In cooperation with the Cuyahoga County Board of Health; Northeast Ohio Regional Sewer District; Ohio Lake Erie Office; and U.S. Environmental Protection Agency, Advanced Monitoring Initiative

Testing and Refining the Ohio Nowcast at Two Lake Erie Beaches—2008

Open-File Report 2009–1066

U.S. Department of the Interior
U.S. Geological Survey

Cover image: View of Edgewater Beach, Cleveland, Ohio. (Photograph by Donna S. Francy, U.S. Geological Survey.)

Testing and Refining the Ohio Nowcast at Two Lake Erie Beaches—2008

By Donna S. Francy, Erin E. Bertke, and Robert A. Darner

In cooperation with the Cuyahoga County Board of Health; Northeast Ohio Regional Sewer District; Ohio Lake Erie Office; and U.S. Environmental Protection Agency, Advanced Monitoring Initiative

Open-File Report 2009–1066

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
KEN SALAZAR, Secretary

U.S. Geological Survey
Suzette M. Kimball, Acting Director

U.S. Geological Survey, Reston, Virginia: 2009

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment, visit http://www.usgs.gov or call 1-888-ASK-USGS

For an overview of USGS information products, including maps, imagery, and publications, visit http://www.usgs.gov/pubprod

To order this and other USGS information products, visit http://store.usgs.gov

Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this report is in the public domain, permission must be secured from the individual copyright owners to reproduce any copyrighted materials contained within this report.

Suggested citation:
Francy, D.S., Bertke, E.E., and Darner, R.A., 2009, Testing and refining the Ohio Nowcast at two Lake Erie beaches—2008: U.S. Geological Survey Open-File Report 2009–1066, 19 p.

Contents

Abstract 1
Introduction 1
Methods 2
Data Collection and Operation of the Nowcast 3
Variables for Predictive Models 3
Testing and Refinement of Predictive Models 6
Results 7
Performance of the Models in 2008 7
Exploratory Data Analysis 9
Model Refinement and Selection for 2009 13
Edgewater 14
Huntington 14
Next Steps 16
Acknowledgments 16
References Cited 17
Appendix 1. Candidate models used to predict *Escherichia coli* concentrations at two Lake Erie beaches 19

Figures

1. Location of study area. 2
2. Locations of sampling sites and instrumentation 4
3. Boxplots showing comparison of lake levels in subseason 2 for 2004–7 and 2008 at Edgewater, Cleveland, Ohio 7
4. Wave heights measured by use of the wave rod (manual) or wave buoy (automated), Edgewater, Cleveland, Ohio, 2008 12
5. Turbidity measured by use of a grab sample and turbidimeter (manual) or probe attached to a nearshore buoy (automated), Edgewater, Cleveland, Ohio, 2008 12
6. Wave heights measured by use of the wave rod (manual) or wave buoy (automated), Huntington, Bay Village, Ohio, 2008 13
7. Comparison of morning and afternoon *Escherichia coli* concentrations at Huntington, Bay Village, Ohio, 2008 14

Tables

1. Models developed for the Ohio Nowcast in 2008 5
2. Nowcast model responses as compared to using previous day's *Escherichia coli* concentration (persistence model) in 2008 8
3. Pearson's r correlations between $\log_{10}$ *Escherichia coli* (*E. coli*) concentrations and explanatory variables for subseason 1 (May 19–June 17), subseason 2 (June 18–Aug. 13), subseason 3 (Aug 14–Sept. 11), and the recreational year at Edgewater, Cleveland, Ohio, 2008 10
4. Pearson's r correlations between $\log_{10}$ *Escherichia coli* (*E. coli)* concentrations and explanatory variables for subseason 1 (May 19–July 23), subseason 2 (July 24–Sept. 5), the recreational year, and for afternoon samples at Huntington, Bay Village, Ohio, 2008 11
5. Examination of candidate Edgewater 2005–8 models 15
6. Examination of candidate Huntington 2005–8 models 16

Conversion Factors

Multiply	By	To obtain
inch (in.)	2.54	centimeter (cm)
inch (in.)	25.4	millimeter (mm)
foot (ft)	0.3048	meter (m)
yard (yd)	0.9144	meter (m)
mile (mi)	1.609	kilometer (km)
kilometer (km)	0.6214	mile (mi)
milliliter (mL)	0.06102	cubic inch (in^3)

Temperature in degrees Celsius (°C) may be converted to degrees Fahrenheit (°F) as follows:

$$°F=(1.8\times°C)+32$$

Temperature in degrees Fahrenheit (°F) may be converted to degrees Celsius (°C) as follows:

$$°C=(°F-32)/1.8$$

Concentrations of *Escherichia coli* are given in colony-forming units per 100 milliliters (CFU/100 mL).

Turbidities are given in nephelometric turbidity units (NTU).

Testing and Refining the Ohio Nowcast at Two Lake Erie Beaches—2008

By Donna S. Francy, Erin E. Bertke, and Robert A. Darner

Abstract

The Ohio Nowcast has been providing real-time beach advisories to the public on the basis of predictive models since 2006. In support of the nowcast, data were collected during the recreational season of 2008 to validate and refine predictive models at two Lake Erie beaches. Predictive models yield data on the probability that the single-sample bathing-water standard for *E. coli* will be exceeded. Field personnel collected or compiled data on *Escherichia coli* (*E. coli*) concentrations as well as variables expected to affect these concentrations, including manual and automated measurements of turbidity, wave height, and water temperature; lake level; and radar and airport rainfall amounts. Two new variables were measured during 2008—photosynthetically-active radiation at Huntington (Bay Village) and foreshore head at Edgewater (Cleveland).(The foreshore is a strip of land along a body of water between low and high water marks.)

The performance of the nowcast was monitored during 2008. The Huntington nowcast yielded a greater percentage of correct responses (84.9 percent) than did the previous day's *E. coli* concentration (75.2 percent). In contrast, at Edgewater, the nowcast yielded a slightly higher percentage of correct responses (61.0 percent) as compared to the previous day's *E. coli* concentration (56.5 percent), but both percentages were relatively low. Lake levels in 2008 were significantly higher than levels in the data used to develop the Edgewater models (2004−7), confounding their abilities to provide correct responses. At Edgewater during 2008, the strongest relation (as measured by Pearson's correlation) was between *E. coli* concentrations and the difference in foreshore head over the past 24 hours (r=0.48), a variable not included in the models. At Huntington, photosynthetically-active radiation on the previous day showed a significant negative relation to *E. coli* concentrations (r=-0.33) during 2008.

Refined models were developed for Huntington and Edgewater using data collected from 2005−8. The refined models included the variables wave height, log turbidity, radar or airport rainfall, and day of the year in various combinations for different dated segments of the recreational season. Water-resource managers will determine which models to apply to the Ohio Nowcast for issuing water-quality advisories in 2009.

Introduction

In Ohio, concentrations of the fecal-indicator bacterium *Escherichia coli* (*E. coli*) are used as the basis for issuing water-quality advisories at beaches. If the *E. coli* concentration exceeds the single-sample bathing water standard of 235 colony-forming units per 100 milliliters (CFU/100 mL), an advisory is issued for the beach (Ohio Department of Health, 2007). The widely acknowledged shortcoming of using this approach is that standard culture methods for *E. coli* take at least 18−24 hours before results are available. The beach is posted with an advisory or is determined to be acceptable for swimming on the basis of the previous day's concentration of *E. coli*. Sanitary conditions may change overnight and even throughout the day (Boehm and others, 2002), so beach managers may issue water-quality advisories based on outdated information of current public-health risk. As an alternative, some water-resource managers are using predictive models to provide near-real-time estimates of recreational water quality to better protect public health. Predictive models use water-quality and environmental variables, such as rainfall and turbidity, to provide these estimates.

The Ohio Nowcast has been providing near-real-time beach advisories to the public on the basis of predictive models since May 2006 (*http://ohionowcast.info*). The Ohio Nowcast is an Internet-based system that displays data on the probability that the *E. coli* concentration at a beach will exceed safe levels—similar to a weather forecast, except for a shorter timeframe. During the recreational season (May−Aug.) of 2007 at Huntington Reservation (Bay Village), the bathing-water standard was exceeded on 14 percent of the days sampled (15 out of 104 days). The nowcast for Huntington resulted in a greater percentage of correct responses (82.7 percent) than did the use of the previous day's *E. coli* concentrations (80.2 percent) (the method used for many advisories in the United States, often referred to as the "persistence model"). Although this difference was smaller than expected, the nowcast model better predicted exceedance of the standard (53 percent correct) than did the persistence model (31 percent correct) (Francy and Darner, 2007). Data from 2007 were added to the existing dataset (2000−2006), and two new models were developed for Huntington for use in the nowcast

in 2008. Work was also done to develop predictive models for a second beach, Edgewater State Park (Cleveland) (Francy and Darner, 2007), and Edgewater was added to the Ohio Nowcast in 2008. During 2007 at Edgewater, the bathing-water standard was exceeded on 28 percent of the days sampled (22 out of 78 days).

The U.S. Geological Survey (USGS), in cooperation with the Cuyahoga County Board of Health (CCBH), Northeast Ohio Regional Sewer District (NEORSD), Ohio Lake Erie Office, and U.S. Environmental Protection Agency, Advanced Monitoring Initiative, continued to operate and evaluate the Ohio Nowcast during 2008 at Huntington and Edgewater. This project involved the daily operation of the nowcast and the validation and refinement of predictive models. Equipment was installed to measure and transmit near-real-time-data on wave height at Huntington and Edgewater and to measure turbidity and foreshore head at Edgewater. A device to measure photosynthetically active radiation (PAR) was installed at Huntington to determine whether predictive models could be improved by adding PAR as an explanatory variable. Afternoon sampling was added to daily morning sampling at Huntington to determine whether this additional sampling would result in more accurate nowcast advisories. Finally, the data from 2008 were compiled and new models for Huntington and Edgewater were developed for 2009.

