BEACH ADVISORIES: IMPROVING WATER QUALITY NOTIFICATION WITH REGRESSION MODELING AND NHDPLUS

Michele Cutrofello (presenter), RTI International; Timothy Bondelid, Consulting Engineer; Breda Munoz, RTI International
RTI International
701 13th Street NW, Suite 750
Washington, DC 20005

ABSTRACT

Beach advisories/closures are used to protect public health because elevated bacteria and pathogen levels can cause serious health effects; however, the economic impacts of unnecessary beach closures can be substantial. In an effort to protect human health and reduce the number of unnecessary beach closures, researchers have investigated methods to predict, without direct measurement, when bacteria concentrations will exceed standards. Statistical models completed for U.S. coastal beaches use explanatory variables based on meteorological, water quality, and beach data to predict bacteria concentrations. Although developments in statistical modeling for estimating bacteria levels have improved prediction capabilities, we believe the false negative and false positive rates can be further decreased.

We used multivariate linear regression to improve probabilistic predictive models by incorporating watershed information from the National Hydrography Dataset Plus (NHDPlus) to create a watershed-based spatial modeling framework, without analytical data, for beach advisories. The research was designed to improve, not duplicate, existing approaches that consider only beach characteristics (e.g., turbidity, number of birds, chlorophyll). The method allows for examination of pathogen source areas within the watershed and impacts of flow over a range of magnitudes from the surface water contributions to the beach.

We chose East Beach in Santa Barbara County, CA, as a test site. Using new watershed-based variables and readily available marine and meteorological data, we created models to estimate beach bacteria concentrations. Overall, these models produced an 18% false negative rate and a 3% false positive rate (out of 532 estimations), statistical levels that are on par with, or better than, current regression models created from intensively sampled and mined data. Probabilistic thresholds placed around the model estimates allow further insight and improvement of the model prediction capabilities for advisories. Using the multiyear data set, we also created models for 3-year increments to inspect the temporal variability of regression modeling, and discuss the false positive and negative rates resulting from applying each incremental model to the entire data set. Each of these four models (one for the entire data set and three 3-year incremental models) highlighted mixed forest lands near the beach as a possible source of pathogens to the beach waters. Future work will extend the method to multiple test sites to investigate NHDPlus characteristics that may be used to explain geographic variability in regression modeling at beaches. We also plan to further explore statistical methods and evaluations.

KEYWORDS

Bacteria, regression, NHDPlus, beach advisories, beach closures, pathogen