Predictive Modeling at Beaches
Volume II: Predictive Tools for Beach Notification

Richard G. Zepp, Mike Cyterski, Rajbir Parmar, Kurt Wolfe, Emily M. White, and Marirosa Molina

U.S. Environmental Protection Agency
National Exposure Research Laboratory

EPA-600-R-10-176

November 22, 2010

Contents

Tables

Figures

This page is intentionally blank.

Abbreviations and Acronyms

ADCP	acoustic Doppler current profiler
AIC	Akaike Information Criterion
AICC	Corrected Akaike Information Criterion
BEACH Act	Beaches Environmental Assessment and Coastal Health Act
BIC	Schwarz Bayesian Information Criterion
CCE	calibrator cell equivalents
CDOM	colored dissolved organic matter
CFU	colony forming units
Cp	Mallows' Cp
cm	centimeter(s)
CSO	combined sewer overflow
DOC	dissolved organic carbon
EC	*Escherichia coli* or *E. coli*
EPA	U.S. Environmental Protection Agency
FIB	fecal indicator bacteria
GA	Genetic Algorithm
IV	independent variable
km	kilometer(s)
m	meter(s)
MEF	mean squared error of fitting
MEP	mean squared error of prediction
mL	milliliter(s)
MLR	multiple (or multivariable) linear regression
MSE	mean squared error
µm	micrometer
nm	nanometer
NEEAR	EPA National Epidemiological and Environmental Assessment of Recreational Water study

NOAA	National Oceanic and Atmospheric Administration
PA	publicly available
PREQB	Puerto Rico Environmental Quality Board
PREMIER	Predictive Modeling of Indicator Exposure Research
PRESS	Predicted Residual Sum of Squares
qPCR	quantitative Polymerase Chain Reaction
R^2	coefficient of determination
RMSE	root mean square error
RMSEP	root mean square error of prediction
SS	site-specific
TSA	temporal synchronization analysis
TSC	target sequence copies
USGS	U.S. Geological Survey
UV	ultraviolet
VB	Virtual Beach Manager Toolset or Virtual Beach
VIF	variance inflation factor
WWTP	wastewater treatment plant

Executive Summary

The U.S. Environmental Protection Agency (EPA) is in the process of developing new or revised recreational water quality criteria as required in the Beaches Environmental Assessment and Coastal Health (BEACH) Act of 2000. Microbial contamination is often assessed by monitoring for indicator bacteria such as fecal coliforms, *Escherichia coli* or fecal enterococci. Epidemiological studies have demonstrated a correlation between the levels of those bacteria in water and rates of gastrointestinal and other illnesses in swimmers. Those fecal indicator bacteria (FIB) generally are not a significant cause of illness, but their presence at levels exceeding certain criteria indicates that health effects are likely to occur from pathogenic bacteria, viruses, or protozoans that are associated with fecal matter. One means to supplement, not replace, monitoring results and to make same-day public health decisions is to use predictive tools such as statistical models to evaluate beach water quality by providing estimates of FIB densities on the basis of current environmental conditions.

Pioneering studies at beaches on the West Coast and Great Lakes showed that site-specific predictive models can be used to determine beach closure notifications. This report supplements those earlier studies by refining and evaluating modeling tools for building multiple linear regression (MLR) models that can predict FIB densities measured by culturable and quantitative Polymerase Chain Reaction (qPCR) techniques at freshwater and marine beaches. The Great Lakes component of the research was linked to EPA's Advanced Monitoring Initiative. The research detailed here also was guided by a panel of experts who met in Airlie, Virginia in 2007 to discuss and recommend critical research needs for developing new or revised recreational water quality criteria {USEPA, 2007 #121.

This document is Volume II of a two-volume report. Volume I summarizes current uses of predictive tools that provide beach managers with basic concepts to develop predictive tools for same-day beach notifications at coastal, Great Lakes, and inland waters.

Volume II provides results of research conducted by EPA's Office of Research and Development to develop statistical predictive models at research sites. It also presents Virtual Beach—a software package that builds statistical multiple linear regression (MLR) predictive models.

A highlight of the report discussed in Chapter 2 is the development of a user-friendly software tool (Virtual Beach [VB]) that can be used to build and evaluate predictive MLR models. VB uses a collection of culturable and qPCR microbial FIB data, referred to in the report as dependent variables, and concurrently collected parameters that quantified environmental conditions at the beach sites. Such parameters are called *independent variables* (IVs) in the report. The software systematically relates FIB densities to the IVs to produce an optimal fit, i.e., a *predictive model*. The software's other capabilities include improving MLR modeling by creating interaction terms, transforming variables to maximize response linearity, and filtering out highly correlated IVs. The software chooses models (i.e., selects IVs) by optimizing metrics such as Root Mean Square Error (RMSE), Akaike Information Criterion, Bayesian Information Criterion, and Predicted Residuals Sum of Squares (PRESS) and includes a genetic search algorithm for handling a large number of IVs. In any endeavor involving predictive modeling, variability and uncertainty is associated with the model output. Such uncertainty arises from a variety of factors that are impossible to completely eradicate from the modeling exercise. VB addresses the uncertainty issue by providing a probability of exceedance for any regulatory standard that the user wishes to investigate. Even so, there is no guarantee that every model prediction will be correct, and a situation where the model predicts water quality to be good enough for public recreation could be erroneous. Decisions to open or close the beach must be made, however, and in the best case scenario, the regression models developed using VB will outperform less rigorous predictive efforts.

VB facilitates the use of different approaches to evaluate model performance. On the basis of analyses of data obtained at a freshwater beach using VB, we recommend having at least 50 observations for model development but, having 100 or more is preferable. However, shorter-term dynamic models might still provide useful guidance for beach managers as the longer term data sets are obtained. That conclusion should be evaluated using long-term data sets from other locations, especially from coastal marine beaches. Chapter 3 of this report discusses several different approaches that are used to evaluate model performance. Volume I and other parts of this report includes discussion of the use of model builders such as VB version 2.0 to select the *best* model for a given data set. The VB tool can be used to examine techniques for evaluating the performance, or *trustworthiness*, of this best model. Among other approaches to evaluate whether a chosen model is *good* in an absolute sense rather than best in a comparison to other potential models (see Chapter 3 on model evaluation methods), we used the model's adjusted coefficient of determination (adjusted R^2) and the RMSE to evaluate the fit achieved in regression of predicted versus observed values. Another approach involves evaluating a model's ability to predict observations known as *exceedances*, i.e., observations that are greater than a given value (e.g., an EPA regulatory threshold for FIB levels at freshwater beaches). Another assessment of a model's goodness-of-fit uses a process called cross-validation, in which the data set is split into a training data set and a testing data set. A model built by fitting the former is used to make predictions for the latter. Analysis of a decade-long data set collected by the U.S. Geological Survey (USGS) at Huntington Beach, Ohio, indicated that the ratio of predictive to fitting errors moves toward unity as the sample size increases.

To provide the data needed for refining and evaluating statistical models, we developed a program designed to enhance the *predictive modeling of indicator exposure research* (PREMIER). Using automated instruments to obtain data enhanced modeling results by facilitating application of new modeling techniques such as temporal synchronization analysis (Chapter 9). The dependent variables and IVs obtained in those studies were used for model refinement at selected freshwater beaches on the Great Lakes and at coastal marine beaches of the eastern United States and in the tropics. That part of the research is discussed in Chapter 4 and the appendices. We chose culturable enterococci and enterococci qPCR as the dependent variables because previous studies have shown that enterococci are the best FIB for assessing risk at both freshwater and marine beaches. We deliberately patterned the spatial and temporal sampling patterns of those studies after those used in EPA's National Epidemiological and Environmental Assessment of Recreational Water (NEEAR) studies. At the PREMIER beach sites, we used automated techniques for IV measurements that were not used in most NEEAR studies or in other modeling studies. We also used data provided by EPA's NEEAR epidemiological studies team.EPA data from a total of five freshwater and seven marine beaches were used in the studies. In addition, our model evaluation studies used the extensive data set obtained over 10 years by USGS and collaborators at Huntington Beach, Ohio.

Comparisons of the MLR modeling results provided in Chapter 5 indicate that, on the basis of adjusted R^2 values for predicted versus observed levels of the FIB, model performance was better for the freshwater beaches than for the marine beaches (freshwater average adjusted $R^2 = 0.5$, marine average adjusted $R^2 = 0.39$). Also, modeling results for the culturable FIB data are somewhat better than for the qPCR data (colony forming units [CFU] average adjusted $R^2 = 0.46$, qPCR average adjusted $R^2 = 0.42$). The lower values for marine beaches likely reflect the interplay of several factors. Those factors include the effects of currents and tides at the beaches, sunlight-induced inactivation, and inputs of FIB from bird and dog droppings, bather shedding, runoff, groundwater and desorption from sand and decaying vegetation. Waves also can be an important factor that reduces model performance; however, all but one of the marine beaches examined in the study were enclosed bays or estuaries that had subdued wave action.

Another general finding of Chapter 5 is that the models are much more accurate in predicting non-exceedances than exceedances at the beaches included in the study. On the other hand, on the basis of general ability of the models to predict observed FIB densities throughout the data sets (judged by the adjusted R^2), little evidence was apparent of a relationship between FIB densities and model performance, although we had expected that there might be. One contributing factor to the findings, especially in regards to the marine sites we studied, is that the data sets include very few exceedances. Accurately predicting phenomena (such as exceedances) that are rarely seen in the training data set is a very difficult task for a statistical model. In the case of freshwater beaches modeled in the study, CFU modeling results for the two cleanest beaches, South Shore Beach and West Beach, were among the best. Likewise, satisfactory modeling results were obtained at marine beaches such as Surfside Beach where almost no exceedances were observed. Although comparisons of results using culturable enterococci or *E. coli* were limited, we did find that model performance was about the same for both indicators, e.g., at Huntington Beach, Ohio, during 2004.

The VB tool was successfully used to evaluate the effectiveness of various IVs in predicting culturable-based and qPCR-based observations of enterococci at freshwater and marine beaches (Chapter 5). The analysis showed that

1. **Turbidity and antecedent rainfall were top IVs for culturable enterococci at both freshwater and marine beaches. This conclusion reinforces the findings of earlier studies.**
2. **The number of swimmers is an important IV, perhaps in part because part of our data sets were obtained at the NEEAR sites that were selected to ensure that large numbers of people were present for epidemiological studies at the beaches.** Presumably, shedding could have contributed to this finding.
3. **The effectiveness of IVs was dependent on factors such as the method used to measure the FIBs and site-specific factors.** For example, chlorophyll was the top IV for culturable measurements of FIB, but water UV absorption coefficient was the top IV for qPCR measurements. Chlorophyll and dissolved oxygen were top IVs for freshwaters, but turbidity, salinity, absorption coefficients, and bird abundance were the top variables

for marine beaches. Additional process-based research is required to understand the various factors that underlie the results of the statistical models.

Models developed using the VB tool can provide useful analyses of (1) the effect of the data set's length on model performance (Chapters 6 and 7). It could be valuable for optimizing and updating dynamic models based on short-term data sets.; (2) effects of data source location on the accuracy of models developed for a freshwater beach (Chapter 8); and (3) sub-grouping of data sets to help improve modeling results (Chapter 9).

Taking into account time lags and time windows using a *temporal synchronization analysis* can significantly improve MLR model results. Analyses of data from a freshwater and marine beach (Chapter 9) showed that temporal synchronization analysis is a noteworthy modeling technique that should be pursued. That methodology computes mean values of the IVs over temporal windows and lags relative to the time at which the response variable is measured. That is done to improve the statistical relationship between the IVs and the response. Applying the technique at a freshwater beach in the Great Lakes (South Shore Beach, Milwaukee) produced better MLR models compared to models using IVs measured at the time of FIB sampling.

1 Introduction

Contamination of recreational waters and drinking water supplies by pathogenic microorganisms has attracted the attention of environmental groups, public health officials, and water resource managers in coastal areas. Microbial contamination is often assessed by monitoring for indicator bacteria such as fecal coliforms, *E. coli* or fecal enterococci. Epidemiological studies have demonstrated a correlation between the levels of those bacteria in water and rates of gastrointestinal and other illnesses in swimmers. Those fecal indicator bacteria (FIB) generally are not a significant cause of illness, but their presence at levels exceeding certain criteria indicates that health effects are likely to occur from pathogenic bacteria, viruses, or protozoans that are associated with fecal matter. Fecal contamination originates from many sources, including coastal and shoreline development, wastewater collection and treatment facilities, septic tanks, urban runoff, disposal of human waste from boats, bathers themselves, animal feeding operations, and natural animal sources like wildlife and pet wastes.

State and local public health agencies use beach advisories and closings to communicate to the public that the level of pathogens in the water could be unsafe for swimming or other body contact recreation. The advisories and closings are based on water quality information and typically occur when monitoring results show that fecal bacteria levels exceed an applicable water quality criterion. The model can be very effective in locations where conditions change slowly, i.e., where conditions persist. But FIB densities often are highly uncorrelated with those of the previous day.

The Beaches Environmental Assessment and Coastal Health (BEACH) Act of 2000 requires coastal states to submit to the U.S. Environmental Protection Agency (EPA) monitoring, notification, and other information concerning their beaches. Recreational water quality assessments are based primarily on enumerating FIB; such measurements can take up to a day to complete. When beach advisories rely on sampling techniques that require about a day to analyze, a closure decision is made using a *persistence model,* which assumes that the last measured FIB density accurately reflects current contamination levels. Microbial densities can be highly dynamic because they are sensitive to factors such as changing meteorological conditions, water hydrodynamics, solar irradiance and in-water composition such as temperature and salinity (Boehm et al. 2007; Boehm 2003; Whitman et al. 2004). Thus, water-quality advisories based on the persistence model are likely of limited relevance to current bacterial densities (Kim and Grant 2004). The emerging use of rapid monitoring techniques such as quantitative Polymerase Chain Reaction (qPCR) techniques can help reduce that time lag (Haugland et al. 2005), but up to 4 hours or more are still required to use such techniques. An alternative approach for evaluating beach water quality uses models to predict indicator densities in recreational waters. Predictive models can provide useful and timely estimates that have been the basis for advisories at several locations in the Great Lakes and coastal marine beaches. The utility of predictive models is reflected by the fact that the BEACH Act included a *Beach Action Plan* that calls for research on developing predictive models to assess recreational water quality.

An overview of tools that have been used to develop predictive models is provided in Volume I of this report. It emphasizes statistical models and provides background information on (1) types of predictive models and tools that can be used to make beach notification decisions; (2) influences that different hydrologic environments have on biological contamination of beaches and how consideration of those influences affect model development; (3) predictive tools that health departments and other responsible agencies are using to make timely decisions on beach notifications; (4) procedures for developing statistical predictive models such as MLR models, rainfall thresholds, and notification protocol; and (5) trends in predictive tools for beach notifications, including deterministic tools.

In Volume II, we present specific considerations of research on how to improve and evaluate tools for creating statistical models that predict FIB levels at freshwater and marine beaches. The research was fostered initially through interactions with EPA Region 5 and scientists from the U.S. Geological Survey (USGS), National Oceanic and Atmospheric Administration (NOAA) and other partners in a Great Lakes project sponsored by EPA's Advanced Monitoring Initiative. That project focused on exchanging information and comparing modeling approaches at selected beaches of the Great Lakes. With that information in hand, a pilot study was conducted to improve and evaluate the statistical component of a model-building tool called Virtual Beach (VB). The results of that effort are described partly in Section 2.2 of this report. After software improvements, VB version 1.0 (VB 1.0) was created. Communicating VB 1.0 to end users was accomplished through several workshops conducted by the Wisconsin Department of Natural Resources, which was doing its own extensive testing of the software. Monitoring personnel from across the Great Lakes attended its two hands-on training workshops, which provided direct technical assistance to potential end users (Mednick and Watermolen 2009). The effort resulted in a number of recommendations for enhancing VB from a local operations perspective. Modifications made to version 1.0 before the 2009 beach season enabled a successful use case at Upper Lake Park Beach in Ozaukee County, Wisconsin (http://dnr.wi.gov/org/es/science/pdf/OzaukeeCountyWisconsin.pdf). Feedback from workshop participants and end users has informed the development of VB version 2.0 (VB 2.0). Personnel at other agencies, such as the USGS, NOAA and local governments, such as Lake County, Illinois, have been helpful in developing and refining the tool.

The research detailed here also was guided by a panel of experts who met in Airlie, Virginia, in 2007 to discuss and recommend critical research needs for developing new or revised recreational water quality criteria {USEPA, 2007 #121}. A sub-group of the panel met to discuss modeling applications for FIB prediction. The group's recommendations that are relevant to the research reported here are the following:

- Developing and testing simple (heuristic statistical) notification models on different recreational water types with a wide range of sources and geographical locales.
- Training recreational water managers.
- Creating a user-friendly portable package for developing local models.
- Investigating the effect of meteorological factors (e.g., rainfall, evapotranspiration) on nonpoint sources.
- Conducting modeling studies concurrently with planned epidemiological studies to help link statistical models to health effects.

A key aspect of predictive modeling involves collecting the data used to create the models. MLR models are the main statistical model used for making beach closure or advisory decisions. Building such models requires relevant microbial data (the dependent variable in MLR models) and other concurrently observed hydrometeorological and biogeochemical data (independent variables or IVs) that characterize the beach. Although monitoring data for culturable FIB have become more abundant in recent years, citable information involving qPCR measurements of FIB densities in recreational waters are scarce. Because qPCR data provide useful information about the sources and health impacts of fecal contamination, developing predictive model techniques for this type of measurement is also of interest. Compared to the amount of microbial data now available for beaches, it is difficult to find beach studies in which microbial data and relevant IVs have been concurrently observed. When these studies were initiated, predictive models had been developed primarily at beaches of the West Coast and Great Lakes (Boehm 2007; Francy et al. 2006b; Nevers and Whitman 2005). Although those pioneering studies showed that statistical models were quite useful for beach notifications, the predictive approach needed to be evaluated using culturable- and qPCR-based data from other types of recreational waters, including data collected as part of epidemiological studies.

Our objectives in this study were the following:

1. Develop a user-friendly software tool for building and evaluating predictive MLR models that can be used for beach notifications.
2. Use various strategies for model evaluation, e.g., cross-validation or exceedances of regulatory criteria, to assess the effects of differing data set sizes on model performance.
3. Refine and evaluate the capabilities of this modeling tool using concurrently collected culturable- and qPCR-based data and IVs from selected freshwater and marine beaches in the Great Lakes, eastern and southeastern United States, and the tropics (Puerto Rico).
4. Use data from epidemiological study sites to refine and evaluate the modeling tool.
5. Collect IVs using automated techniques to provide sufficient data to evaluate the dependence of model performance on the period over which the IVs are collected.
6. Use data from objective 4 above, evaluate the dependence of model performance on the period over which the IVs are collected (going beyond a focus on precipitation only).
7. Evaluate the relationship between the degree of contamination at a beach and model performance.
8. Compare model performance and the most important IVs at both freshwater and marine beaches in which similar or the same types of microbial and IV data were obtained.
9. Compare model performance for culture-based and qPCR-based methods that were used to measure beach contamination.
10. Provide model results that can be used to help evaluate the variability of model performance at a beach over differing periods of data collection.
11. Compare the results of models constructed using IV data obtained from sources that are varying distances from a beach, i.e., at the beach or several miles away from the beach.
12. Evaluate the use of data sub-grouping to improve model performance.

Chapter 2 of the report provides a general discussion of the VB software. The software and its user guide are provided on the Web separately from this report. Chapter 3 discusses several different approaches used to evaluate model performance and applies the approaches in an assessment of sample size on model performance. In Chapter 4 and the appendices, we provide information about characteristics of the various beaches that were the subjects of the modeling efforts and approaches used in *predictive modeling of indicator exposure research* (PREMIER) to acquire the data for the modeling studies. Chapter 5 presents and discusses MLR modeling results for the various freshwater and marine beaches in this report where the data were collected by EPA (PREMIER and NEEAR [National Epidemiological and Environmental Assessment of Recreational Water] research). The discussion includes evaluation of the performance of VB 2.0 at the various beaches and identifies the most effective IVs for the beaches. The results are compared with results from other regions, and possible impacts of beach conditions on the results are discussed. The remainder of the report details efforts that have been pursued in this study to improve statistical modeling techniques. Chapter 6 provides results on the variability of model performance at a beach over differing periods of data collection. We provide evidence that short-term dynamic models—those that are re-fit periodically to new data as they become available—can provide satisfactory short-term predictions (nowcasts) of FIB densities and that predictive models can be structured to provide 24-hour forecasts of FIB levels. Chapter 7 presents evidence that models built from extensive long-term data for a freshwater beach provide results superior to those based on short-term data sets. Chapter 8 shows the utility of the modeling approach by providing comparisons of models constructed using IV data obtained from sources that are varying distances from a beach, i.e., at the beach or several miles away from the beach (or both). Chapter 9 introduces the concept of temporal synchronization to refine MLR model results. That technique was used to improve modeling results at both a freshwater beach in Milwaukee and a marine beach in Miami. Chapter 10 summarizes and discusses the modeling results detailed in Volume II.

2 Virtual Beach

2.1 Introduction

VB 2.0 is a software package designed to construct site-specific MLR models to predict pathogen indicator levels at recreational beaches. The VB 2.0 modeling interface is designed to help find the best model amongst a large number of candidate models, based on criteria selected by the user. As the number of IVs increases, the number of possible models in the solution space increases exponentially. The user is able to select all, or a subset of, the IVs for consideration in the model to reduce the size of the solution space.

A pilot study for developing the VB tool was conducted using historical data for *E. coli* and various relevant IVs for Huntington Beach, Ohio (Frick et al. 2008). The site and data collection techniques are described in Section 4.2.5 of this report and in Francy et al. (2006a; Francy and Darner 2007). The data were collected by the USGS Ohio Water Science Center and its partners, the Cuyahoga County Board of Health, and others.

The early version of the VB tool(VB 1.0) was designed to help develop statistical models (equation 2.1) for predicting beach FIB densities (denoted by *EC*):

$$E[\ln(EC)] = \beta_0 + \sum_{i=1}^{p} \beta_i x_i \,, \qquad \text{(equation 2.1)}$$

where $E[\ln(EC)]$ represents the expected value of the natural logarithm of the mean of the dependent variable (EC), β_0 and β_i are the regression coefficients, x_i the IVs and p the number of variables used in the model. MLR analysis is based on the least squares method to fit models and is subject to several considerations, notably variable interactions, multi-collinearity and model selection. Their relevance to beach bacteria modeling was described further by Ge and Frick (Ge and Frick 2007). VB 1.0 used a backward elimination process, described in the parsimonious model section below, to help the user select the most promising model from a number of candidates that rapidly increase with the number of IVs in the analysis. That process offered a way to rank the usefulness of the selected IVs, with the best variable being the last recommended for elimination.

By definition, MLR equations (equation 2.1) are linear. That property can limit the value of IVs if the response variable is not a linear function of the variable. Often, variable transformations are used to overcome this limitation. In fact, the response variable itself is routinely natural log-transformed, a transformation that, alone, substantially increased the adjusted coefficient of determination (adjusted R^2) value of the model. VB 1.0 offered a number of common transformations, including square root, square, and others that, when selected, automatically transform the values of the IVs. Transformations can greatly increase the predictive value of vector quantities, such as wind and current (Nevers et al. 2007). For example, wind speed and direction are two IVs that, untransformed, often rank low as MLR IVs (Olyphant and Whitman 2004). The problem is that wind and current are vector quantities, and direction is a harmonic function. By transforming vector variables into their components, however, e.g., into along-shore and cross-shore, their value often is enhanced. VB 1.0 included the trigonometric transformations for converting wind (or current) speed and direction into their vector components, including axis rotation. It also offered asymmetric transformations, because, for example, the effects of wind speed can be completely different for onshore winds than offshore winds. Thus, the offshore component might be square-root transformed, while the onshore component remains unchanged.

Most importantly, VB 1.0 helped to identify the best IVs from the suite of potential variables available for fitting. In a typical study, dozens of candidate variables might be considered (Olyphant 2005; Olyphant and Whitman 2004). Each variable tended to increase the explained variance but, in successively smaller increments, compared the best variables. Fit to noise or to chance occurrences present in the data, and risking multi-collinearity, the marginal variables can degrade prognostic performance, and model maintenance also can increase. In practice, a smaller set of variables from which a parsimonious model is finally selected is preferred. In this work, the number of variables recommended for the parsimonious model ranged from two to five, but, for uniformity and comparability, four variables were retained. To help the user identify the most significant variables, VB 1.0 offered an automated model selection facility. VB 1.0 used backward elimination with Mallows' *Cp* (Frick et al. 2008)

$$C_p = p + (n - p)(\sigma^2 - \sigma^2_{full}) / \sigma^2_{full} \qquad \text{(equation 2.2)}$$

as the selection criterion, where n is the number of samples, p is the number of included variables, and σ^2 and $\sigma_{\mathrm{full}}{}^2$ are the residual variances of the reduced and full models. The parsimonious model is defined as the one with the minimum *Cp* value among the full (all variables) and the reduced (fewer variables) models. The *automated model selection* command of VB 1.0 computed the *Cp* statistic for all possible models and ranked the variables for elimination. If the default parsimonious model recommended by VB 1.0 is not selected, a process of guided variable elimination based on other factors that might influence variable selection, can be performed to arrive, stepwise, at the final model.

VB 1.0 offered other features, such as warnings when the data matrix is singular and when residuals demonstrate significant serial correlations. Users could also check for influential cases (i.e., those that greatly influence the values of the regression coefficients) and data outliers. Although VB 1.0 provided a useful tool for building MLR models, it did not have the full range capabilities of VB 2.0 described in Section 2.3.

2.2 Multiple Linear Regression Model Development

Variability and uncertainty are intrinsically associated with the model output of any predictive modeling endeavor. . Such uncertainty arises from a variety of factors that are impossible to completely eradicate from the modeling exercise. VB addresses the uncertainty issue by providing a probability of exceedance for any regulatory standard that the user wishes to investigate. Even so, there is no guarantee that every model prediction will be correct, and a situation where the model predicts water quality to be good enough for public recreation might be erroneous. Decisions to open or close the beach must be made, however, and in the best case scenario the regression models developed using VB will outperform less rigorous predictive efforts.

VB 2.0 is a software package designed to construct site-specific (SS) MLR models to predict pathogen indicator levels at recreational beaches. MLR has been shown to outperform persistence models (using only FIB concentrations at time t-1 to predict FIB levels at time t) beaches where conditions such as weather, hydrology, and human and animal traffic levels change significantly day to day (Frick et al. 2008). On the basis of direct user input and participants' feedback at the VB1 workshops conducted by Wisconsin's Department of Natural Resources, several enhancements were made to VB 1.0. For example, VB 1.0 was written in a computer language called Delphi; VB 2.0 has been written in Microsoft.NET language C# which makes it more inter-operable with other open source and commercial software libraries. Other improvements in VB 2.0 include better data import/export, improved data preparation methods, enhanced user interface, more model selection criteria, better reporting of modeling results, handling of larger number of IVs, improved prediction capabilities once a model has been developed, and new project management functions.

2.2.1 Automated Retrieval of Data over the Internet

VB 2.0 automatically retrieves geospatial data from the Internet and displays it on a map. That allows the user to locate the beach of interest and determine its orientation, which is important when processing wind and current data (Figure 2.1). Cross-shore and along-shore wind and current components affect pathogen indicator levels in some beach waters. Locating the beach on a map also allows VB 2.0 users to identify meteorological and water quality monitoring stations in the vicinity of the beach.

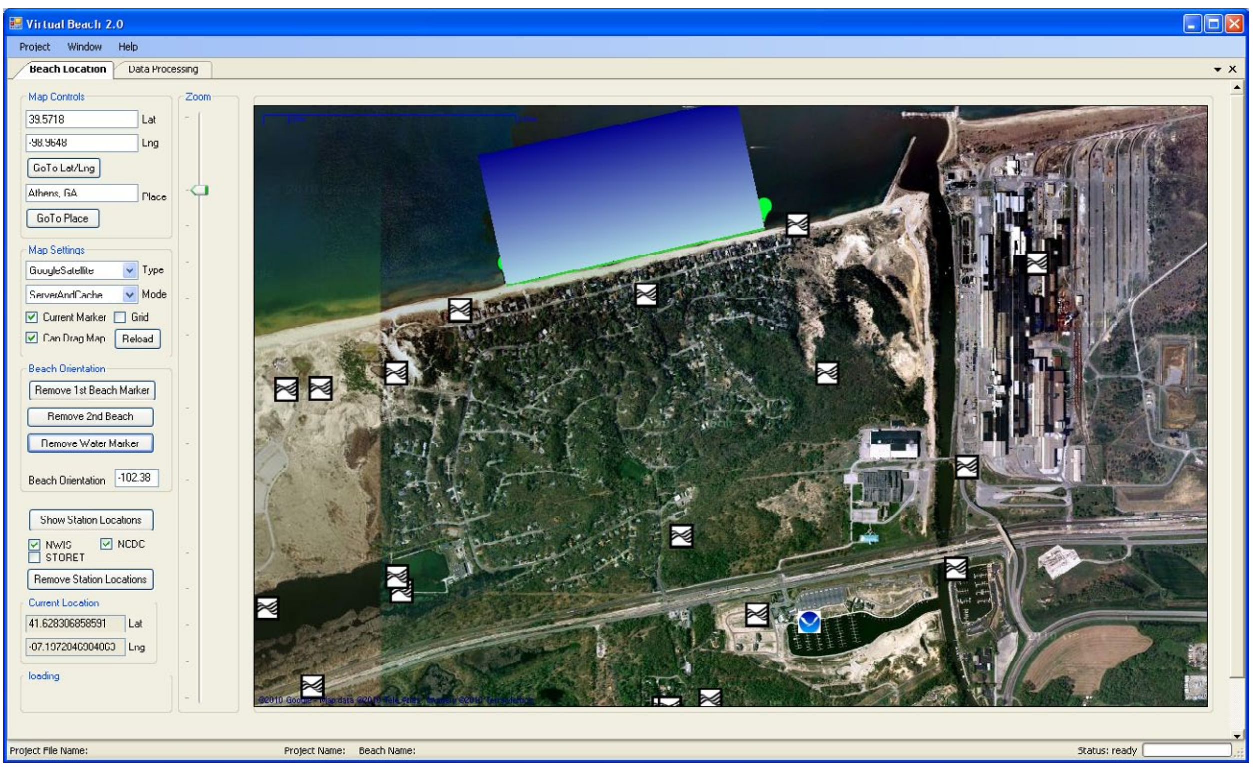


Figure 2.1. The VB 2.0 Beach Location interface showing a Google Hybrid map. Beach orientation is marked by the blue rectangle at the top of the screen.

2.2.2 Model Development

VB 2.0 uses a multilinear regression technique to relate a dependent variable (FIB densities) to IVs such as meteorological conditions, hydrological and water variables. The user can provide data for any number of potential model IVs. VB 2.0 allows the user to import data from Microsoft Excel spreadsheets. FIB concentration values can vary from 0 to several million or more. Such a large variation in dependent variable values often violates the assumption of linear relationship between the dependent and IVs in MLR. To make the relationship between dependent variables and IVs a linear one, it is recommended that the FIB concentration values be transformed, using a log transformation. The dependent variable values can be transformed either before or after importing data into VB 2.0.

VB 2.0 alerts the user for missing values of IVs. The user can then decide whether to exclude data with missing values from the analysis or enter estimates for missing values. VB 2.0 ensures that sufficient data points are available before a model can be developed. To develop a meaningful model, it requires a minimum ratio of 1:5 (although 1:10 is preferable) between the number of IVs and FIB concentration values.

Processing wind and current data also is facilitated by VB 2.0. Using beach orientation (beach angle), VB 2.0 processes wind and current data into cross-shore and along-shore components (Figure 2.2). Two new data columns corresponding to the two components are automatically added to the data table. Beach orientation is automatically calculated if the beach is defined using the mapping interface in VB 2.0. Otherwise, the user must specify the beach orientation using the rules given in VB 2.0 user's guide.

The user can opt to develop a model with any subset of IVs. The user also specifies model evaluation (goodness of fit) criteria. In addition, the user specifies the maximum value of variance inflation factor (VIF) to be tolerated. VIF measures multi-collinearity in MLR. From the user's choices, VB 2.0 develops and ranks models by the specified model selection criterion. The top models are presented with various types of model statistics and graphs for each, allowing users to decide on the best model for their application.

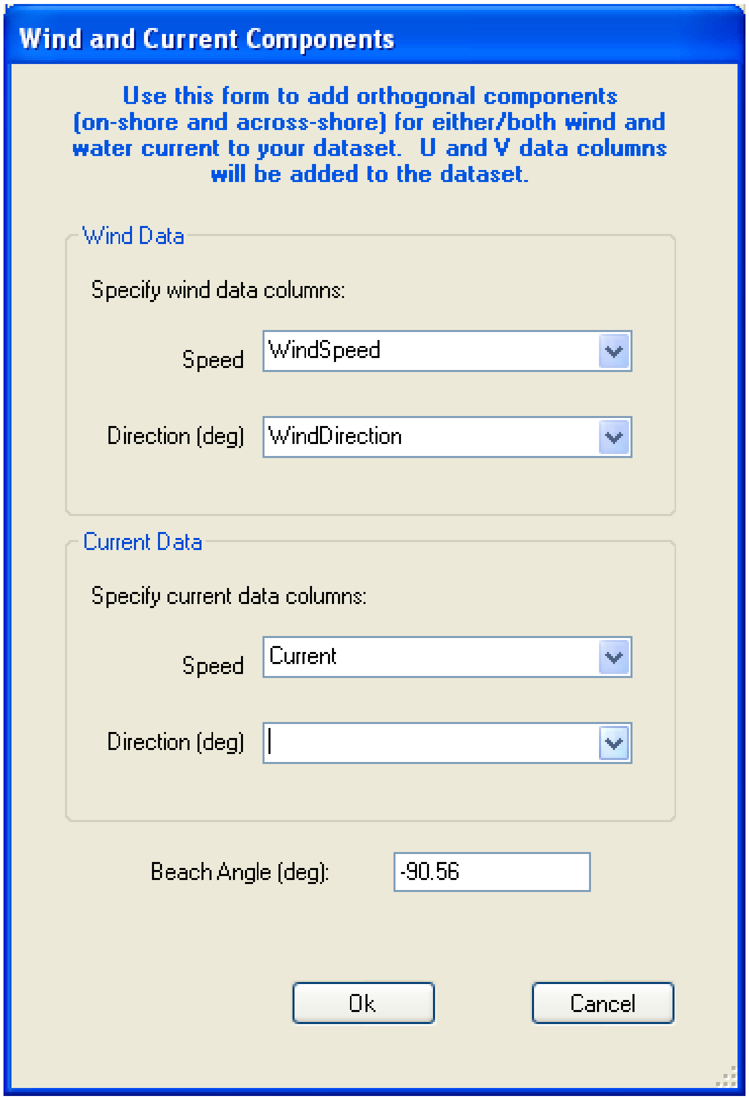


Figure 2.2. VB 2.0 wind and current data decomposition dialog.

2.2.3 Improving MLR Models Using VB 2.0 Software

VB 2.0 presents many ways to improve MRL modeling. Decomposition of wind and current data into cross-shore and along-shore parts, interactions between IVs, transformation of dependent and IVs, and filtering out highly correlated IVs all help to improve MLR modeling exercises.

Creating Interaction Terms. The user is presented with an option to include two-way interactions between IVs. Interaction between two IVs implies that the relationship between the first IV and the dependent variable is influenced by the second IV. If the user believes there are no interactions between the IVs, this step can be skipped.

Transforming Variables to Maximize Response Linearity. If a relationship between the response variable and an IV is nonlinear, the IV can be transformed to make the relationship linear. VB 2.0 provides five types of IV transformations: polynomial, square root, inverse, natural log, and log to the base 10. For a comparison of the response variable to each IV, a univariate correlation statistic (Pearson Correlation Coefficient) is calculated for each transformation (Figures 2.3 and 2.4). By default VB 2.0 uses the transformation with the highest value of Pearson Correlation Coefficient, but the user can accept any other transformation or to pick a transformation based on univariate relationship plots.

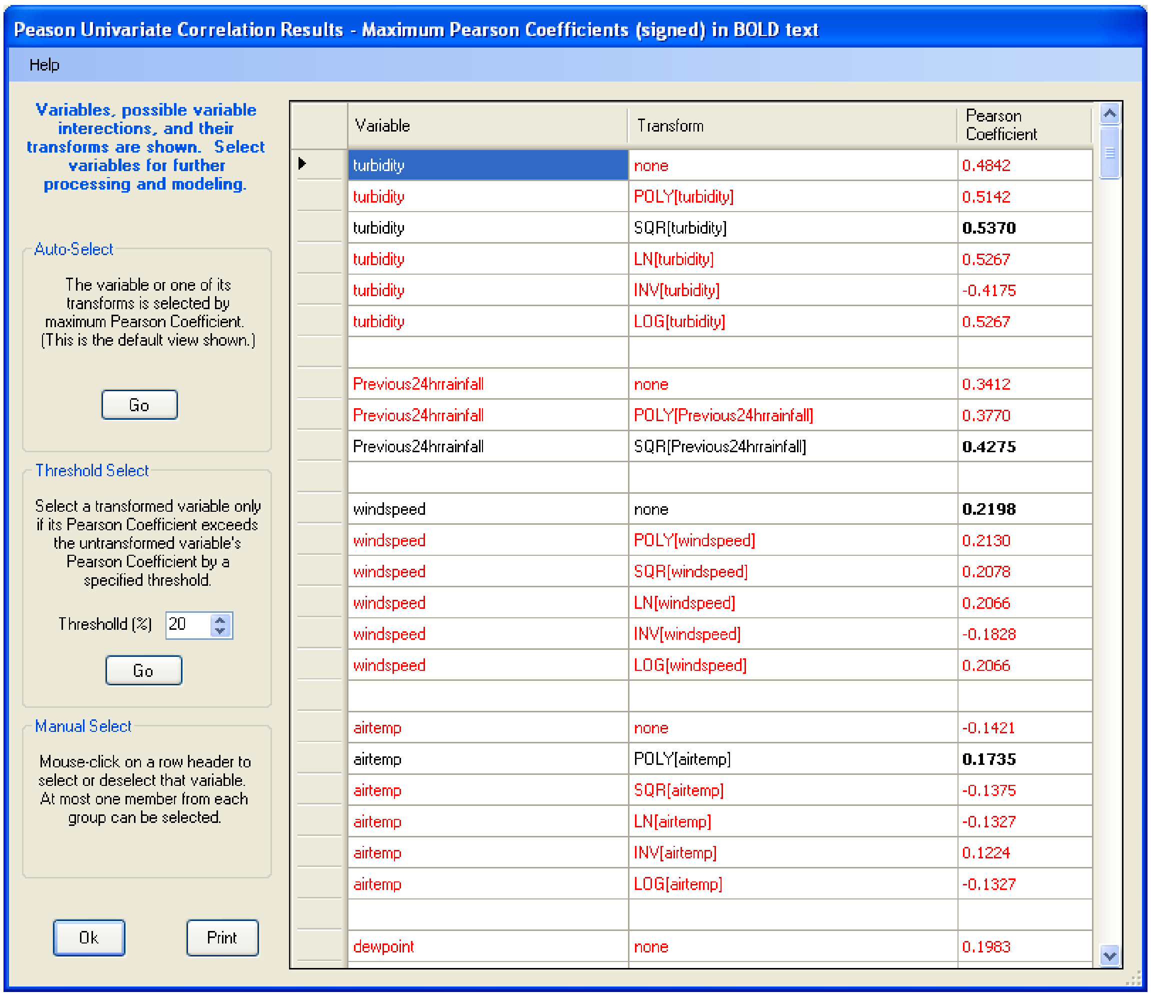


Figure 2.3. Pearson scores dialog for transformation of IVs.

Figure 2.4. Response and IV plots.

Filtering out highly correlated IVs. MLR has an underlying assumption that the IVs are not highly correlated with each other. VB 2.0 uses VIF to filter out highly correlated IVs. The higher the value of the VIF, the stronger the correlation between the variables:

$$VIF = 1 / (1 - R_i^2) \qquad \text{(equation 2.3)}$$

where R_i^2 is the coefficient of determination of a regression where the IV in question is used as the dependent variable, and all other IVs in the model are used as predictors. As such, a VIF value is calculated for every IV in the regression model. The user specifies the maximum value of VIF, and VB 2.0 drops any model with an IV whose VIF value exceeds that threshold..

2.2.4 Best Model Selection

VB 2.0 uses two methods to develop possible MLR models. If the number of IVs is not large, an exhaustive search method is used. In the exhaustive method, MLR models are developed with all possible combinations of the IVs. Invalid models, such as those including highly correlated IVs, are dropped. The remaining models are then ranked, on the basis of user-specified model selection criterion such as Akaike Information Criterion (AIC), Corrected AIC (AICC), Bayesian Information Criterion (BIC), R Squared, Corrected R Squared, Predicted Residual Sum of Squares (PRESS) Statistics, Root Mean Square Error (RMSE), number of true positives (Sensitivity), number of true negatives (Specificity), and overall correct number of positives and negatives (Accuracy). The top models are then presented to the user along with various model statistics. Users can then examine the statistics to select the best model for their application. The user has the following statistics available for each IV in a model: standard error, t-statistics, and p-value. Although models are ranked by a single user-specified criterion, the following statistics available for each ranked model help the user pick the model best suited to his or her application: R Squared, Adjusted R Squared, AIC, AICC, RMSE, BIC, PRESS statistics, Sensitivity, Specificity, and Accuracy. A graph of observations versus predictions can be used to further assist model selection (Figure 2.5).

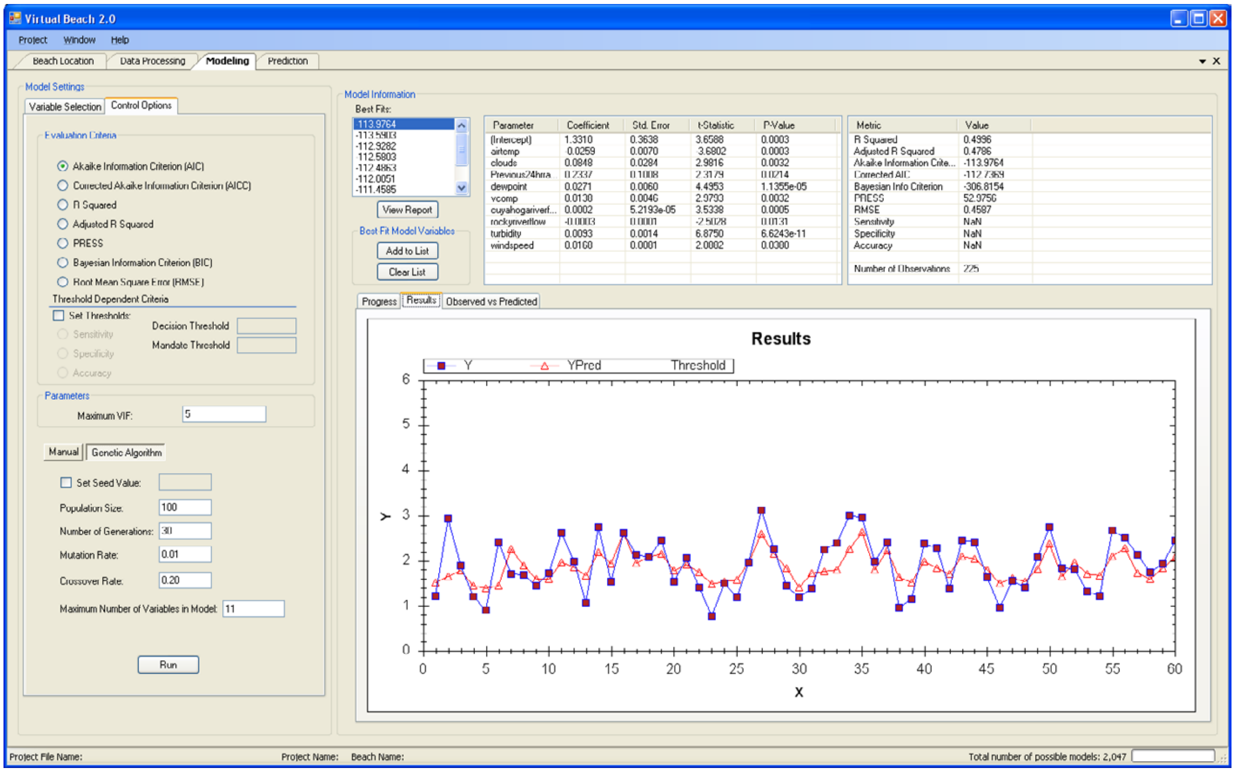


Figure 2.5. Model selection interface.

As the number of IVs increases, the number of possible models to be developed increases exponentially. VB 2.0 uses an artificial intelligence technique called Genetic Algorithms (GAs) to reduce the number of models developed when the number of IVs is large. GAs are search methods inspired by evolutionary biology. They are based on genetic processes such as inheritance, selection, crossover, and mutation that are seen in nature for species evolution. While GAs cannot be used directly for modeling beach pathogens, they can be used to evaluate and select models developed by other modeling techniques. Rather than developing and evaluating every possible model, GAs intelligently select IVs to be used, resulting in fewer models that need development and evaluation. GAs often result in finding the near-best model, as opposed to finding the best model.

The user specifies GA performance parameters such as initial population size, stopping criterion, crossover and mutation rates, random number generator seed, and the like. The model ranking and selection criteria remain the same as the exhaustive search method. GAs and exhaustive methods can also be combined to achieve better models for situations with a large number of IVs. First, the user takes a collection of all IVs represented in the top few (five or so) models ranked by GAs then selects the exhaustive method using the IVs collection from step one. In other words, GAs are used to select a set of potential IVs to be used in the exhaustive search method.

2.2.5 Using MLR Models to Provide Nowcast and Forecast of FIB Concentrations

Once a suitable model has been selected as described in the previous sections, VB 2.0 can be used to make FIB concentration predictions. Two types of predictions can be made with a model: predictions based on current and recently observed values of IVs (nowcast), and prediction based on forecasted values of IVs (forecast). VB 2.0 lets the user either directly enter values of IVs or import data from an Excel spreadsheet. Interactions and transformations, if required, are calculated automatically in the background. Predicted values of the response variable, and the probability of exceedance, are presented to the user. Predictions can be made for a single set or multiple sets of IVs. Predictions can be exported to Microsoft Excel. Predictions also can be appended to existing predictions for recordkeeping (Figure 2.6).

