ABSTRACT

DEVELOPING MATHEMATICAL MODELS FOR PREDICTING *Escherichia coli* CONCENTRATIONS AT WISCONSIN BEACHES USING REGRESSIONS

By Sashidhar R. Yerram

The traditional methods to quantitatively determine *Escherichia coli* concentrations in beach water are the membrane filtration and defined substrate methods. These methods take 18-24 hours for enumeration of *E. coli* concentrations. This may result in improper beach closures and openings, as authorities base their decisions on previous day *E. coli* concentrations. To overcome these problems mathematical models were developed, using the data collected from the 2007 swimming season, to predict the *E. coli* concentrations using various explanatory variables. Beaches from Douglas, Ashland, Bayfield and Door Counties, Wisconsin were chosen for developing the mathematical models. Mathematical models were developed using the United States Environmental Protection Agency (USEPA) “Virtual Beach” software, an application that uses multiple linear regressions.

The developed mathematical models were tested with model fitting (i.e. model is fitted with *E. coli* concentrations used for developing the predictive model). Mathematical models for six beaches were able to predict the $\log_e$ *E. coli* concentrations with 100% accuracy. Explanatory variables that were included in the predictive models were unique for each beach. Overall it was concluded that predictive models should be beach specific. In future more robust mathematical models should be developed using larger data sets and mathematical models should be tested with real time *E. coli* concentrations.

DEVELOPING MATHEMATICAL MODELS FOR PREDICTING *Escherichia coli* CONCENTRATIONS AT WISCONSIN BEACHES USING REGRESSIONS

by

Sashidhar R. Yerram

A Thesis Submitted
In Partial Fulfillment of the Requirements
For the Degree of

Master of Science-Biology

at

The University of Wisconsin Oshkosh
Oshkosh WI 54901-8621

August 2009

COMMITTEE APPROVAL

Advisor

8/10/09 Date Approved

Member

8-11-09 Date Approved

Member

8/14/09 Date Approved

PROVOST AND VICE CHANCELLOR

8/19/2009 Date Approved

FORMAT APPROVAL

Julie Kumbier

8/19/2009 Date Approved

ACKNOWLEDGEMENTS

I want to thank all the people who made my graduate school possible. First, I would like to thank my parents, who always encouraged and supported me for doing my Masters. Then, I would like to thank Dr. Colleen McDermott, my thesis advisor, for her support in the Masters program. I also would like to thank my committee members Dr. Greg Kleinheinz and Dr. Bob Pillsbury. I would like to thank my mother, brother, sister and Dr. Garipalli for their support. I also would like to thank all my friends for their constant support and encouragement they have given to me over the years. Finally, thanks to my B-Tech Biotechnology 2002-2006 batch for being such great friends.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (Continued)

Page

LIST OF TABLES (Continued)

Page

LIST OF TABLES (Continued)

Page

LIST OF FIGURES

LIST OF FIGURES (Continued)

Page

LIST OF FIGURES (Continued)

Page

INTRODUCTION

The United States of America has nearly 23,000 miles of shoreline and more than 5,500 miles of Great Lakes shoreline (3). Beaches are one of the most popular vacation destinations for people in America. The United States Environmental Protection Agency (USEPA) had estimated that Americans make a total of 910 million trips and spend about $44 billion for recreation at these coastal areas (16). Although drinking water quality is rightly given more importance than recreational water quality, outbreaks of water-borne diseases through contact with bathing water motivated the federal government to focus on recreational water quality. Recreational waters include natural waters such as marine and fresh water beaches, as well as artificial waters such as swimming pools and spas. Uses of recreational water include activities ranging from whole body water contact sports, such as swimming, diving, and surfing to non-contact activities such as fishing, walking and picnicking (15). In 2004 the Natural Resources Defense Council (NRDC) estimated that there were at least 18,284 beach closures and water quality advisories in the US during 2003 attributed to microbial contamination, which increased from the previous year by 6,206 beach closure days (8). These closings and advisories lead to economic loss, but hopefully, prevented waterborne illnesses.

Most beach closings and advisories are based on water quality monitoring that detects elevated concentrations of bacteria that indicate the presence of pathogenic microbes derived from sources such as human or animal waste. This waste may enter recreational water through stormwater runoff, leaking septic water systems, and untreated or partially treated waste from sewage plants

Health Concerns at the Beaches?

The primary route of exposure to disease-causing microorganisms in recreational waters is through accidental ingestion of water while swimming (21). The microorganisms causing waterborne diseases are bacteria, viruses, and protozoa. These pathogens potentially spread through the fecal-oral route (21) and cause diseases such as eye, skin, ear, nose, and throat infections, and most importantly, gastrointestinal (GI) disease. The symptoms of GI disease include diarrhea, nausea, vomiting, abdominal cramps, and loss of appetite. During 2001-2002 in the United States, a total of 65 cases of water-borne disease were associated with recreational water usage in 23 states, which was the largest number reported since 1978 (34).

Bacterial Contamination:

In humans, bacteria are harbored on skin and in the digestive tract. Bacterial cells numbers are found to be ten times greater than human cells in the human body (25). Bacterial pathogens such as *E. coli* (enteropathogenic), *Salmonella, Shigella,* and *Vibrio* are associated with waterborne gastroenteritis (21). *Salmonella enterica* serotype *typhimurium* is frequently associated with diarrheal disease syndromes in the United States (27). Most *Salmonella* serotypes cause gastrointestinal illnesses with symptoms of nausea, vomiting, headache and diarrhea. It has been shown that *Salmonella* is present in aquatic habitats and correlates with indicators of fecal pollution. This means that as the concentration of indicator organisms, like total coliforms, fecal coliforms or *E. coli*, increases the chance for the presence of *Salmonella* also increases (23).

Outbreaks of *Shigella* occur commonly through contaminated food or water ingestion; however, it has been seen that *Shigella* also is spread through recreational water usage. In 1982 it was reported that six children who were exposed to recreational water were infected with *Shigella* (29).

Protozoal Contamination:

Protozoans like *Cryptosporidium, Entamoeba histolytica, Giardia lamblia*, and *Balantidium coli* are known to cause waterborne diseases. The parasitic protozoa belonging to the genus *Cryptosporidium* cause cryptosporidiosis, (watery diarrhea, vomiting, loss of appetite and weight loss). *Cryptosporidium* also can cause opportunistic infections in individuals who are immune compromised. In 2006, five states, (Colorado, Illinois, Louisiana, South Carolina, and Wyoming) had cryptosporidiosis outbreaks, associated with recreational water use (20). During 2003-2004, *Cryptosporidium* caused the largest number of waterborne diseases associated with recreational water (number of outbreaks = 11) (30). *Entamoeba histolytica* causes amoebic dysentery and infections in other organs. *Giardia lamblia*, an intestinal parasite, cause diarrhea. *Balantidium coli* cause dysentery and intestinal ulcers.

Viral Contamination:

Enteric viruses are the main group of viruses that cause significant health risk to humans using recreational water, and are usually transmitted through the fecal-oral route. They primarily infect and replicate in the gastrointestinal tract of the host (18). The enteric viruses belong to the families *Adenoviridae, Caliciviridae* (Norwalk virus, astroviruses), *Picornaviridae* (polio viruses, enteroviruses, and hepatitis A viruses)*, and Reoviridae* (reoviruses and rotaviruses) (7). Enteric viruses are shed in high numbers from infected individuals, with about 10^5 to 10^{11} virus particles per gram of stool, and they have a very low infective dose. Outbreaks related to Norovirus and Hepatitis A virus have occurred in recreational waters. These enteric viruses survive for long periods and occur frequently in cold waters; hence, they are potential pathogens during early fall when water temperatures are beginning to decrease (13) .

Beaches Environmental Assessment and Coastal Health Act (BEACH Act):

In 1972 the USEPA began to study the relationship between quality of beach water and health of bathers after they were exposed to beach water. It was observed that swimmers were more likely to develop waterborne diseases than nonswimmers (5, 26). Studies conducted in the 1970s and 1980s showed that swimmers who bathe in water contaminated with sewage are at greater risk for getting gastroenteritis than nonswimmers; and as water quality deteriorates, the chance of getting swimming related illness increases (28). The rate of illness is greater in marine water swimmers than the fresh water swimmers, and most swimming-related diseases are of unknown cause (32). These findings made USEPA recognize the need to improve protection of public health at beaches. The *B*eaches *E*nvironmental *A*ssessment and *C*oastal Health Act (BEACH Act) was passed by the US Congress, amending the Clean Water Act (CWA) on October 10, 2000. The main objectives of the BEACH Act are to monitor public beaches for microbial contamination, inform the public of health concerns at these beaches, reduce the risk of diseases to users of beaches, and promote scientific research for better protection of the health of coastal and inland recreational water users (21).

Wisconsin Waters:

Wisconsin's residents are blessed with many natural resources, including lakes, streams, wetlands, and springs. Wisconsin is bounded by two Great Lakes, Lakes Michigan and Superior. Wisconsin's share of these lakes totals about 2.6 million hectares, and their combined coast line is about 800 miles. The Upper Mississippi River borders about 38,455 hectares and its coast line is about 200 miles. The state has more than 15,000 lakes, almost 3 percent of Wisconsin's area. This is nearly a million acres of lakes and 13,500 miles of navigable streams and rivers (19). Water recreation is a major economic factor in the state providing about $11.5 billion in revenue from tourism and 300,000 full-time job equivalents (19). To monitor these recreational waters, the State of Wisconsin implemented BEACH Act recommendations in 2003. *E. coli* conentrations between 0-234 *E. coli*/100 mL of water indicate that water quality is safe for swimming and the beach remains open. If *E. coli* concentrations are between 235-999 *E. coli*/100 mL of water then a water quality advisory is issued. At this *E. coli* concentration 8/1000 people who would swim would likely develop GI tract disease. *E. coli* concentrations above 1000 *E. coli*/100 mL of water result in a beach closure. At this *E. coli* concentrations, approximately 14/1000 people who would swim would likely develop GI tract disease (2, 6, and 9).

Indicator Organisms:

It is impractical to monitor for all of the pathogens potentially present in water; hence, monitoring of the recreational waters for microbial contamination is conducted by using indicator organisms. These indicator organisms are employed to detect fecal contamination and thereby to correlate with the presence of pathogenic microbes in contaminated water. The ideal indicator organisms should be universally present in the feces of humans and other mammals, be present in large numbers compared to pathogenic microbes, be rapidly detected by easy and inexpensive techniques, should not multiply outside the host when they leave and enter the water, and should have a die-off rate slower than the die-off rate of the pathogens of concern (10). Beginning in 1976, the USEPA recommended using fecal coliforms as indicator organisms for the presence of pathogenic organisms in recreational waters based on studies conducted on Lake Michigan and on inland lakes of Ohio (33). In the mid-1970s and 1980s, studies conducted by Cabelli *et al*. for the USEPA on marine and fresh waters determined a direct linear relation between the frequency of gastrointestinal illnesses in swimmers in marine waters and the enterococci present in the water. They observed no relationship between fecal coliform bacteria and gastrointestinal illnesses in these swimmers (6) The same group performed experiments at fresh water bathing beaches using three bacterial indicators of fecal pollution, *E. coli*, enterococci and fecal coliform bacteria (9). The results showed that there was a strong correlation between swimming-associated gastrointestinal illnesses and *E. coli* and enterococci, but not with fecal coliform bacteria. From these results it was suggested that either *E. coli* or enterococci could be used to determine the water quality of fresh water bathing beaches (12). From 1986, fecal coliforms were replaced by either enterococci or *E. coli* as indicator organisms. The other indicator organisms that have been suggested for use are *Bacteroides* and coliphages (11), but the isolation and identification of these organisms is technically difficult and the coliphages are not host specific.

Techniques for Enumerating Indicator Organisms:

The traditional methods to quantitatively determine *E. coli* concentrations in water are the membrane filtration and defined substrate methods. In the former method, a water sample is filtered through a sterile 0.45μm filter, which retains the bacteria, as their size does not allow passage through the filter. The filter containing the target organism is then placed on a selective medium and incubated for 24 hours to identify the bacteria. After 24 hours the numbers of colonies are counted and expressed in colony forming units (CFU)/100 mL. This method has disadvantages; it is laborious, time consuming and error prone. This method is error prone as the water sample to be tested should be diluted accurately, to get countable number of colonies on the agar plates. The defined substrate method detects the enzymes, β-galactosidase and β-glucuronidase produced by fecal coliforms and *E. coli,* respectively (24), and changes the color of a chromogenic substrate which is added to the water to be analyzed (10). Even though this method is a relatively uncomplicated process, it takes 24 hours of incubation before enumeration of *E. coli* numbers in most probable number (MPN)/100 mL water.

As these techniques take 24+ hours to complete, they may result in unnecessary beach closures and unhealthy reopenings. Beaches may be closed when water quality is acceptable and open when the water quality is unacceptable, because the beach advisories are issued based on yesterday's *E. coli* concentrations. Studies have shown that the lag in obtaining results from water quality tests results in "correct" closures or openings only 50 percent of the time (17). These findings have interested scientists in searching for alternate tools to assess beach water quality. Rapid, sensitive and specific methods like multiplex polymerase chain reaction, gene probe detection, and quantitative polymerase chain reactions (target *lacz* and *uidA* genes from total coliform bacteria and *Escherchia coli,* respectively) (4, 31) are expensive and require more sophisticated technology than the defined substrate and membrane filtration methods. Local health departments cannot afford these techniques, nor provide the personnel to operate the high-tech instruments.

Lately, to overcome the above mentioned obstacles, beach managers have focused on predictive models that could be used to estimate beach water quality based on beach environmental and meteorological data. One of the goals of the USEPA BEACH program is development of mathematical models for real-time quantification of *E. coli* concentrations (16).

Previous Predictive Models:

There are many examples of predictive models developed by scientists in the recent past. During the summer of 2000 at 63rd Street Beach, Chicago, Illinois, 57 water samples were taken and were used to develop a mathematical model and thereby predict the *E. coli* concentrations at the beach. The variables which were included in the best-fit model were wind direction and speed, rain fall, insolation, lake stage, water temperature, and turbidity. These accounted for 71 percent of the variability in the log *E. coli* concentration (22).

At Huntington Beach, Ohio, data collected from 2000-2004 were used for developing a predictive model. The explanatory variables in the model, wave height, weighted rainfall in past 48 hours, and $\log_{10}$ turbidity explained 38 percent of the variability in *E. coli* concentrations. This model predicted 210 total correct predictions, which include correct exceedences and non exceedences, 20 false positive predictions and 18 false negative predictions (14).

Some of the predictive models were based on single variable. In Southern California, at the Santa Monica Bay beaches, a rainfall greater than 6 mm correlated with elevated *E. coli* concentrations and was used to preemptively close beaches (1). Predictive models which were developed using multiple linear regressions have shown satisfactory results in forecasting recreational water quality and have given insight to other researchers to develop predictive models for their local beaches.

A review of the literature describing previously developed predictive models identifies few attempts to develop predictive models for beaches in Wisconsin. This project focuses on developing predictive models for various beaches in Wisconsin, using data collected in summer 2007 from May 15th to August 31st. Specifically, the objectives of this project were:

1) To collect beach environmental weather data during summer of 2007 to be used to predict *E. coli* concentrations at these beaches.

2) To develop mathematical models for predicting the *E. coli* concentrations at the beaches, using multiple linear regression.
3) To identify the environmental factors associated with high concentrations of *E. coli* in beach waters.
4) To compare the environmental factors at beaches from different geographical locations in Wisconsin for similarities and differences in predicting *E. coli* concentrations.
5) To determine whether generalized models can be developed for Wisconsin beaches or whether each beach requires a unique model.

MATERIALS AND METHODS

Study Locations:

Beaches of Lake Superior and Lake Michigan were used for developing the predictive models. Seven beaches located in Ashland, Bayfield, and Douglas Counties, on Lake Superior and eight beaches located in Door County, on Lake Michigan, were used for the study. These beaches were chosen as they were monitored during 2007 beach season and to compare beaches on Lake Michigan and Lake Superior.

