



Predictive Model for Fecal Bacteria Contamination at Rincón Public Beach, Puerto Rico

Priscila Vargas Babilonia^{1,2,3} and Sylvia Rodríguez-Abudo^{2,3}

¹Department of Civil Engineering and Surveying, University of Puerto Rico at Mayagüez

²UPRM Center for Applied Ocean Sciences and Engineering, Department of Engineering Science and Materials, University of Puerto Rico at Mayagüez

³Caribbean Coastal Ocean Observing System, University of Puerto Rico at Mayagüez



CARICOOS

Abstract

Beach water quality has become a frequent problem in Puerto Rico. Within the last year, 85% of the beaches sampled by the Environmental Quality Board exceeded, at least once, the allowable limits for fecal indicator bacteria (FIB). The current operational strategy to assess swimming conditions consists of detecting FIB from in situ collected samples. This process requires at least 18 hours of incubation, which by the time a result is produced may not be representative of the actual state of the site. Furthermore, due to the laborious nature of this process, the common sampling frequency is scarcely once a week, which does not detect abrupt events with periods of hours to days. This project focuses on developing a statistical model for the presence of FIB at Rincón Public Beach using existing site-specific meteorological, oceanographic, and microbial data as input for EPA's Virtual Beach. Results show 3-hr rain accumulation consistently correlating with the \log_{10} of bacteria colony forming units (CFUs), suggesting that surface water runoff may play a crucial role on beach water quality. A multiple linear regression model using rain accumulation, water levels, wind, tides, and wave characteristics was developed. This model provides daily nowcasts of beach water quality of Rincón Public Beach at CariCOOS web portal.