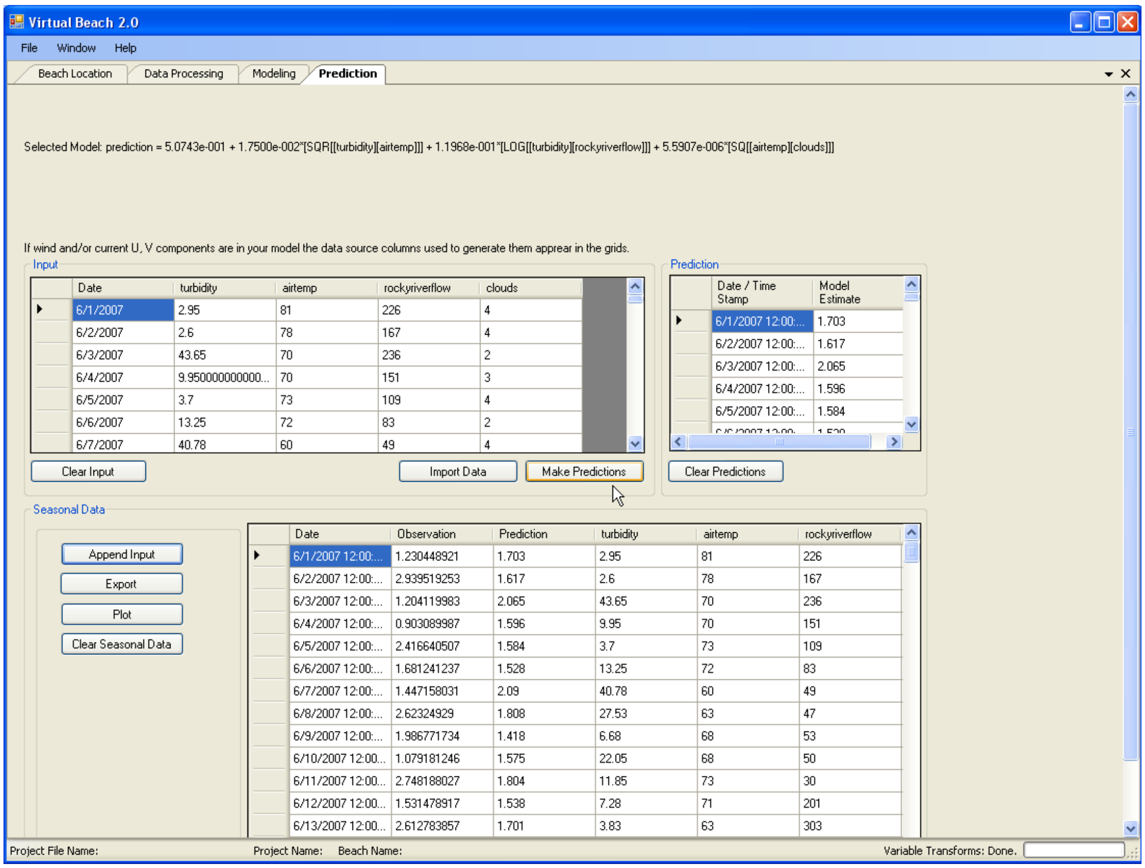


Figure 2.6. Prediction interface showing populated input data grid, prediction grid, and seasonal data grid.

2.3 Features Comparison: VB 2.0 Versus VB 1.0

VB1.0 and VB 2.0 are SS pathogen-predictor software packages designed to explore and analyze users' data sets to produce a *best-fit* MLR model and employ such a model for estimating future pathogen levels using environmental data collected at the beach site. Both versions use regression methodologies to measure the appropriateness of variables for inclusion in models.

VB 1.0 can be characterized as an MLR model building tool that supports a primarily manual analysis of data sets by visual inspection of data plots and manipulation of variables (e.g. transformations, creating interaction terms), followed by an iterative process of testing, comparing and evaluating models. The fitness of developed models is computed and tracked, allowing for comparison and eventual selection of a *best* model for the data set under consideration. This model can then be used to produce estimates of pathogen levels with current or forecasted environmental data from the site.

VB 2.0 enhances the functionality of its predecessor. VB 2.0 performs similar functions as VB 1.0 (visual inspection of univariate data plots, manual transformations of individual variables, MLR model building, prediction, and such) but also automates and extends the functionality in several ways:

- The Map component provides access to localized data sources (NOAA/NCDC data) through the map interface. Such data sources can provide recently collected or forecasted data, or both, for generating predictions using a chosen MLR model.
- The Map component provides a convenient method for defining the beach orientation by overlaying the beach on shore-line layers (satellite images, Google Maps, MS Virtual Earth, and the like). Given the orientation, VB 2.0 can calculate wind or current components (one component is parallel to shore, and one is perpendicular to the beach), which can be important predictor variables.
- Although the manual processing and analysis of imported data (visual inspection of univariate data plots and the transformations/interactions of variables) has been retained, the *Data Processing* component of VB 2.0 provides automated generation of all possible second-order interaction terms among IVs, as well as automated testing of a suite of variable transformations for improved model linearity. That functionality increases the number of models to evaluate during later model selection routines and removes the burden/difficulty of manual assessment placed on a user of VB 1.0.
- Multi-collinearity among predictor variables is handled automatically in the *Model Building* component; any model containing an IV with a high degree of correlation with other IVs (as measured by a large VIF; the threshold value is user-specified, but defaulted to 5) is removed from consideration during the model selection process.
- During the model selection process, MLR models are ranked by a user-selected evaluation criterion. Possible criteria include R^2, adjusted R^2, AIC, AICC, PRESS, BIC, Accuracy, Sensitivity, Specificity, or the model's RMSE. Regardless of what criterion is chosen, the software records the ten best models in terms of that criterion. In comparison, VB 1.0 had only a single comparative criterion, Mallows' Cp.
- As the number of IVs in a data set increases, the number of possible MLR models increases factorially (considering transforms/interactions) resulting in trillions of possible models from a modest number (12–13) of IVs. VB 2.0 implements a GA to effectively and efficiently search for the best possible MLR model. Instead of using the GA, VB 2.0 users can optionally perform an exhaustive calculation in which all possible combinations of IVs are used and tested (if the number of possible models is reasonably small). Both the GA and exhaustive approaches greatly expands the modeling building capabilities compared to VB 1.0.

- Users no longer have to enter data values in transformed, interacted, or component-decomposed form to make a prediction with a chosen MLR model. On the VB 2.0 *Prediction* tab, a user-selected model is coded into an input grid with data entry columns matching the model's main effects. Any mathematical manipulation of the IVs is then automatically performed before making predictions.

VB 1.0 was written in Object Pascal (Delphi) for DotNet frameworks usable on Windows-based PC systems. Excel v4.0 data files and text files with conforming data organizational specifications can be imported into the application.

VB 2.0 is written in C# for the DotNet framework 3.5 and is targeted for Windows desktop computers. VB 2.0 reads and writes Excel 2003, Excel 2007, and text-formatted data files. VB 2.0 uses several open-source, licensed and custom-written components: Extreme Numerics mathematical and statistical library, ZedGraph for plotting and charts, WeifenLuo's Windows Docking UI as a basis for the application's user interface, GMap.Net and Google's GeoCoding network services for the map interface, EPA's D4EM for NOAA/NCDC station locations, and an OLEDB library for Excel file access.

This page is intentionally blank.

3 Model Evaluation Strategies

3.1 INTRODUCTION

Many statistical metrics (AIC, BIC, PRESS, and such) allow a model developer to choose the *best* model from a large suite of potential models; they are discussed in Section 6.6 of Volume I of this report. This section focuses on a different, but important, question regarding modeling FIB densities at beaches: is the best model good enough? We know the chosen model is best in a relative sense, but how trustworthy are its predictions in an absolute sense? Several approaches can provide information that allows us to make such a determination.

Among other approaches to evaluate whether a chosen model is *good* in an absolute sense rather than *best* in a comparison to other potential models (see Chapter 3 on model evaluation methods), we used the model's adjusted R^2 and the RMSE to evaluate the fit achieved in regression of predicted versus observed values. Adjusted R^2 is similar to R^2, which is the proportion of variability in the response variable that can be explained by the model. However, in a model with more than one parameter, the adjusted R^2 is always lower than the R^2, as it incorporates a penalty for additional parameters (see Section 5.6, Volume 1). The RMSE is the square root of the sum of the squared residuals of the model, divided by the degrees of freedom for error in the model:

$$RMSE = [\Sigma(y_i - \hat{y}_i)^2 / (n - p)]^{0.5} \qquad \text{(equation 3.1)}$$

where y_i is the i_{th} observation in the data set; $\hat{y}_i$ is the fitted value of the i_{th} observation (found using the regression model); n is the number of observations in the data set; and p is the number of estimated parameters in the regression model. The adjusted R^2 and the RMSE are amenable to comparison between models derived from different data sets. However, in the case of the RMSE, comparisons should be made only if the response variable is the same in each data set.

One statistical method depends on classifying observations and predictions as *exceedances*, meaning they are greater than a given value (e.g., an EPA regulatory threshold for FIB levels at freshwater beaches). In Figure 3.1, the x axis shows observed values of the FIB response variable, and the y axis is corresponding predicted values (made with a regression model) of those observations. The interior of the plot is broken into four quadrants on the basis of two values. The regulatory standard is a vertical line that denotes a value above which water quality is considered to be out of compliance for human recreational use. The horizontal line (often set at the same value of the regulatory standard) represents a decision threshold marking the place on the y axis above which predicted values will be considered exceedances; below that line, predicted values will be considered non-exceedances. Therefore, the regulatory standard defines exceedances for observations, while the decision threshold defines exceedances for predictions.

The upper-right quadrant of the graph is where correct positives are found. Both the observation and the prediction are exceedances. Given that result, a beach manager rightfully would close the beach to the public. The lower-left quadrant denotes correct negatives. Both predictions and observations are non-exceedances, so the beach correctly could be opened for public use. Errors occur in the other two quadrants: the upper left quadrant indicates false positives. The model makes a prediction above the decision threshold, but the actual observation is below the regulatory standard, so the beach might be closed in error when no danger existed. The bottom-left quadrant indicates false negatives, when predictions fall below the threshold, but observations are actually above the standard. If the beach were open in such cases, public health would be threatened.

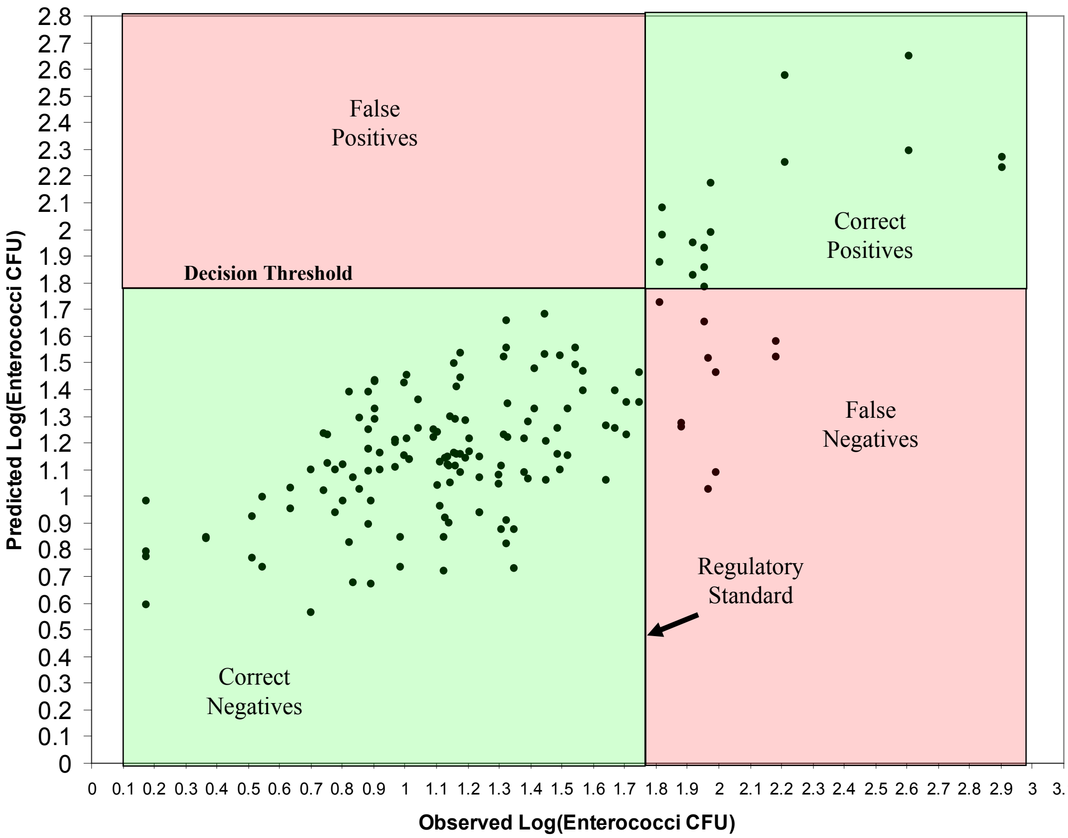


Figure 3.1. Plot of predictions versus observations for a regression model.

To assess how well a model is performing, three statistics have been developed on the basis of the data in Figure 3.1 (Francy and Darner 2006, 2007; Frick et al. 2008).

Specificity describes how well the model predicts non-exceedances:

> Specificity = Number of Correct Negatives / (Number of Correct Negatives + Number of False Positives)

Sensitivity defines how well the model predicts exceedances:

> Sensitivity = Number of Correct Positives / (Number of Correct Positives + Number of False Negatives)

Accuracy is the total percentage of correct predictions:

> Accuracy = (Number of Correct Positives + Number of Correct Negatives) / Total Number of Observations

All three metrics are fractions that can vary from 0 to 1, with 1 being ideal. It would be up to the analyst or beach manager to decide if a model's specificity, sensitivity, and accuracy are high enough to be considered *good*. Note that analysts can choose to raise or lower the decision threshold to alter the sensitivity and specificity of the model. They might do that after considering the relative costs of false positives (lost economic revenue from the beach being closed) to false negatives (public health endangered). For example, they might decide to close the beach if the model makes a prediction above 1.5 instead of above 1.8, as in Figure 3.1. That would take some observations out of the false negative quadrant and put them into the correct positive quadrant. However, it would also take observations out of the correct negative quadrant and put them into the false positive quadrant. In such a scenario, a manager has to decide whether to sacrifice specificity for improved sensitivity. VB 2.0 allows for the use of specificity, sensitivity, and accuracy as decision criteria in model selection.

We note that when a data set is composed of many FIB measures that fall below an instrument detection limit, it might be advisable to reformulate the response variable as a binary exceedance measure, and then use other statistical techniques designed for that purpose (e.g., binary logistic regression). However, modeling exceedances at a beach with only a small percentage of exceedances is extremely difficult. For statistical modeling, it is ideal for a beach to have between 25 and 75 percent exceedances.

We investigated another way to assess model goodness-of-fit using a process called cross-validation in which a data set is split into two groups. One is called the *training* data set, and the analyst will choose a model (using a model selection approach of choice) that best fits the training data. It will then be used to make predictions for the second group, termed the *testing* data set. Doing so easily shows how well the best model makes predictions for data it has not seen before. If done hundreds of times by randomly splitting the data into training and testing sub-groups, one can develop robust statistics that quantify the predictive capabilities of the chosen model. Our objective was to investigate relationships between the errors seen in the training data set, errors seen in the testing data set, the measured goodness-of-fit, sample size considerations, and other characteristics of a data set to see if useful generalizations could be drawn. The fundamental question is whether characteristics of a data set (and the best model for that data set) exist that will determine if that model will provide good predictions.

3.2 METHODS

We define the mean squared error of fitting (MEF) for a training data set as

$$MEF = \Sigma(y_i - y_{ifit})^2 / n_{train} \qquad \text{(equation 3.2)}$$

where the sum is over all observations in the training data set, y_i is the i_{th} observation in that data set, y_{ifit} is the fitted value for the i_{th} observation, and n_{train} is the number of observations in the training data set. The MEF is similar to the commonly used mean squared error, or MSE, except that the denominator of the MSE is n_{train} minus the number of parameters in the regression model. We define the MSE of a testing data set as MEP (mean squared error of prediction):

$$MEP = \Sigma(y_i - y_{ipred})^2 / n_{test} \qquad \text{(equation 3.3)}$$

where the sum is over all observations in the testing data set, y_i is the i_{th} observation in that data set, y_{ipred} is the predicted value for the i_{th} observation, and n_{test} is the number of observations in the testing data set.

Because of its large sample size, we performed our analysis on the Huntington Beach, Ohio, 2000–2009 data set. The response variable was *E. coli* colony forming units (CFU) measurements taken by Ohio (for a site description, see Section 4.2.5). The data set has many more observations (709) than IVs (11). When that is not the case, MLR techniques might not be the most appropriate; instead, Partial Least Squares regression techniques are often implemented(Haenlein and Kaplan 2004; Hou et al. 2006). For our study, we performed the following steps:

1. Take a random sample of size *j* (*j* varied from 30 to 700) from the original data set.
2. Randomly split the subsample into training (50 percent) and testing (50 percent) data sets.
3. Use a backwards-stepwise model selection routine (based on minimizing the AIC statistic) to choose a linear regression model that best fits the training data.
4. Record the MEF, AIC, adjusted R^2, and number of parameters for the model.
5. Use the model to make predictions on the testing data, and record the MEP of the predictions.
6. Go back to step 1.

For each value of *j*, we performed the steps a total of 150 times and then examined the output. Note that by making the number of observations in the training and testing data sets equal, the denominators of equations 3.2 and 3.3 are equivalent so, essentially, we compared the sum of squared errors within both data sets. Code written in the R statistical language was used to generate the results because VB 2.0 does not have the required functionality to produce the desired output.

3.3 Results at Various Sample Sizes

As might be expected, the single most important factor in determining sample statistics such as the MEF, MEP, and adjusted R^2 was the size of the sample taken from the data set. Figures 3.2 through 3.7 demonstrate those relationships. The x axis in the graphs refers to the size of the training (and testing) data set. It is important to note that, for all following graphs, studies done with data sets from other sites would display very similar trends. What would change are the values on the y axis, which are site-dependent. In Figure 3.2, we see that taking larger samples from the population leads to an increase in the number of significant parameters in the final model. That increase seems to level off at a sample size of 350, and there is great variability in the smallest sample sizes (n = 15 and n = 20). For another site, number of parameters could vary between 2 and 6, or 12 and 18; however, the upward trend of more significant parameters with increasing sample size would hold.

The MEF increases with increasing sample size (Figure 3.3), meaning small samples are easier to fit than large samples. That is supported by Figure 3.4, which shows that the adjusted R^2 declines as sample size increases. Both statistics leveled off at a sample size of around 100. Those findings are important for interpreting later results that detail modeling efforts across many marine and freshwater sites. We found a wide range of adjusted R^2 values between sites, and some of that variability is due to differences in the complexities of water quality dynamics at the various beaches; however, some of the variability in model goodness-of-fit must be attributed to sample size differences in the data sets: a small data set is expected to have a higher adjusted R^2 compared to a much larger data set.

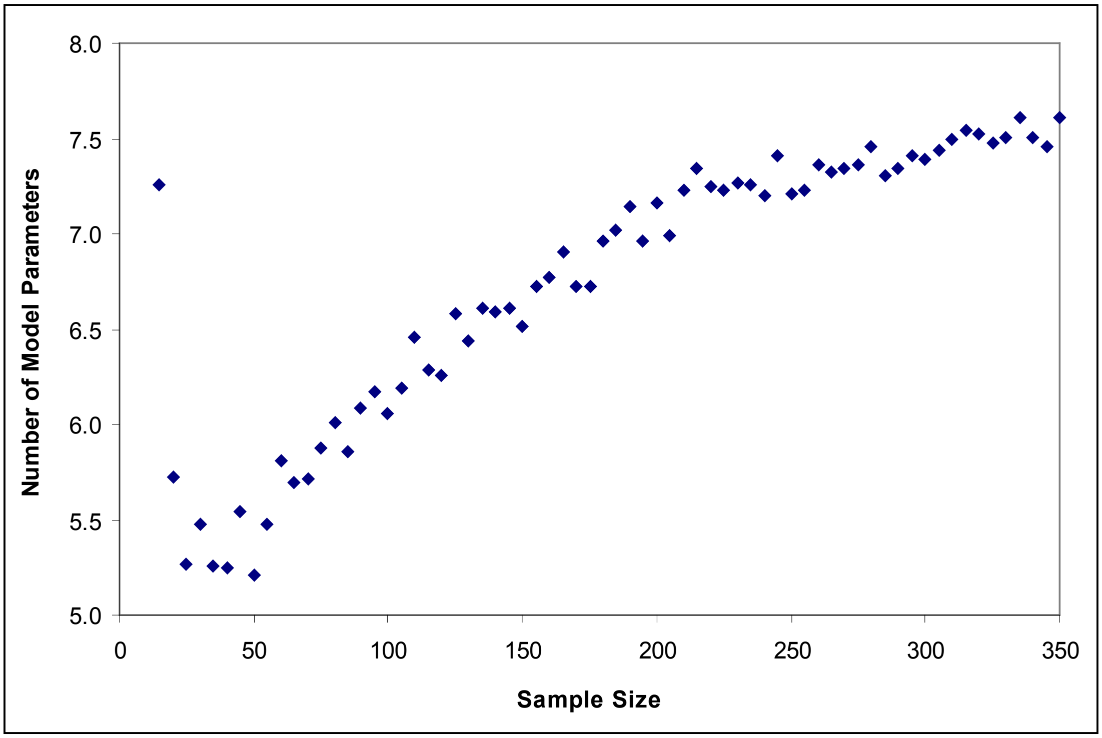


Figure 3.2. Increasing sample size leads to more significant parameters in the chosen model.

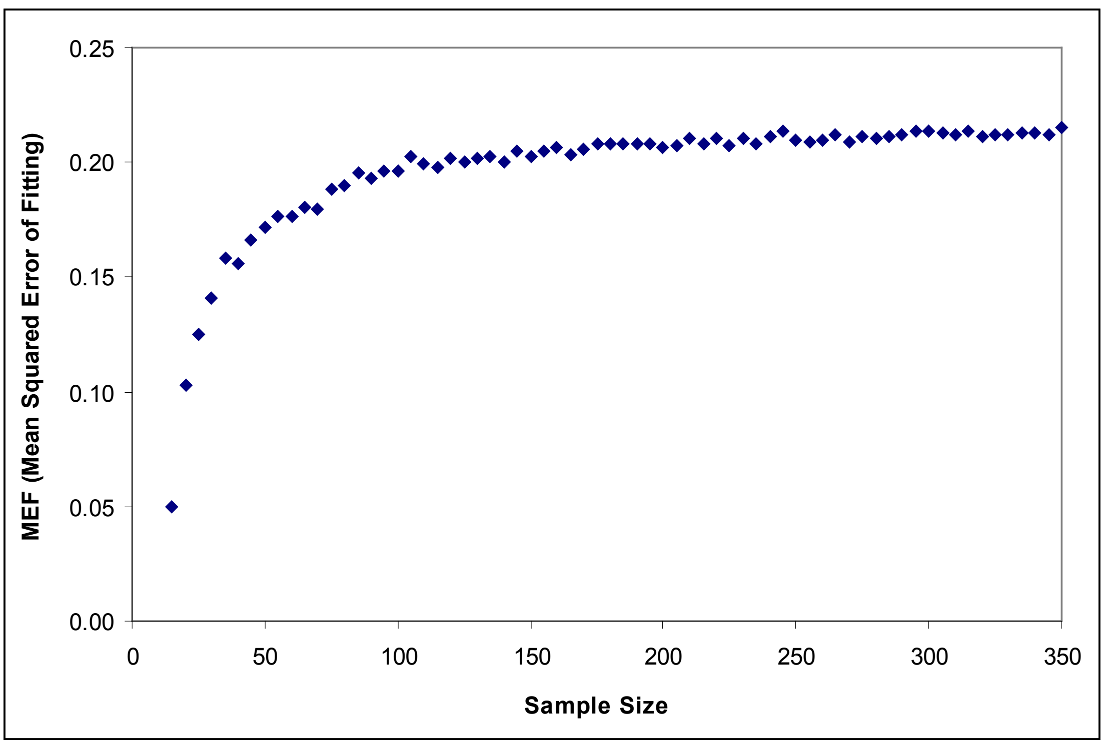


Figure 3.3. The mean squared error of fitting (MEF) increases with increasing sample size.

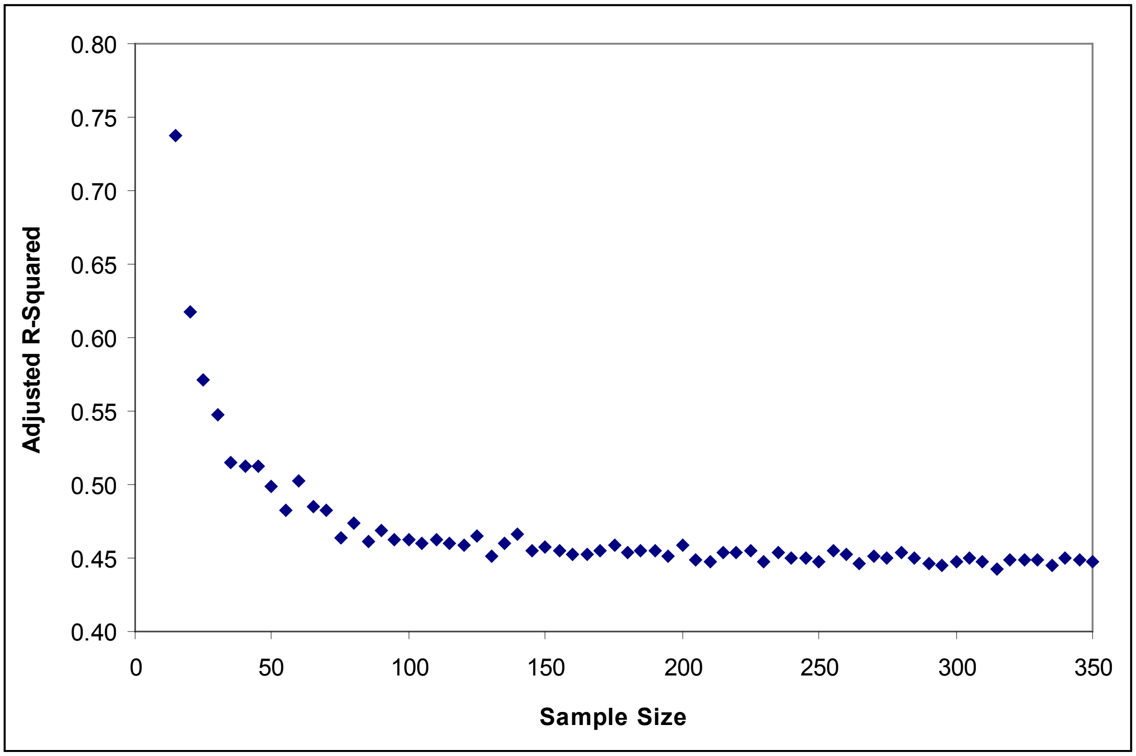


Figure 3.4. The adjusted R^2 declines as sample size increases.

The MEP declines as sample size increases (Figure 3.5). That statistic is close to leveling off at a sample size of about 50. Prediction errors are expected to be much higher for samples sizes less than about 20. Again, at another site, the magnitude of prediction errors would differ depending on the complexity of FIB dynamics at the site; however, the pattern shown in Figure 3.5 should be the same.

The effect of the sample size on the ratio of MEP to MEF is shown in Figure 3.6. As sample size increases, the ratio moves toward 1.0, meaning that predictive errors are about equal to fitting errors. Figure 3.7 zooms into a small portion of the y axis to show in greater detail how the ratio is changing for samples sizes greater than 50. At a sample size of 50, the MEP is approximately 90 percent greater than the MEF. At the largest sample size we could examine (350), the MEP is about 10 percent larger than the MEF.

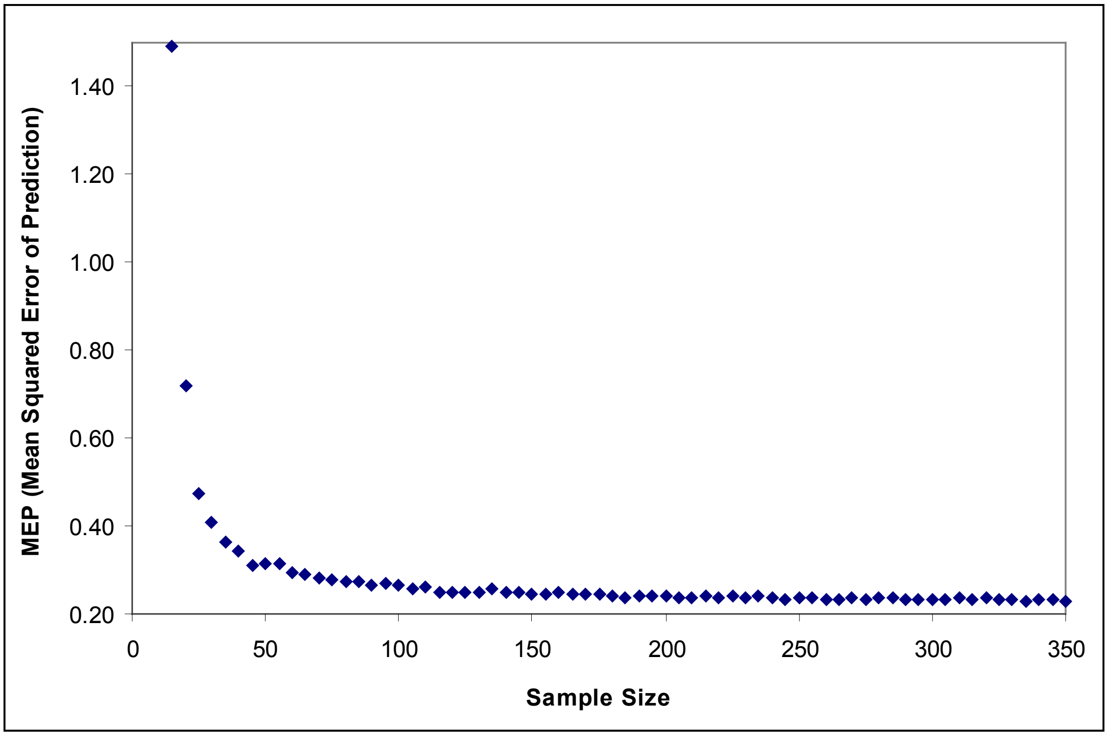


Figure 3.5. Declining mean squared errors of prediction with increasing sample size.

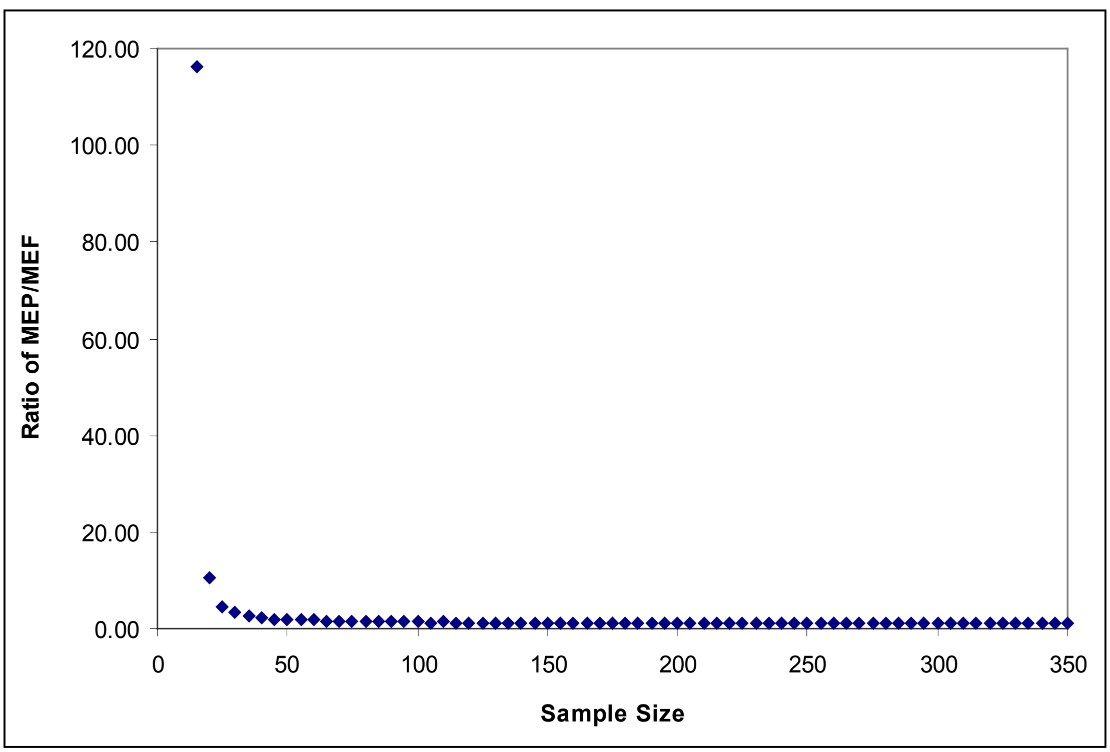


Figure 3.6. Sample size effect on the MEP/MEF ratio.

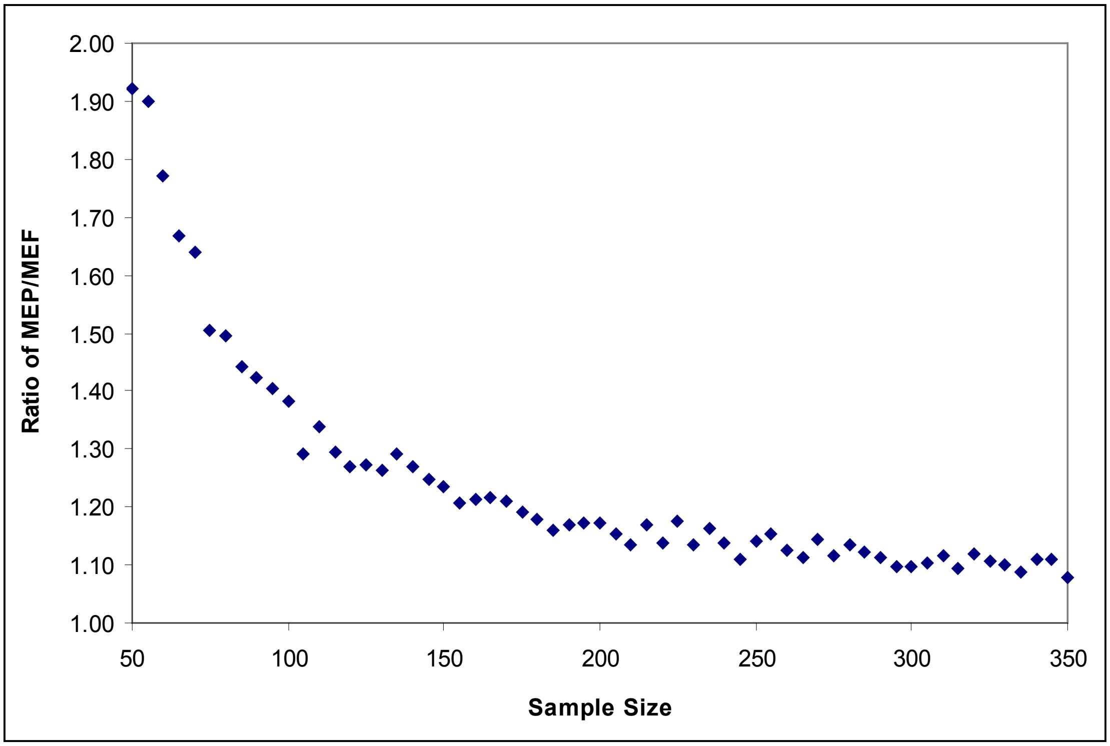


Figure 3.7. Effect of sample size on MEP/MEF for sample sizes greater than 50.

3.4 RESULTS WITHIN A SAMPLE SIZE

Increasing sample size has an observed effect on the various model statistics, but how do those quantities vary within a sample of a certain size? These results, because of sample variability, are more representative of what can be expected when models across many different beach sites are examined. Results are presented for a small training sample (n = 20), an intermediate-sized training sample (n = 75), and a large training sample (n = 200). For each of the sample sizes, 150 random samples of size 40, 150, and 300 were drawn from the population of more than 700 observations, and then randomly and equally split into training and testing subgroups. A model was fit to the training data, and then predictive errors were calculated for the testing data. The adjusted R^2 for the model fit to the training data was plotted against the MEP for the corresponding testing data. That enables evaluation of whether a model's level of fit to the training data can be used to indicate its expected success in making predictions.

3.4.1 Small Sample Size

For the smallest sample size, as the adjusted R^2 of the model fit to the training data increased, prediction errors and their variability increased (Figure 3.8). That result is somewhat counterintuitive and deserves restating: for a small sample size, the better the model fit the training data, the greater was the probability that prediction errors from using that model could be very high. Indeed, the greatest prediction error was seen in a sample with an adjusted R^2 over 0.9. How can such a counterintuitive result be explained? Basically, when such a small sample size is taken from the population, there is a possibility that members of that sample can be fit very well to a very specific model. However, those individuals can be very different from others in the population, meaning that when that model is used for predictive purposes, large errors are found.

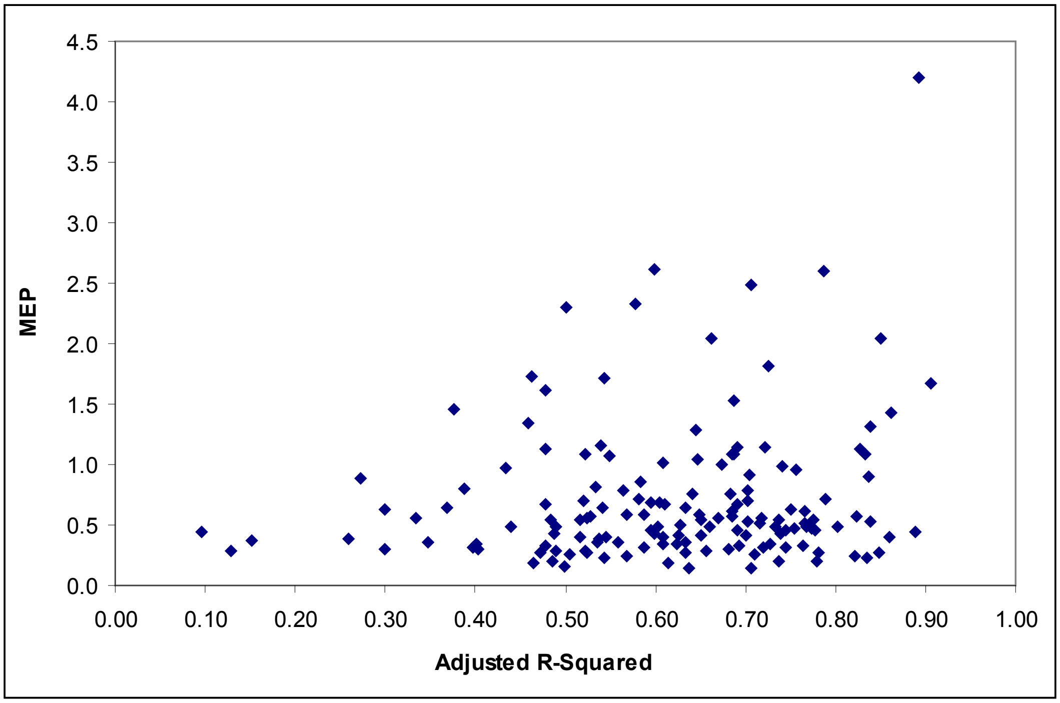


Figure 3.8. Predictive errors increase with increased adjusted R^2. Each of the 150 points represents the predictive errors of a testing data set of size = 20 observations.

3.4.2 Intermediate and Large Sample Sizes

For intermediate-sized and large samples, no discernible relationship exists between adjusted R^2 and the MEP (Figures 3.9 and 3.10). Thus, at those sample sizes, the fit of the model to the training data provides no information about how well the model will make predictions.

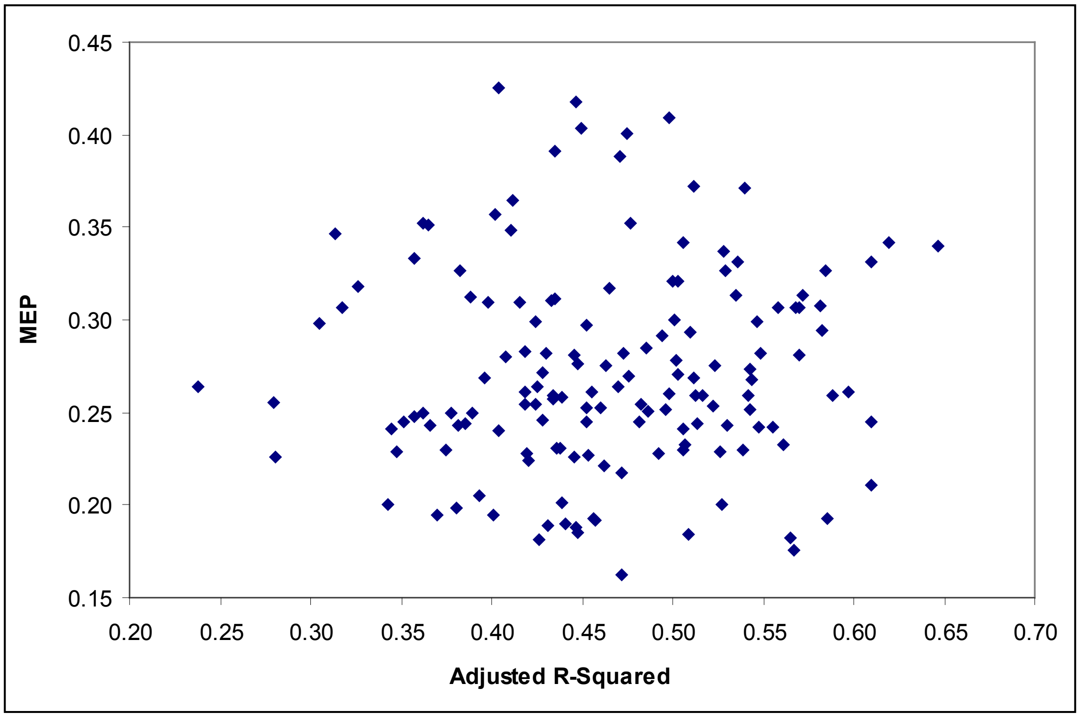


Figure 3.9. Relationship between adjusted R^2 and MEP for 150 samples of size 75.

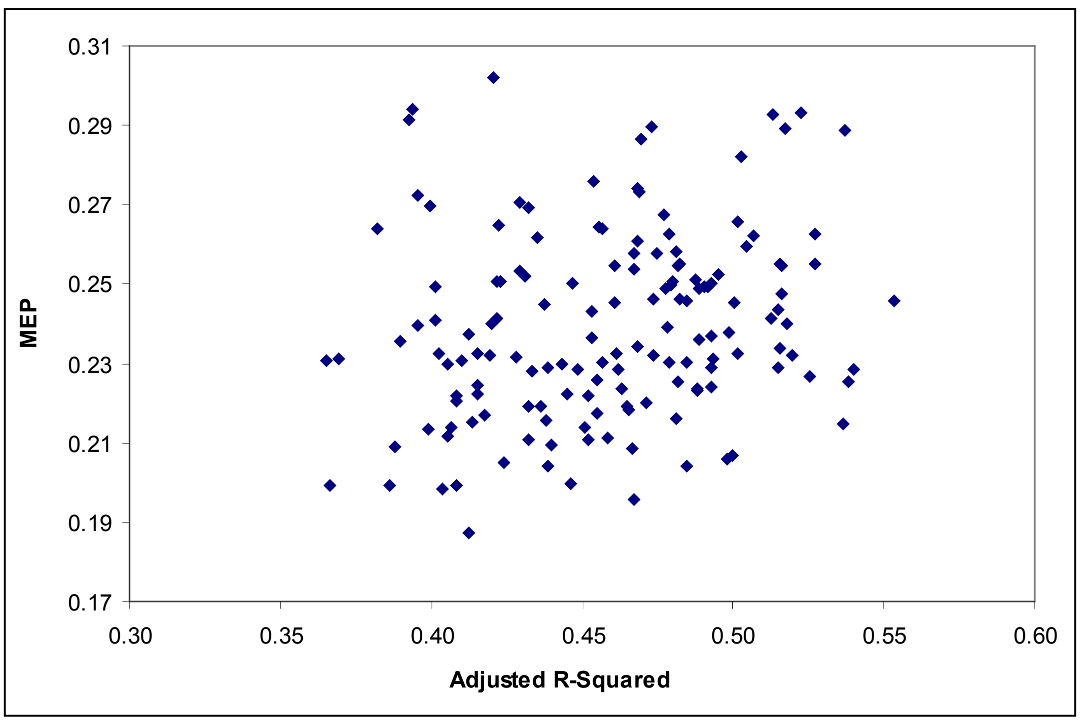


Figure 3.10. Relationship between adjusted R^2 and MEP for 150 samples of size 200.

3.4.3 Effect of Response Levels on Predictions

Finally, we used a relatively small training sample size (n = 30) to examine if any relationship existed between the magnitude and variability of the response (the mean and standard deviation of the FIB levels in the sample) in the training data and the magnitude of the predictive errors. In other words, do models developed for beaches with greater FIB levels (or less variable FIB levels) make better predictions than models developed for beaches with low FIB levels? If such relationships exist, they might lead to guidelines for determining if regression models are appropriate for a specific beach site.

Unfortunately, Figures 3.11 and 3.12 show no strong relationships between either the mean value or the standard deviation of the response variable and the mean errors of prediction. That indicates that a manager presented with a data set could not infer much about the predictive ability of a model developed using the data set from the mean or variation of the FIB levels in that data set.

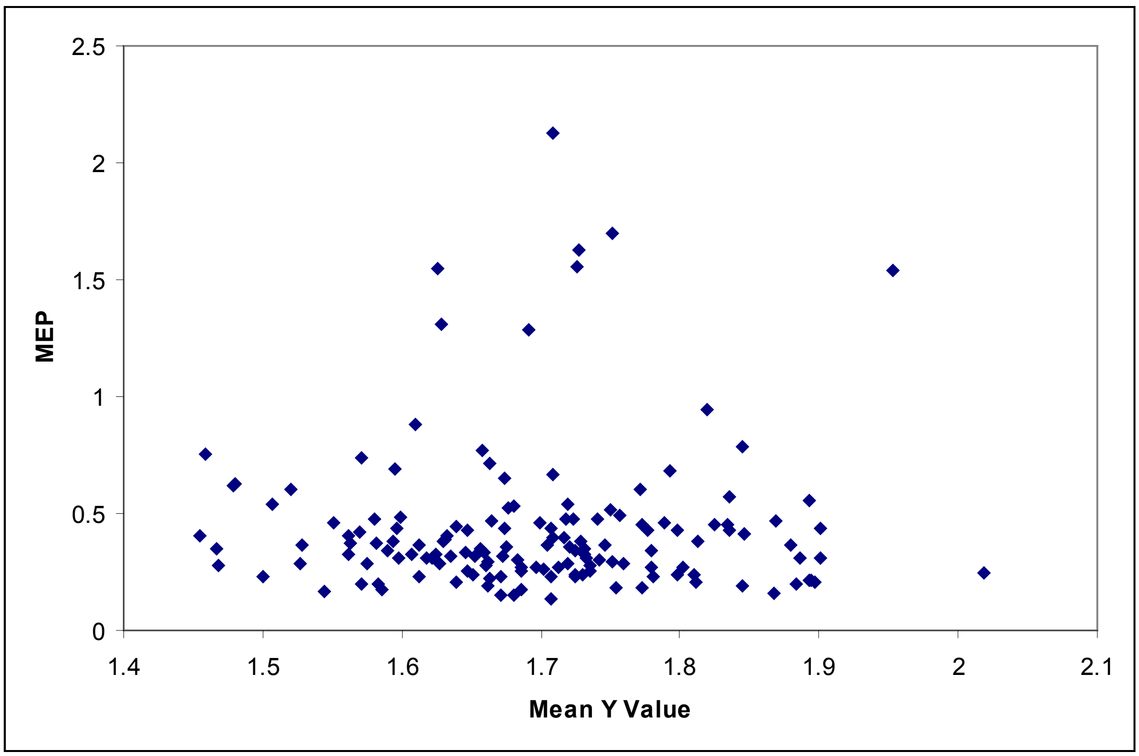


Figure 3.11. No clear relationship between the mean Y value in a training data set (n_{train} = 35) and the resultant MEP of the testing data set.