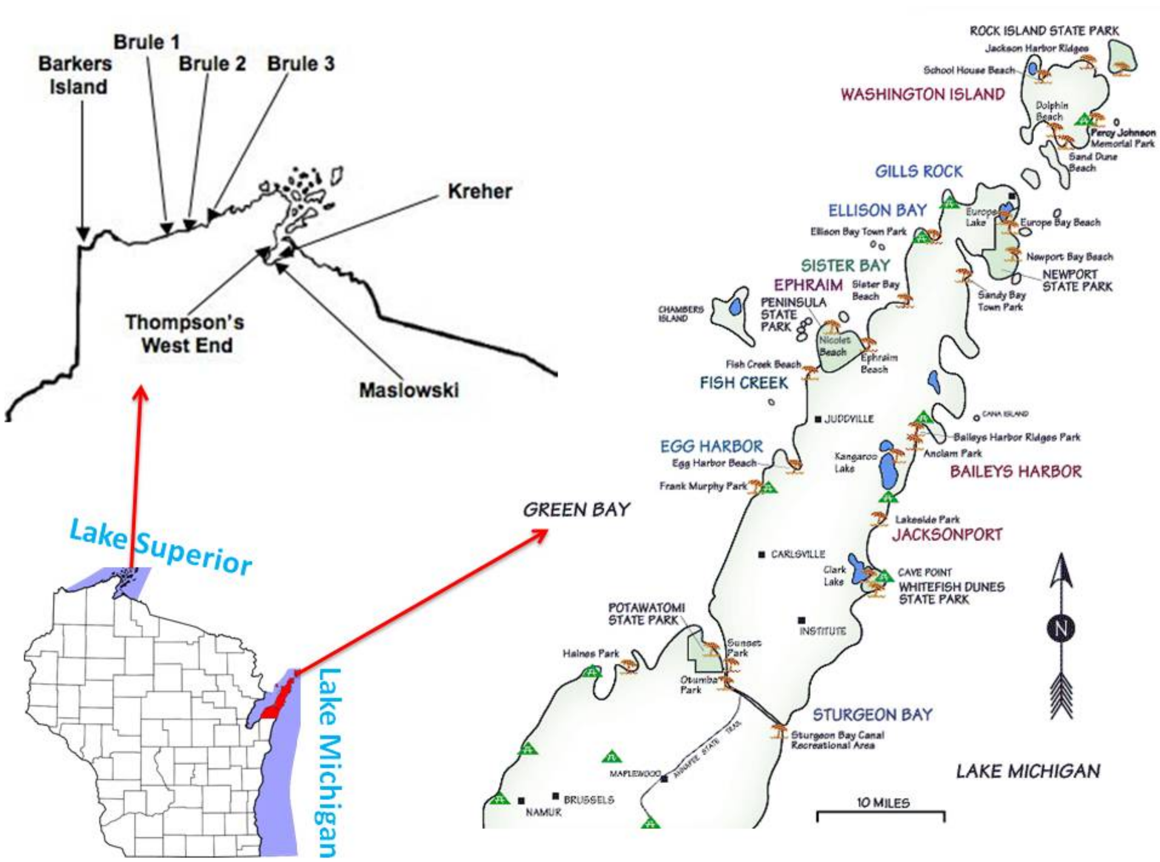

Figure 1: Map showing the beaches at Ashland, Bayfield and Door Counties

Water Sample Collection:

Water samples were collected from the center of beach at a water depth of 24-30 inches, approximately knee depth, as specified by the Wisconsin BEACH monitoring program. Water samples were taken approximately 6-12 inches below the surface of the water. Care was taken to avoid disturbing the bottom sediment at the sampling site. All water samples were collected into sterile, 100 mL polystyrene bottles which were labeled by the sampler with date, time, and location of the sample taken. After the samples were collected they were immediately placed in a cooler at 4ºC and transported to a UW-Oshkosh run, state certified lab where *E. coli* analysis was conducted.

Collection and Compilation of Explanatory Variables:

Explanatory variables were used to predict the response variable (i.e. *E. coli* concentration). These variables were collected either in the field or through weather-related websites (www.wunderground.com). Explanatory variables that were collected and used for model development are:

Wind Direction and Wind Speed:

Wind direction and wind speed were collected in the field with an anemometer, in degrees and miles per hour, respectively. Based on the wind direction in degrees, another explanatory variable was determined. If there was an onshore wind blowing towards the beach, then it was coded as 1, along shore winds were coded as 0.5, and all other wind directions as 0. In addition, wind directions were coded as east = 1, south east = 2, south = 3, south west = 4, west = 5, north west = 6, north = 7, and north east = 8 and were included as variables.

Turbidity and Wave Height:

Turbidity was measured in the field using a turbidometer in Nephelometric Turbidity Units (NTU) and wave height was estimated in inches using a yard stick. Wave height was used as a continuous variable for analysis. To determine the trend in *E. coli* concentrations with wave height, wave heights were grouped and labeled as low (0-<1 feet), medium (≥1-<2, feet) and high (≥2, feet) wave heights. Later, boxplots were created for each group against *E. coli* concentration.

Algae and Gulls:

Presence of algae (i.e. stranded *Cladophora*) on the beach was classified either as none (0), low (1), medium (2), or high (3), based upon amount of algae stranded on the beach. Samplers were given photographs which showed algal amounts on the beaches to use for categorizing the algae amounts on the beaches. Before collection of water samples the number of gulls was counted manually.

Water Temperature, Air Temperature and Barometric Pressure:

Water and air temperature were measured using thermometer in degrees Celsius. Barometric pressure was recorded in inches of mercury, from a weather station website, for the previous day and on the day of sampling. From this, the change in barometric pressure was obtained and included as an explanatory variable.

Rain Fall Variables:

As was evident from previous predictive models, rain fall can be an important variable accounting for elevated concentrations of *E. coli*; hence, many rainfall variables were included in developing these mathematical models (1). Twenty-four hour rain (tfrain) was measured as the amount of rain, in centimeters, that fell the day before sampling. Similarly, 48 (ferain) and 72 (strain) hour rainfall were measured as the amount of rain that fell 2 days and 3 days before sampling, respectively. Event rain (everain) was the amount of rain that fell on the day of sampling, preceding the sampling event. These explanatory variables were collected from the Wunderground weather forecasting website (www.Wunderground.com).

Additional, weighted rainfall variables also were calculated using the variables collected above. Weighted 72 hour rain fall (wstrain) was the amount of rain that fell in the total 72 hour period preceding the sampling, with weight given to previous day (24 hour) rain fall before sampling. Weighted 48 rain fall (wferain) was the amount of rain that fell in the total 48 hour period preceding the sampling, with weight given to previous day rain fall before sampling. The weighted rainfalls are calculated as rainfall weighted 72 hours (wstrain) = (3*(tfrain + everain) + 2*ferain + strain) and rainfall weighted 48 hours (wferain) = (2*(tfrain + everain) + ferain). Combined 72 hour rainfall (cstrain) was the amount of total rain that fell in the 72 hour period preceding sampling, which was calculated as follows: (tfrain + ferain + strain + everain). These variables were sometimes transformed using square root function if they were not linear.

Water Sample Analysis:

After the samples were brought to the lab, the defined substrate (DS) technique, Colilert^{tm} (IDEXX Corp., Portland, ME), was used to analyze the samples for *E. coli* concentration.

Data Management:

All the data collected in the field and through the Wunderground website, explanatory variables and *E. coli* concentrations, were entered into a Microsoft Excel 4.0 worksheet. Date and time of the samples collected were entered in the first and second column of the worksheet, respectively. The third column of the worksheet contained the response variable (*E. coli* concentrations) enumerated using defined substrate technique in the collected water samples. In the remaining columns explanatory variables data was entered.

Predictive Model Development:

Predictive model development included explanatory data analysis, model development and model diagnostics. Explanatory data analysis was accomplished by first examining the summary statistics, median, minimum and maximum *E. coli* concentrations and the number of the days the *E. coli* concentration limit (235 *E. coli*/100mL) was exceeded. Boxplots were constructed for *E. coli* concentrations, which helps in summarizing the statistics graphically. Summary statistics provided information regarding the general water quality of the beach being studied and aided in comparing the water quality of the beaches being studied. After examining the summary statistics of *E. coli* concentrations for the beaches, scatterplots were constructed to examine the relation between the response variable ($\log_e$ *E. coli* concentrations) and explanatory variables. The response variable generally was $\log_e$ transformed because of the wide range of expected *E. coli* concentration values. Each of the explanatory variables (continuous variables) was plotted on the x-axis and response variable on the y-axis. These plots showed whether the relations were linear or nonlinear. If the relation was nonlinear, then the explanatory variable was transformed to linearize the relation. If it couldn't be linearized, then the variable was omitted from the model.

Correlation analysis (Pearson's r correlation analysis [r]), a statistical test, was done to calculate the quantitative measure of the linear association between the explanatory variable and response variable. As the Pearson's correlation coefficient (r) approached 1 or -1, it meant there was a strong positive or negative relation, respectively between the explanatory variable and response variable. If r approached zero, it meant there was a weak relation or no relation between the explanatory variable and response variable. Correlation coefficients (r) were considered statistically significant if the p-value was <0.05. Boxplots and Analysis of variance (ANOVA) were used to understand the relation between categorical variables, such as wave height and wind direction, and the response variable ($\log_e$ *E. coli* concentration). Using these statistics the best explanatory variables which could explain $\log_e$ *E. coli* concentration are determined and mathematical models were developed, using Virtual Beach (VB) software and R program, for each beach. After this, the R squared (R^2) was calculated for the model. This statistic measured the percent of the variation in the response variable ($\log_e$ *E. coli* concentration) that was accounted for by the variation in explanatory variables.

Virtual Beach Software:

Software called "Virtual Beach" (33) was used to develop the predictive models; this software internally uses the multiple linear regressions to develop the predictive models. The general form of a multiple linear regression equation is

$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon$, where y is the response variable (i.e. $\log_e$ *E. coli* concentration), β_0 is the intercept, β_1 is the slope coefficient of first explanatory variable, β_2 is the slope coefficient of second explanatory variable, β_k is the slope coefficient of kth explanatory variable, and ε is the remaining unexplained noise in the data (error). Data in an Excel spread sheet was first imported into the VB software, and then the response variable (*E. coli* concentration) was $\log_e$ transformed. Next, the scatter plots between the $\log_e$ *E. coli* concentration and each of the explanatory variables were inspected for linearity. If the relation was nonlinear, then the explanatory variables are transformed to get the best possible linear relation. The backward elimination method was used for developing the mathematical models. In this method all the explanatory variables were first included for analysis, and then each of the explanatory variables was eliminated based upon their t-value (the value for statistical significance of the coefficient of explanatory variable) and p-value. The explanatory variable with the smallest t-value and $p>0.05$ was eliminated, and the process continued until a three or four variable model was obtained or until only significant variables remain.

Model diagnostics:

1) To estimate whether the developed mathematical model was predicting the known beach *E. coli* concentrations (i.e. data used for the development of predictive model) accurately, the following regression diagnostics were conducted. First, scatter plots were constructed between observed *E. coli* concentration and predicted *E. coli* concentration. These diagrams helped reveal the quality of the relationship between the predicted and observed *E. coli* concentration. Pearson's

correlation coefficient (r) value was calculated to quantify the strength of the relationship.

2) The number of false positives, false negatives, correct predictions above the *E. coli* concentration limit and correct predictions below the *E. coli* concentration limit were computed. Ideally, a good predictive model should be able to predict all the *E. coli* concentrations above the limit and should not predict any false negatives.
3) Model statistics were calculated based on fitting the model to the data used in developing the predictive model.
4) Model predictions were not calculated for data not included in the development of predictive model.

RESULTS

Douglas County Beaches:

In 2003 there were nine beach closings/advisories for Barker's Beach, none for Brule 1 Beach, and one for Brule 2 and Brule 3 Beaches (8). In 2007 the overall $\log_e$ *E. coli* concentrations (MPN ± SD) for the Douglas County Wisconsin beaches (Table 1, 6, 11 and 16) were maximum at Barker's Beach (3.61 ± 0.99) and were similar for all beaches, except Brule 3 Beach (1.73 ± 1.91) ($p = 0.01$, $n = 19$). Mean (MPN ± SD), median, and maximum $\log_e$ *E. coli* for Barker's Beach are 3.61 ± 0.99, 3.38, and 5.73, respectively (Table 1). Explanatory variables which showed significant correlations with $\log_e$ *E. coli* concentrations at Barker's Beach were wind speed ($r = 0.50$, $p < 0.05$) and the interaction term wind speed*onshore wind ($r = 0.58$, $p < 0.01$). These were the only variables included in the mathematical model (Table 2). This mathematical model predicted *E. coli* concentrations exceeding the bathing water quality standards (1 day), as seen in Figure 2 and Table 3. R^2 for this model was 0.33 and the model was statistically significant (Tables 4 and 5).

The mean (MPN ± SD), median, and maximum $\log_e$ *E. coli* for Brule1 Beach are 2.16 ± 2.43, 1.13, and 7.59, respectively, as shown in Table 6. Explanatory variables that showed significant correlations to the $\log_e$ *E. coli* concentrations were water temperature ($r = 0.53$, $p < 0.05$), onshore wind ($r = 0.59$, $p < 0.01$) and wave height ($r = 0.58$, $p < 0.05$). Pearson's correlation coefficients for all the variables can be seen in Table 7. Variables that were included in the mathematical model were onshore wind and wave height (Table 9). The wind variable looks at onshore winds, along-shore winds and offshore winds. When onshore winds are present, then the onshore wind variable is coded as 1. During along-shore winds the variable is coded as 0.5, and during offshore winds it is coded as 0. As it can be seen in Figure 4, on days with along-shore winds *E. coli* concentrations were greater than on days with offshore winds. *E. coli* concentrations were also greater during higher wave height periods (≥2 feet) as compared to low (0 - <1) and medium (≥1 and <2) wave height periods (Figure 3). This mathematical model could predict when the *E. coli* concentrations exceeded the bathing water quality standards, as seen in Figure 5 and Table 9. R^2 for this model was 0.63 and the model was statistically significant (Tables 9 and 10).

The mean (MPN ± SD), median, and maximum $\log_e$ *E. coli* for Brule 2 Beach were 2.16 ± 1.57, 2.00, and 6.42, respectively (Table 11). Explanatory variables which showed significant correlations to the $\log_e$ *E. coli* concentrations were turbidity ($r = 0.57$, $p < 0.001$), onshore wind ($r = 0.52$, $p < 0.01$) and wave height ($r = 0.64$, $p < 0.05$) (Table 12). All significant variables were included in the mathematical model (Figures 6 and 7). This predictive model predicted the only day when the *E. coli* concentrations exceeded the bathing water quality standards, as seen in Figure 8 and Table 13. R^2 for this model was 0.63 and was statistically significant (Tables 14 and 15).

The mean (MPN ± SD), median, and maximum $\log_e$ *E. coli* concentrations for Brule 3 Beach were 1.73 ± 1.91, 1.84, and 7.25, respectively (Table 16). Explanatory variable that showed significant correlation to the $\log_e$ *E. coli* concentrations was turbidity ($r = 0.67$, $p < 0.001$) (Table 17), but variables included in the predictive model were both turbidity and wave height (Figure 9). This model predicted when the *E. coli* concentrations exceeded the bathing water quality limit, as seen in Figure 10 and Table 18. R^2 for this model was 0.66, and the model was statistically significant (Tables 19 and 20).

Barkers Beach:

Predictive equation: Predicted $\log_e$ *E. coli* concentration = 3.35 + 0.2 (Wind speed * onshore wind).

Table 1: Summary statistics of *E. coli* concentrations at Barkers Beach, 2007. Bathing water standard concentration, for *E. coli* is 235 *E. coli*/100mL water.

Number of samples	Daily log *E. coli* concentrations (MPN/100mL)				Number (percent) of days bathing water standard was exceeded
	Minimum	Mean	Median	Maximum	
19	2	3.61	3.38	5.73	1 (5)

Table 2: Pearson's correlation coefficient (r) between $\log_e$ *E. coli* concentrations and explanatory variables for Barkers Beach, numbers in bold are significant ($p<0.05$).