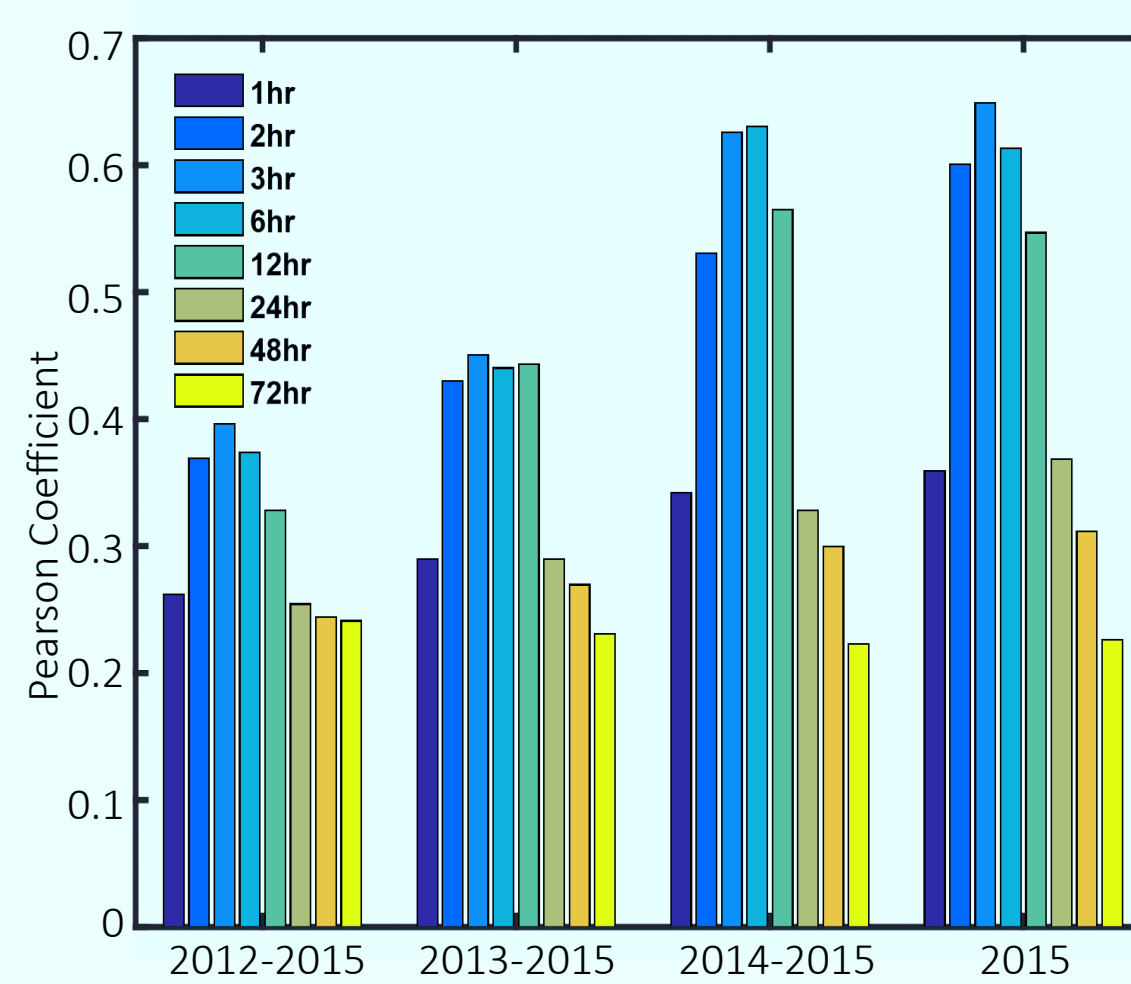
Study Area



Model Development

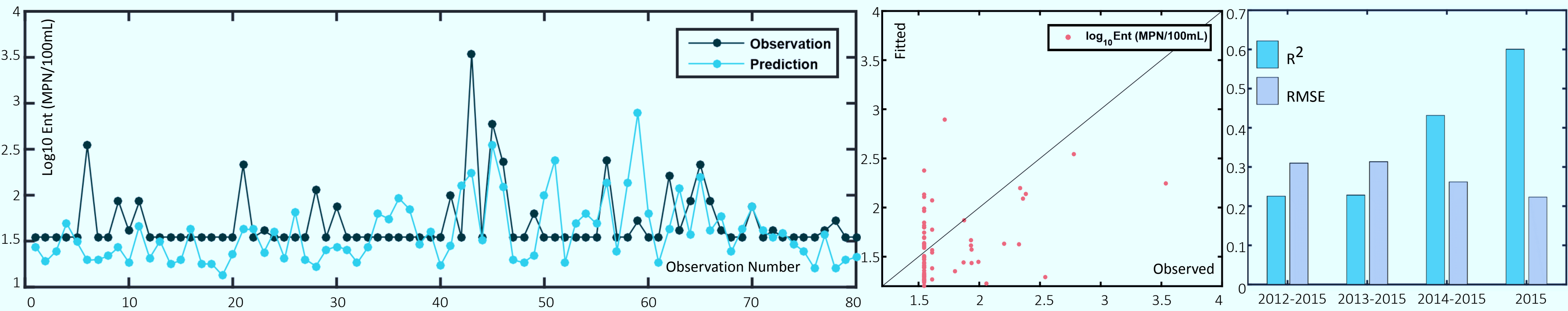
EPA's Virtual Beach 3.0.4 Input Variables	
Independent Variables	Source of Variables
Accumulated Precipitation	Weather Underground STA KPRRINCO3
Wave Height & Direction	CariCOOS Rincón Waverider Buoy
Wind Speed & Direction	PTRP4 CariCOOS Meteorological Station
Tide Predictions	Tidal Model Driver (TMD) Atlantic Ocean

EPA's Virtual Beach software was used to develop a site-specific statistical model for the prediction of pathogen indicator levels at this recreational beach. With historical data used as input, the correlation between independent variables and levels of bacteria was identified, followed by the development of a multiple linear regression model.



Correlations between independent and dependent variables were tested for four sampling periods (2012-2015, 2013-2015, 2014-2015, 2015) and various statistical quantities (averages / medians / accumulation over 1, 2, 3, 6, 12, 24, 48, and 72 hours, depending on the nature of the variable). The highest correlating statistics were chosen as input to generate the model (see left for rainfall example). Collinearity between the independent variables was automatically checked by the software package. An equation for the model was then generated based on the highest correlating transformed variables and their p-value.

Model Results



(Left) Observations and predictions based on the 2015 model. (Middle) Predictions vs observations based on the 2015 model. (Right) R^2 and RMSE of models generated for different data periods.

$$\log_{10} Ent = \beta_0 + \Delta\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4 + \beta_5x_5 + \beta_6x_6$$

Variable	$\beta \pm \epsilon$	Transformation	P-value
β_0 Intercept	-7.30 ± 2.28	-	0.0021
$\Delta\beta_0$ Offset	$-\log_{10}(1.12)$	-	-
x_1 Accumulated Precipitation	$0.68 \pm .08$	$x^{1/4}$	2.06×10^{-13}
x_2 Wind Speed	$-0.12 \pm .06$	$1/x$	0.0284
x_3 Wind Direction	1.43 ± 0.55	$1.642 + 0.002x - 3.314 \times 10^{-6}x^2$	0.0108
x_4 Wave Height	2.66 ± 1.67	$1.820 - 0.261x - 0.114x^2$	0.0148
x_5 Wave Direction	$0.22 \pm .06$	$1/x$	0.0005
x_6 Tides	$1.18 \pm .40$	$1.682 - 1.728x + 12.767x^2$	0.0042