Methods

Data were collected for the Ohio Nowcast during the recreational season of 2008 (May−August) at two Lake Erie beaches in northeast Ohio—Huntington Reservation ("Huntington") and Edgewater State Park ("Edgewater") (fig. 1). Huntington is in a suburb of Cleveland (Bay Village) and is operated by the Cleveland Metroparks. Edgewater, in Cleveland, is an urban beach operated by the Ohio Department of

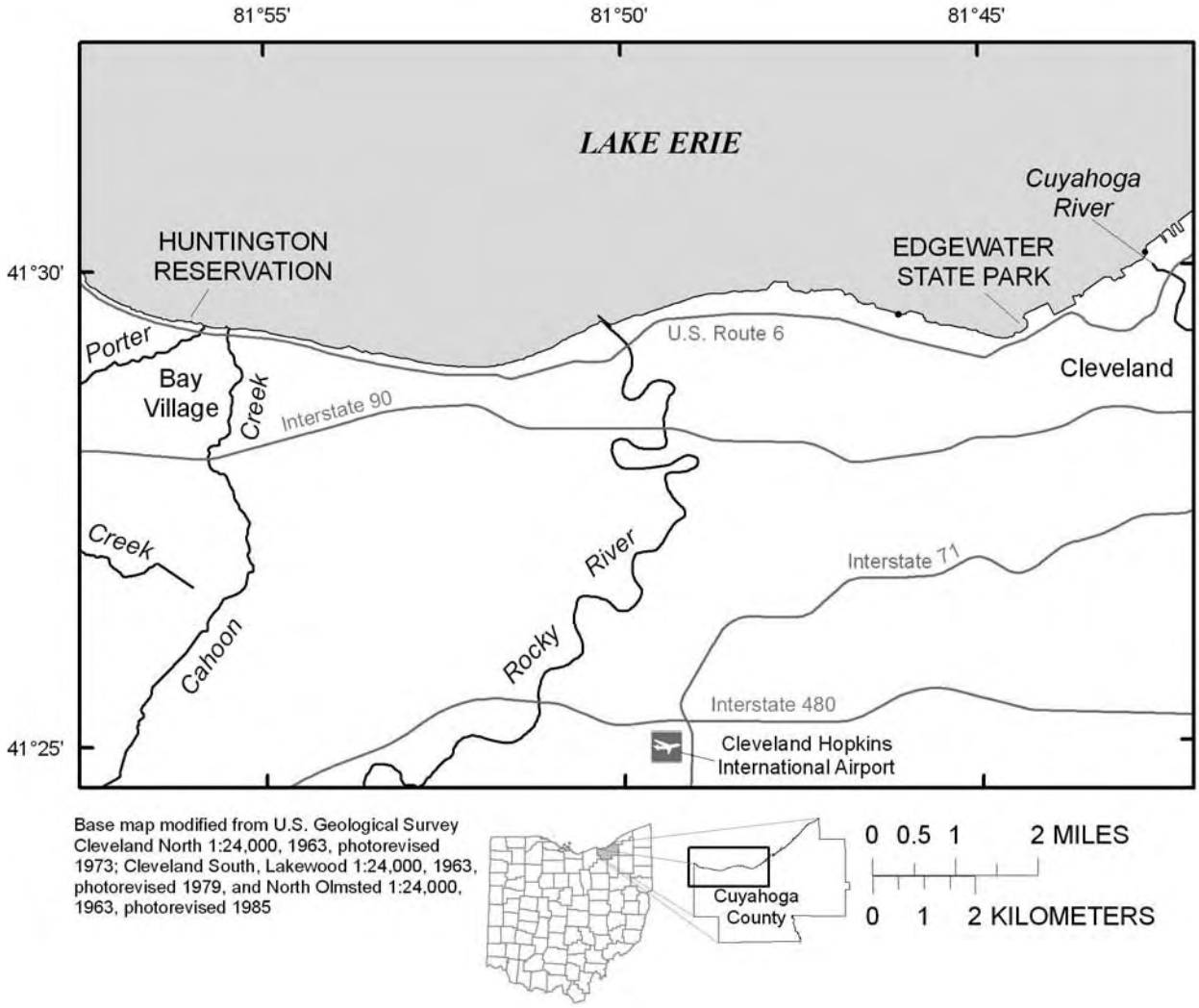

Figure 1. Location of study area.

Natural Resources. Both beaches are popular bathing beaches, used extensively during the recreational season. The NEORSD and CCBH are responsible for monitoring Edgewater and Huntington, respectively; these agencies updated the Ohio Nowcast Web site daily by 9:30 a.m. with water-quality advisories based on the models and other information (*http://ohionowcast.info*). Nowcast results were also listed by the Ohio Department of Health through their "Beach Monitoring" Web site (Ohio Department of Health, 2007).

Data Collection and Operation of the Nowcast

Lake-water samples were collected by CCBH and NEORSD for the nowcast between 7:30 and 9:30 a.m., 7 days per week, at a depth of 2−3 ft in two designated areas of each beach (fig. 2) by means of a grab-sampling technique (Myers and others, 2007). Additional samples were collected at Huntington between 1:00 and 3:00 p.m., Monday through Thursday; these sample results were not used in the nowcast during 2008 but may be used in future modeling efforts. Water samples were immediately placed on ice while the field crew completed field measurements and observations.

After the daily water-sample collection, equipment for measuring turbidity and a field laptop computer were set up on site. Rainfall from the nearest airport and lake-level data were obtained via the World Wide Web. Radar rainfall data and wave heights measured by use of a site-specific buoy (see "Variables for Predictive Models" section of this report) were received through an automated email from the USGS. Field personnel entered water-quality and environmental data into a computer program. The program provided results for the nowcast as the probability that today's water sample would exceed 235 CFU/100 mL. On the basis of a threshold probability established for each model from historical data (Francy and Darner, 2007), the nowcast was updated with current water-quality and advisory information (*http://ohionowcast.info*). Threshold probabilities were determined before the swimming season by taking the dataset used to develop the model and finding the probability that was a reasonable balance between achieving a high number of correct responses and a low number of false negative responses; this procedure is described in detail by Francy and Darner (2006, example 10). Computed probabilities less than the threshold indicated that bacterial water quality was most likely acceptable for swimming. Computed probabilities equal to or greater than the threshold probability indicated that the water quality was most likely not acceptable, and an advisory was issued.

The models used in 2008 were developed from data collected during 2000−2007 for Huntington and 2004−7 for Edgewater. Two models for Huntington and three models for Edgewater were used in the 2008 nowcast, based on specific time periods and denoted as subseason 1, subseason 2, or subseason 3 models (table 1). Threshold probabilities for the models ranged from 27 to 34 percent. Different models were used for each subseason because, in the data used to develop the models, it was noted that the percentages of false negatives were high in the beginning of the season and the percentages of false positive were high late in the season (Francy and Darner, 2007). To minimize the number of incorrect model responses, the 2000−2007 data for Huntington and 2004−7 data for Edgewater were, therefore, split into segments based on the patterns of false negatives and positives that would have occurred in responses from a model. In addition, because of poor performance of the Edgewater nowcast during subseason 2 of 2008, two modified models were developed for use in the later part of subseason 2 and in subseason 3.

Samples were collected and analyzed for *E. coli* concentrations within 3 hours. Morning samples were analyzed by use of the modified mTEC membrane-filtration method (U.S. Environmental Protection Agency, 2006). Because of the required 2-hour resuscitation for the modified mTEC method, afternoon samples were analyzed by use of the Colilert Quanti-Tray/2000 (IDEXX Laboratories, Inc., Westbrook, Maine), a commercially available enzyme-substrate liquid-broth medium that can be processed in less than 5 minutes. For Edgewater, composite samples were analyzed. To form a composite sample, 100-mL aliquots from each of the east and west sampling points (fig. 2*A*) were combined. For Huntington, the central and west sampling point samples (fig. 2*B*) were analyzed separately, and average concentrations were used to compute daily *E. coli* concentrations.

Quality-assurance and quality-control (QA/QC) procedures were implemented to ensure the collection of high-quality data for the nowcast. The USGS did several onsite QA/QC checks of field and laboratory operations. Duplicates, field blanks, and positive-control reference cultures for *E. coli* were analyzed as described in Francy and others (2008). Duplicate measurements of each water sample for turbidity were made. The turbidity measurements that did not agree within 1 NTU for values <10 NTU or 10 percent for values >10 NTU were repeated. Turbidity reference standards were sent to participating agencies. Results from QA/QC procedures were documented and retests were done or corrective measures taken when needed.

Variables for Predictive Models

Field personnel collected or compiled daily data for environmental and water-quality variables expected to affect *E. coli* concentrations. All these data were used in exploratory data analysis done after the 2008 recreational season, whereas some data were collected specifically for the 2008 nowcast.

Wave height and wave period. Wave heights were measured by automated and manual techniques. For Edgewater, where automated data were collected and tested in the previous year (2007), automated wave-height data measured at 7:30 a.m. were used in the nowcast. If automated data were not available because of equipment or transmission problems, manual measurements were used. For Huntington, because this was the first year for automated data collection, only manually measured wave-height data were used in the nowcast.