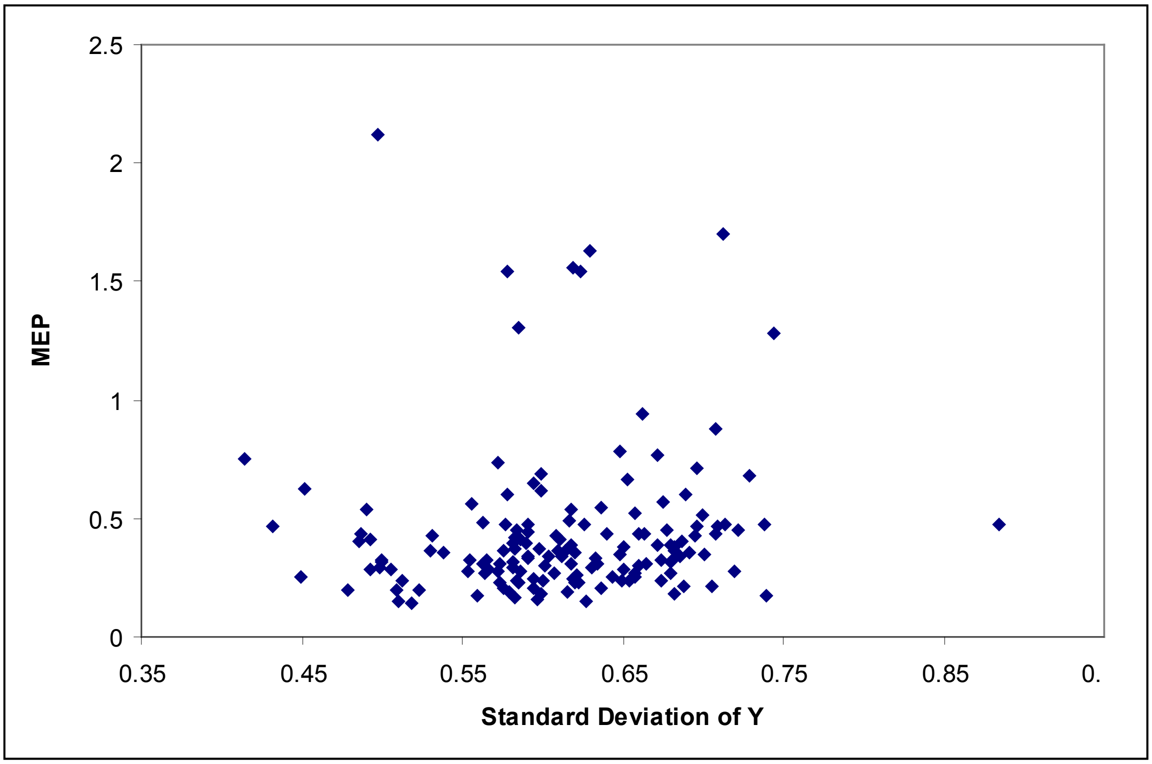


Figure 3.12. No clear relationship between the standard deviation of Y values in a training data set (n_{train} = 35) and the resultant MEP of the testing data set.

In the end, there is no substitution for careful monitoring of the predictions made by statistical models. As predictions are made, it is important to go back later and validate them against real data to see if the model is performing adequately. A few weeks' worth of MEP values at least twice as large as the MEF values, or more than a handful of false negatives or false positives (or both) over a month's time, can be sufficient to mandate developing a new empirical model/alternative modeling approach, or at the very least ceasing to rely on the current one. On the basis of the data in Figures 3.5 and 3.7, we recommend that at least 50 observations be used for model development, but 100 or more are preferable.

4 Study Sites and Data Acquisition

4.1 INTRODUCTION

To provide the data for our modeling research, we developed a program designed to enhance the *predictive modeling of indicator exposure research* (PREMIER). PREMIER includes a concurrent collection of culturable and qPCR microbial data, as well as IVs for model refinement at selected freshwater beaches on the Great Lakes and at coastal marine beaches of the eastern United States and in the tropics. Several of the beaches have nonpoint sources of fecal contamination, but most were contaminated by nearby publicly owned treatment works. We deliberately patterned the spatial and temporal sampling patterns of the studies after those used in EPA's National Epidemiological and Environmental Assessment of Recreational Water (NEEAR) studies (Haugland et al. 2005; Wade et al. 2006). At PREMIER beach sites in Milwaukee, Wisconsin; Surfside Beach, South Carolina; Miami, Florida; Luquillo, Puerto Rico; and Boquerón, Puerto Rico, we used automated techniques for IV measurements that were not used in most NEEAR studies or in other modeling studies. Such techniques were prompted in part by consideration of processes that are known to affect the fate and transport of microorganisms (Figure 4.1) (Boehm 2007; Haugland et al. 2005; USEPA 2007). Examples are concurrent measurements of underwater solar ultraviolet (UV) radiation, currents and waves (using acoustic Doppler current profilers or ADCPs); meteorological data; and water quality parameters, such as turbidity, dissolved organic carbon (DOC), UV-visible spectra, dissolved oxygen, salinity, or conductivity. Automated collection of the data permitted subsequent analysis of the time-dependence of modeling results (see Section 6.3 on temporal synchronization).

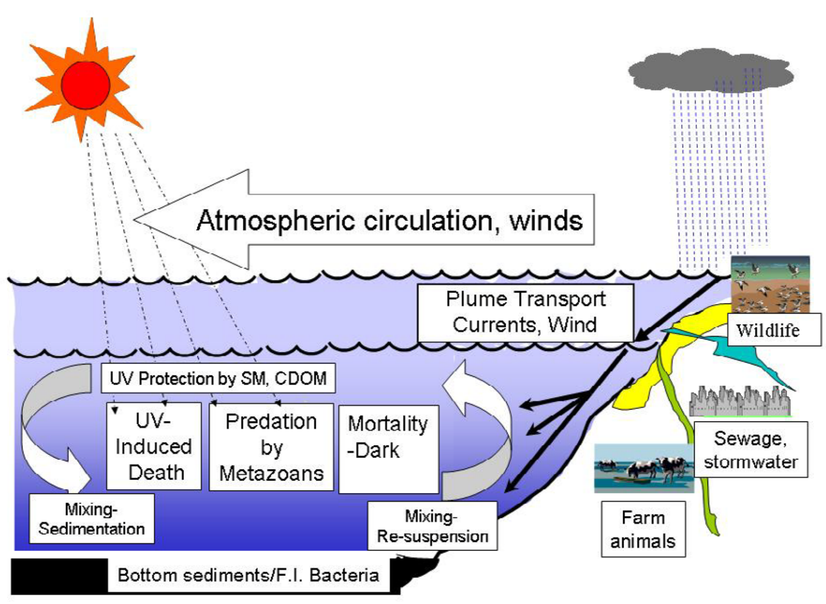
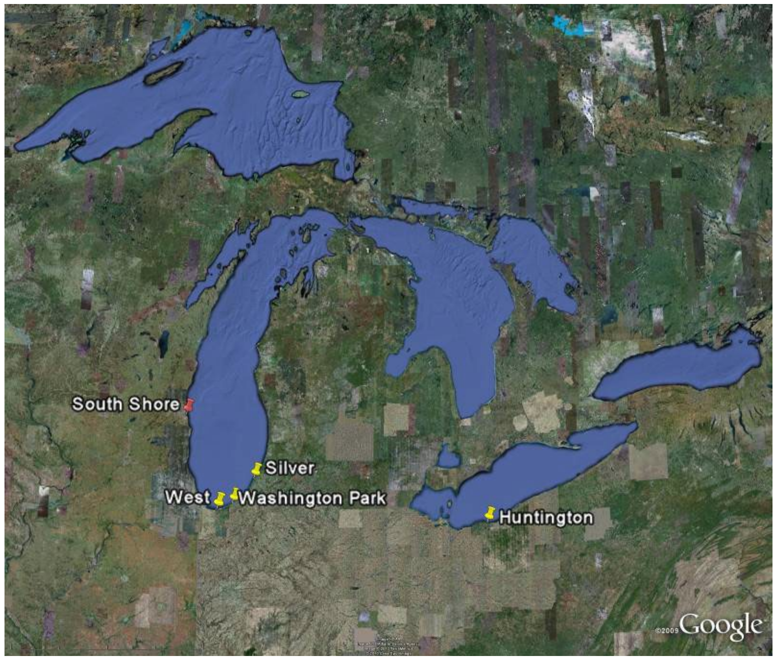


Figure 4.1. Processes considered in collecting IVs for predictive modeling studies.

During the 2008 and 2009 summer beach seasons, samples were collected at the PREMIER recreational water sites to measure culturable and qPCR-based *Enterococcus*, while automated field equipment was deployed to monitor various environmental parameters. The resulting data, and existing data sets from the NEEAR studies, were used in the predictive modeling studies presented here. Beaches in both the PREMIER and NEEAR studies include six freshwater Great Lakes beaches (Figure 4.2A), a temperate marine beach, four subtropical marine beaches, and two tropical marine beaches (Figure 4.2B). In addition to differences in water type (fresh versus marine) and climate (temperature, rainfall), factors such as tidal change, wave energy, and location relative to contamination sources likely will also influence the usefulness of certain IVs. The beaches represent a range of different contamination sources and include sites that potentially are affected by point sources (i.e., treated wastewater effluent) and nonpoint sources (e.g., urban runoff, birds) of FIB.

Descriptions of the freshwater and marine study sites, including location, potential sources of fecal contamination, and historical water quality information are given in the following sections and Appendix A to provide background and context for the modeling results. Such information will help to evaluate and explain the usefulness of certain IVs in predictive statistical models. Details regarding the weather stations, ADCPs, water quality sondes, and underwater light sensors deployed at the beaches are provided, as are the methods used to determine culturable and qPCR-based enterococci.

(A)



(B)

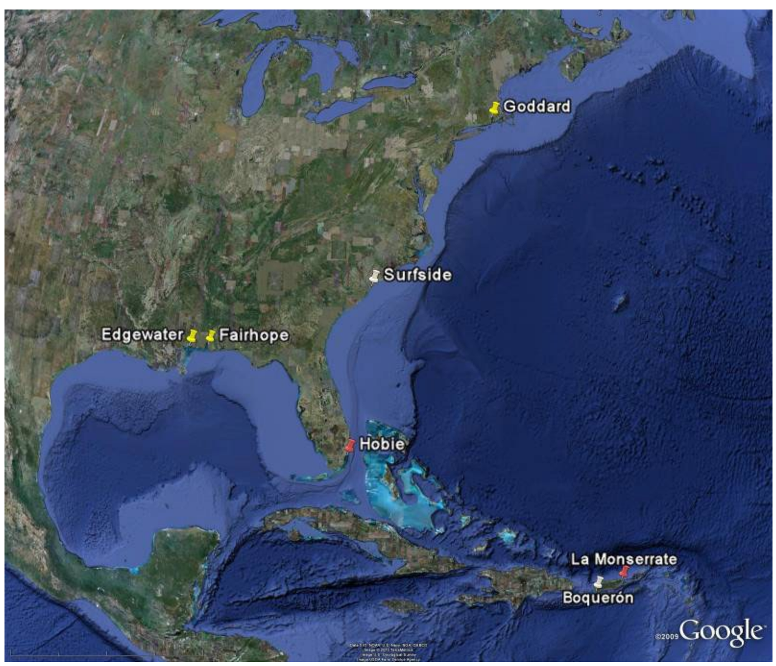


Figure 4.2. Location of (A) freshwater and (B) marine beach sites. Yellow and red pins indicate the NEEAR and PREMIER study sites, respectively. Joint NEEAR/PREMIER studies were carried out at Surfside Beach and Boquerón (white pins).

4.2 FRESHWATER BEACHES (GREAT LAKES)

4.2.1 South Shore Beach, Milwaukee, Wisconsin

South Shore Beach is located on the western shore of Lake Michigan (Figure 4.2A) in South Shore Park in Milwaukee, Wisconsin. South Shore Beach is a public beach area (~ 0.01 square kilometer [km^2]) with 150 meters (m) of sandy shoreline. Its shore is low sloping, and the benthic zone is muddy and sandy. The beach is adjacent to the South Shore Yacht Club and a ~ 0.02 km^2 paved parking area that drains into the lake. A 20–m-long rock embankment juts into the lake, separating the sandy beach area from a 185-m-long cobble/pebble beach area with a high sloping shore (South Shore Rocky Beach). The beach is occupied consistently by large flocks of roosting waterfowl and shore birds. Ring-billed gulls are the predominant species, but Canada geese, Mallard ducks, and pigeons are also present (McLellan and Salmore 2003; Scopel et al. 2006). The entire beach and marina area is partially enclosed by a breakwall ~ 300 m offshore, which limits wave action, water circulation and exchange with the outer harbor**.** Water depths within the breakwall are < 5 m, with depths < 2 m within 50 m of shore. The beach is ~ 4 km south of Milwaukee Harbor, which is the site of the Milwaukee Metropolitan Sewerage District Jones Island Water Reclamation Facility. The Milwaukee, Menomonee, and Kinnickinnic rivers also discharge to Lake Michigan inside the Milwaukee Harbor breakwall.

Historically, South Shore has poor water quality, with 34 percent of samples collected from 2003 to 2009 exceeding water quality criteria standards (Appendix A). Potential sources of fecal contamination include combined sewer overflows (CSOs); urban/suburban and agricultural runoff from the Milwaukee River Basin; runoff from impervious surfaces including parking lots and the beach face; and gulls (Scopel et al. 2006). However, high *E. coli* counts are not always attributable to rainfall and CSO events (McLellan et al. 2001; Scopel et al. 2006). A detailed spatial assessment found that poor beach water quality was mostly a local phenomenon, with contamination originating at the shoreline (McLellan and Salmore 2003). Previous modeling efforts at South Shore Beach include hydrodynamic and water quality models developed to describe the fate and transport of fecal coliform in Milwaukee Harbor and nearshore Lake Michigan. Calibration and validation of the models was accomplished using Milwaukee Metropolitan Sewerage District and Great Lakes WATER Institute field data. Modeling results indicate that the fecal coliform load from rivers and CSO/sanitary sewer overflow events had only slightly more than marginal impact on the beach site, with local sources (e.g., stormwater runoff and birds) being more important (MMSD 2005). South Shore Beach was also one of the 55 beaches included in a regional forecast model for southern Lake Michigan, from Milwaukee, Wisconsin, to Michigan City, Indiana (Whitman and Nevers 2005). The City of Milwaukee is considering adopting a predictive model. In addition to issuing advisories on the basis of the monitoring of *E. coli* levels (where < 235 CFU/100 milliliters (mL) is acceptable, 235–1,000 CFU/100 mL results in a water quality advisory, and > 1,000 CFU/100 mL results in a closure advisory), a rainfall threshold of 2.5 centimeters (cm) in 24 hours is also used to predict poor water quality and issue advisories at South Shore.

4.2.2 West Beach, Porter, Indiana

West Beach is on the south shore of Lake Michigan (Figure 4.2A), ~ 40 km southeast of Chicago Harbor; it is part of the Indiana Dunes National Lakeshore Park in Porter, Indiana. This public beach has ~ 2 km of sandy shoreline and contributes to the total 10-km length of contiguous swimming beaches along this section of Lake Michigan. The shore has a gradual slope, and water depths are < 2.5 m within 250 m of shore. As an open, unprotected, lakefront beach, West Beach is subject to moderate wave exposure. The beach area is west of the Portage-Burns Waterway (Burns Ditch), a man-made channel that serves as the outfall point for the Little Calumet River to Lake Michigan. Discharge from Burns Ditch is directed west by a breakwall and then dispersed by lake currents and waves (Appendix A).

The West Beach and Burns Ditch sites have been the subject of several predictive modeling studies. Multiple regression analysis found that a combination of water quality variables could account for the observed variability in *E. coli* at Burns Ditch outlet during storm events (Olyphant et al. 2003). Further efforts led developing regression models that predict beach *E. coli* levels from wave and precipitation data, as well as water quality parameters. Because of the role of wind direction on the transport of Burns Ditch discharge to the beaches (Nevers and Whitman 2005), model performance was improved by separately modeling days with onshore and offshore winds. Regional forecast models also were developed for southern Lake Michigan, which includes West Beach (Whitman and Nevers 2005; Nevers and Whitman 2008). In addition to *E. coli*, more recent predictive models for beaches affected by Burns Ditch include culturable and qPCR-based enterococci as dependent variables. Differences in the models developed for each indicator (i.e., IVs identified for culturable compared to qPCR-based enterococci) provide evidence of the different processes that determine their fate (Byappanahalli et al. 2010; Telech et al. 2009). Beach monitoring practices rely on culturable *E. coli* where levels > 235 CFU/100 mL result in a contamination advisory or closure, as well as on a rainfall threshold. In addition, the Project S.A.F.E. (Swimming Advisory Forecast Estimate) predictive model was implemented as a USGS pilot study (Whitman 2008) and is used in conjunction with the NOAA/Great Lakes Environmental Research Laboratory Indiana Dunes Nowcast hydrodynamic model (Schwab et al. 2010) to provide real time *E. coli* estimates at West Beach.

4.2.3 Washington Park Beach, Michigan City, Indiana

Washington Park Beach is in Michigan City, Indiana on the south shore of Lake Michigan (Figure 4.2A). The public beach area is ~ 1.1 km long and is immediately east of a breakwall that directs discharge from Trail Creek to the west. Trail Creek is a source of *E. coli* to the lake with the potential to affect water quality at nearby beaches (Nevers et al. 2007). The Trail Creek watershed drains urban, agricultural, and residential areas, with a number of human and animal nonpoint sources including agricultural field drainage and runoff, cattle/steer grazing, failing septic systems, illicit connections, and urban stormwater runoff (Triad Engineering Incorporated 2003). In addition, the Michigan City Sanitary District Wastewater Treatment Plant (WWTP) (~ 3 km upstream of Lake Michigan), which applies chlorine disinfection during the summer months, is a major discharger to the creek (Wade et al. 2008), although plant improvements have practically eliminated CSO events (Nevers et al. 2007). Water quality at Washington Park Beach is poor, with the recreational water quality criteria for *E. coli* exceeded in 23 percent of the samples collected from 2005 to 2009. In addition, pathogens (adenoviruses and enteroviruses) and a marker for human sewage were detected in samples collected from the swimming area during summer 2004 (Wong et al. 2009).

Predictive modeling of *E. coli* has not been reported at Washington Park Beach, but work was done at another Trail Creek-influenced beach, Mount Baldy Beach, ~ 2.5 km west of the mouth of Trail Creek (Nevers et al. 2007). The Mount Baldy site was also incorporated in a southern Lake Michigan regional forecast model (Whitman and Nevers 2005; Nevers and Whitman 2008). The observed variation in *E. coli* levels at Mount Baldy was further explained, based on loadings from Trail Creek and nearby Kintzele Ditch, using a process-based hydrodynamic model (Liu et al. 2006; Thupaki et al. 2009).

4.2.4 Silver Beach, St. Joseph, Michigan

Silver Beach is in St. Joseph, Michigan, on the southern end of the eastern shore of Lake Michigan, ~ 55 km northeast of Washington Park Beach (Figure 4.2A). The ~ 0.6-km-long sandy beach area is just south of where the St. Joseph River flows into Lake Michigan. At its mouth, the river is lined by two parallel navigational piers that extend into the lake, guiding riverine discharge ~ 0.5 km out, roughly perpendicular to the shoreline.

Possible sources of fecal contamination to the St. Joseph River watershed, which encompasses both urban and agricultural areas, include seven WWTPs (that use chlorine disinfection), four of which are on the river (Wade et al. 2008), CSOs, stormwater discharges, agricultural inputs, and illicit discharges (MDEQ 2003). The river is the receiving water for the St. Joseph-Benton Harbor WWTP, which is located ~ 2.7 km upstream of Lake Michigan. Only 1 percent of samples collected from 2001 to 2009 exceeded local water quality criteria standards for *E. coli* (Appendix A).

Regression modeling was employed at Silver Beach to investigate relationships between culturable and qPCR-based enterococci and various environmental conditions in an effort to identify useful predictors of water quality (Telech et al. 2009). Routine water quality monitoring at the beach is based on a culturable *E. coli* criteria of 300 CFU/100 mL and is supplemented by the use of a *rainfall plus 48-hour health advisory*.

4.2.5 Huntington Beach, Bay Village, Ohio

Huntington Beach is part of Huntington Reservation of Cleveland Metroparks, which is in Bay Village, Ohio (a western suburb of Cleveland), on the southern shore of western Lake Erie (Figure 4.2A). The swimming area is situated just west of Porter Creek, and the beach is ~ 8 km west of the Rocky River mouth. The ~ 0.5-km-long sandy beach area is broken into segments by a series of rock jetties (< 100-m long) that run perpendicular to the shoreline (Appendix A). The breakwalls limit water circulation in the swimming area.

Huntington Beach is in the Black-Rocky Watershed, within which are 10 sewage discharge locations that could affect water quality at the beach. Three of those flow directly into Lake Erie: Avon Lake WWTP, ~ 11 km west of the beach; Rocky River WWTP, ~ 6 km east of the beach; and Westerly WWTP, ~ 18 km east of the beach. The others are discharged to the Rocky River or its tributaries, the closest being Lakewood WWTP (< 3 km from the mouth of the Rocky River). The majority of the treatment plants use UV or chlorine disinfection during the summer

(Wade et al. 2008; Wade et al. 2006). Porter Creek, on the east end of the beach, was identified as a likely contributor to high *E. coli* counts at beach (Huang and Sigler 2006). In addition, two outfalls discharge storm runoff from the parking lot to the beach (Francy et al. 2003).

Water quality at Huntington Beach is poor, with 16 percent of samples collected from 2000 to 2009 exceeding water quality criteria standards for *E. coli*. During the summer beach season, the Cuyahoga County Board of Health and Cuyahoga County Sanitary Engineers Laboratory, monitor Huntington Beach water quality (i.e., *E. coli*). In addition, a number of ancillary environmental parameters (i.e., water temperature, turbidity, and wave height) are also measured at the time of sample collection. Those data (real-time and historical since 2006) are publicly available through the Ohio Nowcast website, along with the associated water quality predictions and advisories (http://www.ohionowcast.info/nowcast_huntington.asp). Real-time and historical discharge and gage height data are available from the USGS Rocky River gauging station near Berea, Ohio (04201500), which is ~ 20 km upstream of Lake Erie (http://waterdata.usgs.gov/nwis/uv?04201500).

As a result of ongoing efforts of the USGS Water Science Center and its partners in daily routine sampling and data collection, an extensive, multiyear data set is available for Huntington Beach. The focus of the work was to develop an MLR model that continues to be tested and refined over time by incorporating additional years of data. The Ohio Nowcast predictive model is employed to supplement *E. coli* data (235 CFU/100 mL criteria) when making decisions about swimming advisories.

4.3 Marine Beaches

4.3.1 Goddard Beach, West Warwick, Rhode Island

Goddard Beach is in Goddard State Memorial Park in West Warwick, Rhode Island (Figure 4.2B). The beach stretches ~ 1.2 km along the southern shoreline of Greenwich Bay, just east of the mouth of Greenwich Cove. Greenwich Bay is a small (13 km^2) estuary on the western side of Narrangansett Bay, which is ~ 6.5 km southwest of the mouth of the Providence River. The Maskerchugg River discharges to the head of Greenwich Cove, while a number of smaller brooks and creeks flow into Greenwich Bay, either directly or through one of the bay's four other coves. Tributary streams that discharge to the cove and bay transport fecal contamination and direct stormwater discharge. Those vectors are along the west coast of the cove and include the East Greenwich Wastewater Treatment Facility, which discharges treated effluent to the middle of the channel about halfway down the cove, < 2 km from the beach. Faulty septic systems, waterfowl that gather at the beach, wildlife, and domestic pets are other potential sources of contamination at Goddard Beach. Nine percent of samples collected from 2002 to 2009 exceeded water quality criteria standards for enterococci (Appendix A). To our knowledge, formal predictive models have not been previously employed at Goddard Beach. Beach management decisions are based on enterococci monitoring (using the 104 CFU/100 mL criteria) and consideration of water quality history and other environmental conditions such as rain (RIDEM 2005).

4.3.2 Surfside Beach, Surfside Beach, South Carolina

Surfside Beach is in the town of Surfside Beach, South Carolina, just south of Myrtle Beach in Horry County (Figure 4.2B). The beach is 3.4 km long and is only a small fraction of the South Carolina coastline that encompasses more than 100 km of uninterrupted, open beachfront known as the Grand Strand. The Grand Strand is the northernmost part of the South Carolina Coastal Plain. Several impoundments were created in Surfside Beach to serve as retention ponds during storm events. The associated watersheds were designed and constructed for stormwater management purposes. The Myrtle Basin designed in 2005–2006, with construction starting in 2007, was followed by lake dredging. Two small lakes (each with an area of ~ 3000 m^2) are between North Myrtle Drive, North Dogwood Drive, 2nd Avenue North, and 5th Avenue North. They are connected by a 120-m-long channel, with another 200-m-long channel running between them and the beach. The swash at 5th Avenue North receives runoff from this area, which is then discharged directly to the beach. A sign is permanently posted at the swash, stating that swimming is not recommended within 30.5 m because stormwater runoff can result in elevated levels of bacteria. The Public Works Department digs out the swash outlets on the beach as needed (as often as three times per week) to ensure proper water flow.

The Grand Strand Water and Sewer Authority Schwartz WWTP is ~ 8 km northwest of the beach. Surfside Beach is not affected by effluent from the WWTP because the effluent is discharged to the Intracoastal Waterway where the outlet to the Atlantic Ocean is 50 km from the beach. Several campgrounds along the coast north of the beach are within 4 km of Surfside Beach. In addition to the swash at 5th Avenue North, which is at the section of the beach considered here, additional swashes are up and down the coast. Those closest to the beach area include the 11th Avenue North Dogwood Swash (0.6 km north), the Surfside Drive outfall (0.5 km south by the pier), and the 3rd Avenue South Swash (0.9 km south of the 5th Avenue North Swash). Given the lack of known point sources, the most likely source of fecal contamination to Surfside Beach is runoff from the surrounding urban areas. Wildlife can also contribute to observed fecal indicator levels in the swash as birds (i.e., geese, ducks, gulls) frequent the lake and surrounding areas. Water quality at the 5th Avenue North Swash of Surfside Beach has been poor, with 29 percent of samples collected from 2005 to 2007 exceeding water quality criteria standards for enterococci (Appendix A). However, water quality has improved over time (from 70 percent in 2005 to 13 percent in 2007) because of improvements made to the stormwater management watershed. Beach monitoring practices at Surfside Beach rely on culturable enterococci, where levels > 500 CFU/100 mL or repeated measurements > 104 CFU/ 100 mL lead to advisories. In addition, preemptive rainfall advisories were issued on the basis of a rainfall threshold. More recently, a rain model for advisory predictions has been in development. The beach was used for a combined PREMIER-NEEAR study in 2009.

4.3.3 Edgewater Beach, Biloxi, Mississippi

Edgewater Beach is in Biloxi, Mississippi, on the Mississippi Sound along the Gulf of Mexico (Figure 4.2B). The Mississippi Sound is separated from the Gulf by a chain of barrier islands, the closest of which is Ship Island, ~ 20 km south of the beach. This lagoon is < 5-m deep in most areas and runs 124 km along the southern coasts of Mississippi and Alabama. Beaches in the area generally are subject to low energy wave conditions. The beach shore is gently sloped (i.e., 5 to 10 degrees) and consists of well- to very-well-sorted medium sand. Major rain events form runoff channels (Otvos 1999). Seven Harrison County Utility Authority WWTPs are within 10 km of the beach. Water quality at Edgewater Beach (near Eisenhower Drive) is poor, with 16 percent of samples collected from 2004 to 2009 exceeding water quality criteria standards for enterococci (Appendix A). Decisions regarding water quality are based on culturable enterococci where levels > 104 CFU/100 mL result in beach advisories.

4.3.4 Fairhope Beach, Fairhope, Alabama

Fairhope Municipal Beach is in Fairhope, Alabama, ~ 22 km southeast of Mobile, on the eastern shore of Mobile Bay (Figure 4.2B). The Mobile Bay Estuary is an inlet of the Gulf of Mexico, with an average depth of 3 m. The small mouth of the bay is shaped on the east by the Fort Morgan Peninsula and the Dauphin Island barrier island on the west. The sandy beach area at Fairhope Beach is ~ 0.6 km long.

The Fairhope WWTP, < 1 km northeast of the beach, serves as a potential source of continuous fecal contamination. The facility employs UV disinfection as the final treatment step. Additional sources include sanitary sewer overflows, failing septic tanks, and urban runoff (ADEM 2010). Water quality at Fairhope Beach is poor, with 17 percent of samples collected from 2006 to 2009 exceeding water quality criteria standards for enterococci (Appendix A). Decisions regarding water quality are based on culturable enterococci, where repeated measurement of levels > 104 CFU/100 mL result in a public health advisory.

4.3.5 Hobie Beach, Miami, Florida

Hobie Beach in Miami, Florida is on Virginia Key in the southern part of Biscayne Bay, off the east coast of mainland Miami (Figure 4.2B). Biscayne Bay is a subtropical estuary that receives freshwater inputs from the Miami River and small creeks, as well as from a network of drainage canals. The Miami River is ~ 4 km northwest of the beach, and its freshwater can influence the beach under certain conditions. Hobie Beach is ~ 1.6 km long and runs along the south side of Rickenbacker Causeway, between the William Powell Bridge and Miami Seaquarium. Hobie is also known as *Dog Beach* because it is the only Miami-Dade County beach where pets are allowed; the ratio of dogs to humans at the beach is on the order of 1:6. The beach is narrow, with a distance of 5 m and12 m between the water line and the outer edge of the sand line during high and low tide, respectively. Vehicles park right along the sand line. The benthic zone is silty and muddy, and the shoreline typically is covered with seaweed. The slope is relatively shallow, and natural runoff channels form following heavy rainfall events. Hobie Beach is shallow with water depths < 2 m at the buoy line, ~ 130 m from the shoreline. Due to its location in a shallow cove, water circulation at the beach is poor, and movement near shore is controlled by tidal action (with an average tidal height fluctuation of 58 cm) rather than waves (Shibata et al. 2004; Wright 2008). During ebb tide (the period in between high and low tide), water flows out of Biscayne Bay to the Atlantic through the Norris Cut and Bear Cut inlets. During flooding tide, water enters the bay. Flow is parallel to the shoreline with velocities of 0.2 m/s and < 0.1 m/s during ebbing and flooding tide, respectively. Easterly winds prevail with a weak southerly component in the summer, accompanied by a strong local sea-breeze and thunderstorms (Zhu 2009).

During July 1999 to June 2000 at Hobie Beach an EPA/Florida Department of Environmental Health (FDOH) Beach Monitoring Study, indicated that 29 percent of samples exceeded the enterococci water quality criteria (Solo-Gabriele et al. 2002). Since that period, however, only 6 percent of samples collected from 2001 to 2009 exceeded water quality criteria standards for enterococci.

There are no known point sources to the beach. No storm drains are at the beach, but other suspected sources of enterococci include runoff during heavy rainfall events and animals, particularly dogs (Wright et al. 2009). Spatially intensive sampling efforts, including surveys of beach sand and water, identified the shoreline as a source of fecal indicators (Shibata et al. 2004; Solo-Gabriele et al. 2002). A water quality model developed to evaluate sources of nonpoint pollution found runoff to be the most important source of enterococci, followed by dogs, sand, birds, and bathers (Elmir 2006). Research efforts also have focused on the associations between indicator microbes, pathogens, and environmental conditions (Abdelzaher et al. 2010).

Decisions regarding water quality are based on enterococci and fecal coliform levels, with counts > 104 CFU/100 mL resulting in a poor enterococci rating. A process-based predictive numerical hydrodynamic water quality model for Hobie Beach is being developed as a tool to improve assessing nonpoint source fecal contamination (Zhu 2009).

4.3.6 La Monseratte Beach, Luquillo, Puerto Rico

La Monserrate public beach, better known as Luquillo Beach, is on the northeast coast of Puerto Rico in the town of Luquillo (Figure 4.2B). The public beach is ~ 400 m long and runs along the eastern side of the bay. The mouth of the Mameyes River is ~ 2.2 km west of the main swimming area at Luquillo. During periods of low flow, the river mouth can be closed off completely from the ocean. Two streams fed by stormwater runoff are potential sources of fecal contamination to the beach; one is just west of the beach, and the other is east of the beach, along the northern shoreline.

The Puerto Rico Environmental Quality Board (PREQB) monitors enterococci and fecal coliform levels at Luquillo Beach twice a month year-round by. Some data for the current year are publicly available at the PREQB water quality website (http://www.prtc.net/~jcaagua). Historical data indicate that 8 percent of samples collected at Luquillo Beach from 2006 to 2008 exceeded the water quality criteria standard of 104 CFU/100 mL for enterococci. Decisions regarding water quality are based on both culturable enterococci and fecal coliforms, where the level of 35 CFU/100 mL is used as the criteria for enterococci. Puerto Rico's standard is more stringent than that recommended by EPA, and its use results in a total of 18 percent exceedances for 2006 to 2008.

4.3.7 Boquerón Beach, Cabo Rojo, Puerto Rico

Boquerón Beach is in southwestern Puerto Rico, in the town of Boquerón in Cabo Rojo (Figure 4.2B). The 1.6-km-long beach is along the eastern shore of Boquerón Bay, on the western coast of Puerto Rico. The bay is ~ 4.7-km wide at its mouth and the beach is ~ 4 km from the mouth. Potential sources of fecal contamination to the beach include a sewage treatment plant's outfall in the bay ~ 1.3 km northwest of the beach, and two package plants that operate during periods of high demand. The two plants discharge treated effluent into the mangrove lagoon south of the beach. The mouth of the lagoon is ~ 1.5 km from the beach. A marina and condominium complex, just north of the beach, is also a potential source of contamination. Urban runoff is also likely because afternoon storms are common during the summer, with heavy rains resulting in flooding. The beach was used for a combined PREMIER-NEEAR study in 2009.

4.4 METHODS AND DATA ACQUISITION

For more details on methods used for data acquisition, see Appendix B of this report.

4.4.1 Sample Collection

Details regarding the NEEAR study sample collection protocol have been described previously (Haugland et al. 2005; Wade et al. 2006). That general approach was taken at all NEEAR study sites. Briefly, samples were collected on weekends and holidays during the single summer study carried out at each beach. Samples were collected three times a day (8 a.m., 11 a.m., and 3 p.m.) at six locations at each beach. The locations included two depths (shin and waist deep) collected along three transects, ≥ 60 m apart. Shin-depth samples were collected ~ 0.15 m below the surface in 0.3-m-deep water, and waist deep samples were collected ~ 0.3 m below the surface in 1-m-deep water. Grab samples were collected in sterilized polypropylene bottles in accordance with Standard Methods Section 9060 (Clesceri et al. 1998). Samples were mixed to create three composite samples per time point: a shin composite, waist composite, and total composite.

Sample collection for the 2008 PREMIER studies (at South Shore, Hobie, and La Monserrate) was designed deliberately to match the NEEAR studies, where possible. However, more frequent sampling was performed to provide a larger data set for modeling purposes. Waist-deep samples were collected four days a week (Mondays, Wednesdays, Thursdays, and Saturdays), three times a day (9 a.m., 11:30 a.m., and 3 p.m.) at South Shore and Hobie. Shin-deep samples were also collected on Saturdays at South Shore and on Thursdays and Saturdays at Hobie. No shin-deep samples were collected at La Monserrate, and waist-deep samples were collected only once a day (10 a.m.), three days a week (Mondays, Thursdays, and Saturdays) over a longer period. At each sampling location and time, three 500-mL grab samples were collected and mixed to give composite samples for each of the three transects; unlike the NEEAR studies, no beach-wide composites were collected. The distance between sampling transects at the beaches ranged from 125 to 250 m.

The 2009 PREMIER studies at Surfside Beach and Boquerón were specifically designed to complement the concurrent NEEAR studies by expanding the spatial and temporal scale of data collection. In addition to the NEEAR sampling scheme (i.e., weekend and holiday samples) described previously, waist-deep samples were collected on Fridays (three times per day) as part of the PREMIER study. Samples were also collected from two locations outside the beach area (at 8 a.m. and 3 p.m. on Fridays, Saturdays, Sundays, and holidays to better evaluate potential sources of contamination to the beaches (i.e., runoff and effluent at Surfside and Boquerón, respectively). At Surfside Beach, samples were collected from the 3rd Avenue North swash channel along North Ocean Boulevard and from the lake at North Dogwood Drive. Boquerón samples were collected at the outfall of the WWTP and just outside the mouth of the mangrove lagoon. For each sampling event at the sites, three grab samples were collected and composited. The distance between sampling transects at both beaches was ~ 150 m.

In addition to the samples collected for microbial analysis (described above), water samples were also collected during the PREMIER studies to analyze DOC and colored dissolved organic matter (CDOM). The same collection scheme was used for these samples, as detailed above for the 2008 sites, without any sample compositing. At Surfside and Boquerón, beach water samples for the analyses were collected only at the waist-deep site of the middle transect. Single grab samples were also collected from the potential contamination source locations using the microbial sampling scheme. The samples were collected in glass bottles, which were cleaned as described in Standard Methods Section 5310B (Clesceri et al. 1998).

Samples collected from EPA's PREMIER and NEEAR studies provided most of the large data set that was used to refine and evaluate predictive modeling tools, discussed later in this report. In addition, the unique long-term (10-year) data set from Huntington Beach, Ohio (Section 4.2.5) was provided for our use by Donna Francy, USGS. Details regarding the beach field studies that generated data for the modeling studies discussed in this report are summarized in Appendix B.

4.4.2 Dependent Variables

As the dependent variable, FIB density data are the most important component of predictive model development. The current acceptable standard by which recreational water quality is evaluated is culturable enterococci; *E. coli* is also acceptable for freshwater only.

As a result of the BEACH Act of 2000, more than 7 years of monitoring data now exist for many beaches in the United States. Most of the data are available on the Internet, providing an extensive database of historical information. In the PREMIER and NEEAR studies described here, culturable enterococci were enumerated by membrane filtration according to EPA Method 1600 (USEPA 2006). For that method, the detection limit was 1 CFU/volume filtered (i.e., 0.01 CFU/mL for 100 mL volumes) and a value of 0.5 CFU/100 mL (one-half the limit of detection) was used as the lower limit for data analysis. All culturable data were log-transformed before modeling. Given the interest in more rapid monitoring techniques (i.e., qPCR-based enterococci) and the paucity of existing qPCR monitoring data that includes the environmental parameters required for modeling, a goal of the PREMIER studies was to build on the existing NEEAR data set by obtaining qPCR-based enterococci measurements at additional beach sites. Procedures used for the qPCR-based enterococci measurements have been detailed elsewhere (Haugland et al. 2005; USEPA 2010) and are described in Appendix B. We used the same qPCR data for our modeling research that were employed in the NEEAR studies conducted at Boquerón Bay and at Surfside Beach.

Table 4.1. Summary of beach sites (i.e., water type and climate) and field studies that served as a source of data for modeling studies

	Water		Climate			Data source			
Beach	**F**	**M**	**TEMP**	**SUB**	**TROP**	**PREMIER**	**NEEAR**	**USGS**	**YEAR**
South Shore	X	X	X			X			2008
West	X		X				X		2003
Washington Park	X		X				X		2004
Silver	X		X				X		2004
Huntington-EPA	X		X				X		2003
Huntington-USGS								X	2000–2009
Goddard		X		X			X		2007
Surfside		X		X		X	X		2009
Edgewater		X		X			X		2005
Fairhope		X		X			X		2007
Hobie		X		X		X			2008
La Monserrate		X			X	X			2008
Boquerón		X			X	X	X		2009

Notes:
Huntington refers to Huntington Beach, Ohio.
F = freshwater, M = marine, TEMP = temperate, SUB = subtropical, TROP = tropical, PREMIER = EPA modeling studies, NEEAR = EPA National Epidemiological and Environmental Assessment of Recreational Water studies, USGS = USGS, Ohio Water Science Center (provided by D. Francy).

4.4.3 Independent (Explanatory) Variables

Because water quality criteria are based on FIB levels, monitoring data are crucial for beach management. However, models can be useful in predicting indicator densities, which is increasingly important, given issues associated with capability of monitoring data to reflect beach conditions accurately at a given time. More easily measured environmental conditions, often obtained with automated techniques, can be used as IVs in statistical models to explain observed variability in FIB levels. Those include physical hydrologic measurements (e.g., water temperature, turbidity, current, and wave information; tidal phase; and stream discharge); chemical and biological parameters (e.g., pH, dissolved oxygen, conductivity, salinity, and chlorophyll); meteorological conditions (e.g., rainfall, solar radiation, air temperature, and wind information); and ancillary beach conditions (e.g., number of bathers and birds). A variety of such parameters were measured concurrently with FIB measurements during the NEEAR and PREMIER studies. While measurements were taken at the time of sample collection and in discrete samples during the NEEAR studies, one purpose of the PREMIER studies was to obtain more detailed IV data by deploying automated instruments at the beach sites to monitor ambient conditions. Those measurements were also supplemented with variables mined from public databases, where available.

Details about collecting environmental data during the NEEAR studies were previously discussed for the four freshwater sites (Heaney et al. 2009); the same procedures were followed for the subsequent marine studies. Measurements recorded at each sampling event included the following parameters: air and water temperature; percent cloud cover; UV irradiance; wave height; current direction; wind speed and direction; number of bathers on the beach and in the water; total number of birds and animals within 20 m of the sampling area; number of boats within 500 m of the sampling area; and the presence of debris. Rainfall data for the study period were obtained from weather stations at nearby airports and from on-site weather stations installed at some of the beaches. Turbidity, pH, salinity, and conductivity were measured in collected water samples. Limited ancillary data were collected during the 2008 PREMIER studies because efforts were focused on in situ measurements. Local water and beach conditions were noted, including wave conditions (South Shore and Hobie), presence of birds (South Shore) or dogs (Hobie), and number of people and dogs within 50 m of the sampling transects (Hobie).

IV data at South Shore Beach, Milwaukee; Hobie Beach, Miami; Surfside Beach, SC; La Monserrate Beach, Luquillo, Puerto Rico; and Boquerón Beach, Cabo Rojo, Puerto Rico were obtained using procedures and equipment, which are listed with more detail in Appendix B.

Field equipment was deployed at the PREMIER and NEEAR/ PREMIER beach sites to obtain more detailed and beach-relevant IVs from which to develop predictive water quality models. Field equipment locations are indicated in Figure 4.3 for the PREMIER beaches at Milwaukee, Miami, and Luquillo. Figure 4.4 provides locations of the equipment locations for the combined PREMIER/NEEAR studies at Boquerón and Surfside Beach. In addition to using automated field instruments, a unique aspect of the PREMIER studies was measuring underwater UV radiation with the analysis of DOC and CDOM (i.e., UV-visible absorption spectra). By characterizing the optical properties of beach waters, the amount of light to which bacteria are exposed in the water column was more accurately determined in investigating the effects of light on the inactivation of FIB.

Meteorological conditions were monitored by installing HOBO (U30 NRC, Onset Computer Corporation) weather stations at or near each beach site. Weather stations were equipped with sensors to measure air temperature; relative humidity; dew point (determined from temperature and relative humidity); barometric pressure; wind speed and direction; gust speed (i.e., highest 3-second wind recorded during logging interval); rain; photosynthetically active radiation , solar radiation (silicon pyranometer); and UV radiation (Apogee instruments sensors were used in 2009 only) approximately every 15 or 30 minutes. Data were routinely downloaded in the field by connecting a laptop to the weather station data logger. Wind speed and direction were used to determine cross-shore (u-component) and along-shore (v-component) winds at each beach site. The on-site weather data were supplemented with other available meteorological data (http://cdo.ncdc.noaa.gov/ulcd/ULCD).

Current and wave information was obtained by deploying Nortek Aquadopp Profilers (2 MHz, right angle sensor head) at each beach site. The ADCPs were installed on the lake or sea floor using a weighted cross-frame with a mounting height of ~ 0.3 m (Mooring Systems, Inc.). With the exception of the studies at Boquerón Bay, where we deployed two ADCPs, one ADCP was deployed at each beach. A University of Puerto Rico-Mayagüez ADCP was also installed at the Boquerón Bay site.

A B

C

Figure 4.3. Location of 2008 PREMIER studies at (A) South Shore Beach in Milwaukee, Wisconsin; (B) Hobie Beach, Miami, Florida; (C) La Moserrate beach, Luquillo, Puerto Rico Yellow pins indicate locations of the beach and sampling transects; red pins show locations of the field equipment.

A B

Figure 4.4. Location of 2009 PREMIER/NEEAR study at (A) Boquerón Beach, Puerto Rico; (B) Surfside beach, South Carolina. Yellow pins indicate the locations of the beach and sampling transects; red pins show locations of the field equipment.

Multi-parameter water quality sondes (YSI 6600V2-2) were deployed at each beach site for the duration of the studies. The sondes were deployed at fixed locations at a depth of < 2 m. Although actual sonde depths varied at most sites because of tidal changes, sonde readings were taken to represent surface conditions. Water temperature, specific conductance, salinity, dissolved oxygen (using a Clark oxygen electrode), pH, turbidity, and chlorophyll (as relative fluorescence) were measured every 15 minutes. The sondes were typically retrieved every one to two weeks for cleaning, calibration, and data retrieval. Fouling was significant and, in some cases, could be identified as having influenced data quality, specifically for optical measurements (i.e., turbidity and chlorophyll). Fouling of the optical sensors results in high turbidity and chlorophyll readings with a high frequency of spikes. Real (i.e., event-driven) spikes in data progress in a natural upward trend and are short in duration. Examples of bad data from fouling (available from the instrument manufacturer) were used to help identify questionable data during manual review by an experienced individual.

Underwater downwelling solar irradiance (E_d) was measured at each beach using pairs of Satlantic multispectral radiometers (OCR-504 ICSW) with 305, 325, 340, and 380 nanometer (nm) channels. The sensors were placed at two depths to evaluate the attenuation of UV radiation in the water column. The top sensor was < 0.5 m below the water surface (at low tide), and the bottom sensor was placed ~ 0.6 m to 1.5 m below the top sensor, depending on water clarity. To reduce biofouling, each sensor was equipped with a copper Satlantic Bioshutter. In addition to the sensors and Bioshutters, each UV sensor instrument package was equipped with a data logger, battery pack (51Ah), wireless telemetry system with GSM modem, and dual band marine-grade cellular antenna. The equipment was deployed on temporarily installed tower structures. UV irradiance data were processed using the Satlantic SatCon data conversion software. The rate at which irradiance at a given wavelength decreases as a function of depth can be described by a diffuse attenuation coefficient for downwelling irradiance ($K_d(\lambda)$). K_d values were estimated from the irradiance measured at the top and bottom sensors, assuming exponential decrease with increasing depth. The $K_d(\lambda)$ value is the slope of a natural log plot of irradiance versus depth.

DOC was measured in filtered water samples as non-purgeable organic carbon by high temperature combustion using a Shimadzu 5050A or TOC-V Total Organic Carbon Analyzer. UV-visible absorption spectra were determined from 200 to 800 for 0.2-micrometer (µm) filtered water samples, using a Perkin-Elmer LAMBDA™ 35 UV/Vis Spectrophotometer. Measured A_λ were used to calculate absorption coefficients (i.e., $a_\lambda = 2.303(A_\lambda/L)$ where L is the path length in m) at 350 nm.

This page is intentionally blank.

