WATER RESEARCH 67 (2014) 105–117

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/watres

Predicting water quality at Santa Monica Beach: Evaluation of five different models for public notification of unsafe swimming conditions

W. Thoe [a,*], M. Gold [b], A. Griesbach [c], M. Grimmer [c], M.L. Taggart [c], A.B. Boehm [a]

[a] Department of Civil and Environmental Engineering, Environmental and Water Studies, Stanford University, Stanford, CA 94305, USA

[b] Institute of the Environment and Sustainability, University of California, Los Angeles, CA 90095, USA

[c] Heal the Bay, Santa Monica, CA 90401, USA

ARTICLE INFO

Article history:
Received 8 April 2014
Received in revised form
19 July 2014
Accepted 1 September 2014
Available online 16 September 2014

Keywords:
Coastal beach
Water quality
Statistical model
Performance evaluation

ABSTRACT

Bathing beaches are monitored for fecal indicator bacteria (FIB) to protect swimmers from unsafe conditions. However, FIB assays take ~24 h and water quality conditions can change dramatically in that time, so unsafe conditions cannot presently be identified in a timely manner. Statistical, data-driven predictive models use information on environmental conditions (i.e., rainfall, turbidity) to provide nowcasts of FIB concentrations. Their ability to predict real time FIB concentrations can make them more accurate at identifying unsafe conditions than the current method of using day or older FIB measurements. Predictive models are used in the Great Lakes, Hong Kong, and Scotland for beach management, but they are presently not used in California – the location of some of the world's most popular beaches. California beaches are unique as point source pollution has generally been mitigated, the summer bathing season receives little to no rainfall, and in situ measurements of turbidity and salinity are not readily available. These characteristics may make modeling FIB difficult, as many current FIB models rely heavily on rainfall or salinity. The current study investigates the potential for FIB models to predict water quality at a quintessential California Beach: Santa Monica Beach. This study compares the performance of five predictive models, multiple linear regression model, binary logistic regression model, partial least square regression model, artificial neural network, and classification tree, to predict concentrations of summertime fecal coliform and enterococci concentrations. Past measurements of bacterial concentration, storm drain condition, and tide level are found to be critical factors in the predictive models. The models perform better than the current beach management method. The classification tree models perform the best; for example they correctly predict 42% of beach postings due to fecal coliform exceedances during model validation, as compared to 28% by the current method. Artificial neural network is the second best model which minimizes the number of incorrect beach postings. The binary logistic regression model

* Corresponding author.
E-mail addresses: awthoe@stanford.edu, awthoe@gmail.com (W. Thoe).
http://dx.doi.org/10.1016/j.watres.2014.09.001
0043-1354/© 2014 Elsevier Ltd. All rights reserved.

also gives promising results, comparable to classification tree, by adjusting the posting decision thresholds to maximize correct beach postings. This study indicates that predictive models hold promise as a beach management tool at Santa Monica Beach. However, there are opportunities to further refine predictive models.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Epidemiological studies show that swimming in fecal contaminated waters may result in gastrointestinal and respiratory diseases (Prüss, 1998; Haile et al., 1999). To protect bathers from swimming in polluted waters, bathing beaches are typically monitored for fecal indicator bacteria (FIB), such as enterococci and *Escherichia coli*. Traditional methods to detect FIB take 18–24 h; even if rapid detection methods, such as qPCR (Noble et al., 2010), are used, there is still at least a 6-hour lag time between exposures of swimmers and public notification of beach water quality. Ample evidence has shown that waterborne FIB concentrations change in a matter of days or even hours (Boehm et al., 2002). Beach management based on out-dated sampling results can lead to contaminated beaches left open and clean beaches being posted or closed (Kim and Grant, 2004; Whitman and Nevers, 2004). A recent epidemiology study by Colford et al. (2012) showed that swimmer illness is associated with FIB concentrations measured on the same day but not the day before, further substantiating the need for better beach water quality warning systems.