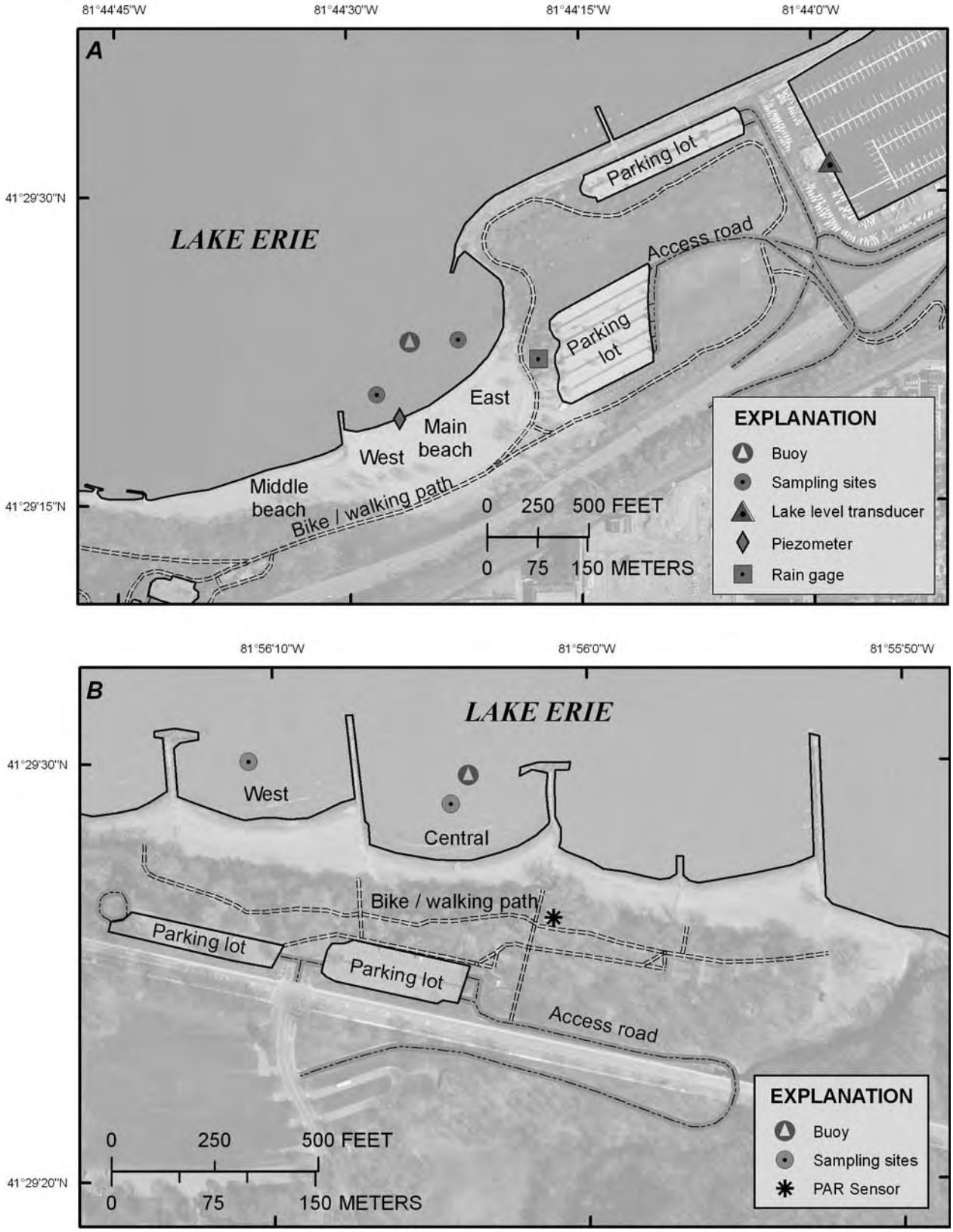

Figure 2. Locations of sampling sites and instrumentation. *A*, Edgewater, Cleveland, Ohio. *B*, Huntington Reservation, Bay Village, Ohio.

Table 1. Models developed for the Ohio Nowcast in 2008.

[Threshold probability is based on meeting or exceeding the single-sample bathing standard for *E. coli* (Francy and Darner, 2007)]

Model	Range of dates used	Variables	Threshold probability
		Huntington	
Subseason 1	May 19–July 23	Wave height, Rw48[a], log turbidity, antecedent dry days[b], day of the year	28
Subseason 2	July 24–Sept. 1	Wave height, Rw48[a], log turbidity, day of the year	27
		Edgewater	
Subseason 1	May 13–June 17	Wave height, Radar2cell-w48[c], log turbidity	27
Subseason 2a	June 18–July 18	Wave height, Radar2cell-w48[c], log turbidity, lake level	30
Subseason 2b (modified)	July 19–Aug. 13	Wave height, Radar2cell-w48[c], log turbidity, day of the year	34
Subseason 3	Not used	Wave height, log turbidity, lake level	32
Subseason 3 modified	Aug. 14–Sept. 11	Wave height, Radar2cell-w48[c], log turbidity	32

[a] Rw48 was the amount, in inches, at Hopkins International Airport, Cleveland, Ohio, in the 48-hour period before sampling, with the most recent rainfall receiving the most weight.

[b] Antecedent dry days were the number of days with zero precipitation preceding the current day.

[c] Radar2cell-w48 was the maximum amount of rainfall in the previous 48-hour period that fell in one of two 4-km grids, as determined from radar data, with the most recent rainfall receiving the most weight.

For automated data collection, wave heights and wave periods were compiled every 30 minutes from buoys placed outside the swimming areas at Huntington and Edgewater in approximately 6 ft of water. These data are available seasonally in near real time through the USGS National Water Information System Web-interface (NWISWeb), for Lake Erie at Edgewater Beach (USGS site ID 412923081442600) and Lake Erie at Huntington Reservation (USGS site ID 412928081560220). The buoys were custom made by an environmental equipment manufacturer for use in nearshore shallow waters. Buoys were equipped with pressure transducers to measure wave heights and wave periods; a datalogger to store the data; a radio, amplifier, and antenna to transmit the data; and a battery and solar panel for power.

For manual wave-height measurements, a graduated rod was placed at the sampling location and minimum and maximum water heights, in inches, were noted for 1 minute. Wave-height measurements were obtained in approximately 3 ft of water. Minimum was subtracted from maximum water height to obtain the manual wave height.

Water temperature. Water temperature was measured by field personnel at the sampling location at the time of sample collection by use of a digital thermometer. Although these data were entered into the nowcast Web site, they were not used in the predictive models. Additionally, water temperature, measured by use of a temperature probe attached to the Edgewater buoy, was recorded hourly and available on NWISWeb.

Turbidity. For the nowcast, duplicate manual measurements of turbidity were made with portable turbidimeters at Edgewater (Orion AQUAfast AQ4500) and Huntington (Orbeco-Hellige Model 966, Orbeco Analytical Systems, Inc., Famingdale, N.Y.). Additionally, at Edgewater, a turbidity probe (DTS-12 Digital Turbidity Sensor, FTS Forest Technology Systems Ltd., Victoria, British Columbia, Canada) attached to the buoy recorded hourly turbidity measurements. Turbidity data from the buoy were available on NWISWeb; however, the automated turbidity data were not used in the nowcast.

Foreshore head. At Edgewater, a temporary piezometer (shallow water well) was installed 20 ft inland from the edge of water at a depth of 3.5 ft (designated as "PZ-2") (fig. 2*A*). (The foreshore is defined as the strip of land along a body of water between low and high water marks.) The 2-in.-internal-diameter piezometer with a 0.5-ft-long screen was driven to the desired depth by hand. The relative elevation of the piezometer was determined by conventional surveying technique. The piezometer was equipped with a pressure transducer and data logger to measure and record water levels at 15-minute intervals. Water-level data at 8:30 a.m. from the piezometer were used in exploratory data analysis and are referred to as "foreshore head." The foreshore head is not equal to an actual change in water level; rather, it reflects the increased pressure from waves on water levels (data not shown). The head measured in PZ-2 can best be characterized as a combination of the pressures changes due to the local water table, fluctuations in barometric pressure, the water level in the lake, and the pressure variance due to wave action.

Lake level. Lake-level data were obtained from a station in Cleveland, approximately 400 yards offshore (National Oceanic and Atmospheric Administration [n.d.], station 9063063). These data were used in the Edgewater nowcast. Lake levels were also measured by the USGS by use of a pressure transducer placed on a pier in the harbor to the east of Edgewater (fig. 2*A*). The data from the pressure transducer were postprocessed to correct for barometric-pressure fluctuations. The elevation of the pressure transducer was surveyed with differential GPS to tie it to the same datum as the piezometer.

Photosynthetically active radiation. A PAR sensor (LI-COR Biosciences, Lincoln, Nebr.) and data logger were mounted on top of a 50-ft water tower at Huntington to record PAR data every 10 minutes (fig 2*B*). For exploratory data analysis, PAR data from midnight to midnight were summed to obtain daily PAR values.

Rainfall. Rainfall data were obtained from the nearest airport, from radar, or from a site-specific USGS-installed gage near Edgewater.

Airport rainfall data were obtained for Cleveland Hopkins International Airport (National Oceanic and Atmospheric Administration, 2002) (fig. 1) for Huntington and Edgewater. The total rainfall, in inches, for the 24-hour period preceding the morning sampling (8:51 to 7:51) was used in calculations for the nowcast (R_{d-1}). Similarly, the total rainfall in the 24-hour period preceding the afternoon sampling (14:51 to 13:51) or since the morning sampling (8:51 to 13:51) were used for data analysis of afternoon data.

Radar rainfall data were compiled in order to obtain information from a more widespread area. Hourly radar rainfall data were provided for 4-km grids ("cells") by the National Weather Service. For Edgewater and Huntington, previous analysis identified the radar rainfall variable that had the strongest relation to *E. coli* concentrations (Francy and Darner, 2007). For Edgewater, this variable was identified as the maximum value from two squares for the previous 48 hours and weighted as shown below (radar2cell-w48) (Edgewater is near the center of two 4-km squares that are bounded by a box with the coordinates lat 41°29′07.7″, long -081°48′29.4″; lat 41°31′13.8″, long -081°47′16.7″; lat 41°29′25.4″, long -081°41′39.9″; and lat 41°27′19.3″, long -081°42′52.7″.) For Huntington, radar rainfall data from the previous 24 hours from six squares were summed (radar-6cell-24) and used in the model (Huntington is near the center of six 4-km squares that are bounded by a box given the coordinates lat 41°28′30.5″, long -082°02′08.0″; lat 41°32′42.2″, long -081°59′42.8″; lat 41°30′01.1″, long -081°51′17.8″; and lat 41°25′49.5″, long -081°53′43.5″.)

