United States Environmental Protection Agency

Tools for Assessing and Managing Environmental Waters

Marirosa Molina
,USEPA, ORD, National Exposure Research Laboratory



Office of Research and Development
NERL/ Ecosystems Research Division

June 5, 2019

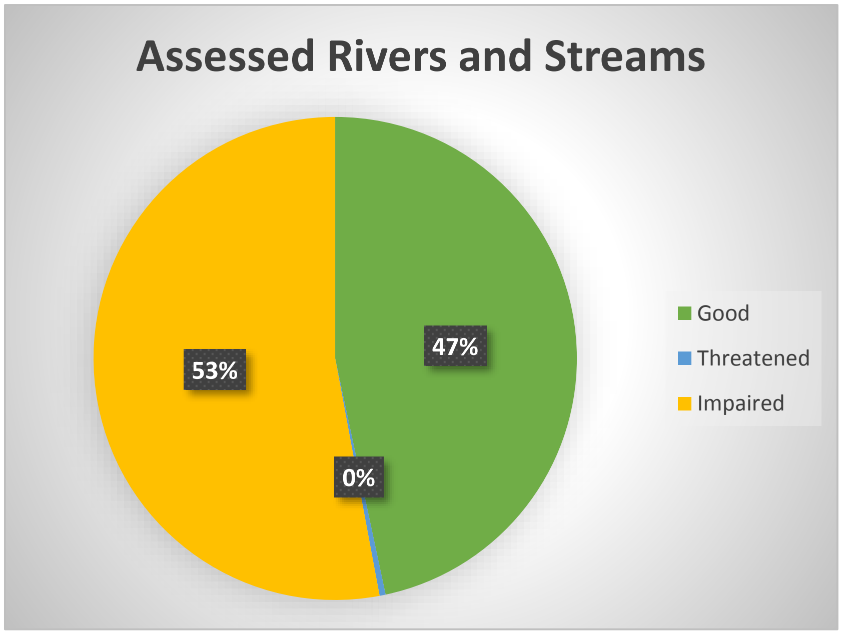
Introduction

- Water resources across the nation are regularly contaminated by unacceptably high levels of fecal indicator bacteria (FIB) that exceed EPA water quality criteria levels.



National Summary of Water Quality Assessments

http://ofmpub.epa.gov/waters10/attains_nation_cy.control



Total miles assessed 1,110,961

Cause of Impairment Group	Miles Threatened or Impaired
Pathogens	187872
Sediment	138874
	118831
	98037
	94488
	94384
	82311
	72554
	63019
Turbidity	47750

List of Pathogens	Miles Threatened or Impaired
Escherichia Coli (E. Coli)	110,964
Fecal Coliform	55,099
Enterococcus Bacteria	11,213
Bacteria	7,107
Pathogens	5,524
Total Coliform	5,399
Indicator Bacteria	3,820
Fecal Bacteria	108
Bacterial Slimes	25
Viruses (Enteric)	6
Coliforms	

Problems

- Pathogens ≠ indicators
- No source information
- Increased economic burden
- Increased health risk

Cost of Chesapeake Bay TMDL for Virginia

	Projected Total Cost ($ in billions)	Who Pays	Potential State Costs ($ in billions)	Potential Sources of Funding
Wastewater (including CSOs)	$1.4	State Govt./Local Govt./Rate-payers	$0.3 (plus $78 million for CSOs?)	WQIF, State GF, Bonds /Local GF, Bonds/Tax Assessments, Sewer Rates
Agriculture	$1.2+	State Govt./ Farmers	$0.8+	WQIF, State GF/ Agribusinesses
Stormwater	$9.4 to $11.5 (including VDOT)	Local Govt./ Property Owners/ VDOT	$2.1 (VDOT Share)	Local GF, Bonds/Utility Fees, Assessments/ Transportation Trust Fund
Onsite/ Septic Systems	$1.6	Property Owners	Unknown What Role State May Play	"Betterment loans", Potential for Tax Credits or Grants
Bay TMDL Total	**$13.6 to $15.7**			

From Senate Finance Committee Report Nov 18, 2011

NEED: Develop Microbial Risk Assessment Tools

=

Understand the risk of exposure to waterborne pathogens

→

Information may help states and stakeholders prioritize use of resources

2012/2017 RWQC Revision
More tools for assessing and managing recreational waters

Rapid qPCR Assays

- Enterococcus qPCR technique (EPA Method 1611),
- use of qPCR on a site-specific basis for monitoring.

Microbial Source Tracking (MST):

- allows development of alternative site-specific criteria and identifies opportunities for source remediation.

Predictive models

- offer states, territories, and tribes the potential for same-day notification and public health protection
- considerably lower capital investment and unit costs.

2012/2017 RWQC Revision
More tools for assessing and managing recreational waters

Rapid qPCR Assays
- Enterococcus qPCR technique (EPA Method 1611),
- use of qPCR on a site-specific basis for monitoring.

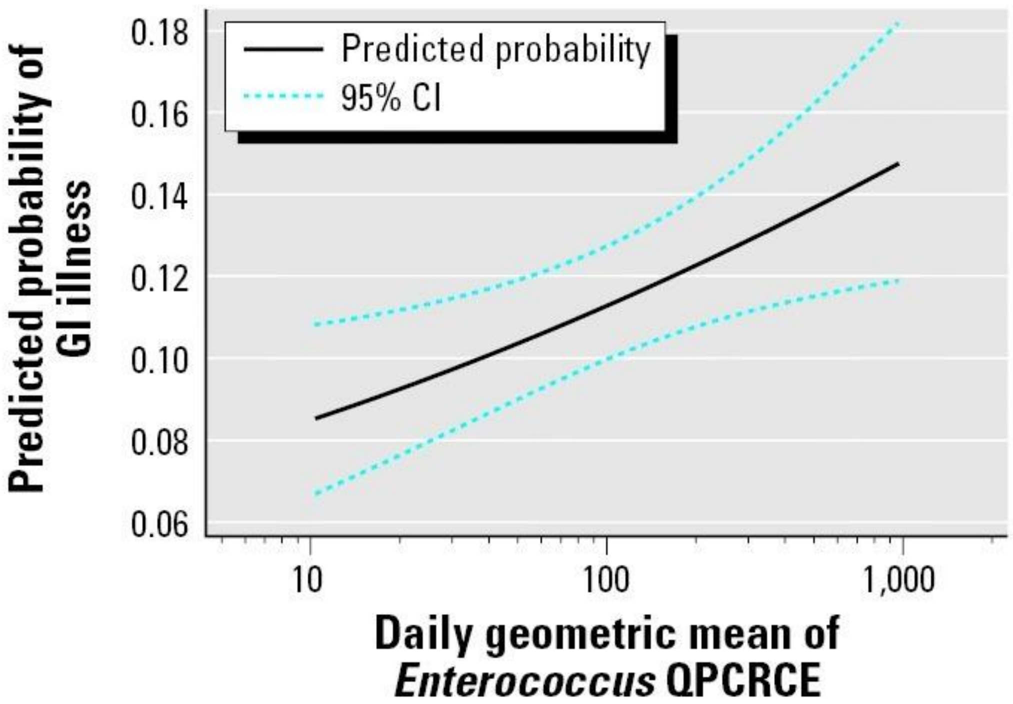
Microbial Source Tracking (MST):
- allows development of alternative site-specific criteria and identifies opportunities for source remediation.

Predictive models
- offer states, territories, and tribes the potential for same-day notification and public health protection
- considerably lower capital investment and unit costs.

Culture vs. Real Time PCR Based
Fecal Indicator Bacteria Measurements
to Determine Beach Water Quality
Filter
Water Sample
Filter
Water Sample
Grow Indicators
on Filter Membranes
Extract DNA
from Filter
Count Indicator Colonies on
Filter to Determine Water Quality
Amplify & Quantify Indicator DNA
by PCR to Determine Water Quality
24 Hours
2-6 Hours
Pathogen (Virus
Parasite, or Bacteria)
Indicator
(Fecal Bacteria)

qPCR Enterococci - better predictor of illness

- *Enterococcus* spp. measured by qPCR is more predictive of swimming-associated health effects.
- Results are more timely than traditional methods used for culturing bacterial indicators.



Predicted probabilities of GI illness as a function of *Enterococcus* QPCRCE, predicted from the logistic regression model, adjusted for age and beach.

