Empirical Modeling of Microbial Indicators at a South Carolina Beach

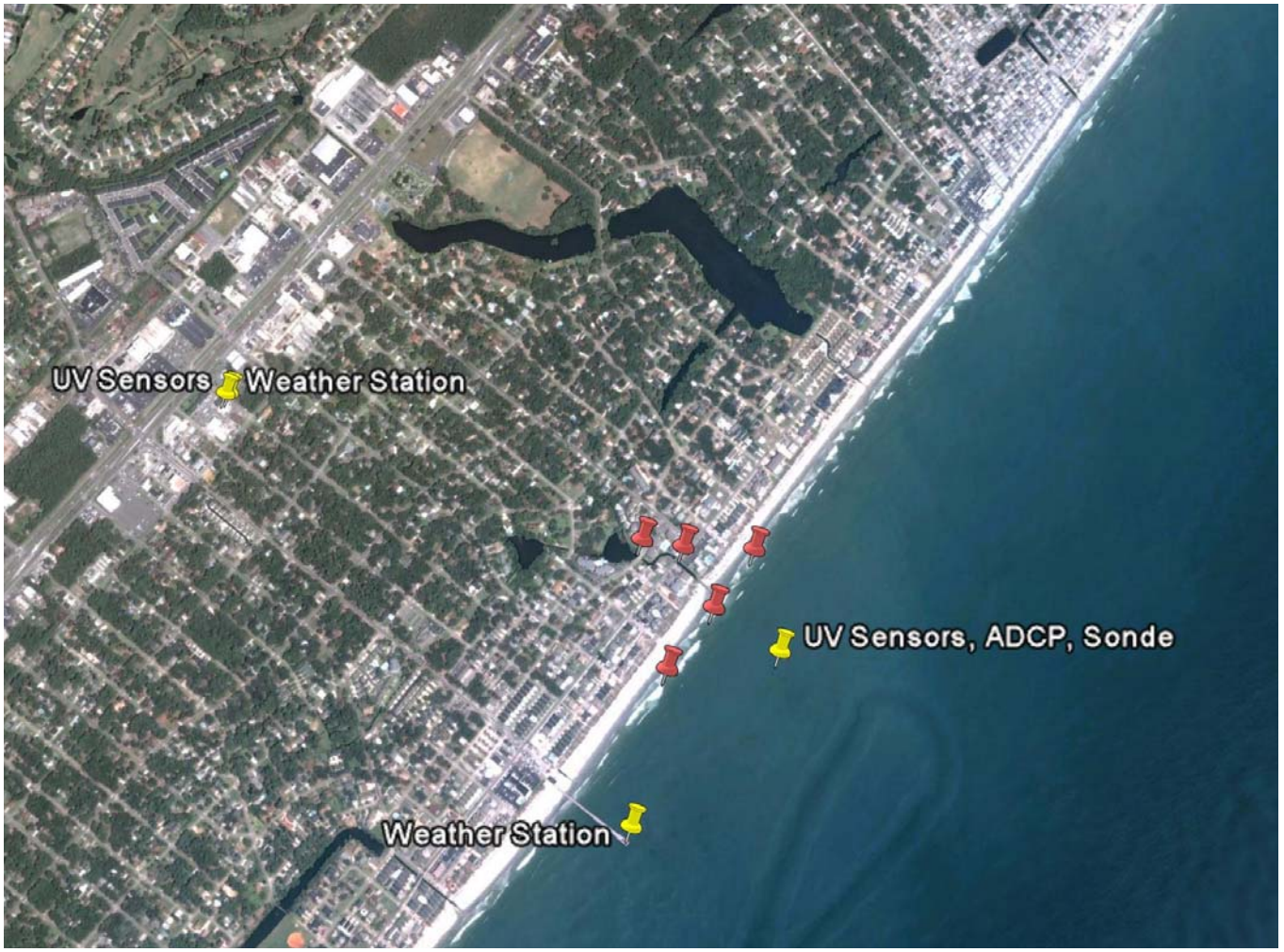
R. Zepp[1], M.Cyterski[1], E.White[1], M. Molina[1], K.Wolfe[1], R. Parmar[1]
S. Zhang[2], S. Hunter[2], E. Duvall[2], O. Tedrow[2]
C. Fitzpatrick[2], E. Brouillette[2]; C. Sloan[4]

[1]U.S. EPA, NERL/ERD, Athens GA;
[2]Student Services, U.S. EPA, Athens GA

Objectives

- General Objective: Refine and evaluate procedures for building water quality models used for notification and advisories/closures
- Specific Objectives: (1) Compare and contrast the effectiveness of empirical models developed using onsite and near-site datasets for Surfside Beach, SC; (2)Compare models for culturable enterococci and enterococci qPCR

Surfside Beach Monitoring Locations




06.13.2009 14:20


06.13.2009 10:14


06 13 2009 14:18

Study Approach

- Obtain concurrent fecal indicator bacteria (FIB) concentrations and independent variables at or near Surfside Beach
- Use these data to build empirical models for prediction of indicator concentrations and criteria exceedances

