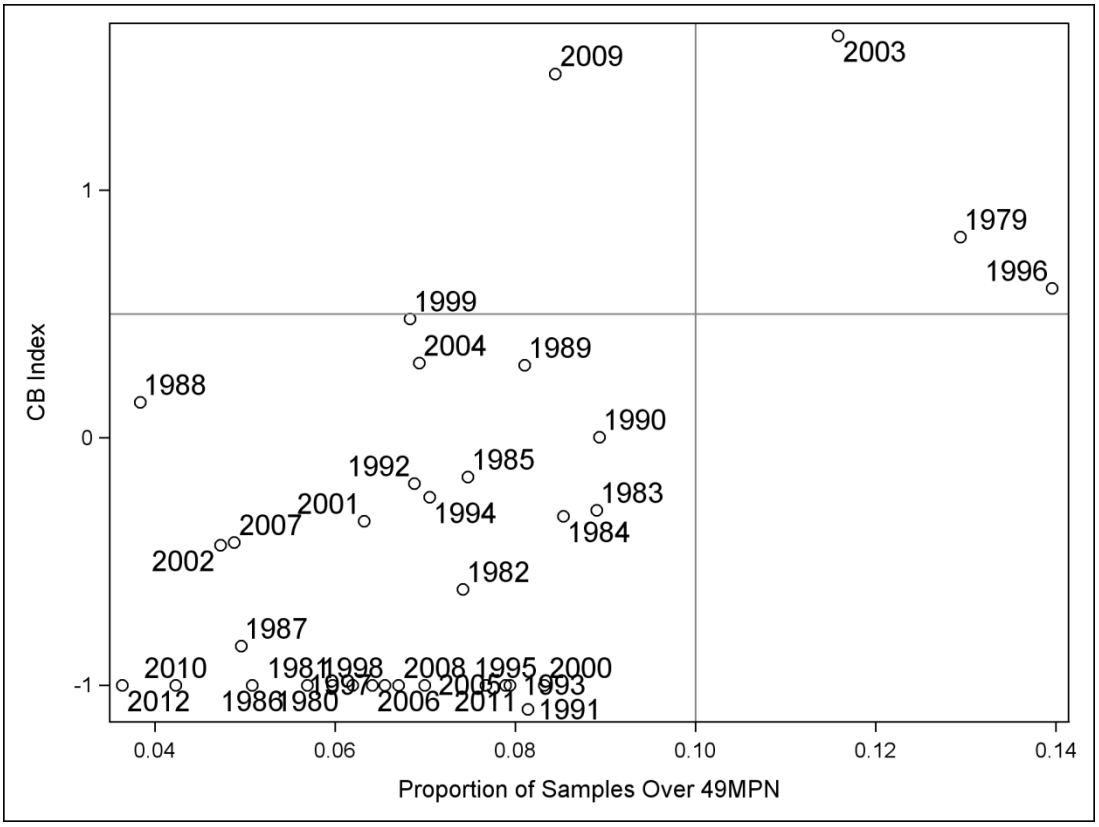

Figure A.4. The CB Index compared to the annual proportion of samples over the FDA criterion.

Appendix B: Long term trends in fecal coliform densities in Maryland shellfish harvest waters

The Maryland Department of Environment (MDE) has collected surface water samples at shellfish harvest grounds since the 1950s. However, monitoring stations for shellfish waters in Maryland have been added and removed over time. Therefore, a subset of monitoring stations that have been continuously sampled were selected for our analysis. A significant change in the number and locations of sampling stations occurred in 1978 and 1979, making that a necessary time period to begin our analysis. For the purposes of this study, we considered any station as continuously monitored if it had been sampled at least four months out of each year from 1979 to the 2013, with at least six samples for each year. Even within this subset, there were a number of stations missing several years of data. However, these gaps did not tend to be for a large number of stations within the same year. Therefore, stations were included that had no more than four non-contiguous years in which no samples were collected. These criteria resulted in the selection of 355 continuous sampling stations, which included most of the rivers and bays of Maryland Chesapeake Bay waters (Figure 2.1).

Surface water estimates of fecal coliform bacteria densities from the selected stations were obtained from an existing MDE database. These data were generated in a consistent manner for the entire period of sample collection, using the three-tube broth dilution protocol as recommended by the American Public Health Association (APHA 1998) and FDA (FDA 2013). Data were excluded from the analysis if it was

from targeted sampling around specific weather events or samples collected from more than one meter below the surface of the water.

The proportion of samples exceeding the 49MPN/100mL criterion showed great interannual variability, but also showed a general decreasing trend over time (Figure B.1). Although this trend was insignificant ($p>0.10$) when year was considered as a stand-alone explanatory variable, when included with total annual precipitation and average temperature of the hottest month (Chapter 2), year was significant. (Table B.1). However, it only explained a small percentage of the overall variability. Including year raises the adjusted R^2 from 0.72 to 0.75. A more focused look at the data shows a relatively low proportion of samples that were over the criterion for 1980 and 1981, and again from 1986 through 1988. Drought conditions over much of the US during the 1980's likely contributed to these lower fecal coliform levels. These periods were followed by a general increase and more stable proportions in the early 1990s, which were followed in-turn by a more variable but distinct decrease in the proportion of samples over the FDA criterion. These decreases over the last two decades showed a general improvement in water quality resulting from efforts to reduce the introduction of fecal matter into the Bay from both point and nonpoint sources.

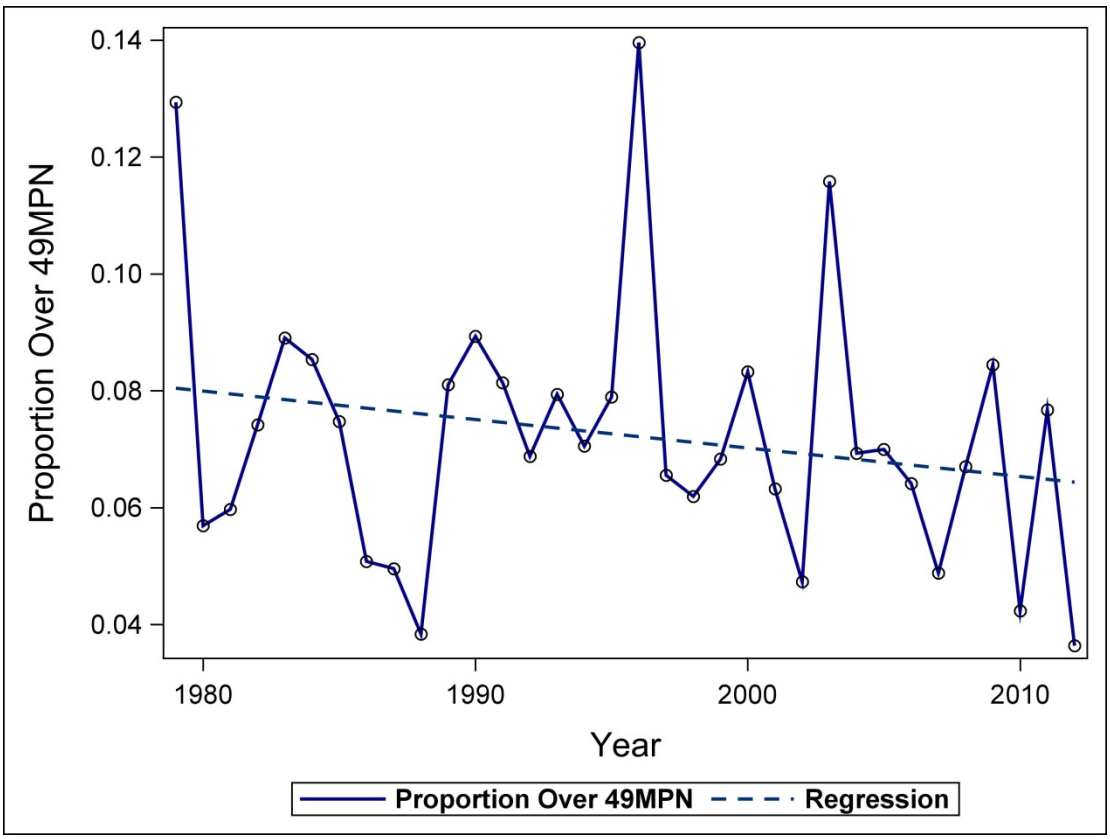

Figure B.1. Annual proportion of samples over the 49MPN/100mL criterion (solid blue line), with linear trend line (dotted line).