Weighted rainfall variables were calculated from airport or radar rainfall data. Rainfall weighted 48 hours (Rw48) was 48 hours of cumulative rainfall, giving more weight to the most recent rainfall amount as follows:

$$Rw48 = (2*R_{d-1}) + R_{d-2}$$

where R_{d-1} is the amount of rain, in inches, that fell in the 24-hour period preceding the morning sampling and R_{d-2} is the amount of rain that fell in the 24-hour period 2 days preceding the morning sampling.

A rain gage was installed near Edgewater (fig. 2*A*). The raingage was installed on top of a nearby building at a location of lat 41°29′22″, long -081° 44′ 17.5″ and consisted of a tipping bucket calibrated for 0.01 in. per tip and electronic datalogger that recorded the number of tips in each 10-minute period.

Antecedent dry days. Antecedent dry days were calculated by counting the number of consecutive days without measurable rainfall up to and including 7:51 a.m. on the date of sampling.

Swimmers. At the time of the afternoon sampling, the number of swimmers was counted and recorded by field personnel at Huntington.

Day of the year. Day of the year was the number representing the date beginning with 1 for January 1 and 365 or 366 for December 31 (the latter being a leap year).

Testing and Refinement of Predictive Models

The models previously developed for the Huntington and Edgewater nowcast were validated during 2008 by monitoring the correct predictions and sensitivies and specifities of the model responses compared to the use of the previous day's *E. coli* concentrations (persistence model). Looking at these data in hindsight, a correct response is based on the actual *E. coli* concentration. The sensitivity is the percentage of exceedances of the bathing-water standard that were correctly predicted by the nowcast or the persistence model. The specificity is the percentage of nonexceedances where the nowcast or the persistence model correctly predicted nonexceedance of the standard. To further evaluate performance of the models in 2008, Pearson's r correlations (Helsel and Hirsch, 2002) were

calculated to determine the strength of linear relations between *E. coli* concentrations and the nowcast variables or other variables.

After validation of the models in 2008, the 2008 data were added to the existing dataset, and refined models were developed for 2009. Steps to developing new models involved the use of multiple linear regression techniques, described in detail in Francy and Darner (2006), and consist of (1) exploratory data analysis and (2) model development, diagnostics, and selection.

Results

Performance of the Models in 2008

The Huntington nowcast models (table 1) yielded a greater percentage of correct responses (84.9 percent) than did the persistence model (75.2 percent) for the recreational season of 2008 (table 2). The nowcast models had a greater sensitivity than the persistence model, but specificity differences were small. The improved sensitivity of the nowcast models over the persistence model is especially important; for the 14 days the bathing-water standard was exceeded in 2008, the nowcast predicted 8 of them (57.1 percent) whereas the persistence model predicted only 1 (7.1 percent).

Although the Edgewater nowcast models yielded a slightly higher percentage of correct responses (61.0 percent) as compared to the persistence model (56.5 percent), both percentages were relatively low (table 2). Specificity was improved by using the nowcast models, but the nowcast models did not improve sensitivity (31.7 percent) over the persistence model (37.5 percent). The subseason 2a nowcast model was unable to predict any exceedance events correctly. The subseason 2a model included lake level as an explanatory variable. A t-test (Helsel and Hirsch, 2002) was used to evaluate the difference between lake levels in 2008 for subseason 2 and the lake levels used to develop the subseason 2a model (2004−7) for the same dates. Lake levels in 2008 for subseason 2 were significantly higher than those in 2004−7 ($p<0.0001$, fig. 3). Because of the negative relation between *E. coli* concentrations and lake levels for 2004−7 (Francy and Darner, 2007), the subseason 2a model was underpredicting the probability that the standard was exceeded during 2008. Consequently, a modified subseason 2 model was developed without lake level (subseason 2b model) and was adopted for use in the nowcast on July 19. The subseason 2b model improved the sensitivity of the nowcast over the subseason 2a model (54.5 and 0 percent, respectively) but resulted in lower specificity than the subseason 2a model (60.0 and 90.5 percent).

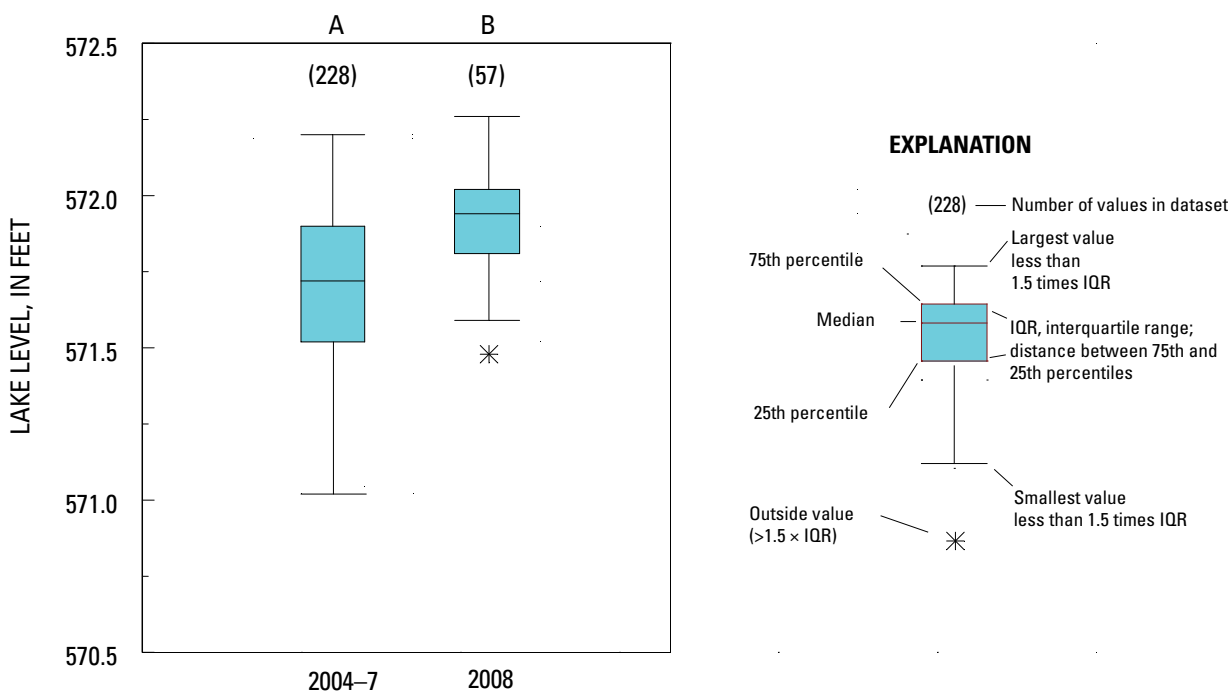

Figure 3. Boxplots showing comparison of lake levels in subseason 2 for 2004–7 and 2008 at Edgewater, Cleveland, Ohio

Table 2. Nowcast model responses as compared to using previous day's *Escherichia coli* concentration (persistence model) in 2008.

Model	Subseason	Sample size	No. Days >235[a]	Percentage of responses		
				Correct	Sensitivity[b]	Specificity[c]
Huntington [Subseason 1 is May 19–July 23, subseason 2 is July 24–Sept. 5]						
Nowcast	1	66	8	81.8	50.0	86.2
Persistence	1	65	8	76.9	0.0	87.7
Nowcast	2	40	6	90.0	66.7	94.1
Persistence	2	40	6	72.5	16.7	82.3
TOTAL NOWCAST	1 and 2	106	14	84.9	57.1	89.1
TOTAL PERSISTENCE	1 and 2	105	14	75.2	7.1	85.7
Edgewater [Subseason 1 is May 19–June 17, subseason 2a is June 18–July 18, subseason 2b is July 19–Aug. 13, and subseason 3 is Aug. 14–Sept. 11]						
Nowcast	1	34	13	64.7	30.8	85.7
Persistence	1	32	12	59.4	41.7	70.0
Nowcast	2a	31	10	61.3	0.0	90.5
Persistence	2a	31	10	45.2	20.0	57.1
Nowcast modified	2b	26	11	57.7	54.5	60.0
Persistence	2b	26	11	61.5	54.5	66.7
Nowcast	3	27	7	59.2	42.9	65.0
Persistence	3	26	7	61.5	28.6	73.7
TOTAL NOWCAST	1,2, and 3	118	41	61.0	31.7	76.6
TOTAL PERSISTENCE	1,2, and 3	115	40	56.5	37.5	66.7

[a] 235 colony-forming units per 100 milliliters is the single-sample bathing-water standard.

[b] Sensitivity was the proportion of exceedance responses that were correctly predicted as as an exceedance.

[c] Specificity was the proportion of nonexceedance responses that were correctly predicted as as an nonexceedance.

Exploratory Data Analysis

Exploratory data analysis was done to better understand the performance of the models in 2008 and identify the most promising variables for inclusion in refined models. Comparisons of data on automated and manual measurements were made to determine whether they could be used interchangeably in the nowcast models. In addition, afternoon and morning *E. coli* data were examined in terms of meeting or exceeding the standard and for their relations to explanatory variables to determine whether afternoon data can provide more accurate estimates of changing water-quality conditions.

E. coli and explanatory variables. The correlations between *E. coli* and explanatory variables are shown for data collected in 2008 at Edgewater (table 3) and Huntington (table 4). Correlations significant at $p<0.05$ are bolded, and the relations for the variables used in the nowcast are shaded.

At Edgewater, none of the relations examined were significant during subseason 1, including the variables used in the nowcast; however, wave height measured with the wave rod was nearly significant (p=0.0545). For the subseason 2 nowcast variables, two were significantly related to *E. coli* concentrations (radar2cell-w48 and wave height)and log turbidity was nearly significant, but lake level and day of the year were not significantly related to *E. coli*. For subseason 3, only one of the nowcast variables—wave height—was significantly related to *E. coli*. For the recreational year (subseasons 1, 2, and 3 combined), the strongest relation was between *E. coli* concentrations and the difference in foreshore head in PZ-2 over the past 24 hours, a new variable first measured during 2008. The Edgewater rain gage for the previous 24 hours showed a slightly higher correlation to *E. coli* concentrations (r=0.34) than the comparable radar (r=0.32) or Hopkins rainfall (r=0.29).