Variable	**R**
Wind direction (degrees)	0.34
Wind direction	0.37
Wind speed	**0.50**
Water temperature	-0.25
Onshore wind	0.31
Wind speed * onshore wind	**0.58**
Air temperature	-0.29
Turbidity	-0.30
Wave height	0.43
Algae	0.08
Gulls	-0.34
24hr rainfall	0.01
48hr rainfall	0.01
72hr rainfall	0.01
Event rainfall	0.03
Yesterday's barometric pressure	0.30
Today's barometric pressure	-0.28

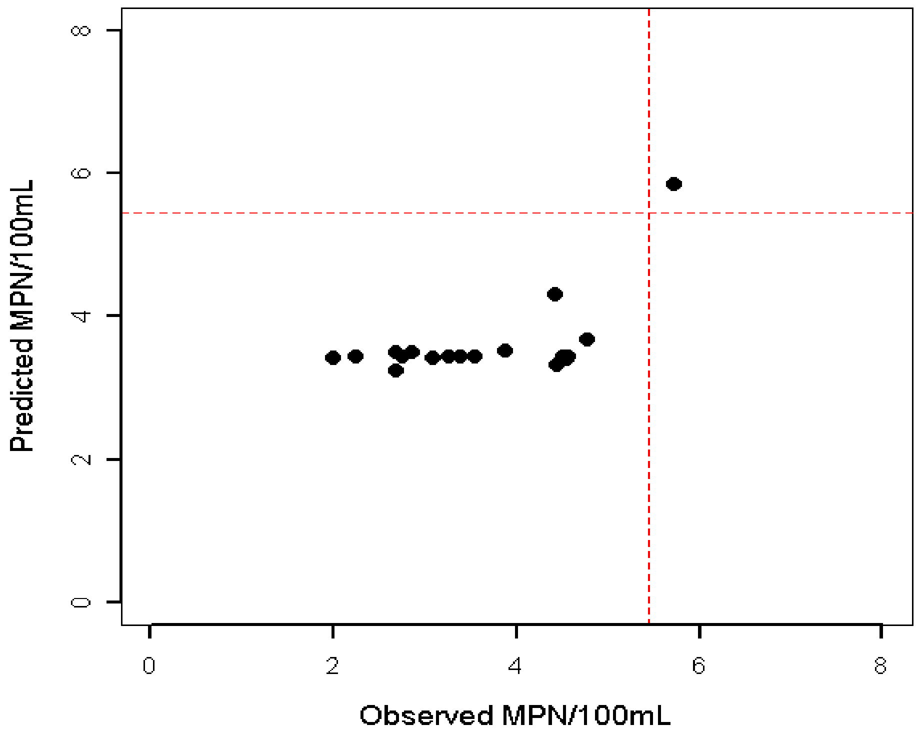

Figure 2: Scatter plot between predicted and observed $\log_e$ *E. coli* concentrations. Red dotted line indicates bathing water quality standard ($\log_e$ 235(5.45) *E. coli*/100mL water).

Table 3: Barkers Beach predictions and accuracy data.

Predictions	Number	Percent
Total predictions	19	-
False positives	0	0
False negatives	0	0
Correct above standard	1	5.3
Correct below standard	18	94.7

Table 4: Barkers Beach mathematical model parameter estimates. p-values (<0.05) of t-tests indicate that each of parameter estimates is statistically different from zero.

Variable	Parameter estimate	Standard error	t-value	p-value
Intercept	3.35	0.212	15.799	0.001
Onshore wind * wind speed	0.20	0.068	2.902	0.001

Table 5: Barkers Beach statistical parameters. R-squared indicates the fraction of variation in *E. coli* concentration explained by the model. p-value (<0.05) indicates the model is statistically significant

Statistic	Values
F-statistic	8.424
Degrees of freedom	17
Residual S.E	0.838
R^2	0.33
p-value	0.04

Brule 1 Beach:

Predictive equation: Predicted $\log_e$ *E. coli* concentration = 0.058 + 5.309 * onshore wind + 1.484 * wave height.

Table 6: Summary statistics of *E. coli* concentrations at Brule 1 Beach, 2007. Bathing water standard concentration, for *E. coli* is 235 *E. coli*/100mL water.

Number of samples	**Daily log *E. coli* concentrations (MPN/100mL)**				**Number (percent) of days bathing water standard was exceeded**
	Minimum	**Mean**	**Median**	**Maximum**	
19	0	2.16	1.13	7.59	2 (10.5)

Table 7: Pearson's correlation coefficient (r) between $\log_e$ *E. coli* concentrations and explanatory variables for Brule1 Beach, numbers in bold are significant ($p<0.05$).

Variable	**R**
Wind direction (degrees)	0.13
Wind direction	0.27
Wind speed	0.04
Water temperature	**0.53**
Onshore wind	**0.59**
Air temperature	0.26
Turbidity	0.47
Wave height	**0.58**
Gulls	-0.03
48hr rainfall	0.22

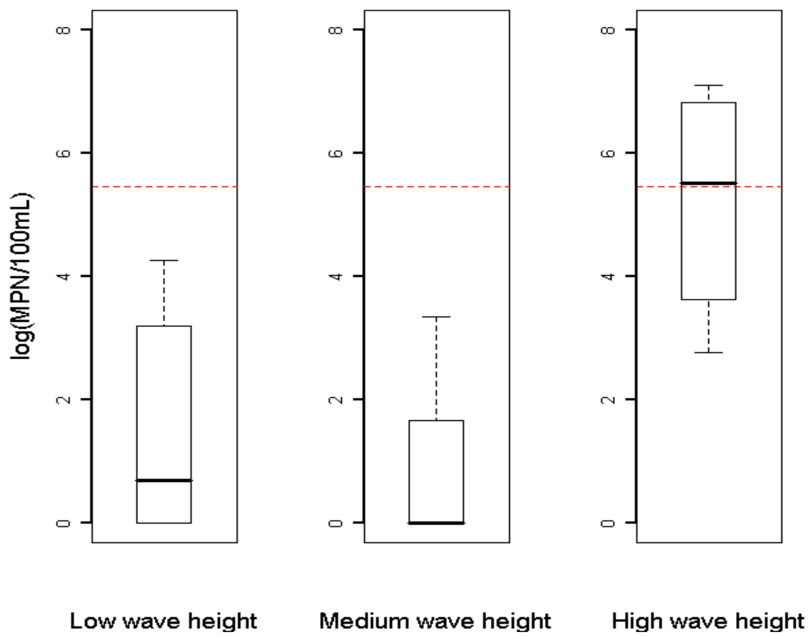

Figure 3: Waves were categorized based on their heights for Brule 1 Beach. Low wave height (<1 feet), medium wave height (≥1 and <2 feet) and high wave height (≥2 feet).

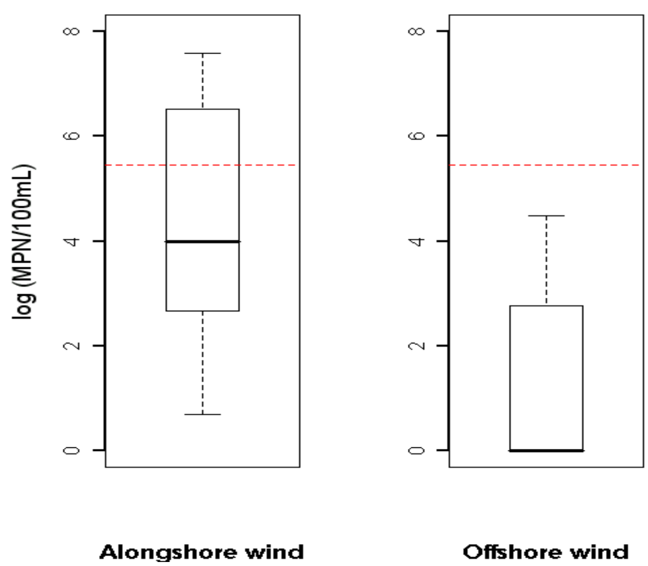

Figure 4: Winds at Brule 1 Beach were categorized and coded as offshore (0), along shore (0.5), and onshore winds (1).

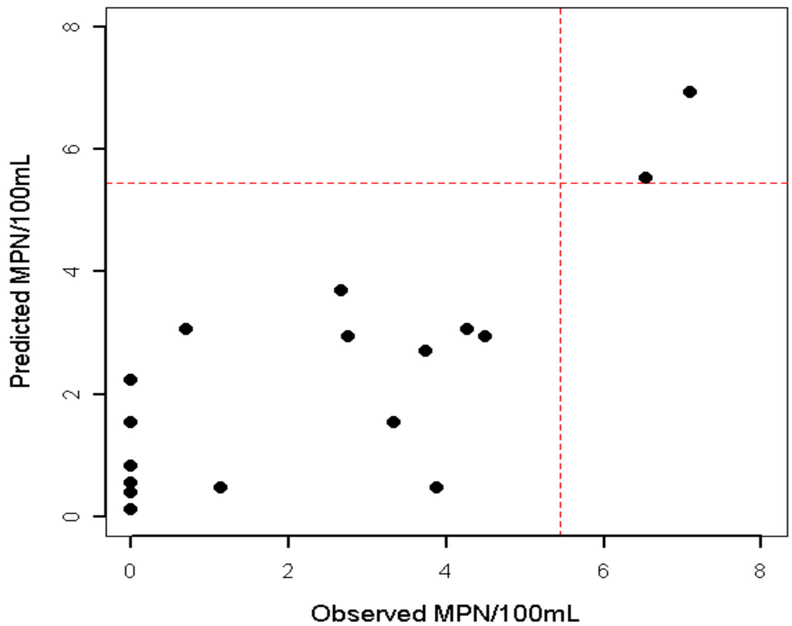

Figure 5: Scatter plot between predicted and observed $\log_e$ *E. coli* concentrations for Brule 1 Beach. Red dotted line indicates bathing water quality standard (log 235(5.45) *E. coli*/100mL water).

Table 8: Brule1 Beach predictions and accuracy data.

Predictions	Number	Percent
Total predictions	19	-
False positives	0	0
False negatives	0	0
Correct above standard	2	10.5
Correct below standard	17	89.5

Table 9: Brule1 Beach mathematical model parameter estimates. p-values of t-tests indicate that each of parameter estimates is statistically different from zero.

Variable	Parameter estimate	Standard error	t-value	p-value
Intercept	0.058	0.545	0.106	0.05
Onshore wind	5.309	1.559	3.404	<0.005
Wave height	1.484	0.429	3.459	<0.005

Table 10: Brule 1 Beach statistical parameters. R-squared indicates the fraction of variation in *E. coli* concentration explained by the model. p-value indicates the model is statistically significant.

Statistic	Values
F-statistic	13.78
Degrees of freedom	16
Residual S.E	1.563
R^2	0.63
p-value	<0.001

Brule 2 Beach:

Predictive equation: Predicted $\log_e$ *E. coli* concentration = 0.660 + 2.139 * onshore wind + 0.873 * wave height + 0.077 * turbidity.

Table 11: Summary statistics of *E. coli* concentrations at Brule 2 Beach, 2007. Bathing water standard concentration, for *E. coli* is log 235 (5.45) *E. coli*/100mL water.

Number of samples	**Daily log *E. coli* concentrations (MPN/100mL)**				**Number (percent) of days bathing water standard was exceeded**
	Minimum	**Mean**	**Median**	**Maximum**	
19	0	2.16	2.0	6.42	1 (5.2)

Table 12: Pearson's correlation coefficient (r) between $\log_e$ *E. coli* concentrations and explanatory variables for Brule 2 Beach, numbers in bold are significant ($p < 0.05$).

Variable	**R**
Wind direction (degrees)	0.22
Wind direction	0.13
Wind speed	0.22
Water temperature	-0.32
Onshore wind	**0.52**
Air temperature	-0.20
Turbidity	**0.57**
Wave height	**0.64**
Gulls	-0.14
One day lag gulls	0.30

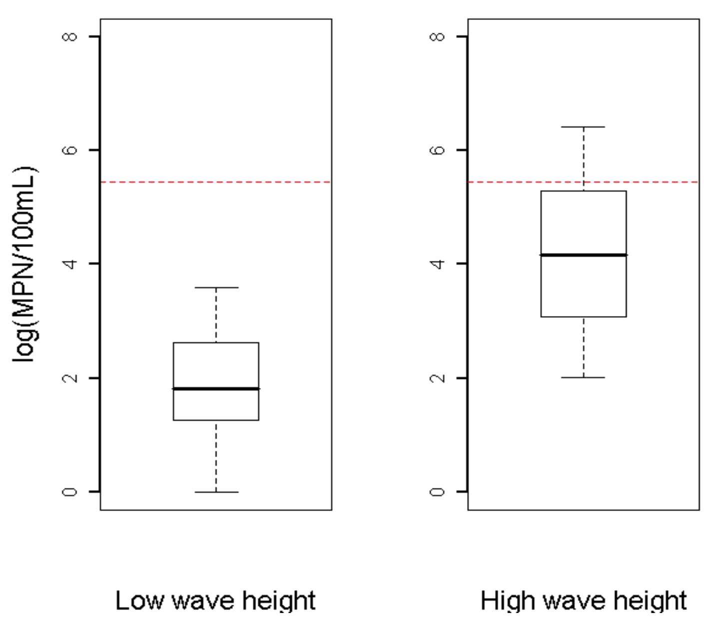

Figure 6: Waves at Brule 2 Beach were categorized as low (<2 feet) and high wave heights (>=2 feet).

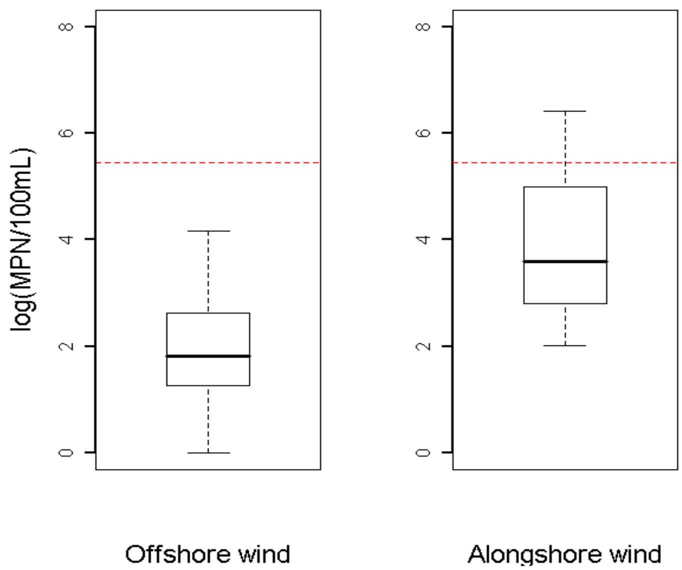

Figure 7: Winds at Brule 2 Beach were categorized and coded, as offshore (0), along shore (0.5), and onshore winds (1).

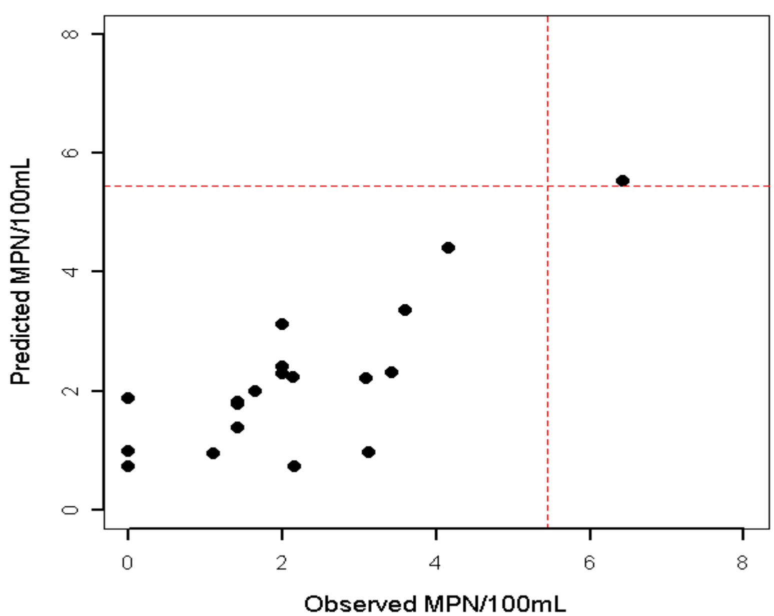

Figure 8: Scatter plot between predicted and observed $\log_e$ *E. coli* concentrations at Brule 2 Beach. Red dotted line indicates bathing water quality standard ($\log_e$ 235(5.45) *E. coli*/100mL water).

Table 13: Brule 2 Beach predictions and accuracy data

Predictions	Number	Percent
Total predictions	19	-
False positives	0	0
False negatives	0	0
Correct above standard	1	5.3
Correct below standard	18	94.7

Table 14: Brule 2 Beach mathematical model parameter estimates. p-values (<0.05) of t-tests indicate that each of parameter estimates is statistically different from zero.

Variable	Parameter estimate	Standard error	t-value	p-value
Intercept	0.660	0.408	1.616	0.12
Onshore wind	2.139	1.444	1.480	0.15
Wave height	0.873	0.339	2.577	0.02
Turbidity	0.077	0.034	2.250	0.04

Table 15: Brule 2 Beach statistical parameters. R^2 indicates the fraction of variation in *E. coli* concentration explained by the model. p-value (<0.05) indicates the model is statistically significant.

Statistic	Values
F-statistic	8.54
Degrees of freedom	15
Residual S.E	1.05
R^2	0.63
p-value	0.001

Brule 3 Beach:

Predictive equation: Predicted $\log_e$ *E. coli* concentration = 0.064 + 0.713 * wave height + 0.027 * turbidity.

Table 16: Summary statistics of *E. coli* concentrations at Brule 3 Beach, 2007. Bathing water standard concentration, for *E. coli* is 235 *E. coli*/100mL water.

Number of samples	Daily log *E. coli* concentrations (MPN/100mL)				Number (percent) of days bathing water standard was exceeded
	Minimum	Mean	Median	Maximum	
19	0	1.729	1.84	7.254	1 (5.2)

Table 17: Pearson's correlation coefficient (r) between log *E. coli* concentrations and explanatory variables for Brule 3 Beach, numbers in bold are significant ($p < 0.05$).