5 Predictive Modeling of Beaches

5.1 FRESHWATER SITES

5.1.1 Data Sources and Methods

We used MLR analyses to examine data sets from five freshwater beaches. At most sites, the response variable was either *Enterococcus* CFU or qPCR values, but we also had *E. coli* CFU data for 2000–2009 from Huntington Beach, Ohio (Section 4.2.5). FIB data are primarily log-normally distributed, so we always $\log_{10}$ transformed the raw FIB measurements (CFU and qPCR) before model development. Site characteristics, the collection of water samples, laboratory bacterial measurement methods, and deployment of on-site instrumentation to collect environmental data are described in Chapter 4, and Appendices A and B of this report.

VB 2.0 was used for developing MLR models for each of the data sets. As described in Chapter 2, VB 2.0 performs data preprocessing to improve the predictive capabilities of MLR models. IVs are tested to see if transformations (logarithm, polynomial, inverse, square root) can induce a more linear relationship to the response variable. Later, as model selection is occurring, VB 2.0 automatically checks for large correlations between IVs (using the VIF), thus avoiding problems associated with multi-collinearity among IVs. When selecting MLR models, many different metrics can be used to choose a best model from a large number of candidates (see Section 5.6, Volume I). For the data sets and analyses, we used the AICC (McQuarrie and Tsai (1998). Model evaluation techniques used here are discussed in Chapter 3.

5.1.2 MLR Model Results

Tables 5.1a (significant IVs) and 5.1b (summary model statistics) show the results of our MLR modeling on the freshwater beach data sets. In all, 23 significant IVs appeared across the various data sets gathered at the five freshwater beach locations. The frequency of occurrence shown in Table 5.1a and in Table 5.2a does not account for the fact that not all the IVs were measured at all beaches. Results corrected for whether the IV was measured are in Section 5.3. Turbidity and antecedent rainfall were most often seen as significant variables. Those were followed by air pressure, the number of bathers, and air temperature. Water temperature, across-shore winds, wave height, dew point, wind speed, and the number of boats were each seen as significant in three of the analyses. Along-shore wind, dissolved oxygen, algae, cloud cover, and chlorophyll were significant for two data sets. Excluding the very large data set from Huntington Beach, Ohio, the adjusted R^2 values for cultured FIB models (CFU in the table) ranged from 0.32 to 0.78, with a mean of 0.55. For these data, RMSE values ranged from 0.24 to 0.48, with a mean of 0.36. For qPCR values, the adjusted R^2 ranged from 0.2 to 0.69, with an average of 0.46. RMSE values for qPCR models ranged from 0.36 to 0.59, with an average of 0.44. The qPCR model produced a higher adjusted R^2 than the CFU model at two of five sites where both qPCR and CFU data were obtained. Details of the regression models chosen for each site are in Appendix C.

5.2 MARINE SITES

5.2.1 Data Sources and Methods

Data sets from seven different marine beach sites were modeled, using MLR analyses. The response variable for the data sets was the logarithm (base 10) of enterococci CFU/qPCR values. Site characteristics, the collection of water samples, laboratory bacterial measurement methods, and deployment of on-site instrumentation for collecting environmental data are described in Chapter 4, and Appendices A and B of this report. VB 2.0 was used to develop MLR models for each of the data sets. For a description of our MLR analytical techniques, see Section 5.1. because the methodology used was very similar to that for freshwater sites.

5.2.2 MLR Model Results

Tables 5.2a (significant IVs) and 5.2b (summary model statistics) show the results for the marine site MLR modeling. Twenty-seven IVs were found to be significant across analyses at the seven sites. Water temperature, humidity, and antecedent rainfall (typically cumulative over the past 48 hours) were most often significant—four analyses for each. Salinity/conductivity, UV intensity, the number of birds seen on the beach, turbidity, and absorbance were next with three appearances each. The qPCR model's adjusted R^2 exceeded that of the CFU model for only one of four data sets where nearly equal numbers of qPCR and CFU observations were taken. Comparisons were not made between qPCR and culturable data at Hobie, Boqueron, and Surfside because of the wide discrepancy in sample sizes between the two data sets (see Chapter 3). If we ignore the sample size discrepancies, the mean adjusted R^2 values for CFU models ranged from 0.19 to 0.48, with a mean of 0.39. For qPCR models, the adjusted R^2 ranged from 0.14 to 0.68, with an average of 0.39. The RMSE values for the qPCR data sets ranged from 0.06 (although that is for the very small Hobie data set) to 0.7, with a mean of 0.41. For the CFU models, the RMSE varied from 0.43 to 0.6, with a mean of 0.52. Details of the regression models for each of the marine sites are included in Appendix C.

Table 5.1a. Results of MLR modeling on freshwater beach sites. A "+" in the cell means that the regression coefficient for that IV was positive. A "-" means the regression coefficient was negative. "P" means the IV was modeled as a second-degree polynomial ($ax^2 + bx + c$). The response variable is the $\log_{10}$ of enterococci levels except for the *E.coli* data set for Huntington Beach, Ohio.

			Independent Variables																						
Beach	Location	Response	Turbidity	Antecedent Rainfall	Barometric Pressure	# of Swimmers	Air Temperature	Water Temperature	Wind V-Component	Wave Height	Dewpoint	Wind Speed	# of Boats	Wind U-Component	Dissolved Oxygen	Algae	Cloud Cover	Chlorophyll	# of Dogs	Nearby River Flow	Debris	# of Birds	Alongshore Currents	Conductivity	Dissolved Organic Carbon
South Shore	Milwaukee, WI	qPCR		P											-			+		P				+	+
		CFU	P	+	P						P			+	-			-					P		
Huntington	Bay Village, OH	qPCR	+	+		+							P				+								
		CFU	+	+		-							+								P				
		CFU (2003 *E. coli*)	+	+			+			+															
		CFU (*E.coli*)	+	+	-				-	+	+														
Washington Park	Michigan City, IN	qPCR	P				P			-															
		CFU	+	-		P			+							-	+		P						
Silver	St. Joseph, MI	qPCR					+	-	-			+		-											
		CFU			P	P		+														+			
West	Porter, IN	qPCR					P	+				P	-												
		CFU	+		P						+	+				-									
Total Appearances per Independent Variable			8	7	4	4	4	3	3	3	3	3	3	2	2	2	2	2	1	1	1	1	1	1	1

Table 5.1b. Summary statistics for the MLR models for freshwater beach sites. The response variable is the $\log_{10}$ of enterococci levels except for the *E.coli* data set for Huntington Beach, Ohio. *Sensitivity* is the number of correctly predicted exceedances over the total actual number of exceedances; *Specificity* gives the number of correctly predicted non-exceedances over the total actual number of non-exceedances, *Accuracy* gives the total number of correctly predicted values over the total count of the data set.

Beach	Location	Response	n	Mean Log FIB Value	St.Dev. of Log FIB Values	RMSE	Adjusted R-Square	Sensitivity	Specificity	Accuracy
South Shore	Milwaukee, WI	qPCR	81	2.05	0.46	0.36	0.39	-	-	-
		CFU	79	1.26	0.5	0.33	0.56	7/13	66/66	73/79
Huntington	Bay Village, OH	qPCR	44	2.25	0.75	0.42	0.69	-	-	-
		CFU	45	1.85	0.79	0.48	0.62	22/23	20/22	42/45
		CFU (2003 *E. coli*)	46	1.69	0.57	0.38	0.54	2/5	39/41	41/46
		CFU (*E.coli*)	709	1.71	0.63	0.47	0.44	35/115	587/594	622/709
Washington Park	Michigan City, IN	qPCR	66	1.47	0.66	0.59	0.2	-	-	-
		CFU	66	1.41	0.32	0.24	0.45	2/5	61/61	63/67
Silver	St. Joseph, MI	qPCR	58	1.45	0.54	0.41	0.41	-	-	-
		CFU	58	1.48	0.44	0.36	0.32	4/14	40/44	44/58
West	Porter, IN	qPCR	49	2.02	0.67	0.41	0.63	-	-	-
		CFU	49	0.87	0.86	0.4	0.78	1/3	44/46	45/49

Table 5.2a. Results of MLR modeling on marine beach sites. A "+" in the cell means that the regression coefficient for that IV was positive. A "-" means the regression coefficient was negative. "P" means the IV was modeled as a second-degree polynomial ($a \cdot x^2 + b \cdot x + c$). The response variable is the $\log_{10}$ of *Enterococcus* CFU or qPCR measures.

			Independent Variables																										
Beach	Location	Response	Water Temperature	Humidity	Antecedent Rainfall	Salinity/Conductivity	UV Intensity/Solar Radiation	# of Birds	Turbidity	Absorbance	Water Depth/Tides	Air Temperature	Dewpoint	Wave Height	# Swimmers	Cloud Cover	Debris	Chlorophyll	Wind Speed	Alongshore Winds	Barometric Pressure	Acrossshore Winds	# Boats	pH	Time of Day	Shin-Deep Water Temp	Alongshore Currents	# of Dogs	Algae
Edgewater	Biloxi, MS	qPCR			P											+													
		CFU					-							+						-								+	+
Fairhope	Mobile, AL	qPCR	P				-					P	P																
		CFU	-	+				-				+	P																
Goddard	West Warwick, RI	qPCR	P														+									P			
		CFU						+			-								P										
Hobie	Miami, FL	qPCR		P			P			+																			
		CFU		+					-									+											
La Monseratte	Puerto Rico	qPCR			+					-																	-		
		CFU				-					+							-											
Boqueron	Puerto Rico	qPCR	+													+	+					+							
		CFU				-			+						+														
Surfside	Myrtle Beach, SC	qPCR		+	+			P	+	P																			
		CFU			+	P								P	P						-		+	+	-				
Total Appearances per Independent Variable			4	4	4	3	3	3	3	3	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1

Table 5.2b. Summary statistics for the MLR models for marine beach sites. The response variable is the $\log_{10}$ of enterococci levels. *Sensitivity* is the number of correctly predicted exceedances over the total actual number of exceedances; *Specificity* gives the number of correctly predicted non-exceedances over the total actual number of non-exceedances, *Accuracy* gives the total number of correctly predicted values over the total count of the data set.

Beach	Location	Response	n	Mean Log FIB	St.Dev. of Log FIB	RMSE	Adjusted R-Square	Sensitivity	Specificity	Accuracy
Edgewater	Biloxi, MS	qPCR	55	2.43	0.48	0.44	0.14	-	-	-
		CFU	55	0.69	0.78	0.57	0.48	1/3	52/52	53/55
Fairhope	Mobile, AL	qPCR	66	2.09	0.61	0.49	0.35	-	-	-
		CFU	66	1.07	0.75	0.57	0.43	0/5	60/61	60/66
Goddard	West Warwick, RI	qPCR	69	1.71	0.77	0.7	0.18	-	-	-
		CFU	69	0.55	0.75	0.6	0.36	0/1	68/68	68/69
Hobie	Miami, FL	qPCR	17	2.08	0.11	0.06	0.67	-	-	-
		CFU	97	0.33	0.52	0.47	0.19	0/1	96/96	96/97
La Monseratte	Puerto Rico	qPCR	24	1.63	0.62	0.45	0.45	-	-	-
		CFU	32	0.5	0.76	0.59	0.41	0/1	31/31	31/32
Boqueron	Puerto Rico	qPCR	79	2.05	0.33	0.29	0.23	-	-	-
		CFU	44	0.99	0.57	0.43	0.43	0/1	43/43	43/44
Surfside	Myrtle Beach, SC	qPCR	40	1.99	0.67	0.43	0.68	-	-	-
		CFU	85	0.72	0.58	0.43	0.46	0/2	83/83	83/85

5.3 Comparison of MLR Modeling Results Across Freshwater and Marine Sites

Comparisons of the MLR modeling results indicate that, on the basis of adjusted R^2 values for predicted versus observed levels of the FIB, model performance was better for the freshwater beaches than for the marine beaches (Tables 5.1b and 5.2b; freshwater average adjusted $R^2 = 0.5$, marine average adjusted $R^2 = 0.39$). Also, modeling results for the culturable FIB data were somewhat better than for the qPCR data (Tables 5.1b and 5.2b; CFU average adjusted $R^2 = 0.46$, qPCR average adjusted $R^2 = 0.42$). The lower values for marine beaches likely reflect the interplay of several factors that have recently been reviewed by Grant and Sanders (2010). Those factors include the effects of currents and tides at the beaches (Grant and Sanders 2010), sunlight-induced inactivation (Boehm et al. 2009) and inputs of FIB from bird and dog droppings, bather shedding, runoff, groundwater and desorption from sand and decaying vegetation (Grant and Sanders 2010; Yamahara et al. 2007). Waves also can be an important factor that reduces model performance; however, all but one of the marine beaches examined in the study were enclosed bays or estuaries that had subdued wave action. Given the results shown in Tables 5.1b and 5.2b, the models were much more accurate in predicting non-exceedances than exceedances at the beaches. One contributing factor, especially in regards to the marine sites we studied, is that there were very few exceedances in the data sets. Accurately predicting phenomena (such as exceedances) that are rarely seen in the training data set is a very difficult task for a statistical model. For the calculations of sensitivity and specificity, we used the same decision thresholds as the EPA regulatory standards (61 CFU for enterococci in freshwaters, 104 for enterococci in marine waters, and 235 for *E.coli* in freshwaters). A beach manager always has the option of lowering the decision threshold (i.e., a model predicted value over which the beach is closed). Depending on how low the decision threshold is set, the result is fewer false negatives (higher sensitivity), but more false positives (lower specificity).

Another general finding was that there was little evidence of a relationship between FIB densities and model performance, although we had expected that there might be. In the case of freshwater beaches modeled in this study, CFU modeling results (adjusted R^2) for the two cleanest beaches, South Shore Beach and West Beach, were among the best. Likewise, satisfactory modeling results were obtained at marine beaches such as Surfside Beach where almost no exceedances were observed. Although comparisons of results using culturable enterococci or *E. coli* were limited, we found that model performance was about the same for both indicators, e.g., at Huntington Beach, Ohio, during 2004.

Table 5.3 shows the results of examining every available IV for all the freshwater and marine sites that were examined. The table indicates variables that were available in each data set, and whether they were significant contributors to the empirical modeling of qPCR or CFU responses. The top half of the table lists the five freshwater sites, and the seven marine sites are shown in the bottom half. The *Rating* in the last row of the table is the ratio of the number of times an IV was found to be significant (filled circles), over the maximum number of times it could have been significant (all circles). Take the IV *algae* as an example. It was present in 12 data sets and found to be significant in three instances: Washington Park CFU, West CFU, and Edgewater CFU; thus, its overall rating is 3/12 = 0.25. The assumption here is that variables with the highest ratings would be the most worthwhile for measuring because there is a higher likelihood they will be significant contributors to an empirical model.

Table 5.4 displays three lists of the IVs sorted by their rating for freshwater sites, their marine site rating, and the difference between their freshwater and marine ratings. Variables at the top of the (freshwater-marine) list are relatively more important at freshwater sites, while variables at the bottom are relatively more important at marine sites. Table 5.5 displays three lists of the IVs sorted by their rating for culturable (CFU) data, for qPCR data, and the difference between those ratings. Variables at the top of the (CFU-qPCR) list are relatively more important for modeling CFU data, while variables at the bottom are relatively more important for modeling qPCR data.

The results in Tables 5.4 and 5.5 are combined for both culturable and qPCR enterococci models. To further explore the modeling differences between the two measurements of beach water quality, we evaluated IV effectiveness in four categories: culturable /freshwater, qPCR/freshwater, culturable/marine, and qPCR/ marine. That analysis showed that there were distinct differences between culture-based and qPCR-based models. For predictions of culturable enterococci at the freshwater beaches, the top IVs were turbidity, number of swimmers, barometric pressure, dew point, algae, and antecedent rainfall. The top IVs for qPCR predictions at freshwater sites were air temperature, wind speed, boats, turbidity, antecedent rainfall, and water temperature. At the marine beaches, the most effective IVs for predicting culturable enterococci were salinity, turbidity, bird counts, algae, swimmers, and wave height. For qPCR levels at marine beaches, the top IVs were the UV absorption coefficient of the water, antecedent rainfall, water temperature, debris on beach, relative humidity, and cloud cover.

The occurrence of turbidity and antecedent rainfall as top IVs for culturable enterococci at both freshwater and marine beaches reinforces the findings of earlier studies (Boehm et al. 2007; USEPA 2007). The number of swimmers was an important IV, perhaps in part because the NEEAR sites were selected to ensure that large numbers of people were present for epidemiological studies at the beaches. Presumably, shedding might have contributed to that finding. The fact that algal levels in the water also were important IVs suggests that remote sensing of chlorophyll could be a useful tool for evaluating beach water quality. Our studies also indicate that UV absorption coefficients of the water and, thus, CDOM can be important IVs in predicting qPCR levels at marine beaches. That finding is probably related to co-variation of this IV with inputs of contaminated waters from nearby streams or wetlands. CDOM also can be quantified by remote sensing. Additional process-based research is required to understand the various factors that underlie the results of these statistical models.

We acknowledge that the accuracy of the ratings calculated for IVs that appear in only a few data sets is suspect, and the robustness of the results can be improved by incorporating information from future empirical modeling efforts.

Table 5.3. Summary of variables used for modeling beach sites.

			Independent Variables																													
Dataset	Location	Response	Turbidity	Chlorophyll	Antecedent Rainfall	Conductivity/Salinity	Swimmers/Bathers	Current Direction	Absorbance	Water Temperature	Nearby River Flow	Dewpoint	Algae	Birds	Directional Wind Components	Air Temp	Barometric Pressure	Wave Height	Boats	Wind Speed	Dissolved Oxygen	Relative Humidity	Cloud Cover	UV Intensity/Solar Radiation	Dogs	Debris	Water Depth/Tides	Dissolved Organic Carbon	pH	Visibility	Jellyfish	Nitrate/Nitrite
South Shore	Milwaukee, WI	QPCR	○	●	●	●		○	○	○	●	○			○	○	○	○		○	●	○		○			○	●	○	○		○
		CFU	●	●	●	○		●	○	○	○	●			●	○	●	○		○	●	○		○			○	○	○	○		○
Huntington	Bay Village, OH	QPCR	●		●		●			○			○	○	○	○		○	●				●		○	○			○			
		CFU	●		●		●			○			○	○	○	○		○	●				○		○	●			○			
		CFU (2003 E. coli)	●		●							○			○	●	○	●				○								○		
		CFU (E. coli)	●		●							●			●	○	●	●				○								○		
Washington Park	Michigan City, IN	QPCR	●		○		○			○			○	○	○	●	○	●	○	○			○	○	○	○			○			
		CFU	●		●		●			○			●	○	●	○	○	○	○	○			●	○	●	○			○			
Silver	St. Joseph, MI	QPCR	○		○		○			●		○		○	●	●	○	○	○	●		○	○			○			○			
		CFU	○		○		●			●		○		●	○	○	●	○	○	○		○	○			○			○			
West	Porter, IN	QPCR	○		○		○			●		○	○	○	○	●	○	○	●	●		○	○		○	○			○	○		
		CFU	●		○		○			○		●	●	○	○	○	●	○	○	●		○	○		○	○			○	○		
Edgewater	Biloxi, MS	QPCR			●		○			○			○	○	○	○		○	○				●	○	○	○					○	
		CFU			○		○			○			●	○	●	○		●	○				○	●	●	○					○	
Fairhope	Mobile, AL	QPCR			○		○			●		●	○	○	○	●	○	○	○	○		○	○	●	○	○	○				○	
		CFU			○		○			●		●	○	●	○	●	○	○	○	○		●	○	○	○	○	○				○	
Goddard	West Warwick, RI	QPCR			○		○			●		○	○	○	○	○	○	○	○	○		○	○	○	○	●	○				○	
		CFU			○		○			○		○	○	●	○	○	○	○	○	●		○	○	○	○	○	●				○	
Hobie	Miami, FL	QPCR	○	○	○	○		○	●	○	○	○			○	○	○			○	○	●	○	●			○	○	○			
		CFU	●	●	○	○		○	○	○	○	○			○	○	○			○	○	●	○	○			○	○	○			
La Monseratte	Puerto Rico	QPCR	○	○	●	○		●	●	○		○			○	○	○			○	○	○		○			○	○	○			
		CFU	○	●	○	●		○	○	○		○			○	○	○			○	○	○		○			●	○	○			
Boqueron	Puerto Rico	QPCR	○	○	○	○	○		○	●		○			●	○	○	○	○	○	○	○	●	○		●	○	○	○			
		CFU	●	○	○	●	●		○	○		○			○	○	○	○	○	○	○	○	○	○		○	○	○	○			
Surfside	Myrtle Beach, SC	QPCR	●		●	○	○		●	○		○		●	○	○	○	○	○	○	○	●	○	○		○	○	○	○			
		CFU	○		●	●	●		○	○		○		○	○	○	●	●	●	○	○	○	○	○		○	○	○	●			
		Rating	0.55	0.50	0.42	0.40	0.33	0.33	0.30	0.29	0.25	0.25	0.25	0.25	0.23	0.23	0.23	0.23	0.22	0.20	0.20	0.20	0.20	0.17	0.17	0.17	0.14	0.10	0.06	0.00	0.00	0.00

Notes: ○ = an IV was in the data set for a specific site but did not appear in the site's chosen MLR model; ●. = the IV did appear in the chosen model; *Rating* = how often each IV, when available, appeared in chosen models (for each column, this is equal to the number of filled circles divided by the total number of circles).

Table 5.4. Importance ratings for IVs at freshwater versus marine sites.

Independent Variable	Freshwater Rating	Independent Variable	Marine Rating	Independent Variables	Fresh - Marine Rating
Chlorophyll	1.00	Turbidity	0.38	Dissolved Oxygen	1.00
Dissolved Oxygen	1.00	Conductivity/Salinity	0.38	Chlorophyll	0.67
Turbidity	0.67	Absorbance	0.38	Dissolved Organic Carbon	0.50
Antecedent Rainfall	0.58	Birds	0.38	Nearby River Flow	0.50
Conductivity/Salinity	0.50	Chlorophyll	0.33	Barometric Pressure	0.32
Swimmers/Bathers	0.50	Relative Humidity	0.33	Swimmers/Bathers	0.30
Current Direction	0.50	Antecedent Rainfall	0.29	Antecedent Rainfall	0.30
Nearby River Flow	0.50	Water Temperature	0.29	Wind Speed	0.29
Dissolved Organic Carbon	0.50	Current Direction	0.25	Turbidity	0.29
Barometric Pressure	0.40	UV Intensity/Solar Radiation	0.21	Boats	0.28
Dewpoint	0.38	Swimmers/Bathers	0.20	Current Direction	0.25
Boats	0.38	Wave Height	0.20	Dewpoint	0.21
Wind Speed	0.38	Debris	0.20	Air Temp	0.19
Algae	0.33	Dewpoint	0.17	Directional Wind Components	0.19
Directional Wind Components	0.33	Algae	0.17	Algae	0.17
Air Temp	0.33	Cloud Cover	0.17	Conductivity/Salinity	0.13
Water Temperature	0.30	Dogs	0.17	Cloud Cover	0.08
Wave Height	0.25	Water Depth/Tides	0.17	Wave Height	0.05
Cloud Cover	0.25	Directional Wind Components	0.14	Water Temperature	0.01
Dogs	0.17	Air Temp	0.14	Dogs	0.00
Birds	0.13	pH	0.13	Debris	-0.08
Debris	0.13	Boats	0.10	pH	-0.13
Absorbance	0.00	Barometric Pressure	0.08	Water Depth/Tides	-0.17
Relative Humidity	0.00	Wind Speed	0.08	UV Intensity/Solar Radiation	-0.21
UV Intensity/Solar Radiation	0.00	Nearby River Flow	0.00	Birds	-0.25
Water Depth/Tides	0.00	Dissolved Oxygen	0.00	Relative Humidity	-0.33
pH	0.00	Dissolved Organic Carbon	0.00	Absorbance	-0.38
Visibility	0.00	Jellyfish	0.00	Jellyfish	-
Nitrate/Nitrite	0.00	Visibility	-	Nitrate/Nitrite	-
Jellyfish	-	Nitrate/Nitrite	-	Visibility	-

Note: In each section of the table, the variables are sorted in descending order by their rating. The last section gives the result of subtracting the marine rating from the freshwater rating. Variables at the top of this list are *relatively* much more important at freshwater sites, while variables at the bottom are *relatively* much more important at marine sites. A variable that is very important (or completely unimportant) at both freshwater and marine sites would appear in the middle of this last list.

Table 5.5. Importance ratings for IVs for culturable (CFU) versus qPCR data.

Independent Variable	CFU Rating	Independent Variable	qPCR Rating	Independent Variable	CFU - qPCR Rating
Chlorophyll	0.75	Absorbance	0.60	Chlorophyll	0.50
Turbidity	0.67	Nearby River Flow	0.50	Algae	0.50
Conductivity/Salinity	0.60	Antecedent Rainfall	0.42	Swimmers/Bathers	0.44
Swimmers/Bathers	0.56	Water Temperature	0.42	Conductivity/Salinity	0.40
Algae	0.50	Turbidity	0.33	Barometric Pressure	0.40
Barometric Pressure	0.40	Current Direction	0.33	Turbidity	0.33
Birds	0.38	Air Temp	0.33	Dogs	0.33
Antecedent Rainfall	0.33	Cloud Cover	0.30	Water Depth/Tides	0.29
Current Direction	0.33	Chlorophyll	0.25	Birds	0.25
Dewpoint	0.33	Boats	0.22	Dewpoint	0.22
Dogs	0.33	Relative Humidity	0.22	pH	0.11
Water Depth/Tides	0.29	UV Intensity/Solar Radiation	0.22	Wave Height	0.10
Directional Wind Components	0.25	Debris	0.22	Directional Wind Components	0.08
Boats	0.22	Conductivity/Salinity	0.20	Current Direction	0.00
Relative Humidity	0.22	Wind Speed	0.20	Boats	0.00
Wave Height	0.20	Dissolved Oxygen	0.20	Wind Speed	0.00
Wind Speed	0.20	Dissolved Organic Carbon	0.20	Dissolved Oxygen	0.00
Dissolved Oxygen	0.20	Directional Wind Components	0.17	Relative Humidity	0.00
Water Temperature	0.17	Birds	0.13	Visibility	0.00
UV Intensity/Solar Radiation	0.11	Swimmers/Bathers	0.11	Jellyfish	0.00
Debris	0.11	Dewpoint	0.11	Nitrate/Nitrite	0.00
pH	0.11	Wave Height	0.10	Antecedent Rainfall	-0.08
Cloud Cover	0.10	Algae	0.00	UV Intensity/Solar Radiation	-0.11
Air Temp	0.08	Barometric Pressure	0.00	Debris	-0.11
Absorbance	0.00	Dogs	0.00	Cloud Cover	-0.20
Nearby River Flow	0.00	Water Depth/Tides	0.00	Dissolved Organic Carbon	-0.20
Dissolved Organic Carbon	0.00	pH	0.00	Water Temperature	-0.25
Visibility	0.00	Visibility	0.00	Air Temp	-0.25
Jellyfish	0.00	Jellyfish	0.00	Nearby River Flow	-0.50
Nitrate/Nitrite	0.00	Nitrate/Nitrite	0.00	Absorbance	-0.60

Note: In each section of the table, the variables are sorted in descending order by their rating. The last section gives the result of subtracting the qPCR rating from the CFU rating. Variables at the top of this list are *relatively* much more important for CFU data, while variables at the bottom are *relatively* much more important for qPCR data. A variable that is very important (or completely unimportant) for both qPCR and CFU data would appear in the middle of this last list.

This page is intentionally blank.

6 Evaluation of Dynamic Modeling and Forecasts of Biological Contamination

The utility of the VB tool was demonstrated initially by a pilot study that evaluated and assessed the feasibility of dynamic models—those that are refit periodically to new data as they become available. Relatively short-duration, dynamic models developed using VB 1.0 (see Chapter 2) were compared in this pilot study, using the adjusted R^2 statistic as a performance measure. A second objective of the modeling study using VB 1.0 was to evaluate 24-hour forecasts of microbial contamination at beaches. The study also tested the usefulness of publicly available data. The results of this section have been presented, in part in a peer-reviewed journal article by Frick et al. (2008).

6.1 Materials and Methods

Section 4.2.5 of this report describes the Huntington Beach, Ohio, site used for the study and our data collection techniques. The data were collected by the USGS Ohio Water Science Center and its partners, the Cuyahoga County Board of Health, and others.

6.2 Results and Discussion

6.2.1 Various Approaches to Developing MLR Models for Nowcasts

MLR models can be fit to data sets of durations ranging from a few days (the actual minimum number is a function of the number of variables included) to a season or longer. Some authors maintain that MLR models should be based on long-term data sets (Francy and Darner 2006; Nevers and Whitman 2005), producing models that are referred to as *static models* in this document. With a static MLR model, the regression coefficients are invariant over the period of its application. For example, USGS used 2000–2005 summer data to fit a unique MLR model that then was used to produce the 2006 beach advisories that are discussed here. In this subsection, such a unique model based on several years of data is called *static*.

The dynamic modeling approach anticipates that models will likely vary as the regression responds to recent trends in the data. Periods as short as 10 days have been used to fit models at marine beaches (Hou et al. 2006). Both the model coefficients and best variables can vary over time because of changing environmental and other conditions, such as treatment plant flow rates, land use changes, storm events, and even climatic fluctuations, which are not well represented by static models (Boehm et al. 2002; Hou et al. 2006). In fact, as the database grows, the best predictive capacity for beaches with changing conditions over time might be obtained using models based on limited subsets of the data record (Hou et al. 2006). A dynamic MLR model can be defined as one in which the IVs are updated periodically with data generated within a limited recent period—usually on the order of 30 to 60 days.

6.2.2 The Performance of Dynamic Nowcast Models of Variable Duration

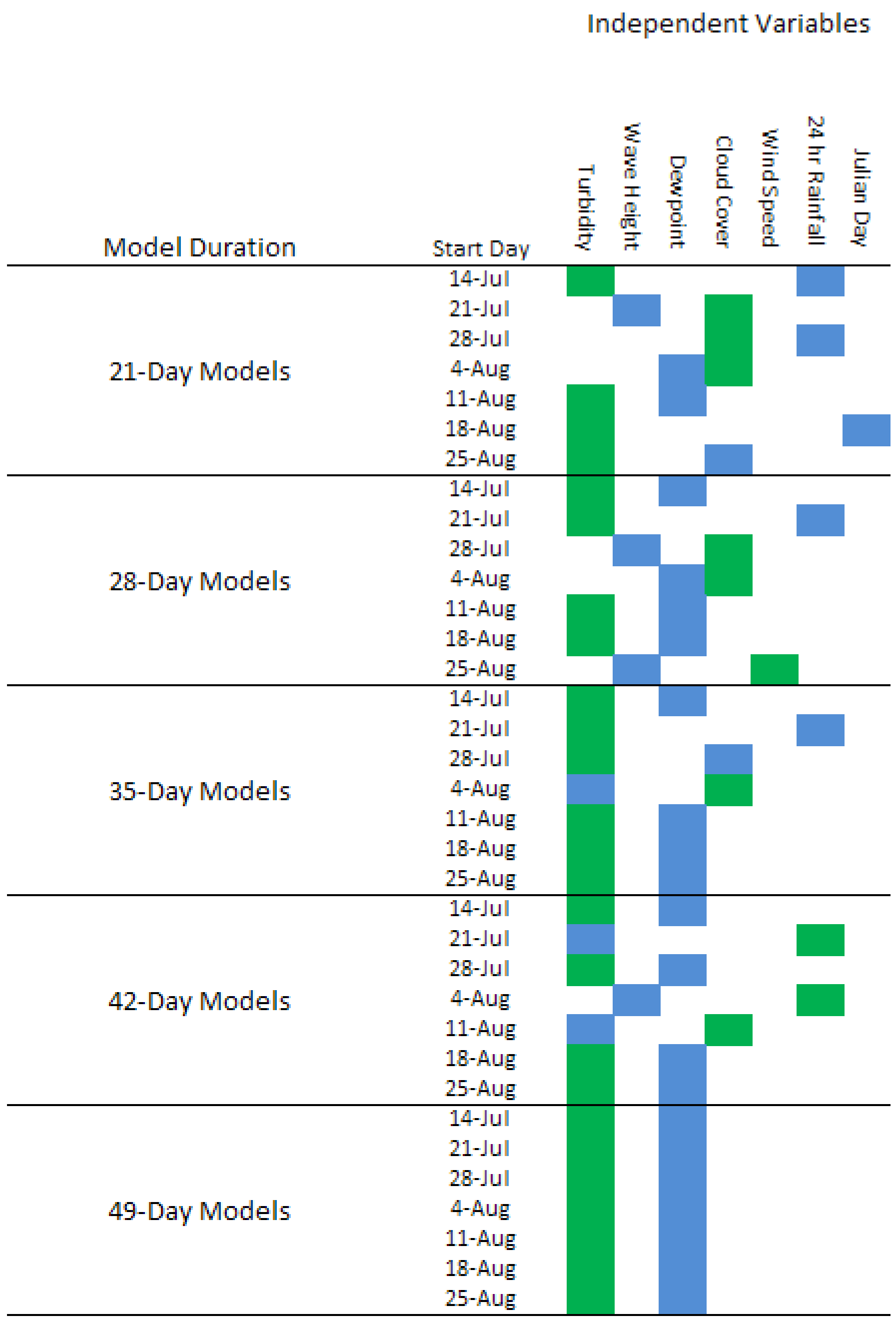
To evaluate the possible use of dynamic nowcast models, VB 1.0 was used to develop a series of models for multi-week subsets of 2006 data for Huntington Beach, Ohio (Figure 6.1). The models were evaluated by comparing predicted and observed *E. coli* densities for data subsets corresponding to 21, 28, 35, 42 and 49 days of data collection. Those results were compared to predictions obtained from the USGS using a model based on the entire 2006 summer season. For the 2006 season, the overall adjusted R^2 for the USGS model was reported to be 42 percent (Francy and Darner 2006). The dynamic modeling results (Table 6.1) for the final 49 days of the 2006 season compared favorably with that result. Like the USGS model, the models built using VB 1.0 identified turbidity as one of the most significant IVs.

Table 6.1. Statistics for the predictions made by models of five temporal durations (seven models were fit in each duration category and each one made seven predictions, thus n = 49 for each category).

Model Duration	Adj. R^2	SE	CE	FN	FP
21-Day Models	50.0	1.22	9	6	0
28-Day Models	45.7	1.08	8	8	1
35-Day Models	61.0	0.99	12	4	2
42-Day Models	53.0	1.00	11	5	2
49-Day Models	60.7	0.89	11	5	2

Note: Given is the adjusted R^2 value for the training data, the standard error of the predictions (SE), the number of correctly predicted exceedances (CE), the number of false negative predictions (FN), and the number of false positive predictions (FP).

Figure 6.1 categorizes models by their duration of fit to the data and best IVs. The IVs used in each case are listed across the top of the figure. The rows of the figure show different categories of models based on five different fitting durations: 21, 28, 35, 42, and 49 days. The *start day* column provides the date of the start of each model's predictive implementation. For example, the 21-day model used to nowcast daily bacteria densities for the 7 days starting July 14 and ending July 20 was fitted to the previous 21 days (June 23 to July 13) of known data. That was followed by the second 21-day model in the category fitted to the 21 days before July 21, and so on. Thus, each of the seven 21-day models contributed seven predictions for a total of 49 predictions (n = 49 for each category). For each duration category, statistics appear in Table 6.1: adjusted R^2 (%) of the training data, the standard error of the predictions, number of correct predictions above the health standard, false negatives, and false positives.



Note: The data were collected by the USGS Ohio Water Science Center, and its partners, the Cuyahoga County Board of Health and others. Data sets were composed of 21-, 28-, 35-, 42-, and 49-day durations. The significance of the variables: green = most significant IV, blue = 2nd most significant IV is based on a model selection process using *Cp* as the selection criterion (figure modified from Frick et al. (2008).

Figure 6.1. Results of models built by VB 1.0 using data at Huntington Beach, Ohio, during 2006.

Figure 6.1 also shows that turbidity, with a few exceptions, was the best IV in 2006. While turbidity was a valuable IV in the 2006 nowcast test, useful models were developed incorporating off-site IVs. Off-site IVs also were found to be valuable predictors for other beaches (Chapter 8). Here at Huntington, Ohio, we found that models incorporating dew point, temperature, and cloud cover variables from the Cleveland airport routinely outperformed those incorporating on-site water temperature and wave height variables. That is valuable knowledge because the off-site data from NOAA weather stations can be obtained with little effort. Moreover, adjusted R^2 values generally increased as the fitting period for the data sets increased from 21 to 35 days but were more or less constant, even decreasing as the duration of fit increased further. The standard errors followed a similar, but inverse, relationship. Overall, models of limited (35 to 42 days) duration appeared to perform best. Also, models developed over the short time frames show great stability; i.e., turbidity, dew point, and cloud cover are consistently among the best variables.

The results of that research suggest that models based on short-term data sets at Huntington Beach, Ohio, perform approximately as well as static models that are based on much longer term data from the beach. However, a detailed analysis of a more extensive data set (2000–2009) from Huntington Beach (Chapter 7) indicates that inclusion of more historic data improved the models for the beach. Thus, as suggested in Chapter 7, random climate-driven fluctuations in conditions might be the strongest determinant of FIB levels at Huntington Beach over an extended period. For other beaches that experience more rapid changes in nearby land use or climate than in northern Ohio, it is possible that models based on longer-term data sets would be less reliable than those based on more recently observed data (Boehm et al. 2002).

6.2.3 Dynamic Forecasting Models

The term *forecast* is defined in this section as making predictions that are typically 24 to 48 hours into the future, using primarily Cleveland airport forecasts as IVs. More details about this procedure are included in Frick et al. (2008). The experimental forecast period covered 42 days in the second half of the 2006 beach season. From nowcasting results that showed 5-week models performed best, no attempt was made to fit forecast models with data sets of other durations. For the 42-day forecast period (July 21 to August 31), the adjusted R^2 was 42.3 percent, similar to nowcast results that exclude turbidity. Eight exceedances were correctly predicted, and there were six false negatives and no false positives. The 24-hour forecast performance compared favorably to the corresponding nowcast performance (omitting turbidity), despite the reduction in available IVs.

The results of the research were an initial indication that the VB tool could facilitate and optimize the development of statistical models used for nowcasting and forecasting FIB densities at recreational water sites, and it could be valuable for optimizing and updating dynamic models based on short-term data sets.

7 Evaluating the Predictive Capabilities of Models for *E. coli* Levels at Huntington Beach, Ohio, Using Varying Amounts of Historical Data

7.1 INTRODUCTION

The previous section, which discusses results presented in Frick et al. (2008), raises interesting questions about data used for MLR model development. They discuss the concept of dynamic modeling, in which the temporal characteristics of the data used to generate FIB models come under scrutiny. Using USGS data for 2006 at Huntington Beach, Ohio, they investigated whether short-term data sets, such as those spanning from 1 to 7 weeks before the desired prediction date, could produce models that outperform models using longer-term data sets, i.e., entire previous years of data. They found that short-term data sets could, indeed, produce useful statistical models for the site and that it was not necessary to accumulate multiple years of modeling data before statistical relationships could be fruitfully explored.

Our objective was to expand on the analyses of Chapter 6. Given a robust 10-year (2000–2009) USGS data set documenting environmental conditions and *E. coli* counts at Huntington Beach, Ohio, we set out to record the predictive capabilities of models developed using successively longer periods of data. We did not investigate short-term intra-annual data sets as in Frick et al. (2008), instead focusing on modeling results for inter-annual data sets ranging from 1 to 9 years of observations. Our initial hypothesis was that more data would mean better models, i.e., predictions would improve as more historical data were used for model development. However, if factors that control FIB levels at a site change through time continuously—possibly from land use change in the surrounding watershed or climate change—model performance could degrade according to the period of data used. If periodic or cyclical environmental conditions affect the site, there might be an optimum historic period of record to use for making predictions. When forces driving FIB contamination at a site change annually in an unpredictable manner, models developed using a prior year's data could be good or poor, depending on the similarity between the historical data period and the current period for which predictions are needed.

We formalized these issues with the following research questions:

1. Will using longer historical periods of data lead to better predictive models for *E. coli* at Huntington Beach?
2. If not, is there another period of record that optimizes model predictive performance?
3. Does year-to-year variability in model predictive ability show any discernible patterns?
4. Do the answers to 1, 2, and 3 provide insight about FIB dynamics at Huntington Beach?

7.2 METHODS

Huntington Beach, including sampling protocols, is described in detail in Section 4.2.5, Appendix A, and in Francy et al. (2003), Francy and Darner (2006), and Haugland et al. (2005). *E. coli* levels were measured during the 2000–2009 swimming seasons (late May through late August), along with a suite of potential IVs (water temperature, turbidity, wave height, wind speed and direction, solar radiation, and the like). Wind speed and direction were converted into along-shore (u) and cross-shore (v) components, as discussed in Frick et al. (2008). Relative to each *E. coli* measurement, raw rainfall data were aggregated into 48-hour antecedent cumulative precipitation. Log_{10} values of *E. coli* were used as the response variable.

To address our first two questions, we tested the predictions of MLR models, using *E. coli* levels as the response, based on successively longer historic data sets (Table 7.1).

Table 7.1. The matrix for recording MSE of models developed using a variable number of previous years' data (M_{1prev}–M_{9prev}) applied to each single year of data (Y_{2001}–Y_{2009}).

	M_{1prev}	M_{2prev}	M_{3prev}	M_{4prev}	M_{5prev}	M_{6prev}	M_{7prev}	M_{8prev}	M_{9prev}
Y_{2001}	$MSE_{1,1}$	-	-	-	-	-	-	-	-
Y_{2002}	$MSE_{2,1}$	$MSE_{2,2}$	-	-	-	-	-	-	-
Y_{2003}	$MSE_{3,1}$	$MSE_{3,2}$	$MSE_{3,3}$	-	-	-	-	-	-
Y_{2004}	$MSE_{4,1}$	$MSE_{4,2}$	$MSE_{4,3}$	$MSE_{4,4}$	-	-	-	-	-
Y_{2005}	$MSE_{5,1}$	$MSE_{5,2}$	$MSE_{5,3}$	$MSE_{5,4}$	$MSE_{5,5}$	-	-	-	-
Y_{2006}	$MSE_{6,1}$	$MSE_{6,2}$	$MSE_{6,3}$	$MSE_{6,4}$	$MSE_{6,5}$	$MSE_{6,6}$	-	-	-
Y_{2007}	$MSE_{7,1}$	$MSE_{7,2}$	$MSE_{7,3}$	$MSE_{7,4}$	$MSE_{7,5}$	$MSE_{7,6}$	$MSE_{7,7}$	-	-
Y_{2008}	$MSE_{8,1}$	$MSE_{8,2}$	$MSE_{8,3}$	$MSE_{8,4}$	$MSE_{8,5}$	$MSE_{8,6}$	$MSE_{8,7}$	$MSE_{8,8}$	-
Y_{2009}	$MSE_{9,1}$	$MSE_{9,2}$	$MSE_{9,3}$	$MSE_{9,4}$	$MSE_{9,5}$	$MSE_{9,6}$	$MSE_{9,7}$	$MSE_{9,8}$	$MSE_{9,9}$

For example, we developed a model (model selection always done using backwards stepwise procedure with AIC as the criterion) with data from 2004, and then used this model to predict the *E. coli* values in 2005. We termed the model M_{1prev} because it was based on one previous years' worth of data. We would then record the MSE of those predictions in row "Y_{2005}," column "M_{1prev}" in Table 7.1. Next, we would construct a model with data from 2003 and 2004, use that model to predict the *E. coli* values in 2005, and record the MSE of the predictions in row "Y_{2005}," column "M_{2prev}." That methodology was used to find all the MSE values in Table 7.1. Y_{2000} values do not appear in the table because it was the first year of our data record.

To address our third research question, we created 10 data sets of IVs (X) and corresponding *E. coli* measurements (Y). Those are labeled as (X_{2000}, Y_{2000}), (X_{2001}, Y_{2001}), …, (X_{2009},Y_{2009}). Next, we developed MLR models (again using backwards stepwise procedures with AIC as the model criterion) using each of these data sets, and denoted those models as M_{2000}, M_{2001}, …, M_{2009}. Next, we fit each of the models to every other year of data during the period of record and noted the MSE for that application (Table 7.2).

Table 7.2. The matrix for recording MSEs of models developed using a single year of data, then applying that model to all other years of observations.

	M_{2000}	M_{2001}	M_{2002}	M_{2003}	M_{2004}	M_{2005}	M_{2006}	M_{2007}	M_{2008}	M_{2009}
(X_{2000},Y_{2000})	$MSE_{1,1}$	$MSE_{1,2}$	$MSE_{1,3}$	$MSE_{1,4}$	$MSE_{1,5}$	$MSE_{1,6}$	$MSE_{1,7}$	$MSE_{1,8}$	$MSE_{1,9}$	$MSE_{1,10}$
(X_{2001},Y_{2001})	$MSE_{2,1}$	$MSE_{2,2}$	$MSE_{2,3}$	$MSE_{2,4}$	$MSE_{2,5}$	$MSE_{2,6}$	$MSE_{2,7}$	$MSE_{2,8}$	$MSE_{2,9}$	$MSE_{2,10}$
(X_{2002},Y_{2002})	$MSE_{3,1}$	$MSE_{3,2}$	$MSE_{3,3}$	$MSE_{3,4}$	$MSE_{3,5}$	$MSE_{3,6}$	$MSE_{3,7}$	$MSE_{3,8}$	$MSE_{3,9}$	$MSE_{3,10}$
(X_{2003},Y_{2003})	$MSE_{4,1}$	$MSE_{4,2}$	$MSE_{4,3}$	$MSE_{4,4}$	$MSE_{4,5}$	$MSE_{4,6}$	$MSE_{4,7}$	$MSE_{4,8}$	$MSE_{4,9}$	$MSE_{4,10}$
(X_{2004},Y_{2004})	$MSE_{5,1}$	$MSE_{5,2}$	$MSE_{5,3}$	$MSE_{5,4}$	$MSE_{5,5}$	$MSE_{5,6}$	$MSE_{5,7}$	$MSE_{5,8}$	$MSE_{5,9}$	$MSE_{5,10}$
(X_{2005},Y_{2005})	$MSE_{6,1}$	$MSE_{6,2}$	$MSE_{6,3}$	$MSE_{6,4}$	$MSE_{6,5}$	$MSE_{6,6}$	$MSE_{6,7}$	$MSE_{6,8}$	$MSE_{6,9}$	$MSE_{6,10}$
(X_{2006},Y_{2006})	$MSE_{7,1}$	$MSE_{7,2}$	$MSE_{7,3}$	$MSE_{7,4}$	$MSE_{7,5}$	$MSE_{7,6}$	$MSE_{7,7}$	$MSE_{7,8}$	$MSE_{7,9}$	$MSE_{7,10}$
(X_{2007},Y_{2007})	$MSE_{8,1}$	$MSE_{8,2}$	$MSE_{8,3}$	$MSE_{8,4}$	$MSE_{8,5}$	$MSE_{8,6}$	$MSE_{8,7}$	$MSE_{8,8}$	$MSE_{8,9}$	$MSE_{8,10}$
(X_{2008},Y_{2008})	$MSE_{9,1}$	$MSE_{9,2}$	$MSE_{9,3}$	$MSE_{9,4}$	$MSE_{9,5}$	$MSE_{9,6}$	$MSE_{9,7}$	$MSE_{9,8}$	$MSE_{9,9}$	$MSE_{9,10}$
(X_{2009},Y_{2009})	$MSE_{10,1}$	$MSE_{10,2}$	$MSE_{10,3}$	$MSE_{10,4}$	$MSE_{10,5}$	$MSE_{10,6}$	$MSE_{10,7}$	$MSE_{10,8}$	$MSE_{10,9}$	$MSE_{10,10}$

For example, using the model developed from X_{2002},Y_{2002} data (M_{2002}), plugging in the X_{2004} IVs and taking the MSE of the fit of this model compared to the actual Y_{2004} observations, would produce the $MSE_{5,3}$ value (the highlighted cell in Table 7.2).