Table B.1. Results of linear model comparing the annual proportion of samples over 49MPN/100mL to year, total annual precipitation, and the temperature of the hottest month.

Variable	Parameter Estimate	Standard Error	p
Intercept	1.236	0.4106	0.005
Year	-0.0005	0.0002	0.040
Total Precipitation (inch)	0.0024	0.0003	<.0001
Hottest Monthly Temperature	-0.004	0.001	0.006

Appendix C: Stations included in analyses

The following table lists stations included in the retrospective analyses conducted in Chapters 2 and 3. These stations were selected from a larger set of over 1,900 stations which have been monitored by Maryland Department of the Environment since the year 1979.

Station	Latitude	Longitude	Chapter 2	Chapter 3
0101004	39.22563	-76.39548	X	X
0101004B	39.24333	-76.35444	X	
0101005	39.20167	-76.42056		X
0101006	39.19028	-76.4425	X	X
0101007	39.20111	-76.45972	X	X
0101011	39.21526	-76.53044	X	X
0101011A	39.22163	-76.52206	X	X
0101012	39.17194	-76.515		X
0101013	39.15778	-76.49667		X
0101015	39.255	-76.57472	X	X
0101018	39.23056	-76.49861	X	X
0101019	39.25021	-76.5388	X	X
0101020	39.16944	-76.4725	X	X
0101025	39.1975	-76.49361	X	X
0101027	39.22028	-76.50167	X	X
0101028	39.25223	-76.57867	X	X
0101029	39.25917	-76.57194	X	X
0101032	39.188247	-76.516075	X	
0102001	39.31778	-76.36389		X
0102002	39.3025	-76.35778		X
0102006	39.33194	-76.34222		X
0102010	39.37111	-76.33306		X
0102011	39.35139	-76.32861		X
0102012	39.29611	-76.39111		X
0102013	39.31028	-76.40472		X
0102014	39.31139	-76.41778		X
0102016	39.24167	-76.39778		X
0102017	39.25861	-76.44333		X
0102018	39.29	-76.46722		X
0102203HM	39.25806	-76.37083		X
0103002B	39.44737	-76.05577		X
0103005	39.4843	-75.92573		X
0103008	39.47693	-75.94316		X
0103009	39.47549	-75.88633		X
0103010	39.45951	-75.86064		X
0103011	39.45417	-75.86226		X
0103014	39.3688	-75.97934		X
0103015	39.38091	-75.94744		X
0103016	39.36558	-75.89446		X
0201001	39.28528	-76.19944	X	
0201002	39.23452	-76.24043	X	
0201007	39.21306	-76.26361	X	
0201010	39.14669	-76.29369	X	
0202005	39.14667	-76.26111	X	X
0202006	39.13917	-76.26319	X	X
0202010	39.13028	-76.24801	X	X
0202010A	39.1275	-76.25056	X	X
0202014	39.10111	-76.26056	X	
0202014A	39.09414	-76.29811	X	
0202018	39.06667	-76.27306	X	
0301001	39.081833	-76.516167		X
0301001A	39.07133	-76.51428		X
0301001C	39.07593	-76.50838		X
0301003	39.06889	-76.48111	X	X
0301003C	39.07639	-76.49167		X

Station	Latitude	Longitude	Chapter 2	Chapter 3
0301004	39.0775	-76.481111	X	X
0301005A	39.0625	-76.46069	X	X
0301005B	39.08056	-76.46347		X
0301005C	39.0875	-76.46083		X
0301006B	39.09609	-76.441		X
0301007	39.08	-76.45542		X
0301009	39.05556	-76.43667	X	X
0301009A	39.05333	-76.45		X
0301011	39.06069	-76.49		X
0301800	39.075	-76.49722		X
0301801	39.08444	-76.45792		X
0301802	39.08667	-76.44889		X
0302005	39.13306	-76.43861	X	X
0302006A	39.15306	-76.465		X
0302006B	39.15556	-76.46833		X
0302007	39.125	-76.42639	X	
0302008	39.085	-76.41611	X	
0302009	39.15556	-76.40917	X	
0302009A	39.07694	-76.37333	X	
0302009B	39.10472	-76.37417	X	
0302009C	39.14306	-76.3725	X	
0302010	39.03889	-76.40611	X	
0302011A	39.01444	-76.39389		X
0302011B	39.0225	-76.38944		X
0303001A	39.00806	-76.39667		X
0303004	38.98667	-76.4375	X	X
0303005	39.00056	-76.43167	X	X
0303005A	39.00056	-76.42528	X	X
0303006	38.99306	-76.44778	X	X
0303010	38.92833	-76.45833	X	
0303103	38.91333	-76.46417	X	X
0303200	38.96833	-76.45389	X	X
0303202	38.9525	-76.43806	X	
0303204	38.94361	-76.44778	X	X
0303205	38.9334861	-76.430114	X	
0304002A	39.04222	-76.557306		X
0304005	39.04056	-76.53917	X	X
0304008	39.02917	-76.52611	X	X
0304011	39.017	-76.5165	X	X
0304020	38.99333	-76.48333		X
0304028	38.98	-76.45917	X	X
0304029	38.96667	-76.46806	X	X
0304150	39.05996	-76.56642	X	X
0304152	39.06889	-76.57528	X	X
0305004	38.88	-76.46167	X	
0305005	38.86336	-76.4502	X	
0305006	38.89828	-76.4358	X	
0305201	38.87071	-76.4813	X	
0305204	38.8875	-76.4825	X	
0306002	38.95	-76.53637	X	X
0306006	38.93556	-76.52306	X	X
0306011	38.920417	-76.505944	X	X
0306012	38.91904	-76.50109	X	X
0306013A	38.9269444	-76.488328	X	X
0306015	38.90917	-76.50417	X	X

