

Real-Time Assessments of Water Quality—Expanding Nowcasting Throughout the Great Lakes

The overall mission of this work is to provide science-based information and methods that will allow beach managers to more accurately make beach closure and advisory decisions, understand the sources and physical processes affecting beach contaminants, and understand how science-based information can be used to mitigate and restore beaches and protect the public.

The U.S. Geological Survey (USGS), in collaboration with many Federal, State, and local agencies and universities, has conducted research on beach health issues in the Great Lakes Region for more than a decade. The work consists of four science elements that align with the USGS Beach Health Initiative Mission: real-time assessments of water quality; coastal processes; pathogens and source tracking; and data analysis, interpretation, and communication. The ongoing or completed research for real-time assessments of water quality is described in this fact sheet.

Why was the Study Done?

- Fecal indicator bacteria (FIB), such as *Escherichia coli* (*E. coli*) and enterococci, are currently used in beach water-quality monitoring to determine whether to post a beach advisory or closing. Although not necessarily pathogenic themselves, FIB indicate the possible presence of pathogens (disease-causing microorganisms). Traditional culture methods for FIB can take up to 18 hours to obtain results so that advisory decisions are made based on the previous day's FIB concentration. Water quality can change rapidly at beaches in response to varying meteorological conditions and other factors. This may result in beach advisories that are issued too late to protect visitor health, or unnecessarily issued the next day, after the water quality has improved.

- Agencies that monitor beaches need tools that can provide real-time, reliable measures of recreational water-quality conditions. Rapid methods for FIB, such as the quantitative polymerase chain reaction (qPCR) molecular method included in new recreational water-quality criteria (U.S. Environmental Protection Agency, 2012a), are one such tool, but are not always practical or cost effective for inclusion in local monitoring programs. Predictive modeling is a tool for assessing water quality in real-time using easily measured environmental and (or) water-quality variables—such as turbidity (water clarity) or rainfall—

to make management decisions in 1 or 2 hours on whether to post beach advisories or closures. The USGS and collaborators have been investigating predictive modeling for more than a decade (Francy and Darner, 1998; Nevers and Whitman, 2005; Zimmerman, 2008). Nowcast systems that use predictive models to post advisories or closings are one means to provide real-time assessments of water quality.

What are Nowcasts?

Nowcasts are systems that inform the public of current bacterial water-quality conditions at beaches and are based on predictive models. A nowcast estimates conditions in real-time, similar to the way a weather forecast looks to the future. Nowcasts use environmental and water-quality variables that are easily and quickly measured, such as rainfall and turbidity (water clarity), as surrogates in beach-specific statistical models to estimate densities of bacterial indicators, such as *E. coli*, or to provide the probability that a State standard will be exceeded. At the time this study began in 2010, nowcasts were used for beach closure or advisory decisions (operational nowcasts) at only a few locations within the Great Lakes Region including three beaches in Ohio (<http://www.ohionowcast.info>), three beaches in Illinois (Lake County Health Department, 2013), and two beaches in Wisconsin (<http://www.wibeaches.us/>).

When and Where was the Study Done?

During the 2010–12 recreational seasons, the USGS, in cooperation with 23 local and State agencies, worked to refine existing operational nowcast systems at four beaches and to expand the use of operational predictive models at 45 beaches throughout the Great Lakes Region (fig. 1). These included beaches (with numbers of beaches in parentheses) in Illinois (8), Indiana (1), Michigan (4), New York (7), Ohio (12), Pennsylvania (6), and Wisconsin (11).

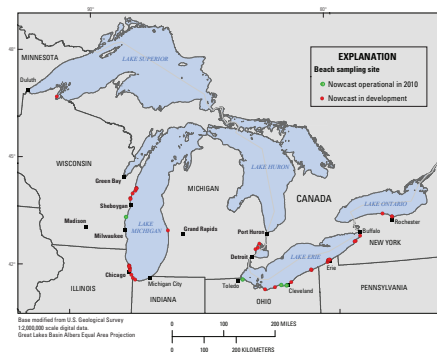


Figure 1. Locations of beach study sites.



Beach manager changing the nowcast sign before 9:30 a.m. at Edgewater Beach, Cleveland, Ohio.

How Were Data Collected and Compiled?