Microbial Data

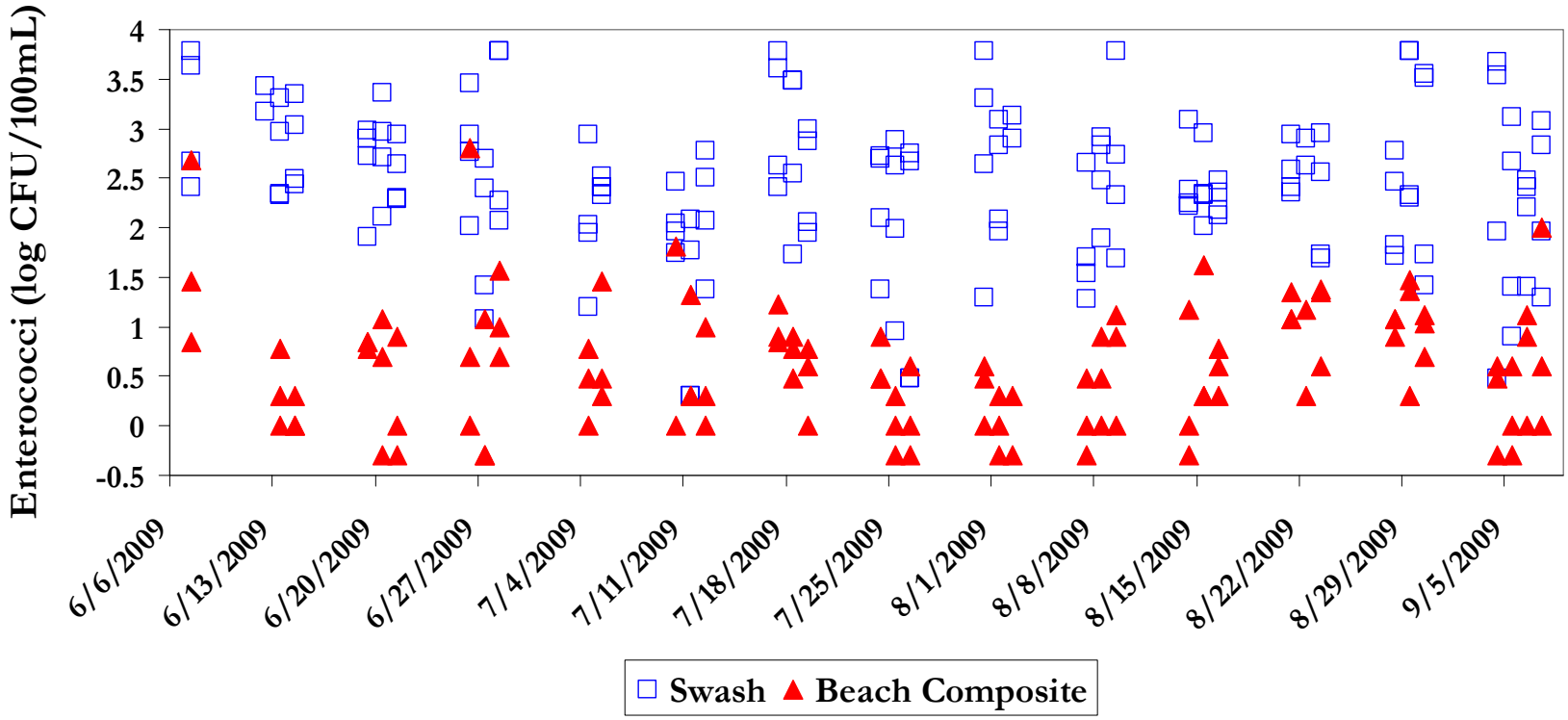
- Contract (via Westat) provided microbial and water sampling:3 transects sampled 3 times/day, 3 days week (Friday and weekends) over 3 month period during summer 2008; contaminated sites in nearby swash
- Culturable and qPCR techniques used