Variable	R
Wind direction (degrees)	0.15
Wind direction	0.13
Wind speed	0.17
Water temperature	0.31
Onshore wind	0.26
Air temperature	0.05
Turbidity	**0.67**
Wave height	0.48
Dogs	-0.26
48hr rainfall	0.29
72hr rainfall	0.19

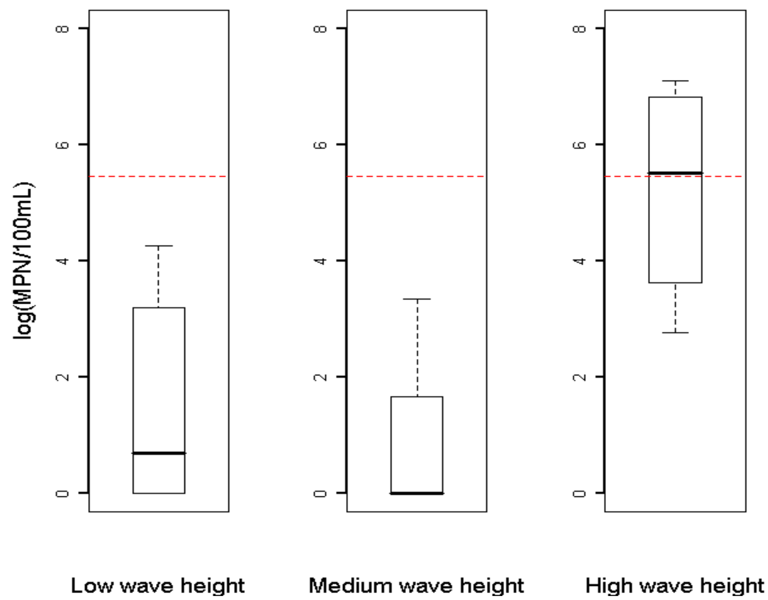

Figure 9: Waves were categorized based on their heights. Low wave height (<1 feet), medium wave height (≥1 and <2 feet) and high wave height (≥2 feet).

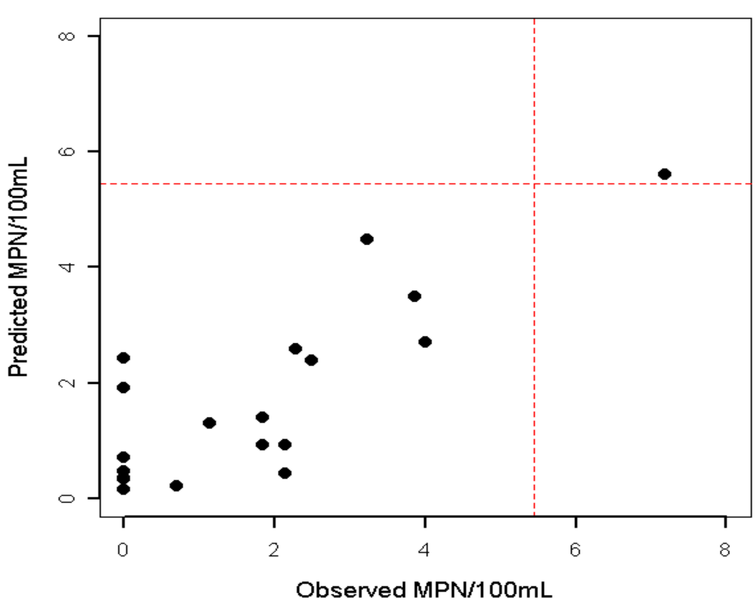

Figure 10: Scatter plot between predicted and observed $\log_e$ *E. coli* concentrations at Brule 3 Beach. Red dotted line indicates bathing water quality standard ($\log_e$ 235(5.45) *E. coli*/100mL water).

Table 18: Brule 3 Beach predictions and accuracy data.

Predictions	Number	Percent
Total predictions	19	-
False positives	0	0
False negatives	0	0
Correct above standard	1	5.3
Correct below standard	18	94.7

Table 19: Brule 3 Beach mathematical model parameter estimates. p-values (<0.05) of t-tests indicate that each of parameter estimates is statistically different from zero.

Variable	Parameter estimate	Standard error	t-value	p-value
Intercept	0.064	0.427	0.151	0.09
Wave height	0.713	0.227	3.136	0.1
Turbidity	0.027	0.006	4.523	<0.005

Table 20: Brule 3 Beach statistical parameters. R-squared indicates the fraction of variation in *E. coli* concentration explained by the model. p-value (<0.05) indicates the model is statistically significant.

Statistic	Values
F-statistic	15.59
Degrees of freedom	16
Residual S.E	1.183
R^2	0.66
p-value	<0.001

Ashland and Bayfield Counties:

In 2003 Kreher Beach had two beach closures/advisories, Maslowski Beach had four beach closures/advisories and Thompson beach had none (8). In 2007 the overall $\log_e$ *E. coli* concentrations (MPN ± SD), which can be seen in Figure 10, Tables 21, 26, and 31, was maximum for Thompson Beach (3.44 ± 2.11) and was significantly different from Kreher Beach (1.31 ± 1.76) ($p < 0.001$, $n = 29$) but not from Maslowski beach (3.32 ± 1.76) ($p > 0.9$, $n = 30$). The $\log_e$ *E. coli* concentrations, at Kreher Beach was significantly different ($p < 0.0003$) from Maslowski Beach.

The mean (MPN ± SD), median, and maximum $\log_e$ *E. coli* concentrations for Thompson's Beach are 3.44 ± 2.11, 3.12, and 7.46, respectively (Table 21). Explanatory variables that showed significant correlation with Thompson Beach's $\log_e$ *E. coli* concentrations were onshore wind direction ($r = 0.63$, $p < 0 .05$), wind speed ($r = -0.44$, $p < 0.05$), event rainfall ($r = 0.70$, $p < 0.05$), √event rainfall ($r = 0.74$, $p < 0.05$), weighted 72hr rainfall ($r = 0.58$, $p < 0.05$), √weighted 72hr rainfall ($r = 0.56$, $p < 0.05$), weighted 48hr rainfall ($r = 0.68$, $p < 0.05$), and √weighted 48hr rainfall ($r = 0.64$, $p < 0.05$) as shown in Table 22. The variables that were included in the predictive model were onshore wind (90° - 180°), √event rainfall and weighted 48hr rainfall. This predictive model could predict all the days (four) when *E. coli* concentrations exceeded the bathing water quality standard ($\log_e$ 235 (5.45) *E. coli*/100mL water), as seen in Figure 10 and Table 23. R^2 for this model was 0.77 and the model was statistically significant ($p < 0.001$), as seen in Table 26.

The mean (MPN ± SD), median, and maximum log *E. coli* concentrations for Maslowski Beach are 3.32 ± 1.76, 3.12, and 7.21, respectively (Table 26). The *E. coli* concentrations exceeded bathing water quality standards for 5 days. Explanatory variables that showed significant association with the log *E. coli* concentrations were 2 days lag algae ($r = 0.54$, $p < 0.05$), gulls ($r = 0.43$, $p < 0.05$), 1 day lag gulls ($r = 0.63$, $p < 0.05$), event rainfall ($r = 0.49$, $p < 0.05$), weighted 72hr rainfall ($r = 0.36$, $p < 0.1$), and weighted 48hr rainfall ($r = 0.39$, $p < 0.05$), which can be seen in Table 27. The variables that were included in the predictive model were 2 days lag algae, 1 day lag gulls, and weighted 48hr rainfall (Table 29). This model predicted, 3 out of 5 days, when *E. coli* concentrations exceeded the bathing water quality standard ($\log_e$ 235 (5.45) *E. coli*/100mL water), as seen in Figure 11 and Table 28.

Kreher Beach had the lowest log *E. coli* concentrations (1.31 ± 1.76) of the three beaches (Table 32). The explanatory variables that showed significant correlation were event rainfall ($r = 0.81$, $p < 0.05$), √event rainfall ($r = 0.81$, $p < 0.05$) weighted 72hr rainfall ($r = 0.66$, $p < 0.05$), weighted 48hr rainfall ($r = 0.66$, $p < 0.05$), and combined 72hr rainfall ($r = 0.58$, $p < 0.05$), as seen in Table 32. The variables that were included in the predictive model are weighted 72hr rainfall and √event rainfall, which could explain the 0.70 (R^2) of total variation in *E. coli* concentrations (Table 35). This model predicted specific *E. coli* concentrations (1 day) when they exceeded the bathing water quality standard and predicted one false positive event (Figure 12 and Table 33).

Thompson Beach:

Predictive equation: Predicted $\log_e$ *E. coli* concentration = 1.939 + 1.824 onshore wind+ 2.146 $\sqrt{}$event rainfall + 0.460 weighted 48hr rainfall.

Table 21: Summary statistics of *E. coli* concentrations at Thompson Beach, 2007. Bathing water standard concentration, for *E. coli* is 235 *E. coli*/100mL water.

Number of samples	**Daily log *E. coli* concentrations (MPN/100mL)**				**Number (percent) of days bathing water standard was exceeded**
	Minimum	**Mean**	**Median**	**Maximum**	
22	0	3.44	3.12	7.46	4 (18)

Table 22: Pearson's correlation coefficient (r) between $\log_e$ *E. coli* concentrations and explanatory variables for Thompson Beach, numbers in bold are significant ($p<0.05$).

Variable	R
Wind direction (degrees)	-0.28
Onshore wind	**0.63**
Wind speed	**-0.44**
Water temperature	-0.01
Air temperature	-0.14
Turbidity	0.37
Long shore current speed	-0.16
Algae	-0.18
24hr rainfall	0.20
48hr rainfall	-0.07
72hr rainfall	-0.30
Event rainfall, √Event rainfall	**0.70 , 0.74**
Yesterday's B.P	0.03
Today's B.P	-0.17
Difference B.P	0.25
Weighted 72hr rainfall, √weighted 72hr rainfall	**0.58, 0.56**
Weighted 48hr rainfall, √weighted 48hr rainfall	**0.68, 0.64**
Combined 72hr rainfall	0.35
Wind vector	0.15

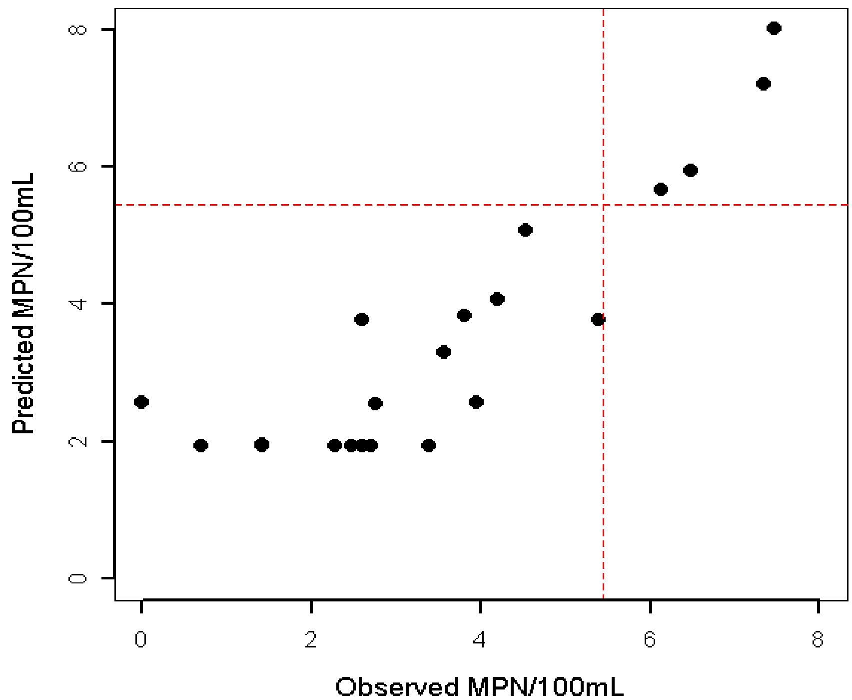

Figure 11: Scatter plot between predicted and observed $\log_e$ *E. coli* concentrations. Red dotted line indicates bathing water quality standard ($\log_e$ 235(5.45) *E. coli*/100mL water).

Table 23: Thompson Beach predictions and accuracy data.

Predictions	Number	Percent
Total predictions	22	-
False positives	0	0
False negatives	0	0
Correct above standard	4	18.2
Correct below standard	18	81.8

Table 24: Thompson Beach mathematical model parameter estimates. p-values (<0.05) of t-tests indicate that each of parameter estimates is statistically different from zero.

Variable	**Parameter estimate**	**Standard error**	**t-value**	**p-value**
Intercept	1.939	0.325	5.972	<0.0001
Onshore wind	1.824	0.506	3.607	0.002
√Event rainfall	2.146	0.777	2.763	0.01
Weighted 48hr rainfall	0.460	0.258	1.785	0.05

Table 25: Thompson Beach statistical parameters. R-squared indicates the fraction of variation in *E. coli* concentration explained by the model. p-value (<0.05) indicates the model is statistically significant.

Statistic	**Values**
F-statistic	20.82
Degrees of freedom	18
Residual S.E	1.079
R^2	0.77
p-value	<0.001

Maslowski Beach:

Predictive equation: Predicted $\log_e$ *E. coli* concentration = 1.321 + 0.895 two days lag algae + 0.025 one day lag gulls + 0.459 weighted 48hr rainfall.

Table 26: Summary statistics of *E. coli* concentrations at Maslowski beach, 2007. Bathing water standard concentration, for *E. coli* is 235 *E. coli*/100mL water.

Number of samples	**Daily log *E. coli* concentrations (MPN/100mL)**				**Number (percent) of days bathing water standard was exceeded**
	Minimum	**Mean**	**Median**	**Maximum**	
30	0	3.32	3.12	7.21	5 (16.7)

Table 27: Pearson's correlation coefficient (r) between $\log_e$ *E. coli* concentrations and explanatory variables for Maslowski Beach, numbers in bold are significant ($p < 0.05$) and * indicates significant ($p < 0.1$).

Variable	R
Wind direction (degrees)	0.02
Wind direction	-0.25
Wind speed	-0.07
Water temperature	-0.20
Air temperature	-0.14
Turbidity	0.04
Wave height	-0.16
Algae	0.18
Two days lag algae	**0.54**
Gulls	**0.43**
One day lag gulls	**0.63**
24hr rainfall	0.12
48hr rainfall	0.14
72hr rainfall	-0.05
Event rainfall	**0.49**
Weighted 72hr rainfall	0.36*
Weighted 48hr rainfall	**0.39**
Combined 72hr rainfall	0.29
Yesterday's barometric pressure	-0.01
Today's barometric pressure	-0.03
Difference of barometric pressure	0.03
Wind vector	-0.14

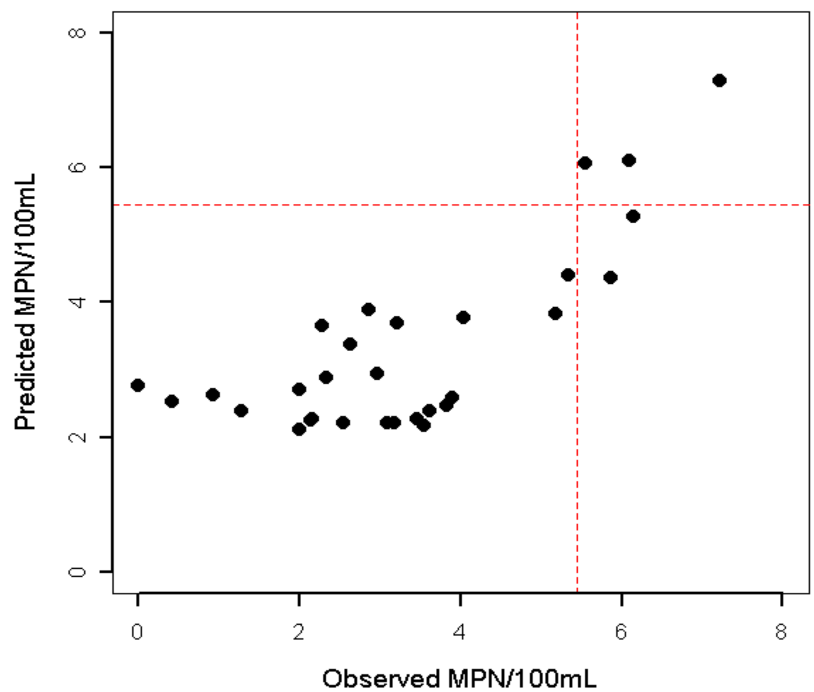

Figure 12: Malowski Beach scatter plot between predicted and observed $\log_e$ *E. coli* concentrations. Red dotted line indicates bathing water quality standard ($\log_e$ 235(5.45) *E. coli*/100mL water).