Station	Latitude	Longitude	Chapter 2	Chapter 3
0306016	38.91056	-76.49389	X	X
0306017	38.915	-76.48944	X	X
0306104	38.93056	-76.48389	X	X
0306106	38.94222	-76.52194	X	X
0306110	38.95083	-76.55778	X	X
0306111	38.93472	-76.53778	X	X
0306115	38.90278	-76.50833	X	X
0306115A	38.90111	-76.49972	X	X
0306203	38.93167	-76.5175	X	X
0306204	38.93	-76.52167	X	X
0306205	38.94556	-76.54083	X	X
0306206	38.90556	-76.49139	X	X
0306208	38.94396	-76.53286	X	X
0306211	38.94944	-76.54588	X	X
0306212	38.94167	-76.53306	X	X
0306801	38.90917	-76.51083		X
0307011	38.84722	-76.50556	X	X
0307011B	38.85167	-76.50583	X	X
0307012A	38.8647222	-76.490944	X	
0307019	38.88056	-76.51472	X	X
0307108	38.86083	-76.51194	X	X
0307117	38.8775	-76.51833	X	X
0307120	38.88283	-76.5227	X	X
0307120A	38.88833	-76.52556	X	X
0307122	38.88389	-76.52889	X	X
0307122A	38.88694	-76.52889		X
0307204	38.85468	-76.53138	X	X
0307205	38.84444	-76.53472		X
0307206	38.8697	-76.51657		X
0307900	38.86861	-76.51306		X
0307901	38.86806	-76.51111		X
0401002	39.15444	-76.06806	X	X
0402006	39.1225	-76.10306	X	X
0402013A	39.1025	-76.14333	X	X
0402018	39.08083	-76.13222	X	X
0402022	39.08219	-76.11178	X	X
0403001A	39.06774	-76.16425	X	X
0403004A	39.0875	-76.16556		X
0403006A	39.04556	-76.20556	X	X
0403011	39.01611	-76.19639	X	X
0403018A	39.08611	-76.19861	X	X
0403019A	39.0975	-76.20917		X
0403023	39.10083	-76.17639	X	X
0403027	39.11194	-76.16861	X	X
0403028	39.11835	-76.16953		X
0404003	38.9975	-76.165	X	X
0404019A	39.02361	-76.26556	X	
0404020	39.0375	-76.28833	X	
0501001	38.83	-76.48222	X	
0501001B	38.81694	-76.47528	X	
0501001F	38.79583	-76.50889	X	X
0501003	38.77444	-76.53306	X	X
0501003A	38.77044	-76.54922	X	X
0501004	38.77059	-76.56063	X	X
0501004A	38.76689	-76.55466	X	X

Station	Latitude	Longitude	Chapter 2	Chapter 3
0501005	38.75482	-76.54878	X	X
0501006A	38.74194	-76.545	X	X
0501006B	38.73369	-76.53843	X	X
0501007A	38.74028	-76.51417	X	
0501201	38.76863	-76.49506	X	
0501202	38.80306	-76.49889		X
0502004	38.69583	-76.52722		X
0502005	38.68944	-76.52722		X
0601002	38.945	-76.379	X	
0601005	38.83361	-76.39194	X	
0601006	38.86778	-76.37028	X	X
0601007	38.8775	-76.36694	X	X
0601008	38.88361	-76.36639	X	X
0601011	38.91472	-76.3675	X	X
0601016	39.0075	-76.32889	X	X
0601202	38.97146	-76.38857	X	
0801006	38.8432778	-76.230917	X	
0801011	38.82561	-76.23586	X	
0801019A	38.79821	-76.21588		X
0801021	38.79719	-76.19479	X	X
0801022	38.81083	-76.18222	X	X
0801028	38.76458	-76.17628	X	X
0801030	38.75556	-76.17278	X	X
0801033	38.78846	-76.13954	X	X
0801033B	38.77472	-76.15333		X
0801034	38.79306	-76.12722		X
0801060	38.78389	-76.17694	X	X
0801060A	38.77505	-76.17156		X
0801207	38.78883	-76.21619		X
0801209	38.78839	-76.20786		X
0802001	38.84722	-76.20111	X	X
0802004	38.86583	-76.19167	X	X
0802010	38.88806	-76.18028	X	X
0802013A	38.90917	-76.17	X	X
0802014	38.90528	-76.16	X	X
0802017	38.86413	-76.17949	X	X
0802019	38.87389	-76.16583	X	X
0802023	38.88222	-76.1475	X	X
0802026	38.88056	-76.11861	X	X
0802304	38.9076	-76.13197	X	X
0803005	38.89111	-76.20194	X	X
0803005B	38.89889	-76.18778	X	X
0803007	38.90056	-76.23333	X	
0803012	38.953	-76.243139	X	X
0803013	38.96	-76.23778	X	X
0803014B	38.96639	-76.24556	X	X
0803202	38.965169	-76.24333	X	X
0803203	38.97833	-76.24225	X	X
0804002	38.87583	-76.25167	X	
0804014	38.84222	-76.30361	X	
0804019	38.89833	-76.31333	X	X
0804024	38.90861	-76.33861	X	X
0804028	38.91972	-76.31167	X	
0804031	38.93972	-76.31806	X	X
0804035	38.93583	-76.28917	X	X

Station	Latitude	Longitude	Chapter 2	Chapter 3
0804035C	38.9225	-76.28833	X	X
0804801	38.9379	-76.28273		X
0805003T	38.72172	-76.34041		X
0805008T	38.69564	-76.34763		X
0805019T	38.71941	-76.33315		X
0805104	38.76083	-76.37083		X
0805800	38.71887	-76.34438		X
0805801	38.71772	-76.34291		X
0806004	38.7527778	-76.1175	X	X
0806006A	38.73336	-76.13172		X
0806009	38.72579	-76.1415	X	X
0806012	38.70806	-76.13944	X	X
0806014	38.7125	-76.12278	X	X
0806016	38.7033333	-76.146472	X	X
0806018	38.70167	-76.16417	X	X
0806019A	38.70889	-76.16917	X	X
0806022	38.69667	-76.16694	X	X
0806025	38.69325	-76.179444	X	X
0806028	38.66444	-76.18861	X	
0806120	38.70917	-76.17722	X	X
0806801	38.75956	-76.11081		X
0807001	38.77889	-76.25694		X
0807004	38.75583	-76.22722	X	X
0807005	38.76692	-76.22784	X	X
0807006C	38.77139	-76.23111	X	X
0807010	38.73611	-76.21	X	X
0807010A	38.73111	-76.19972		X
0807011	38.745	-76.23778	X	X
0807014	38.745	-76.257	X	X
0807020	38.72	-76.271389	X	X
0807023	38.71	-76.25861	X	
0807026	38.70333	-76.22142	X	X
0807103	38.77139	-76.22111	X	X
0807700	38.76444	-76.24861		X
0808001T	38.71556	-76.32807	X	X
0808012T	38.67611	-76.33417	X	X
0808027	38.67611	-76.27972	X	
0808038	38.71306	-76.30556	X	X
0808039	38.73083	-76.30472	X	X
0808042A	38.74139	-76.31667	X	X
0808044	38.75333	-76.30778	X	X
0808046	38.76556	-76.30028	X	X
0808048	38.77167	-76.28667	X	X
0808048B	38.77833	-76.28972	X	X
0808109	38.79389	-76.27		X
0808110	38.807361	-76.266611		X
0901001	38.54083	-76.67361	X	X
0901002	38.53667	-76.68556	X	X
0901005	38.52611	-76.67	X	X
0901012	38.50286	-76.68177		X
0901013	38.49528	-76.68389	X	X
0901014	38.49139	-76.66806	X	X
0901018	38.46361	-76.66639	X	X
0901019	38.4575	-76.66306	X	X
0901021	38.46389	-76.64972	X	X