At Huntington, relations were significant between *E. coli* concentrations and the nowcast variables for subseasons 1 and 2, except for day of the year. For the recreational year (subseasons 1 and 2 combined), day of the year was nearly significant with respect to *E. coli* concentrations (p=0.0626). The radar data resulted in a slightly higher correlation than the airport data for subseason 1 and the recreational year, but not for subseason 2. For example, for subseason 1, Radar6cell-w48 had a higher correlation coefficient (r=0.45) than the comparable Rw48 (r=0.33). The sum of the PAR on the previous day showed a significant negative relation to *E. coli* concentrations (r=-0.33); this was a new variable introduced during 2008.

Manual and automated measurements. The relations between *E. coli* concentrations and wave height or turbidity measured manually (using the wave rod or turbidimeter) and automatically (using the buoys) were compared by means of datasets that included both measurements. These datasets were slightly smaller than those in tables 3 and 4 because some automated measurements were missing. It should be noted that there were differences in the ways these data were collected. Automated measurements were made in approximately 6 ft of water and consistently recorded at 7:00 (turbidity) or 7:30 (waves) each morning. Manual measurements were made in approximately 3 ft of water and made at the time and location of sampling, which varied between 7:30 and 9:30 a.m.

A comparison of manual and automated wave heights at Edgewater is shown in fig. 4. The relation to E. coli concentrations was slightly stronger for the manual measurement (fig. 4*A*, r=0.43) than for the automated measurement (fig. 4*B*, r=0.33). The slopes and y-intercepts that resulted from the regression relations between E. coli concentrations and wave heights were similar for manual and automated measurements. Even though these relations were similar and the manual and automated measurements were highly correlated (r=0.80), the plot of manual and automated measurements shows considerable scatter around the 1:1 and also shows that manual measurements were often higher than automated measurements (fig. 4*C*).

At Edgewater, relations were not significant between E. coli concentrations and either manual turbidity (r=0.24) or automated turbidity (r=0.05) (fig. 5*A*). Closer examination of two outliers measured by the buoy on July 24 and August 10 revealed that turbidities were fluctuating on both days, wave heights were elevated, and several data values were missing on August 10. Strong waves had moved the buoy closer to shore on July 24, and the buoy had to be placed back to the original nearshore location on July 25. After removing the data for July 24 and August 10, the correlation between E. coli concentrations and automated turbidity was still low (r = 0.05, p=0.6130). Fig. 5*C* shows that automated values were biased high as compared to manual values for turbidity.

A comparison of manual and automated wave heights at Huntington is shown in figure 6. The relation to E. coli concentrations was slightly stronger for the automated measurement (fig. 6*B*, r=0.55) than for the manual measurement (fig. 6*A*, r=0.47). The slope that resulted from the regression relation between E. coli concentrations and wave heights was steeper for the automated measurement than manual measurement. The plot of manual and automated paired measurements (fig. 6*C*) further illustrates the differences between these two measurements. Although the relation between the two measurements was significantly correlated (r=0.75, $p<0.001$), the placement of data around the 1:1 line shows that the manual measurements were often the higher.

Comparisons can also be made between wave heights measured by the Huntington and Edgewater buoys. The automated wave heights measured at Edgewater were significantly related to *E. coli* concentrations at Huntington (table 4, r=0.59). Similarly, the automated wave heights measured at Huntington were significantly related to *E. coli* concentrations at Edgewater (table 3, r=0.45). The correlation between wave heights measured at Huntington and Edgewater by each respective buoy was very high (r=0.92, $p<0.0001$).

Table 3. Pearson's r correlations between $\log_{10}$ *Escherichia coli* (*E. coli*) concentrations and explanatory variables for subseason 1 (May 19–June 17), subseason 2 (June 18–Aug. 13), subseason 3 (Aug. 14–Sept. 11), and the recreational year at Edgewater, Cleveland, Ohio, 2008.

[The *p* values are in parenthesis; relations that were significant at $p < 0.05$ are in bold; variables used in the Ohio nowcast are shaded; ND is not determined; PAR is photosynthetically active radiation; n indicates number of samples and those different from heading numbers are specified]

Variable	Subseason 1 n=34	Subseason 2 n=57	Subseason 3 n=27	Recreational year n=118
Radar2cell-24[a]	0.20 (0.2669)	**0.42 (0.0013)**	**0.42 (0.0279)**	**0.32 (0.0005)**
Radar2cell-48[a]	0.08 (0.6686)	0.10 (0.4805)	-0.14 (0.4921)	0.06 (0.5554)
Radar2cell-w48[b]	0.20 (0.2683)	**0.42 (0.0012)**	0.28 (0.1569)	**0.29 (0.0017)**
R_{d-1}[c]	0.24 (0.1653)	**0.29 (0.0274)**	**0.58 (0.0014)**	**0.29 (0.0013)**
Rainfall, total, Edgewater rain gage, previous 24 hrs	ND	ND	ND	**0.34 (0.0050)** n=66
Lake level at 8:30 a.m., measured by NOAA	0.15 (0.4040)	0.13 (0.3265)	0.13 (0.5217)	0.09 (0.3232)
Lake level at 8:30 a.m., measured by USGS				0.05 (0.6913) r=67
Foreshore head measured in Piezometer 2 at 8:30 a.m.	0.52 (0.1023) n=13	0.29 (0.0280)	0.97 (0.0068) n=5	**0.31 (0.0084)** n=73
Difference in foreshore head in Piezometer 2, 24 hrs	ND	ND	ND	**0.48 (<0.0001)** n=72
Turbidity, manual	0.25 (0.1508)	0.20 (0.1294)	0.20 (0.3055)	**0.21 (0.0208)**
Turbidity, automated, measured by buoy at 7:00 a.m.	0.06 (0.7942) n=20	0.06 (0.6859)	-0.01 (0.9537) n=22	0.05 (0.6094) n=98
Log Turbidity, manual	0.23 (0.1905)	0.26 (0.0543)	0.22 (0.2711)	**0.24 (0.0096)**
Log Turbidity, automated, measured by buoy at 7:00 a.m.	0.04 (0.8532) n=20	0.07 (0.6079)	-0.04 (0.8567) n=22	0.05 (0.6484) n=98
Wave height, manual, measured with wave rod	0.33 (0.0545)	**0.53 (<0.0001)**	**0.39 (0.0460)**	**0.44 (<0.0001)**
Wave height, automated, measured by Edgewater buoy at 7:30 a.m.	0.22 (0.3422) n=20	**0.37 (0.0042)**	0.33 (0.1295) n=22	**0.33 (0.0009)** n=99
Wave height, automated, measured by Huntington buoy at 7:30 a.m.	ND	ND	ND	**0.45 (0.0015)** n=46
PAR measured at Huntington, sum of previous day	ND	ND	ND	-0.16 (0.1682)
Day of the year	-0.01 (0.9594)	0.11 (0.4106)	0.04 (0.8583)	0.04 (0.6465)

[a] Radar2cell-24 and radar2cell-48 were the maximum amounts of rainfall that fell in one of two 4-km grids, as determined from radar data in the previous 24- and 48-hour periods, respectively.

[b] Radar2cell-w48 was the maximum amount of rainfall that fell in one of two, 4-km grids as determined from radar data in the previous 48-hour period, with the most recent rainfall receiving the most weight.

[c] R_{d-1} was the amount of rainfall that fell, in inches, at Hopkins International Airport Cleveland, Ohio in the 24-hour period preceding sampling.

Table 4. Pearson's r correlations between $\log_{10}$ *Escherichia coli* (*E. coli*) concentrations and explanatory variables for subseason 1 (May 19–July 23), subseason 2 (July 24–Sept. 5), the recreational year, and for afternoon samples at Huntington, Bay Village, Ohio, 2008.

[The p values are in parenthesis; relations that were significant at $p < 0.05$ are in bold; variables used in the Ohio nowcast are shaded; ND is not determined; PAR is photosynthetically active radiation, n indicates number of samples and those different from heading numbers are specified]

Variable	Morning 2008			Afternoon 2008 recreational year n=46
	Subseason 1 n=66	Subseason 2 n=40	Recreational year n=106	
Radar6cell-24[a]	**0.42 (0.0006)**	**0.38 (0.0159)**	**0.37 (0.0001)**	ND
Radar6cell-48[a]	0.16 (0.2044)	0.27 (0.0883)	0.16 (0.1120)	ND
Radar6cell-w48[b]	**0.45 (0.0002)**	**0.45 (0.0034)**	**0.39 (<0.0001)**	ND
R_{d-1}[c]	**0.32 (0.0080)**	**0.45 (0.0035)**	**0.29 (0.0027)**	0.19 (0.1992)
R_{d-2}[c]	0.09 (0.4604)	0.22 (0.1796)	0.07 (0.4534)	-0.04 (0.7935)
Rw48[d]	**0.33 (0.0075)**	**0.48 (0.0015)**	**0.28 (0.0035)**	0.15 (0.3107)
Number of antecedent dry days since last rainfall (R_{d-1})	**-0.45 (0.0002)**	**-0.55 (0.0002)**	**-0.39 (<0.0001)**	-0.25 (0.1012)
Day of the year	0.20 (0.1111)	-0.02 (0.8793)	0.18 (0.0626)	0.02 (0.9108)
PAR, day of sampling (midnight to midnight)	ND	ND	0.02 (0.8788)	-0.01 (0.9421)
PAR, previous day (midnight to midnight)	ND	ND	**-0.33 (0.0029)**	-0.26 (0.1137)
log PAR, previous day (midnight to midnight)	ND	ND	**-0.29 (0.0111)**	ND
PAR, day of sampling (8:50 to 13:50)	ND	ND	**ND**	-0.03 (0.8682)
Turbidity, manual	**0.37 (0.0022)**	**0.60 (<0.0001)**	**0.46 (<0.0001)**	**0.52 (0.0002)**
Log Turbidity, manual	**0.42 (0.0004)**	**0.64 (<0.0001)**	**0.50 (<0.0001)**	**0.51 (0.0003)**
Wave height, manual, measured by wave rod	**0.38 (0.0017)**	**0.68 (<0.0001)**	**0.49 (<0.0001)**	**0.58 (<0.0001)**
Wave height, automated, measured by Huntington buoy at 7:30 (AM samples) or 13:30 (PM sample)	**0.48 (0.0029)** n = 37	**0.85 (0.0035)** n = 9	**0.55 (<0.0001)** n = 46	**0.49 (0.0077)** n = 28
Wave height, automated, measured by Edgewater buoy at 7:30 AM	ND	ND	**0.59 (<0.0001)** n=96	ND
Number of swimmers	ND	ND	ND	-0.13 (0.3955)

[a] Radar6cell-24 and radar6cell-48 were the summed amounts of rainfall that fell in six 4-km grids, as determined from radar data in the previous 24- and 48-hour periods, respectively.