7.3 Results and Discussion

The MSEs given in Table 7.3 support the notion that using as much historic data as possible leads to better overall predictions, i.e., the mean column values decline from left to right. Within any single row, the pattern is not always maintained, but on average, it is clearly seen. Such results are consistent with some previous studies (Francy et al. 2006b; Nevers and Whitman 2005).

Table 7.3. Results of modeling the MSEs of Table 7.1.

	M_{1prev}	M_{2prev}	M_{3prev}	M_{4prev}	M_{5prev}	M_{6prev}	M_{7prev}	M_{8prev}	M_{9prev}
Y_{2001}	2.39	-	-	-	-	-	-	-	-
Y_{2002}	2.99	2.12	-	-	-	-	-	-	-
Y_{2003}	3.21	1.91	1.53	-	-	-	-	-	-
Y_{2004}	1.64	1.77	1.44	1.60	-	-	-	-	-
Y_{2005}	2.84	1.88	2.04	2.02	1.92	-	-	-	-
Y_{2006}	1.92	1.92	1.70	1.82	1.67	1.78	-	-	-
Y_{2007}	1.23	1.17	1.24	1.28	1.19	1.26	1.20	-	-
Y_{2008}	1.62	1.20	1.16	1.02	1.15	1.22	1.13	1.09	-
Y_{2009}	0.77	0.67	0.67	0.66	0.67	0.64	0.64	0.66	0.62
Mean	2.07	1.58	1.40	1.40	1.32	1.22	0.99	0.87	0.62

Note: The mean MSE values in the last row indicate that using more historic data generally leads to better predictions.

FIB levels at beaches respond to a suite of environmental factors such as storm events, local land use, nearby WWTPs, beach attendance, and such. If the set of factors important for determining FIB levels at a site changes through time, it would seem that predictions made using short-term data sets could, indeed, be more responsive to these changes. However, when using a long-term data set, the model is basically *averaged* over all the data, and representative of many different conditions experienced at the site. In other words, predictions from this model will be decent over a large range of possible environmental scenarios. With models derived from short-term data sets, the likelihood of boom and bust predictions arises. When conditions at the site have not changed recently, using the past 5 weeks of data to make predictions should work well. In the days following an important change in conditions, however, predictions based on the past several weeks of data could be terrible. A third scenario, in which conditions at the site are slowly changing along some gradient (global climate change leading to increasing temperatures or precipitation, or impervious surfaces in the local watershed continually increasing) could favor short-term data sets. In such a case, abrupt changes to the system are not present, and conditions in the future will be unlike those seen in the past, so using long, historic data sets could be counterproductive.

The MSE values in Table 7.4 support the notion that random climate fluctuation determined FIB levels at Huntington Beach over the period.

Table 7.4. Results of modeling the MSEs of Table 7.2.

	M_{2000}	M_{2001}	M_{2002}	M_{2003}	M_{2004}	M_{2005}	M_{2006}	M_{2007}	M_{2008}	M_{2009}	Mean
(X_{2000},Y_{2000})	0.56	2.03	1.97	1.42	2.43	1.79	1.56	1.39	1.07	0.99	1.52
(X_{2001},Y_{2001})	2.39	1.26	2.42	1.94	2.64	1.99	1.98	2.62	1.62	1.68	2.05
(X_{2002},Y_{2002})	3.49	3.15	1.27	2.44	2.27	2.36	2.70	2.19	2.77	2.06	2.47
(X_{2003},Y_{2003})	2.71	1.74	1.76	0.61	1.46	0.83	1.39	1.23	1.67	0.95	1.44
(X_{2004},Y_{2004})	2.14	1.82	2.25	1.62	1.26	1.96	2.49	1.64	2.67	1.87	1.97
(X_{2005},Y_{2005})	6.79	3.05	2.56	1.97	2.84	1.69	2.22	2.93	2.68	2.34	2.91
(X_{2006},Y_{2006})	3.22	2.74	2.11	1.90	2.15	1.92	1.21	1.92	1.77	1.52	2.05
(X_{2007},Y_{2007})	4.24	2.19	1.93	1.81	1.54	1.86	1.23	0.77	1.20	1.14	1.79
(X_{2008},Y_{2008})	4.84	1.20	3.08	1.50	1.32	1.88	1.13	1.62	0.72	0.93	1.82
(X_{2009},Y_{2009})	2.72	0.94	2.24	1.12	0.74	1.09	1.31	1.08	0.77	0.52	1.25
Mean	3.31	2.01	2.16	1.63	1.86	1.74	1.72	1.74	1.70	1.40	

Note: The lack of an easily discernible pattern in these data is consistent with a hypothesis that random annual climatic variation governs FIB dynamics at Huntington Beach, Ohio.

Those data show no easily discernible pattern (although no statistical test was performed), other than values along the diagonal are the lowest for each row, as would be expected—i.e., the model derived from the data of a given year has to fit that same year's observations better than any model developed from another year's data. The MLR regression coefficients are determined by minimizing the sum of the squared error term. Any other set of coefficients mathematically could not produce a lower MSE. That does not mean, however, that the value along the diagonal has to be the lowest value within any column. For example, the 2004 model fits the 2009 data even better than it fits its own 2004 observations. But the 2009 model still fits the 2009 data best of all, and the 2004 model fits the 2004 data best (lowest values within each row).

If a gradient effect reflecting gradually changing conditions existed, the MSE values in Table 7.4 would rise moving left or right, away from the diagonals within each row, but that pattern is not seen. The table makes it clear that certain pairs of years are more alike than other pairs, leading to better predictions. For example, the model based on 2006 data fits 2008 observations very well, but it does not fit 2002 observations well. Overall, the 10 models had the hardest time fitting 2005 observations, and the easiest time fitting 2009 observations, based on the MSE row means given in the far right column. The model from 2000 did the poorest job of fitting the 10 data sets, because of its largest mean value in the last row of the table.

7.4 CONCLUSIONS

Using MLR modeling on a series of annual data sets (2000–2009) for Huntington Beach, Ohio, better predictions of *E. coli* were made as more historic data were used in developing predictive models. Also, random annual climatic fluctuations appeared to be driving FIB levels at this site because no easily discernible pattern or gradient was found when examining temporal trends within model predictions on a year-to-year basis. In other words, a model developed using data from year T did not make better predictions for years T-1 and T+1, as compared to predictions made for years more temporally removed from T. The temporal positions of good and bad predictions between individual years showed no apparent pattern.

We must note that these results, having been derived from a study at a single site, are not nearly robust enough to draw general conclusions regarding other beaches, but they do provide clear evidence that in at least this case, more historic data leads to better predictions. FIB dynamics at other beaches, driven by different sources, hydrodynamics, and climatic conditions, might react quite differently over time and short-term models might be able to provide more effective predictions in some cases.

8 The Importance of Site-Specific Environmental Data for Modeling Enterococci Densities at South Shore Beach, Wisconsin

8.1 INTRODUCTION

This study focused on comparing the fit and predictive ability of MLR models developed using two sets of IVs for a site—one data set containing only publicly available environmental data collected near the beach, and a second data set that contains the first, but then adds additional environmental data collected at the beach itself. In both cases, on-site microbial data were used as the dependent variables in model development. We did not include in this comparison a data set consisting of data collected on-site only, because publicly available data could always be added at little cost, so the condition itself is not worth examining when costs and benefits are at issue. Such a type of analysis can allow a beach manager to make an objective determination of whether the resources used for equipment installation and monitoring are worthwhile, as measured by improvements in model predictive abilities.

8.2 MATERIALS AND METHODS

8.2.1 Site Details

South Shore Beach's location is shown in Figure 4.2, and the beach is described in detail in Section 4.2.1. Given the history of biological contamination at South Shore, plus the public availability of nearby meteorological and other IVs for modeling microbial levels, we decided to conduct a detailed modeling study at this beach during the summer of 2008.

In keeping with the objectives of our study, all IVs were categorized into two groups (Table 8.1): publicly available (PA) and SS. *SS* refers to all additional data collected by EPA or the University of Wisconsin-Milwaukee's Great Lakes WATER Institute despite the fact that the EPA's weather station was a mile from the beach. For our response variable (enterococci CFU/100 mL), we then developed and compared MLR models that used only the PA variables versus a model that used both PA and SS variables.

8.2.2 Data Management

An important step in MLR analysis is to check for high correlations among the IVs to avoid inflated standard errors and instability when estimating regression coefficients. We checked for IV collinearity, using Pearson correlation coefficients. Given a pair of highly correlated IVs (correlation coefficient above 0.8, (Cohen 1988), collinearity is remedied by dropping one of the two from the analysis. The primary criterion for this choice was to choose the IV with fewer missing values in the data set. When PA and SS variables were found to be highly correlated, and if the missing-value criterion was equivalent, PA variables were retained over SS in

Table 8.1. Environmental variables used in model development

Data Source	Type	Variable Description	Data Collection Duration	Data Collection Frequency
EPA Sonde	SS	Absorption Coefficient (325 nm) Dissolved Organic Carbon	7/2 - 9/22	9:00, 11:30, 15:00
		Water Temperature Conductivity Salinity Dissolved Oxygen pH Turbidity Chlorophyll Ammonium Ammonia Nitrate	7/17 - 10/18	Every 15 minutes
		Current Speed Current Direction Water Depth	7/17 - 10/21	Every 10 minutes
		Wave Height Wave Direction	7/17 – 10/21	Every 60 minutes
EPA UV Sensor	SS	Attenuation Coefficient (325 nm) Mean Attenuation at 0.3 meters (325 nm)	7/17 - 10/22	Every 60 minutes
EPA Weather Station	SS	Solar Radiation Photosynthetic Active Radiation Wind Speed Wind Direction Gust Speed Cumulative 48hr Rainfall Air Temperature Humidity Dewpoint Pressure	7/21 - 10/18	Every 15 minutes
UW Sonde	SS	Water Temperature Conductivity pH Dissolved Oxygen Chloride Turbidity	7/17 - 10/19	Every 30 minutes
UW Weather Station	SS	Air Temperature Cumulative 48hr Rainfall Wind Speed Wind Direction	7/17 - 10/19	Every 60 minutes
General Mitchell Airport	PA	Visibility Dry Bulb Temperature Wet Bulb Temperature Dewpoint Humidity Wind Speed Wind Direction Barometric Pressure Pressure at Sea Level Cumulative 48hr Rainfall Altimeter Pressure	7/1 - 9/30	Every 60 minutes
USGS	PA	Flow at Mouth of Milwaukee River	6/25 - 8/28	Once per day

developing models based on both data sets. If correlated IVs were either both PA or both SS—and no missing-value advantage was seen—ease of interpretation and measurement was used to decide which IV to retain. After that processing, 9 IVs remained for the PA data set, and 21 IVs for the PA+SS data set.

As in previous modeling studies (Frick et al. 2008; Hou et al. 2006; Francy and Darner 2006; Olyphant and Whitman 2004), the logarithm (base 10) of the *Enterococcus* count was taken to help ensure a more linear relationship to the IVs. Measurements describing wind speed and direction, and water current speed and direction, were transformed into an along-shore component (u) and perpendicular-to-shore component (v) (Nevers et al. 2007).

8.2.3 Model Development

After removing missing values from our data set, 79 measurements of enterococci CFUs remained for model development. Many metrics allow an analyst to judge the relative merit of how well statistical models fit data, such as the AIC, AICC, BIC, adjusted R^2, PRESS and Mallows' Cp (Cp). While the adjusted R^2, AIC, AICc, BIC and Cp are different measures of the fit of a model to a set of training data, with a penalty imposed for over-fitting, PRESS uses one-observation-at-a-time cross-validation to optimize the out-of-sample prediction performance. PRESS sums the prediction residuals of a model, so lower PRESS values mean better models. Because of the public health ramifications of not making accurate predictions of fecal indicator levels at beaches, we chose the PRESS statistic to define the best predictive model.

Model selection was accomplished using VB 2.0 (see Chapter 2). Because we had 79 observations, and knowing that 10 observations per IV provide adequate power to estimate regression coefficients with precision (Aguinis and Harden 2009), we evaluated only those models with seven IVs or fewer. While calculating every possible 1- to 7-parameter regression model was possible for the PA data set (a total of 9 IVs), the PA+SS data set had too many IVs (21) for this to be done, so we used VB 2.0's GA instead.

8.2.4 Model Validation/Evaluation

After the two *best* models (one using the PA data set and one using the PA+SS data set) were identified, they were compared. First, the ability of these models to fit the observed *Enterococcus* counts was evaluated using the model's adjusted R^2 value (Frick et al. 2008; Nevers et al. 2007; Francy and Darner 2006). Second, the model's PRESS values were compared. Finally, we examined the *sensitivity* and *specificity* of each model (Francy and Darner 2006); see definitions in Chapter 3) using the 61 CFU/100 mL EPA national advisory level for *Enterococcus* in freshwater as the critical threshold.

Generally, the probability of exceeding a threshold or standard, S, can be calculated as:

$$P_{exceed} = Prob\ (T \geq [\hat{y} - \log(S)] \ /\ s_{\hat{y}})$$

where T has the Student's t distribution with $n–p$ degrees of freedom (n is the number of observations in the data set and p is the number of estimated regression parameters), $\hat{y}$ is the predicted value of the logarithm of the FIB count and $s_{\hat{y}}$ is the standard error of this prediction. We could alternatively define a probability threshold (e.g. 30 percent), such that any $\hat{y}$ whose corresponding P_{exceed} surpasses the threshold (meaning $\hat{y}$ has a greater than 30 percent chance of exceeding log[*S*]) is considered an unacceptable risk to public health and would lead to beach closure. We specified three *Enterococcus* count thresholds (30, 61 and 100 CFU/100 mL) and three probability thresholds (30, 50, and 70 percent) to calculate the specificity and sensitivity of each model.

8.3 RESULTS

8.3.1 PA Analysis

The primary objective of this report is to compare the ability of the PA and PA+SS data sets to fit observed bacteria counts, and interpretation of significant IVs is beyond the scope of this report. Analysis of the PA data set, which included measurements from the General Mitchell Airport and the USGS gage on the Milwaukee River, produced the following model:

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-52.4	9.96	-5.26	< 0.001
Dewpoint	0.020	0.01	2.40	0.019
Wind U-Component	-0.018	0.01	-2.91	0.005
Barometric Pressure	1.79	0.33	5.38	< 0.001
Cumulative 48-hr Rainfall	104.7	14.02	7.47	< 0.001

The adjusted R^2 of this model is 0.47, and the PRESS statistic is 11.02.

8.3.2 Combined PA + SS Analysis

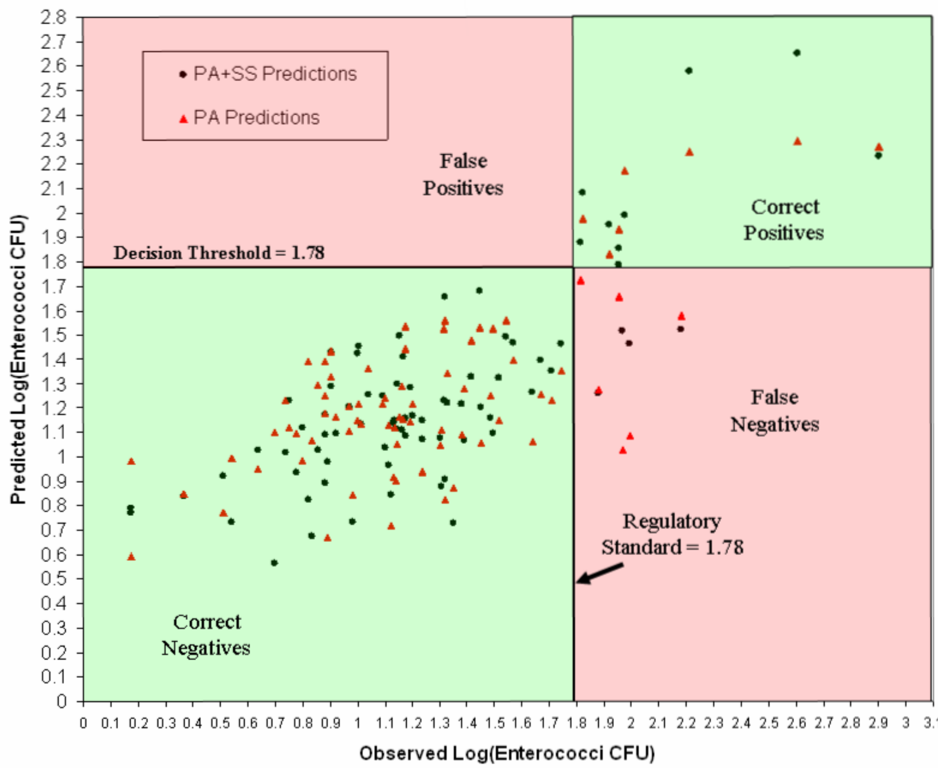
The analysis used the publicly available General Mitchell Airport and USGS data as well as the data collected by EPA and the UWM-GLWI at South Shore Beach. The VB 2.0 GA selected this model:

Parameter	Data Source	Coefficient	Std. Error	t-Statistic	P-Value
Intercept		-12.12	2.73	-4.44	< 0.001
Cumulative 48-hr Rainfall	PA - Airport	94.38	12.67	7.45	< 0.001
pH	SS - EPA Sonde	0.65	0.24	2.70	0.0087
Dissolved Oxygen	SS - UW Sonde	-0.21	0.03	-6.25	< 0.001
Nitrate	SS - EPA Sonde	0.20	0.06	3.43	0.001
Mean Attenuation Coefficient (0.3m)	SS - EPA Sensor	-0.04	0.01	-3.43	0.001
Wind U-Component	PA - Airport	-0.02	0.01	-3.00	0.0037
Water Depth	SS - EPA Sonde	3.74	0.64	5.83	< 0.001

The adjusted R^2 obtained was 0.59, and the PRESS statistic was 9.1 for the model. The PA model uses four IVs, while the PA+SS model has seven—five SS and two PA variables. Using the PA+SS data set increased the adjusted R^2 by 25 percent and decreased the PRESS statistic by 17 percent over the model that used only PA variables. That indicates that on-site environmental data can help predict FIB at South Shore Beach but that publicly available meteorological data alone can explain a significant proportion of the variability in FIB levels.

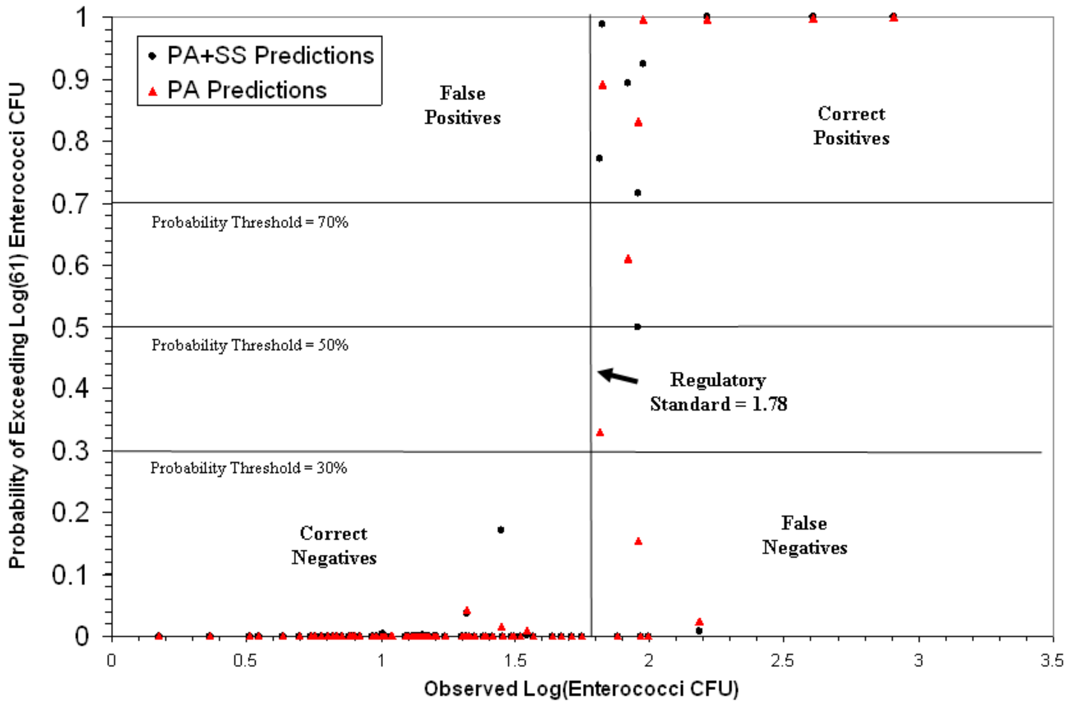
8.3.3 Model Comparisons

Plots, using a count threshold (Figure 8.1) and a probability threshold (Figure 8.2), of model predictions versus actual observations demonstrate a graphical approach to model evaluation.



Note: The colored quadrants are formed by the vertical line (EPA regulatory standard) and the horizontal line (beach manager's decision criterion). The decision criterion can be shifted up and down to meet an objective of minimizing false positives or false negatives.

Figure 8.1. Plotting model predictions versus observations for the PA and PA+SS data sets.



Note: As in Figure 8.1, the vertical line is the regulatory standard. The three horizontal lines represent three choices of probability threshold (30 percent, 50 percent, and 70 percent).

Figure 8.2. Plotting the probability that a predicted bacteria count will exceed a threshold value (log[61] CFU/100 mL) versus actual observations.

Horizontal thresholds, set by the analyst, and vertical standards, typically set by a regulatory agency (e.g., EPA's freshwater *Enterococcus* standard of 61 CFU/100 mL), separate the space into four quadrants: false positives (upper left), correct positives (upper right), correct negatives (lower left) and false negatives (lower right). The plots are used to count observations in each quadrant directly and provide the summary data entered in Table 8.2.

In addition, the analyst can determine the impact of changing the decision threshold (horizontal line) on the numbers of false negatives and false positives that a model produces. For example, in looking at PA predictions in Figure 8.1, lowering the decision threshold to 1.57 would result in three fewer false negatives, while not incurring any additional false positives.

Table 8.2 shows that the minimum specificities of both models were near 90 percent, meaning actual counts of bacteria below the standard were well-predicted. Some threshold-specific sensitivities, however, were much lower, indicating both models had more difficulty distinguishing false negatives from correct positives. Upon examination of Table 8.2, we conclude that the PA data set produced a model that performed well when compared to the PA+SS data set. Sensitivity for the PA model was on average 87 percent as large as sensitivity for the PA+SS model, and specificity values for the two models were nearly identical. So, while comparisons using adjusted R^2 and PRESS statistics definitely favored the PA+SS model, the threshold analysis, focusing only on placing prediction/observation pairs into one of four quadrants (as in Figure 8.2), showed fewer discrepancies between the two models.

Table 8.2. Results of the threshold analysis for the PA and PA+SS models.

Model	Threshold Type	Threshold	CN	FP	Specificity	CP	FN	Sensitivity
PA	Count	Log(30)	59	7	0.894	10	3	0.769
		Log(61)	66	0	1.000	7	6	0.538
		Log(100)	66	0	1.000	4	9	0.308
	Probability	30%	66	0	1.000	8	5	0.615
		50%	66	0	1.000	7	6	0.538
		70%	66	0	1.000	6	7	0.462
PA+SS	Count	Log(30)	62	4	0.939	11	2	0.846
		Log(61)	66	0	1.000	9	4	0.692
		Log(100)	66	0	1.000	4	9	0.308
	Probability	30%	66	0	1.000	9	4	0.692
		50%	66	0	1.000	8	5	0.615
		70%	66	0	1.000	8	5	0.615

Notes:
EPA's regulatory threshold for *Enterococcus* is 61 CFU/100 mL.
CN = number of correct negatives, FP = number of false positives, CP = number of correct positives, and FN = number of false negatives. Specificity = CN/(CN+FP), Sensitivity = CP/(CP+FN)

8.4 DISCUSSION

Our preliminary analyses indicate that the *best* models chosen by metrics that measure fit to training data (Mallows Cp, AIC, BIC, and such) are very similar to the models chosen by the PRESS statistic. That is not completely unexpected, given that our data are taken from a single swimming season. Environmental processes dictating the fate and transport of fecal indicators at the beach probably were stable over the period. At beach sites where there are intra-seasonal or annual shifts in the processes important to fate and transport of bacterial contaminants, models fit to a long-term series of previously collected data might not be best at predicting new observations (Frick et al. 2008). For example, if a new bacterial source near the beach is introduced (i.e., a breakwall is constructed), hydrologic patterns at the beach might be altered, or if an exceptional climatic period occurs, using a weighted regression that emphasizes the most recently collected data would be our recommended statistical technique.

Adopting higher decision thresholds (as shown in Figure 8.2) imposes more risk to public health than if lower horizontal thresholds were to be used; however, lowering the horizontal threshold can result in more frequent false positives. Such a strategy is more protective of public health, but it results in economic losses because of more frequent beach closures. When setting the horizontal thresholds, the manager must seek a balance between protection of public health and economic concerns. From a health perspective, the consequences of false negatives often are seen as more serious than those of false positives; therefore, horizontal count thresholds below the regulatory standard (corresponding to probability thresholds lower than 50 percent) are used. For example, the model developed for the combined (PA+SS) data set had a sensitivity of about

70 percent when using a probability threshold of 30 percent. That means, if a manager closed the beach whenever the model predicted a greater than 30 percent chance of exceeding the standard (61 CFU/100 mL), the standard would actually be exceeded on 70 percent of those days. An even lower probability threshold might improve the sensitivity, but it also could cause the model's specificity to decrease. Lower model sensitivities (relative to specificities) in the analysis were likely the result of a data set with very few large FIB counts. Regression models generally will fit the dominant type of observation well, especially when using the PRESS statistic as the selection criterion because it negates the influence of high leverage outliers; in this case, the majority of observations were small FIB counts (< 50 CFU/100 mL). A more balanced data set of small and large FIB levels would produce a model better suited to fit either type of observation. If the analyst is more interested in making good predictions when bacteria levels are high, one tactic could be a weighted regression with greater weights given to the larger FIB observations.

Managers attempting to model FIB beach levels with publicly available near-site data must consider the distance from the beach site to the monitoring location. In this case, approximately 4 miles separate South Shore Beach from the General Mitchell Airport. The geographic variability of meteorological data depends on the parameter (Rogers and Vanloon 1982; Koprov et al. 1998; Grants and Gerbeth 2003; Baigorria et al. 2007), but one would expect the predictive capabilities of meteorological data to decline as the distance between the monitoring location and beach site increases. Unfortunately, we cannot provide definitive advice on a maximum distance that should be considered.

8.5 CONCLUSIONS

We found that the PA data set could not match the performance of the PA+SS data set at South Shore Beach, in terms of the adjusted R^2 and PRESS statistics. However, given funding constraints or lack of availability of monitoring equipment at the beach of interest, a PA data set can provide a feasible alternative for developing an acceptable model. In our study, the PA data produced a model that explained about 47 percent of the variability in FIB levels. The predictive performance (as measured by the PRESS statistic) of the PA model was 11 versus 9 for the PA+SS models. To draw general conclusions about the utility of on-site data, however, more studies of this type need to be performed, including at marine and riverine beach sites where different processes could control the fate and transport of pathogens.

9 Advanced Techniques to Refine MLR Model Results

9.1 Introduction to Temporal Synchronization

For most IVs used in an MLR analysis, a value is recorded at the moment the water sample for FIB levels is taken, but rainfall is an exception. Cumulative precipitation over some period before the FIB water sample is often used as an IV in predictive beach bacteria modeling efforts (Francy and Darner 2006; Hose et al. 2006; Neumann et al. 2006; Nevers et al. 2009). MLR models might be improved if all measured environmental data, not just precipitation, were summarized over a longer temporal *window* rather than relying on the instantaneous value taken at the time of FIB sampling. In addition to taking a mean value of an IV over some temporal window, a temporal lag might also increase the relationship between an IV and the FIB response. For example, consider a water sample collected at 9 a.m. on a Thursday. For the IV *water temperature*, you wish to investigate a window of 24 hours and a lag of 12 hours. That means you will select a time 12 hours earlier (the lag) to 9 p.m. on Wednesday, then examine every water temperature measurement taken from that point to 24 hours earlier (the window), 9 p.m. on Tuesday. You would then compute the average of all the water temperature measurements over the window/lag combination. The mean value would become the new value of the water temperature for the Thursday 9 a.m. water sample. Note that using a window of zero and a lag of zero is equivalent to the traditional approach of using a single instantaneous value of the IV measured as the water sample is collected.

We gave this technique a name—*Temporal Synchronization Analysis* (TSA). Essentially, TSA seeks to maximize the correspondence between each IV and the FIB response by examining different combinations of temporal windows and lags over which the mean value of the IV is calculated. Those windows and lags are defined, relative to when the FIB concentration is measured. We set out to test whether a TSA could improve both the ability of an MLR model to fit a data set and, more importantly, the ability of an MLR model to make predictions of new observations. Such a technique assumes that the potential effects of the IVs on the response are hard to detect using single instantaneous measures of each IV, but the examination of mean values of IVs over a longer duration at some point in the past will lead to stronger empirical relationships between the IVs and the response variable.

Using such methods, one can compute an infinite number of different *enumerations*, or window/lag combinations, for each IV. One area of our research focused on criteria for selecting the enumeration that would be the best for empirical modeling. Note that if you begin the analysis with 10 IVs and, for each IV you investigate 10 enumerations, after selecting one enumeration for each IV, you would still have 10 IVs—but each would be temporally synchronized to the response variable.

9.2 TEMPORAL SYNCHRONIZATION AT SOUTH SHORE BEACH

9.2.1 Methods

One site where we applied TSA methods was South Shore Beach (Figure 4.2A). The period we examined was the swimming season of 2009. Site characteristics, laboratory methods, and our data collection efforts are detailed in Section 4.2.1. After data preparation, 44 FIB observations and 13 IVs remained: water temperature, conductivity, water depth, pH, turbidity, chloride, NH_4, NO_3, dry bulb air temperature, wet bulb air temperature, relative humidity, wind speed and wind direction. The last two variables were combined into two wind components: parallel to the shore (u component) and perpendicular to the shore (v component).

Cross-Validation. To begin the analysis, we randomly split the raw data into two parts: 75 percent of the 44 observations (33) were used as training data, and the methods described in the following sections were applied. The remaining 25 percent of the observations (11) were set aside as testing data. The randomization was done 500 times to create a population of 500 paired training/testing data sets that subsequently could be analyzed.

Temporal Synchronization. Each FIB measurement has a *time stamp* in the data set. That is a real number to represent the day and time each observation was collected. For instance, the time stamp of an observation made on July 13, 2008, at 9:00 a.m. would have a time stamp of 39642.375. The integer 39642 is the number of days since day 1 (January 1, 1900, in our convention, but it could be any day for calculation purposes.) The decimal 0.375 represents the time of day (9/24 hours). The time stamp for every IV measurement was also known. However, the IVs were measured much more frequently than the response variable, typically at intervals of every 5, 15, or 30 minutes. For our TSA, we wrote code (using the R statistical package) that would generate an array of data columns for every IV variable. Within each column, the IV would be averaged over a specific temporal window and lag, with respect to the time stamp of the response variable. We used four windows (0 days, 0.5 days, 1.5 days and 3 days) and four lags (0 days, 1 day, 1.5 days, and 3 days). Note that the 0-day window and 0-day lag correspond to traditional FIB regression modeling. We called each window-lag combination an *aspect*, generating 16 aspects (4 windows x 4 lags) of each IV.

Next, we choose a single aspect of the possible 16 for each IV to model the response variable. We considered two criteria to determine the aspect selection: a Pearson correlation coefficient between the response variable and each aspect, with the highest coefficient indicative of the best linear relationship; and the PRESS statistic calculated from a univariate regression of the response on each potential IV aspect. The Pearson coefficient emphasizes the best fit of a model to the data, while the PRESS statistic emphasizes prediction of observations not seen by the model. Lower PRESS values indicate greater predictive accuracy.

Model Selection. At this point, each of the 500 training data sets had been changed into three data sets: one where none of the IVs had been synchronized (the UNS data set); one *synchronized* data set that had IV aspects chosen by a Pearson coefficient (the PCC data set); and one synchronized data set where IV aspects were chosen using the PRESS statistic (the PRS data set). We then used MLR methods to develop predictive models of FIB densities using the three data sets. In developing our MLR models, we used two selection procedures:

1. For the UNS and PCC data sets, we used the AIC as a variable selection criterion.

2. For the UNS, PCC, and PRS data sets, we used the PRESS statistic to perform model selection.

For each analysis, we employed a backwards-stepwise variable selection algorithm.

Measuring Model Performance. To measure how well a chosen model fit its corresponding training data, we calculated the mean squared error of fitting (MEF) using the 33 training observations (which had been natural-log transformed). However, to measure predictive model performance, a model was selected using the 33 observations in the training data set, and then an MEP was calculated using the 11 observations in the testing data set (again, natural-log transformed).

To summarize, the following four-step process was repeated 500 times:

1. 75–25 percent randomized splitting of the original raw data into training and testing sub-groups
2. IV aspect selection (both Pearson coefficient and PRESS) using the training data.
3. Model selection (PRESS or AIC) using the training data.
4. Model performance metrics calculated (MEF for training data, MEP for testing data).

Afterwards, we made the following comparisons:

A. MEF of UNS versus PCC when AIC was used for model selection
B. MEP of UNS versus PRS when PRESS was used for model selection
C. MEP of UNS versus PCC when PRESS was used for model selection
D. MEP of PCC versus PRS when PRESS was used for model selection

The comparisons allowed us to measure the effectiveness of TSA for improving MLR fitting and predictive capabilities of the best-fit models.

9.2.2 Results

Comparison A. Across the 500 data sets (Figure 9.1, Table 9.1), the mean of the MEF for models developed using temporally synchronized data (IV aspects chosen using correlation coefficients) was 35 percent smaller than the mean of the MEF using unsynchronized data (16.9 versus 26.2, $p < 0.001$). When the statistical objective is fitting a set of training data, rather than prediction of new observations, selecting IV aspects using the maximum correlation coefficient should be ideal.

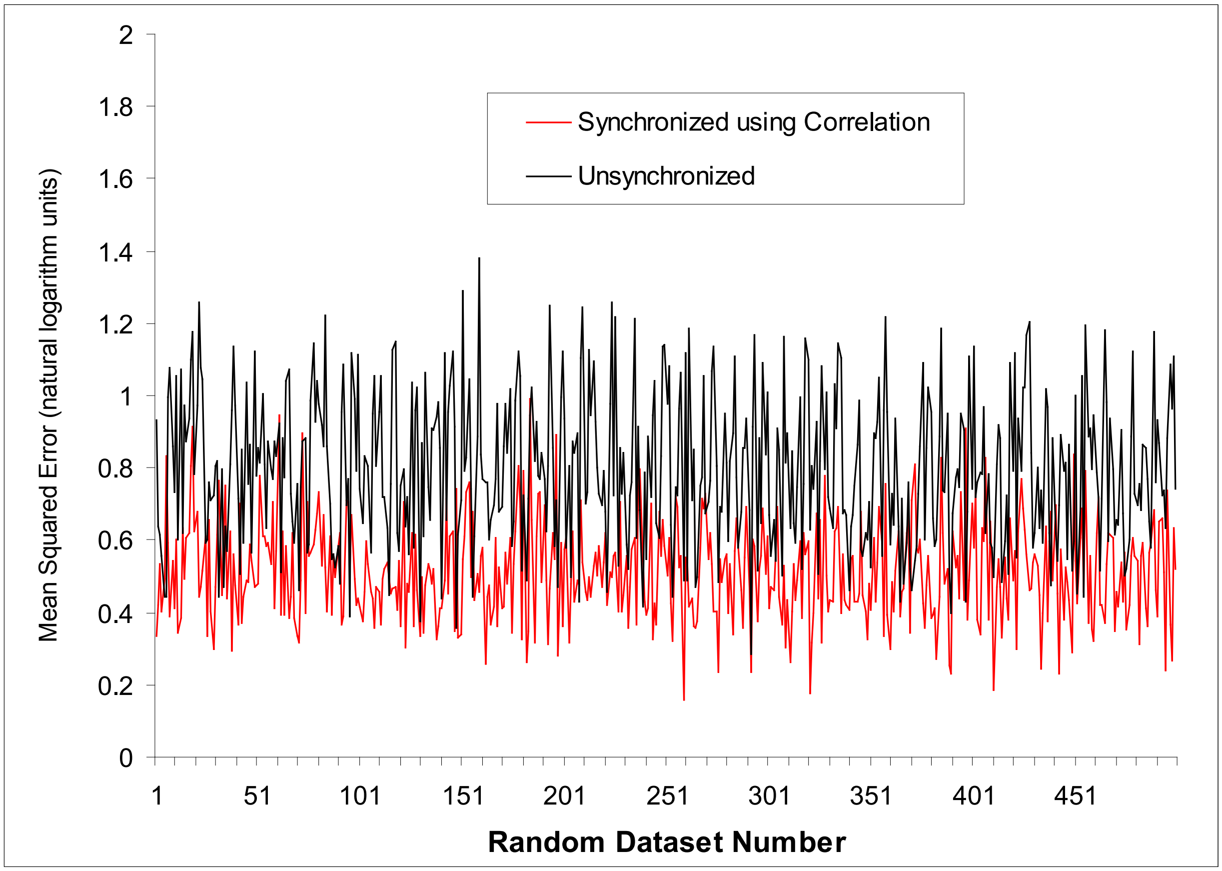


Figure 9.1. A comparison between the MEF values of model developed using temporally synchronized data (IV aspects selected using correlation coefficients) and unsynchronized data. Five-hundred randomly generated data sets were examined. Each data set had 33 observations. Note that we modeled the natural logarithm of *Enterococcus* CFU measurements in these models.

Table 9.1. Statistics for the 500 MEF and MEP values obtained from models developed using temporally synchronized data (both PRESS-selected IV aspects, PRS, and correlation coefficient-selected IV aspects, PCC) and unsynchronized data (UNS). The natural logarithm of *Enterococcus* CFU measurements were modeled in this study.

				95% Confidence Interval on the Mean	
Metric	Dataset	Mean	St.Dev.	Lower Bound	Upper Bound
MEF	UNS	0.795	0.206	0.777	0.813
	PCC	0.513	0.137	0.501	0.525
MEP	UNS	3.670	4.990	3.230	4.110
	PCC	2.710	1.290	2.590	2.820
	PRS	1.850	0.940	1.760	1.930

Comparison B. Figure 9.2 shows the 500 MEP values for models, using the PRS data (red line) and the UNS data (black line). The mean MEP for the PRS data was 1.85, and for the UNS data it was 3.67—nearly twice as large ($p < 0.001$). For many data sets, the UNS data produced a MEP similar to that of the PRS data, but there were also many spikes; the MEP for the UNS data was much higher than that of the PRS data. We can conclude that temporal synchronization helped to improve predictions of new observations when fitting a MLR model.

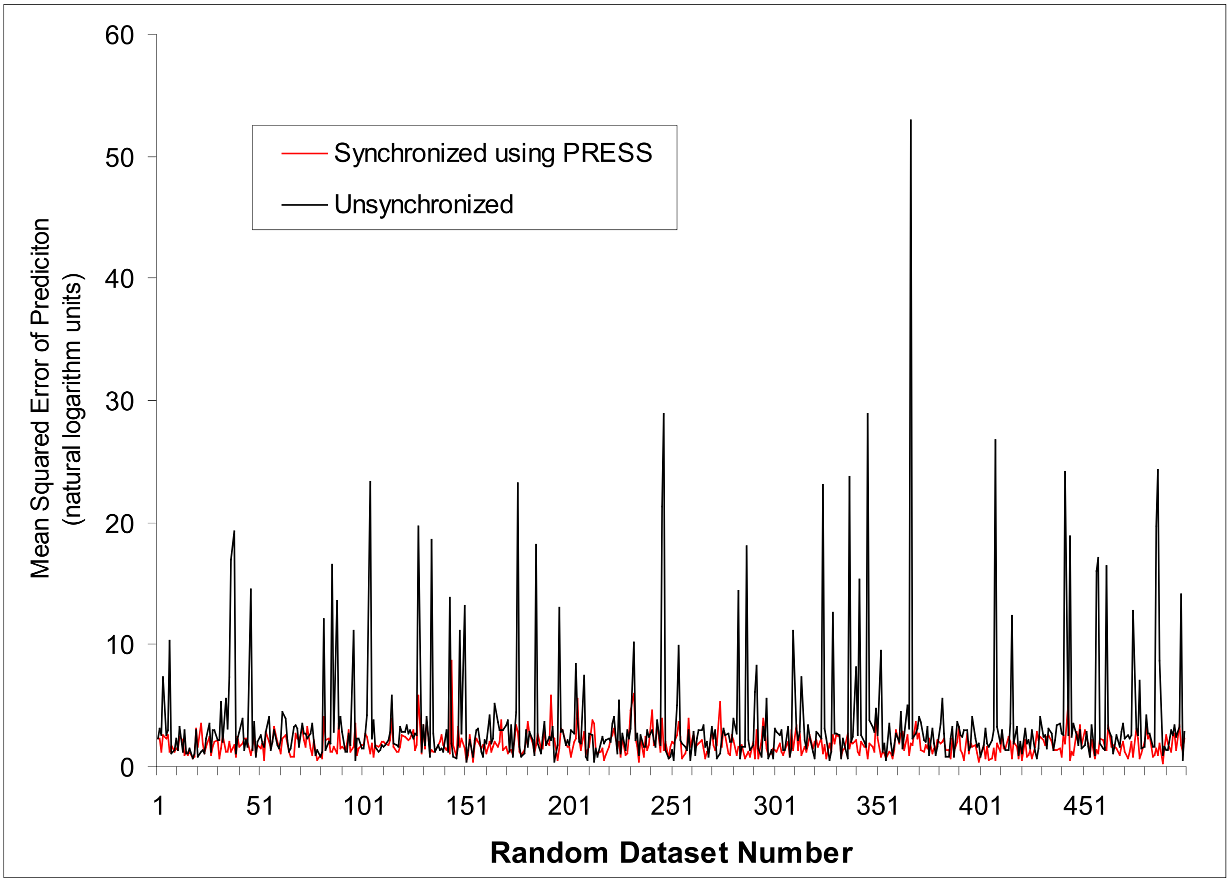


Figure 9.2. A comparison of the 500 MEP values for the PRS data and the UNS data.

Comparison C. The MEP values produced by models fit to the PCC data sets are also smaller, on average, than the mean MEP values for the UNS data sets (Figure 9.3). The mean MEP value for the PCC data sets was 2.71 versus 3.67 for the UNS data sets ($p < 0.001$).

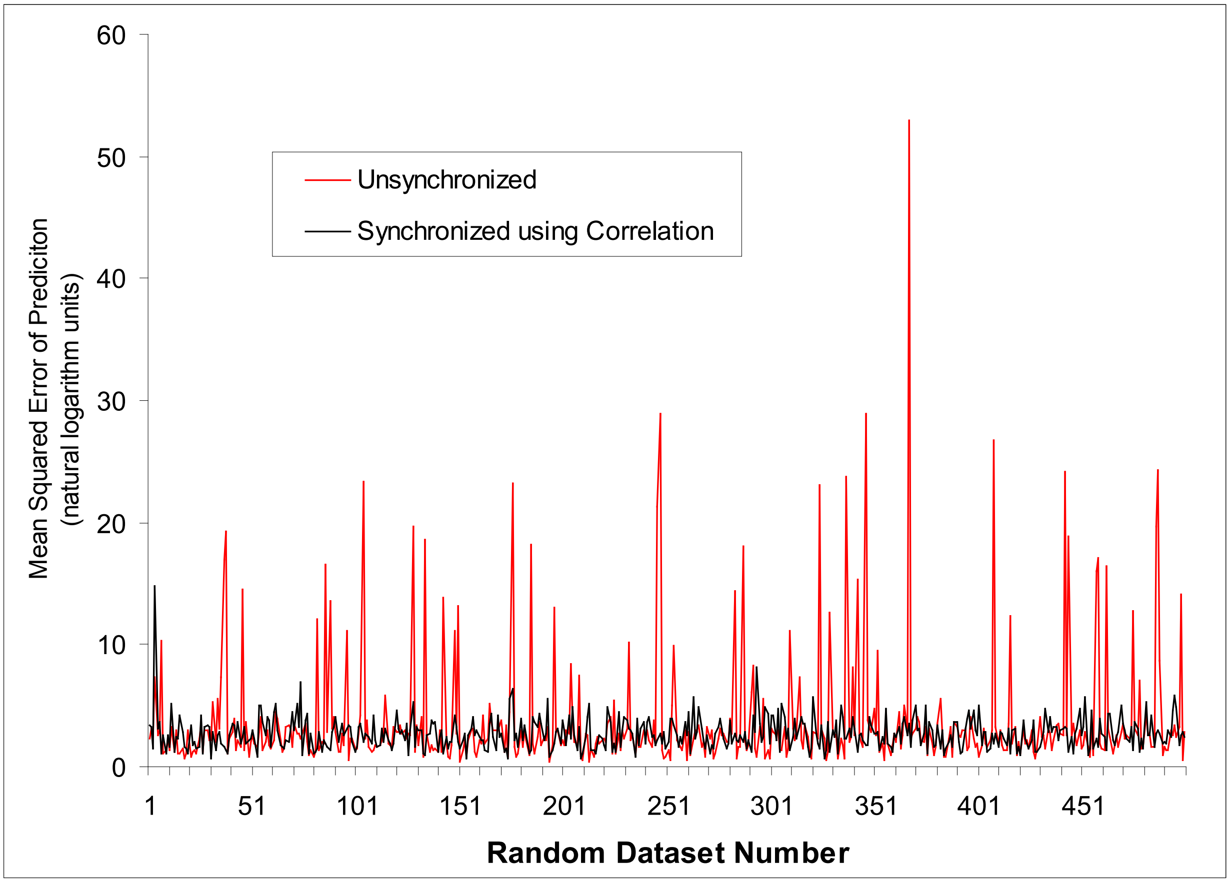


Figure 9.3. A comparison of the 500 MEP values for the PCC data and the UNS data.

Comparison D. Figure 9.4 shows that the MEP values for the PRS data sets were indeed smaller, on average, than the MEP values for the PCC data sets (mean of 1.85 versus 2.71, $p < 0.001$). Thus, choosing IV aspects with a method that emphasizes predictive modeling (PRESS) will result in better predictions than synchronizing data (choosing IV aspects) with a simple correlation coefficient that emphasizes data fitting.