Station	Latitude	Longitude	Chapter 2	Chapter 3
0901101A	38.4825	-76.67528	X	X
0901101E	38.49492	-76.65338		X
0902007A	38.43	-76.60889	X	X
0902019	38.38611	-76.53667	X	X
0902020A	38.40806	-76.54444	X	X
0902022	38.38668	-76.54994		X
0902023	38.3822	-76.55722		X
0902023F	38.38143	-76.56022		X
0902105D	38.44722	-76.60361	X	X
0902107A	38.45056	-76.6	X	X
0902109	38.41333	-76.54333	X	X
0903006	38.39	-76.5	X	X
0903100	38.33542	-76.50419	X	X
0903101	38.34778	-76.50556	X	X
0903103	38.35944	-76.47806	X	X
0903106	38.39389	-76.49	X	X
0903107	38.40472	-76.485	X	X
0903201	38.38361	-76.50694	X	X
0903203	38.3415	-76.50039	X	X
0903204	38.35278	-76.4875	X	X
0904005B	38.3125	-76.455	X	X
0904007	38.315	-76.47944	X	X
0904015A	38.32167	-76.45139	X	X
0904104	38.32972	-76.4575	X	X
0904109A	38.33417	-76.44861	X	X
0904300	38.31889	-76.48139	X	X
0905007	38.31717	-76.40069	X	
0905007A	38.30333	-76.35333	X	
1001010	38.62167	-75.98019	X	X
1001017	38.59756	-75.98642	X	X
1001701	38.642139	-75.965194		X
1001702	38.59044	-76.00608		X
1001800	38.63363	-75.9636		X
1001801	38.61186	-75.97031		X
1001802	38.59894	-75.97611		X
1001803	38.58636	-75.97701		X
1002004A	38.574833	-76.026889	X	X
1002005	38.57317	-76.03456	X	X
1002005D	38.58639	-76.03861	X	X
1002005F	38.58528	-76.04306	X	X
1002008	38.58444	-76.055	X	X
1002011	38.59139	-76.08194	X	X
1002013	38.58167	-76.07403	X	X
1002015	38.57361	-76.07111	X	X
1002020	38.59	-76.08639	X	X
1002021	38.58	-76.0625	X	X
1002022	38.591	-76.07467	X	X
1002804	38.5684	-76.0188		X
1003005	38.5831	-76.11305	X	X
1003008B	38.60947	-76.12983	X	
1003011	38.63153	-76.13833	X	X
1003014	38.6305556	-76.116111	X	X
1003015	38.64364	-76.10956	X	X
1003016E	38.59956	-76.11467	X	
1003016H	38.6155	-76.09864	X	X

Station	Latitude	Longitude	Chapter 2	Chapter 3
1003704	38.60064	-76.08464		X
1003707F	38.6261	-76.16775		X
1004003	38.64611	-76.305	X	
1004005	38.66556	-76.27306	X	
1004006	38.65694	-76.23361	X	
1004009	38.63694	-76.21278	X	
1004010	38.65389	-76.18306	X	
1004013	38.65778	-76.15722	X	X
1004015	38.66389	-76.13833	X	X
1006007	38.54528	-76.26444	X	X
1006010C	38.499969	-76.272361	X	X
1006703	38.53861	-76.26111		X
1007002A	38.53083	-76.22556	X	X
1007003	38.54125	-76.24083	X	X
1007006F	38.56667	-76.24583		X
1007009	38.53556	-76.21056	X	X
1007012	38.529528	-76.176583	X	X
1007015	38.5525	-76.2125	X	X
1007016	38.55833	-76.19194	X	X
1007704	38.52028	-76.17556		X
1007705	38.55917	-76.21556		X
1103003	38.18278	-76.54444		X
1103004	38.17528	-76.53889		X
1103004B	38.18011	-76.53556		X
1103004C	38.17356	-76.53342		X
1103005	38.18056	-76.52944		X
1103008	38.23	-76.675		X
1103011	38.14556	-76.54111		X
1103015	38.14778	-76.53361		X
1103201	38.17861	-76.54389		X
1104001	38.16083	-76.51472	X	X
1104001B	38.16583	-76.50722	X	X
1104003	38.15373	-76.50423	X	X
1104003A	38.15444	-76.50139	X	X
1104004	38.14944	-76.50028	X	X
1104012	38.1225	-76.46611	X	X
1104013	38.11139	-76.46361	X	X
1104017	38.15139	-76.47111	X	X
1104017A	38.15472	-76.47028	X	X
1104022A	38.18806	-76.43889	X	X
1104030A	38.16577	-76.41732		X
1104041	38.11	-76.41028		X
1104041A	38.11667	-76.40278		X
1104041B	38.11722	-76.40917		X
1104042	38.123731	-76.41205		X
1104044B	38.12126	-76.39363		X
1104202	38.13806	-76.51417	X	X
1104203	38.13222	-76.505		X
1104205	38.14167	-76.49833	X	X
1104701	38.12361	-76.45833		X
1104701A	38.13167	-76.48194		X
1104702	38.14111	-76.455		X
1104703	38.15	-76.44583		X
1104704	38.15826	-76.44645		X
1104705	38.14556	-76.46111		X

Station	Latitude	Longitude	Chapter 2	Chapter 3
1104706	38.16111	-76.42861		X
1104707	38.16806	-76.45111		X
1104708	38.19444	-76.44944		X
1105008	38.12225	-76.343528		X
1105018	38.119611	-76.341861		X
1105019	38.126556	-76.338556		X
1105020	38.134028	-76.347917		X
1105022A	38.127472	-76.348056		X
1105024	38.119806	-76.353		X
1105208	38.06444	-76.32972	X	X
1105701	38.12194	-76.34611		X
1105701AF	38.119833	-76.35025		X
1105701F	38.119056	-76.345972		X
1105703	38.06194	-76.32556		X
1201002	38.42267	-76.30364		X
1301001B	38.36583	-76.85806		X
1301001C	38.37056	-76.87306		X
1301003A	38.3525	-76.81917		X
1301004A	38.34389	-76.83194		X
1301006	38.34583	-76.85167		X
1301008	38.33194	-76.85917		X
1301014	38.31694	-76.845		X
1301017	38.29694	-76.83806		X
1301018A	38.28944	-76.85	X	X
1301024	38.27667	-76.82417		X
1301024A	38.26528	-76.83917		X
1301027A	38.27028	-76.86333		X
1301033	38.25028	-76.79139		X
1301034	38.24264	-76.78739		X
1301037	38.23306	-76.78139		X
1301045A	38.26528	-76.86389		X
1301202	38.24794	-76.80267		X
1302001	38.29528	-76.72056	X	X
1302004	38.28667	-76.71528	X	X
1302006B	38.27056	-76.71278	X	X
1302009A	38.2525	-76.72917	X	X
1302010	38.25722	-76.73333	X	X
1302014	38.23353	-76.74222	X	X
1302014C	38.235	-76.75347	X	X
1302015	38.22944	-76.72	X	
1302016	38.22472	-76.74472	X	X
1302021B	38.26103	-76.67869		X
1302023A	38.25667	-76.665	X	X
1302024	38.26303	-76.65338	X	X
1302025	38.26361	-76.64944	X	X
1302029A	38.27333	-76.63389	X	X
1302202	38.25472	-76.69333	X	X
1302203	38.23944	-76.69111	X	
1302701	38.24861	-76.68917		X
1302702	38.25778	-76.71306		X
1303801	38.38813	-77.13095		X
1303802	38.35135	-77.20467		X
1304041	38.33972	-76.98		X
1304058	38.20833	-76.96194		X
1304060A	38.23167	-76.94806		X