Microbial Sampling

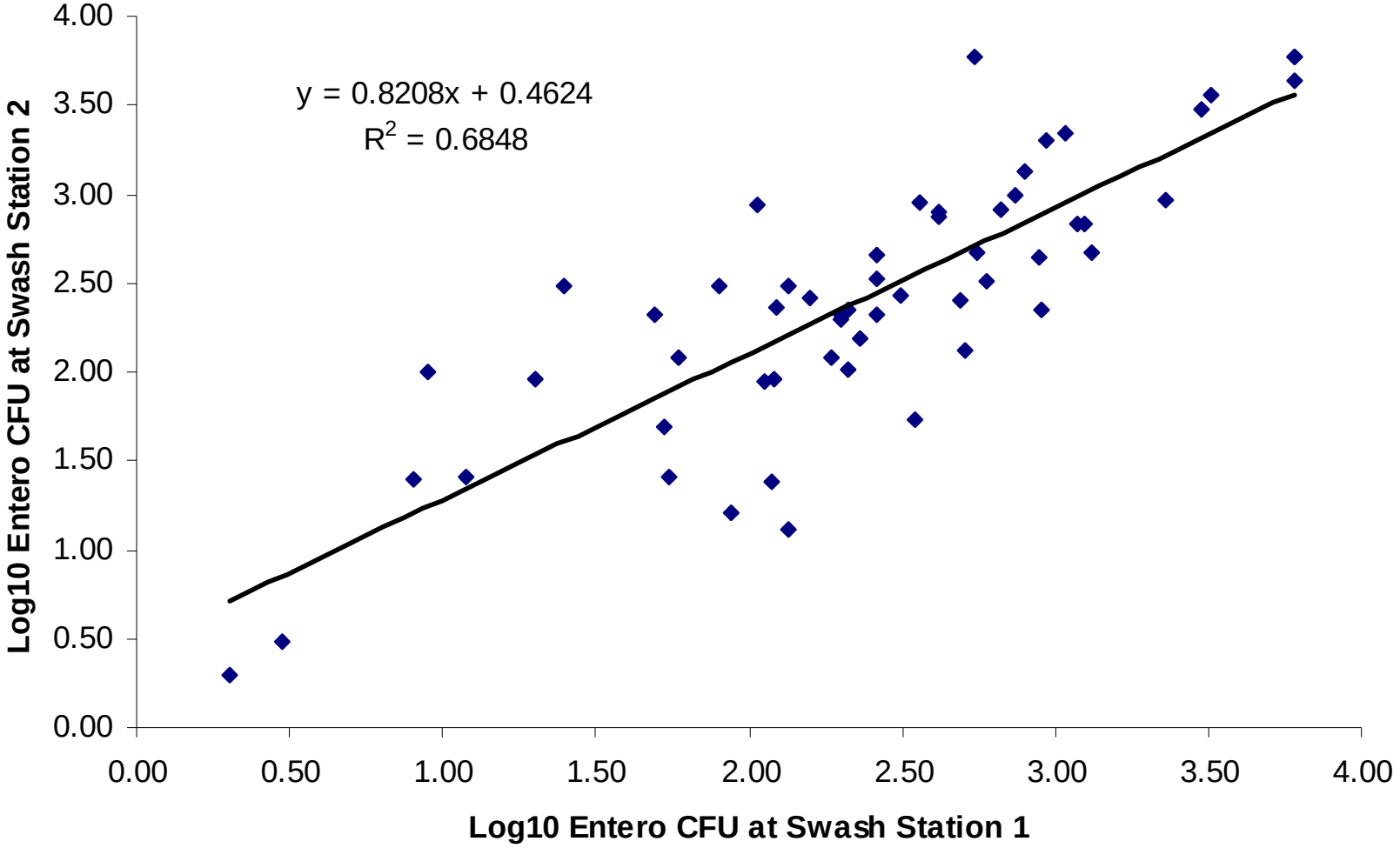

06 13 2009 08:24


06.13.2009 10:10

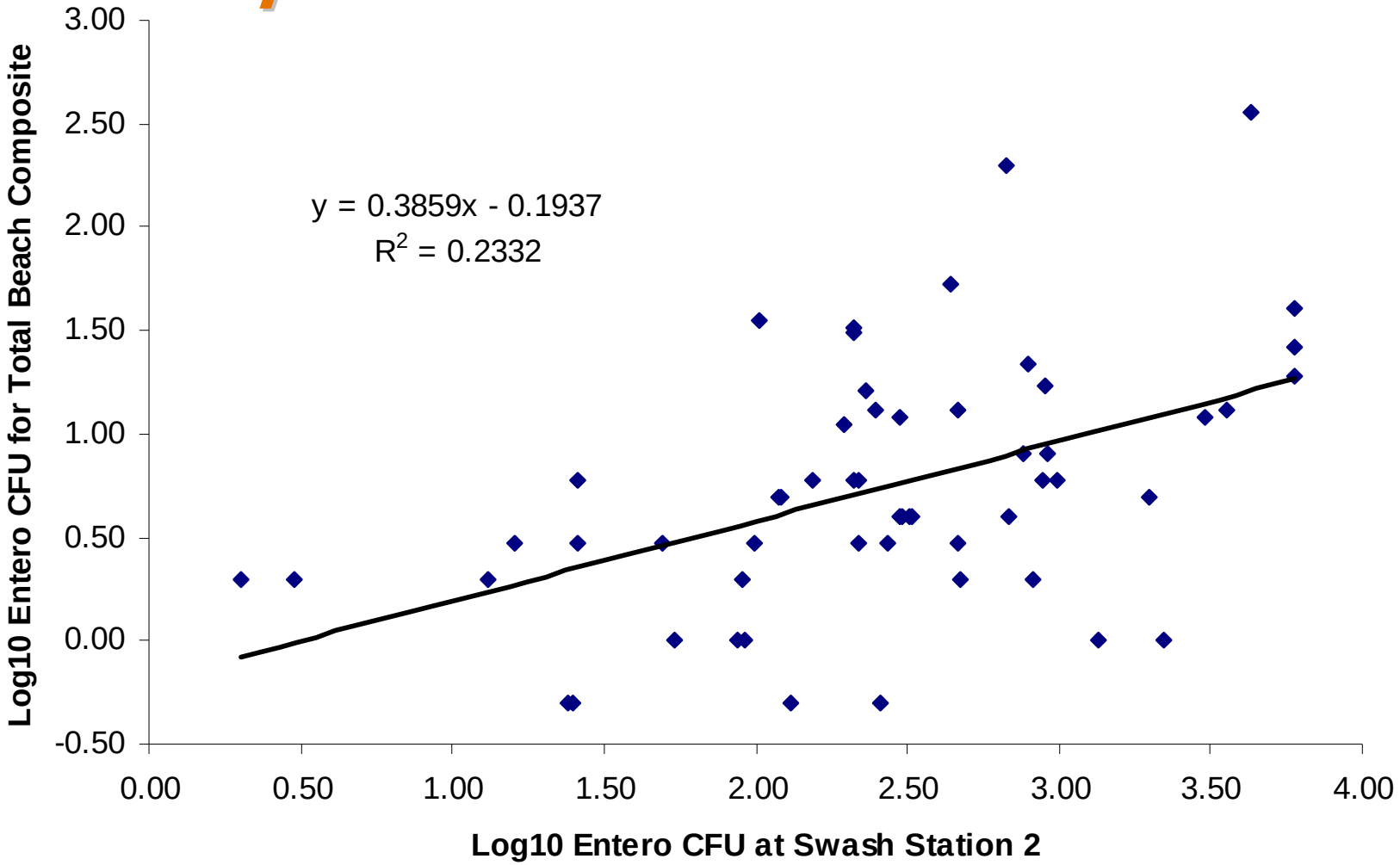
Enterococci Counts During Summer 2009 Swash and Beach Composite (Waist Only) Samples



Culturable Enterococcus Concentrations At Two Points in Swash



Culturable Enterococcus in Swash Compared to Swim Area at Beach



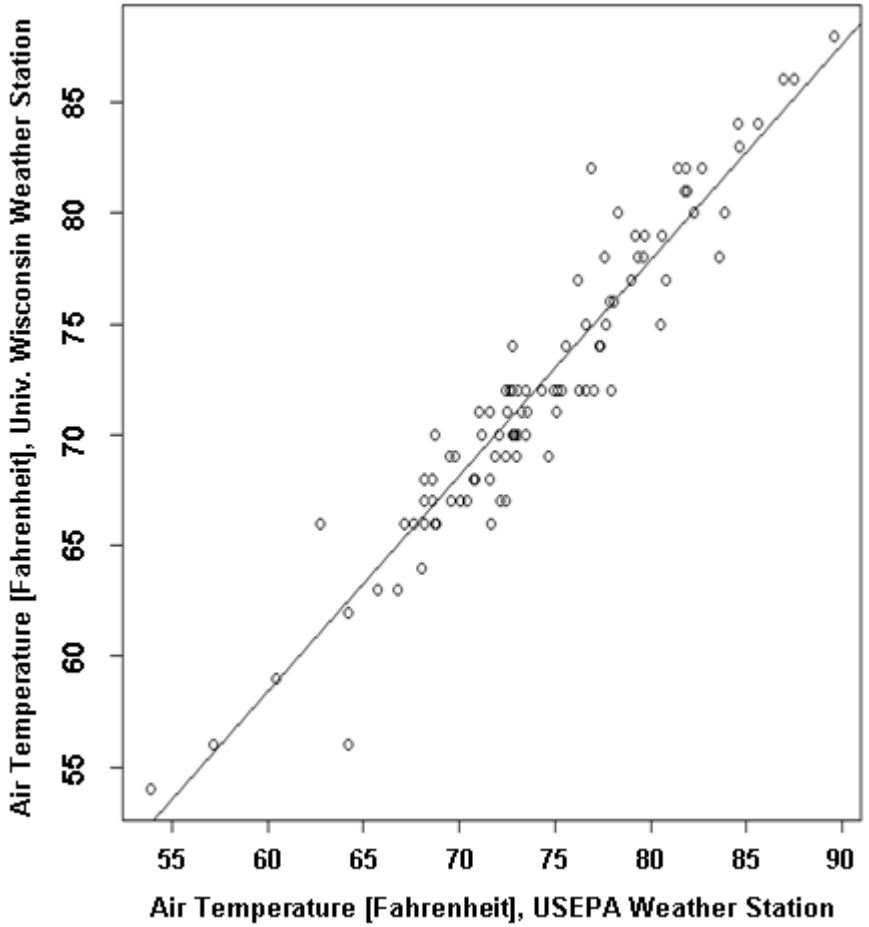
Instruments Deployed

- Data sonde for continuous water quality measurements-
- UV sensors for underwater UV
- Acoustic doppler current profilers for current speed and direction-onsite
- Weather stations for meteorological and solar irradiance –on-site, at Surfside fire station
- Optical properties and DOC of water from onsite- EPA