Wade et al., 2006

2012/2017 RWQC Revision
More tools for assessing and managing recreational waters

Rapid qPCR Assays

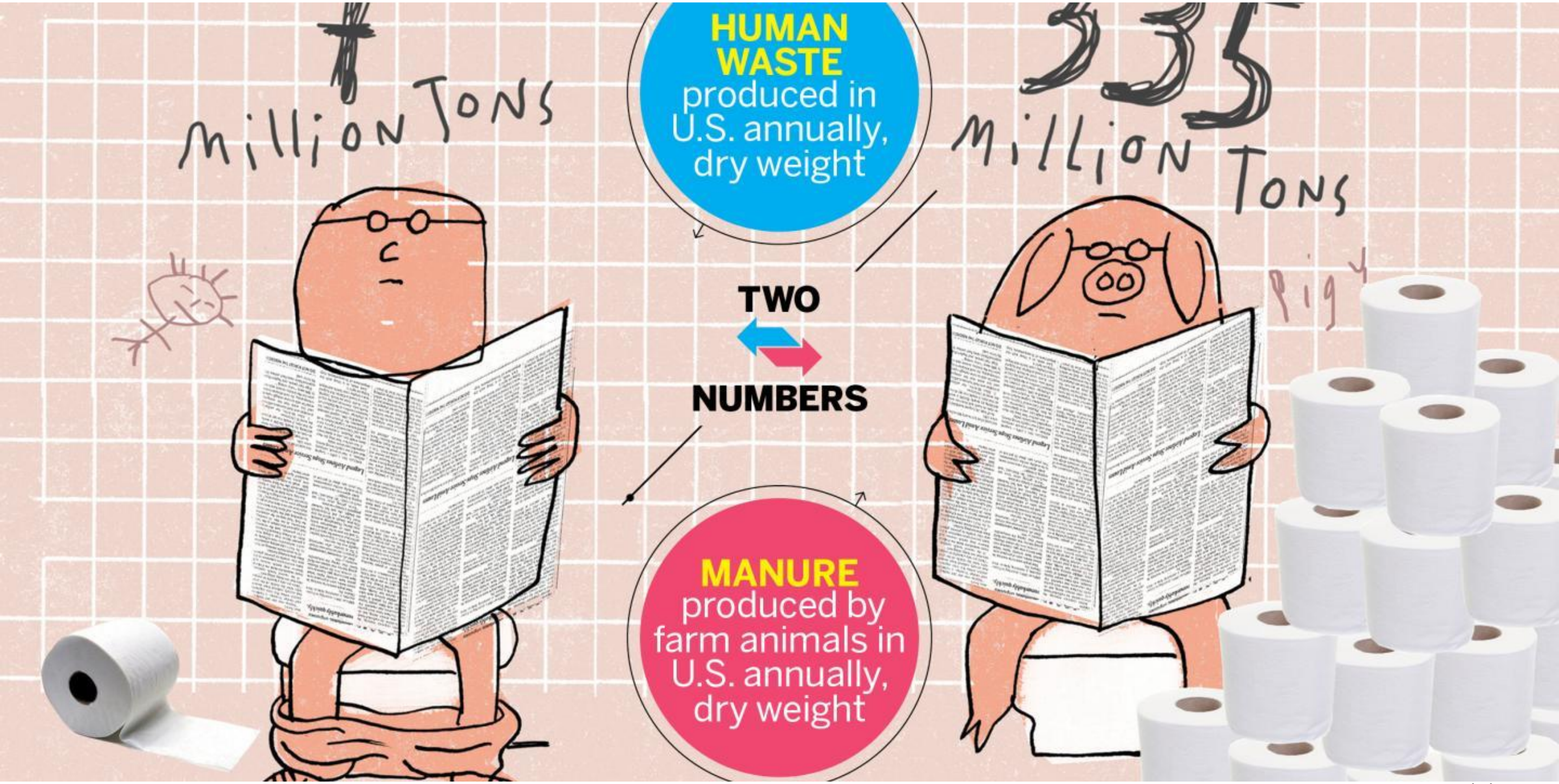
- Enterococcus qPCR technique (EPA Method 1611),
- use of qPCR on a site-specific basis for monitoring.

Microbial Source Tracking (MST):

- allows development of alternative site-specific criteria and identifies opportunities for source remediation.

Predictive models

- offer states, territories, and tribes the potential for same-day notification and public health protection
- considerably lower capital investment and unit costs.



FROM **NEWSWEEK MAGAZINE** 12/8/15

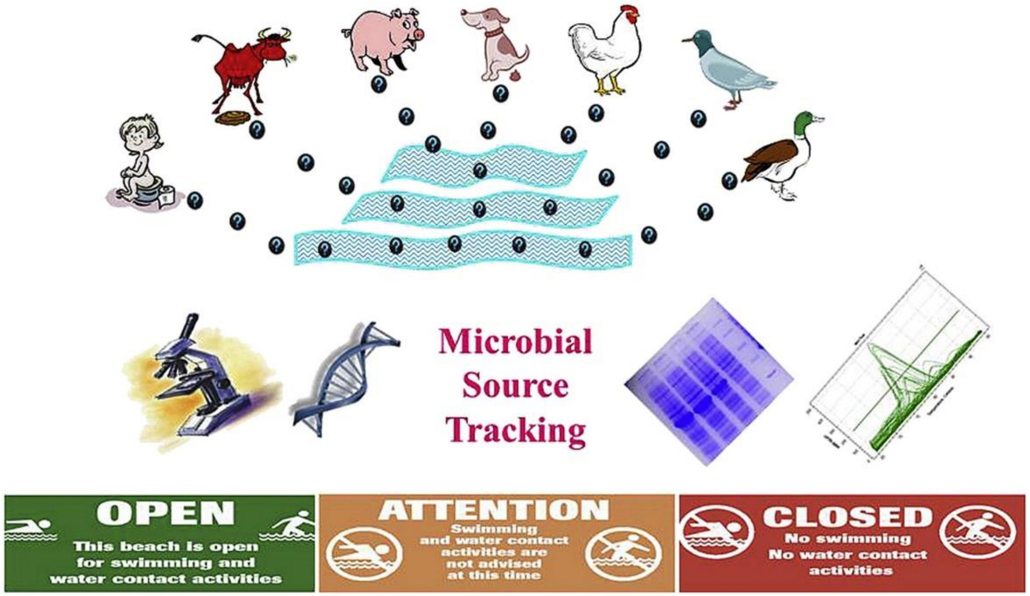
- Animal feeding operations (CAFOS, AFOs, feedlots produce > 335 million tons of manure/yr.
- Humans produce ~ 7 million tons of feces/yr (or 3.9×10^{12} kg/yr)

https://www.ars.usda.gov/ARSUserFiles/np206/NP206FY2005AnnualReport.pdf
Berendes et al., 2018

MST Definition

Microbial source tracking (MST) refers to methodology designed to determine sources of fecal pollution in environmental waters by relating fecal microorganisms with an associated host.