Station	Latitude	Longitude	Chapter 2	Chapter 3
1304061	38.28583	-76.99444	X	
1304062	38.2275	-76.965		X
1304063	38.31583	-77.0304	X	X
1304070	38.31333	-76.92278		X
1304072	38.30611	-76.90639		X
1304073	38.3029	-76.905		X
1305051	38.13833	-76.72917		X
1305081	38.09825	-76.58167		X
1306076	38.00167	-76.45167		X
1306077A	37.94639	-76.325		X
1401001	38.29083	-76.15611	X	X
1401002	38.30994	-76.16321	X	X
1401004	38.32556	-76.20972	X	X
1401006	38.315	-76.22444	X	X
1401009	38.3375	-76.17694	X	X
1401014	38.35889	-76.21278	X	X
1401015	38.35	-76.22417	X	X
1401020	38.33694	-76.19806	X	X
1401023	38.31806	-76.18083	X	X
1401024	38.29861	-76.20083	X	X
1401030	38.37528	-76.2225	X	X
1402004	38.2475	-76.11667		X
1402008	38.28589	-76.10549		X
1402009	38.29081	-76.12349		X
1402021	38.21556	-76.04444		X
1402033	38.24806	-76.0775		X
1402036	38.195	-76.09361		X
1402201	38.26694	-76.15083		X
1403002	38.28104	-76.20241		X
1403003A	38.29492	-76.21282		X
1403004	38.31103	-76.233		X
1403005	38.341797	-76.241319		X
1404007	38.30213	-76.02859	X	
1404008	38.30938	-76.03808	X	X
1404016A	38.33718	-75.9841	X	
1404031	38.22369	-75.98912	X	
1404036	38.26743	-76.00753	X	
1404038	38.2806	-76.01409	X	
1404106	38.33696	-76.00711	X	X
1405022	38.32913	-75.87819	X	X
1405024	38.35278	-75.88083	X	X
1405025	38.35667	-75.85583	X	X
1405144	38.32139	-75.88944	X	X
1405144A	38.32139	-75.90333	X	
1405144B	38.32361	-75.91111	X	X
1405202	38.26361	-75.92611	X	X
1405208	38.23472	-75.92722	X	
1405701	38.34167	-75.89028		X
1405702	38.29483	-75.92967		X
1406004	38.24444	-75.84278	X	X
1406007	38.25389	-75.80278	X	X
1406008	38.26661	-75.79644		X
1406201	38.23333	-75.86944	X	X
1406208	38.23917	-75.85889	X	X
1406211	38.25222	-75.835	X	X

Station	Latitude	Longitude	Chapter 2	Chapter 3
1601204	37.9967	-76.025		X
1601205	37.99218	-76.01407		X
1601207	37.97216	-76.02258		X
1601208	37.96	-76.0207		X
1601801	37.971389	-76.041556		X
1801010	38.23007	-75.82208	X	X
1801013	38.20352	-75.87949	X	X
1801019	38.22359	-75.86102	X	X
1801108A	38.22338	-75.84269	X	X
1801202	38.21771	-75.94258	X	
1801203	38.22294	-75.88113	X	X
1802101	38.16722	-75.94306		X
1802201	38.04778	-76.01528		X
1803007	38.12278	-75.85056	X	
1803010	38.12028	-75.90889	X	
1803015	38.13194	-75.82167	X	X
1803017	38.09806	-75.88167	X	
1803110	38.1475	-75.81889	X	X
1803112	38.15611	-75.77889	X	X
1803202	38.14167	-75.85111	X	X
1803204	38.13795	-75.80305	X	X
1803205	38.145	-75.92611	X	X
1804001	38.07583	-75.78389	X	X
1804002	38.06778	-75.79833	X	X
1804007	38.0525	-75.87	X	X
1804008	38.05639	-75.80444	X	X
1804200	38.055	-75.82639	X	X
1804201	38.08534	-75.78909		X
1804700B	38.05322	-75.79907		X
1804700C	38.0533	-75.79913		X
1805004	37.97611	-75.865	X	X
1805006	37.96778	-75.87278	X	X
1805202	37.98361	-75.86361	X	X
1805203	37.99694	-75.855	X	X
1805204	37.96111	-75.87917	X	X
1805206	38.00972	-75.84917	X	X
1805207	38.02583	-75.83972	X	X
1805701	37.96722	-75.9025		X
1806008	37.95028	-75.81144	X	X
1806011	37.92972	-75.775	X	
1806113	37.95556	-75.7125	X	
1806201	37.96566	-75.69114	X	X
1806700	37.96667	-75.67833		X
1806702	37.91611	-75.84611		X
1807010	37.98034	-75.63021	X	X
1807012	37.97958	-75.63711	X	X
1807014	37.96323	-75.65088	X	X
1807015	37.95369	-75.64881	X	X
1807111	37.95692	-75.665	X	X

Appendix D: Table of soil classifications

Soil Type	Characteristics
A	Soils having a high infiltration rate (low runoff potential) when thoroughly wet. These consist mainly of deep, well drained to excessively drained sands or gravelly sands. These soils have a high rate of water transmission.
B	Soils having a moderate infiltration rate when thoroughly wet. These consist chiefly of moderately deep or deep, moderately well drained or well drained soils that have moderately fine texture to moderately coarse texture. These soils have a moderate rate of water transmission.
C	Soils having a slow infiltration rate when thoroughly wet. These consist chiefly of soils having a layer that impedes the downward movement of water or soils of moderately fine texture or fine texture. These soils have a slow rate of water transmission.
D	Soils having a very slow infiltration rate (high runoff potential) when thoroughly wet. These consist chiefly of clays that have a high shrink-swell potential, soils that have a high water table, soils that have a claypan or clay layer at or near the surface, and soils that are shallow over nearly impervious material. These soils have a very slow rate of water transmission.
A/D, B/D, & C/D	If a soil is assigned to a dual hydrologic group (A/D, B/D, or C/D), the first letter is for drained areas and the second is for undrained areas. Only the soils that in their natural condition are in group D are assigned to dual classes.

Based on United States Department of Agriculture (USDA) Soil Survey (http://websoilsurvey.nrcs.usda.gov)

Bibliography

Ackerman, D., and S. B. Weisberg. 2003. Relationship between rainfall and beach bacterial concentrations on Santa Monica Bay beaches. Journal of Water and Health 1:85-89.

Anderson, M. L., J. E. Whitlock, and V. J. Harwood. 2005. Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. Applied and Environmental Microbiology 71:3041-3048.

APHA. 1998. Standard Methods for the Examination of Water and Wastewater, 20th Edition. American Public Health Association, Washington, DC. 1325pp.

Ashbolt, N. J., W. O. K. Grabow, and M. Snozzi. 2001. Indicators of microbial water quality. Pages 289-315 *in* L. Fewtrell and J. Bartram, editors. Water Quality: Guidelines, Standards and Health. IWA Publishing, London, UK.

Baquero, F., J. L. Martinez, and R. Canton. 2008. Antibiotics and antibiotic resistance in water environments. Current Opinion in Biotechnology 19:260-265.

Barry, R. G. 2005. Synotpic Climatology. Pages 700-704 *in* J. E. Oliver, editor. Encyclopedia of Climatology. Springer, Netherlands.

Bernhard, A. E., and K. G. Field. 2000. A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Applied and Environmental Microbiology 66:4571-4574.

Brasher, C. W., A. DePaola, D. D. Jones, and A. K. Bej. 1998. Detection of microbial pathogens in shellfish with multiplex PCR. Current Microbiology 37:101-107.

Burkhardt, W., K. R. Calci, W. D. Watkins, S. R. Rippey, and S. J. Chirtel. 2000. Inactivation of indicator microorganisms in estuarine waters. Water Research 34:2207-2214.

Campos, C. J. A., S. R. Kershaw, and R. J. Lee. 2013. Environmental Influences on Faecal Indicator Organisms in Coastal Waters and Their Accumulation in Bivalve Shellfish. Estuaries and Coasts 36:834-853.

Carney, J. F., C. E. Carty, and R. R. Colwell. 1975. Seasonal occurrence and distribution of microbial indicators and pathogens in the Rhode River of Chesapeake Bay. Applied and Environmental Microbiology 30:771-780.

CBP. 2003. Chesapeake Bay Segments. The Chesapeake Bay Program. Map.

CDC. 2014. Recreational Water Associated Disease Outbreaks — United States, 2009–2010. Centers for Disease Control, USA. 6-10pp.