Weather Stations

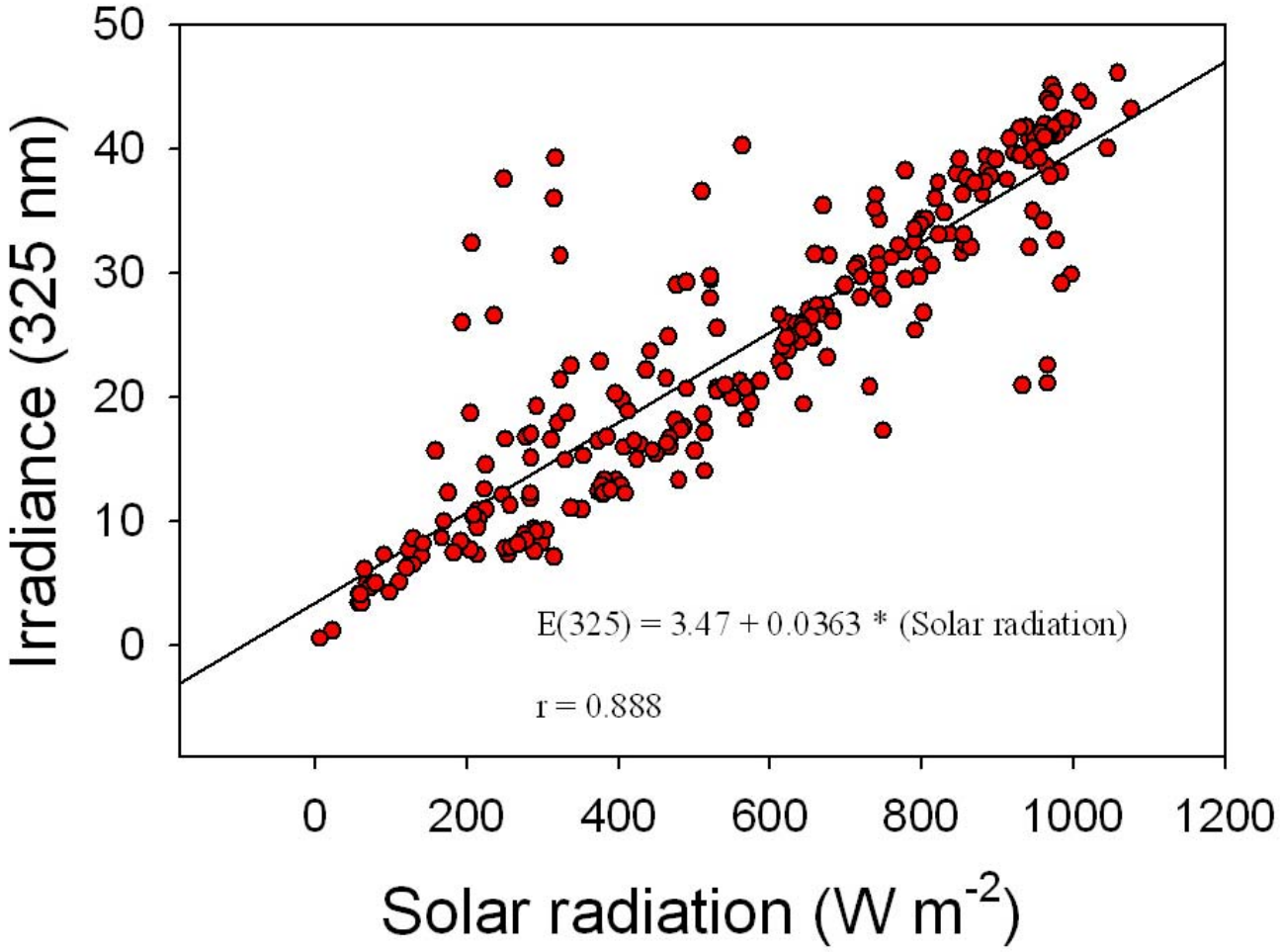
Correlation Between Weather Variables Measured at Different Sites