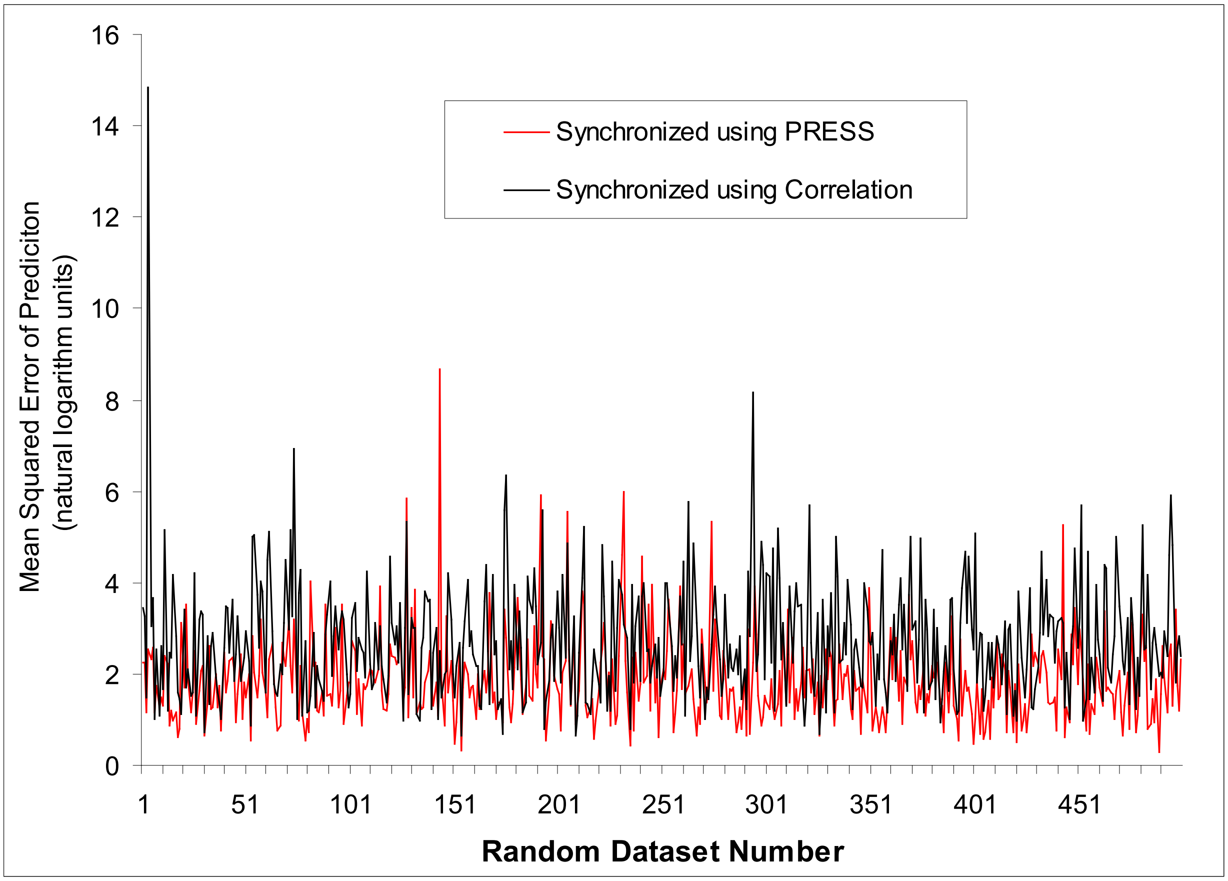


Figure 9.4. The comparison between the 500 MEP values for the models developed using the two different temporal synchronized data, PRESS-selected IV aspects and correlation-coefficient-selected IV aspects.

Methodological Comments. For simplicity and clarity in this South Shore TSA, we investigated a limited number of windows and lags, the longest of which was 72 hours. Expanding to a week or more, if the data allow, could produce additional predictive improvements. In addition, we computed a mean value of the IV over each window, but other statistics such as the median, standard deviation, maximum or minimum value of the IV might be a more potent correlate to the response variable.

The TSA can be likened to a type of transformation of each IV, but instead of using a logarithmic or square root transformation to linearize the relationship between the IV and the response, we temporally transform the IV to discover if windows and lags can improve its predictive power. The most useful windows and lags most certainly would be dependent on SS relationships of pathogen sources to each individual beach site. The best aspect of each IV might even vary annually or seasonally, depending on SS dynamics of pathogen fate and transport.

9.3 ADVANCED MODELING TECHNIQUES AT HOBIE BEACH, MIAMI

9.3.1 Introduction

Standard MLR techniques were not able to produce successful models at Hobie Beach, Miami (Figure 4.2B) for CFU data collected in the summer of 2008 (adjusted R^2 of 20 percent). Conditions and historical data for the beach are described in Section 4.3.5. For data collected by Florida in 2005–2008, the adjusted R^2 of the best CFU regression model was only 7 percent (see Appendix C). We decided to apply more sophisticated techniques to the summer 2008 CFU data in hopes of improving the statistical models. The first method was a TSA of the IVs. The second technique can be best described as data sub-setting, or a form of *hierarchical modeling*.

9.3.2 Temporal Synchronization Analysis

Because we performed TSA before making refinements in the approach, the TSA for Hobie Beach data was not as sophisticated as that for South Shore, Milwaukee. For Hobie, we examined 14 different windows ranging from one hour to 28 days, but we did not incorporate lags. We used the Pearson correlation coefficient for IV *aspect* selection and did not consider the PRESS statistic. After running a MLR analysis with the new IVs, we produced a five-parameter model (water depth, air temp, water temp, photosynthetically active radiation, dew point) that had an adjusted R^2 of 0.41, twice that of the original adjusted R^2 of 0.2 obtained using unsynchronized IVs. Thus, even without using lags in the TSA, we saw a large increase in the fit achieved by the empirical model.

9.3.3 Data Sub-Setting

As stated, the best regression model fit to 2008 Hobie Beach CFU data (137 observations, no TSA used on the IVs), was not very convincing, with an adjusted R^2 of about 0.2 and residuals that failed a normality test. In hopes of improving the model, we examined the DFFIT values of each observation. The DFFIT value for an observation is a statistic that measures how influential a single observation is on a regression model. It is defined as the change in the predicted value of an observation (i.e., a comparison of the predicted value when the observation is removed from the dataset versus its predicted value when it remains in the dataset), divided by the standard deviation of the observation's predicted value when the data point has been excluded.

$$DFFIT = (\hat{y}_{\mathrm{i}} - \hat{y}_{\mathrm{-i}}) / s_{\mathrm{-i}} \times h_{\mathrm{ii}}^{0.5}$$

where $\hat{y}_{\mathrm{i}}$ is the predicted value of the *i*th observation, $\hat{y}_{\mathrm{-i}}$ is the predicted value of the *i*th observation computed after removing the *i*th observation from the data, $s_{\mathrm{-i}}$ is the standard error of $\hat{y}_{\mathrm{-i}}$, and h_{ii} is the leverage of the *i*th observation. Observations with large DFFIT values (negative or positive) have a large influence on the fitted regression coefficients and are often interpreted as outliers.

To improve the model fit, we could remove observations with the largest DFFIT value, one by one, re-fitting the model after each removal and then rechecking the remaining DFFIT values. As the process is repeated, the adjusted R^2 rises and the residuals might become more normally distributed (Figure 9.5). However, the analyst must make a decision at some point to stop the process. Usually that is done using a general threshold on what exactly constitutes a DFFIT value that is too large. The decision, however, might also be made on the basis of a leveling off of the adjusted R^2—that occurred in our data set after about 65 observations had been removed. Alternatively, it could be done when the residuals appear to be normally distributed (when the p value for the residual normality test exceeds 0.05)—which for this data set also occurred after removing about 65 observations.

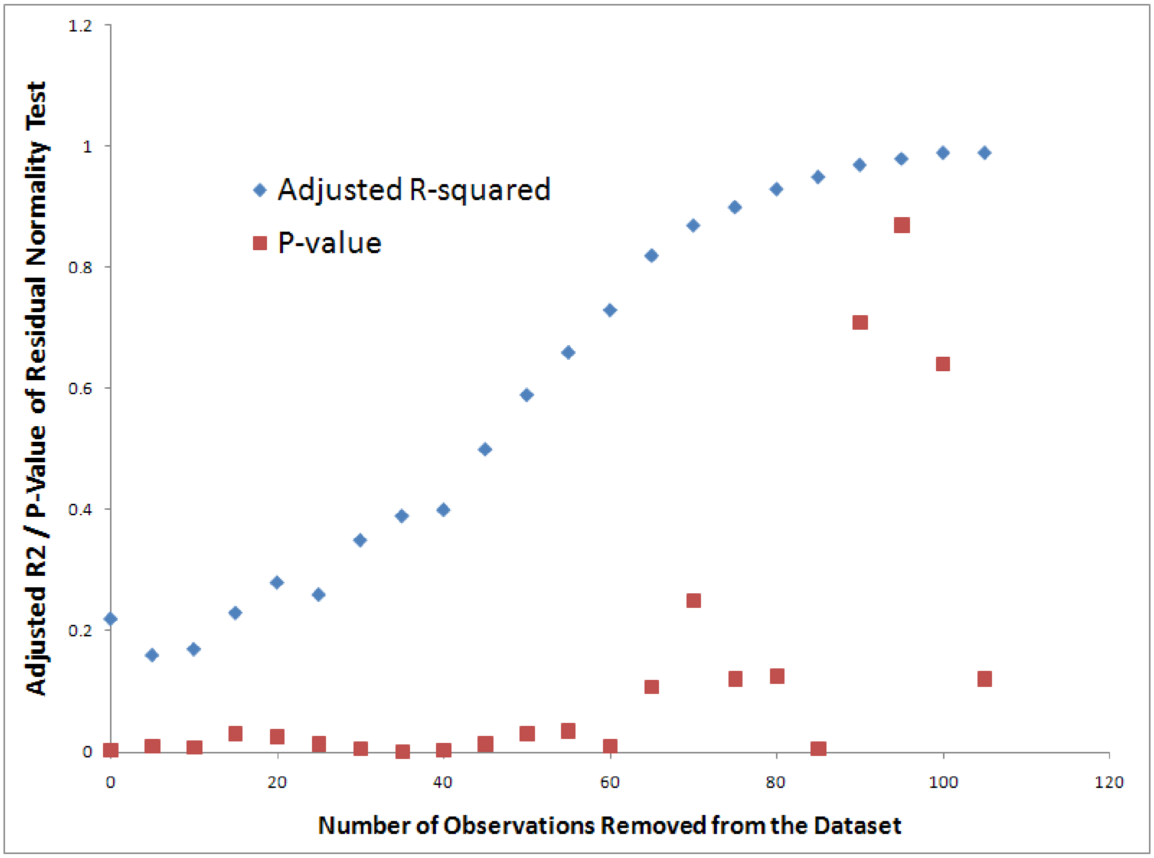


Figure 9.5. Effects on the regression adjusted R^2 and the significance of the residual normality test when successively removing observations based on the largest remaining DFFIT value in the data set.

Looking at the data set after 65 points had been removed, 72 observations remain. The adjusted R^2 is about 0.81, a very impressive fit. However, some questions arise. We now have a drastically different data set compared to what we started with. A model using completely different IVs fit to the 72 points might be even better. Also, what should be done with the 65 points that had been removed? It might be possible to achieve a good fit to those observations with another regression model, or perhaps no model can be found to fit them well. In our case, after removing 5 observations with very large DFFIT values, we had a good regression model for the remaining 60 observations (adjusted $R^2 = 0.5$, residuals normally distributed). Whatever the case, we need to answer the following question: when new data are collected and a prediction needs to be made, what should we do? Do we use the model for the 72 observations, or do we decide that new observation is more similar to the 65 removed observations?

To answer the question, we could use an advanced statistical technique like a discriminant analysis, which is a multivariate method whose objective is to find variables that play the most significant roles in discriminating between two or more groups of objects. For that analysis we would use the same IVs as used for the regression models but focus on finding a combination of IVs that can best distinguish if an observation belongs to Group 1 (the 72 observations) or Group 2 (the 65 observations). For example, such a technique might tell us that one model fits observations taken on warm, sunny days and another model should be used for cooler, cloudy days. That information would then allow us to choose between the use of various models depending on environmental conditions, and we would have added confidence in their predictions. Such a technique is related to *hierarchical modeling*, where subsets of observations are analyzed with different models to achieve better results.

After performing a discriminant analysis on the Hobie Beach data, we found a discriminant function that used three IVs—water depth, air temperature, and UV radiation—and it was 65 percent successful in categorizing our observations as belonging to Group 1 or Group 2 (64 percent successful for Group 1 and 66 percent successful for Group 2). Thus, if any observation were chosen from the data set, we would have a decent chance of correctly classifying it as a Group 1 or Group 2 observation, on the basis of those three IVs. Without that information (if done completely randomly), we would have about a 53 percent chance of success for classifying Group 1 observations and a 47 percent chance for Group 2 observations. If a new data point were predicted to be in Group 1, we could then apply the Group 1 regression model to achieve a much higher adjusted R^2 than if we had attempted to model all 137 observations using a single regression function. The danger of the technique is that if we applied that regression model to a Group 2 observation, the prediction would be very poor. In general, we would like the discriminant function to have a high rate of classification success (85–90 percent) for us to have confidence in using the method.

We should note that the example discussed in this section is not typical of the analysis we performed on the other data sets presented in this report. One usually finds only a handful of observations with large enough DFFIT values to be considered outliers, and thus possibly warrant exclusion from the data set. Indeed, with the Hobie Beach data, only a few observations (out of the original 137) actually had DFFIT values large enough to be considered for exclusion under normal circumstances. The problem was that after excluding the data points, we were still left with a poor regression model (adjusted R^2 under 0.2, and non-normal residuals). At that point, we wanted to go beyond the typical and demonstrate an analysis that can produce much better regression models for a given data set. However, analysts must exercise caution when using the technique so as not to completely overlook sound statistical principles.

10 Acknowledgements

We gratefully acknowledge the assistance of many institutions and individuals in obtaining the data, refining and evaluating the predictive models, and communicating the use of the software.

We especially thank Candida West, U.S. EPA, ORD, National Exposure Research Laboratory for her support and many inputs to this research. Bruce Mintz of U.S. EPA, National Exposure Research Laboratory assisted in many ways with communicating the research to ORD management and the Office of Water. Jim Kitchens, Brenda Kitchens, Lourdes Prieto and Carlyn Haley (U.S. EPA, ORD, National Exposure Research Laboratory) all made substantive contributions to administrative aspects of the project. We appreciate receiving data from the NEEAR epidemiological research team for this project and particularly thank Tim Wade, Elizabeth Sams, Al Dufour, Richard Haugland, Kristen Brenner, and Larry Wymer for their assistance. Many useful review comments were received from John Wathen and Beth Leamond of the U.S. EPA Office of Water.

Analyses of the data at the U.S. EPA, ORD, National Exposure Research Laboratory Ecosystems Research Division, were conducted by O'Niell Tedrow, Jack Varner, Shayla Hunter, Eva Duvall, Ellen Price, Caitlin Sloan, Chris Fitzgerald, and Anjali Viswakumar. Shuyan Zhang, Jon Wong, Walter Frick, and Zhongfu Ge all contributed to the modeling studies.

We had many fruitful interactions related to the U.S. EPA Advanced Monitoring Initiative (AMI). We thank the AMI program for its support of part of this research. Richard Zdanowitz from Region 5, Chicago was helpful in many different ways. He helped to coordinate workshops and a series of teleconferences with scientists from USGS, NOAA and EPA that provided many useful insights into modeling activities in the Great Lakes region. Also, thanks go to Holly Wirick of Region 5 for her assistance with developing liaisons with others who were conducting beach modeling studies in the Great Lakes region. We also thank Adam Mednick and Dreux Watermolen, State of Wisconsin, Dept. of Natural Resources for their outstanding efforts with communicating the use of the Virtual Beach software.

At South Shore Beach, Milwaukee we were assisted by Lindsay Olson (Student Services Associate); Sandra McLellan, Sabrina Mueller-Spitz, Robert Paddock, Geoff Anderson, Jack Orchard, and Greg Barske of University of Wisconsin- Milwaukee Great Lakes WATER Institute; Daniel Feinstein of U.S.G.S.- Madison WI; Richard Henry and Alan Humphrey of U.S. EPA, Region 2; Don Conlee of Nortek USA; and ERG with sample collection and analysis and equipment maintenance. At Hobie Beach, Miami Diana Aranda, Chris Sinigalliano, and Maribeth Gidley of NOAA-Atlantic Oceanographic Marine Laboratory; Erich Bartels of Mote Marine Laboratory; Jed Campbell of U.S. EPA, National Health and Environmental Effects Laboratory, Gulf Ecology Division; Manuel Collazo, Lora Fleming, John Wang and Helena Solo-Gabriele of University of Miami); Samir Elmir of Miami-Dade County Health Department); and ERG helped collect and analyze samples and maintain the equipment. Likewise, at La Monserrate we were assisted by Maria Vega-Rodriguez(Student Services Associate), Milton Carlo of University of Puerto Rico-Mayaguez); Angel Meléndez of PREQB; Mildred Matos, Juan Pacheco, and Eddie Nieves of Puerto Rico Department of Natural Resources); Jose Font of U.S. EPA, Region 2; and Luquillo lifeguards. At Surfside Beach Henry Styron of University of North Carolina at Wilmington; Shannon Berry, Sean Torrens, and

Ted Ambrose of South Carolina Department of Health and Environmental Control; Micki Fellner, Ed Booth, and beach personnel of the Town of Surfside Beach, SC; Carolyn Ross and Kellah Webster (Westat subcontractors) and Westat all provided assistance. At Boquerón, able assistance was provided by Milton Carlo and Ernesto Otero-Morales of University of Puerto Rico- Mayaguez; Angel Meléndez of PREQB; and Eddie Nieves, beach staff and life guards of Puerto Rico Department of Natural Resources; Gary Toranzos of University of Puerto Rico; Raisha Cornier (Westat sub-contractors), and Westat personnel.

11 References

Abdelzaher, A.M., M.E. Wright, C. Ortega, H.M. Solo-Gabriele, G. Miller, S. Elmir, X. Newman, P. Shih, J.A. Bonilla, T.D. Bonilla, C.J. Palmer, T. Scott, J. Lukasik, V.J. Harwood, S. McQuaig, C. Sinigalliano, M. Gidley, L.R.W. Plano, X. Zhu, J.D. Wang, and L.E. Fleming. 2010. Presence of pathogens and indicator microbes at a non-point source subtropical recreational marine beach. *Applied and Environmental Microbiology* 76(3):724-732.

ADEM (Alabama Department of Environmental Management). 2010. *DRAFT Total Maximum Daily Load (TMDL) for Mobile Bay*. Alabama Department of Environmental Management, Montgomery, AL.

Aguinis, H., and E. Harden. 2009. Sample size rules of thumb: evaluating three common practices. In *Statistical and Methodological Myths and Urban Legends: Received Doctrine, Verity and Fable in the Organizational and Social Sciences*, ed. C. Lance and R. Vandenberg. Routledge, New York.

Baigorria, G.A., J.W. Jones, and J.J. O'Brien. 2007. Understanding rainfall spatial variability in southeast USA at different timescales. *International Journal of Climatology* 27(6):749-760.

Boehm, A., R. Whitman, M. Nevers, D. Hou, and S. Weisberg. 2007. Now-Casting Recreational Water Quality. In *Statistical Framework for Water Quality Criteria and Monitoring*, ed L.J. Wymer. John Wiley & Sons, Hoboken, NJ.

Boehm, A.B. 2003. Model of microbial transport and inactivation in the surf zone and application to field measurements of total coliform in Northern Orange County, California. *Environmental Science & Technology* 37(24):5511-5517.

Boehm, A.B. 2007. Enterococci concentrations in diverse coastal environments exhibit extreme variability. *Environmental Science & Technology* 41(24):8227-8232.

Boehm, A.B., S.B. Grant, J.H. Kim, S.L. Mowbray, C.D. McGee, C.D. Clark, D.M. Foley, and D.E. Wellman. 2002. Decadal and shorter period variability of surf zone water quality at Huntington Beach, California. *Environmental Science & Technology* 36(18):3885-3892.

Boehm, A.B., K.M. Yamahara, D.C. Love, B.M. Peterson, K. McNeill, and K.L. Nelson. 2009. Covariation and Photoinactivation of Traditional and Novel Indicator Organisms and Human Viruses at a Sewage-Impacted Marine Beach. *Environmental Science & Technology* 43(21):8046-8052.

Byappanahalli, M.N., R.L. Whitman, D.A. Shively, and M.B. Nevers. 2010. Linking non-culturable (qPCR) and culturable enterococci densities with hydrometeorological conditions. *Science of the Total Environment* 408(16):3096-3101.

Clesceri, L.S., A.E. Greenberg, and A.D. Eaton, eds. 1998. *Standard Methods for the Examination of Water and Wastewater*. 20th ed. American Water Works Association and Water Environment Federation, Washington, DC.

Cohen, J. 1988. *Statistical power analysis for the behavioral sciences*. 2nd ed. Academic Press, New York.

Elmir, S.M. 2006. *Development of a water quality model which incorporates non-point microbial sources*. University of Miami, Civil Engineering, Coral Gables, FL.

Francy, D.S., A.M. Gifford, and R.A. Darner. 2003. *Escherichia coli at Ohio Bathing Beaches - Distribution, Sources, Wastewater Indicators, and Predictive Modeling*. U.S. Department of the Interior, U.S. Geological Survey, Water-Resources Investigations Report 02-4285, Columbus, OH.

Francy, D.S., and R.A. Darner. 2006. *Procedures for developing models to predict exceedances of recreational water-quality standards at coastal beaches*. U.S. Department of the Interior, U.S. Geological Survey, Techniques and Methods 6–B5.

Francy, D.S., and R.A. Darner. 2007. *Nowcasting beach advisories at Ohio Lake Erie Beaches*. U.S. Department of the Interior, U.S. Geological Survey, 2007-1427.

Francy, D.S., E.E. Bertke, D.P. Finnegan, C.M. Kephart, R.A. Sheets, J. Rhoades, and L. Stumpe. 2006a. *Use of spatial sampling and microbial source-tracking tools for understanding fecal contamination at two Lake Erie beaches*. U.S. Department of the Interior, U.S. Geological Survey, 2006-5298.

Francy, D.S., R.A. Darner, and E.E. Bertke. 2006b. *Models for predicting recreational water quality at Lake Erie beaches*. U.S. Department of the Interior, U.S. Geological Survey, Scientific Investigations Report 2006-5192.

Frick, W.E., Z. Ge, and R.G. Zepp. 2008. Nowcasting and forecasting concentrations of biological contamination at beaches: A feasibility and case study. *Environmental Science & Technology* 42:4818-4824.

Ge, Z., and W.E. Frick. 2007. Some statistical issues related to multiple linear regression modeling of beach bacteria concentrations. *Environmental Research* 103:358-364.

Ge, Z., and W.E. Frick. 2009. Time-frequency analysis of beach bacteria variations and its implication for recreational water quality modeling. *Environmental Science & Technology* 43:1128-1133.

Grant, S.B., and B. Sanders. 2010. The beach boundary layer: A framework for addressing recreational water quality impairment at enclosed beaches. *Environmental Science & Technology* in press.

Grants, I., and G. Gerbeth. 2003. Experimental study of non-normal nonlinear transition to turbulence in a rotating magnetic field driven flow. *Physics of Fluids* 15(10):2803-2809.

Haenlein, M., and A.M. Kaplan. 2004. A beginner's guide to partial least squares analysis. *Understanding Statistics* 3(4):283-297.

Haugland, R.A., S.C. Siefring, L.J. Wymer, K.P. Brenner, and A.P. Dufour. 2005. Comparison of *Enterococcus* measurements in freshwater at two recreational beaches by quantitative polymerase chain reaction and membrane filter culture analysis. *Water Research* 39:559-568.

Heaney, C.D., E. Sams, S. Wing, S. Marshall, K. Brenner, A.P. Dufour, and T.J. Wade. 2009. Contact with beach sand among beachgoers and risk of illness. *American Journal of Epidemiology* 170(2):164-172.

Hose, G.C., B.R. Murray, M.L. Park, B.P. Kelaher, and W.F. Figueira. 2006. A meta-analysis comparing the toxicity of sediments in the laboratory and in situ. *Environmental Toxicology and Chemistry* 25(4):1148-1152.

Hou, D., S.J. Rabinovici, and A.B. Boehm. 2006. Enterococci predictions from a partial least squares regression model can improve the efficacy of beach management advisories. *Environmental Science & Technology* 40(6):1737-1743.

Huang, X., and V. Sigler. 2006. *Population-Based Molecular-Tracking of E. Coli at Lake Erie Beach and Huntington Beach (Ohio)*. Paper read at the 49th Annual Conference of the International Association of Great Lakes Research, May 22–26, Windsor, Ontario.

IDEM (Indiana Department of Environmental Management). 2004. *Little Calumet and Portage Burns Waterway TMDL for E. coli Bacteria*. Final TMDL Report. Indiana Department of Environmental Management, Indianapolis, IN.

Kim, J.H., and S.B. Grant. 2004. Public mis-notification of coastal water quality: A probabilistic evaluation of posting errors at Huntington Beach, California. *Environmental Science & Technology* 38(9):2497-2504.

Koprov, B.M., S.L. Zubkovsky, V.M. Koprov, M.I. Fortus, and T.I. Makarova. 1998. Statistics of air temperature spatial variability in the atmospheric surface layer. *Boundary-Layer Meteorology* 88(3):399-423.

Liu, L., M.S. Phanikumar, S.L. Molloy, R.L. Whitman, D.A. Shively, M.B. Nevers, D.J. Schwab, and J.B. Rose. 2006. Modeling the transport and inactivation of *E. coli* and Enterococci in the near-shore region of Lake Michigan. *Environmental Science and Technology* 40:5022-5028.

McLellan, S. 2004. *Sources of E. coli at South Shore Beach Final Research Report*. Great Lakes WATER Institute, Milwaukee, WI.

McLellan, S.L., A.D. Daniels, and A.K. Salmore. 2001. Clonal populations of thermotolerant enterobacteriaceae in recreational water and their potential interference with fecal *Escherichia coli* counts. *Applied and Environmental Microbiology.* 67(10):4934-4938.

McLellan, S.L., E.J. Hollis, M.M. Depas, M. Van Dyke, J. Harris, and C.O. Scopel. 2007. Distribution and fate of *Escherichia coli* in Lake Michigan following contamination with urban stormwater and combined sewer overflows. *Journal of Great Lakes Research* 33(3):566-580.

McLellan, S.L., and E.T. Jensen. 2005. *Identification and Quantification of Bacterial Pollution at Milwaukee County Beaches*. Great Lakes WATER Institute, Milwaukee, WI.

McLellan, S.L., and A.K. Salmore. 2003. Evidence for localized bacterial loading as the cause of chronic beach closings in a freshwater marina. *Water Research* 37(11):2700-2708.

McQuarrie, A., and C.L. Tsai. 1998. *Regression and Time Series Model Selection*. World Scientific, Hackensack, NJ. MDEQ (Mississippi Department of Environmental Quality). 2002. *Fecal Coliform TMDL for the Back Bay of Biloxi and Biloxi Bay*. Mississippi Department of Environmental Quality, Jackson, MS.

MDEQ (Michigan Department of Environmental Quality). 2003. *Total Maximum Daily Load for Escherichia coli for the St. Joseph River Berrien County*. Michigan Department of Environmental Quality, Lansing, MI.

Mednick, A., and D. Watermolen. 2009. Beach pathogen forecasting tools: Pilot testing, outreach, and technical assistance. Edited by Bureau of Science Services. Wisconsin Department of Natural Resources, Madison, WI.

MMSD (Milwaukee Metropolitan Sewerage District). 2005. *Bacteria Source, Transport and Fate Study - Phase 1, Milwaukee Harbor Estuary Hydrodynamic & Bacteria Modeling*. Milwaukee Metropolitan Sewerage District, Milwaukee, WI.

Neumann, C.M., A.K. Harding, and J.M. Sherman. 2006. Oregon Beach Monitoring Program: Bacterial exceedances in marine and freshwater creeks/outfall samples, October 2002-April 2005. *Marine Pollution Bulletin* 52(10):1270-1277.

Nevers, M.B., D.A. Shively, G.T. Kleinheinz, C.M. McDermott, W. Schuster, V. Chomeau, and R.L. Whitman. 2009. Geographic relatedness and predictability of *Escherichia coli* along a peninsular beach complex of Lake Michigan. *Journal of Environmental Quality* 38(6):2357-2364.

Nevers, M.B., and R.L. Whitman. 2005. Nowcast modeling of *Escherichia coli* concentrations at multiple urban beaches of southern Lake Michigan. *Water Research* 39:5250-5260.

Nevers, M.B., and R.L. Whitman. 2008. Coastal strategies to predict *Escherichia coli* concentrations for beaches along a 35 km stretch of Southern Lake Michigan. *Environmental Science & Technology* 42:4454-4460.

Nevers, M.B., R.L. Whitman, W.E. Frick, and Z. Ge. 2007. Interaction and influence of two creeks on *Escherichia coli* concentrations of nearby beaches: Exploration of predictability and mechanisms. *Journal of Environmental Quality* 36:1338-1345.

NIRPC (Northwestern Indiana Regional Planning Commission). 2005. *Watershed Management Plan for Lake, Porter, and LaPorte Counties*. Northwestern Indiana Regional Planning Commission, Portage, IN.

Olyphant, G.A. 2005. Statistical basis for predicting the need for bacterially induced beach closures: Emergence of a paradigm? *Water Research* 39:4953-4960.

Olyphant, G.A., J. Thomas, R.L. Whitman, and D. Harper. 2003. Characterization and statistical modeling of bacterial (*Escherichia coli*) outflows from watersheds that discharge into southern Lake Michigan. *Environmental Monitoring and Assessment* 81:289-300.

Olyphant, G.A., and R.L. Whitman. 2004. Elements of a predictive model for determining beach closures on a real time basis: The case of 63rd Street Beach Chicago. *Environmental Monitoring and Assessment* 98(1-3):175-190.

Otvos, E.G. 1999. Rain-induced beach processes; landforms of ground water sapping and surface runoff. *Journal of Coastal Research* 15(4):1040-1054.

RIDEM (Rhode Island Department of Environmental Management). 2005. *Total Maximum Daily Load Analysis for Greenwich Bay Waters: Pathogen/Bacteria Impairments*. Rhode Island Department of Environmental Management, Providence, RI.

Rogers, J.C., and H. Vanloon. 1982. Spatial Variability of Sea-Level Pressure and 500 Mb Height Anomalies over the Southern-Hemisphere. *Monthly Weather Review* 110(10):1375-1392.

Schwab, D.J., D. Beletsky, and G.A. Lang. 2010. *Indiana Dunes Nowcast* 2010. <http://www.glerl.noaa.gov/res/glcfs/bd>. Accessed July 2010.

Scopel, C.O., J. Harris, and S.L. McLellan. 2006. Influence of nearshore water dynamics and pollution sources on beach monitoring outcomes at two adjacent Lake Michigan beaches. *Journal of Great Lakes Research* 32(3):543-552.

Shibata, T., H.M. Solo-Gabriele, L.E. Fleming, and S. Elmir. 2004. Monitoring marine recreational water quality using multiple microbial indicators in an urban tropical environment. *Water Research* 38:3119-3131.

Siefring, S., M. Varma, E. Atikovic, L. Wymer, and R.A. Haugland. 2008. Improved real-time PCR assays for the detection of fecal indicator bacteria in surface waters with different instrument and reagent systems. *Journal of Water and Health* 6:225-237.

Smith, R.C., and K.S. Baker. 1981. Optical properties of the clearest natural waters (200-800 nm). *Applied Optics* 20(2):177-184.

Solo-Gabriele, H., T. Shibata, M. Al-Kendi, Y. St. Fort, L. Fleming, D. Squicciarini, W. Quirino, M. Arguello, S. Elmir, and M. Rybolowik. 2002. *A Pilot Study Evaluation and Sanitary Survey of Microbial Recreational Water Quality Indicators in the Subtropical Marine Environment*. National Institute of Environmental Health Sciences Marine and Freshwater Biomedical Sciences Center, Rosensteil School of Marine and Atmospheric Sciences, Miami-Dade County Department of Health.

Telech, J.W., K.P. Brenner, R. Haugland, E. Sams, A.P. Dufour, L. Wymer, and T.J. Wade. 2009. Modeling Enterococcus densities measured by quantitative polymerase chain reaction and membrane filtration using environmental conditions at four Great Lakes beaches. *Water Research* 43:4947-4955.

Thupaki, P., M.S. Phanikumar, D. Beletsky, D.J. Schwab, M.B. Nevers, and R.L. Whitman. 2009. Budget Analysis of *Escherichia coli* at a Southern Lake Michigan Beach. *Environmental Science & Technology* 44(3):1010-1016.

Triad Engineering Incorporated. 2003. *Trail Creek Escherichia coli TMDL Report*. Indiana Department of Environmental Management, Indianapolis, IN.

USEPA (U.S. Environmental Protection Agency). 2007. *Report of the Experts Scientific Workshop on Critical Research Needs For the Development of New or Revised Recreational Water Quality Criteria*. EPA 823-R-07-006, EPA 823-R-07-006. U.S. Environmental Protection Agency, Washington, DC.

USEPA (U.S. Environmental Protection Agency). 2006. *Method 1600: Enterococci in Water by Membrane Filtration Using Membrane-Enterococcus Indoxyl-β-D-Glucoside Agar (mEI)*. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

USEPA (U.S. Environmental Protection Agency). 2010. *Method A: Enterococci in Water by TaqMan® Quantitative Polymerase Chain Reaction (qPCR) Assay*. U.S. Environmental Protection Agency, Washington, DC.

Wade, T.J., R.L. Calderon, K.P. Brenner, E. Sams, M. Beach, R. Haugland, L. Wymer, and A.P. Dufour. 2008. High sensitivity of children to swimming-associated bastrointestinal illness: Results using a rapid assay of recreational water quality. *Epidemiology* 19:375-383.

Wade, T.J., R.L. Calderon, E. Sams, M. Beach, K.P. Brenner, A.H. Williams, and A.P. Dufour. 2006. Rapidly measured indicators of recreational water quality are predictive of swimming-associated gastrointestional illness. *Environmental Health Perspectives* 114(1):24-28.

Whitman, R. 2010. *About Project S.A.F.E.* 2008. <http://www.glsc.usgs.gov/main.php?content=research_projectSAFE_about&title=Project%20S.A.F.E.0&menu=research_initiatives_projectSAFE>. Accessed July 2010.

Whitman, R., and M. Nevers. 2005. *Regional and local factors affecting patterns of E. coli distribution in southern Lake Michigan*. U.S. Geological Survey, Porter, IN.

Whitman, R.L., M.B. Nevers, G.C. Korinek, and M.N. Byappanahalli. 2004. Solar and temporal effects on *Escherichia coli* concentration at a lake Michigan swimming beach. *Applied and Environmental Microbiology* 70(7):4276-4285.

Wong, M., L. Kumar, T.M. Jenkins, I. Xagoraraki, M.S. Phanikumar, and J.B. Rose. 2009. Evaluation of public health risks at recreational beaches in Lake Michigan via detection of enteric viruses and a human-specific bacteriological marker. *Water Research* 43:1137-1149.

Wright, M.E. 2008. *Evaluation of enterococci, an indicator microbe, and the sources that impact the water quality of a subtropical non-point source recreational beach*. University of Miami, Civil Engineering, Coral Gables, FL.

Wright, M.E., H.M. Solo-Gabriele, S. Elmir, and L.E. Fleming. 2009. Microbial load from animal feces at a recreational beach. *Marine Pollution Bulletin* In press, corrected proof.

Yamahara, K.M., B.A. Layton, A.E. Santoro, and A.B. Boehm. 2007. Beach sands along the California coast are diffuse sources of fecal bacteria to coastal waters. *Environmental Science & Technology* 41(13):4515-4521.

Zhu, X. 2009. *Modeling microbial water quality at a non-point source subtropical beach*. University of Miami, Applied Marine Physics, Coral Gables, FL.

Appendix A. Additional Site Information

A.1 FRESHWATER BEACHES (GREAT LAKES)

A.1.1 South Shore Beach, Milwaukee, Wisconsin

South Shore Park in Milwaukee, Wisconsin, is within the residential area of Bayview on the west coast of Lake Michigan (Figure 4.2A). The park includes South Shore Beach, a public beach area (~ 0.01 square kilometer [km^2]) with 150 meters (m) of sandy shoreline. Its shore is low sloping, and the benthic zone is muddy and sandy. The beach is adjacent to the South Shore Yacht Club and a ~ 0.02 km^2 paved parking area, which drains into the lake. A 20-m-long rock embankment juts into the lake, separating the sandy beach area from a 185-m-long cobble/pebble beach area with a high sloping shore (South Shore Rocky Beach). Large flocks of roosting waterfowl and shore birds consistently occupy the beach. Ring-billed gulls are the predominant species, but Canada geese, Mallard ducks, and pigeons are also present (McLellan and Salmore 2003; Scopel et al. 2006). The entire beach and marina area is partially enclosed by a breakwall ~ 300 m offshore, which limits wave action, water circulation, and exchange with the outer harbor (Figure 4-2). Water depths within the breakwall are < 5 m, with depths < 2 m within 50 m of shore. The beach is ~ 4 km south of Milwaukee Harbor, where the Milwaukee Metropolitan Sewerage District Jones Island Water Reclamation Facility is. The Milwaukee, Menomonee, and Kinnickinnic rivers also discharge to Lake Michigan inside the Milwaukee Harbor breakwall.

Historically, South Shore has poor water quality, with 34 percent of samples collected from 2003 to 2009 exceeding water quality criteria standards (Table A.1). Potential sources of fecal contamination include combined sewer overflows (CSOs); urban/suburban and agricultural runoff from the Milwaukee River Basin; runoff from impervious surfaces including parking lots and the beach face; and gulls (Scopel et al. 2006). However, high *Escherichia coli* counts are not always attributable to rainfall and CSO events (McLellan et al. 2001; Scopel et al. 2006). A detailed spatial assessment found that poor beach water quality was mostly a local phenomenon, with contamination originating at the shoreline (McLellan and Salmore 2003). While CSO and sanitary sewer overflow events resulted in increased *E. coli* levels in the Milwaukee Estuary and Harbor (compared to rainfall without overflow), that source of sewage contamination did not influence the beach site (McLellan et al. 2007). The Russell Avenue outfall, about 0.5 km north of the beach inside the northwest corner of the marina breakwall is a closer source. The influence of discharge from the outfall is variable, with wind-driven and along-shore currents being important to water exchange in the beach area (Scopel et al. 2006). Antibiotic resistance profiles further confirmed that the major sources for *E. coli* found at the beach were shoreline runoff or stormwater, not CSO or sanitary sources (McLellan 2004). In addition, human-specific *Bacteroides* were not found at the beach or the Russell Avenue outfall (McLellan and Jensen 2005) (Figure A.1).

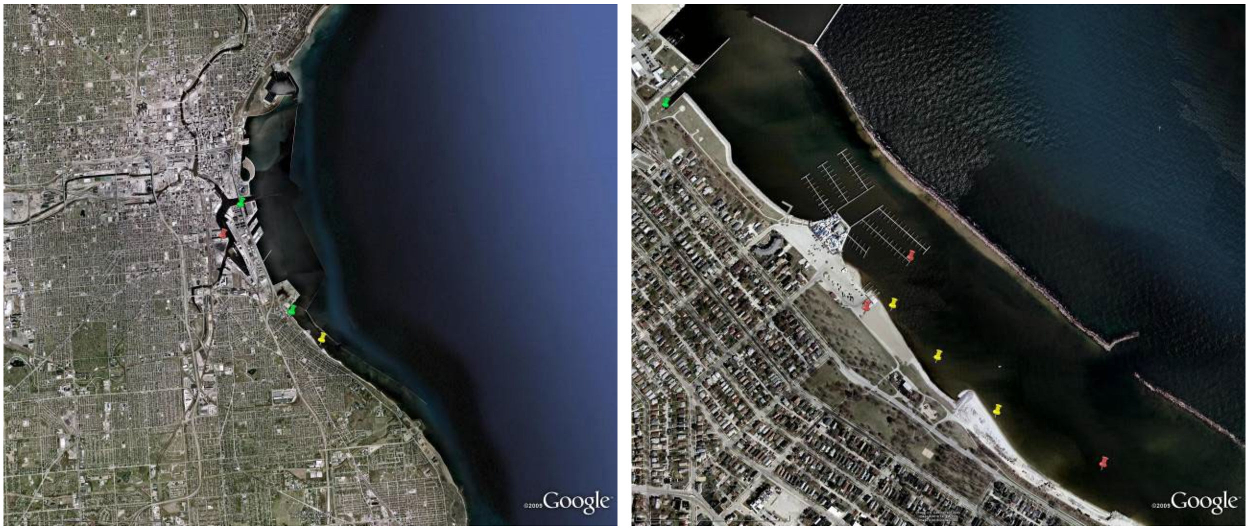


Figure A.1. Location of 2008 PREMIER study at South Shore Beach in Milwaukee, Wisconsin. Yellow pins indicate locations of beach and sampling transects and green pins mark possible sources of fecal contamination (i.e., WWTP on the Milwaukee River and Russell Avenue outfall). Details regarding field equipment (red pins) are given in Section 4.4.3. The beach is on the west shore of Lake Michigan (Figure 4.2A).

During the 14-week swimming season (Memorial Day through Labor Day), the Milwaukee County Health Department routinely monitors the beach for *E. coli*. Water quality (water temperature, specific conductance, pH, dissolved oxygen, chlorophyll, and turbidity) and meteorological conditions (air temperature, rainfall, wind speed, and wind direction) are continuously recorded by a sonde and weather station maintained by the University of Wisconsin-Milwaukee Great Lakes WATER Institute. Data are transmitted every hour (weather) or half hour (water quality) to a website via Ethernet communication. The sonde is deployed just northwest of the beach, off the east end of the South Shore Yacht Club dock, and the weather station is installed on a post at the beach, near the boat launch. Real-time and historical bacterial (since 2003), meteorological, and in-lake sonde data are publicly available through the Wisconsin Beach Health website (http://www.wibeaches.us).

Table A.1. Historical water quality monitoring details and criteria exceedances for freshwater beaches based on publicly available data from local monitoring agencies. Great Lakes beaches are sampled during the summer beach season, the start and end of which typically corresponds to the Memorial Day (last Monday in May) and Labor Day (first Monday in September) holiday weekends, respectively. The number of exceedances is expressed per total number of samples, with the corresponding percentage given in parentheses.

	South Shore	Rocky	West	Ogden Dunes	Washington Park	Silver	Huntington
Water Body	L. Michigan	L. Michigan	L. Michigan	L. Michigan	L. Michigan	L. Michigan	L. Erie
Location	Wisconsin	Wisconsin	Indiana	Indiana	Indiana	Michigan	Ohio
Possible Sources	Runoff	Runoff	Effluent	Effluent	Effluent	Effluent	Effluent
	Birds	Birds	Burns Ditch	Burns Ditch	Trail Creek	St. Joseph R.	Porter Creek
							Rocky River
Beach Length	150 m	200 m	2000 m	2000 m	1100 m	600 m	500 m
Data Source	MCHD	MCHD	IDEM	IDEM	IDEM	MDEQ	CCBH
Sampling Frequency	4-7 /wk	1-4 /wk	1 /wk	3-7 /wk	3-7 /wk	1-2 /wk	3-7 /wk

Water Quality Criteria for *E. coli* (CFU / 100 mL) and Number of Exceedances

Year	> 235	> 235	> 235	> 235	> 235	> 300	> 235
2000							12/51 (24)
2001						1/21 (5)	10/50 (20)
2002						0/13 (0)	11/52 (21)
2003	36/124 (29)					0/11 (0)	6/54 (11)
2004	63/120 (53)	18/114 (16)		2/104 (2)		0/20 (0)	7/54 (13)
2005	56/117 (48)	18/118 (15)		2/77 (3)	18/107 (17)	0/14 (0)	8/58 (14)
2006	27/87 (31)	4/15 (27)		1/45 (2)	13/49 (27)	0/15 (0)	26/96 (27)
2007	36/53 (68)	11/54 (20)		10/73 (14)	24/57 (42)	0/14 (0)	15/106 (14)
2008	27/54 (50)	6/28 (21)	0/10 (0)	1/48 (2)	7/48 (15)	0/15 (0)	14/106 (13)
2009	16/50 (32)	2/15 (13)	0/15 (0)	0/68 (0)	10/48 (21)	0/15 (0)	10/113 (9)
TOTAL	320/949 (34)	59/344 (17)	0/25 (0)	16/415 (4)	72/309 (23)	1/138 (1)	119/740 (16)

Data sources: South Shore Beach (http://www.wibeaches.us), Ogden Dunes Beach (https://extranet.idem.in.gov/beachguard), Washington Park Beach (https://extranet.idem.in.gov/beachguard), Silver Beach (http://www.deq.state.mi.us/beach), Huntington Beach, Ohio (2000–2005 data are from Francy et al. (2006) (Francy and Darner 2006); 2006–2009 data are from http://www.ohionowcast.info/nowcast_huntington.asp).

Previous modeling efforts at South Shore Beach include hydrodynamic and water quality models developed to describe the fate and transport of fecal coliform in Milwaukee Harbor and nearshore Lake Michigan. Calibration and validation of the models was accomplished using Milwaukee Metropolitan Sewerage District and Great Lakes WATER Institute field data. Modeling results indicated that the fecal coliform load from rivers and CSO/sanitary sewer overflow events had only slightly more than marginal impact on the beach site, with local sources (e.g., stormwater runoff and birds) being more important (MMSD 2005). South Shore Beach was also one of the 55 beaches included in a regional forecast model for southern Lake Michigan, from Milwaukee, Wisconsin, to Michigan City, Indiana (Whitman and Nevers 2005). Milwaukee is considering adopting a predictive model. In addition to issuing advisories on the basis of the monitoring of *E. coli* levels (where < 235 CFU/100 mL is acceptable, 235–1,000 CFU/100 mL results in a water quality advisory, and > 1,000 CFU/100 mL results in a closure advisory), a rainfall threshold of 2.5 cm in 24 hours is also used to predict poor water quality and issue advisories at South Shore.

A.1.2 West Beach, Porter, Indiana

West Beach is on the south shore of Lake Michigan (Figure 4.2A), ~ 40 km southeast of Chicago Harbor; it is part of the Indiana Dunes National Lakeshore Park in Porter, Indiana. The public beach has ~ 2 km of sandy shoreline and contributes to the total 10-km length of contiguous swimming beaches along this section of Lake Michigan. The shore has a gradual slope, and water depths are < 2.5 m within 250 m of shore. As an open, unprotected, lakefront beach, West Beach is subject to moderate wave exposure. The beach area is west of the Portage-Burns Waterway (Burns Ditch), a man-made channel that serves as the outfall point for the Little Calumet River to Lake Michigan. Discharge from Burns Ditch is directed west by a breakwall and then dispersed by lake currents and waves (Figure A.2B).

E. coli levels in Burns Ditch are typically higher than those in the lake and adjacent beach sites. Levels greater than 10,000 CFU/100 mL have been observed at the Burns Ditch outlet following storm events, and *E. coli* levels at the beach were found to be associated with rainfall (Olyphant et al. 2003). Outfall from Burns Ditch has the potential to affect water quality in the surf zone of the adjacent swimming beaches, particularly when prevailing north winds force the plume onshore (Nevers and Whitman 2005). Land use in the surrounding Little Calumet-Galien Basin is primarily agricultural with some forested, urban, and industrial areas. Seven wastewater treatment plants (WWTPs) are within the river basin and Lake Michigan watershed that provide disinfection with chlorine and ultraviolet (UV) radiation during the summer (Wade et al. 2008). The Portage, Indiana, WWTP, discharges UV-treated wastewater to the west branch of the Little Calumet River, ~ 5 km upstream of Lake Michigan. The Little Calumet River also receives inputs from a variety of nonpoint sources, including urban runoff, runoff from farms and livestock operations, failing septic systems, and wildlife (IDEM 2004; NIRPC 2005). Coliphage was detected in the ditch and at the beaches following heavy rainfall that resulted in CSO, providing evidence of sewage sources (Nevers and Whitman 2005). Water quality is good with water quality criteria exceeded in 4 percent of the samples collected from 2004 to 2009 at Ogden Dunes, and with no exceedances at West Beach in 2008 and 2009 (Table A.1).