Chigbu, P., S. Gordon, and T. Strange. 2004. Influence of inter-annual variations in climatic factors on fecal coliform levels in Mississippi Sound. Water Research 38:4341-4352.

Chigbu, P., S. Gordon, and P. B. Tchounwou. 2005. The seasonality of fecal coliform bacteria pollution and its influence on closures of shellfish harvesting areas in Mississippi Sound. International Journal of Environmental Research and Public Health 2:362-373.

Comte, J., and P. A. del Giorgio. 2010. Linking the patterns of change in composition and function in bacterioplankton successions along environmental gradients. Ecology 91:1466-1476.
Craun, M. F., G. F. Craun, R. L. Calderon, and M. J. Beach. 2006. Waterborne outbreaks reported in the United States. Journal of Water and Health 4 Suppl 2:19-30.
Cronin, T. M., G. S. Dwyer, T. Kamiya, S. Schwede, and D. A. Willard. 2003. Medieval Warm Period, Little Ice Age and 20th century temperature variability from Chesapeake Bay. Global and Planetary Change 36:17-29.
Crump, B. C., G. W. Kling, M. Bahr, and J. E. Hobbie. 2003. Bacterioplankton community shifts in an arctic lake correlate with seasonal changes in organic matter source. Applied and Environmental Microbiology 69:2253-2268.
Curriero, F. C., J. A. Patz, J. B. Rose, and S. Lele. 2001. The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948-1994. American Journal of Public Health 91:1194-1199.
Davies, C. M., J. A. H. Long, M. Donald, and N. J. Ashbolt. 1995. Survival of fecal microorganisms in marine and fresh-water sediments. Applied and Environmental Microbiology 61:1888-1896.
DiDonato, G. T., J. R. Stewart, D. M. Sanger, B. J. Robinson, B. C. Thompson, A. F. Holland, and R. F. Van Dolah. 2009. Effects of changing land use on the microbial water quality of tidal creeks. Marine Pollution Bulletin 58:97-106.
Ducklow, H. W. 1982. Chesapeake Bay nutrient and plankton dyanimcs. 1. Bacterial biomass and production during spring tidal destratification in the York River, Virginia, estuary Limnology and Oceanography 27:651-659.
Eleria, A., and R. M. Vogel. 2005. Predicting fecal coliform bacteria levels in the Charles River, Massachusetts, USA. Journal of the American Water Resources Association 41:1195-1209.
Elsner, J. B. 2003. Tracking hurricanes. Bulletin of the American Meteorological Society 84:353-356.
EPA. 1976. Quality Criteria for Water. EPA-440976023, Environmental Protection Agency. 534pp.
EPA. 1986. Ambient water quality criteria for bacteria - 1986. Tech. Memo. 440/5-84-002, Environmental Protection Agency. 18pp.
EPA. 1991. Guidance for Water Quality-Based Decisions: The TMDL Process. . 440/4-91-001, United States Environmental Protection Agency, Office of Water. 62pp.
EPA. 2010a. Chesapeake Bay Phase 5.3 Community Watershed Model. EPA 903S10002 - CBP/TRS-303-10, U.S. Environmental Protection Agency, Chesapeake Bay Program Office, Annapolis, Maryland.
EPA. 2010b. Predictive Tools for Beach Notification Volume I: Review and Technical Protocol. EPA-823-R-10-003, US Environmental Protection Agency. 71pp.
Erkenbrecher, C. W., Jr. 1981. Sediment bacterial indicators in an urban shellfishing subestuary of the lower Chesapeake Bay. Applied and Environmental Microbiology 42:484-492.

Fadrosh, D. W., B. Ma, P. Gajer, N. Sengamalay, S. Ott, R. M. Brotman, and J. Ravel. 2014. An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. Microbiome 2.

FAO. 2011. Risk assessment of *Vibrio parahaemolyticus* in seafood: interpretative summary and technical report. Page 193 *in* M. R. A. S. N. Food and Agriculture Organization, editor.

FDA. 2013. National Shellfish Sanitation Program (NSSP) Guide for the Control of Molluscan Shellfish: 2013 Revision.438.

FDA. 2015. National Shellfish Sanitation Program (NSSP) Guide for the Control of Molluscan Shellfish: 2015 Revision.438.

Ferguson, C. M., B. G. Coote, N. J. Ashbolt, and I. M. Stevenson. 1996. Relationships between indicators, pathogens and water quality in an estuarine system. Water Research 30:2045-2054.

Firth, D. 1993. Bias reduction of maximum-likelihood estimates. Biometrika 80:27-38.

Fortunato, C. S., and B. C. Crump. 2011. Bacterioplankton community variation across river to ocean environmental gradients. Microbial Ecology 62:374-382.

Fortunato, C. S., L. Herfort, P. Zuber, A. M. Baptista, and B. C. Crump. 2012. Spatial variability overwhelms seasonal patterns in bacterioplankton communities across a river to ocean gradient. Isme Journal 6:554-563.

FWPCA. 1972. Federal Water Pollution Control Act. 33 U.S.C. § 1251 et seq.

Glasoe, S., and A. Christy. 2004. Coastal urbanization and microbial contamination of shellfish growing areas - literature review and analysis. PSAT04-09, Puget Sound Action Team, State of Washington. 29pp.

Gonzalez, R. A., K. E. Conn, J. R. Crosswell, and R. T. Noble. 2012. Application of empirical predictive modeling using conventional and alternative fecal indicator bacteria in eastern North Carolina waters. Water Research 46:5871-5882.

Goyal, S. M., C. P. Gerba, and J. L. Melnick. 1977. Occurrence and distribution of bacterial indicators and pathogens in canal communities along the Texas coast. Applied and Environmental Microbiology 34:139-149.

Grant, S. B., B. F. Sanders, A. B. Boehm, J. A. Redman, J. H. Kim, R. D. Mrse, A. K. Chu, M. Gouldin, C. D. McGee, N. A. Gardiner, B. H. Jones, J. Svejkovsky, and G. V. Leipzig. 2001. Generation of enterococci bacteria in a coastal saltwater marsh and its impact on surf zone water quality. Environmental Science & Technology 35:2407-2416.

Green, C. M., and T. P. Tracy. 2013. Fostering shellfish aquaculture production in Maryland and other states. Department of Legislative Services, Office of Policy Analysis, Annapolis, Maryland. 29pp.

Green, M. B., P. Griffin, J. K. Seabridge, and D. Dhobie. 1997. Removal of bacteria in subsurface flow wetlands. Water Science and Technology 35:109-116.

Hacker, J., G. BlumOehler, I. Muhldorfer, and H. Tschape. 1997. Pathogenicity islands of virulent bacteria: Structure, function and impact on microbial evolution. Molecular Microbiology 23:1089-1097.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd Edition edition. Springer, New York. 745pp.

Heidelberg, J. F., K. B. Heidelberg, and R. R. Colwell. 2002. Seasonality of Chesapeake Bay bacterioplankton species. Applied and Environmental Microbiology 68:5488-5497.

Herbert, P. J. 1980. Atlantic hurricane season of 1979. Monthly Weather Review 108:973-990.

Hood, M. A., G. E. Ness, and N. J. Blake. 1983. Relationship among fecal coliforms, *Escherichia coli*, and *Salmonella spp*. in shellfish. Applied and Environmental Microbiology 45:122-126.

Howell, J. M., M. S. Coyne, and P. L. Cornelius. 1996. Effect of sediment particle size and temperature on fecal bacteria mortality rates and the fecal coliform/fecal streptococci ratio. Journal of Environmental Quality 25:1216-1220.