Other Weather Variables Also Well-Correlated

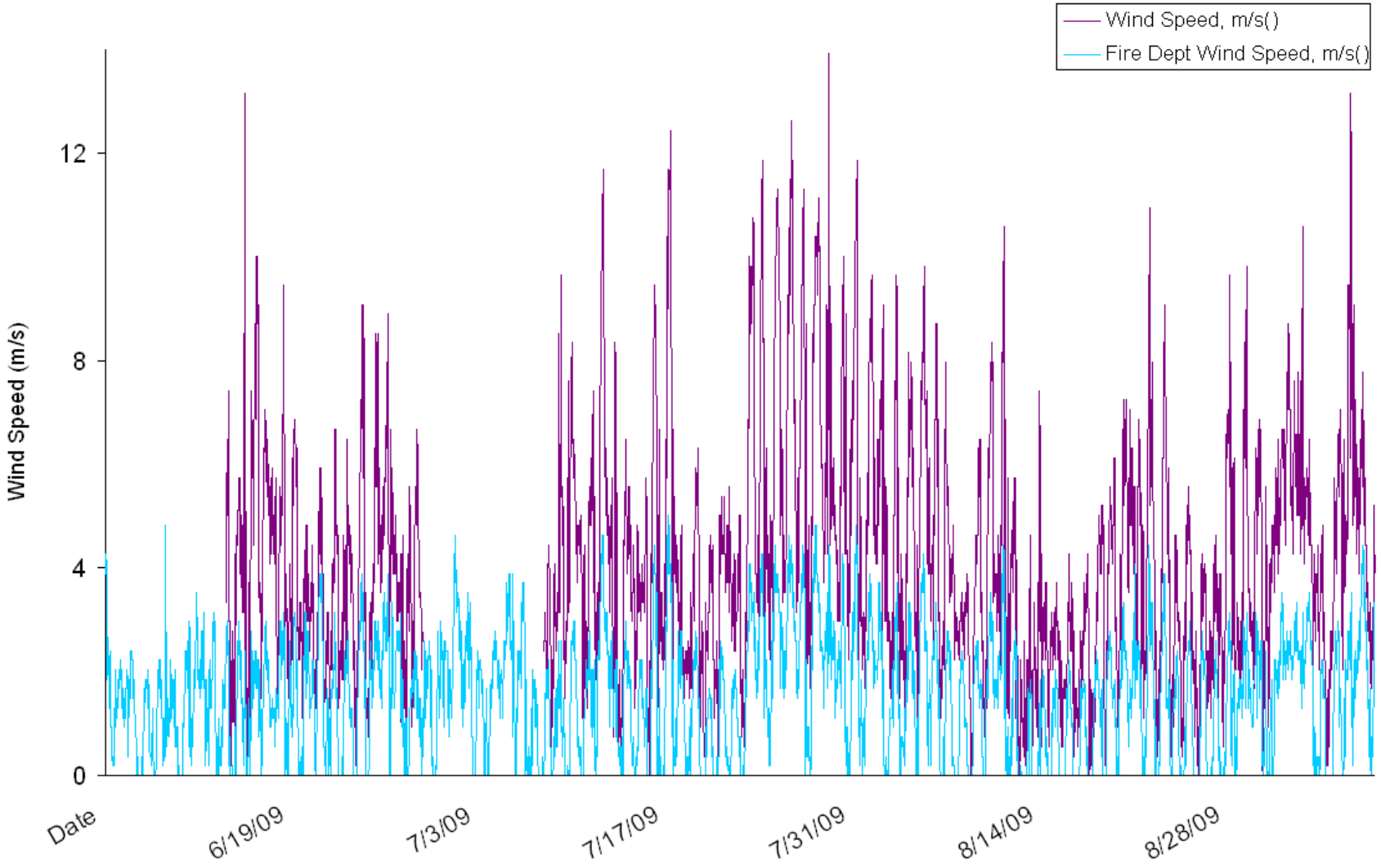


r=0.9496

UV (325 nm) Compared to Solar Radiation



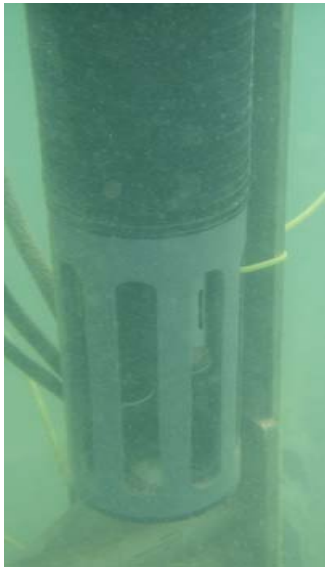
Wind Speeds at Beach Vs. Fire Station

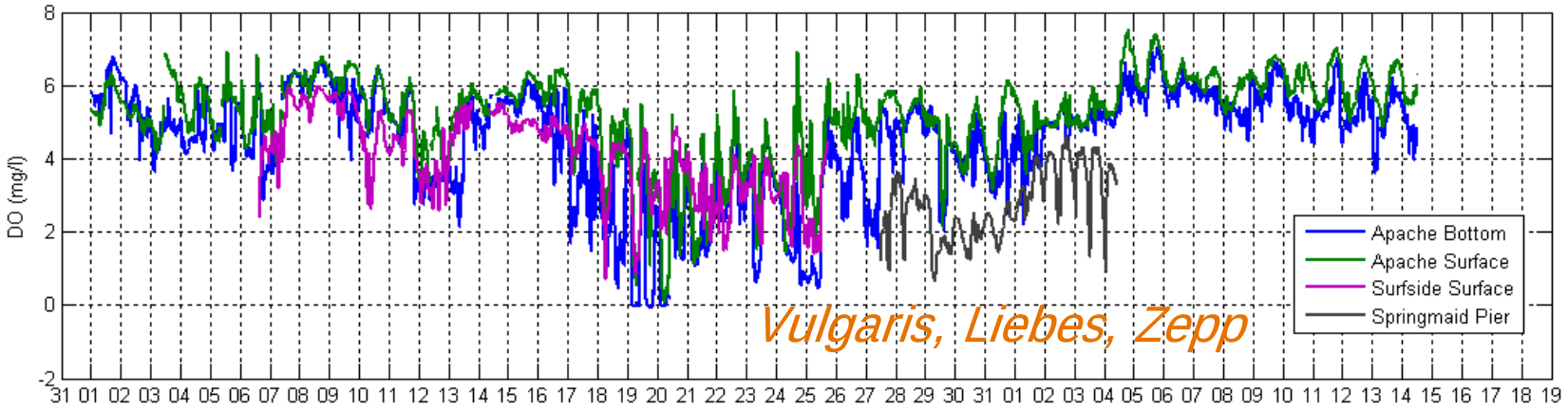


Tower for UV Sensors and Sonde

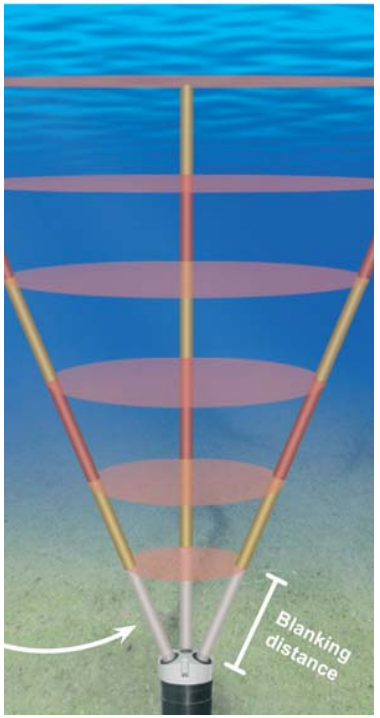
Water Quality Sonde

Water Temperature
Specific Conductivity
Salinity
Dissolved Oxygen
pH
Turbidity
Chlorophyll
Ammonia/Ammonium
Nitrate





Nortek Aquadopp ADCP for Currents



Multiple Linear Regression Model Development

- **Phases to developing and evaluating MLR models for recreational waters**

  - Diagnostic phase – Microbial data and independent or forecasted variables are systematically related to produce an optimal fit