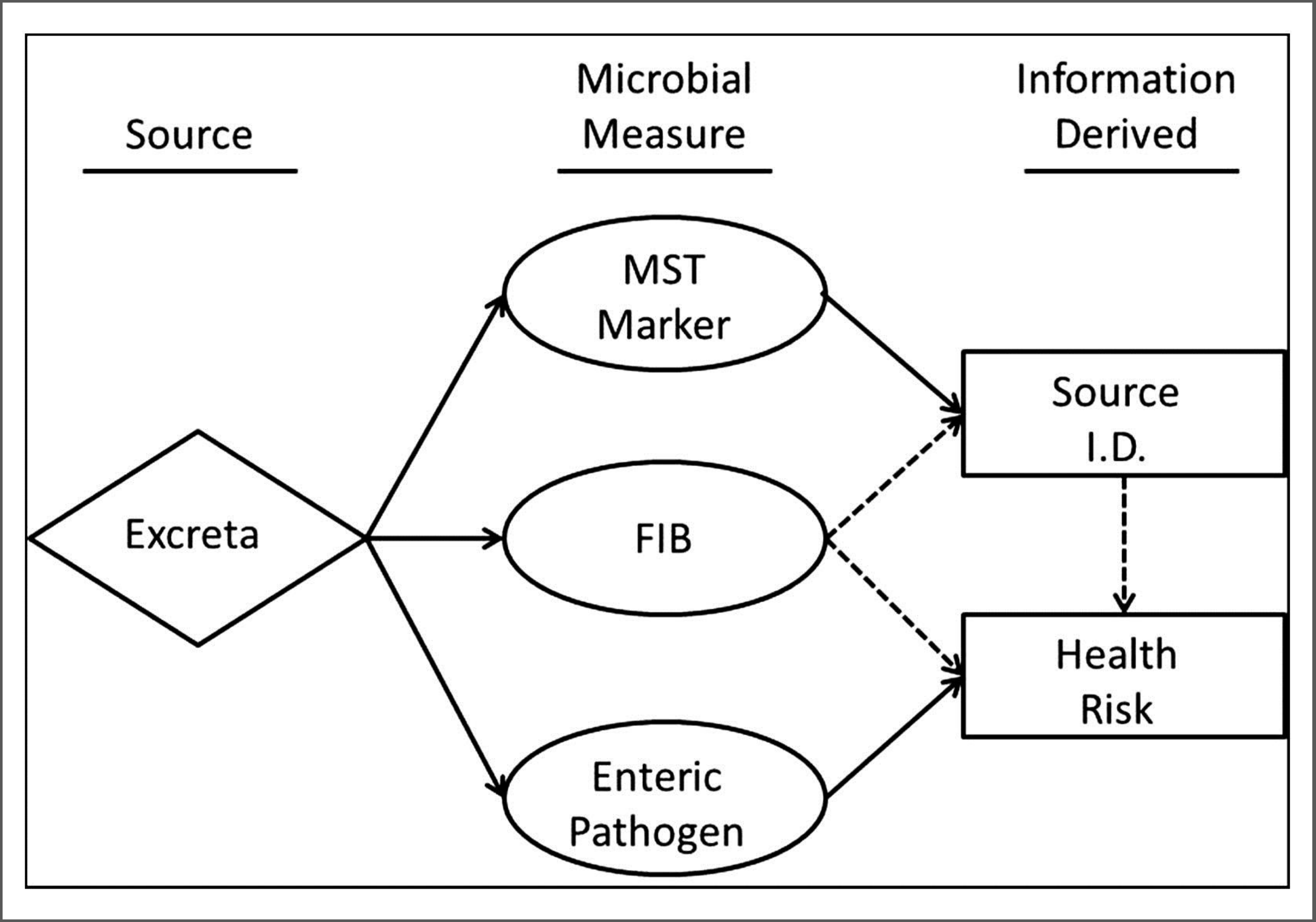
United States Environmental Protection Agency



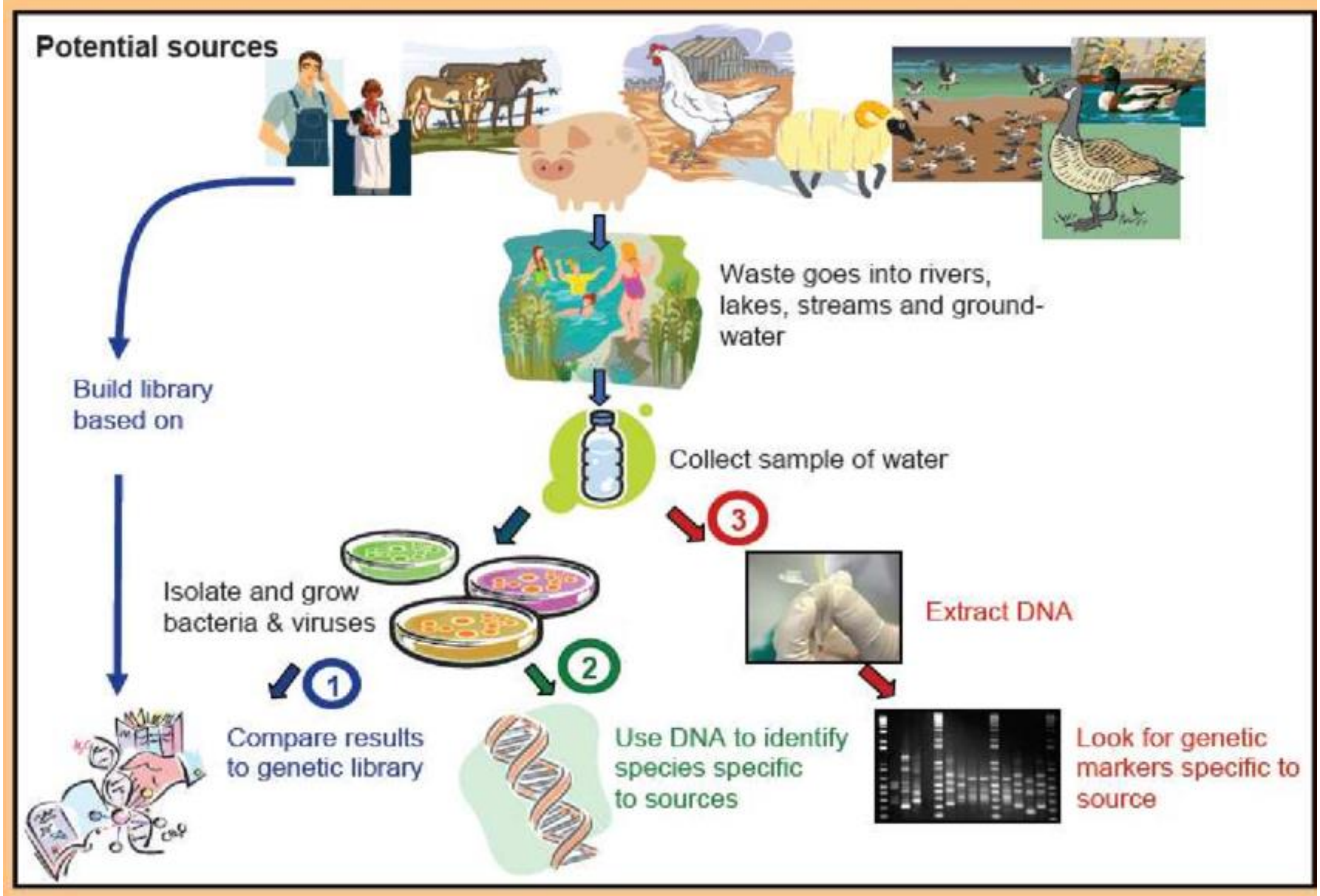
From Stewart et al., 2013

MST tools can be applied for:

- Development of Total Maximum Daily Loads (TMDL)
- Evaluation of the effectiveness of best management practices.



From Stewart et al., 2013


Potential sources
Waste goes into rivers, lakes, streams and ground-water
Collect sample of water
Build library based on
Isolate and grow bacteria & viruses
1
2
3
Extract DNA
Compare results to genetic library
Use DNA to identify species specific to sources
Look for genetic markers specific to source

T table.pdf - Adobe Acrobat Reader DC

Edit View Window Help

me Tools technical_manuscri... Boehn 2013 MST m... MST table.pdf ×

1 / 1 106%

Table 5 – Performance of quantitative MST assays normalized by CFU ENT. The number and percent of non-target samples measuring copies per CFU in the range measured in target samples is provided in column labeled 'non-target in target range'; assays for which this is 0 are deemed specific. The median copies/CFU measured in target challenge samples is provided. The second number for the human assays (the number following the semi-colon) shows the median if DNQs and NDs are not included in the median calculation. Assays with a median greater than 50 copies/CFU are deemed sensitive. Both sensitive and specific assays are indicated. An "*" indicates that the N would change to Y if the median that does not include DNQs and NDs is used. The number of laboratories (N) that ran the assay is provided. Assays run by a single lab will require further testing.

Assay	Host	N	# (%) non-target in target range	Median in target (copies/CFU)	Specific	Sensitive	Spec & sens
BacH	Human	1	0 (0)	87; 375	Y	Y	Y
BacHum	Human	7	20 (11)	331; 374	N	Y	N
BsteriF1	Human	4	10 (10)	123; 130	N	Y	N
Btheta	Human	1	0 (0)	11; 13	Y	N	N
gyrB	Human	1	0 (0)	0.003; 279	Y	N*	N*
HF183SYBR	Human	4	7 (7)	52; 71	N	Y	N
HF183Taqman	Human	5	0 (0)	138; 140	Y	Y	Y
HumM2	Human	6	0 (0)	7; 48	Y	N	N
nifH	Human	5	24 (19)	33; 167	N	N*	N
BacCow	Cow	5	12 (7)	13,490	N	Y	N
CowM2	Cow	5	0 (0)	15	Y	N	N
CowM3	Cow	1	0 (0)	1	Y	N	N
BacR	Ruminant	2	0 (0)	955	Y	Y	Y
Rum2Bac	Ruminant	1	0 (0)	832	Y	Y	Y
Gull2SYBR	Gull	4	14 (10)	0.4	N	N	N
Gull2Taqman	Gull	6	21 (10)	7	N	N	N
LeeSeaGull	Gull	1	0 (0)	55	Y	Y	Y

2012/2017 RWQC Revision
More tools for assessing and managing recreational waters

Rapid qPCR Assays
- Enterococcus qPCR technique (EPA Method 1611),
- use of qPCR on a site-specific basis for monitoring.