Huang, J. 2010. Spatial and temporal analysis of fecal coliform distribution in Virginia coastal waters. Dissertation. College of William and Mary. 181pp.

Hunt, D. A. 1977. Indicators of Quality for Shellfish Waters. Pages 337-345 *in* A. W. Hoadley and B. J. Dutka, editors. Bacterial Indicators I Health Hazards Associated With Water. American Society for Testing and Materials.

Ibekwe, A. M., M. Leddy, and S. E. Murinda. 2013. Potential human pathogenic bacteria in a mixed urban watershed as revealed by pyrosequencing. Plos One 8.

Interlandi, S. J., and C. S. Crockett. 2003. Recent water quality trends in the Schuylkill River, Pennsylvania, USA: a preliminary assessment of the relative influences of climate, river discharge and suburban development. Water Research 37:1737-1748.

Jacobs, J. M., M. Rhodes, C. W. Brown, R. R. Hood, A. Leight, W. Long, and R. Wood. 2014. Modeling and forecasting the distribution of *Vibrio vulnificus* in Chesapeake Bay. Journal of Applied Microbiology 117:1312-1327.

Jensen, P., and Y. Su. 1992. Characterization of Selected Public Health Issues in Galveston Bay. GBNEP-21, Galveston Bay National Estuary Program. 147pp.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph. 1996. The NCEP/NCAR 40-year reanalysis project. Bulletin of the American Meteorological Society 77:437-471.

Kan, J. J., B. C. Crump, K. Wang, and F. Chen. 2006. Bacterioplankton community in Chesapeake Bay: Predictable or random assemblages. Limnology and Oceanography 51:2157-2169.

Kelsey, H., D. E. Porter, G. Scott, M. Neet, and D. White. 2004. Using geographic information systems and regression analysis to evaluate relationships between land use and fecal coliform bacterial pollution. Journal of Experimental Marine Biology and Ecology 298:197-209.

Kelsey, R. H., G. I. Scott, D. E. Porter, T. C. Siewicki, and D. G. Edwards. 2010. Improvements to shellfish harvest area closure decision making using GIS, remote sensing, and predictive models. Estuaries and Coasts 33:712-722.

Kembel, S. W., M. Wu, J. A. Eisen, and J. L. Green. 2012. Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. Plos Computational Biology 8.

Kimmel, D. G., W. D. Miller, and M. R. Roman. 2006. Regional scale climate forcing of mesozooplankton dynamics in Chesapeake Bay. Estuaries and Coasts 29:375-387.

Klappenbach, J. A., J. M. Dunbar, and T. M. Schmidt. 2000. rRNA operon copy number reflects ecological strategies of bacteria. Applied and Environmental Microbiology 66:1328-1333.

Klein, S. L., C. K. Gutierrez West, D. M. Mejia, and C. R. Lovell. 2014. Genes similar to the *Vibrio parahaemolyticus* virulence-related genes tdh, tlh, and vscC2 occur in other Vibrionaceae species isolated from a pristine estuary. Applied and Environmental Microbiology 80:595-602.

Knox, A. K., R. A. Dahlgren, K. W. Tate, and E. R. Atwill. 2008. Efficacy of natural wetlands to retain nutrient, sediment and microbial pollutants. Journal of Environmental Quality 37:1837-1846.

Kozich, J. J., S. L. Westcott, N. T. Baxter, S. K. Highlander, and P. D. Schloss. 2013. Development of a dual-Index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Applied and Environmental Microbiology 79:5112-5120.

Krometis, L. A. H., G. W. Characklis, P. N. Drummey, and M. D. Sobsey. 2010. Comparison of the presence and partitioning behavior of indicator organisms and *Salmonella spp*. in an urban watershed. Journal of Water and Health 8:44-59.

Leight, A. K., R. Hood, R. Wood, and K. Brohawn. 2016. Climate relationships to fecal bacterial densities in Maryland shellfish harvest waters. Water Research 89:270-281.

Lemon, S. C., J. Roy, M. A. Clark, P. D. Friedmann, and W. Rakowski. 2003. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. Annals of Behavioral Medicine 26:172-181.

Lipp, E. K., R. Kurz, R. Vincent, C. Rodriguez-Palacios, S. R. Farrah, and J. B. Rose. 2001a. The effects of seasonal variability and weather on microbial fecal pollution and enteric pathogens in a subtropical estuary. Estuaries 24:266-276.

Lipp, E. K., N. Schmidt, M. E. Luther, and J. B. Rose. 2001b. Determining the effects of El Nino - Southern Oscillation events on coastal water quality. Estuaries 24:491-497.

Macconkey, A. 1905. Lactose-fermenting bacteria in faeces. Journal of Hygiene (London) 5:333-379.

Maimone, M., C. S. Crockett, and W. E. Cesanek. 2007. PhillyRiverCast: A real-time bacteria forecasting model and web application for the Schuylkill River. Journal of Water Resources Planning and Management 133:542-549.

Mallin, M. A., S. H. Ensign, M. R. McIver, G. C. Shank, and P. K. Fowler. 2001. Demographic, landscape, and meteorological factors controlling the microbial pollution of coastal waters. Hydrobiologia 460:185-193.

MDDNR. 2015. Harvest Reporting. http://dnr2.maryland.gov/fisheries/Pages/aquaculture/harvest-reporting.aspx. Accessed: 12/30/15.

MDE. 2006. Total maximum daily loads of fecal coliform for the restricted shellfish harvest area in the lower Choptank River mainstem in Dorchester and Talbot Counties, Maryland. Maryland Department of the Environment.

MDE. 2014. Detailed Descriptions of MDE's Shellfish Growing Water Classifications as of November 17, 2014. Maryland Department of Environment, Science Services Administration, MDE website (acessed Feb 11, 2015). 21pp.

MDE. 2016. Fecal Coliform TMDL Watersheds Final. Maryland Department of the Environment. http://www.mde.state.md.us/programs/Water/TMDL/DataCenter/Pages/TMDLMapFecalColiform.aspx. Dataset.

Miller, W. D., and L. W. Harding. 2007. Climate forcing of the spring bloom in Chesapeake Bay. Marine Ecology Progress Series 331:11-22.

Montgomery, M. A., and M. Elimelech. 2007. Water and sanitation in developing countries: Including health in the equation. Environmental Science & Technology 41:17-24.

Morgan, R. C., P. Guerry, and R. R. Colwell. 1976. Antibiotic resistant bacteria in the Chesapeake Bay. Chesapeake Science 17:216-219.

Mothur. 2015. MiSeq SOP. www.mothur.org/wiki/MiSeq_SOP. Accessed: 4/6/2015.

Najjar, R. G., C. R. Pyke, M. B. Adams, D. Breitburg, C. Hershner, M. Kemp, R. Howarth, M. R. Mulholland, M. Paolisso, D. Secor, K. Sellner, D. Wardrop, and R. Wood. 2010. Potential climate-change impacts on the Chesapeake Bay. Estuarine Coastal and Shelf Science 86:1-20.

Newton, A., M. Kendall, D. J. Vugia, O. L. Henao, and B. E. Mahon. 2012. Increasing rates of vibriosis in the United States, 1996-2010: review of surveillance data from 2 systems. Clinical Infectious Diseases 54 Suppl 5:S391-395.

Newton, R. J., J. L. VandeWalle, M. A. Borchardt, M. H. Gorelick, and S. L. McLellan. 2011. Lachnospiraceae and Bacteroidales alternative fecal indicators reveal chronic human sewage contamination in an urban harbor. Applied and Environmental Microbiology 77:6972-6981.