During the Memorial Day to Labor Day summer beach season, the beaches are monitored routinely for *E. coli*, typically by the Indiana Dunes National Lakeshore. Real-time and historical bacterial data (since 2008) are publicly available through the Indiana Department of Environmental Management's BeachGuard website (https://extranet.idem.in.gov/beachguard). Real-time and historical discharge, gage height, and stream velocity data are available from the U.S. Geological Survey (USGS) Burns Ditch gauging station at Portage, Indiana (04095090), which is ~ 1.2 km upstream of Lake Michigan (http://waterdata.usgs.gov/nwis/uv?04095090).

The West Beach and Burns Ditch sites have been the subject of several predictive modeling studies. Multiple regression analysis found that a combination of water quality variables could account for the observed variability in *E. coli* at Burns Ditch outlet during storm events (Olyphant et al. 2003). Further efforts led to developing regression models that predict beach *E. coli* levels from wave and precipitation data, as well as water quality parameters. Because of the role of wind direction on the transport of Burns Ditch discharge to the beaches (Nevers and Whitman 2005), model performance was improved by separately modeling days with onshore and offshore winds. Regional forecast models also were developed for southern Lake Michigan, which include West Beach (Whitman and Nevers 2005; Nevers and Whitman 2008). In addition to *E. coli*, more recent predictive models for beaches impacted by Burns Ditch included culturable and quantitative polymerase chain reaction (qPCR)-based enterococci as dependent variables. Differences in the models developed for each indicator (i.e., IVs identified for culturable compared to qPCR-based enterococci) provide evidence of the different processes that determine their fate (Byappanahalli et al. 2010; Telech et al. 2009). Beach monitoring practices rely on culturable *E. coli* where levels > 235 CFU/100 mL result in a contamination advisory or closure, as well as on a rainfall threshold. In addition, the Project S.A.F.E. (Swimming Advisory Forecast Estimate) predictive model was implemented as a USGS pilot study (Whitman 2008) and is used in conjunction with the NOAA/Great Lakes Environmental Research Laboratory Indiana Dunes Nowcast hydrodynamic model (Schwab et al. 2010) to provide real time *E. coli* estimates at West Beach.

A.1.3 Washington Park Beach, Michigan City, Indiana

Washington Park Beach is in Michigan City, Indiana, on the south shore of Lake Michigan (Figure 4.2A). The public beach area is ~ 1.1 km long and is immediately east of a breakwall that directs discharge from Trail Creek to the west (Figure A.2C).

Trail Creek is a source of *E. coli* to the lake with the potential to affect water quality at nearby beaches (Nevers et al. 2007). The Trail Creek watershed drains urban, agricultural, and residential areas, with a number of human and animal nonpoint sources including agricultural field drainage and runoff, cattle/steer grazing, failing septic systems, illicit connections, and urban stormwater runoff (Triad Engineering Incorporated 2003). In addition, the Michigan City Sanitary District WWTP (~ 3 km upstream of Lake Michigan), which applies chlorine disinfection during the summer months, is a major discharger to the creek (Wade et al. 2008), although plant improvements have practically eliminated CSO events (Nevers et al. 2007). Water quality at Washington Park Beach is poor, with the recreational water quality criteria for *E. coli* exceeded in 23 percent of the samples collected from 2005 to 2009 (Table A.1). In addition, pathogens (adenoviruses and enteroviruses) and a marker for human sewage were detected in samples collected from the swimming area during summer 2004 (Wong et al. 2009).

The Michigan City Parks Department or LaPorte County Health Department monitor the beach regularly for *E. coli* during the 14-week beach season (Memorial Day through Labor Day). Real-time and historical bacterial data (since 2005) are publicly available through the Indiana Department of Environmental Management's BeachGuard website (https://extranet.idem.in.gov/beachguard). In addition, daily water quality results that lead to advisories or closures are posted on the Michigan City Parks & Recreation Current Beach Conditions website (http://www.emichigancity.com/cityhall/departments/parks/beach.htm). Real-time and historical discharge, gage height, and stream velocity data are available from the USGS Trail Creek gauging station at Michigan City Harbor, Indiana (04095380), which is < 1 km upstream of Lake Michigan (http://waterdata.usgs.gov/nwis/uv?04095380).

Predictive modeling of *E. coli* has not been reported at Washington Park Beach, but work was done at another Trail Creek-influenced beach, Mount Baldy Beach, ~ 2.5 km west of the mouth of Trail Creek (Nevers et al. 2007). The Mount Baldy site was also incorporated in a southern Lake Michigan regional forecast model (Whitman and Nevers 2005; Nevers and Whitman 2008). The observed variation in *E. coli* levels at Mount Baldy was further explained, based on loadings from Trail Creek and nearby Kintzele Ditch, using a process-based hydrodynamic model (Liu et al. 2006; Thupaki et al. 2009). Given the proximity of Washington Park to Mount Baldy, a similar type of model will likely apply to both beaches. In work specific to the Washington Park site, independent and predictive models were used to investigate the relationship between enterococci (culturable and qPCR-based) and environmental conditions (Telech et al. 2009). Water quality monitoring at the beach is based on culturable *E. coli* where levels > 235 CFU/100 mL result in a contamination advisory or closure when counts ≥ 1,000 CFU/100 mL.

A.1.4 Silver Beach, St. Joseph, Michigan

Silver Beach is in St. Joseph, Michigan, on the southern end of the eastern shore of Lake Michigan, ~ 55 km northeast of Washington Park Beach (Figure 4.2A). The ~ 0.6-km-long sandy beach area is just south of where the St. Joseph River flows into Lake Michigan. At its mouth, the river is lined by two parallel navigational piers that extend into the lake, guiding riverine discharge ~ 0.5 km out, roughly perpendicular to the shoreline (Figure A.2D).

Possible sources of fecal contamination to the St. Joseph River watershed, which encompasses both urban and agricultural areas, include seven WWTPs (that use chlorine disinfection), four of which are on the river (Wade et al. 2008), CSOs, stormwater discharges, agricultural inputs, and illicit discharges (MDEQ 2003). The St. Joseph-Benton Harbor WWTP is on the river, ~ 2.7 km upstream of Lake Michigan. Water quality at Silver Beach is good, with only 1 percent of samples collected from 2001 to 2009 exceeding local water quality criteria standards for *E. coli* (Table A.1). As with Washington Park Beach, pathogens (adenoviruses, enteroviruses, and rotaviruses) and a human sewage marker were detected in samples collected from the Silver Beach swimming area during summer 2004 (Wong et al. 2009).

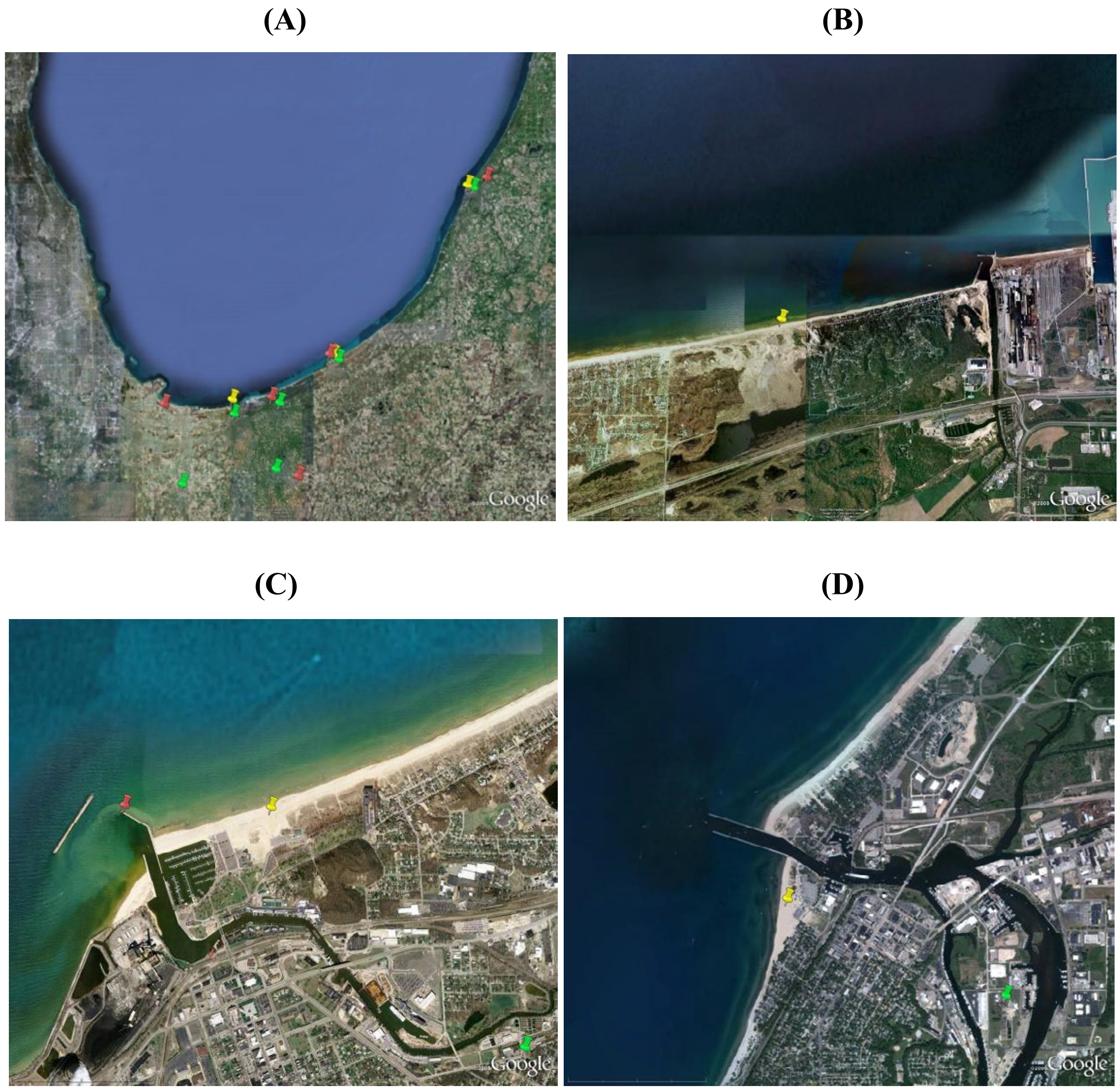


Figure A.2. Locations of (A) Lake Michigan NEEAR studies including (B) 2003 study at West Beach in Porter, Indiana, (C) 2004 study at Washington Park Beach in Michigan City, and (D) 2004 study at Silver Beach in St. Joseph, Michigan. Yellow pins indicate locations of beaches and green pins mark possible fecal contamination sources (i.e., WWTPs). Red pins show nearby USGS, NOAA, and airport weather stations. The beaches are on southern Lake Michigan (Figure 4.2A).

The Berrien County Environmental Health Department monitors the public beach weekly for *E. coli* during the summer season (from Memorial Day to Labor Day).

Real-time and historical bacterial data (since 2001) are publicly available through the Michigan Department of Environmental Quality BeachGuard website (http://www.deq.state.mi.us/beach).

Real-time and historical discharge and gage height data are available from the USGS St. Joseph River gauging station at Elkhart, Indiana (04101000), which is ~ 100 km upstream of Lake Michigan (http://waterdata.usgs.gov/nwis/uv?04101000).

Regression modeling was employed at Silver Beach to investigate relationships between culturable and qPCR-based enterococci and various environmental conditions in an effort to identify useful predictors of water quality (Telech et al. 2009). Routine water quality monitoring at the beach is based on a culturable *E. coli* criteria of 300 CFU/100 mL and the use of a *rainfall plus 48-hour health advisory*.

A.1.5 Huntington Beach, Bay Village, Ohio

Huntington Beach is part of Huntington Reservation of Cleveland Metroparks, which is in the City of Bay Village, Ohio (a western suburb of Cleveland), on the southern shore of western Lake Erie (Figure 4.2A). The swimming area is just west of Porter Creek, and the beach is ~ 8 km west of the Rocky River mouth. The ~ 0.5-km-long sandy beach area is broken into segments by a series of rock jetties (< 100 m long) that run perpendicular to the shoreline (Figure A.3). The breakwalls limit water circulation in the swimming area.

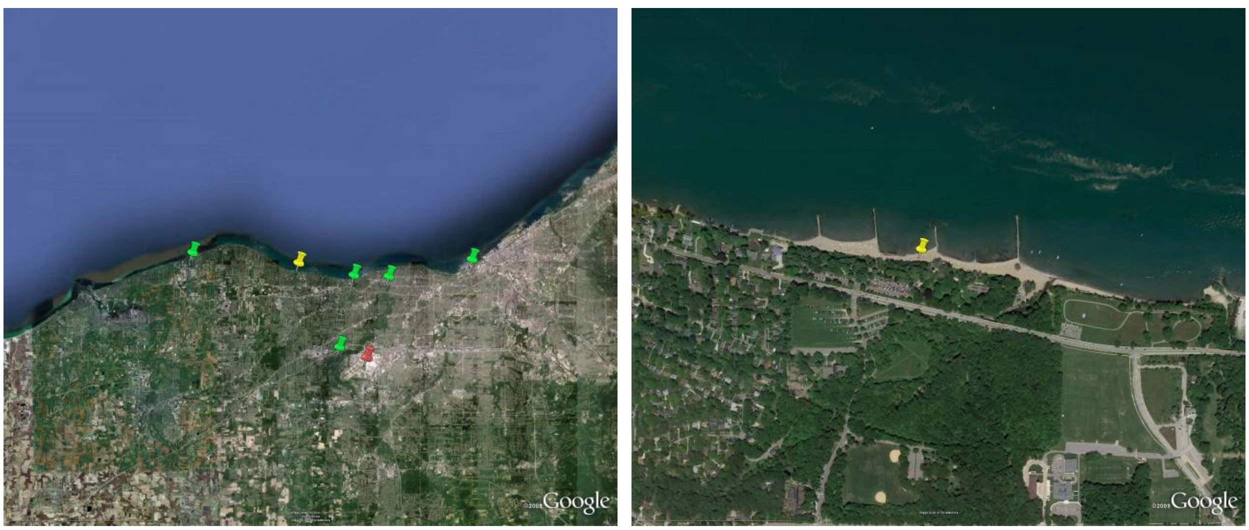


Figure A.3. Location of 2003 NEEAR study at Huntington Beach in Bay Village, Ohio. Yellow pin indicates location of beach and green pins mark possible sources of fecal contamination (i.e., WWTPs). Red pin shows nearby airport weather station. The beach is on Lake Erie (Figure 4.2A).

Huntington Beach is in the Black-Rocky watershed, within which are 10 sewage discharge locations that could affect the beach. Three of those flow directly into Lake Erie: Avon Lake WWTP, ~ 11 km west of the beach; Rocky River WWTP, ~ 6 km east of the beach; and Westerly WWTP, ~ 18 km east of the beach. The others are discharged to the Rocky River or its tributaries, the closest being Lakewood WWTP (< 3 km from the mouth of the Rocky River). The majority of the treatment plants use UV or chlorine disinfection during the summer (Wade et al. 2008; Wade et al. 2006). Porter Creek, on the east end of the beach, was identified as a likely contributor to high *E. coli* counts at beach (Huang and Sigler 2006). In addition, two outfalls discharge storm runoff from the parking lot to the beach (Francy et al. 2003). Water quality at Huntington Beach is poor, with 16 percent of samples collected from 2000 to 2009 exceeding water quality criteria standards for *E. coli* (Table A.1).

During the 14-week summer beach season, the Cuyahoga County Board of Health and Cuyahoga County Sanitary Engineers Laboratory, monitor Huntington Beach water quality (i.e., *E. coli*). In addition, a number of ancillary environmental parameters (i.e., water temperature, turbidity, and wave height) are also measured at the time of sample collection. The data (real-time and historical since 2006) are publicly available through the Ohio Nowcast website, along with the associated water quality predictions and advisories (http://www.ohionowcast.info/nowcast_huntington.asp). Real-time and historical discharge and gage height data are available from the USGS Rocky River gaging station near Berea, Ohio (04201500), which is ~ 20 km upstream of Lake Erie (http://waterdata.usgs.gov/nwis/uv?04201500).

Because of ongoing efforts of the USGS Water Science Center and its partners in daily routine sampling and data collection, an extensive, multiyear data set is available for Huntington Beach. The focus of the work was to develop a multiple linear regression (MLR) model that continues to be tested and refined over time by incorporating additional years of data.

Water quality predictions resulting from the model have been available to the public since 2006 as a real-time, Internet-based nowcasting system (Francy et al. 2003; Francy et al. 2006b; Francy and Darner 2007). The USGS data set (for 2000 to 2004) was used to examine statistical issues related to MLR modeling (Ge and Frick 2007). In addition, publicly available data from 2006 demonstrate the efficacy of VB 1.0 in developing statistical models for predicting *E. coli* at Huntington Beach, while exploring the usefulness of dynamic models and time-frequency analysis (Frick et al. 2008; Ge and Frick 2009). Additional pilot studies focusing on developing VB 1.0, using Huntington Beach as an example study site, are discussed in more detail in Section 2.1. Relationships between culturable and qPCR-based enterococci and various environmental parameters were investigated at Huntington Beach, using regression modeling to identify potentially useful water quality predictors (Telech et al. 2009). The Ohio Nowcast predictive model is used to supplement *E. coli* data (235 CFU/100 mL criteria) when making decisions about swimming advisories.

A.2 MARINE BEACHES

A.2.1 Goddard Beach, West Warwick, Rhode Island

Goddard Beach is in Goddard State Memorial Park in West Warwick, Rhode Island (Figure 4.2B). The beach stretches ~ 1.2 km along the southern shoreline of Greenwich Bay, just east of the mouth of Greenwich Cove (Figure A.4). Greenwich Bay is a small (13 km^2) estuary on the western side of Narragansett Bay, which is ~ 6.5 km southwest of the mouth of the Providence River. The Maskerchugg River discharges to the head of Greenwich Cove, while a number of smaller brooks and creeks flow into Greenwich Bay, either directly or through one of the bay's four other coves.

Figure A.4. Location of 2007 NEEAR study at Goddard Beach in West Warwick, Rhode Island. Yellow pin indicates location of beach, and green pin marks possible source of fecal contamination (i.e., WWTP). Red pin shows nearby airport weather station. The beach is on the western side of Narragansett Bay (Figure 4.2B).

Greenwich Cove and the south inner Greenwich Bay (just north of the beach) are designated for shellfish harvesting and recreational use. As a result, the surrounding waters are held to stricter water quality criteria than the swimming standard at the beach. The Greenwich Bay watershed includes parts of Warwick, East Greenwich, and West Warwick in central Rhode Island where land use is predominantly urban and residential. Potential sources of fecal contamination include tributary streams that discharge to the cove and bay, such as the Maskerchugg River, and direct stormwater discharges. Those sources are along the west coast of the cove and include the East Greenwich Wastewater Treatment Facility, which discharges treated effluent to the middle of the channel about halfway down the cove, < 2 km from the beach. However, dye studies showed sufficient dilution of fecal coliform that suggest that Greenwich Cove and Bay are not significantly affected by the effluent. Faulty septic systems, waterfowl that gather at the beach, wildlife, and domestic pets are other potential sources of contamination at Goddard Beach. In addition to the swimming beach, the state-owned Goddard Park, which makes up half of the 1.6 km^2 Potowomut subwatershed, includes golf courses and forested land; the rest of the subwatershed is residential. No sewers and few freshwater sources drain the area to Greenwich Bay. In the park, culverts direct stormwater from the parking lot onto the beach (RIDEM 2005).

During the 14-week summer season (Memorial Day through Labor Day), the Rhode Island Department of Health routinely monitors Goddard Beach for enterococci. Real-time and historical data (since 2002) are publicly available through the Rhode Island Department of Health Beach Monitoring Program website (http://www.ribeaches.org/beach.cfm?beachID=RI810609). Water quality at Goddard Beach is good, with 9 percent of samples collected from 2002 to 2009 exceeding water quality criteria standards for enterococci (Table A.2). To our knowledge, formal predictive models have not been previously employed at Goddard Beach. Beach management decisions are based on enterococci monitoring (using the 104 CFU/100 mL criteria), and consideration of water quality history and other environmental conditions such as rain (RIDEM 2005).

Table A.2. Historical water quality monitoring details and criteria exceedances for temperate and subtropical marine beaches. Number of exceedances is expressed per total number of samples, with the corresponding percentage given in parentheses.

	Goddard	Surfside	Edgewater	Fairhope	Hobie
Water Body	Greenwich Bay Estuary	Atlantic Ocean	Mississippi Sound	Mobile Bay	Biscayne Bay
Location	Rhode Island	South Carolina	Mississipi	Alabama	Florida
Possible Sources	Effluent	Runoff	Effluent	Effluent	Runoff
	Maskerchugg River	5th Ave. N. Swash			Dogs, Birds
Climate	temperate	sub-tropical	sub-tropical	sub-tropical	sub-tropical
Beach Length	1200 m	3400 m		600 m	1600 m
Data Source	RIDH	SCDHEC	MDEQ	ADEM	FDOH
Sampling Frequency	2-4 /wk	1 /wk	1-4 /wk	2 /wk	2-4 /mo
Sampling Period	late May to early September	mid-May to mid-October	year round	year round	year round

Water Quality Criteria for Enterococcus (CFU / 100 mL) and Number of Exceedances

Year	> 104	> 104	> 104	> 104	> 104
2000					7/24 (29)
2001					1/27 (4)
2002	0/28 (0)				3/40 (8)
2003	8/35 (23)				4/57 (7)
2004	1/41 (2)		12/70 (17)		3/55 (5)
2005	3/49 (6)	14/20 (70)	18/64 (28)		6/58 (10)
2006	7/50 (14)	7/35 (20)	8/89 (9)	13/56 (23)	4/56 (7)
2007	0/42 (0)	4/31 (13)	16/91 (18)	10/58 (17)	4/56 (7)
2008	3/35 (0)		11/61 (18)	8/60 (13)	2/55 (4)
2009	8/52 (15)		3/52 (6)	10/62 (16)	5/61 (8)
TOTAL	30/332 (9)	25/86 (29)	68/427 (16)	41/236 (17)	31/484 (6)

Data sources: Goddard Beach (http://www.ribeaches.org/beach.cfm?beachID=RI810609), Surfside Beach (station 31A, data provided by SCHEC), Edgewater Beach (station 11A, http://www.coms.usm.edu/msbeach/harbmon.cgi), Fairhope Beach (http://www.adem.state.al.us/programs/coastal/beachMonitoring.cnt), Hobie Beach (1999–2000 is taken from Solo-Gabriele et al. (2002); 2001–2009 data are from http://esetappsdoh.doh.state.fl.us/irm00beachwater).

A.2.2 Surfside Beach, Surfside Beach, South Carolina

Surfside Beach is in the town of Surfside Beach, South Carolina, just south of Myrtle Beach in Horry County (Figure 4.2B). The beach is 3.4-km long and is only a small fraction of the South Carolina coastline that makes up more than 100 km of uninterrupted, open beachfront known as the Grand Strand. The Grand Strand is the northernmost part of the South Carolina Coastal Plain. Several lakes were created in Surfside Beach to serve as retention ponds during storm events. The associated watersheds were designed and constructed for stormwater management purposes. The Myrtle Basin designed in 2005–2006, with construction starting in 2007, was followed by lake dredging. Two small lakes (each with an area of ~ 3000 m^2) are between North Myrtle Drive, North Dogwood Drive, 2nd Avenue North, and 5th Avenue North. They are connected by a 120-m-long channel, with another 200-m-long channel running between them and the beach. The swash at 5th Avenue North receives runoff from this area, which is then discharged directly to the beach (Figure A.5). A sign is permanently posted at the swash, stating that swimming is not recommended within 30.5 m because stormwater runoff can result in elevated levels of bacteria. The Public Works Department digs out the swash outlets on the beach as needed (as often as three times per week) to ensure proper water flow.

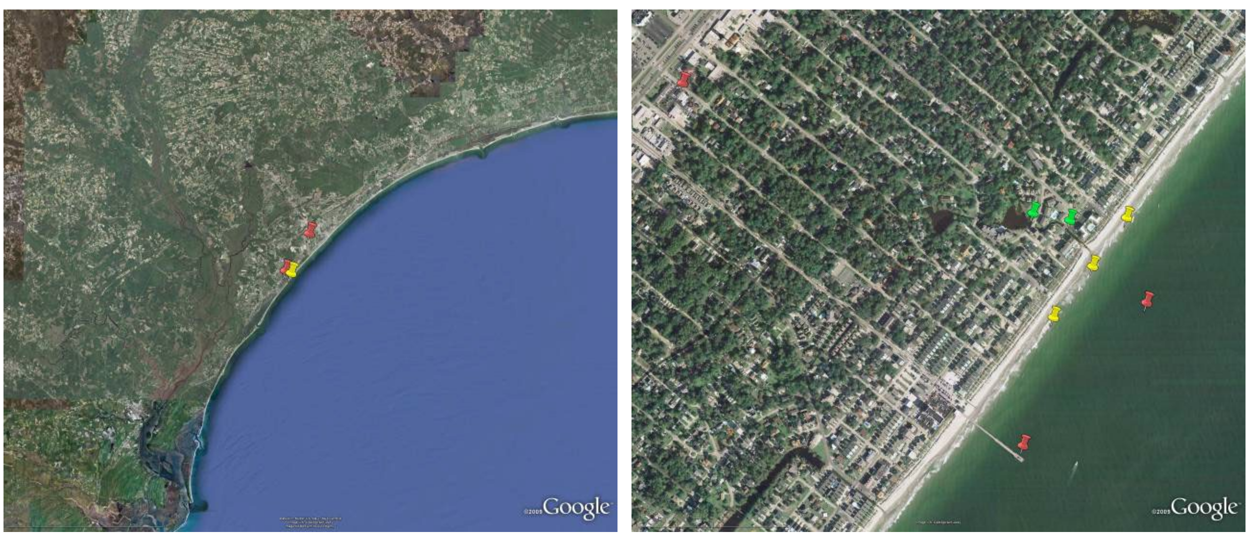


Figure A.5. Location of 2009 NEEAR/ PREMIER study at Surfside Beach in Surfside Beach, South Carolina. Yellow pins indicate locations of beach and sampling transects and green pins mark possible sources of fecal contamination (i.e., swash channel at 5th Avenue North and retention pond). Details regarding field equipment (red pins) are given in Section 4.4.3. The beach is on the Grand Strand on the southeastern coast of the United States (Figure 4.2B).

The Grand Strand Water and Sewer Authority Schwartz WWTP is ~ 8 km northwest of the beach. Surfside Beach is not affected by effluent from the WWTP because it is discharged to the Intracoastal Waterway (~ 3.4 km northwest). The outlet to the Atlantic Ocean is 50 km from the beach. Several campgrounds along the coast north of the beach are within 4 km of Surfside Beach. In addition to the swash at 5th Avenue North, which is at the section of the beach considered here, additional swashes are up and down the coast. Those closest to the beach area include the 11th Avenue North Dogwood Swash (0.6 km north), the Surfside Drive outfall (0.5 km south by the pier), and the 3rd Avenue South Swash (0.9 km south of the 5th Avenue North Swash). Given the lack of known point sources, the most likely source of fecal contamination to

Surfside Beach is runoff from the surrounding urban areas. Wildlife can also contribute to observed fecal indicator levels in the swash as birds (i.e., geese, ducks, gulls) frequent the lake and surrounding areas.

The South Carolina Department of Health and Environmental Control (SCDHEC) monitors Surfside Beach for enterococci regularly, from May 15 to October (during the 20-week summer season). Data from the current year are publicly available through the SCDHEC Beach Monitoring Data website (http://www.scdhec.gov/environment/water/beachmondata.aspx). In addition, a new GIS-based Integrated Monitoring and Assessment Program web application is being developed to provide water quality data and beach status (http://gisweb00.dhec.sc.gov/ImapPublic/beach.html). Water quality at the 5th Avenue North Swash of Surfside Beach is poor, with 29 percent of samples collected from 2005 to 2007 exceeding water quality criteria standards for enterococci (Table A.2). However, water quality has improved over time (from 70 percent in 2005 to 13 percent in 2007) because of improvements made to the stormwater management watershed. Beach monitoring practices at Surfside Beach rely on culturable enterococci, where levels > 500 CFU/100 mL or repeated measurements > 104 CFU/100 mL lead to advisories. In addition, preemptive rainfall advisories were issued on the basis of a rainfall threshold. More recently, a rain model for advisory predictions has been in development.

A.2.3 Edgewater Beach, Biloxi, Mississippi

Edgewater Beach is in Biloxi, Mississippi, on the Mississippi Sound along the Gulf of Mexico (Figure 4.2B). The Mississippi Sound is separated from the Gulf by a chain of barrier islands, the closest of which is Ship Island, ~ 20 km south of the beach. This lagoon is < 5 m deep in most areas and runs 124 km along the southern coasts of Mississippi and Alabama. Beaches in the area generally are subject to low energy wave conditions. The beach shore is gently sloped (i.e., 5 to 10 degrees) and consists of well- to very-well-sorted medium sand. Major rain events form runoff channels (Otvos 1999). The beach is about 16 km west of the mouth of Biloxi Bay. There is a narrow sand island (Deer Island) to the west of the Biloxi Bay and a system of breakwalls direct water from the bay westward along the shoreline toward the beach. The Back Bay of Biloxi runs parallel to the coast as a land-locked, westward continuation of Biloxi Bay (Figure A.6).

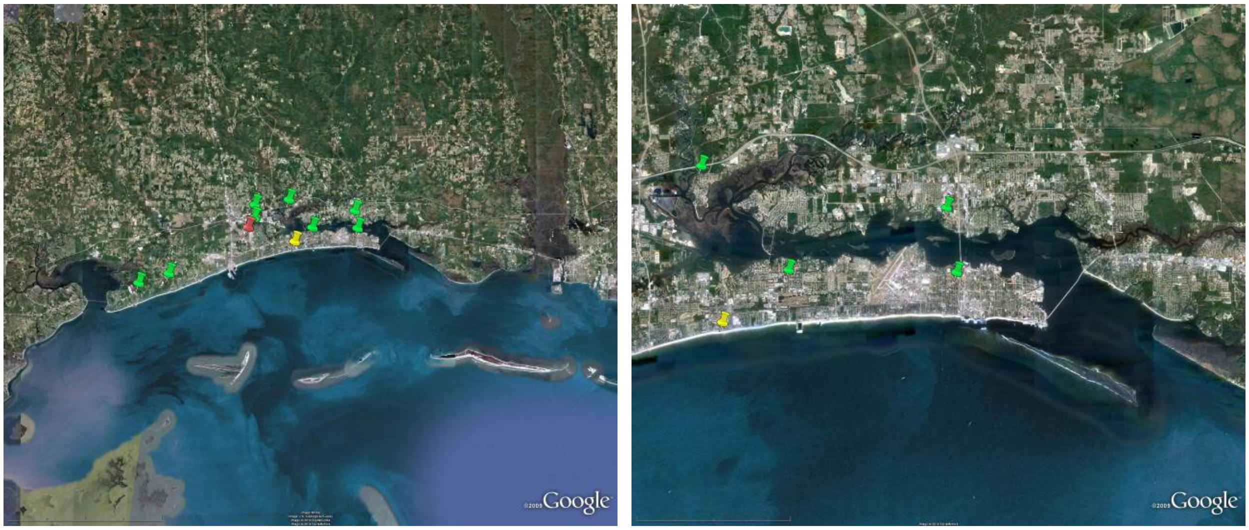


Figure A.6. Location of 2005 NEEAR study at Edgewater Beach in Biloxi, Mississippi. Yellow pin indicates location of beach and green pins mark possible sources of fecal contamination (i.e., WWTPs). Red pin shows nearby airport weather station. The beach is on the Mississippi Sound along the Gulf of Mexico (Figure 4.2B).

The Biloxi Bay watershed has a drainage area of ~ 1,600 km^2. Land use in the watershed consists predominantly of forested and wetland areas. Only a small percentage (<6 percent) is urban, and those areas are concentrated around the Back Bay of Biloxi and Biloxi Bay. Biloxi Bay is designated for shellfishing and, therefore, is subject to stricter water quality criteria than the recreational beach standard. The Coastal Streams Basin is made up of several independent watersheds including a number of rivers, streams, creeks, bayous, and bays. Seven Harrison County Utility Authority WWTPs are within 10 km of the beach. Treated effluent from the facilities is discharged to the Biloxi River, Bernard Bayou (Gulfport Lake), and the Back Bay of Biloxi. The Keegan Bayou East Biloxi Wastewater Treatment Facility is closest to Biloxi Bay (the bayou is ~ 6 km from the mouth of Biloxi Bay). Two additional treatment facilities are closer to the coast ≥ 20 km west of the beach. The Gulfport North Wastewater Treatment Facility, which discharges to Gulfport Lake in the Bernard Bayou, employs UV disinfection, while all the others use chlorine. Possible nonpoint sources of fecal contamination include failing septic systems, wildlife, applied manure, grazing animals, and urban development (MDEQ 2002).

The beach is monitored year-round for enterococci and fecal coliform by the Mississippi Department of Environmental Quality (MDEQ) and the University of Southern Mississippi's Gulf Coast Research Laboratory. Real-time and historical data (since 2004) are publicly available from the MDEQ Mississippi Beach Monitoring Program website (http://www.coms.usm.edu/msbeach/harbmon.cgi). Water quality at Edgewater Beach (near Eisenhower Drive) is poor, with 16 percent of samples collected from 2004 to 2009 exceeding water quality criteria standards for enterococci (Table A.2). Decisions regarding water quality are based on culturable enterococci where levels > 104 CFU/100 mL result in beach advisories.

A.2.4 Fairhope Beach, Fairhope, Alabama

Fairhope Municipal Beach is in Fairhope, Alabama, ~ 22 km southeast of Mobile, on the eastern shore of Mobile Bay (Figure 4.2B). The Mobile Bay Estuary is an inlet of the Gulf of Mexico, with an average depth of 3 m. The small mouth of the bay is shaped on the east by the Fort Morgan Peninsula and the Dauphin Island barrier island on the west. The sandy beach area at Fairhope Beach is ~ 0.6 km long. Several piers are perpendicular to the shoreline on the southwest end of the beach (Figure A.7).

The > 115,500 km^2 Mobile Bay watershed includes seven river systems, with the Mobile River serving as the primary freshwater input to the bay. The seawater portion of Mobile Bay is designated for shellfish harvesting and subject to more stringent water quality standards than the beach. The Fairhope WWTP, < 1 km northeast of the beach, serves as a potential source of continuous fecal contamination. The facility employs UV disinfection as the final treatment step. Additional sources include sanitary sewer overflows, failing septic tanks, and urban runoff (ADEM 2010).

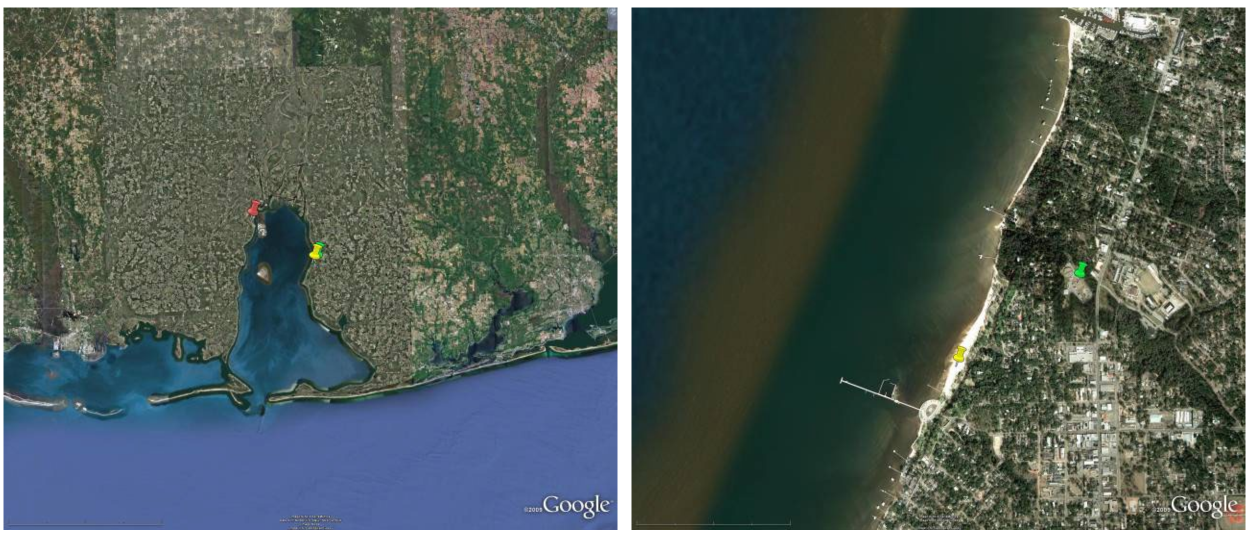


Figure A.7. Location of 2007 NEEAR study at Fairhope Beach in Fairhope, Alabama. Yellow pin indicates location of beach and green pins mark possible sources of fecal contamination (i.e., WWTPs). Red pin shows nearby airport weather station. The beach is on the eastern shore of Mobile Bay, an inlet of the Gulf of Mexico (Figure 4.2B).

Enterococci levels are monitored year-round at Fairhope Beach by the Alabama Department of Environmental Management (ADEM) and the Alabama Department of Public Health (ADPH) with samples collected by the Baldwin County Health Department. In addition, a number of ancillary environmental parameters (i.e., water temperature, dissolved oxygen, pH, conductivity, salinity, turbidity, and the occurrence of rain) are sometimes measured at the time of sample collection. The data (real-time and historical since 2006) are publicly available at the ADEM/ADPH Coastal Alabama Beach Monitoring Program website (http://www.adem.state.al.us/programs/coastal/beachMonitoring.cnt). Water quality at Fairhope Beach is poor, with 17 percent of samples collected from 2006 to 2009 exceeding water quality criteria standards for enterococci (Table A.2). Decisions regarding water quality are based on culturable enterococci, where repeated measurement of levels > 104 CFU/100 mL result in a public health advisory.

A.2.5 Hobie Beach, Miami, Florida

Hobie Beach in Miami, Florida is on Virginia Key in the southern part of Biscayne Bay, off the east coast of mainland Miami (Figure 4.2B). Biscayne Bay is a subtropical estuary that receives freshwater inputs from the Miami River and small creeks, as well as from a network of drainage canals. The Miami River is ~ 4 km northwest of the beach, and its freshwater can influence the beach under certain conditions. Hobie Beach is ~ 1.6 km long and runs along the south side of Rickenbacker Causeway, between the William Powell Bridge and Miami Seaquarium (Figure A.8). Hobie is also known as *Dog Beach* because it is the only Miami-Dade County beach where pets are allowed; the ratio of dogs to humans at the beach is on the order of 1:7. The beach is narrow, with a distance of 5 m and 12 m between the water line and the outer edge of the sand line during high and low tide, respectively. Vehicles park right along the sand line. The benthic zone is silty and muddy, and the shoreline typically is covered with seaweed. The slope is relatively shallow and natural runoff channels form following heavy rainfall events. Hobie Beach is shallow with water depths < 2 m at the buoy line, ~ 130 m from the shoreline. Due to its location in a shallow cove, water circulation at the beach is poor and movement near shore is controlled by tidal action (with an average tidal height fluctuation of 58 cm) rather than waves (Shibata et al. 2004; Wright 2008). During ebb tide (the period in between high and low tide), water flows out of Biscayne Bay to the Atlantic through the Norris Cut and Bear Cut inlets. During flooding tide, water enters the bay. Flow is parallel to the shoreline with velocities of 0.2 m/s and < 0.1 m/s during ebbing and flooding tide, respectively. Easterly winds prevail with a weak southerly component in the summer, accompanied by a strong local sea-breeze and thunderstorms (Zhu 2009).

Figure A.8. Location of 2008 PREMIER study at Hobie Beach in Miami, Florida. Yellow pins indicate locations of beach and sampling transects and green pins mark possible sources of fecal contamination (i.e., WWTP). Details regarding field equipment (red pins) are given in Section 4.4.3. The beach is on Virginia Key in the southern part of Biscayne Bay (Figure 4.2B).

Hobie Beach was found to have poor water quality during an EPA/Florida Department of Environmental Health (FDOH) Beach Monitoring Study, when 29 percent of samples collected from July 1999 to June 2000 exceeded the enterococci water quality criteria (Solo-Gabriele et al. 2002). There are no known point sources to the beach, but the Central District WWTP, 2 km northeast of beach, is a potential source of fecal contamination. Treated, chlorine-disinfected effluent is discharged to the Atlantic Ocean through an outfall ~ 4 km east of the plant. No storm drains are at the beach, but other suspected sources of enterococci include runoff during heavy rainfall events and animals, particularly dogs (Wright et al. 2009). Spatially intensive sampling efforts, including surveys of beach sand and water, identified the shoreline as a source of fecal indicators. Bacterial counts were higher at the shoreline than offshore. Higher levels were observed at the east end of the beach, likely reflecting reduced flushing due to the adjacent peninsula. Indicator densities were also higher at high tide compared to low tide, suggesting a correlation with tidal stage and pointing to the inter-tidal zone (i.e., the sand area that is wetted and dried between high and low tide) as the source of contamination (Shibata et al. 2004; Solo-Gabriele et al. 2002). A water quality model developed to evaluate sources of nonpoint pollution found runoff to be the most important source of enterococci, followed by dogs, sand, birds, and bathers (Elmir 2006). Research efforts also have focused on the associations between indicator microbes, pathogens, and environmental conditions (Abdelzaher et al. 2010).

The Miami-Date County Health Department routinely monitors Hobie Beach year-round for enterococci and fecal coliform. Real-time and historical data (since 2000) are publicly available through the FDOH Beaches website (http://esetappsdoh.doh.state.fl.us/irm00beachwater). Water quality at the beach is good, with 6 percent of samples collected from 2001 to 2009 exceeding water quality criteria standards for enterococci (Table A.2). Decisions regarding water quality are based on enterococci and fecal coliform levels, with counts > 104 CFU/100 mL resulting in a poor enterococci rating. A process-based predictive numerical hydrodynamic water quality model for Hobie Beach is being developed as a tool to improve assessing nonpoint source fecal contamination (Zhu 2009).

A.2.6 La Monseratte Beach, Luquillo, Puerto Rico

La Monserrate public beach, better known as Luquillo Beach, is on the northeast coast of Puerto Rico in the town of Luquillo (Figure 4.2B). The public beach is ~ 400 m long and runs along the eastern side of the bay (Figure A.9). The mouth of the Mameyes River is ~ 2.2 km west of the main swimming area at Luquillo. During periods of low flow, the river mouth can be closed off completely from the ocean. Two streams fed by stormwater runoff are potential sources of fecal contamination to the beach; one is just west of the beach and the other is east of the beach, along the northern shoreline.

The Puerto Rico Environmental Quality Board (PREQB) monitors enterococci and fecal coliform levels at Luquillo Beach twice a month year-round. Some data for the current year are publicly available at the PREQB water quality website (http://www.prtc.net/~jcaagua). Water quality is good at Luquillo Beach, with 8 percent of samples collected from 2006 to 2008 exceeding the water quality criteria standard of 104 CFU/100 mL for enterococci (Table A.3). Decisions regarding water quality are based on both culturable enterococci and fecal coliforms, where 35 CFU/100 mL is used as the criteria for enterococci. Puerto Rico's standard is more stringent than that recommended by EPA and its use results in a total of 18 percent exceedances for 2006 to 2008.

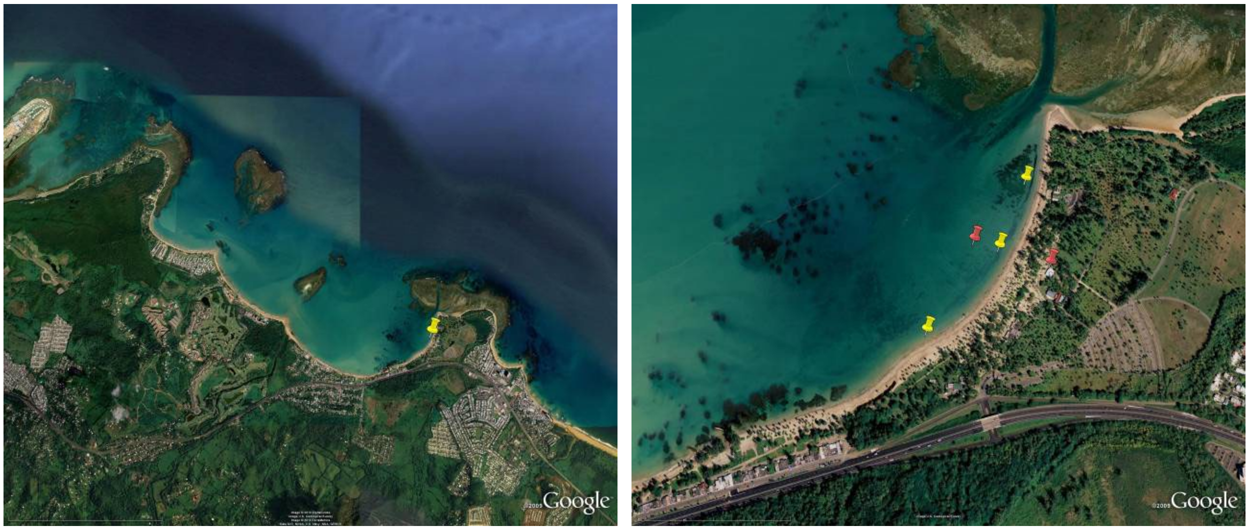


Figure A.9. Location of 2008 PREMIER study at La Monserrate Beach in Luquillo, Puerto Rico. Yellow pins indicate locations of beach and sampling transects. Details regarding field equipment (red pins) are given in Section 4.4.3. The beach is on the northeast coast of Puerto Rico (Figure 4.2B).

A.2.7 Boquerón Beach, Cabo Rojo, Puerto Rico

Boquerón Beach is in southwestern Puerto Rico, in the town of Boquerón in Cabo Rojo (Figure 4.2B). The 1.6-km-long beach is along the eastern shore of Boquerón Bay, on the western coast of Puerto Rico. The bay is ~ 4.7 km wide at its mouth and the beach is ~ 4 km from the mouth. Potential sources of fecal contamination to the beach include a sewage treatment plant's outfall in the bay, ~ 1.3 km northwest of the beach, and two package plants that operate during periods of high demand (Figure A.10). The plants discharge treated effluent into the mangrove lagoon south of the beach. The mouth of the lagoon is ~ 1.5 km from the beach. A marina and condominium complex, just north of the beach, is also a potential source of contamination. Urban runoff is also likely because afternoon storms are common during the summer, with heavy rains resulting in flooding.

PREQB monitors both enterococci and fecal coliform levels at Boquerón Beach every other week during the entire year. Data for the current year are posted on the web (http://www.prtc.net/~jcaagua), but we have not yet analyzed the data. Beachgoers at Boquerón are advised to avoid direct contact with waters in the 24 hours following heavy rainfall events because of the potential risk of exposure to pathogens.

Figure A.10. Location of 2009 NEEAR study at Boquerón Beach in Cabo Rojo, Puerto Rico. Yellow pins indicate locations of beach and sampling transects and green pins mark possible sources of fecal contamination (i.e., WWTP outfall and package plants). Details regarding field equipment (red pins) are given in Section 4.4.3. The beach is on the eastern coast of Puerto Rico (Figure 4.2B).