Table 28: Maslowski Beach predictions and accuracy data.

Predictions	Number	Percent
Total predictions	30	-
False positives	0	0
False negatives	2	6.6
Correct above standard	3	10
Correct below standard	25	83.3

Table 29: Maslowski Beach mathematical model parameter estimates. p-values (<0.05) of t-tests indicate that each of parameter estimates is statistically different from zero

Variable	Parameter estimate	Standard error	t-value	p-value
Intercept	1.321	0.455	2.902	0.001
Two days lag algae	0.895	0.360	2.487	0.01
One day lag gulls	0.025	0.007	3.340	0.001
Weighted 48hr rainfall	0.523	0.232	2.254	0.01

Table 30: Maslowski Beach statistical parameters. R-squared indicates the fraction of variation in *E. coli* concentration explained by the model. p-value (<0.05) indicates the model is statistically significant.

Statistic	Values
F-statistic	12.53
Degrees of freedom	26
Residual S.E	1.192
R^2	0.54
p-value	<0.001

Kreher Beach:

Predictive equation: Predicted $\log_e$ *E. coli* concentration = 0.555 + 0.295 weighted 72 hr rainfall + 3.157 $\sqrt{\text{event rainfall}}$.

Table 31: Summary statistics of *E. coli* concentrations at Kreher beach, 2007. Bathing water standard concentration, for *E. coli* is $\log_e$ 235 (5.45) *E. coli*/100mL water.

Number of samples	Daily log *E. coli* concentrations (MPN/100mL)				Number (percent) of days bathing water standard was exceeded
	Minimum	Mean	Median	Maximum	
29	0	1.31	0.69	7.34	1 (3.4)

Table 32: Pearson's correlation coefficient (r) between $\log_e$ *E. coli* concentrations and explanatory variables for Kreher Beach, numbers in bold are significant ($p < 0.05$)

Variable	**R**
Wind direction (degrees)	-0.10
Wind direction	-0.43
Wind speed	-0.22
Water temperature	0.08
Air temperature	-0.22
Turbidity	-0.04
Wave height	-0.04
Algae	0.28
Gulls	0.14
Long shore current speed	-0.24
Long shore current direction	0.18
24hr rainfall	0.05
48hr rainfall	-0.01
72hr rainfall	0.17
Event rainfall, √event rainfall	**0.81, 0.81**
Weighted 72hr rainfall	**0.66**
Weighted 48hr rainfall	**0.66**
Combined 72hr rainfall	**0.58**
Y. barometric pressure	0.01
T. barometric pressure	-0.12
Difference of barometric pressure	0.15

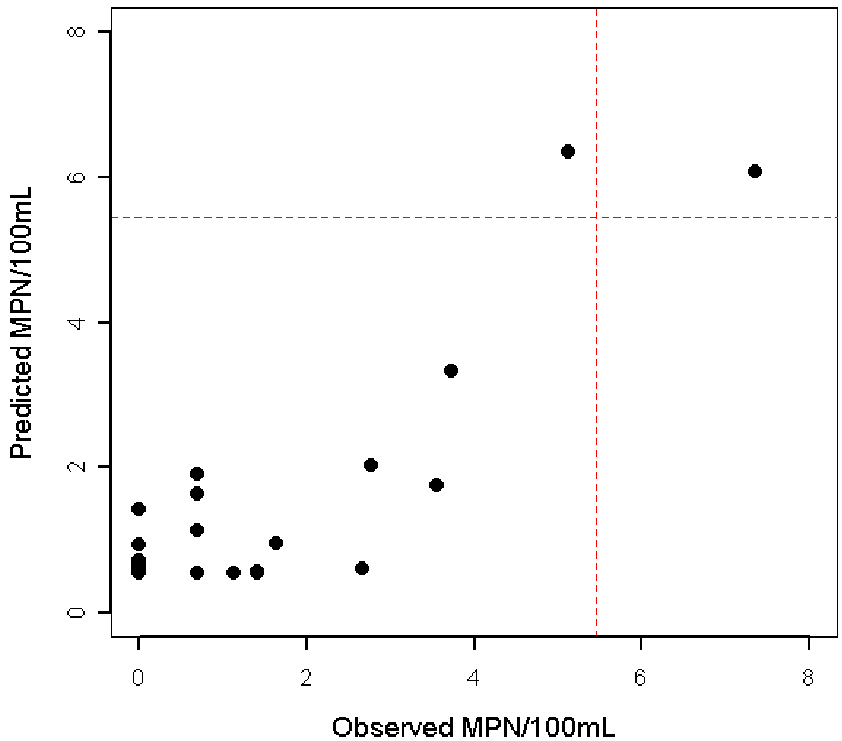

Figure 13: Kreher Beach scatter plot between predicted and observed $\log_e$ *E. coli* concentrations. Red dotted line indicates bathing water quality standard ($\log_e$ 235(5.45) *E. coli*/100mL water).

Table 33: Kreher Beach predictions and accuracy.

Predictions	Number	Percent
Total predictions	29	-
False positives	1	3.4
False negatives	0	0
Correct above standard	1	3.4
Correct below standard	27	93.2

Table 34: Kreher Beach mathematical model parameter estimates. p-values (<0.05) of t-tests indicate that each of parameter estimates is statistically different from zero

Variable	Parameter estimate	Standard error	t-value	p-value
Intercept	0.555	0.213	2.605	0.01
Weighted 72hr	0.295	0.120	2.453	0.02
rainfall	3.157	0.610	5.170	<0.001
√Event rainfall				

Table 35: Kreher Beach statistical parameters. R-squared indicates the fraction of variation in *E. coli* concentration explained by the model. p-value (<0.05) indicates the model is statistically significant.

Statistic	Values
F-statistic	34.76
Degrees of freedom	26
Residual S.E	0.953
R^2	0.70
p-value	<0.001

Door County:

In 2003 there were 6 beach closings/advisories at Baileys Harbor Park Beach, 1 beach closing/advisory at Murphy Park Beach, and 8 beach closings/advisories at Otumba Park Beach. Closings/advisories at Lakeside Park Beach were 2, 13 at Whitefish Dunes State Park Beach, 7 at Fish Creek Beach and 2 at Ephraim Beach. In 2005 and 2006 combined beach closings/advisories at Bailey's Harbor Park Beach were 6, 10 at Murphy Park Beach, 18 at Otumba Park Beach, 6 at Lakeside Park Beach, 15 at Whitefish Dunes State Park Beach, 8 at Fish Creek Beach and 3 at Ephraim Beach. Eight beaches from Door County, Wisconsin, were studied during the 2007 swimming season. Sunset Park Beach water had the highest mean $\log_e$ *E. coli* concentrations (MPN ± SD) 3.01 ± 1.94 and was significantly different from Ephraim (1.26 ± 1.45) ($p < 0.0001$, n = 41), Whitefish Dunes State Park (2.00 ± 1.82) ($p = 0.04$, n = 54), and Murphy Park Beaches (1.64 ± 1.73) ($p < 0.05$, n = 46). The lowest mean loge *E. coli* concentrations were observed at Ephraim Beach (1.26 ± 1.45) and was significantly different from Lakeside Park Beach (2.58 ± 1.74) ($p < 0.05$, n = 47), Sunset Park Beach (3.01 + 1.94) ($p < 0.0001$, n = 4) and Otumba Park Beach (2.56 ± 1.56) ($p = 0.01$, n = 44).

The mean (MPN ± SD), median and maximum $\log_e$ *E. coli* concentrations for Bailey's Harbor Beach are 2.32 ± 1.89, 2.3, and 7.8, respectively (Table 36). Explanatory variables that showed significant correlation with the $\log_e$ *E. coli* concentrations were onshore wind (r = 0.62), wave height (r = 0.72) and turbidity (r = 0.46), as seen in Table 37. The variables that were included in the predictive model were wave height and onshore wind (Table 39). This mathematical model twice predicted when the $\log_e$ *E. coli* concentrations exceeded the bathing water quality standard (log_e 235 (5.45) *E. coli*/100mL water) and predicted 1 false negative, as seen in Figure 15 and Table 38.

The mean (MPN ± SD), median and maximum log_e *E. coli* concentrations for Murphy Park Beach are (1.64 ± 1.72), 1.27 and 7.59, respectively (Table 41). Explanatory variables that showed significant correlation with the log_e *E. coli* concentrations were square root of long shore current speed, which was transformed to linearize the variable, 1 day lag gull numbers and turbidity, and were included in the mathematical model (Table 42 and 44). This predictive model predicted the only day when the *E. coli* concentrations exceeded the bathing water quality standards, which is seen in Figure 20 and Table 43.

Otumba Park Beach mean (MPN ± SD), median and maximum *E. coli* concentrations were (2.56 ± 1.56), 2.55 and 6.08, respectively (Table 46). Explanatory variables that showed significant correlation with the log_e *E. coli* concentrations were onshore wind and turbidity (Table 47). The variables that were included in the predictive model were onshore wind, turbidity and water temperature (Table 49). This model once predicted when *E. coli* concentrations exceeded the bathing water standards and predicted one false negative, as seen in Figure 24 and Table 48.

Lakeside Park Beach mean (MPN ± SD), median, and maximum log *E. coli* concentrations were (2.58 ± 1.74), 2.58 and 5.37, respectively (Table 51). Explanatory variables that showed significant correlation with the log_e *E. coli* concentrations were long shore current speed, turbidity, wave height, algae, and barometric pressure (Table 52). Variables that were included in the predictive model were turbidity, wave height, and algae (Table 54). This predictive model predicted two false positive $\log_e$ *E. coli* concentrations (Figure 27 and Table 53).

The mean (MPN ± SD), median, and maximum $\log_e$ *E. coli* concentrations at Whitefish Dunes State Park Beach were (2.00 ± 1.82), 1.84, and 7.25, respectively (Table 56). Explanatory variables that showed significant correlation with the $\log_e$ *E. coli* concentrations were long shore current speed, wave height, and gull numbers (Table 57). Variables that were included in the predictive model were $\log_e$ water temperature, and wave height (Table 59). This predictive model predicted 2 days when the $\log_e$ *E. coli* concentrations exceeded the bathing water quality standards and predicted two false negative $\log_e$ *E. coli* concentrations, as seen in Figure 30 and Table 58.

The mean (MPN ± SD), median and maximum $\log_e$ *E. coli* concentrations for Fish Creek Beach were 2.00 ± 1.43, 1.84, and 6.03, respectively. Explanatory variables that showed significant correlation with $\log_e$ *E. coli* concentrations were air temperature, turbidity, wave height, and long shore current speed. A mathematical model developed for this beach was not included because the best model developed was not statistically significant.

The mean (MPN ± SD), median and maximum $\log_e$ *E. coli* concentrations for Ephraim Beach are 1.26 ± 1.45, 0.69, and 4.93, respectively. Explanatory variables that showed significant correlation with the $\log_e$ *E. coli* concentrations were air temperature, wave height, turbidity, algae, and the day's barometric pressure. A mathematical model developed for this beach was not included because the best model developed was not statistically significant.

Bailey's Harbor Beach:

Predictive equation: Predicted $\log_e$ *E. coli* concentration = 0.67 + 1.87*wave height + 1.69*onshore wind.

Table 36: Summary statistics of *E. coli* concentrations at Bailey's Harbor Beach, 2007. Bathing water limit concentration, for *E. coli* is ($\log_e$ 235 (5.45) *E. coli*/100mL water).

Number of samples	Daily log *E. coli* concentrations (MPN/100mL)				Number (percent) of days bathing water standard was exceeded
	Minimum	Mean	Median	Maximum	
40	0	2.3	2.3	7.8	3 (7.5)

Table 37: Pearson's correlation coefficient (r) between log_e *E. coli* concentrations and explanatory variables for Bailey's Harbor beach, numbers in bold are significant ($p < 0.05$).

Variable	R
Wind direction (degrees)	-0.20
Wind direction	-0.15
Wind speed	0.12
Long shore current speed	0.32
Long shore current direction	0.28
Water temperature	0.10
Onshore wind	**0.62**
Air temperature	-0.25
Turbidity	0.46
Wave height	**0.72**
Algae	0.17
One day lag algae	-0.29
Gulls	0.45
One day lag gulls	0.04
Two days lag gulls	-0.07
24hr rainfall	0.03
48hr rainfall	-0.15
72hr rainfall	0.02
Event rainfall	0.21
Weighted 48hr rainfall	-0.13
Weighted 72hr rainfall	-0.03
Combined 72hr rainfall	-0.06

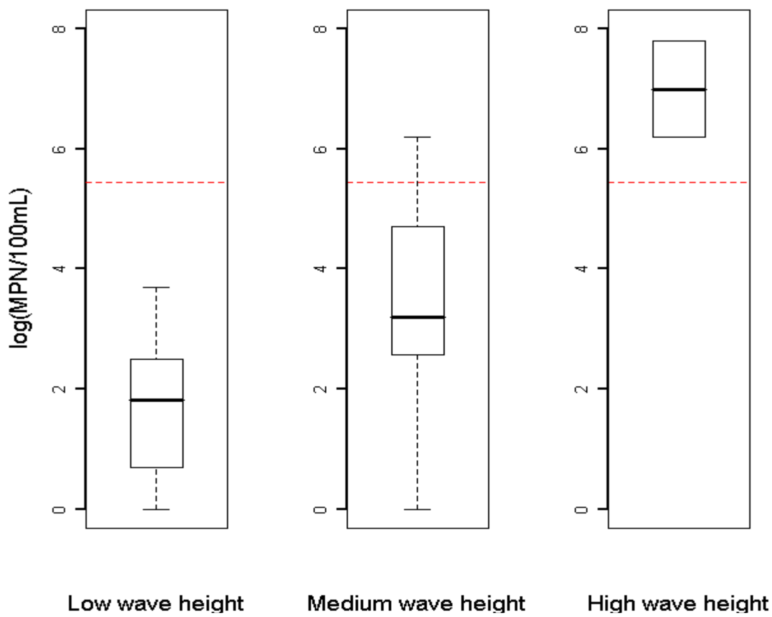

Figure 14: Waves were categorized based on their heights for Bailey's Harbor Beach. Low wave height (<1 feet), medium wave height (≥1 and <2 feet) and high wave height (≥2 feet).

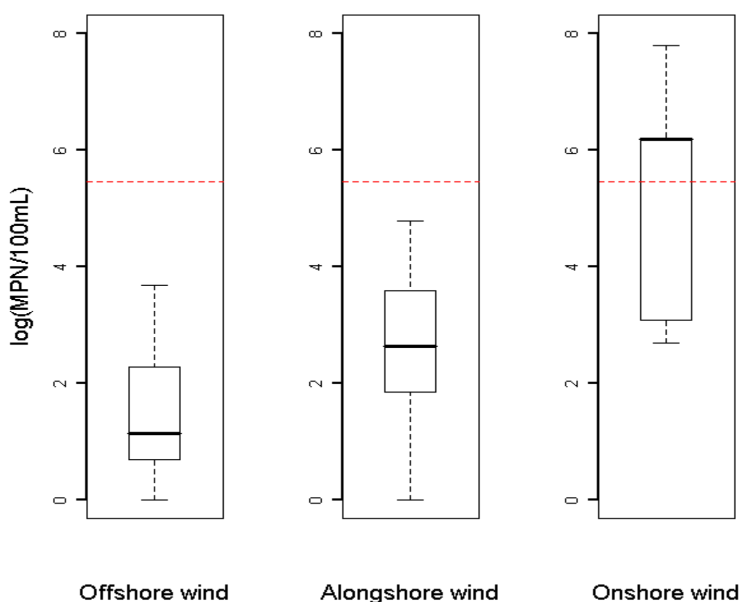

Figure 15: Winds at Bailey's Harbor Beach were categorized and coded as offshore (0), along shore (0.5) and onshore winds (1).