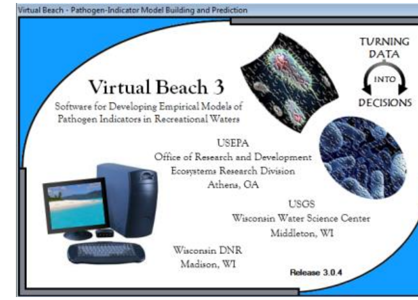
Microbial Source Tracking (MST):
- allows development of alternative site-specific criteria and identifies opportunities for source remediation.

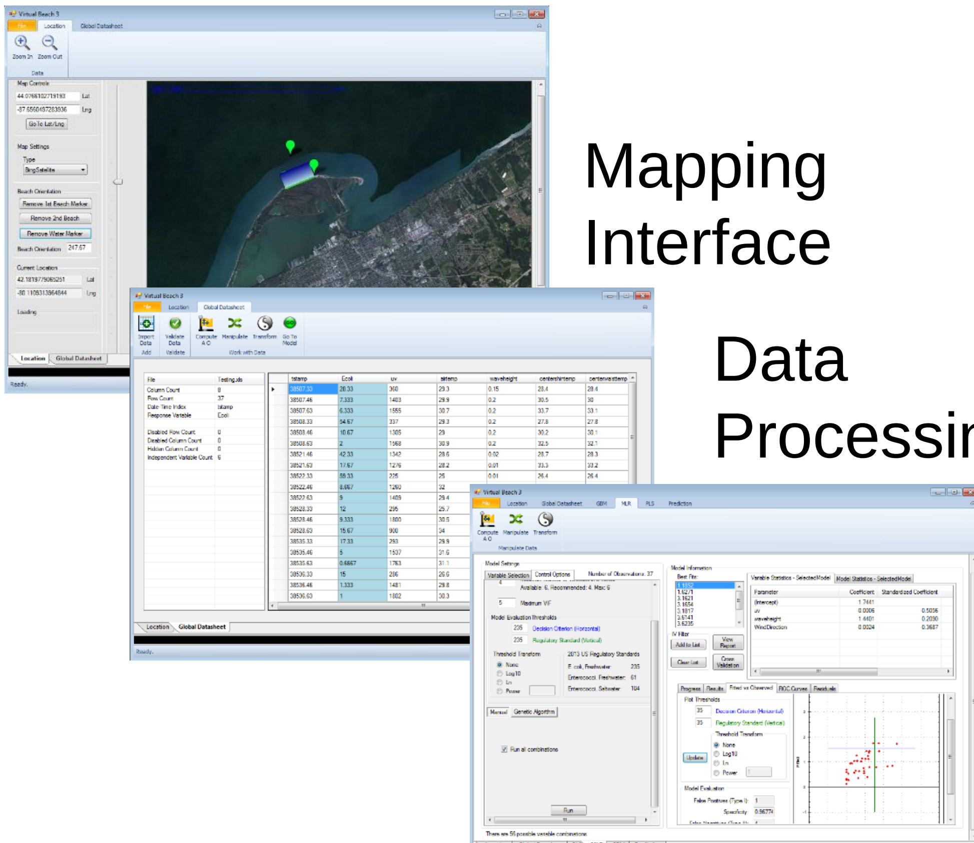
Predictive models
- offer states, territories, and tribes the potential for same-day notification and public health protection
- considerably lower capital investment and unit costs.

EPA United States Environmental Protection Agency

Virtual Beach

A decision-support tool that facilitates the process of building and implementing a statistical model for predicting microbial water quality



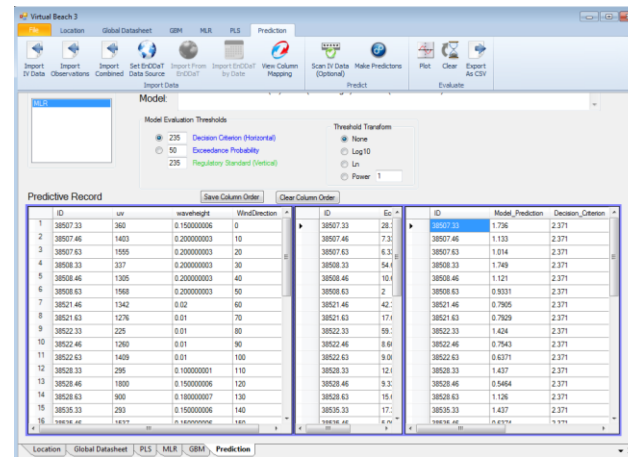


Mapping Interface

Data Processing

Modeling

Prediction



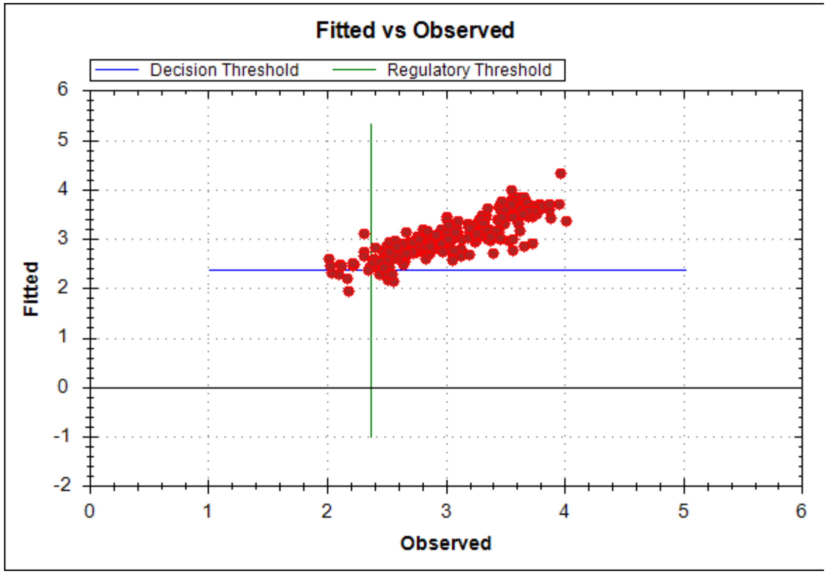
Environmental Parameters

- Rainfall (24-hr, 48-hr)
- Turbidity, conductivity
- Wave height
- Wind speed & direction
- Air & water temperature
- UV/PAR
- Tributary discharge
- Lake currents and stage
- Bathers, birds, debris

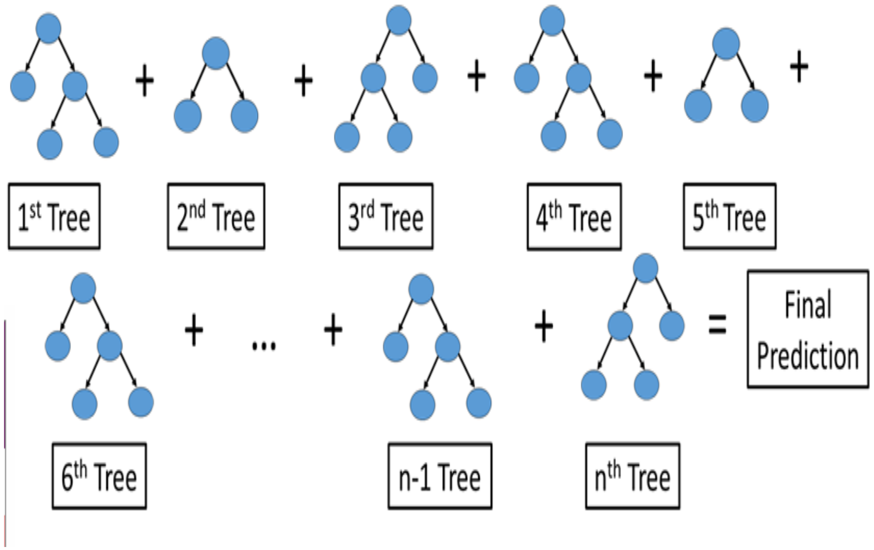
United States Environmental Protection Agency