  - Prognostic phase - Fitted models used to provide **nowcasts** and **forecasts**

- **Model selection-** The best independent variables are identified from the suite of potential variables available for fitting

Multiple Linear Regression

$$E(y) = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \beta_3 \cdot x_3 + ... + \beta_p \cdot x_p$$

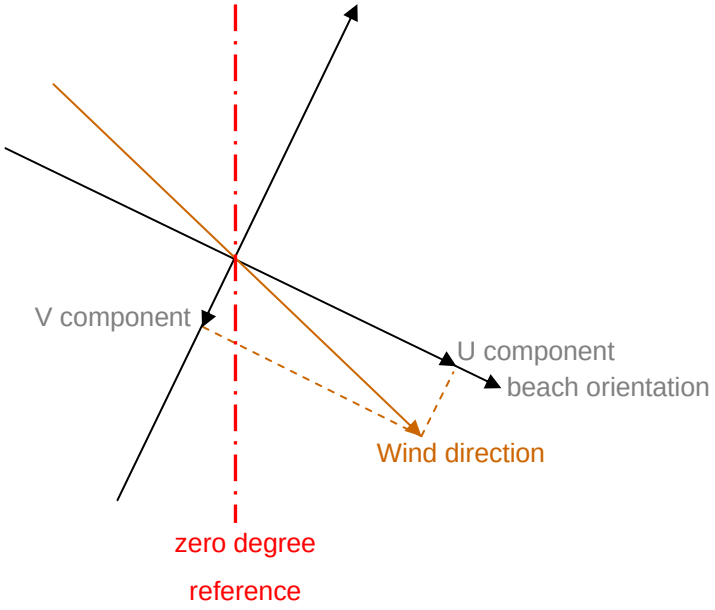
- Where *y* is FIB density (response variable), usually log transformed. x_i's are independent environmental variables (sometime called explanatory variables).
- Transformation on variables may be necessary in order to obtain a linear relationship between the independent variables and the response variable.
- To avoid colinearity, pairwise Pearson's correlation coefficients are examined, and only one variable is kept among a group of highly correlated independent variables.

Model Selection and Diagnostic Criteria

- PRESS (Predicted Residual Sum of Squares) is defined as $\Sigma_i \hat{\varepsilon}_{(i)}^2$, where $\hat{\varepsilon}_{(i)}^2$ are the residuals calculated without using case *i* in the fit. The model with the lowest PRESS criterion is then selected

Transformations of Variables

- FIB concentrations were log transformed
- The wind variables (wind speed and wind direction) transformed into two new components: along the shore (u component) and perpendicular to the shore (v component).



Convention of wind speed and wind direction transformation

MLR Modeling of Enterococci CFU

Objective Criterion = PRESS
(emphasis on predicting new observations)

Adjusted R-square = 0.61

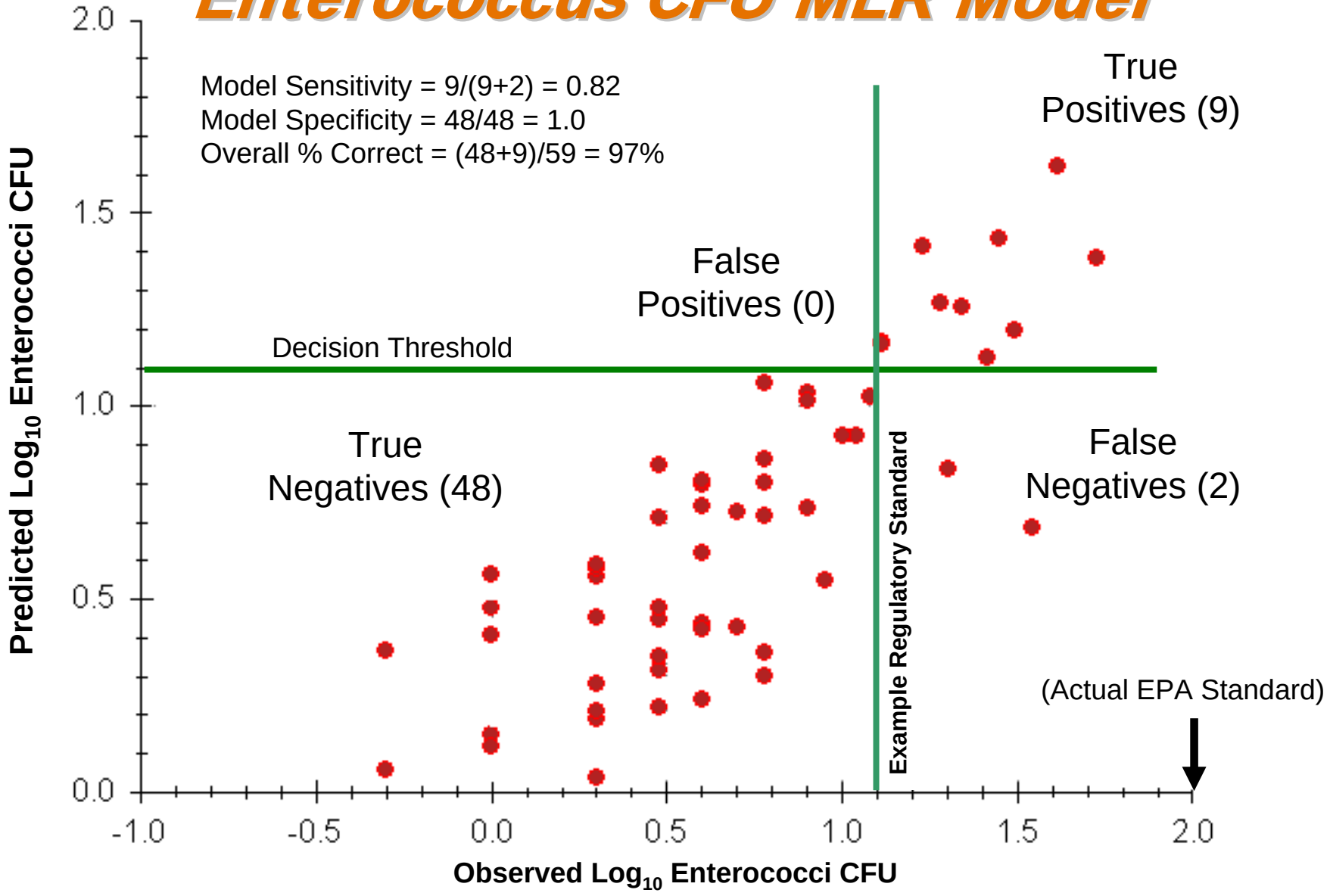
Significant Predictors:

- Wind U-Component (left to right meant higher CFU)
- Wind V-Component (offshore meant lower CFU)
- 48hr Antecedent Rainfall (-)
- Relative Humidity (+)
- Dewpoint (-)
- Air Pressure (-)
- Water Depth (+)

Predicting Exceedances and Non-Exceedances

- Sensitivity: the proportion of the correctly predicted exceedances above the predefined threshold among the total actual exceedances.
- Specificity: the proportion of nonexceedances that are correctly predicted as being below the threshold.