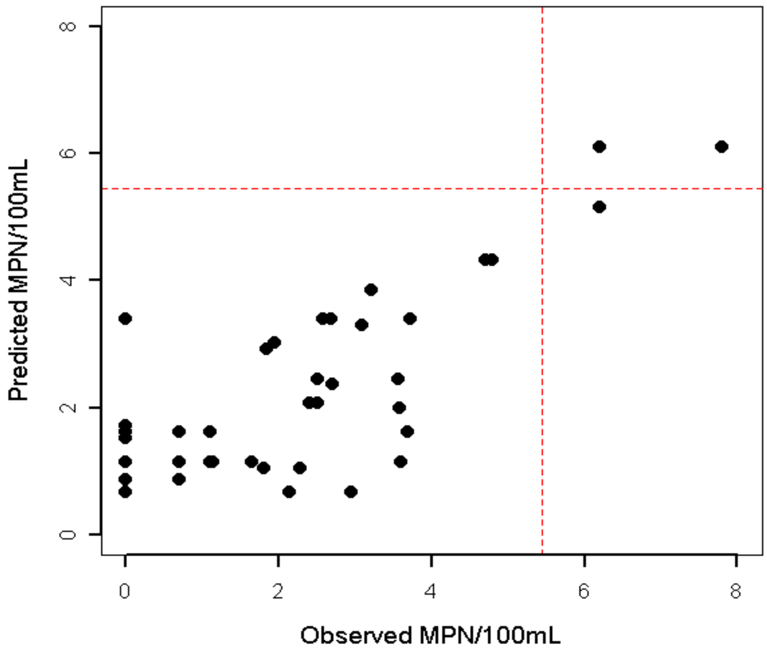

Figure 16: Scatter plot between predicted and observed $\log_e$ *E. coli* concentrations. Red dotted line indicates bathing water quality limit ($\log_e$ 235(5.45) *E. coli*/100mL water).

Table 38: Bailey's Harbor Beach predictions and accuracy data.

Predictions	Number	Percent
Total predictions	40	-
False positives	0	0
False negatives	1	2.5
Correct above standard	2	5
Correct below standard	37	92.5

Table 39: Bailey's Harbor Beach mathematical model parameter estimates. p-values (<0.05) of t-tests indicate that each of parameter estimates was statistically different from zero.

Variable	Parameter estimate	Standard error	t-value	p-value
Intercept	0.67	0.30	2.26	0.02
Wave height	1.87	0.43	4.31	0.001
Onshore wind	1.69	0.67	2.50	0.01

Table 40: Bailey's Harbor Beach statistical parameters. R-squared indicates the fraction of variation in *E. coli* concentration explained by the model. p-value (<0.05) indicate the model was statistically significant.

Statistic	Values
F-statistic	26.74
Degrees of freedom	37
Residual S.E	1.24
R^2	0.57
p-value	0.0001

Murphy Park Beach:

Predictive equation: Predicted $\log_e$ *E. coli* concentration = -0.31 + 0.28* turbidity + 0.14* one day lag gulls + 0.35*√long shore current speed.

Table 41: Summary statistics of *E. coli* concentrations at Murphy Park Beach, 2007. Bathing water limit concentration is $\log_e$ 235 (5.45) *E. coli*/100mL water.

Number of samples	Daily log *E. coli* concentrations (MPN/100mL)				Number (percent) of days bathing water standard was exceeded
	Minimum	Mean	Median	Maximum	
46	0	1.64	1.27	7.59	1 (2.1)

Table 42: Pearson's correlation coefficient (r) between $\log_e$ *E. coli* concentrations and explanatory variables for Murphy Park Beach, numbers in bold are significant ($p<0.05$).

Variable	**R**
Wind direction (degrees)	-0.05
Wind direction	0.03
Wind speed	0.01
Long shore current speed	0.39
√long shore current speed	**0.42**
Long shore current direction	-0.05
Water temperature	0.01
Onshore wind	-0.11
Air temperature	-0.16
Turbidity	**0.36**
Wave height	**0.33**
Algae	0.07
One day lag algae	0.16
Gulls	0.18
One day lag gulls	**0.56**
24hr rainfall	-0.06
48hr rainfall	-0.14
72hr rainfall	-0.11
Event rainfall	-0.03
Weighted 48hr rainfall	-0.15
Weighted 72hr rainfall	-0.16
Combined 72hr rainfall	-0.17

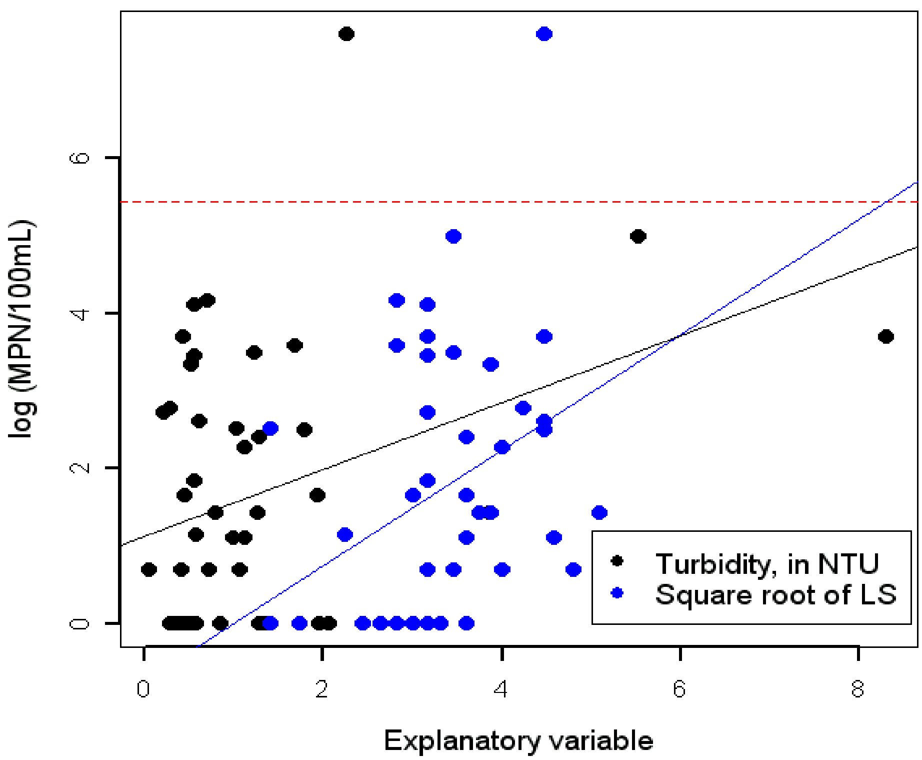

Figure 17: Scatter plots between response variable and square root of long shore current speed (√LS), shown in blue and turbidity shown in black. Red dotted line indicates, bathing water limit ($\log_e$ 235 (5.45) *E. coli*/100mL water).

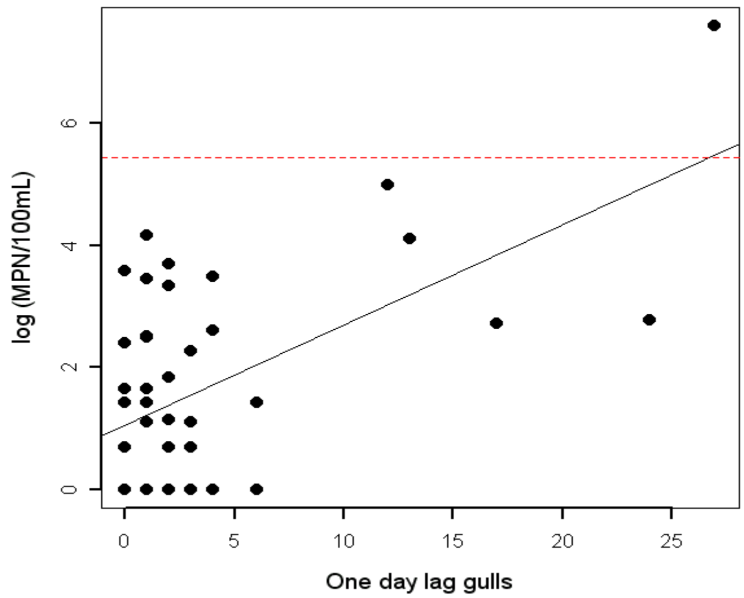

Figure 18: Scatter plot between response variable and 1 day lagged gull numbers. Red dotted line indicates, bathing water limit ($\log_e$ 235 (5.45) *E. coli*/100mL water).

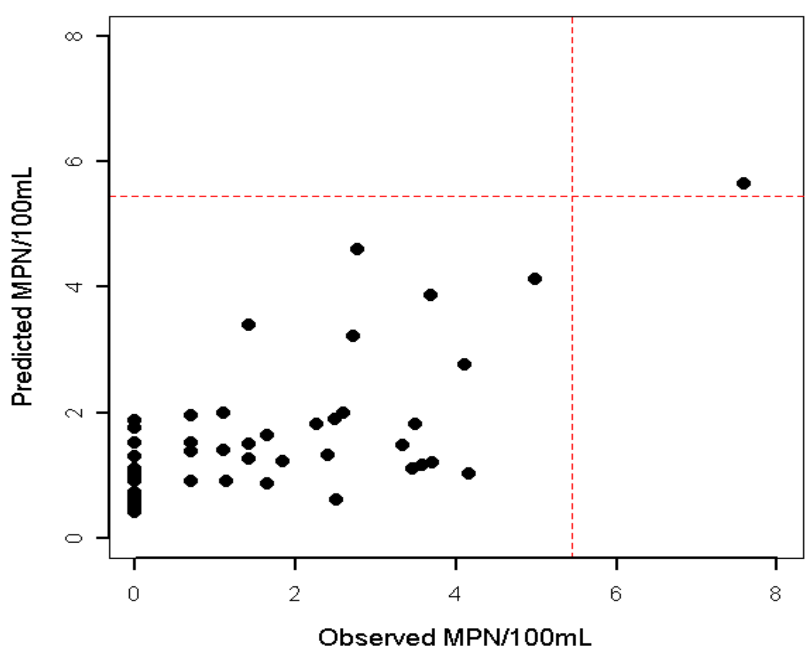

Figure 19: Scatter plot between predicted and observed $\log_e$ *E. coli* concentrations. Red dotted line indicates bathing water limit standard (log 235(5.45) *E. coli*/100mL water).

Table 43: Murphy Park Beach predictions and accuracy data.

Predictions	Number	Percent
Total predictions	46	-
False positives	0	0
False negatives	0	0
Correct above standard	1	2.2
Correct below standard	45	97.8

Table 44: Murphy Park Beach mathematical model parameter estimates. p-values (<0.05) of t-tests indicate that each of parameter estimates is statistically different from zero.

Variable	Parameter estimate	Standard error	t-value	p-value
Intercept	-0.31	0.70	-0.47	0.1
Turbidity	0.28	0.15	1.9	0.05
One day lag	0.14	0.04	3.9	0.0001
gulls	0.35	0.23	1.5	0.1
√long shore current speed				

Table 45: Murphy Park Beach statistical parameters. R-squared indicates the fraction of variation in *E. coli* concentration explained by the model. p-value (<0.05) indicates the model is statistically significant.

Statistic	Values
F-statistic	10.54
Degrees of freedom	42
Residual S.E	1.35
R^2	0.43
p-value	0.001

Otumba Park Beach:

Predictive equation: Predicted $\log_e$ *E. coli* concentration = -1.03 + 0.99* turbidity + 2.15* onshore wind + 0.09*water temperature.

Table 46: Summary statistics of *E. coli* concentrations at Otumba Park Beach, 2007. Bathing water limit concentration, for *E. coli* is $\log_e$ 235 (5.45) *E. coli*/100mL water.

Number of samples	Daily log *E. coli* concentrations (MPN/100mL)				Number (percent) of days bathing water standard was exceeded
	Minimum	Mean	Median	Maximum	
44	0	2.56	2.55	6.08	2 (4.5)

Table 47: Pearson's correlation coefficient (r) between $\log_e$ *E. coli* concentrations and explanatory variables for Otumba Park Beach, numbers in bold are significant ($p < 0.05$).

Variable	**R**
Wind direction (degrees)	0.17
Wind direction	0.24
Wind speed	-0.05
Long shore current speed	0.04
Long shore current direction	-0.36
Water temperature	0.27
Onshore wind	**0.57**
Air temperature	-0.09
Turbidity	**0.41**
Wave height	0.18
Algae	0.11
Gulls	0.18
One day lag gulls	0.02
24hr rainfall	-0.17
48hr rainfall	-0.05
72hr rainfall	-0.26
Event rainfall	0.36
Weighted 48hr rainfall	-0.05
Weighted 72hr rainfall	-0.17
Combined 72hr rainfall	-0.18

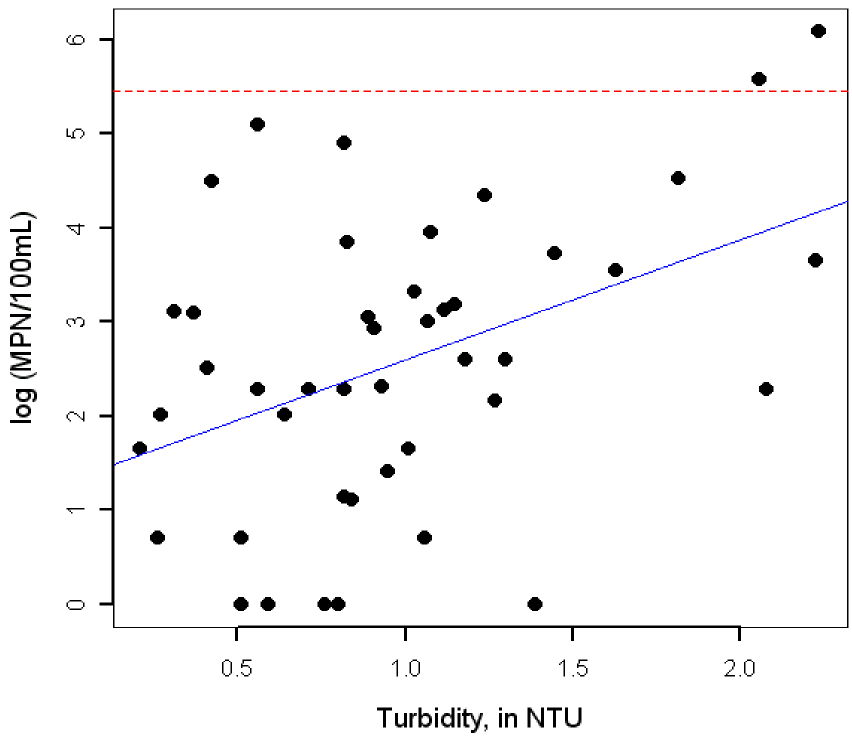

Figure 20: Scatter plot between response variable and turbidity, in NTU. Red dotted line indicates, bathing water standard ($\log_e$ 235 (5.45) *E. coli*/100mL water).

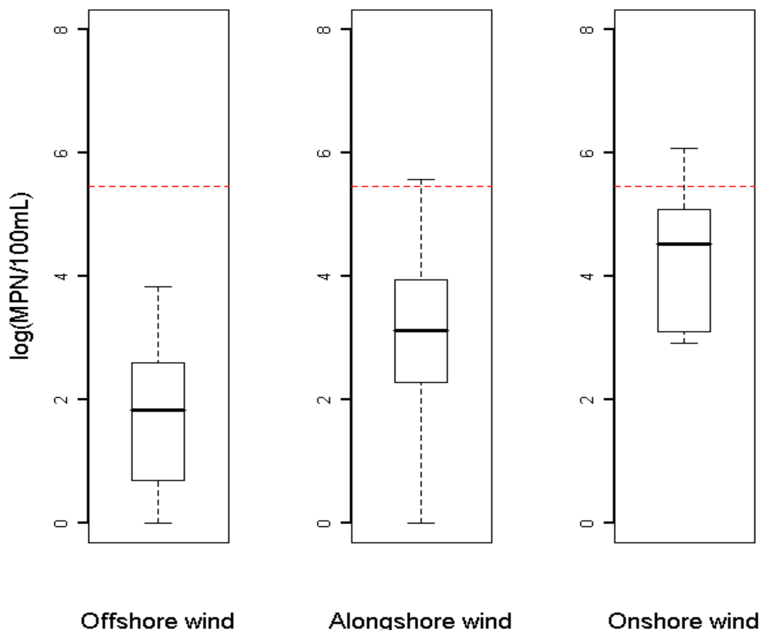

Figure 21: Winds at Otumba Park Beach were categorized and coded as offshore (0), along shore (0.5) and onshore winds (1).

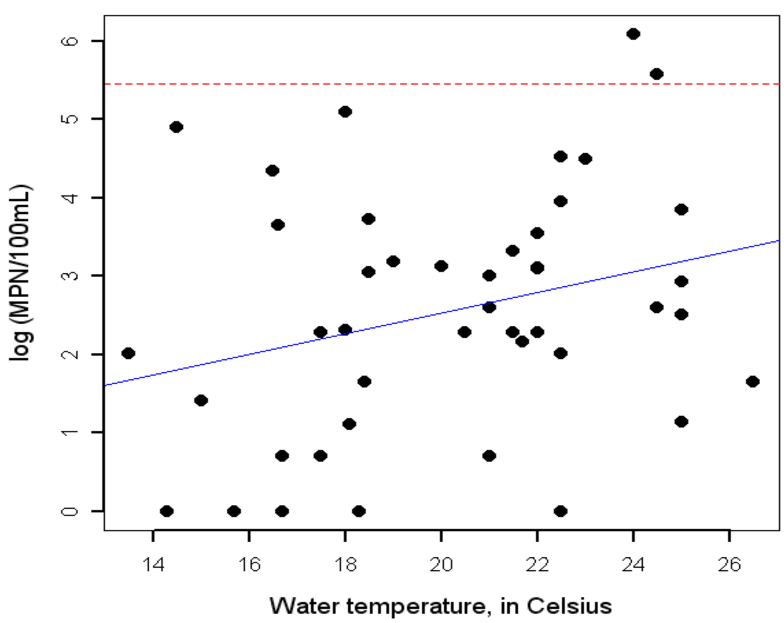

Figure 22: Scatter plot between response variable and water temperature, in Celsius. Red dotted line indicates, bathing water limit ($\log_e$ 235 (5.45) *E. coli*/100mL water).