Virtual Beach Tools: MLR, PLS, GBM

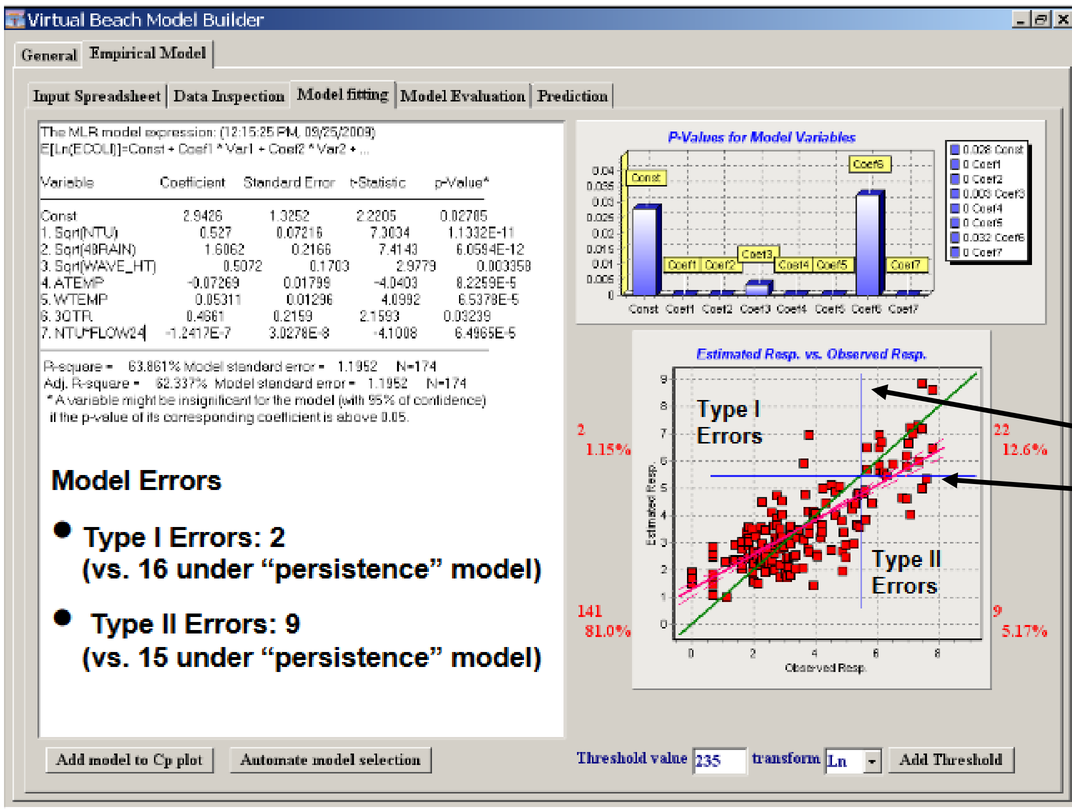
Multiple Linear Regression Models (MLR): fitting a linear equation to observed data



Generalized Boosted Regression Models (GBM): a linked chain of simple decision trees



Probability of exceedance



Type I Error = False Positive; advisory issued under safe conditions

Type II Error = False Negative; no advisory issued under risky conditions

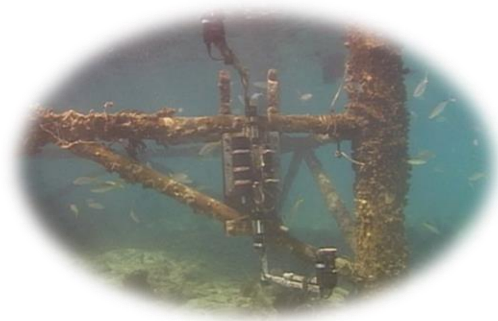
Vertical line = decision criterion

Horizontal line = regulatory standard

Model Application

- Application of a recent version of the software package Virtual Beach tool (VB 3.0) has been performed to predict microbial water quality in beaches located around the Great Lakes, and in South Carolina, Florida, Alabama, Mississippi, Puerto Rico.

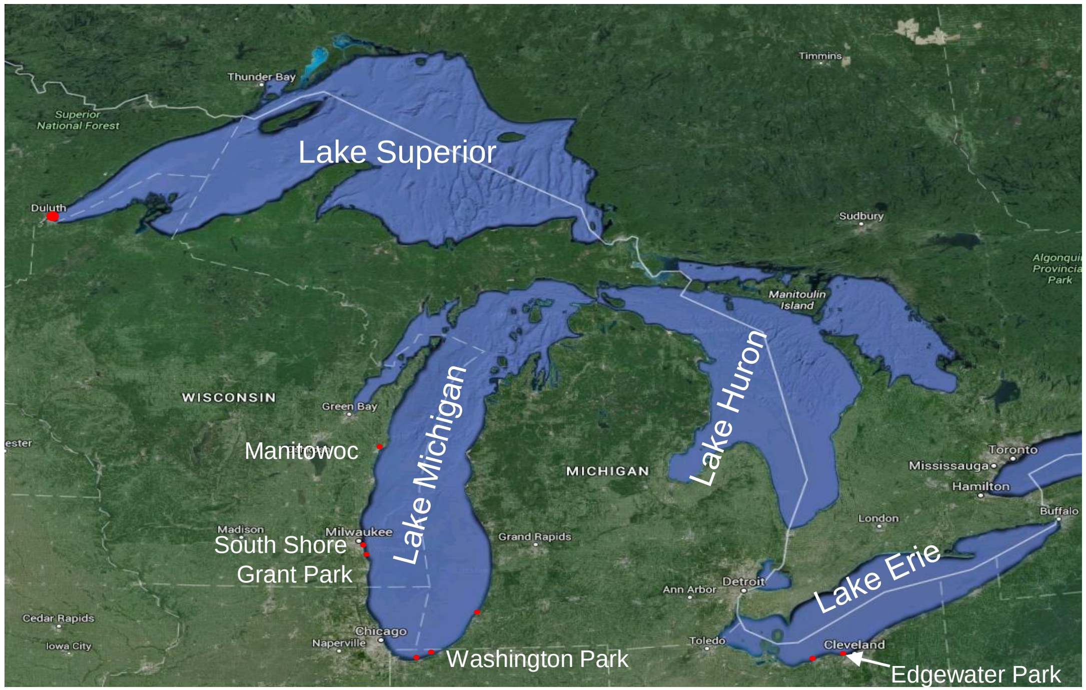




Case Study I

Efficacy of statistical modeling tools to determine stream and river impairment due to fecal contamination.

United States Environmental Protection Agency



- Despite VB widespread use around the Great Lakes (>100 + beaches), application in inland waters is limited.

Assessment of US Waters

	Rivers and Streams (Miles)	Lakes, Reservoirs, and Ponds (Acres)	Bays and Estuaries (Square Miles)	Ocean and Near Coastal (Square Miles)	Wetlands (Acres)	Great Lakes Shoreline (Miles)
Good Waters	518,293	5,390,570	11,516	726	569,328	106
Threatened Waters	4,495	30,309				
Impaired Waters	588,173	13,208,917	44,625	6,218	672,924	4,354
Total Assessed Waters	1,110,961	18,629,795	56,141	6,944	1,242,252	4,460
Total Waters	3,533,205	41,666,049	87,791	54,120	107,700,000	5,202
Percent of Waters Assessed	31.4	44.7	63.9	12.8	1.2	85.7