Local agencies measured *E. coli* concentrations and variables expected to affect *E. coli* concentrations including wave height, turbidity, water temperature, and numbers of birds at the time of sampling (fig. 2). In addition to field measurements, equipment was installed at or near several beaches to measure variables in real-time—nearshore buoys and local weather stations (fig. 3) and streamflow and water-quality gaging stations. Data from existing sources were compiled manually or by use of new tools developed by the USGS, which were designed specifically to compile and process data for predictive model development. These new tools included the following:

- The Lake Level spreadsheet that compiles hourly lake-level data from the National Oceanic and Atmospheric Administration (2012a).
- PROCESSNOAA, a software program that compiles and processes hourly rainfall, wind direction and speed, and barometric pressure data from the nearest airport site. Data are obtained from the National Oceanic and Atmospheric Administration and U.S. Department of Commerce (2012).
- Environmental Data Discovery and Transformation (EnDDaT), which retrieves and processes data from multiple sources including streamflow from the U.S. Geological Survey (2012); radar rainfall from the National Oceanic and Atmospheric Administration (2012b); and forecasted winds, waves, and currents from the National Oceanic and Atmospheric Administration (2012c).

Virtual Beach, a software package designed by the U.S. Environmental Protection Agency (EPA) for model development (U.S. Environmental Protection Agency, 2012b) was used to update or develop new models using multiple linear-regression techniques. The USGS worked with the EPA to develop and enhance the use of Virtual Beach, including the addition of an alternative statistical technique—partial least squares (Brooks and others, 2013)—into the Virtual Beach software.



Figure 2. U.S. Geological Survey employee measuring field variables at Presque Isle State Park Beach 2, Erie, Pennsylvania.



Figure 3. Weather station at Headlands State Park, Ohio (left), and nearshore buoy at Calumet Beach, Chicago, Illinois (right).

PROCESSNOAA for Compiling Weather Data

PROCESSNOAA, a graphical user interface (GUI), assists in processing large quantities of weather data provided by National Oceanic and Atmospheric Administration (NOAA). Hourly precipitation totals and predominant wind speed and direction data are routinely collected at airports across the country. These data often are used in predictive modeling at beaches but could require hours for the modeler to synthesize manually. PROCESSNOAA synthesizes the hourly data and saves these data in a format ready for direct import into Virtual Beach.

For further information on PROCESSNOAA, contact Jeremiah Lant (jglant@usgs.gov).

EnDDaT Tool for Compiling Data from Multiple Sources

A Web tool, Environmental Data Discovery and Transformation (EnDDaT), was developed to facilitate retrieval of data from multiple agencies and combine them into one location (<http://cida.usgs.gov/endedat/>). Currently, this tool includes access to national sources of data for variables including streamflow, radar precipitation, and forecasts of surface-current magnitude and direction, depth-integrated current magnitude and direction, wave height, wave direction, deviation from average lake level, surface-water temperature, and wind speed and direction.

For further information on the EnDDaT, contact Laura De Cicco (endedat@usgs.gov).

What Were the Study Results?

What Types of Variables Did the Models Contain?

Models were refined for existing nowcasts at 4 beaches, models were newly developed for 39 beaches, and data were collected for future predictive models at 6 beaches. The best model for each beach was based on a unique combination of explanatory variables including most commonly, turbidity, day of the year, wave height, wind direction and speed, antecedent rainfall for various time periods (prior to sampling), and change in lake level over 24 hours. Among all variables, turbidity (measured with a turbidimeter) was used most often (79.1 percent of models). Weather data from the nearest airport site were used in 44.2 percent of models. Among unique variables, a lab measurement of qPCR for *E. coli* was used at one beach in Ohio. Variables based on NOAA predictions, such as predicted current or wave direction, often were used in models for Wisconsin beaches. Local sources of data on rainfall, solar radiation, wind direction and speed, air temperature, humidity, and barometric pressure were used in 32.6 percent of models.

What Types of Information Does Each Model Provide for Beach Managers?

There are two types of output available from models: (1) the predicted *E. coli* concentration and (2) the probability of exceeding the single-sample standard or Beach Action Value (BAV). The BAV is a precautionary tool recommended for use in recently published criteria (U.S. Environmental Protection Agency, 2012a). The BAVs were not in use at the time of this study; however, the BAV of 235 colony-forming units per 100 milliliters (CFU/100 mL) is the same as the single-sample bathing-water standards used in all States in this study—except Michigan—for posting beach advisories or closings. At the time of this study, Michigan used 300 CFU/100 mL as their standard. In Illinois and Indiana, the predicted *E. coli* concentration was used as output from the models. At all other beaches, the probabilities of exceeding the single-sample standard were used as output from the models.