[b] Radar6cell-w48 was the summed amount of rainfall that fell in six 4-km grids, as determined from radar data in the previous 48-hour period, with the most recent rainfall receiving the most weight.

[c] R_{d-1} was the amount of rainfall that fell, in inches, at Hopkins International Airport Cleveland, Ohio in the 24-hour period preceding sampling; R_{d-2} was the amount 2 days before sampling.

[d] Rw48 was the amount of rainfall that fell, in inches, at Hopkings International Airport, Cleveland, Ohio in the previous 48-hour period, with the most recent rainfall receiving the most weight.

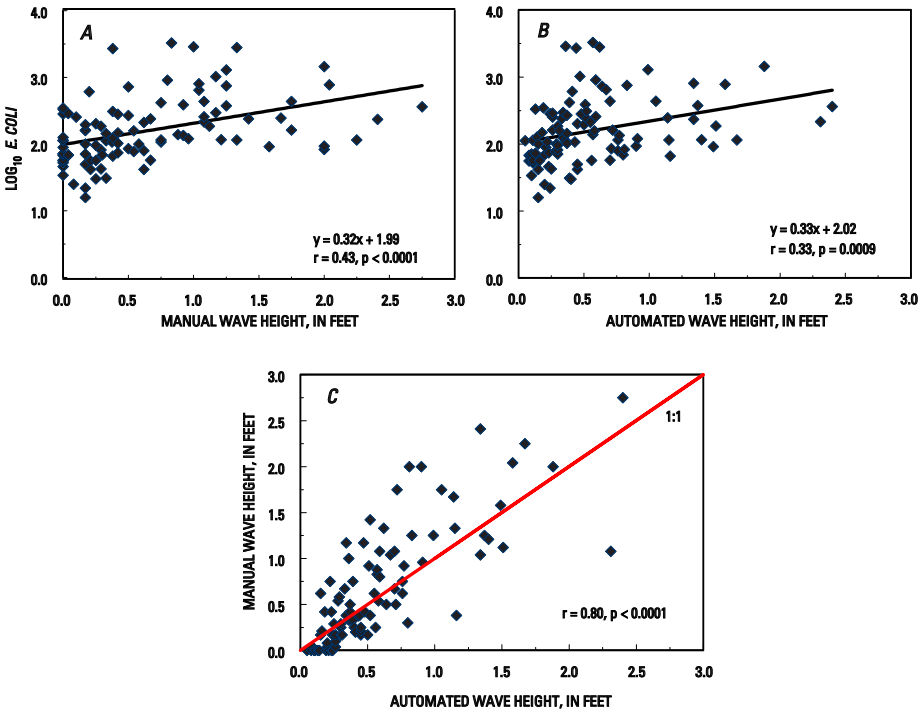

Figure 4. Wave heights measured by use of the wave rod (manual) or wave buoy (automated), Edgewater, Cleveland, Ohio, 2008. Relations between: *A*, Manual wave heights and *E. coli* concentrations. *B*, Automated wave heights and *E. coli* concentrations. C, Automated and manual wave heights. (r is the Pearson's correlation coefficient, p is the significance of the relation.)

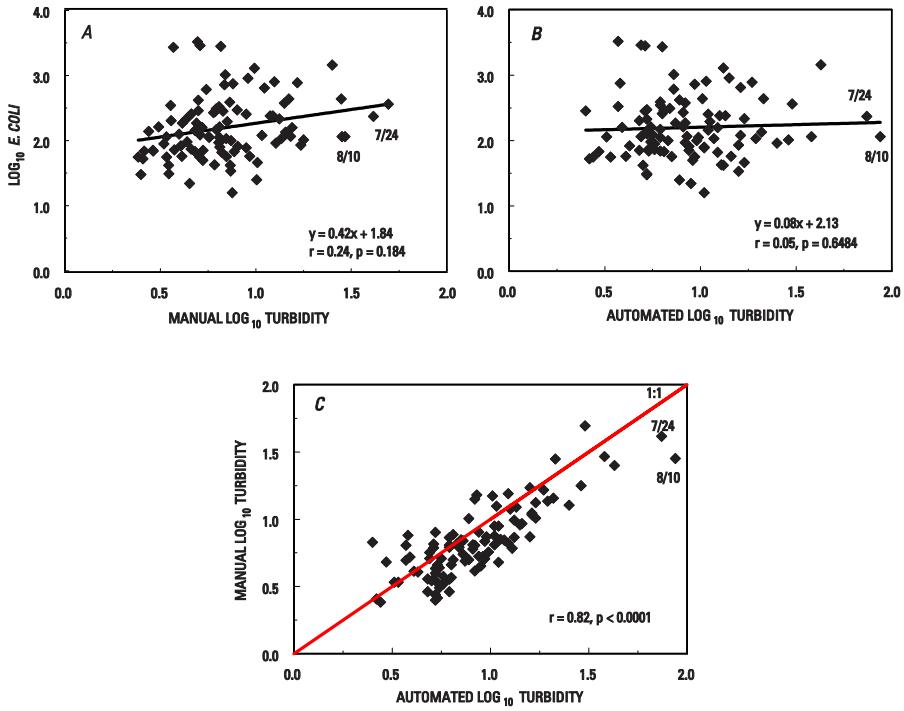

Figure 5. Turbidity measured by use of a grab sample and turbidimeter (manual) or probe attached to a nearshore buoy (automated), Edgewater, Cleveland, Ohio, 2008. Relations between: *A,* Manual turbidity and *E. coli* concentrations. B, Automated turbidity and *E. coli* concentrations. C, Automated and manual turbidity. (r is the Pearson's correlation coefficient, p is the significance of the relation.)

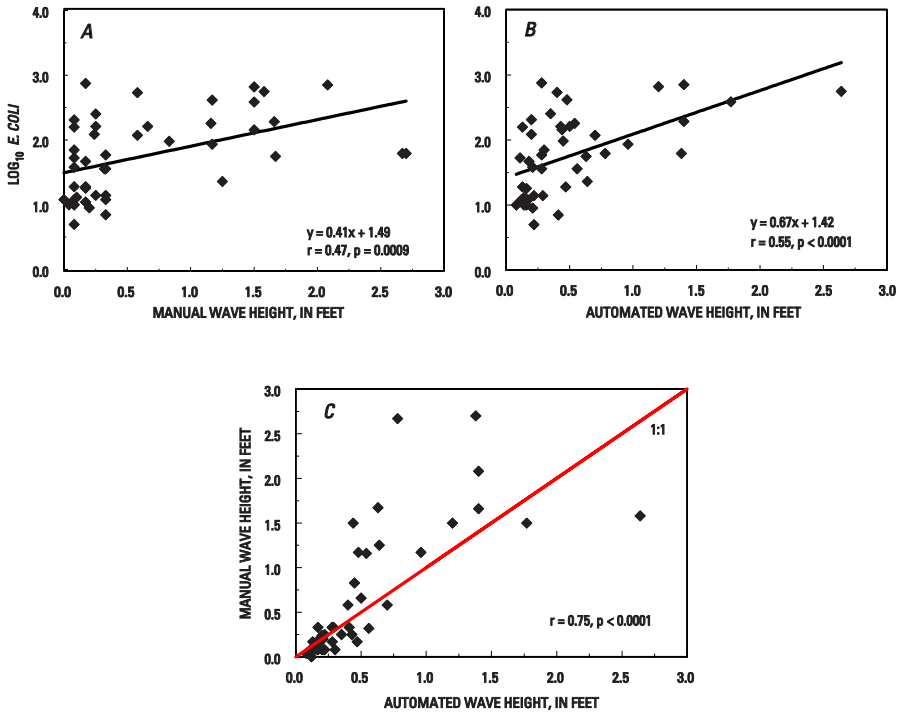

Figure 6. Wave heights measured by use of the wave rod (manual) or wave buoy (automated), Huntington, Bay Village, Ohio, 2008. Relations between: *A*, Manual wave heights and *E. coli* concentrations. *B*, Automated wave heights and *E. coli* concentrations. *C*, Automated and manual wave heights. (r is the Pearson's correlation coefficient, p is the significance of the relation.)

Afternoon sampling. Morning and afternoon *E. coli* concentrations for 46 paired samples are shown in figure 7 with four quadrant designations (A, B, C, and D) based on meeting and exceeding the single-sample bathing-water standard. For samples below the bathing-water standard in the morning (A and B), results for morning and afternoon samples were evenly spaced around the 1:1 line, except for two samples collected in the afternoon that exceeded the standard. In contrast, for samples above the standard in the morning (C and D), morning *E. coli* concentrations were generally higher than afternoon concentrations. Among the 10 samples above the standard in the morning or afternoon, 3 were in agreement in terms of exceeding the standard (C) and 7 were not (A and D); 5 out of 7 had exceedances in the morning, but not in the afternoon (D). On many of the seven days not in agreement, there were differences in turbidity or wave heights measured for morning and afternoon samples (data not shown).