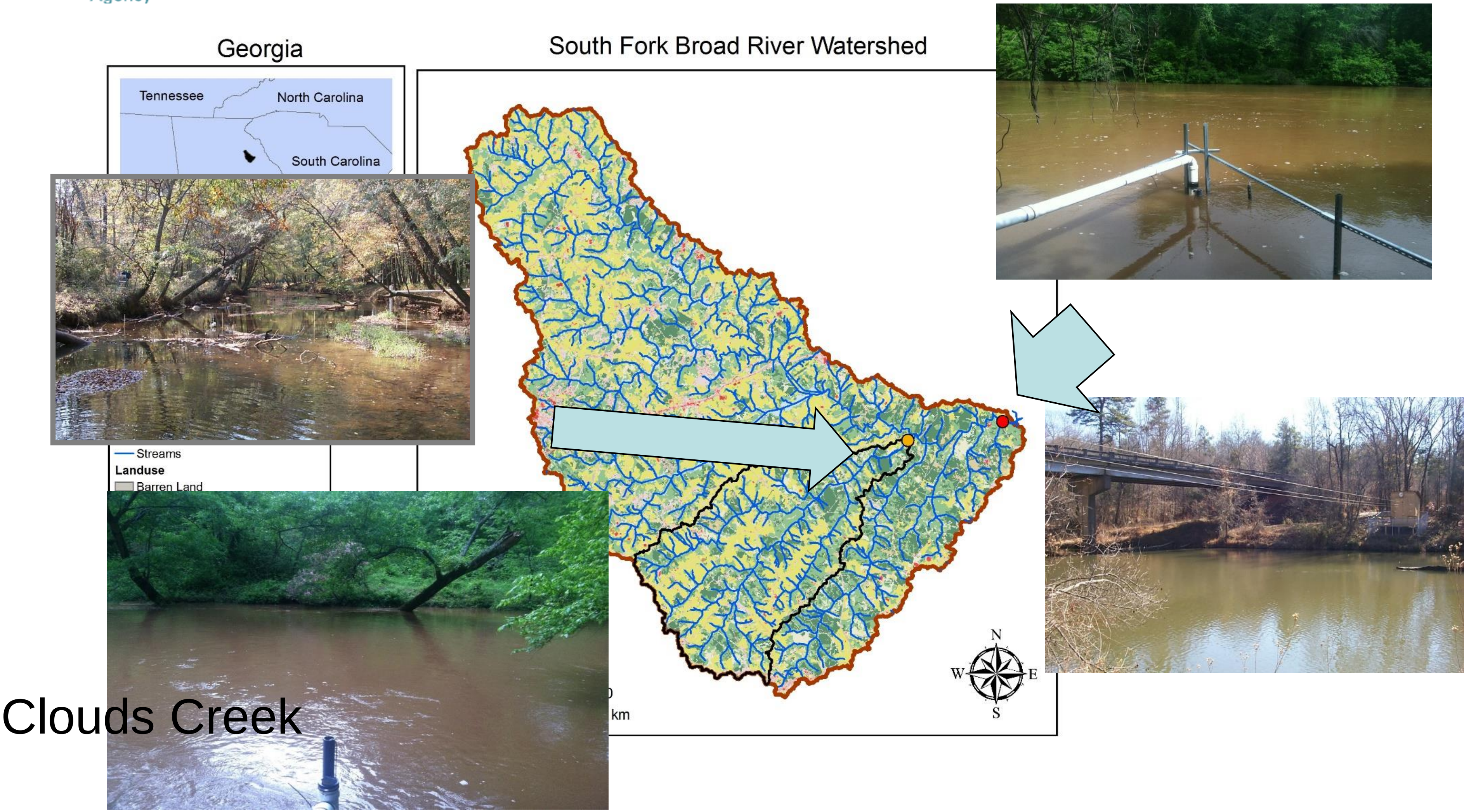
Objective

Evaluate the effectiveness of Multiple Linear Regression (MLR) and Generalized Boosted Regression Models (GBM) for predicting the impairment of inland rivers and streams.

United States Environmental Protection Agency

Study Sites

Carlton Site



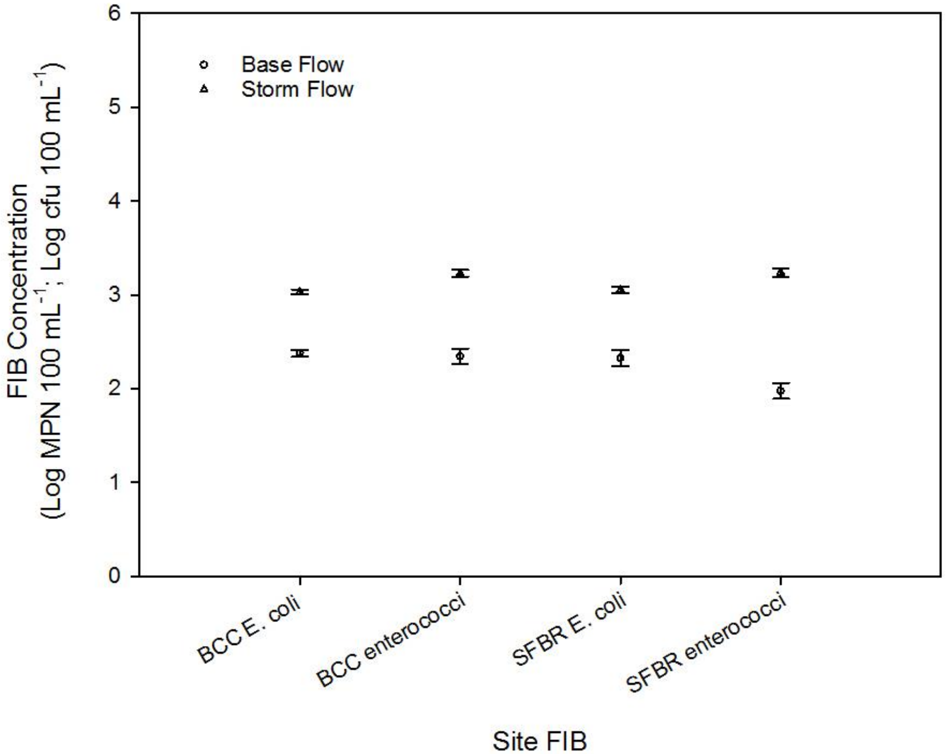
Data Collection

- Intensive field study conducted from 2012-2015

- Water samples collected during rainfall events and baseflow conditions from two sites in the watershed.

- Samples were analyzed for *E. coli,* Enterococci (culturable and qPCR), MST markers and water quality parameters.

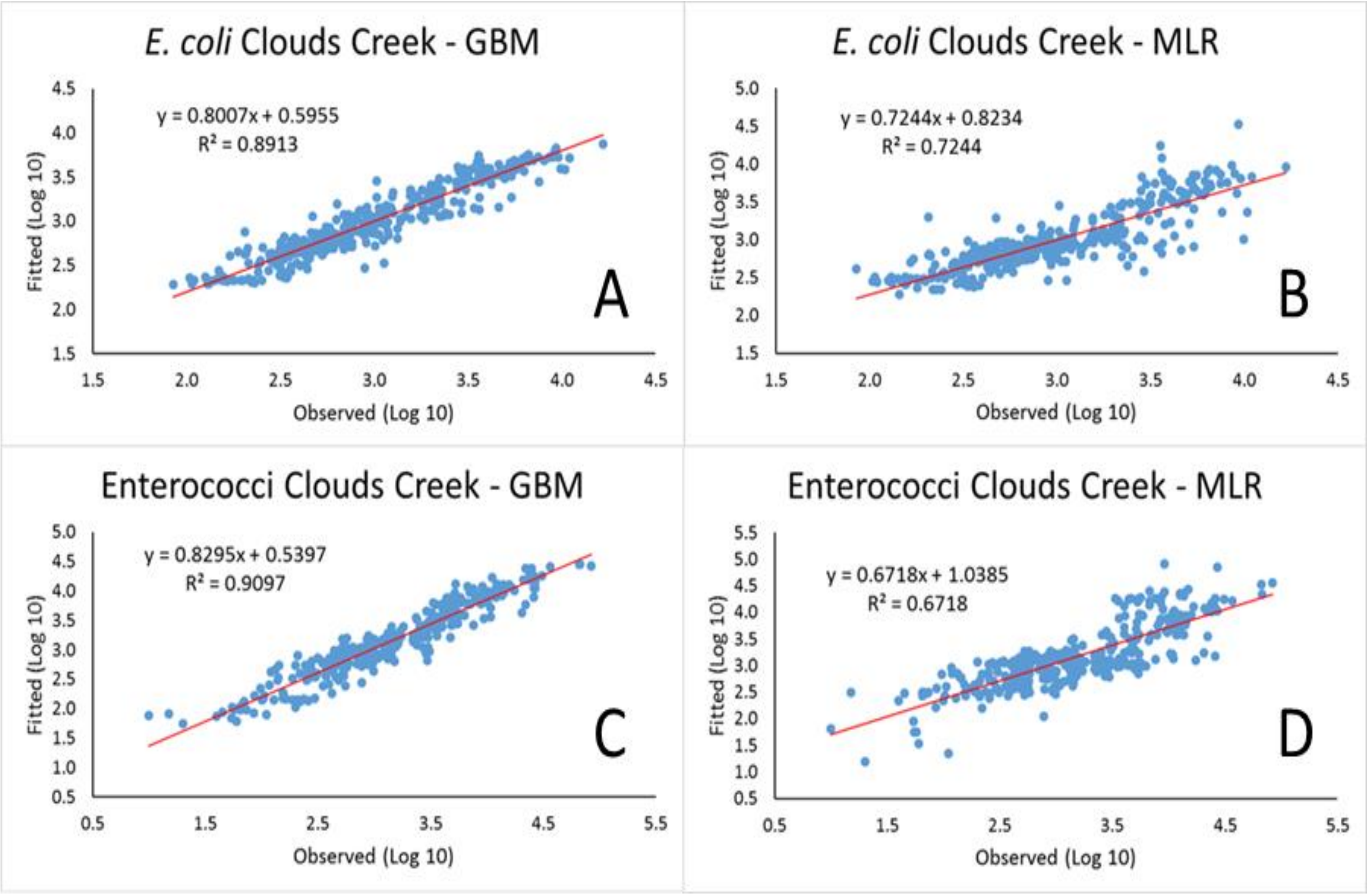
Model Development



- Model training data were randomly selected (approximately 75% of entire dataset)
- IVs included Air Temperature, Barometric Pressure, Dewpoint, Rainfall (15min; 1, 24, 48hr), Rainfall Intensity (1 min), Stream Discharge, Total Suspended Solids (TSS), Turbidity, Water Temperature, Avg. Wind Speed