Evaluation Criteria	
Accuracy	0.6182
Sensitivity	0.6182
Specificity	0.6182
RMSD	0.2230
R^2	0.6004

Ability to predict positive values:
Sensitivity = $tp / (tp + fn)$

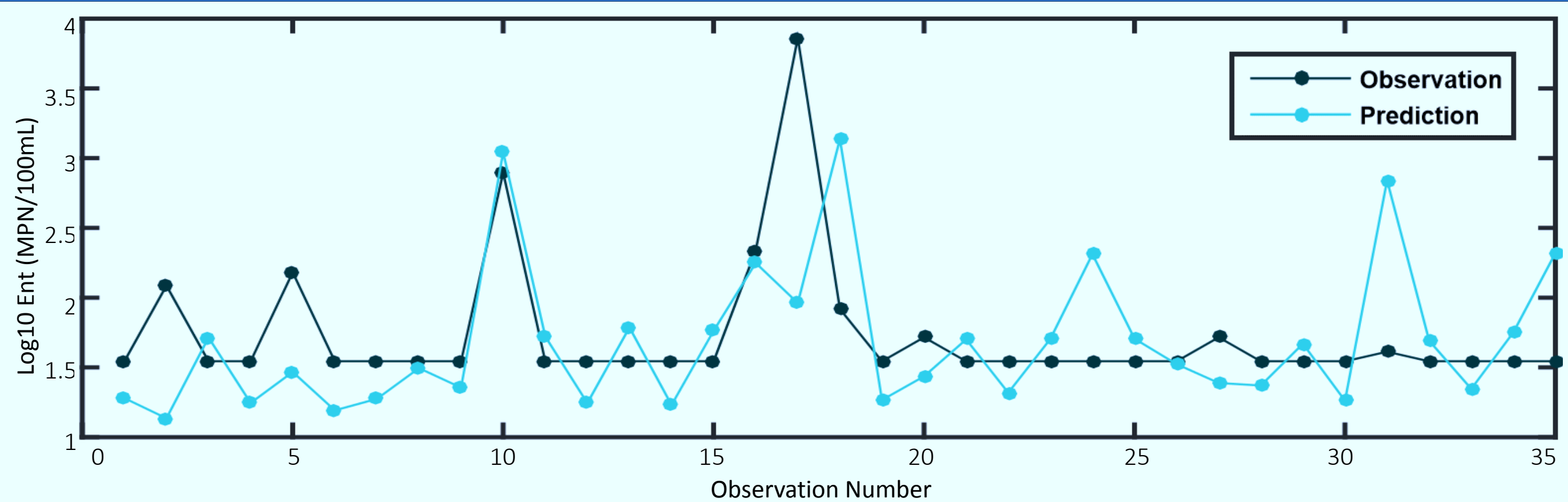
Ability to predict negative values:
Specificity = $tn / (tn + fp)$

Ability to predict overall:
Accuracy = $(tp + tn) / n$

*n = number of observations
*tp = true positives
*tn = true negatives
*fp = false positives
*fn = false negatives

Validation

Evaluation Criteria	
Accuracy	0.5429
Sensitivity	0.5556
Specificity	0.5385



Summary & Ongoing Work

A predictive model for fecal indicator bacteria at Rincon Public Beach was developed. The model uses rainfall, wind, tide, and wave data to provide MPN/100ml of bacteria in seawater. The model shows to be appropriately sensitive, specific, and accurate. This model provides daily nowcasts of beach water quality of Rincón Public Beach at CariCOOS web portal. Further validation and fine-tuning ongoing.

References and Acknowledgements

Cyterski, Mike. (2015). Virtual Beach 3.0.4: User's Manual. U.S. EPA. Predictive Tools for Notification Volume I: Review and Technical Protocol. (2010). U.S. EPA.

Zepp, G. Richard, (2010). Predictive Modeling at Beaches Volume II: Predictive Tools for Beach Notification. U.S. EPA.

This project was funded by NOAA IOOS (Grant #NA11NOS0120035). We thank Steve Tamar from the Surfrider Foundation for providing us with weekly water quality samples in our study site.