Model Refinement and Selection for 2009

Explanatory variables consistently related to *E. coli* during 2008 and previous years were used for model refinement. These variables were the same for Huntington and Edgewater and included radar and Hopkins rainfall, manual wave heights, log turbidity, and day of the year. Automated wave height data were not used because the datasets were too small. Similarly, only one year of PAR and afternoon data were available at Huntington and only year of foreshore head data were available at Edgewater. For both beaches, data from 2005−8 were used to develop refined models because data collected before 2005 included categorical wave-height measurements (Francy and Darner, 2007), which were not as accurate as manual measurements collected with a wave rod. The best model for each subseason and for the recreational year combined (Appendix 1) was identified by multiple linear regression and model diagnostic techniques, and threshold probabilities were established as described by Francy and Darner (2006). Data segments for each subseason were defined by day of the year, and the cutoff date varied slightly from year to year because of leap years. Datasets with Hopkins rainfall data had slightly higher numbers than datasets with radar data because of missing values for the latter.

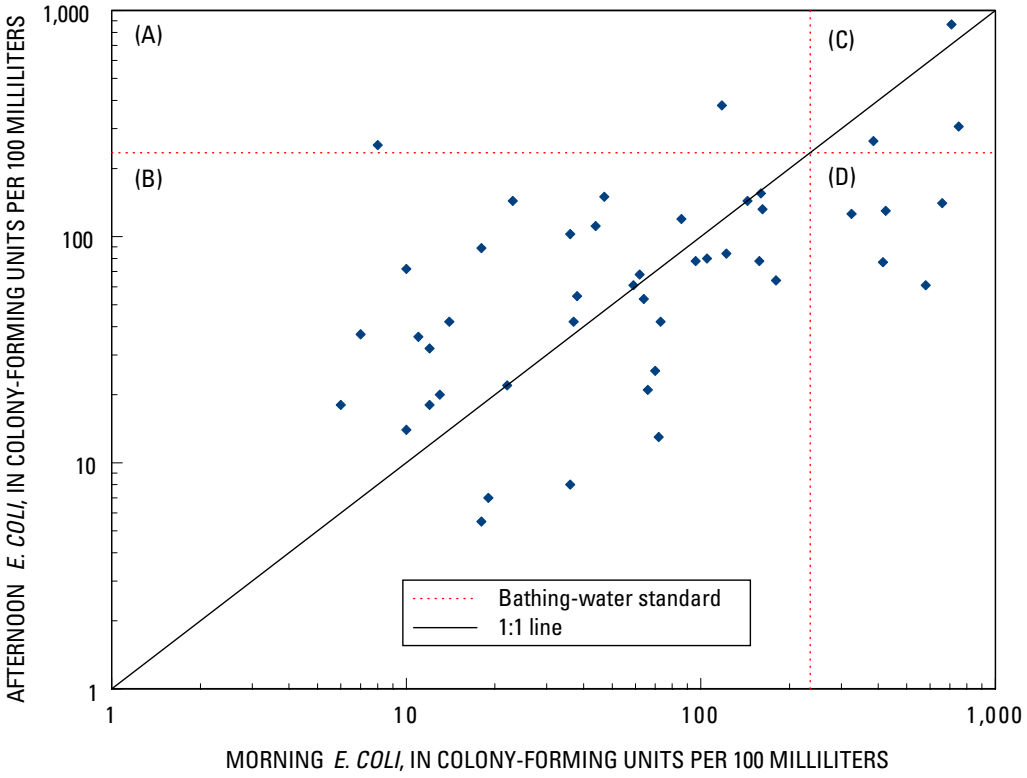

Figure 7. Comparison of morning and afternoon *Escherichia coli* concentrations at Huntington, Bay Village, Ohio, 2008. (Quadrant designations are based on meeting and exceeding the single-sample bathing-water standard of 235 colony-forming units per 100 milliliters.)

Edgewater

For Edgewater, a best model for the recreational year dataset was first identified; this model included Hopkins rainfall (table 5). Because of poor performance of the Edgewater models in the 2008 nowcast, the output from the recreational-year model was examined and split into three new segments, based on day of the year and occurrence of false negatives and positives (Francy and Darner, 2007): (1) Day of the year 134−159, May 13 through June 8 (“subseason 1”), (2) day of the year 160−222, June 9 through August 10 (“subseason 2”), and (3) day of the year 223–258, August 11 through September 15 (“subseason 3”). Candidate models were developed by use of radar or Hopkins data for subseasons 1, 2, and 3. Candidate models were evaluated by examining the numbers and percentages of correct responses, false positives, and false negatives for 2005−8 data (table 5). On the basis of percentage of correct responses, the best individual subseason models included radar rainfall for subseason 1 (81.0 percent) and Hopkins rainfall for subseasons 2 (81.5 percent) and 3 (76.9 percent). The summed responses from the best separate models for subseasons 1, 2, and 3 resulted in a greater percentage of correct responses (80.1) than for the recreational-year model (78.2 percent). Sensitivities were 8.6 percent higher for the summed model as compared to the recreational-year model, and specificities were about the same. The summed model resulted in eight fewer false negatives and two more false positives than the recreational-year model for 2005−8.

Huntington

Because of good performance of the Huntington models in the 2008 nowcast, the same data segments used in 2008 were retained for model refinement: (1) Day of the year 140−205, May 19 through July 23 (“subseason 1”) and (2) day of the year 206−249, July 24 through September 6 (“subseason 2”). Models were developed by use of radar or Hopkins rainfall for subseasons 1 and 2 and compared to the

Table 5. Examination of candidate Edgewater 2005–8 models.

[R^2, fraction of the variation of *E. coli* concentrations that is explained by the model; threshold probability is based on meeting and exceeding the single-sample bathing standard for *E. coli*; subseason 1 is day of the year 134-159 (May 13−June 8), subseason 2 is day of the year 160-222 (June 9−August 10), and subseason 3 is day of the year 223-258 (August 11−September 15); calculated sums are shaded below]

Model	Variables	R^2	Threshold probability	Number			Percentage		
				Correct	False +	False -	Correct	Sensitivity	Specificity
Subseason 1, 2, and 3 models with radar rainfall									
Subseason 1	Wave height, radar2cell-w48[a]	0.41	22	47	7	4	81.0	69.2	84.4
Subseason 2	Wave height, log turbidity, radar2cell-w48, day of the year	0.41	27	154	25	15	79.4	68.8	82.9
Subseason 3	Wave height, radar2cell-w48, day of the year	0.29	32	74	17	8	74.7	74.2	75.0
Subseason 1, 2, and 3 models with Hopkins rainfall									
Subseason 1	Wave height, Rw48[b]	0.27	23	48	8	5	78.7	64.3	83.0
Subseason 2	Wave height, log turbidity, Rw48, day of the year	0.40	26	167	23	15	81.5	68.8	85.4
Subseason 3	Wave height, Rw48, day of the year	0.28	33	80	17	7	76.9	78.1	76.4
Best models									
Subseason 1 (radar), and subseasons 2, 3 (Hopkins)				294	47	26	80.1	72.0	82.8
Recreational year	Wave height, log turbidity, Rw48[b], day of the year	0.37	30	283	45	34	78.2	63.4	83.3

[a]Radar2cell-w48 was the maximum amount of rainfall in the 48-hour period preceding sampling that fell in one of two 4-km grids, as determined from radar data, with the most recent rainfall receiving the most weight.

[b]Rw48 was the amount of rainfall, in inches, at Hopkins International Airport, Cleveland, Ohio, in the 48-hour period before sampling, with the most recent rainfall receiving the most weight.

best recreational-year model which included radar rainfall (table 6). Candidate models were further evaluated by examining the numbers and percentages of correct responses, false positives, and false negatives for 2005−8 data. The summed responses for separate models for subseasons 1 and 2 including radar or Hopkins rainfall resulted in a greater percentage of correct responses (86.1 and 87.1 percent, respectively) than for the recreational year model (82.4 percent). Sensitivities were at least 10 percent higher and specificities only slightly higher for the summed subseason 1 and 2 models as compared to the recreational-year model. In comparing the two summed models, the model with radar rainfall resulted in one more false negative, and the model with Hopkins provided one more false positive; the percentage correct for the summed Hopkins model was higher because the Hopkins rainfall dataset included 12 more days.

Table 6. Examination of candidate Huntington 2005–8 models.

[R^2, fraction of the variation of *E. coli* concentrations that is explained by the model; threshold probability is based on meeting and exceeding the single-sample bathing standard for *E. coli*; subseason 1 is day of the year 140–205 (May 21–July 23) and subseason 2 is day of the year 206–249 (July 24–Sept 6); combined subseason models are shaded below]

Model	Variables	R^2	Threshold probability	Number			Percentage		
				Correct	False +	False -	Correct	Sensitivity	Specificity
Subseason 1 and 2 models with radar rainfall									
Subseason 1	Log turbidity, radar6cell-24[a], day of the year	0.38	23	176	17	12	85.9	62.5	90.2
Subseason 2	Wave height, radar6cell-w48[b]	0.50	26	115	11	7	86.5	74.1	89.6
Subseason 1 plus subseason 2				291	28	19	86.1	67.8	90.0
Subseason 1 and 2 models with Hopkins rainfall									
Subseason 1	Log turbidity, R_{d-1}[c], day of the year	0.36	23	181	15	11	87.4	65.6	91.4
Subseason 2	Wave height, Rw48[d]	0.49	26	122	12	7	86.5	76.7	89.2
Subseason 1 plus subseason 2				303	27	18	87.1	71.0	90.6
Best recreational-year model									
Recreational year	Radar6cell-w48, log turbidity, wave height, day of the year	0.42	26	277	34	25	82.4	57.6	87.7

[a]Radar6cell-24 was the summed amount of rainfall in the previous 24-hour period that fell in six 4-km grids, as determined from radar data.

[b]Radar6cell-w48 was the summed amount of rainfall in the previous 48-hour period that fell in six 4-km grids, as determined from radar data, with the most recent rainfall receiving the most weight.