Beach water quality prediction via a modeling approach for beach management was first introduced by the World Health Organization (WHO) (WHO, 2003). The approach has been incorporated by the European Union in the Bathing Water Directive (European Parliament, 2006) to advise the public against bathing during short term pollution. In 2012, The U.S. Environmental Protection Agency (USEPA) introduced new recreational water quality criteria and endorsed beach water quality predictive modeling as a rapid and inexpensive tool to reduce beach management errors due to the time lag in FIB measurement (USEPA, 2011). Exploratory studies on beach water quality modeling have been carried out worldwide since the early 2000s. These studies typically adopt statistical, data-driven models with hydro-meteorological factors (i.e., rainfall, solar radiation, tide level) as input variables. Multiple linear regression (MLR) is the most widely used model type; it has been tested at beaches in lacustrine (Olyphant and Whitman, 2004; Frick et al., 2008; Nevers and Whitman, 2008; Francy, 2009) and marine (Crowther et al., 2001; Boehm et al., 2007; Gonzalez et al., 2012; Thoe et al., 2012) coastal environments. Other common modeling methods include partial least square regression (Hou et al., 2006; Brooks et al., 2013) and artificial neural networks (Lin et al., 2003; He and He, 2008; Zhang et al., 2012). Categorical models such as decision trees have also been used in some studies (Parkhurst et al., 2005; Boehm et al., 2007; Bae et al., 2010; Stidson et al., 2012).

Predictive models have been successfully applied as management tools in the US Great Lakes (Francy, 2009; Francy et al., 2013), and in Scotland (McPhail and Stidson, 2009; Stidson et al., 2012). A pilot beach water quality model has also been developed in Hong Kong (Thoe and Lee, 2013). In general, predictive models are found to out-perform traditional beach monitoring to capture beach pollution primarily as the latter relies only on outdated FIB measurements. Past studies usually considered only one specific modeling tool, or compared two types of models (e.g., MLR versus ANN) (Mas and Ahlfeld, 2007; Thoe et al., 2012). The strengths and weaknesses of different models at different types of beaches have not been fully addressed. Additionally, there have not been general guidelines to choose the most appropriate type of model, limiting extensive application of predictive tools.

California has some of the most famous beaches in the world. Every year, over 150 million visits are made to Californian beaches, generating over 14 billion USD (Pendleton and Kildow, 2006). In 1997, the California State Legislature passed Assembly Bill 411 (AB411) (CDHS, 1997) which requires monitoring of bathing water at frequently visited beaches adjacent to flowing storm drains and creeks for enterococci (ENT), fecal coliform (FC) and total coliform (TC) during the bathing summer season (April–October). Predictive models are not used for beach management in California, except where the Scripps Institute of Oceanography uses CODAR to provide Tijuana River plume fate and transport information to the City of Imperial Beach in San Diego County (http://www.sccoos.org/data/tracking/IB/). Californian beaches are unique among other coastal beaches. Point pollution sources like wastewater treatment plant discharges have been identified and mitigated at many California beaches, and the bathing season (April–October) corresponds to the dry season. Additionally, in situ measurements of salinity and turbidity, which have been found important in other locations for model development (Nevers and Whitman, 2008; Gonzalez et al., 2012; Thoe et al., 2012) are not typically available for California beaches. Therefore, modeling experiences elsewhere cannot be directly applied to California. A few independent studies have been carried out at specific California beaches using a particular type of predictive model (Hou et al., 2006; Boehm et al., 2007; He and He, 2008; Bae et al., 2010), but it is still unknown if there exists one particular type of model that performs the best at California beaches with different pollution characteristics.

This study provides a comprehensive performance evaluation of five different statistical, data-driven predictive models to predict ENT and FC concentrations at Santa Monica Beach in the summer bathing season. Santa Monica Beach is one of the most visited beaches in California (Morton and Pendleton, 2001). The beach was reported to have the highest excess gastrointestinal illnesses among 28 beaches in Los

Download English Version:

https://daneshyari.com/en/article/6366596

Download Persian Version:

https://daneshyari.com/article/6366596

Daneshyari.com