FIB Prediction at Clouds Creek

Model predictions were made for data not used in model development (approx. 25% of dataset)



United States Environmental Protection Agency

Model Predictive Performance

Parameters	GBM				MLR			
	E.Coli		Enterococci		E.Coli		Enterococci	
	Clouds Creek	Carlton	Clouds Creek	Carlton	Clouds Creek	Carlton	Clouds Creek	Carlton
FP	13	7	15	6	21	19	41	13
FN	22	15	22	17	44	25	52	42
TP	163	115	163	113	141	105	144	88
TN	172	122	170	123	164	110	133	116
Data Points (N)	370	259	370	259	370	259	370	259
Specificity (%)	93	95	92	95	89	85	78	90
Sensitivity(%)	88	88	88	87	76	81	72	67
Accuracy(%)	91	92	90	91	82	83	75	78
R squared	0.89	0.92	0.91	0.91	0.72	0.71	0.67	0.64

- R^2 values for all predictions were > 0.6
- R^2 values were higher for GBM predictions
- Averaged across both models (*E.coli* and enterococci) GBM prediction errors were 16% lower than MLR errors at Clouds Creek, and 7% lower at Carlton

United States Environmental Protection Agency

Selected Environmental Parameters

	Enterococci	*E. coli*
Carlton	Turbidity, 48 hr Rainfall, Dew Point, Humidity and wind speed	Turbidity,48 hr rainfall, dewpoint
Clouds Creek	Turbidity, Rainfall Intensity, 48 hr rainfall, air temperature, dew point, Barometric pressure, wind speed	Water temperature, 1 hr rainfall, 24 hr rainfall

United States Environmental Protection Agency

Take Home Message

- Based on the predictive performance of both GBM and MLR models for FIB at both field sites, our results show statistical models developed with VB can be effective for predicting the impairment of inland rivers and streams.

- GBM produced better *E.coli* and Enterococci predictions than MLR regardless of site differences.

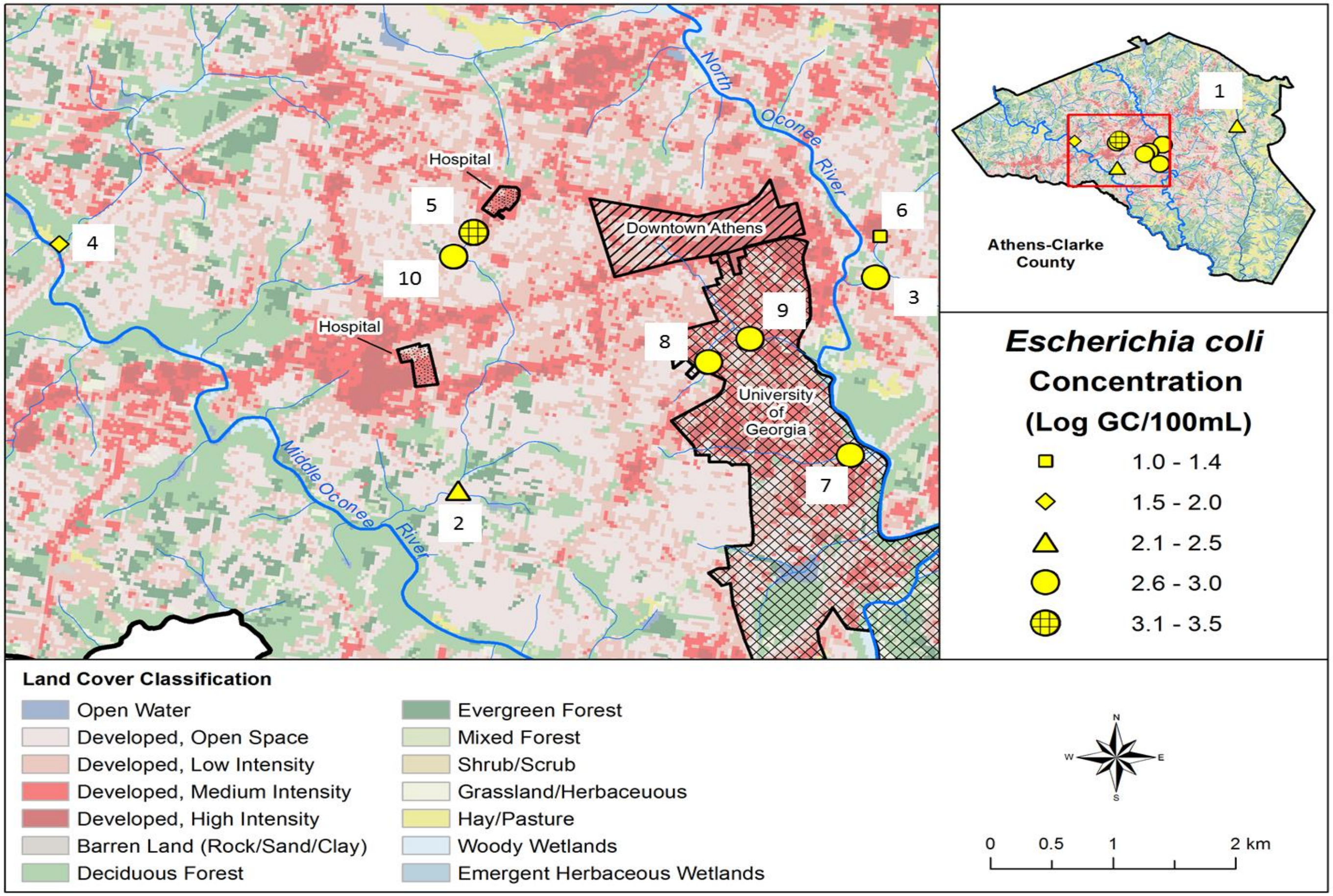
Case Study II

Using a Microbial Citizen Science Initiative As A Tool to Manage Stream Water Quality in An Urban Watershed