Enterococcus CFU MLR Model



Classification Tree Analysis

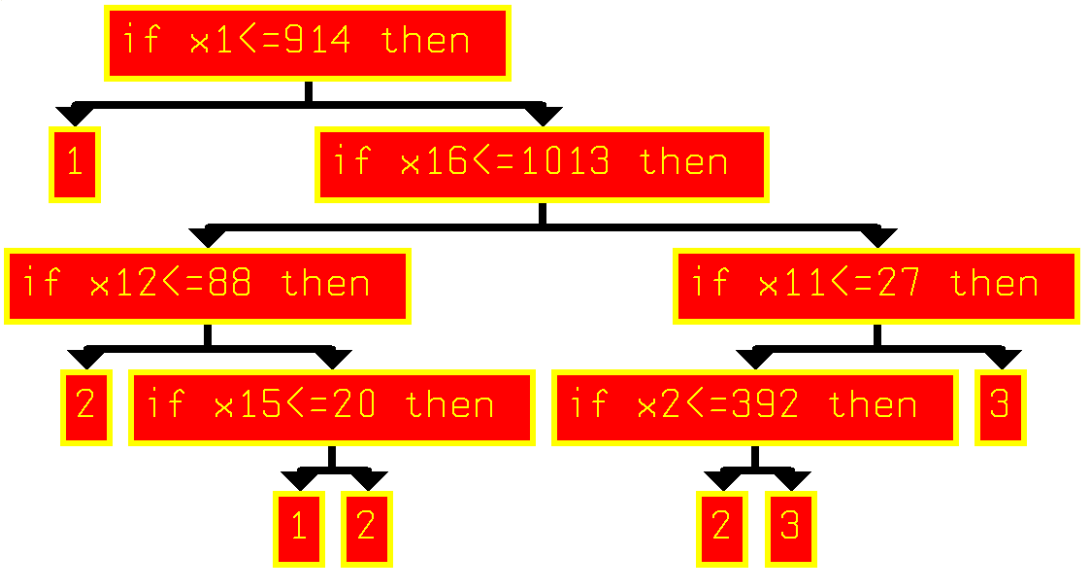
Responses ($y = Log_{10}$[Enterococci CFU]) Split into Three Groups:

Group 1 = $y < 0.4$

Group 2 = $0.4 < y < 0.8$ *Approximately 33% of dataset in each of these groups*

Group 3 = $y > 0.8$

Result: 83% accuracy when classifying training data



X1 = solar radiation (W/m^2)
X16 = air pressure (mbar)
X12 = relative humidity (%)
X11 = air temperature (C)
X15 = dewpoint (C)
X2 = solar radiation at fire department met station (W/m^2)