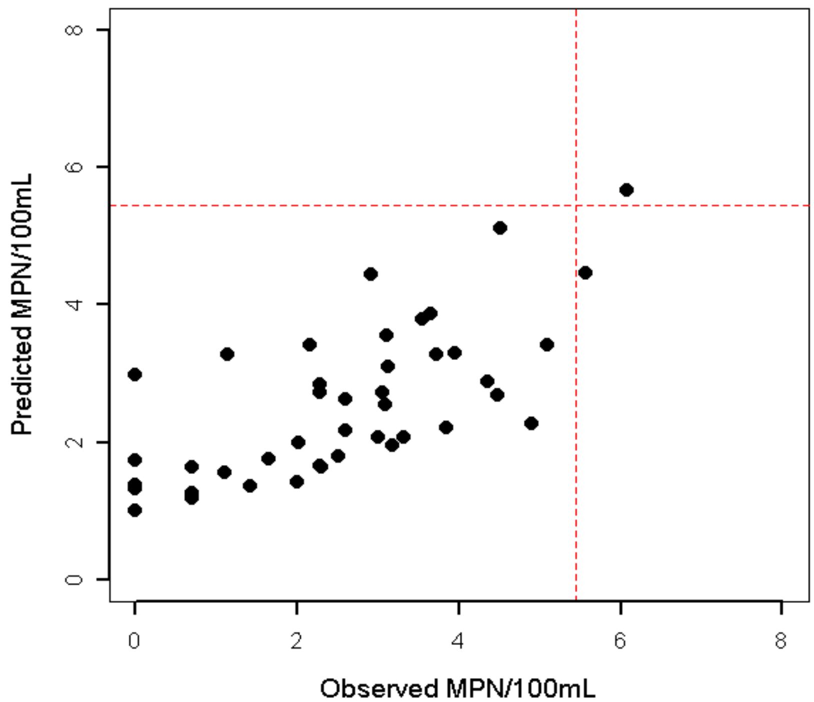

Figure 23: Scatter plot between predicted and observed $\log_e$ *E. coli* concentrations. Red dotted line indicates bathing water quality limit ($\log_e$ 235(5.45) *E. coli*/100mL water).

Table 48: Otumba Park Beach predictions and accuracy data.

Predictions	Number	Percent
Total predictions	44	-
False positives	0	0
False negatives	1	2.2
Correct above standard	1	2.2
Correct below standard	42	95.6

Table 49: Otumba Park Beach mathematical model parameter estimates. p-values (<0.05) of t-tests indicate that each of parameter estimates is statistically different from zero.

Variable	Parameter estimate	Standard error	t-value	p-value
Intercept	-1.03	1.16	-0.88	0.1
Turbidity	0.99	0.35	2.83	0.005
Onshore wind	2.15	0.54	3.97	0.001
Water temperature	0.09	0.05	0.05	0.05

Table 50: Otumba Park Beach R-squared indicates the fraction of variation in *E. coli* concentration explained by the model. p-value (<0.05) indicates the model is statistically significant.

Statistic	Values
F-statistic	12.51
Degrees of freedom	40
Residual S.E	1.16
R^2	0.48
p-value	0.001

Lakeside Park Beach:

Predictive equation: Predicted log *E. coli* concentration = 0.41 + 0.49* turbidity + 0.6* algae + 1.44*wave height.

Table 51: Summary statistics of *E. coli* concentrations at Lakeside Park Beach, 2007. Bathing water concentration limit, for *E. coli* is $\log_e$ 235 (5.45) *E. coli*/100mL water.

Number of samples	Daily log *E. coli* concentrations (MPN/100mL)				Number (percent) of days bathing water standard was exceeded
	Minimum	Mean	Median	Maximum	
47	0	2.58	2.58	5.37	0 (0)

Table 52: Pearson's correlation coefficient (r) between $\log_e$ *E. coli* concentrations and explanatory variables for Lakeside Park Beach, numbers in bold are significant ($p < 0.05$).

Variable	**R**
Wind direction	-0.06
Wind speed	0.24
Long shore current speed	0.47
Long shore current direction	0.17
Water temperature	0.03
Onshore wind	0.13
Air temperature	-0.30
Turbidity	0.54
Wave height	0.47
Algae	0.40
Gulls	-0.39
One day lag gulls	-0.05
24hr rainfall	0.37
48hr rainfall	-0.13
72hr rainfall	0.25
Event rainfall	0.17
Weighted 48hr rainfall	0.37
Weighted 72hr rainfall	0.39
Combined 72hr rainfall	0.37
Today's barometric pressure	-0.40
Yesterdays barometric pressure	-0.21
Difference in barometric pressure	01.7

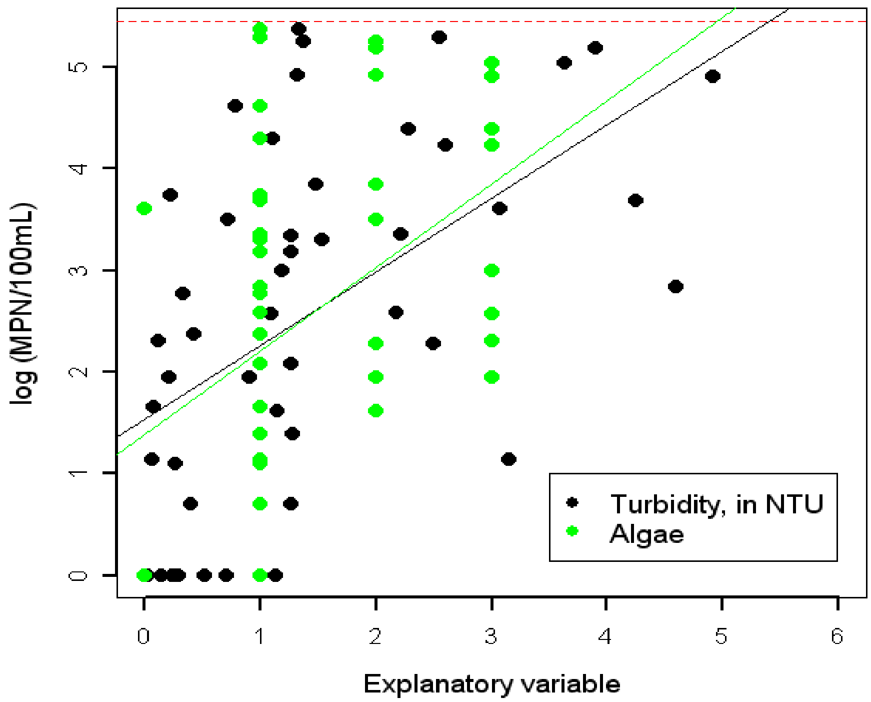

Figure 24: Scatter plot between response variable and turbidity, in (NTU) shown in black. Green shows the scatter plot between response variable and algae. Red dotted line indicates, bathing water standard ($\log_e$ 235 (5.45) *E. coli*/100mL water).

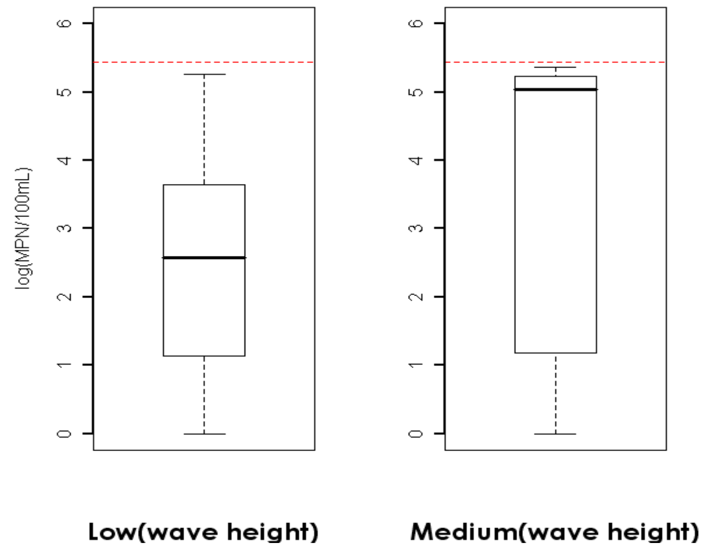

Figure 25: Wave heights at Lakeside Park Beach were categorized as low (<1 feet) and medium (≥1 and <2 feet) wave height only, as there no high wave height periods (≥2 feet) during the sample collection season.

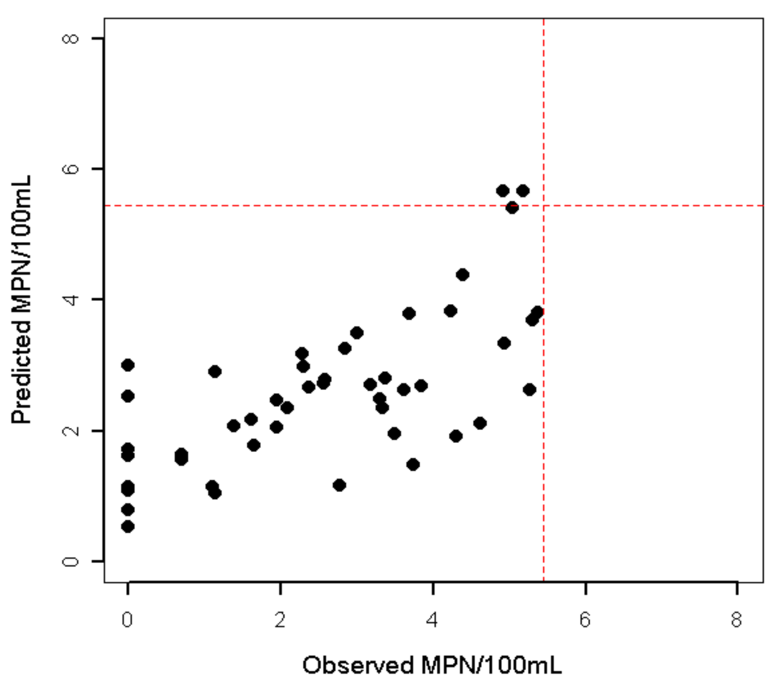

Figure 26: Scatter plot between predicted and observed $\log_e$ *E. coli* concentrations. Red dotted line indicates bathing water quality limit ($\log_e$ 235(5.45) *E. coli*/100mL water).

Table 53: Lakeside Park Beach predictions and accuracy data.

Predictions	Number	Percent
Total predictions	47	-
False positives	3	6.4
False negatives	0	0
Correct above standard	0	0
Correct below standard	44	93.6

Table 54: Lakeside Park Beach mathematical model parameter estimates. p-values (<0.05) of t-tests indicate that each of parameter estimates is statistically different from zero.

Variable	Parameter estimate	Standard error	t-value	p-value
Intercept	0.41	0.43	0.95	0.1
Turbidity	0.49	0.16	2.99	0.005
Algae	0.6	0.23	2.56	0.01
Wave height	1.44	0.52	2.75	0.005

Table 55: Lakeside Park Beach statistical parameters. R-squared indicates the fraction of variation in *E. coli* concentration explained by the model. p-value (<0.05) indicates the model is statistically significant.

Statistic	Values
F-statistic	12.5
Degrees of freedom	43
Residual S.E	1.31
R^2	0.47
p-value	0.005

Whitefish Dunes State Park Beach:

Predictive equation: Predicted $\log_e$ *E. coli* concentration = -5.32 + 1.59* wave height +

2.2*log (water temperature).

Table 56: Summary statistics of *E. coli* concentrations at Whitefish Dunes State Park Beach, 2007. Bathing water standard concentration, for *E. coli* is $\log_e$ 235 (5.45) *E. coli*/100mL water.

Number of samples	**Daily log *E. coli* concentrations (MPN/100mL)**				**Number (percent) of days bathing water standard was exceeded**
	Minimum	**Mean**	**Median**	**Maximum**	
54	0	2	1.84	7.25	4 (7.4)

Table 57: Pearson's correlation coefficient (r) between $\log_e$ *E. coli* concentrations and explanatory variables for White Fish Dunes State Park Beach, numbers in bold are significant ($p<0.05$).

Variable	**R**
Wind direction (degrees)	0.21
Wind direction	0.20
Wind speed	0.06
Long shore current speed	0.49
Long shore current direction	0.31
Water temperature	0.33
Onshore wind	-0.08
Air temperature	0.18
Turbidity	-0.07
Wave height	0.54
Algae	-0.03
Gulls	0.48
One day lag gulls	-0.05
24hr rainfall	-0.19
48hr rainfall	-0.17
72hr rainfall	-0.01
Event rainfall	-0.01
Weighted 48hr rainfall	-0.19
Weighted 72hr rainfall	-0.18
Combined 72hr rainfall	-0.18
Wind vector	0.01

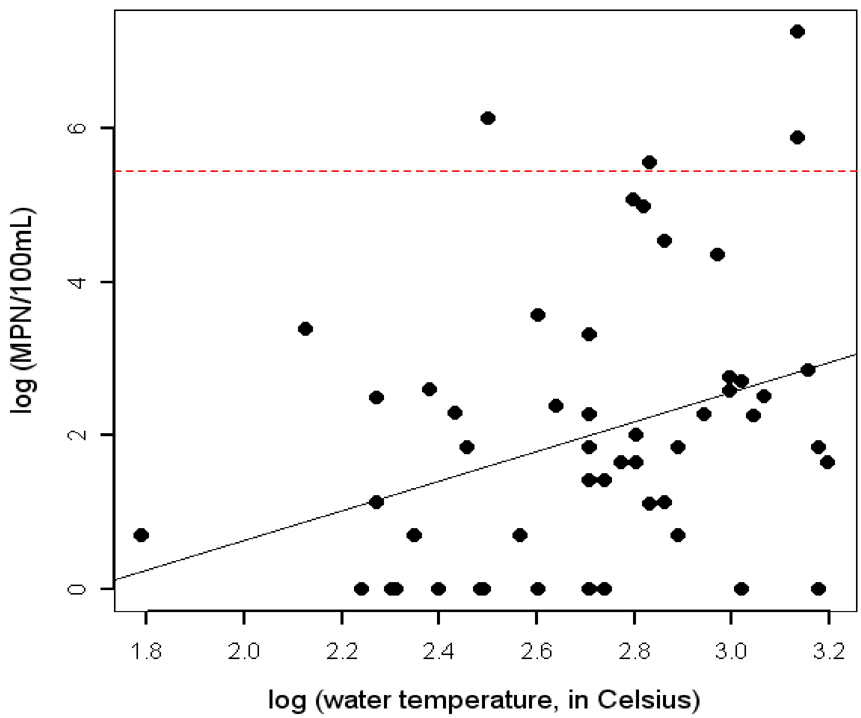

Figure 27: Scatter plot between response variable and $\log_e$ (water temperature), in Celsius. Red dotted line indicates, bathing water standard ($\log_e$ 235 (5.45) *E. coli*/100mL water).

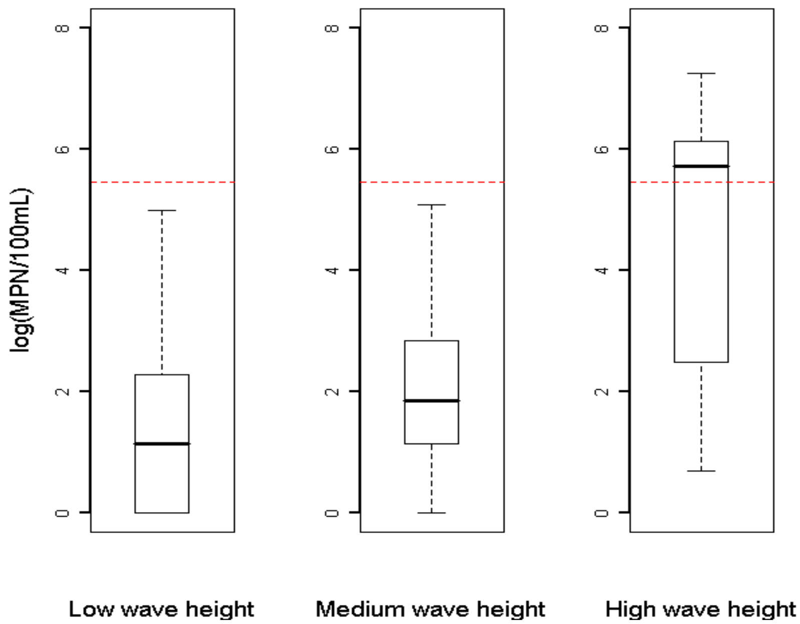

Figure 28: Waves were categorized based on their heights for White Fish Dunes State Park Beach. Low wave height (<1 feet), medium wave height (≥1 and <2 feet) and high wave height (≥2 feet).