Table A.3. Historical water quality monitoring details and criteria exceedances for tropical marine beaches based on data provided by the local monitoring agency, PREQB. Number of exceedances is expressed per total samples, with the corresponding percentage given in parentheses.

	La Monserrate	Boquerón
Water Body		Boquerón Bay
Location	Luquillo, Puerto Rico	Cabo Rojo, Puerto Rico
Possible Sources	Streams	Effluent
	Runoff	Runoff
Climate	tropical	tropical
Beach Length	400 m	1600 m
Data Source	PREQB	PREQB
Sampling Frequency	2x /mo	2x /mo
Sampling Period	year round	year round

	Water Quality Criteria for Enterococcus (CFU / 100 mL) and Number of Exceedances			
Year	> 35	> 104	> 35	> 104
2006	7/27 (26)	4/27 (15)		
2007	4/31 (13)	1/31 (13)		
2008	2/14 (14)	1/14 (7)		
2009				
TOTAL	13/72 (18)	6/72 (8)		

REFERENCES

Abdelzaher, A.M., M.E. Wright, C. Ortega, H.M. Solo-Gabriele, G. Miller, S. Elmir, X. Newman, P. Shih, J.A. Bonilla, T.D. Bonilla, C.J. Palmer, T. Scott, J. Lukasik, V.J. Harwood, S. McQuaig, C. Sinigalliano, M. Gidley, L.R.W. Plano, X. Zhu, J.D. Wang, and L.E. Fleming. 2010. Presence of pathogens and indicator microbes at a non-point source subtropical recreational marine beach. *Applied and Environmental Microbiology*. 76(3):724-732.

ADEM (Alabama Department of Environmental Management). 2010. *DRAFT Total Maximum Daily Load (TMDL) for Mobile Bay*. Alabama Department of Environmental Management, Montgomery, AL.

Byappanahalli, M.N., R.L. Whitman, D.A. Shively, and M.B. Nevers. 2010. Linking non-culturable (qPCR) and culturable enterococci densities with hydrometeorological conditions. *Science of the Total Environment* 408(16):3096-3101.

Elmir, S.M. 2006. *Development of a water quality model which incorporates non-point microbial sources.* University of Miami, Civil Engineering, Coral Gables, FL

Francy, D.S., and R.A. Darner. 2006. *Procedures for developing models to predict exceedances of recreational water-quality standards at coastal beaches*. U.S. Department of the Interior, U.S. Geological Survey, Techniques and Methods 6–B5.

Francy, D.S., and R.A. Darner. 2007. *Nowcasting beach advisories at Ohio Lake Erie Beaches*. U.S. Department of the Interior, U.S. Geological Survey, 2007-1427.

Francy, D.S., R.A. Darner, and E.E. Bertke. 2006. *Models for predicting recreational water quality at Lake Erie beaches*. U.S. Department of the Interior, U.S. Geological Survey, Scientific Investigations Report 2006-5192.

Francy, D.S., A.M. Gifford, and R.A. Darner. 2003. *Escherichia coli at Ohio Bathing Beaches - Distribution, Sources, Wastewater Indicators, and Predictive Modeling*. U.S. Department of the Interior, U.S. Geological Survey, Water-Resources Investigations Report 02-4285, Columbus, OH.

Frick, W.E., Z. Ge, and R.G. Zepp. 2008. Nowcasting and forecasting concentrations of biological contamination at beaches: A feasibility and case study. *Environmental Science & Technology* 42:4818-4824.

Ge, Z., and W.E. Frick. 2007. Some statistical issues related to multiple linear regression modeling of beach bacteria concentrations. *Environmental Research* 103:358-364.

Ge, Z., and W.E. Frick. 2009. Time-frequency analysis of beach bacteria variations and its implication for recreational water quality modeling. *Environmental Science & Technology* 43:1128-1133.

Huang, X., and V. Sigler. 2006. *Population-Based Molecular-Tracking of E. Coli at Lake Erie Beach and Huntington Beach (Ohio)*. Paper read at the 49th Annual Conference of the International Association of Great Lakes Research, May 22–26, Windsor, Ontario.

IDEM (Indiana Department of Environmental Management). 2004. *Little Calumet and Portage Burns Waterway TMDL for E. coli Bacteria*. Final TMDL Report. Indiana Department of Environmental Management, Indianapolis, IN.

Liu, L., M.S. Phanikumar, S.L. Molloy, R.L. Whitman, D.A. Shively, M.B. Nevers, D.J. Schwab, and J.B. Rose. 2006. Modeling the transport and inactivation of *E. coli* and Enterococci in the near-shore region of Lake Michigan. *Environmental Science and Technology* 40:5022-5028.

McLellan, S. 2004. *Sources of E. coli at South Shore Beach Final Research Report*. Great Lakes WATER Institute, Milwaukee, WI.

McLellan, S.L., A.D. Daniels, and A.K. Salmore. 2001. Clonal populations of thermotolerant enterobacteriaceae in recreational water and their potential interference with fecal *Escherichia coli* counts. *Applied and Environmental Microbiology.* 67(10):4934-4938.

McLellan, S.L., E.J. Hollis, M.M. Depas, M. Van Dyke, J. Harris, and C.O. Scopel. 2007. Distribution and fate of *Escherichia coli* in Lake Michigan following contamination with urban stormwater and combined sewer overflows. *Journal of Great Lakes Research* 33(3):566-580.

McLellan, S.L., and E.T. Jensen. 2005. *Identification and Quantification of Bacterial Pollution at Milwaukee County Beaches*. Great Lakes WATER Institute, Milwaukee, WI.

McLellan, S.L., and A.K. Salmore. 2003. Evidence for localized bacterial loading as the cause of chronic beach closings in a freshwater marina. *Water Research* 37(11):2700-2708.

MDEQ (Mississippi Department of Environmental Quality). 2002. *Fecal Coliform TMDL for the Back Bay of Biloxi and Biloxi Bay*. Mississippi Department of Environmental Quality, Jackson, MS.

MDEQ (Michigan Department of Environmental Quality). 2003. *Total Maximum Daily Load for Escherichia coli for the St. Joseph River Berrien County*. Michigan Department of Environmental Quality, Lansing, MI.

MMSD (Milwaukee Metropolitan Sewerage District). 2005. *Bacteria Source, Transport and Fate Study - Phase 1, Milwaukee Harbor Estuary Hydrodynamic & Bacteria Modeling*. Milwaukee Metropolitan Sewerage District, Milwaukee, WI.

Nevers, M.B., and R.L. Whitman. 2005. Nowcast modeling of *Escherichia coli* concentrations at multiple urban beaches of southern Lake Michigan. *Water Research* 39:5250-5260.

Nevers, M.B., and R.L. Whitman. 2008. Coastal strategies to predict *Escherichia coli* concentrations for beaches along a 35 km stretch of southern Lake Michigan. *Environmental Science & Technology* 42:4454-4460.

Nevers, M.B., R.L. Whitman, W.E. Frick, and Z. Ge. 2007. Interaction and influence of two creeks on *Escherichia coli* concentrations of nearby beaches: Exploration of predictability and mechanisms. *Journal of Environmental Quality* 36:1338-1345.

NIRPC (Northwestern Indiana Regional Planning Commission). 2005. *Watershed Management Plan for Lake, Porter, and LaPorte Counties*. Northwestern Indiana Regional Planning Commission, Portage, IN.

Olyphant, G.A., J. Thomas, R.L. Whitman, and D. Harper. 2003. Characterization and statistical modeling of bacterial (*Escherichia coli*) outflows from watersheds that discharge into southern Lake Michigan. *Environmental Monitoring and Assessment* 81:289-300.

Otvos, E.G. 1999. Rain-induced beach processes; landforms of ground water sapping and surface runoff. *Journal of Coastal Research* 15(4):1040-1054.

RIDEM (Rhode Island Department of Environmental Management). 2005. *Total Maximum Daily Load Analysis for Greenwich Bay Waters: Pathogen/Bacteria Impairments*. Rhode Island Department of Environmental Management, Providence, RI.

Schwab, D.J., D. Beletsky, and G.A. Lang. 2010. *Indiana Dunes Nowcast* 2010. . <http://www.glerl.noaa.gov/res/glcfs/bd>. Accessed July 2010.

Scopel, C.O., J. Harris, and S.L. McLellan. 2006. Influence of nearshore water dynamics and pollution sources on beach monitoring outcomes at two adjacent Lake Michigan beaches. *Journal of Great Lakes Research* 32(3):543-552.

Shibata, T., H.M. Solo-Gabriele, L.E. Fleming, and S. Elmir. 2004. Monitoring marine recreational water quality using multiple microbial indicators in an urban tropical environment. *Water Research* 38:3119-3131.

Solo-Gabriele, H., T. Shibata, M. Al-Kendi, Y. St. Fort, L. Fleming, D. Squicciarini, W. Quirino, M. Arguello, S. Elmir, and M. Rybolowik. 2002. *A Pilot Study Evaluation and Sanitary Survey of Microbial Recreational Water Quality Indicators in the Subtropical Marine Environment*. National Institute of Environmental Health Sciences, Marine and Freshwater Biomedical Sciences Center, Rosensteil School of Marine and Atmospheric Sciences, Miami-Dade County Department of Health.

Telech, J.W., K.P. Brenner, R. Haugland, E. Sams, A.P. Dufour, L. Wymer, and T.J. Wade. 2009. Modeling *Enterococcus* densities measured by quantitative polymerase chain reaction and membrane filtration using environmental conditions at four Great Lakes beaches. *Water Research* 43:4947-4955.

Thupaki, P., M.S. Phanikumar, D. Beletsky, D.J. Schwab, M.B. Nevers, and R.L. Whitman. 2009. Budget analysis of *Escherichia coli* at a southern Lake Michigan beach. *Environmental Science & Technology* 44(3):1010-1016.

Triad Engineering Incorporated. 2003. *Trail Creek Escherichia coli TMDL Report*. Indiana Department of Environmental Management, Indianapolis, IN.

Wade, T.J., R.L. Calderon, K.P. Brenner, E. Sams, M. Beach, R. Haugland, L. Wymer, and A.P. Dufour. 2008. High sensitivity of children to swimming-associated bastrointestinal illness: Results using a rapid assay of recreational water quality. *Epidemiology* 19:375-383.

Wade, T.J., R.L. Calderon, E. Sams, M. Beach, K.P. Brenner, A.H. Williams, and A.P. Dufour. 2006. Rapidly measured indicators of recreational water quality are predictive of swimming-associated gastrointestional illness. *Environmental Health Perspectives* 114(1):24-28.

Whitman, R. 2010. *About Project S.A.F.E.* 2008. <http://www.glsc.usgs.gov/main.php?content=research_projectSAFE_about&title=Project%20S.A.F.E.0&menu=research_initiatives_projectSAFE>. Accessed July 2010.

Whitman, R., and M. Nevers. 2005. *Regional and local factors affecting patterns of E. coli distribution in southern Lake Michigan*. U.S. Geological Survey, Porter, IN.

Wong, M., L. Kumar, T.M. Jenkins, I. Xagoraraki, M.S. Phanikumar, and J.B. Rose. 2009. Evaluation of public health risks at recreational beaches in Lake Michigan via detection of enteric viruses and a human-specific bacteriological marker. *Water Research* 43:1137-1149.

Wright, M.E. 2008. *Evaluation of enterococci, an indicator microbe, and the sources that impact the water quality of a subtropical non-point source recreational beach*. University of Miami, Civil Engineering, Coral Gables, FL.

Wright, M.E., H.M. Solo-Gabriele, S. Elmir, and L.E. Fleming. 2009. Microbial load from animal feces at a recreational beach. *Marine Pollution Bulletin* In press, corrected proof.

Zhu, X. 2009. *Modeling microbial water quality at a non-point source subtropical beach*. University of Miami, Applied Marine Physics, Coral Gables, FL.

Appendix B. Additional Data Collection Details

B.1 SAMPLE COLLECTION

Details regarding the NEEAR study sample collection protocol have been described previously (Haugland et al. 2005; Wade et al. 2006). That general approach was taken at all NEEAR study sites. Briefly, samples were collected on weekends and holidays during the single summer study carried out at each beach. Samples were collected three times a day (8 a.m., 11 a.m., and 3 p.m.) at six locations at each beach. The locations included two depths (shin and waist deep) along three transects, ≥ 60 m apart. Shin samples were collected ~ 0.15 m below the surface in 0.3-m-deep water, and waist-deep samples were collected ~ 0.3 m below the surface in 1-m-deep water. Grab samples were collected in sterilized polypropylene bottles in accordance with Standard Methods Section 9060 (Clesceri et al. 1998). Samples were mixed to create three composite samples per time point: a shin composite, waist composite, and total composite. Total composites were primarily used for statistical modeling.

Sample collection for the 2008 PREMIER studies (at South Shore, Hobie, and La Monserrate) was designed deliberately to match the NEEAR studies, where possible. However, more frequent sampling was performed to provide a larger data set for modeling purposes. Waist-deep samples were collected four days a week (Mondays, Wednesdays, Thursdays, and Saturdays), three times a day (9 a.m., 11:30 a.m., and 3 p.m.) at South Shore and Hobie. Shin-deep samples were also collected on Saturdays at South Shore and on Thursdays and Saturdays at Hobie. No shin-deep samples were collected at La Monserrate, and waist-deep samples were collected only once a day (1000), three days a week (Mondays, Thursdays, and Saturdays) over a longer period (June through October 2008). At each sampling location and time, three 500-mL grab samples were collected and mixed to give composite samples for each of the three transects; unlike the NEEAR studies, no beach-wide composites were collected. The distance between sampling transects at the beaches ranged from 125 to 250 m. Because shin- and waist-deep samples were not collected with equal frequency, waist-deep composite samples were used for statistical modeling purposes.

The 2009 PREMIER studies at Surfside Beach and Boquerón were specifically designed to complement the concurrent NEEAR studies by expanding the spatial and temporal scale of data collection. In addition to the NEEAR sampling scheme (i.e., weekend and holiday samples) described previously, waist-deep samples were collected on Fridays (three times per day) as part of the PREMIER study. Samples were also collected from two locations outside the beach area (at 8 a.m. and 3 p.m. on Fridays, Saturdays, Sundays, and holidays to better evaluate potential sources of contamination to the beaches (i.e., runoff and effluent at Surfside and Boquerón, respectively). At Surfside Beach, samples were collected from the 3rd Avenue North swash channel along North Ocean Boulevard and from the lake at North Dogwood Drive. Boquerón samples were collected at the outfall of the wastewater treatment plant and just outside the mouth of the mangrove lagoon. For each sampling event at these sites, three grab samples were collected and composited. The distance between sampling transects at both of these beaches was ~ 150 m. As stated for the 2008 PREMIER sites, waist-deep composite samples were used for statistical modeling purposes.

In addition to the samples collected for microbial analysis (described above), water samples were also collected during the PREMIER studies to analyze dissolved organic carbon (DOC) and colored dissolved organic matter. The same collection scheme was used for those samples, as detailed above for the 2008 sites, without any sample compositing. At Surfside and Boquerón, beach water samples for the analyses were collected at only the waist-deep site of the middle transect. Single grab samples were also collected from the potential contamination source locations using the microbial sampling scheme. The samples were collected in glass bottles, appropriately cleaned as described in Standard Methods Section 5310B (Clesceri et al. 1998).

Samples collected for the studies provided a large data set that was used to develop predictive models, discussed in Chapter 5. Historical monitoring data can also be used for model development, but most local beach monitoring programs do not have the resources to carry out sampling efforts on the scale of the NEEAR and PREMIER studies. For example, while Huntington Beach, Ohio, has been sampled for *E. coli* daily since 2006, sampling is done only at one time (around 9 a.m.) and one location (30 cm below the surface in the surf zone at a depth of 1 m). Routine *E. coli* sampling at the Ogden Dunes Beach site is done less frequently (≤ 5 times per week), but samples are collected from three different locations, which provides better coverage of the larger beach area. Publicly available data from those two beaches were used (as were data obtained from the NEEAR and PREMIER studies) in our efforts to refine and evaluate the development of predictive models. Details regarding the beach field studies that generated data for the modeling studies discussed in this report are summarized in Table 4.1.

B.2 DEPENDENT VARIABLES

As the dependent variable, fecal indicator bacteria (FIB) density data are the most important component of predictive model development. The acceptable standard by which recreational water quality is evaluated is culturable enterococci; *E. coli* is also acceptable for freshwaters only.

As a result of the BEACH Act of 2000, more than 10 years of monitoring data now exist for beaches in the United States. Most of the data are available on the Internet, providing an extensive database of historical information. In the PREMIER and NEEAR studies described here, culturable enterococci were enumerated on membrane-Enterococcus Indoxyl-B-D-Glucoside (mEI) plates according to EPA Method 1600 (USEPA 2006). For that method, the detection limit was 1 CFU/volume filtered (i.e., 0.01 CFU/mL for 100 mL volumes) and a value of 0.5 CFU/100 mL (one-half the limit of detection) was used as the lower limit for data analysis. All culturable data were log-transformed before modeling. Because of the interest in more rapid monitoring techniques (i.e., quantitative polymerase chain reaction [qPCR]-based enterococci) and the paucity of existing qPCR monitoring data that include the environmental parameters required for modeling, a goal of the PREMIER studies was to build on the existing NEEAR data set by obtaining qPCR-based enterococci measurements at additional beach sites.

Enterococcus 23S rRNA target sequences were quantified by the TaqMan® qPCR assay, as detailed in EPA (2010), with several modifications. Briefly, samples were filtered and total DNA was extracted from enterococci collected on the membrane filter by bead beating. PCR amplification of the target DNA sequence was quantified on the basis of the increase in fluorescence resulting from enzymatic hydrolysis of the fluorogenically labeled probe during each amplification cycle. NEEAR samples were analyzed as described in Haugland et al. (2005), using the Sketa2 reverse primer for analysis of the Salmon DNA sample processing control sequence. That assay has been revised (as of April 2010) and now uses the Sketa22 reverse primer (USEPA 2010). qPCR results from the NEEAR studies are reported as calibrator cell equivalents (CCE) per 100 mL volume filtered, relative to calibrator samples containing a known quantity of *Enteroccocus* cells. Samples from the 2008 PREMIER studies were analyzed by the TaqMan® Fast Mix Entero2 assay described by Siefring et al. (2008) using the Sketa22 reverse primer. Reactions were performed in 96-well optical plates in an Applied Biosystems 7500 Fast Real-Time PCR instrument, using a thermal cycling program of 20 seconds (s) at 95 degrees Celsius (°C), followed by 40 cycles of 3 s at 95 °C and 30 s at 60 °C. Threshold cycles (C_T) were calculated by the instrument software, using a threshold setting of 0.03. qPCR results from the 2008 PREMIER studies are reported as *Enteroccocus* target sequence copies (TSC) per 100 mL volume filtered, as described in EPA (2010). Briefly, that involves comparison with TSC recovery in a calibrator standard for which TSC can be calculated from C_T (using a DNA standard curve). The limit of detection for *Enterococcus* was determined to be 5 TSC ($C_T \approx 37$).

As outlined in EPA (2010), samples with Salmon C_T values that were not within 3 C_T units of the calibrator were considered inhibited. Samples with higher C_T values (suggesting poor DNA recovery or PCR inhibition or both) were diluted by an additional factor of 5 (25-fold dilution) and reanalyzed to rule out or eliminate inhibition, wherever possible. If the sample was not inhibited, poor DNA recovery was corrected for using the ΔC_T method (Siefring et al. 2008). Samples with C_T values below detection or non-detects (i.e., no signal after 45 cycles) were assigned a value of one-half the detection limit.

Reporting qPCR results in different units (i.e., CCE versus TSC for the NEEAR and PREMIER studies, respectively) will not influence the modeling results. The units can be converted because the TSC extracted from a calibrator sample should be directly proportional to the number of *Enterococcus* cells. That is, the extraction efficiency for the calibrator sample should be reproducible such that the ratio of TSC to cells is always the same. As with the culturable data, CCE and TSC data were log-transformed before modeling.

B.3 Independent (Explanatory) Variables

Because water quality criteria are based on FIB levels, monitoring data are crucial for beach management. However, models can be useful in predicting indicator densities, which is increasingly important, given issues associated with capability of monitoring data to reflect beach conditions accurately at a given time. More easily measured environmental conditions, often obtained with automated techniques, can be used as independent variables (IVs) in statistical models to explain observed variability in FIB levels. Those include physical hydrologic measurements (e.g., water temperature, turbidity, current and wave information, tidal phase, and stream discharge); chemical and biological parameters (e.g., pH, dissolved oxygen, conductivity, salinity, and chlorophyll); meteorological conditions (e.g., rainfall, solar irradiation, air temperature, and wind information); and ancillary beach conditions (e.g., number of bathers and birds). A variety of such parameters were measured with FIB measurements during the NEEAR and PREMIER studies. While measurements were taken at the time of sample collection and in discrete samples during the NEEAR studies, the purpose of the PREMIER studies was to obtain more detailed IV data by deploying automated instruments at the beach sites to monitor ambient conditions. The measurements were also supplemented with variables mined from public databases, where available.

Details about collecting environmental data during the NEEAR studies were previously discussed for the four freshwater sites (Heaney et al. 2009); the same procedures were followed for the subsequent marine studies. Measurements recorded at each sampling event included the following parameters: air and water temperature; percent cloud cover; ultraviolet (UV) irradiance; wave height; current direction; wind speed and direction; number of bathers on the beach and in the water; total number of birds and animals within 20 m of the sampling area; number of boats within 500 m of the sampling area; and the presence of debris. Rainfall data for the study period were obtained from weather stations at nearby airports and from on-site weather stations installed at some of the beaches. Turbidity, pH, salinity, and conductivity were measured in collected water samples. Limited ancillary data were collected during the 2008 PREMIER studies because efforts were focused on in situ measurements. Local water and beach conditions were noted, including wave conditions (South Shore and Hobie), presence of birds (South Shore) or dogs (Hobie), and number of people and dogs within 50 m of sampling transects (Hobie).

IV data at South Shore Beach, Milwaukee; Hobie Beach, Miami; Surfside Beach, South Carolina; La Monserrate Beach, Luquillo, Puerto Rico; and Boquerón Beach, Cabo Rojo, Puerto Rico were obtained, using procedures and equipment that are described in detail in the balance of this appendix.

General. Field equipment was deployed at the PREMIER and NEEAR/ PREMIER beach sites to obtain more detailed and beach-relevant IVs from which to develop predictive water quality models. Field equipment locations for South Shore, Surfside, Hobie, La Monserrate, and Boquerón beaches are shown in Appendix A. In addition to using automated field instruments, a unique aspect of the PREMIER studies was measuring underwater UV radiation with the analysis of DOC and colored dissolved organic matter (i.e., UV-VIS spectra). By characterizing the optical properties of beach waters, the amount of light to which bacteria are exposed in the water column was more accurately determined in investigating the effects of light on the inactivation of FIB.

Weather stations. Meteorological conditions were monitored by installing HOBO (U30 NRC, Onset Computer Corporation) weather stations at or near each beach site. In Milwaukee, the weather station was on the roof of the Great Lakes WATER Institute, 3.3 km northwest of South Shore beach. At Surfside Beach, a weather station was installed on the roof of the Surfside Beach civic center, ~ 1 km west of the beach. A second, earlier model HOBO station was at the end of Surfside Beach pier, ~ 200 m offshore. The pier is 550 m southwest of the 5th Avenue North Swash sampling location. The weather station at Hobie Beach was deployed southeast of the beach on the grounds of the adjacent Miami Seaquarium property. At Luquillo Beach, the weather station was on the roof of a beach snack bar, ~ 200 m from the middle sampling transect. The Boquerón Beach weather station was deployed on top of the lifeguard station, just south of the handicap-access facilities, ~ 300 m south of the center sampling transect.

Weather stations were equipped with sensors to measure air temperature; relative humidity; dew point (determined from temperature and relative humidity); barometric pressure; wind speed and direction; gust speed (i.e., highest 3-second wind recorded during logging interval); rain; photosynthetically active radiation, solar radiation (silicon pyranometer); and UV radiation (Apogee Instruments sensors were used in 2009 only) every 15 or 30 minutes. Data were routinely downloaded in the field by connecting a laptop to the weather station data logger. Wind speed and direction were used to determine cross-shore (u component) and along-shore (v component) winds at each beach site. The on-site weather data were supplemented with other available meteorological data (http://cdo.ncdc.noaa.gov/ulcd/ULCD).

Acoustic Doppler current profilers. Current and wave information was obtained by deploying Nortek Aquadopp Profilers (2 MHz, right angle sensor head) at each beach site. The acoustic Doppler current profilers (ADCPs) were installed on the lake or sea floor using a weighted cross-frame with a mounting height of ~ 0.3 m (Mooring Systems, Inc.). At South Shore, the ADCP was under the South Shore Yacht Club dock (because of shallow water depths and heavy boat traffic), < 200 m northwest of the beach. At Surfside, an ADCP was deployed ~ 200 m from shore, roughly in line with the middle transect. The Hobie ADCP was on the northwest end of the beach, along the buoy line ~ 300 m northwest of middle sampling transect, 100 m from shore. The Luquillo ADCP was ~ 100 m from the beach shore, at the middle transect. Two ADCPs were used at Boquerón Beach to better characterize the hydrodynamics of the bay, as related to potential contamination sources; one ADCP was deployed at the beach, 55 m from shore along the southern sampling transect, and the second was ~ 400 m southeast of the WWTP outfall, ~ 1 km from the beach. A University of Puerto Rico-Mayagüez ADCP was also installed in the bay, ~ 600 m west of the mouth of the mangrove lagoon. The area immediately beyond the swimming area of both beaches in Puerto Rico was buoyed off from a channel for kayaks. It provided an excellent location for field equipment that was close to the sampling area and where water conditions were likely very representative of water in the swimming area.

Measurement of current velocity by the ADCP is based on the Doppler Effect. Sound waves transmitted by the Aquadopp are reflected by small particles in the water that move at the same speed as the water. Listening to the echo of the transmitted sound, the Aquadopp determines water velocity by measuring changes in frequency. To generate depth profiles of current speed and direction, current measurements are taken in cells along three acoustic beams. Current measurements were taken every 10 minutes, and velocity and pressure data were collected every one or two hours to measure wave bursts; post-processing of the data was done with the Nortek Storm software. Current speed and direction for each cell were averaged to obtain current information for the overall water column. Depth averaged cross-shore (u component) and along-shore (v component) currents were determined from current speed and direction. Approximate water depth (estimated from pressure), temperature, significant wave height, and mean wave direction also were acquired from the ADCP data. ADCPs at South Shore and Luquillo were deployed with a communication cable (from the Aquadopp to a weatherproof enclosure mounted above water on the dock or UV sensor tower) to enable data retrieval in the field with a laptop. The Hobie ADCP was routinely recovered from the water, and data were downloaded before redeployment in the same location. In Surfside Beach and Boquerón, ADCPs were left unattended for the duration of the study period.

Water quality sondes. Multi-parameter water quality sondes (YSI 6600V2-2) were deployed at each beach site for the duration of the studies. At South Shore, Surfside, Luquillo, and Boquerón, sondes were deployed on the support frame of the UV sensor instrument package (described below), although the Surfside sonde was later relocated to its own buoy. The equipment was placed ~ 500 northwest of South Shore Beach, 80 m from shore to avoid boat traffic. At the other beaches, sondes (and UV sensors) were adjacent to the ADCP, allowing enough space so as not to interfere with the ADCP measurements. At Hobie, the sonde was hung from an existing marker buoy (Dade County Public Works Rickenbacker Causeway Buoy 6A) on the northwest end of the study area. The sondes were deployed at fixed locations at a depth of < 2 m. Although actual sonde depths varied at most sites because of tidal changes, sonde readings were taken to represent surface conditions.

Water temperature, specific conductance, salinity, dissolved oxygen (using a Clark oxygen electrode), pH, turbidity, and chlorophyll (as relative fluorescence) were measured every 15 minutes. Nitrate, ammonia, and ammonium were also measured at the freshwater South Shore Beach using ion selective electrodes. The sondes were typically retrieved every one to two weeks for cleaning, calibration, and data retrieval. Probes were calibrated as follows: conductivity (10 mS/cm), dissolved oxygen water-saturated air), pH (pH 7 and 10 buffers), ammonium (1 and 100 mg/L), nitrate (1 and 100 mg/L), chlorophyll (deionized water), and turbidity (deionized water and 123 nephelometric turbidity units [NTU]). Fouling was significant and in some cases could be identified as having influenced data quality, specifically for optical measurements (i.e., turbidity and chlorophyll). Fouling of the optical sensors results in high turbidity and chlorophyll readings with a high frequency of spikes. Real (i.e., event-driven) spikes in data progress in a natural upward trend and are short in duration. Examples of bad data from fouling (available from the instrument manufacturer) were used to help identify questionable data during manual review by an experienced individual. Data for periods during which biofouling was suspected were manually deleted from the final data set to avoid the use of possibly erroneous data. Likewise, data for periods in which calibrations were later found to be unacceptable were manually deleted. Regarding treatment of the turbidity data, the correction with an offset of 0.5 NTU (i.e., adding 0.5 NTU to negative values) was done at the manufacturer's recommendation. Errors resulting in negative values are very common when the sonde is not adequately cleaned (before it is calibrated between deployments) and can result in contamination of the blank. In addition, a longer calibration cup should have been used to minimize interference during calibration when the sonde is to be used in low turbidity waters. While the data could have been excluded, it was decided that negative values should be adjusted (to between 0.05 and < 0.5 NTU) and included because values that low clearly indicate that turbidity was very low (regardless of the exact value). A minimum value of 0.05 NTU was used because it is one-half the smallest value that the turbidity probe can measure (i.e., 0.1 NTU). Using a small positive value rather than 0 prevented the loss of data points following the log-transformation used to process the data before modeling.

Underwater UV radiation. Downwelling irradiance (E_d) was measured at each beach using pairs of Satlantic multispectral radiometers (OCR-504 ICSW) with 305, 325, 340, and 380 nm channels. The sensors were placed at two depths to evaluate the attenuation of UV radiation in the water column. The top sensor was < 0.5 m below the water surface (at low tide) and the bottom sensor was placed ~ 0.6 m to 1.5 m below the top sensor, depending on water clarity; for example, the bottom sensor at Luquillo was placed deeper because UV penetration was deeper in the clearer water. To reduce biofouling on the sensor optics, each sensor was equipped with a copper Satlantic Bioshutter. The shutters opened for hourly measurements (during daylight hours) and data were logged for 60 s. Sensors were cleaned sporadically and, while fouling was not a major issue, the lack of routine cleaning resulted in deterioration of data quality over the course of the study.

In addition to the sensors and Bioshutters, each UV sensor instrument package included the following Satlantic equipment: a STOR-X data logger, battery pack (51Ah), wireless telemetry system with GSM modem, and dual band marine-grade cellular antenna. The modem was housed in a weatherproof enclosure, mounted above the water surface, and connected by cable to the data logger. All other equipment was deployed underwater on temporarily installed tower structures. Irradiance data were transmitted hourly as an email attachment. In some cases, problems with signal strength prevented communication, resulting in intermittent data transmission. In those cases, data were recovered from the loggers when instruments were retrieved at the end of the study. In 2009 surface UV and visible light sensors were also included to provide a better reference for surface light levels.

UV irradiance data were processed using the Satlantic SatCon data conversion software. Data were averaged over each 60-s logging event and irradiance values ≤ 0.01 µW cm^{-2} nm^{-1} were excluded. The rate at which irradiance at a given wavelength decreases as a function of depth can be described by a vertical attenuation coefficient for downwelling irradiance ($K_d(\lambda)$). K_d values were estimated from the irradiance measured at the top and bottom sensors, according to the following equation (Smith and Baker 1981):

$$K_d(\lambda) = -\frac{ln\left(\frac{E_{d,z_2}(\lambda)}{E_{d,z_1}(\lambda)}\right)}{z_2 - z_1} \qquad \text{(equation B.1)}$$

where $E_d(\lambda)$ is the downwelling irradiance measured at each wavelength (λ) by the top (2) and bottom (1) sensors, at a given depth (z) below the water surface. Sensor depth varied during each logging event due to tidal variations at Hobie and Luquillo. Approximate sensor depths were determined as a function of time at each of three beaches, using ADCP pressure readings as a measure of total water depth. Available tide and water level data from nearby NOAA/NOS/CO-OPS stations were also used to relate depths measured during deployment to tidal cycle water level. Data were retrieved (www.tidesandcurrents.noaa.gov) for the study dates from the following stations: Milwaukee, Wisconsin (Station ID: 9087057); Springmaid Pier, South Carolina (Station ID: 8661070); Virginia Key, Florida (Station ID: 8723214); and San Juan, Puerto Rico (Station ID: 9755371). In cases where measured irradiance was higher at the bottom sensor than the top sensor, likely because of fouling or physical blockage of the top (or both) sensors (e.g., seaweed floating over), K_d values were negative. These erroneous data were excluded. Irradiance 0.3 m below the water surface ($E_{d,0.3}$), the depth at which water was sampled for bacterial analyses, was calculated from K_d and top sensor irradiance and depth by rearranging equation B.1.

Dissolved organic material. DOC was measured in filtered water samples as non-purgeable organic carbon by high temperature combustion using a Shimadzu 5050A or TOC-V Total Organic Carbon Analyzer. Samples were acidified to pH ≤ 2 with hydrochloric acid and sparged for 8 minutes before injection. The instrument was calibrated with potassium hydrogen phthalate standards (0.1 to 1 mM carbon) in NANOpure water. A blank and check standard were included with every six samples. Analysis of samples, in triplicate (75 µL injection volume), gave a coefficient of variation of ≤ 5 percent.

UV-visible absorption spectra were determined from 200 to 800 nm for 0.2-µm filtered water samples, using a Perkin-Elmer LAMBDA™ 35 UV/Vis Spectrophotometer equipped with a sipper system and a 5-cm long microvolume, flow-through quartz cell (Hellma). Spectra were referenced against NANOpure water (Barnsted) and then baseline-corrected by adjusting the absorbance (A_λ) to zero between 690 and 710 nm. Measured A_λ were used to calculate absorption coefficients (i.e., $a_\lambda = 2.303(A_\lambda/l)$ where l is the path length in m) at 350 nm.

This page is intentionally blank.

Appendix C. Regression Modeling Results

Tables C.1 through C.26 show the MLR models that were chosen for each site and each response variable (culturable/qPCR data). The first column in each table is the name of the IV. *POLY* means the variable was transformed using a polynomial ($ax^2 + bx + c$) before regression analysis. That was done because such a transformation best linearized the relationship between the IV and the response variable, as measured by a Pearson correlation coefficient (see Chapter 2). The values of a, b, and c were chosen using an ordinary least squares regression approach. We do not show the values of *a*, *b*, and *c*. *INV* means the inverse transformation of the original parameter was instituted for modeling. Columns two through five are the fitted regression coefficient for that parameter, its standard error, t-statistic, and significance level. Each table note gives the adjusted R^2, root mean square error (RMSE) and sample size for each of the models.

C.1 FRESHWATER SITES

Table C.1. Regression model for South Shore Beach enterococci qPCR data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-3.227	1.221	-2.643	0.010
Chlorophyll	0.012	0.004	2.615	0.011
POLY[Riverflow]	0.893	0.302	2.960	0.004
Conductivity	0.003	0.001	3.098	0.003
48hr Rainfall - Site 1	-0.192	0.079	-2.431	0.018
Dissolved Oxygen	-0.060	0.026	-2.288	0.025
Dissolved Organic Carbon	0.001	0.000	2.376	0.020
POLY[48hr Rainfall - Site 2]	0.808	0.251	3.214	0.002

Note: Adjusted R^2 = 0.39, RMSE = 0.36, n = 81.

Table C.2. Regression model for South Shore Beach enterococci culturable data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	0.245	0.921	0.266	0.791
DISO2_WS	-0.124	0.035	-3.588	0.001
INV[CHL_WS]	0.028	0.012	2.381	0.020
POLY[PRESSURE_AP]	0.762	0.377	2.025	0.047
RAIN_AP	112.886	21.692	5.204	0.000
POLY[CURRENTU_ES]	0.910	0.335	2.715	0.008
INV[WINDU_AP]	-0.256	0.078	-3.279	0.002
POLY[DEW_AP]	-1.440	0.418	-3.447	0.001
POLY[TURBIDITY_ES]	1.124	0.389	2.892	0.005

Note: Adjusted R^2 = 0.56, RMSE = 0.33, n = 79.

Table C.3. Regression model for Huntington Beach, Ohio, enterococci qPCR data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-0.871	0.861	-1.011	0.318
Turbidity	0.007	0.002	2.882	0.007
POLY[# of Boats]	0.783	0.373	2.103	0.042
# of Bathers	0.429	0.086	4.972	0.000
Cloud Cover	0.262	0.047	5.565	0.000
Antecedent Rainfall	1.200	0.416	2.883	0.007

Note: Adjusted R^2 = 0.69, RMSE = 0.42, n = 44.

Table C.4. Regression model for Huntington Beach, Ohio, enterococci culturable data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-0.213	0.517	-0.412	0.683
Antecedent Rainfall	1.961	0.458	4.285	0.000
POLY[Amount of Debris]	0.913	0.278	3.288	0.002
Turbidity	0.010	0.002	4.290	0.000
# of Bathers	-0.393	0.119	-3.300	0.002
# of Boats	0.377	0.115	3.278	0.002

Note: Adjusted R^2 = 0.62, RMSE = 0.48, n = 45.

Table C.5. Regression model for Huntington Beach, Ohio, *E.coli* culturable data, 2003.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-1.177	0.996	-1.181	0.245
POLY[Air Temp - Dry Bulb]	0.604	0.271	2.231	0.031
Air Temp - Wet Bulb	0.026	0.013	2.027	0.049
Wave Height	0.213	0.072	2.967	0.005
Antecedent Rainfall	21.413	8.386	2.553	0.015
INV[Turbidity]	-1.454	0.580	-2.508	0.016

Note: Adjusted R^2 = 0.54, RMSE = 0.38, n = 46.

Table C.6. Regression model for Huntington Beach, Ohio, *E.coli* culturable data, 2000–2009.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	14.535	3.790	3.835	0.000
Wave Height	0.147	0.024	6.064	0.000
Air Pressure	-0.491	0.129	-3.812	0.000
Acrossshore Wind	-0.014	0.003	-4.390	0.000
Turbidity	0.009	0.001	9.346	0.000
Dewpoint	0.019	0.003	7.489	0.000
Antecedent Rainfall	24.702	3.423	7.217	0.000

Note: Adjusted R^2 = 0.44, RMSE = 0.47, n = 709.

Table C.7. Regression model for Washington Park enterococci qPCR data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-0.622	0.815	-0.763	0.448
Wave Height	-0.566	0.229	-2.471	0.016
POLY[Air Temp]	0.733	0.364	2.015	0.048
POLY[Turbidity]	0.824	0.404	2.040	0.046

Note: Adjusted R^2 = 0.20, RMSE = 0.59, n = 66.

Table C.8. Regression model for Washington Park enterococci culturable data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-1.132	0.692	-1.636	0.107
Cloud Cover	0.062	0.023	2.696	0.009
INV[Turbidity]	-0.282	0.087	-3.252	0.002
POLY[# of Bathers]	1.046	0.387	2.701	0.009
POLY[# of Dogs]	0.765	0.316	2.420	0.019
Antecedent Rainfall	-0.363	0.128	-2.839	0.006
Acrossshore Wind	0.160	0.071	2.256	0.028
Algae	-0.485	0.145	-3.339	0.002

Note: Adjusted R^2 = 0.45, RMSE = 0.24, n = 66.

Table C.9. Regression model for Silver Beach enterococci qPCR data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	1.021	0.381	2.684	0.010
Alongshore Wind	-0.042	0.012	-3.567	0.001
Air Temp	0.035	0.014	2.461	0.017
INV[Acrossshore Wind]	0.137	0.033	4.137	0.000
Water Temp	-0.042	0.013	-3.316	0.002
Wind Speed	0.073	0.020	3.737	0.001

Note: Adjusted R^2 = 0.41, RMSE = 0.41, n = 58.

Table C.10. Regression model for Silver Beach enterococci culturable data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-3.879	1.010	-3.839	0.000
# of Birds	0.002	0.001	2.989	0.004
POLY[Air Pressure]	1.472	0.417	3.529	0.001
Water Temp	0.025	0.011	2.272	0.027
POLY[# of Bathers]	1.694	0.420	4.030	0.000

Note: Adjusted R^2 = 0.32, RMSE = 0.36, n = 58.

Table C.11. Regression model for West Beach enterococci qPCR data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-2.391	0.603	-3.962	0.000
# of Boats	-0.122	0.059	-2.076	0.044
POLY[Wind Speed]	0.984	0.135	7.294	0.000
Water Temp	0.032	0.010	3.264	0.002
POLY[Air Temp]	0.959	0.257	3.738	0.001

Note: Adjusted R^2 = 0.63, RMSE = 0.41, n = 49.

Table C.12. Regression model for West Beach enterococci culturable data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-3.611	0.550	-6.565	0.000
Algae	-0.666	0.223	-2.996	0.005
Turbidity	0.053	0.007	7.837	0.000
Dewpoint	0.054	0.009	5.877	0.000
POLY[Air Pressure]	0.544	0.201	2.703	0.010
Wind Speed	0.098	0.040	2.444	0.019

Note: Adjusted R^2 = 0.78, RMSE = 0.40, n = 49.

C.2 MARINE SITES

Table C.13. Regression model for Boquerón enterococci qPCR data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-2.550	1.617	-1.577	0.119
Water Temp	0.158	0.055	2.856	0.006
Debris	0.081	0.025	3.205	0.002
INV[Acrossshore Wind]	-0.128	0.064	-2.013	0.048
INV[Cloud Cover]	-0.273	0.121	-2.260	0.027

Note: Adjusted R^2 = 0.23, RMSE = 0.29, n = 79.

Table C.14. Regression model for Boquerón enterococci culturable data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	5.835	2.181	2.675	0.011
Turbidity	0.029	0.011	2.597	0.013
Specific Conductivity	-0.096	0.039	-2.461	0.018
# of Bathers	0.001	0.000	2.113	0.041

Note: Adjusted R^2 = 0.43, RMSE = 0.43, n = 44.

Table C.15. Regression model for Edgewater Beach enterococci qPCR data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	0.393	1.105	0.355	0.724
POLY[Antecedent Rainfall]	0.916	0.450	2.037	0.047
INV[Cloud Cover]	-0.039	0.016	-2.381	0.021

Note: Adjusted R^2 = 0.14, RMSE = 0.44, n = 55.

Table C.16. Regression model for Edgewater Beach enterococci culturable data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	0.905	0.196	4.613	0.000
UV	-0.001	0.000	-5.722	0.000
Wave Height	2.041	0.777	2.626	0.012
Alongshore Wind	-0.049	0.023	-2.122	0.039
# of Dogs	0.825	0.343	2.404	0.020
Algae	0.738	0.187	3.957	0.000

Note: Adjusted R^2 = 0.48, RMSE = 0.56, n = 55.

Table C.17. Regression model for Fairhope Beach enterococci qPCR data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-7.141	2.680	-2.665	0.010
POLY[Air Temp]	2.631	1.089	2.417	0.019
POLY[Dewpoint]	0.678	0.321	2.111	0.039
UV	0.000	0.000	-3.912	0.000
POLY[Water Temp]	1.208	0.433	2.788	0.007

Note: Adjusted R^2 = 0.35, RMSE = 0.49, n = 66.

Table C.18. Regression model for Fairhope Beach enterococci culturable data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-0.521	2.196	-0.237	0.813
Water Temp	-0.238	0.052	-4.539	0.000
Humidity	0.028	0.009	3.227	0.002
POLY[Dewpoint]	0.933	0.308	3.029	0.004
INV[# of Birds]	1.579	0.602	2.625	0.011
Air Temp	0.066	0.033	2.031	0.047

Note: Adjusted R^2 = 0.43, RMSE = 0.57, n = 66.

Table C.19. Regression model for Goddard Beach enterococci qPCR data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	2.827	1.625	1.740	0.087
POLY[Water Temp, Waist Deep]	-3.905	1.601	-2.440	0.017
Amount of Debris	0.539	0.149	3.613	0.001
POLY[Water Temp, Shin Deep]	2.739	0.878	3.119	0.003

Note: Adjusted R^2 = 0.18, RMSE = 0.70, n = 69.

Table C.20. Regression model for Goddard Beach enterococci culturable data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-0.064	0.660	-0.097	0.923
Tide	-0.218	0.038	-5.761	0.000
POLY[Wind Speed]	2.484	1.179	2.107	0.039
# of Birds	0.008	0.003	2.843	0.006

Note: Adjusted R^2 = 0.36, RMSE = 0.60, n = 69.

Table C.21. Regression model for Surfside Beach enterococci qPCR data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-4.519	1.121	-4.031	0.000
Relative Humidity	0.022	0.006	3.570	0.001
Antecedent Rainfall	0.023	0.005	4.426	0.000
Turbidity	0.125	0.041	3.092	0.004
POLY[Absorbance]	1.091	0.256	4.267	0.000
POLY[# of Birds]	0.964	0.370	2.604	0.014

Note: Adjusted R^2 = 0.35, RMSE = 0.60, n = 82.

Table C.22. Regression model for Surfside Beach enterococci culturable data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	35.322	10.955	3.224	0.002
Antecedent Rainfall	0.011	0.005	2.526	0.014
POLY[Wave Height]	0.516	0.279	1.850	0.068
# of Boats	0.055	0.025	2.216	0.030
POLY[Salinity]	1.092	0.255	4.284	0.000
Time of Day	-1.723	0.418	-4.122	0.000
pH	0.591	0.224	2.637	0.010
Air Pressure	-0.040	0.011	-3.666	0.001
POLY[# of Bathers]	0.993	0.493	2.015	0.048

Note: Adjusted R^2 = 0.46, RMSE = 0.43, n = 85.

Table C.23. Regression model for Hobie Beach enterococci qPCR data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-2.221	1.038	-2.140	0.052
POLY[PAR]	0.972	0.314	3.098	0.009
Absorbance	0.029	0.010	3.079	0.009
POLY[Relative Humidity]	0.601	0.244	2.463	0.029

Note: Adjusted R^2 = 0.67, RMSE = 0.06, n = 17.

Table C.24. Regression model for Hobie Beach enterococci culturable data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	-1.825	0.593	-3.076	0.003
Turbidity	-0.016	0.006	-2.745	0.007
Chlorophyll	0.759	0.265	2.864	0.005
Relative Humidity	0.026	0.008	3.271	0.002

Note: Adjusted R^2 = 0.19, RMSE = 0.47, n = 97.

Table C.25. Regression model for La Monseratte Beach enterococci qPCR data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	3.606	0.322	11.196	0.000
Absorbance	-0.748	0.364	-2.055	0.053
Alongshore Current	-102.088	44.678	-2.285	0.033
Antecedent Rainfall	0.005	0.002	2.552	0.019

Note: Adjusted R^2 = 0.45, RMSE = 0.45, n = 24.

Table C.26. Regression model for La Monseratte Beach enterococci culturable data.

Parameter	Coefficient	Std. Error	t-Statistic	P-Value
(Intercept)	41.548	17.795	2.335	0.027
Chlorophyll	-10.755	2.939	-3.659	0.001
Water Depth	17.555	6.247	2.810	0.009
Salinity	-1.264	0.516	-2.448	0.021

Note: Adjusted R^2 = 0.41, RMSE = 0.59, n = 32.