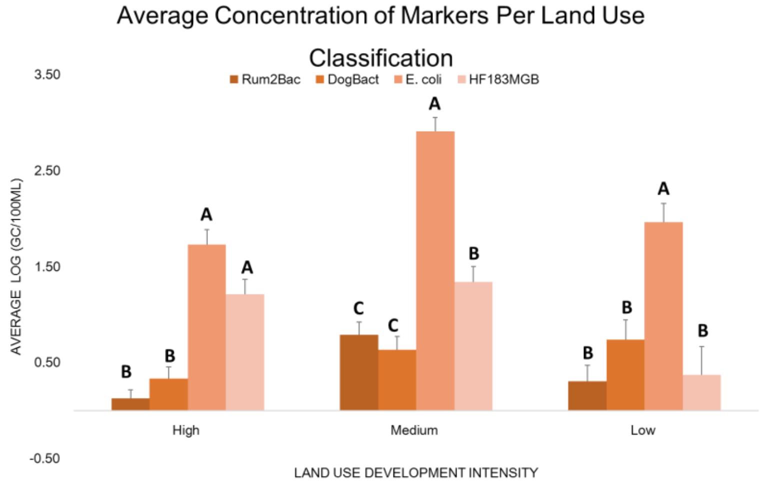
Objective

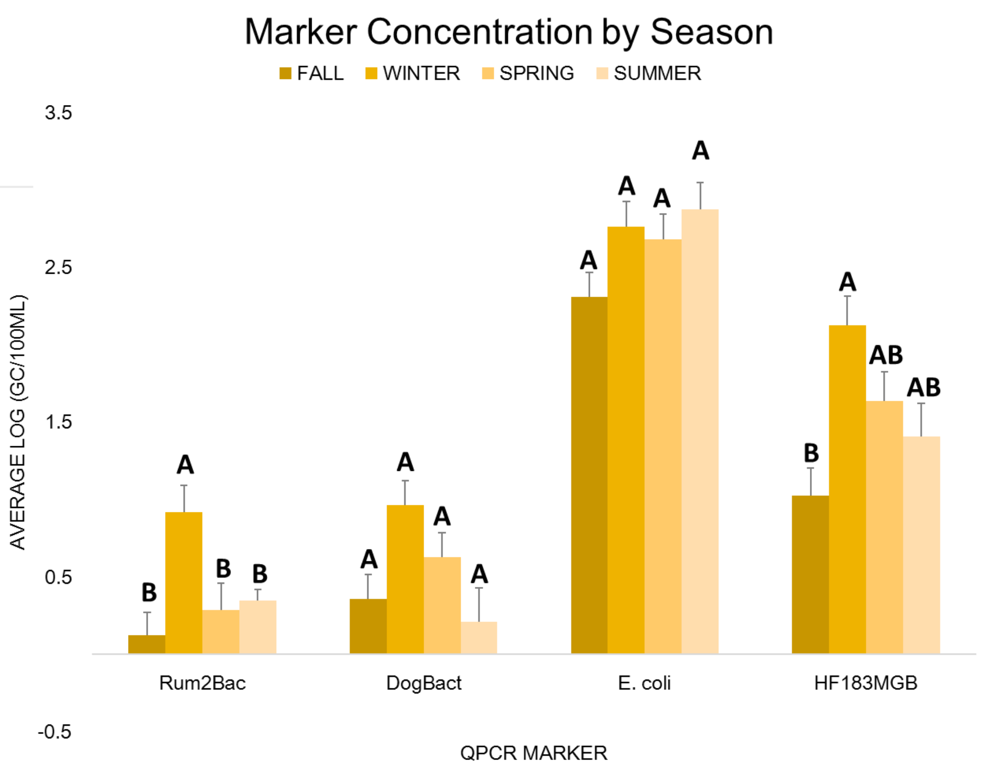
To provide information to local governments and the UGA River Basin Center about common sources of fecal contamination and locate hot spots within local urban streams.

United States Environmental Protection Agency



Land Use Development Intensity of Athens, GA with corresponding average *E. coli* concentrations (Log GC/100mL).


Average Concentration of Markers Per Land Use Classification
Rum2Bac
DogBact
E. coli
HF183MGB
AVERAGE LOG (GC/100ML)
3.50
2.50
1.50
0.50
-0.50
High
Medium
Low
LAND USE DEVELOPMENT INTENSITY


Marker Concentration by Season
FALL
WINTER
SPRING
SUMMER
AVERAGE LOG (GC/100ML)
3.5
2.5
1.5
0.5
-0.5
Rum2Bac
DogBact
E. coli
HF183MGB
QPCR MARKER

United States Environmental Protection Agency

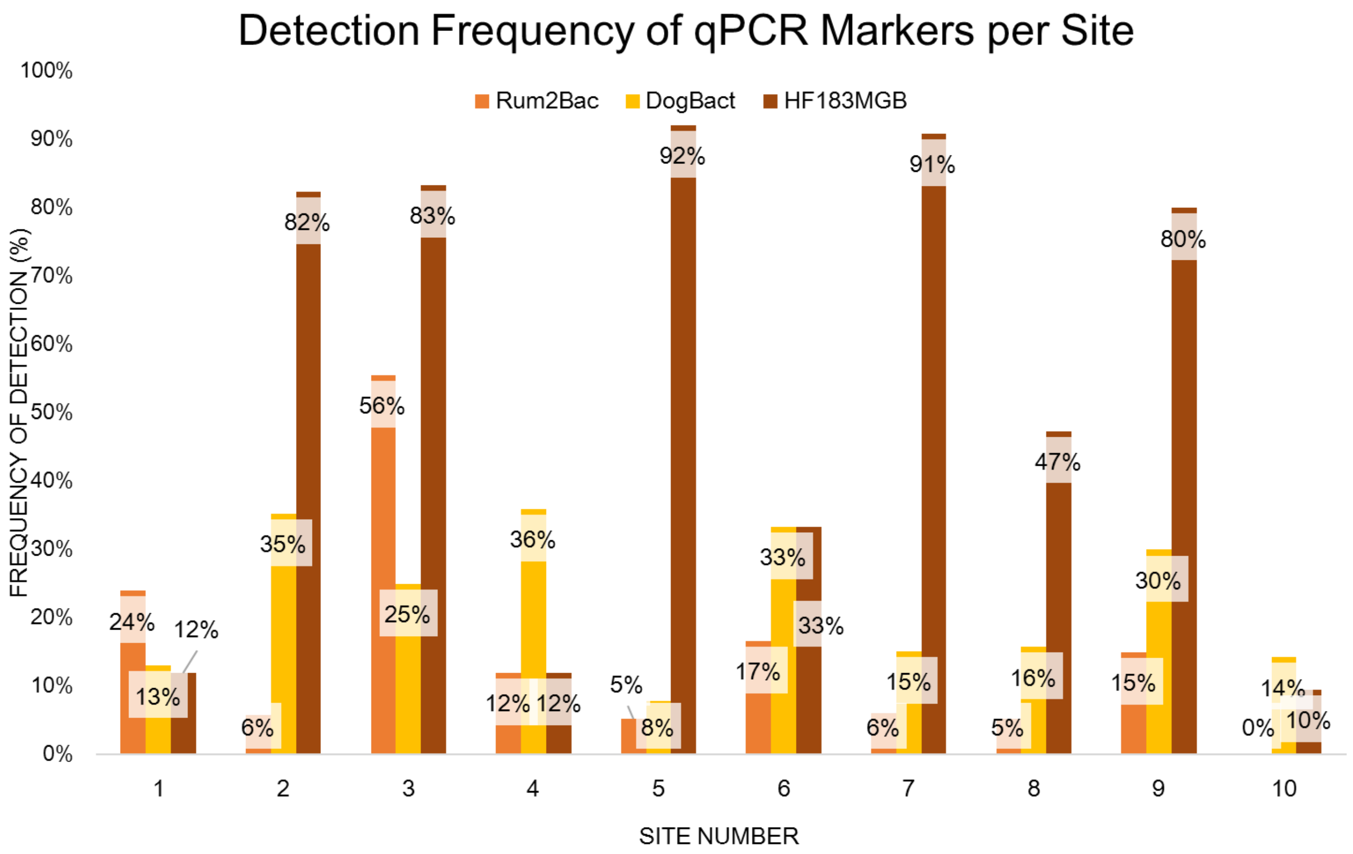


Table 1. Pearson correlation and P-values between qPCR marker concentration and molecular E. coli as a function of land use development.

Land Use Development Intensity	Site	qPCR Marker		
		Rum2Bac	DogBact	HF183MGB
High	5	0.13	0.16	0.64***
High	6	-	0.64	0.64
High	7	0.15	0.50***	0.41**
High	8	-0.428	0.36	0.08
High	9	0.19	0.47**	0.67***
Medium	2	0.21	0.58***	0.56**
Medium	3	0.56***	0.18	0.53***
Medium	10	-	0.11	0.13
Low	1	0.07	0.16	0.34*
Low	4	-0.12	0.56***	0.49**

Sites categorized by Land Use Development Intensity, High >79%, Medium 50-79%, Low<50%). Positive correlation value and significant P-value (*= 0.01), (**=0.05) and (***=0.001) indicated a positive trend between both variables.

United States Environmental Protection Agency

Take Home Message

- Detection of HF183MGB in highly developed areas of Athens may be an indication of failing sewer infrastructure.
- Significant correlations between source tracking markers and *E. coli* concentrations indicate that there are variable sources of contamination contributing with high *E. coli* numbers across the watershed which are independent of land use intensity.
- Little variation of *E. coli* concentration detected between methods used by volunteers and EPA scientists (average CV 3.29%), demonstrates analysis by volunteer scientists meet EPA QA/QC standards.

Next Steps

- Continue validating performance of models using MST markers
- Compare GBM vs. MLR performance across a variety of watersheds (e.g. urban systems)
- Validate model using citizen science collected data.

EPA United States Environmental Protection Agency

Acknowledgements

- Blake Snyder
- David Spidle,
- Mike Cyterski,
- Ken Bradshaw,
- Rania Georgacopolous
- Kelvin Wong
- Robert Sowah
- Kyler Herrington
- Kayleigh Hall
- Reid Brown, SSA-EPA, Athens
- Chris Fitzgerald, SSA-EPA, Athen**s**
- Ellen Price, SSA-EPA, Athens

Thank You!