[c]R_{d-1} was the amount of rainfall, in inches, at Hopkins International Airport, Cleveland, Ohio, in the 24-hour period before sampling.

[d]Rw48 was the amount of rainfall, in inches, at Hopkins International Airport, Cleveland, Ohio, in the 48-hour period before sampling, with the most recent rainfall receiving the most weight.

Next Steps

Water-resource managers and public health officials at Huntington and Edgewater will determine which 2005−8 model(s) to use to issue advisories for the nowcast during 2009. As in previous years, daily performance of the nowcast will be monitored, and models will be validated and refined to improve predictions during subsequent years.

Acknowledgments

The authors acknowledge the following individuals for their hard work in sampling and running the nowcast—Jill Lis and Alex Brown of the Cuyahoga County Board of Health, and Sara English and Tom Vasel of the Northeast Ohio Regional Sewer District. Thanks are also extended to collaborators Mark Citriglia, Eva Hatvani, and Lester Stumpe at the Northeast Ohio Regional Sewer District.

References Cited

Boehm, A.B., Grant, S.B., Kim, J.H., Mowbray, S.L., McGee, C.D., Clark, C.D., Foley, D.M., and Wellman, D.E., 2002, Decadal and shorter period variability of surf zone water quality at Huntington Beach, California: Environmental Science & Technology, v 36, no. 18, p. 3885–3892.

Francy, D.S., and Darner, R.A., 2006, Procedures for developing models to predict exceedance of recreational water-quality standards at coastal beaches: U.S. Geological Survey Techniques and Methods 6–B5, 34 p.

Francy, D.S., and Darner, R.A., 2007, Nowcasting beach advisories at Ohio Lake Erie beaches: U.S. Geological Survey Open-File Report 2007–1427, 13 p.

Francy, D.S., Bushon, R.N., Bertke, E.E., Likirdopolus, C.A., Brady, A.M.G., Kephart, C.M., Stelzer, E.A., and Stoeckel, D.M., 2008, Quality-assurance/quality-control manual for the Ohio Water Microbiology Laboratory, accessed November 2008 at *http://oh.water.usgs.gov/micro_qaqc.htm*

Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources: U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. A3, accessed March 2006 at *http://pubs.er.usgs.gov/pubs/twri/twri04A3*

Ohio Department of Health, 2007, Bathing beach monitoring program, accessed November 2008 at *http://www.odh.ohio.gov/odhPrograms/eh/bbeach/beachmon.aspx*

Myers, D.N., Stoeckel, D.M., Bushon, R.N., Francy, D.S., and Brady, A.M.G., 2007, Fecal indicator bacteria: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chap. A7, section 7.1 (version 2.0), available from *http://pubs.water.usgs.gov/twri9A/*.

National Oceanic and Atmospheric Administration, Center for Operational Oceanic Products and Services [n.d.], NOS water level observation network Great Lakes, accessed November 2008 at *http://glakesonline.nos.noaa.gov/geographic.html*

National Oceanic and Atmospheric Administration, National Weather Service, 2002, National Weather Service, Eastern Region Headquarters home page, accessed November 2008 *http://www.erh.noaa.gov/*

U.S. Environmental Protection Agency, 2006, Method 1603—*Escherichia coli* in water by membrane filtration using modified membrane-thermotolerant *Escherichia coli* agar: Washington, D.C., EPA 821–R–06–011, 42 p.

View of Huntington Beach, Bay Village, Ohio. (Photograph by Robert A. Darner, U.S. Geological Survey.)

Appendix 1. Candidate Models Used To Predict *Escherichia coli* Concentrations at Two Lake Erie Beaches

[Water-resource managers will select from among these models to use the Ohio Nowcast during 2009]

Huntington Reservation (Bay Village, Ohio)

Subseason 1 model with radar rainfall

$$\text{Log } EC = -0.885 + 0.782(LOG\ TURBIDITY) + 0.171(RADAR6CELL\text{-}24) + 0.009(DAY\ OF\ THE\ YEAR)$$

or

$$EC = 0.130 + 6.061^{LOG\ TURBIDITY} + 1.481^{RADAR6CELL\text{-}24} + 1.022^{DAY\ OF\ THE\ YEAR}$$

Subseason 2 model with radar rainfall

$$\text{Log } EC = 1.449 + 0.307(WAVE\ HEIGHT) + 0.033(RADAR6CELL\text{-}W48)$$

or

$$EC = 28.116 + 2.027^{WAVE\ HEIGHT} + 1.080^{RADAR6CELL\text{-}48W}$$

Subseason 1 model with Hopkins rainfall

$$\text{Log } EC = -0.866 + 0.806(R_{D\text{-}1}) + 0.783(LOG\ TURBIDITY) + 0.010(DAY\ OF\ THE\ YEAR)$$

or

$$EC = 0.136 + 6.392^{RD\text{-}1} + 6.071^{LOG\ TURBIDITY} + 1.022^{DAY\ OF\ THE\ YEAR}$$

Susbseason 2 model with Hopkins rainfall $\text{Log } EC = 1.449 + 0.311(WAVE\ HEIGHT) + 0.181(Rw48)$

or

$$EC = 28.149 + 2.046^{WAVE\ HEIGHT} + 1.514^{Rw48}$$

Recreational-year model

$$\text{Log } EC = 0.465 + 0.167(WAVE\ HEIGHT) + 0.045(RADAR6CELL\text{-}W48) + 0.406(LOG\ TURBIDITY) + 0.003(DAY\ OF\ THE\ YEAR)$$

or

$$EC = 2.918 + 1.467^{WAVE\ HEIGHT} + 1.109^{RADAR6CELL48W} + 2.550^{LOG\ TURBIDITY} + 1.007^{DAY\ OF\ THE\ YEAR}$$

Edgewater Park (Cleveland, Ohio)

Subseason 1 model with radar rainfall

$$\text{Log } EC = 1.253 + 0.494(WAVE\ HEIGHT) + 0.910(RADAR2CELL\text{-}W48)$$

or

$$EC = 17.909 + 3.118^{WAVE\ HEIGHT} + 8.126^{RADAR2CELL\text{-}W48}$$

Subseason 2 model with radar rainfall

$$\text{Log } EC = 0.561 + 0.239(WAVE\ HEIGHT) + 0.215(RADAR2CELL\text{-}W48) + 0.335(LOG\ TURBIDITY) + 0.005(DAY\ OF\ THE\ YEAR)$$

or

$$EC = 3.641 + 1.734^{WAVE\ HEIGHT} + 1.641^{RADAR2CELL\text{-}W48} + 2.161^{LOG\ TURBIDITY} + 1.011^{DAY\ \mathrm{OF\ THE\ YEAR}}$$

Subseason 3 model with radar rainfall

$$\text{Log } EC = -0.799 + 0.294(WAVE\ HEIGHT) + 0.115(RADAR2CELL\text{-}W48) + 0.011(DAY\ OF\ THE\ YEAR)$$

or

$$EC = 0.159 + 1.966^{WAVE\ HEIGHT} + 1.304^{RADAR2CELL\text{-}W48} + 1.027^{DAY\ \mathrm{OF\ THE\ YEAR}}$$

Edgewater Park (Cleveland, Ohio)—Continued

Subseason 1 model with Hopkins rainfall

$$\text{Log } EC = 1.381 + 0.613(WAVE\ HEIGHT) + 0.635(Rw48)$$

or

$$EC = 24.031 + 4.103^{WAVE\ HEIGHT} + 4.316^{Rw48}$$

Subseason 2 model with Hopkins rainfall

$$\text{Log } EC = 0.615 + 0.274(WAVE\ HEIGHT) + 0.209(Rw48) + 0.278(LOG\ TURBIDITY) + 0.005(DAY\ OF\ THE\ YEAR)$$

or

$$EC = 4.123 + 1.878^{WAVE\ HEIGHT} + 1.617^{Rw48} + 1.898^{LOG\ TURBIDITY} + 1.011^{DAY\ \text{OF THE YEAR}}$$

Subseason 3 model with Hopkins rainfall

$$\text{Log } EC = -0.828 + 0.311\ (WAVE\ HEIGHT) + 0.089(Rw48) + 0.012(DAY\ OF\ THE\ YEAR)$$

or

$$EC = 0.148 + 2.046^{WAVE\ HEIGHT} + 1.227^{Rw48} + 1.027^{DAY\ \text{OF THE YEAR}}$$

Recreational-year model

$$\text{Log } EC = 0.726 + 0.237(WAVE\ HEIGHT) + 0.158(Rw48) + 0.401(LOG\ TURBIDITY) + 0.004(DAY\ OF\ THE\ YEAR)$$

or

$$EC = 5.325 + 1.725^{WAVE\ HEIGHT} + 1.440^{Rw48} + 2.516^{LOG\ TURBIDITY} + 1.009^{DAY\ \text{OF THE YEAR}}$$

In all these equations,

EC is the *Escherichia coli* concentration, in colony-forming units per 100 milliliters.

LOG TURBIDITY is defined as the $\log_{10}$ of the turbidity in Nephelometeric Turbidity Units.

RADAR6CELL-24 is the summed amount of rainfall in the previous 24-hour period that fell in six 4-km grids, as determined from radar data.

RADAR6CELL-W48 is the summed amount of rainfall in the previous 48-hour period that fell in six 4-km grids, as determined from radar data, with the most recent rainfall receiving the most weight.

DAY OF THE YEAR is the number representing the date beginning with 1 for January 1 and 365 or 366 for December 31 (the latter being a leap year).

WAVE HEIGHT is obtained by manual measurements at the time of sampling.

RADAR2CELL-W48 is the maximum amount of rainfall in the 48-hour period preceding sampling that fell in one of two 4-km grids, as determined from radar data, with the most recent rainfall receiving the most weight.

R_{D-1} is the total rainfall, in inches, at Hopkins International Airport, Cleveland, Ohio, in the 24-hour period before sampling.

RW48 is the amount of rainfall, in inches, at Hopkins International Airport, Cleveland, Ohio, in the 48-hour period before sampling, with the most recent rainfall receiving the most weight.