6 "questions" on the tree are somewhat analogous to a model with 6 parameters

Note: A "yes/true" leads down the right branch of a split

Multiple Linear Regression (MLR) Modeling of Enterococci QPCR

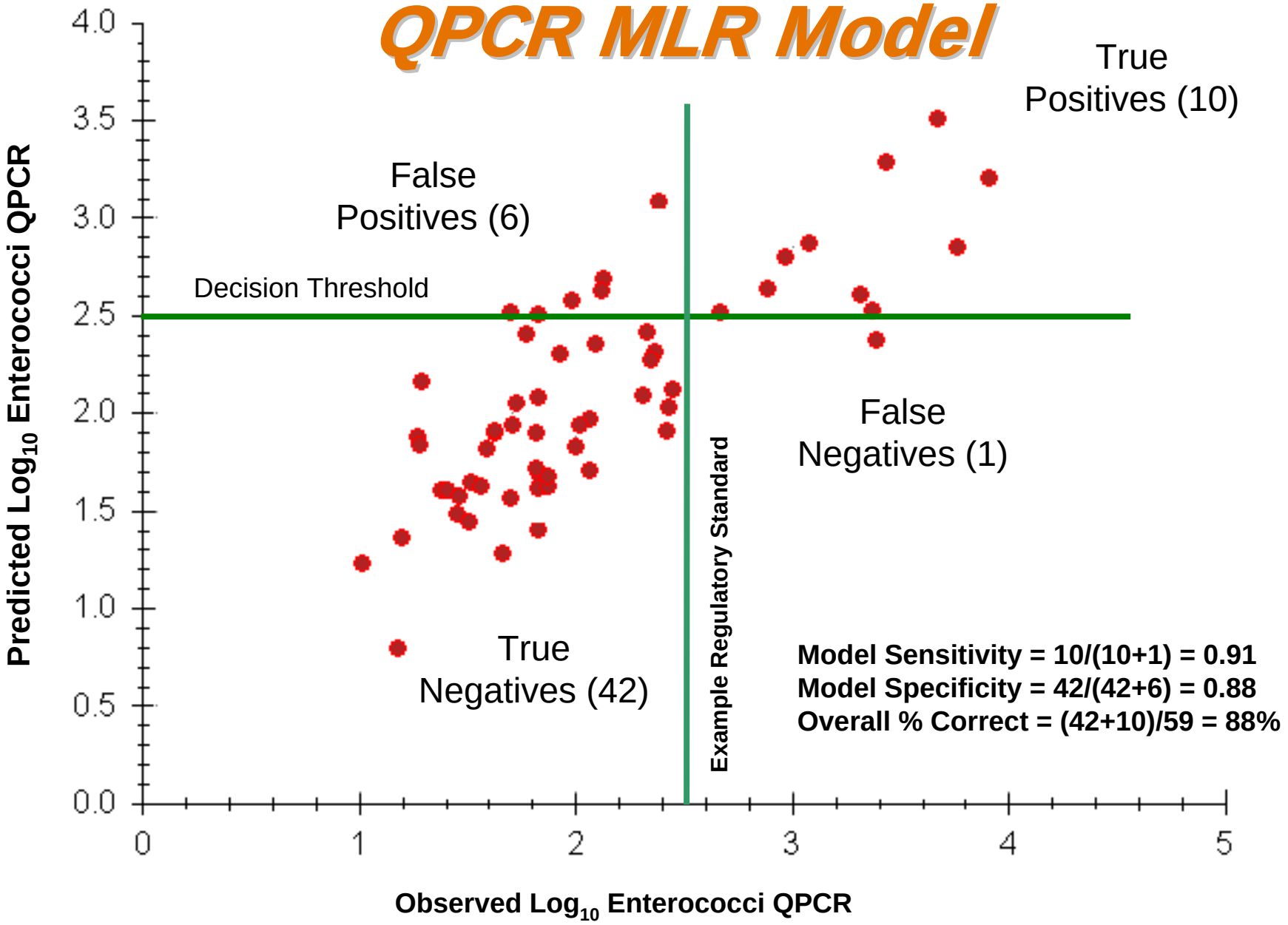
Objective Criterion = PRESS

Adjusted R-square = 0.59

Significant Predictors:

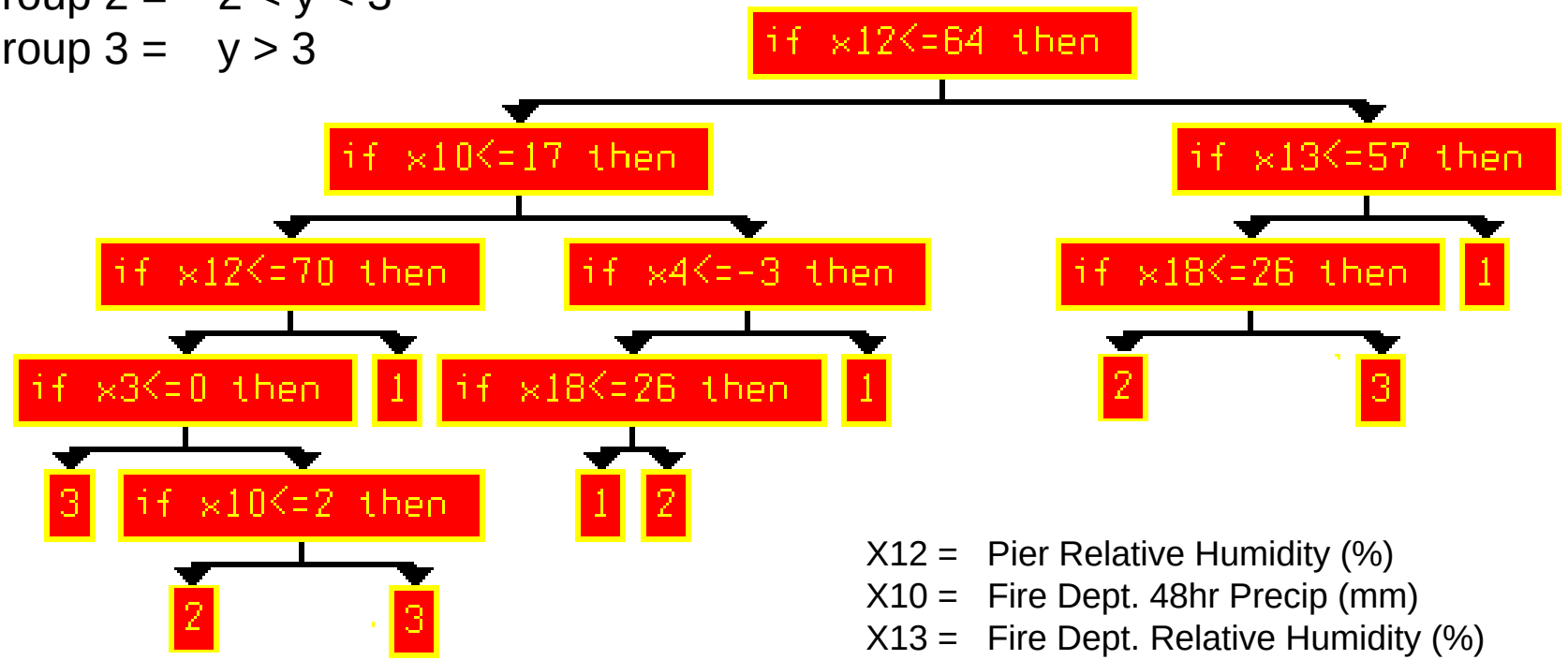
- Wind V-Component (offshore winds meant higher QPCR)
- 48hr Antecedent Precipitation at Fire Department (+)
- 48hr Antecedent Precipitation at Pier (-)
- Air Temperature (+)
- Relative Humidity (+)
- Water Temperature (-)

QPCR MLR Model



Classification Tree for QPCR Data

Response = Log_{10}[Enterococci QPCR])
Group 1 = y < 2
Group 2 = 2 < y < 3
Group 3 = y > 3



X12 = Pier Relative Humidity (%)
X10 = Fire Dept. 48hr Precip (mm)
X13 = Fire Dept. Relative Humidity (%)
X4 = Pier Wind V-Component
X18 = Water Temperature (C)
X3 = Pier Wind U-Component

81% accuracy on training data

Virtual Beach Software for Statistical Modeling

- **Virtual Beach 1.0 - Model Builder** for developing multiple linear regression (MLR) models for indicator prediction and **Beach Advisor for** providing user friendly beach advisory decision support for non technical users
- **Virtual Beach 2.0** – Under development: update software from Delphi to .Net; integrate **Model Builder** and **Beach Adviser** into single tool; add capabilities (e.g. neural networks, CART)

Conclusions

- Both MLR and Tree modeling approaches provide useful predictions at Surfside Beach- both much superior to “persistence” approach
- Meteorological, solar and tide data particularly useful
- Onsite met data, esp. for wind, most useful
- Swash was one of primary sources of contamination at beach, probably via bird sources.

Acknowledgements

- Shannon Berry
- Sean Torrens
- Ted Ambrose
- UNC-Wilmington (Jay Styron et al)
- Town of Surfside Beach (Ed Booth, Micki Fellner, Ty Taylor))
- Tim Wade, Elizabeth Sams – EPA ORD/NHEERL
- Westat (Ben Clickner,Amy Kominski, Sara Hader, Karen Dell Torre))