Ning, L., M. E. Mann, R. Crane, T. Wagener, R. G. Najjar, and R. Singh. 2012. Probabilistic projections of anthropogenic climate change impacts on precipitation for the Mid-Atlantic Region of the United States. Journal of Climate 25:5273-5291.

NOAA. 2015. Tides and Currents. https://tidesandcurrents.noaa.gov/. Accessed: 12/22/15.

NOAA. 2016. Regional Land Cover Classification Scheme. https://coast.noaa.gov/data/digitalcoast/pdf/ccap-class-scheme-regional.pdf. Accessed: 8/8/16.

Noble, R. T., and J. A. Fuhrman. 2001. Enteroviruses detected by reverse transcriptase polymerase chain reaction from the coastal waters of Santa Monica Bay, California: low correlation to bacterial indicator levels. Hydrobiologia 460:175-184.

Nordstrom, J. L., M. C. L. Vickery, G. M. Blackstone, S. L. Murray, and A. DePaola. 2007. Development of a multiplex real-time PCR assay with an internal amplification control for the detection of total and pathogenic *Vibrio parahaemolyticus* bacteria in oysters. Applied and Environmental Microbiology 73:5840-5847.

Novotny, V., and H. Olem. 1994. Water Quality: Prevention, Identification, and Management of Diffuse Pollution. Van Nostrand Reinhold, New York. 1,054pp.

Nye, J. A., M. R. Baker, R. Bell, A. Kenny, K. H. Kilbourne, K. D. Friedland, E. Martino, M. M. Stachura, K. S. Van Houtan, and R. Wood. 2014. Ecosystem effects of the Atlantic Multidecadal Oscillation. Journal of Marine Systems 133:103-116.

Panicker, G., and A. K. Bej. 2005. Real-time PCR detection of *Vibrio vulnificus* in oysters: Comparison of oligonucleotide primers and probes targeting vvhA. Applied and Environmental Microbiology 71:5702-5709.

Pereira, C., L. Santos, A. P. Silva, Y. J. Silva, A. Cunha, J. L. Romalde, M. L. Nunes, and A. Almeida. 2015. Seasonal variation of bacterial communities in shellfish harvesting waters: Preliminary study before applying phage therapy. Marine Pollution Bulletin 90:68-77.

Pond, K. 2005. Water recreation and disease: plausibility of associated infections: acute effects, sequalae and mortality. World Health Organization. 260pp.

Purcell, J. E., and M. B. Decker. 2005. Effects of climate on relative predation by scyphomedusae and ctenophores on copepods in Chesapeake Bay during 1987-2000. Limnology and Oceanography 50:376-387.

Ralston, E. P., H. Kite-Powell, and A. Beet. 2011. An estimate of the cost of acute health effects from food- and water-borne marine pathogens and toxins in the USA. J Water Health 9:680-694.

Rawles, D. D., G. Flick, A. Diallo, and R. Croonenberghs. 1995. Growth of Mixed Cultures of Listeria-Monocytogenes and Listeria-Innocua in Blue-Crab (Callinectes-Sapidus) Meat. Journal of Food Protection 58:1268-1270.

Rhodes, M. W., and H. Kator. 1988. Survival of *Escherichia coli* and *Salmonella* spp. in estuarine environments. Applied and Environmental Microbiology 54:2902-2907.

rrnDB. 2015. The Ribosomal RNA Database, version 4.4.4. rrndb.umms.med.umich.edu. Accessed: 11/3/2015.

Sayler, G. S., J. D. Nelson, Jr., A. Justice, and R. R. Colwell. 1975. Distribution and significance of fecal indicator organisms in the Upper Chesapeake Bay. Applied and Environmental Microbiology 30:625-638.

Scallan, E., R. M. Hoekstra, F. J. Angulo, R. V. Tauxe, M. A. Widdowson, S. L. Roy, J. L. Jones, and P. M. Griffin. 2011. Foodborne illness acquired in the United States--major pathogens. Emerging Infectious Diseases 17:7-15.

Schloss, P. D., S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. A. Lesniewski, B. B. Oakley, D. H. Parks, C. J. Robinson, J. W. Sahl, B. Stres, G. G. Thallinger, D. J. Van Horn, and C. F. Weber. 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology 75:7537-7541.

Schwab, K. J. 2007. Are existing bacterial indicators adequate for determining recreational water illness in waters impacted by nonpoint pollution? Epidemiology 18:21-22.

Silva. 2015. Silva Ribosomal RNA Database, version 1.19. www.arb-silva.de. Accessed: 2/18/2015.

St.Laurent, K. A., V. J. Coles, R. J. Wood, and R. R. Hood. 2016. Climate extremes in the Chesapeake Bay: part 1: past, present, and future. Estuaries and Coasts Submitted 5/31/2016.

Staley, J. T., and A. Konopka. 1985. Measurement of *in situ* activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. Annual Review of Microbiology 39:321-346.

Stoddard, S. F., B. J. Smith, R. Hein, B. R. K. Roller, and T. M. Schmidt. 2015. rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. Nucleic Acids Research 43:D593-D598.

Suzuki, M., M. S. Rappe, and S. J. Giovannoni. 1998. Kinetic bias in estimates of coastal picoplankton community structure obtained by measurements of small-subunit rRNA gene PCR amplicon length heterogeneity. Applied and Environmental Microbiology 64:4522-4529.

TNC. 2013. Bathymetry of the Chesapeake Bay. The Nature Conservancy. Map.

Trtanj, J., L. Jantarasami, J. Brunkard, T. Collier, J. Jacobs, E. Lipp, S. McLellan, S. Moore, H. Paerl, J. Ravenscroft, M. Sengco, and J. Thurston. 2016. Ch. 6: Climate Impacts on Water-Related Illness. Pages 157-188 The Impacts of Climate Change on Human Health in the United States: A Scientific Assessment. U.S. Global Change Research Program, Washington, DC

Wade, T. J., N. Pai, J. N. Eisenberg, and J. M. Colford, Jr. 2003. Do U.S. Environmental Protection Agency water quality guidelines for recreational waters prevent gastrointestinal illness? A systematic review and meta-analysis. Environmental Health Perspectives 111:1102-1109.

Wagenmakers, E., and S. Farrell. 2004. AIC model selection using Akaike weights. Psychonomic Bulletin and Review 11:192-196.

Walters, S. P., V. P. J. Gannon, and K. G. Field. 2007. Detection of Bacteroidales fecal indicators and the zoonotic pathogens E-Coli O157 : H7, *Salmonella*, and *Campylobacter* in river water. Environmental Science & Technology 41:1856-1862.

Ward, J. H. 1963. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 58:236-244.

Wilkes, G., T. A. Edge, V. P. J. Gannon, C. Jokinen, E. Lyautey, N. F. Neumann, N. Ruecker, A. Scott, M. Sunohara, E. Topp, and D. R. Lapen. 2011. Associations among pathogenic bacteria, parasites, and environmental and

land use factors in multiple mixed-use watersheds. Water Research 45:5807-5825.
Winslow, E. A., and L. T. Walker. 1907. Note on the fermentative reactions of the *B. coli* group. Science 26:797-799.
Wood, R. 2000. Synoptic scale climatic forcing of multispecies fish recruitment patterns in Chesapeake Bay. Dissertation. College of William and Mary, Virginia Institute of Marine Sciences.
Wright, A. C., R. T. Hill, J. A. Johnson, M. C. Roghman, R. R. Colwell, and J. G. Morris, Jr. 1996. Distribution of *Vibrio vulnificus* in the Chesapeake Bay. Applied and Environmental Microbiology 62:717-724.
Zhu, J. H., and X. Z. Liang. 2013. Impacts of the Bermuda High on regional climate and ozone over the United States. Journal of Climate 26:1018-1032.