For those beaches where the probability of exceeding the single-sample

standard was used as output from the models, a threshold probability for issuing an advisory was established for each beach. A threshold probability is determined by taking the dataset used to develop the model and finding the probability that maximized the percentage of correctly predicted exceedances and non-exceedances of the standard, while at the same time, found a compromise between false positives and false negatives (Francy and Darner, 2006). Computed probabilities that are less than the threshold indicate that bacterial water quality is most likely acceptable for swimming. Computed probabilities equal to or greater than the threshold probability indicate that the water quality is most likely not acceptable and that a water-quality advisory may be needed. Established thresholds ranged from 15 to 50 percent, with a median of 28 percent.

How Well Did the Models Perform During an Independent Year?

Models were developed from data collected during 2010–11 for most beaches, and included pre-2010 data for a few beaches. Updated or new models for 42 beaches were validated during an independent year (2012); at one beach, the predictive modeling effort was discontinued. During validation, model performance was compared to the current

method for assessing recreational water quality—using the previous day's *E. coli* concentration—and to determine whether model results were good enough to be used for public notifications. Goals for good model performance were correct responses (predictions of exceedances and non-exceedances of the standard) that were at least 5 percent greater than the use of the previous day's *E. coli* concentration and overall correct responses >80 percent, sensitivities >50 percent, and specificities >85 percent.

- The sensitivity was the percentage of exceedances of the bathing-water standard that were correctly predicted by the model or the previous day's *E. coli* concentration.
- The specificity was the percentage of non-exceedances of the standard that were correctly predicted by the model or the previous day's *E. coli* concentration.

During validation in 2012, 24 models provided overall correct responses that were at least 5 percent greater than the use of the previous day's *E. coli* concentration. Predictive models met the above goals for overall correct responses, sensitivities, and specificities more often than the previous day's *E. coli* concentrations; this was especially apparent in terms of sensitivities (fig. 4).

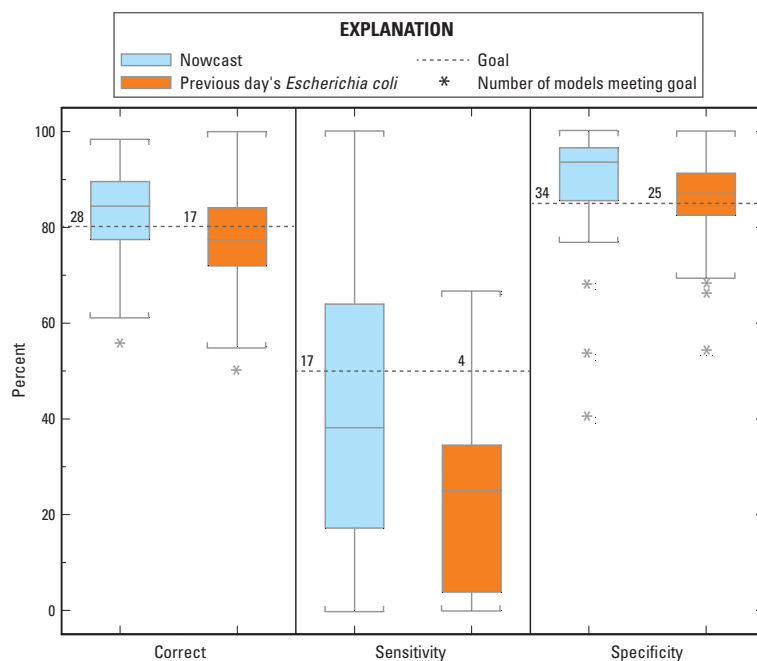


Figure 4. Percentages of overall correct, sensitivities, and specificities during validation of 42 beach-specific predictive models, 2012. Goals for good model performance are overall correct responses ≥ 80 percent, sensitivities ≥ 50 percent, and specificities ≥ 85 percent.

How is Real-Time Information Being Transferred to Beach Managers and Others?