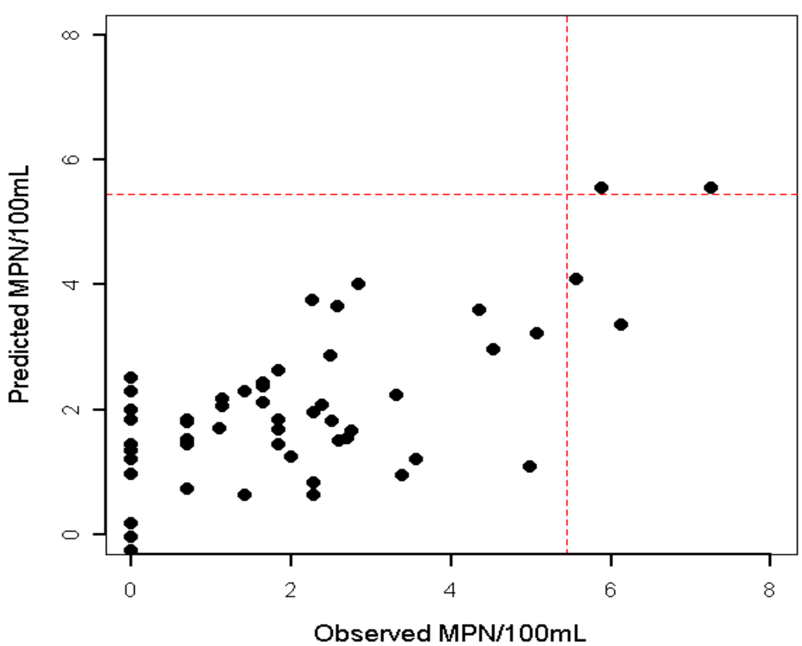

Figure 29: Scatter plot between predicted and observed $\log_e$ *E. coli* concentrations. Red dotted line indicates bathing water quality standard ($\log_e$ 235(5.45) *E. coli*/100mL water).

Table 58: White Fish Dunes State Park Beach predictions and accuracy data.

Predictions	Number	Percent
Total predictions	54	-
False positives	0	0
False negatives	2	3.7
Correct above standard	2	3.7
Correct below standard	50	92.6

Table 59: White Fish Dunes State Park Beach mathematical model parameter estimates. p-values (<0.05) of t-tests indicate that each of parameter estimates is statistically different from zero.

Variable	Parameter estimate	Standard error	t-value	p-value
Intercept	-5.32	1.72	-3.08	0.005
Wave height	1.59	0.3	5.36	0.001
Log (water temperature)	2.2	0.62	3.55	0.001

Table 60: White Fish Dunes State Park Beach statistical parameters. R-squared indicates the fraction of variation in *E. coli* concentration explained by the model. p-value (<0.05) indicates the model is statistically significant

Statistic	Values
F-statistic	19.32
Degrees of freedom	51
Residual S.E	1.4
R^2	0.43
p-value	0.001

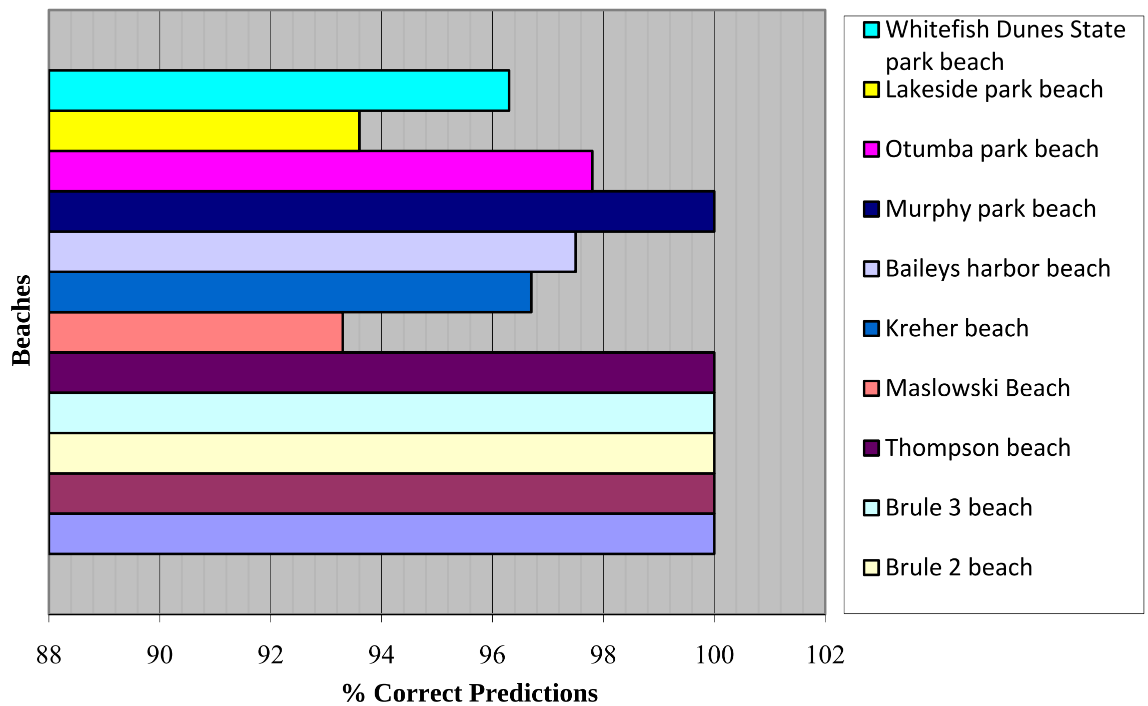

Figure 30: Mathematical model % correct predictions for all beaches combined, which includes predictions of both above and below the water standard limits.

DISCUSSION

Mathematical models were developed to predict *E. coli* concentrations in recreational water for beaches in Wisconsin. Beaches were located in Douglas, Ashland, and Bayfield Counties (Lake Superior) and in Door County (Lake Michigan).

In Douglas County, three beaches (Barkers, Brule 2 and Brule 3) exceeded the bathing water quality limit ($\log_e$ 235 (5.45) *E. coli*/100mL water) once during the 2007 beach season, with Barkers and Brule 2 exceeded the standards on the same day (June 19, 2007). Brule 1 exceeded the bathing water quality standards twice during the beach season (2007), with 1 day coinciding with Brule 3's exceedence of bathing water quality standards (June 20, 2008). Wave height showed significant positive correlation with $\log_e$ *E. coli* concentrations at the three Brule beaches and was included in their mathematical models (Figure 3, 6 and 9). Median $\log_e$ *E. coli* concentrations increased with increase in wave height, probably because as wave heights increased beach water may wash *E. coli* present in the sand and bring in to the water. The *E. coli* concentrations generally showed positive correlation with wave height. Onshore wind variable was included in all the Douglas Counties beaches, except Brule 3 Beach. It was observed that during onshore winds median $\log_e$ *E. coli* concentrations were greater than with long shore and offshore winds. This is because winds bring water towards the sand and wash off the *E. coli,* present in the sand or from the material present on the sand, in to the water. Even though there were coincidences of bathing water quality exceedences on the same days and some variables included in the mathematical models were similar, a single predictive model could not be developed for the Douglas County beaches. Individual mathematical models for the beaches in Douglas County all predicted $\log_e$ *E. coli* concentrations accurately, without predicting any false positives or false negatives (Figure 5, 8 and 10).

Three beaches from Ashland and Bayfield Counties were studied during the 2007 swimming beach season. At least one rainfall variable was included in the mathematical models of these beaches. These rainfall variables were positively correlated to the $\log_e$ *E. coli* concentrations (Table 24, 29 and 34). *E. coli* concentrations were observed to be high following a rainfall event. This may be due to the runoff brought into the beaches by storm water, which has been shown in previous studies by others (1). Maslowski Beach had 2 days lagged algae and 1 day lagged gull numbers as explanatory variables included in its predictive model, which were positively correlated to $\log_e$ *E. coli* concentrations. The influence of algae on *E. coli* concentrations is lagged because, the *E. coli* present in the cladophora mats may take time to get washed off from the cladohora mats. Only Thompson Beach's mathematical model predicted $\log_e$ *E. coli* concentrations accurately, without false positive and false negative predictions (Figure 11 and Table 23).

In Door County wave height was included as a variable for Bailey's Harbor Beach, Lakeside Park Beach and Whitefish Dunes Sate Park Beach mathematical models. Wave height was positively correlated to $\log_e$ *E. coli* concentrations at these beaches. At Bailey's Harbor Beach during high wave height periods (≥2 feet) $\log_e$ *E. coli* concentrations exceeded the bathing water limits (Figure 14). This was not observed at the two other beaches; however, there was an increase in median $\log_e$ *E. coli* concentrations as wave heights increased from low to high wave heights (Figures 21 and 25). A turbidity variable was included in Murphy Park, Otumba Park and Lakeside Park

Beaches' mathematical models, which was positively correlated to the $\log_e$ *E. coli* concentrations (Figures 17, 20 and 24). The Murphy Park Beach mathematical model was the only one which predicted 100% accurately, without false positives and false negatives as seen Figure 19 and Table 43. Overall, it can be concluded that:

1. Mathematical models were developed for all beaches except for two, Fish Creek Beach and Ephraim Beach, because the best mathematical model developed for each was not statistically significant.
2. Six beach's mathematical models predicted the $\log_e$ *E. coli* concentrations with 100% accuracy (no false positives or false negatives) (Figure 30). Four beaches were from Douglas County, 1 from Ashland County and 1 from Door County.
3. There were different sets of explanatory variables included for each beach. There was no common mathematical model for the beaches in any county investigated.

Future Work:

Predictive models were built using data from one beach season and were tested with model fitting (i.e. model fitted with *E. coli* concentrations used for developing the predictive model). However, to use these models for beach monitoring a more robust predictive model should be built using larger data sets containing multiple beach seasons and should be tested with real time *E. coli* concentrations. Building multiple predictive models for a beach should be considered, like using one predictive model for the first half of a beach season and a different predictive model for the rest of beach season.

REFERENCES

1. Ackerman, D. & Weisberg, S (2003). Relationship between rainfall and beach bacterial concentrations on Santa Monica Bay Beaches. J Water Health , 1 (2), 85-89.

2. Ambient Water Quality Criteria for Bacteria (1986). United States Environmental Protection Agency.

3. An Ocean Blueprint for the 21st Century-Final Report (2004). Washington, DC: U.S. Commission on Ocean Policy.

4. Bej, A. K., McCarty, S. C., & Atlas, A. R. (1991). Detection of coliform bacteria and *Escherichia coli* by multiplex polymerase chain reaction: Comparison with defined substrate and plating methods for water quality monitoring. Appl Environ Microb, 57 (8), 2429-2432.

5. Cabelli, J. V., Dufour, P. A., Levin, A. M., & Mccabe, J. L. (1979). Relationship of microbial indicators to health effects at marine bathing beaches. Am J Public Health, 69 (7), 690-696.

6. Cabelli, V. J. (1983). Health effects criteria for marine recreational waters. United States EnvironmentalProtection Agency, Research and Development.

7. Dale, W. G., Kim, A. D., Paul, J. H., & and Joan, B. R. (2003). Pathogenic human viruses in coastal waters. Clin Microbiol Rev, 16 (1), 129-143.

8. Dorfman, M. (2004). Testing the waters 2004. Natural Resources Defense Council - A guide to water quality at vacation beaches.

9. Dufour, A. P. (1984). Health effects criteria for fresh recreational waters. United States Environmental Protection Agency. Cincinnati: U.S. Environmental Protection Agency.

10. Edberg, S. C., Allen, M. J., Smith, D. B., & Kriz, A. N. (1990). Enumeration of total coliforms and *Escherichia coli* from source water by the defined substrate technology. Appl Environ Microb, 56 (2), 366-369.

11. Edberg, S., Rice, E., & Allen, R. K. (2000). *Escherichia coli*: The best biological drinking water indicator for public health protection. J Appl Microbiol, 106-113.

12. Favero, M. S. (1985). Microbiologic indicators of health risks associated with swimming. Am J Public Health, 75 (9), 1051-1053.

13. Fong, T.-T., & Lipp, A. E. (2005). Enteric viruses of humans and animals in aquatic environments: health risks, detection, and potential water quality assessment tools. MicrobioL Mol Biol R, 357-371.

14. Francy, D. S., Darner, R. A., & Bertke, A. E. (2006). Models for predicting recreational water quality at lake Erie beaches. U.S. Geological Survey. Virginia: U.S. Geological Survey.

15. Guidelines for safe recreational water environments(2003). Geneva: World Health Organization.

16. Implementing the BEACH Act of 2000 - Report to Congress(2006). Environmental Protection Agency.

17. Joon, H. K. (2004). Public mis-notification of coastal water quality: A probabilistic evaluation of posting errors at Huntington beach, California. Envir Sci Tech, 2497-2501.

18. Kocwa-Haluch, R. (2001). Waterborne Enteroviruses as a hazard for human health. Polish J Environ Studies, 10 (6), 485-487.

19. Lisa, H. E. (2006). Wisconsin water quality report to Congress. Madison: WIDNR.

20. MMWR and EID. (2007, July 27). Retrieved July 23, 2008, from Centers for Disease Control and Prevention Website: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5629a1.htm

21. National beach guidance and required performance criteria for grants (2002). Washington, DC: U.S. Environmental Protection Agency EPA 823-B-02-004.

22. Olyphant, G. A., & Whitman, A. R. (2004). Elements of a predictive model for determining beach closures on a real time basis: The case of 63rd Street Beach Chicago. Environ Monit Assess, 98, 175-190.

23. Polo, F., Figueras, M., Inza, I., Sala, J., & Guarro, J. F. (1998). Relationship between presence of Salmonella and indicators of faecal pollution in aquatic habitats. FEMS Microbiology Letters , 160 (2), 253-256.

24. Rice, E. W., & Edberg, M. J. (1990). Efficacy of 1-Glucuronidase assay for identification of *Escherichia coli* by the defined-substratet technology. Appl Environ Microb, 56 (5), 1203-1205.

25. Sears, C. L. (2005). A dynamic partnership: Celebrating our gut flora. Anaerobe, 247-251.

26. Seyfried, P. L., Tobin, R. S., Brown, N. E., & Ness, A. P. (1985). A prospective study of swimming-related illness II. Morbidity and the microbiological quality of water. Am J Public Health, 75 (9), 1071-1075.

27. Shuping, Z., Robert A, K., Renato L, S., & Andrews-Polymenis, H. (2003). Molecular pathogenesis of *Salmonella enterica* serotype typhimurium-induced diarrhea. Infect Immun, 71 (1), 1–12.

28. Stephen, J. C., George, L. R., Gregory, K. C., & David, G. K. (1993). The health effects of swimming at Sydney Beaches. Am J Public Health , 83 (12), 1701-1706.

29. Sue, M., Joe, M., & Istre., A. G. (1987). Shigellosis outbreak associated with swimming. Am J Public Health, 77 (2), 166-168.

30. Surveillance for waterborne disease and outbreaks associated with recreational water - United States, 2003–2004. Surveillance Summaries (2006), Centers for Disease Control and Prevention.

31. Timothy, J. W., Rebecca, L. C., Elizabeth, S., Beach, M., Kristen, P. B., Ann, H. W., et al. (2006). Rapidly measured indicators of recreational water quality are predictive of swimming-associated gastrointestinal illness. Environ Health Persp, *114* (1), 24-28

32. USEPA (U.S. Environmental Protection Agency). (1999). Action Plan for Beaches and Recreational Waters.EPA 600/R-98-079. U.S.

33. USEPA (U.S. Environmental Protection Agency). (1976). Fecal coliform bacteria. In: Quality criteria for water. Office of Water and Hazardous Materials, Washington, DC.

34. Yoder, J. S., Blackburn, B. G., & Craun, G. F. (2004). Surveillance for Waterborne-Disease Outbreaks Associated with Recreational Water-United States, 2001–2002. Centers for Disease Control and Prevention. CDC .

35. Zhongfu, G., & Frick., W. E. (2006). Some statistical issues related to multiple linear regression modeling of beach bacteria concentrations. Environ Res, 358-364.