Technology transfer is an important component of USGS activities. The USGS, EPA, and Wisconsin Department of Natural Resources worked together to conduct hands-on workshops in 2011 and 2013 at local USGS offices to train stakeholders on how to develop their own predictive models for use in nowcast systems. Participants had the opportunity to use data from their beach of choice and work through data compilation and analysis and model development, implementation, and validation procedures. Technology transfer also has included working with local agencies on a day-to-day basis to collect real-time data and apply data-management tools and Virtual Beach software to develop and test predictive models for beaches.

What is Next for USGS Real-Time Activities?

During 2013, the USGS will continue to work with local agencies to improve real-time predictions and implement more operational nowcast systems. Although models performed better than the current method to assess recreational water quality (previous day's *E. coli* concentration), work is ongoing to improve their accuracies. The USGS also will continue to work with the team of Virtual Beach developers to enhance model capabilities and provide technology transfer opportunities with stakeholders. The USGS is committed to helping beach managers develop and implement nowcast programs.

References

- Brooks, W.R., Fienen, M.N., and Corsi, S.R., 2013, Partial least squares for efficient models of fecal indicator bacteria on Great Lakes beaches: *Journal of Environmental Management*, v. 114, p. 470–475.
- Francy, D.S., and Darner, R.A., 1998, Factors affecting *Escherichia coli* concentrations at Lake Erie public bathing beaches: U.S. Geological Survey Water-Resources Investigations Report 98–4241, 41 p.
- Francy, D.S., and Darner, R.A., 2006, Procedures for developing models to predict exceedances of recreational water-quality standards at coastal beaches: U.S. Geological Survey Techniques and Methods 6–B5, 34 p.
- Lake County Health Department, 2013, SwimCastData, accessed June 2013 at <http://health.lakecountyil.gov/Population/LMU/Pages/SwimCast-Data.aspx>.
- National Oceanic and Atmospheric Administration Satellite and Information Service and U.S. Department of Commerce National Climatic Data Center, 2012, Unedited local climatological data, accessed January 2012 at <http://cdo.ncdc.noaa.gov/ulcd/ULCD>.
- National Oceanic and Atmospheric Administration, Center for Operational Oceanographic Products and Services, 2012a, Tides and currents, accessed January 2012 at <http://tidesandcurrents.noaa.gov/>.
- National Oceanic and Atmospheric Administration, National Weather Service, 2012b, Advanced Hydrologic Prediction Service, accessed January 2012 at <http://water.weather.gov/precip/download.php>.
- National Oceanic and Atmospheric Administration, Great Lakes Coastal Forecasting System, 2012c, accessed March 2012 at <http://www.glerl.noaa.gov/res/glcfs>.
- Nevers, M.B., and Whitman, R.L., 2005, Nowcast modeling of *Escherichia coli* concentrations at multiple urban beaches of southern Lake Michigan: *Water Research*, v. 39, no. 20, p. 5250–5260.
- U.S. Environmental Protection Agency, 2012a, Recreational water quality criteria: Washington, D.C., Office of Water, EPA-820-F-12-058, 63 p., accessed December 2012 at <http://water.epa.gov/scitech/swguidance/standards/criteria/health/recreation/index.cfm>.
- U.S. Environmental Protection Agency, 2012b, Exposure assessment models—Virtual Beach: Center for Exposure Assessment Modeling, accessed May 2013 at <http://www2.epa.gov/exposure-assessment-models/virtual-beach-vb>.
- U.S. Geological Survey, 2012, National Water Information System—Web interface, USGS water data for the Nation, accessed January 2012 at <http://waterdata.usgs.gov/nwis/>.
- Zimmerman, T.M., 2008, Modeling to predict *Escherichia coli* at Presque Isle Beach 2, City of Erie, Erie County, Pennsylvania: U.S. Geological Survey Scientific Investigations Report 2008–5039, 13 p.

Program Information

Funding for USGS beach projects and research in the Great Lakes comes from a variety of sources including the Ocean Research Priority Plan, the USGS, the Great Lakes Restoration Initiative, the EPA, the NOAA Center for Great Lakes and Human Health, and many State and local partner agencies and organizations throughout the region.

Contact

James R. Morris, Director
USGS Michigan and Ohio
Water Science Centers
614–430–7700
<http://greatlakesbeaches.usgs.gov/>