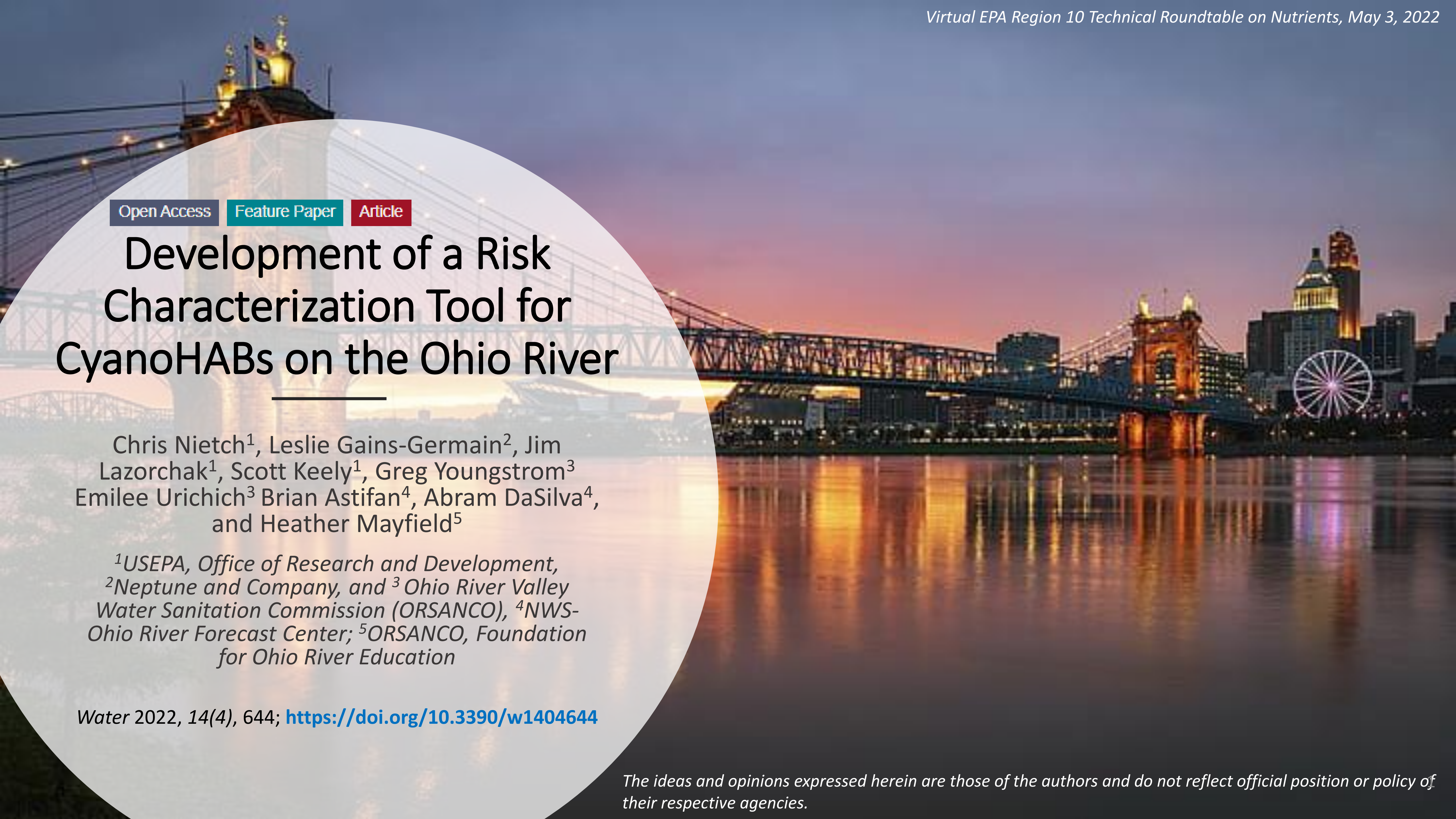


Open Access Feature Paper Article

Development of a Risk Characterization Tool for CyanoHABs on the Ohio River

Chris Nietch[1], Leslie Gains-Germain[2], Jim Lazorchak[1], Scott Keely[1], Greg Youngstrom[3] Emilee Urichich[3] Brian Astifan[4], Abram DaSilva[4], and Heather Mayfield[5]

[1]USEPA, Office of Research and Development, [2]Neptune and Company, and [3] Ohio River Valley Water Sanitation Commission (ORSANCO), [4]NWS-Ohio River Forecast Center; [5]ORSANCO, Foundation for Ohio River Education

Water 2022, *14(4)*, 644; **https://doi.org/10.3390/w1404644**

The ideas and opinions expressed herein are those of the authors and do not reflect official position or policy of their respective agencies.

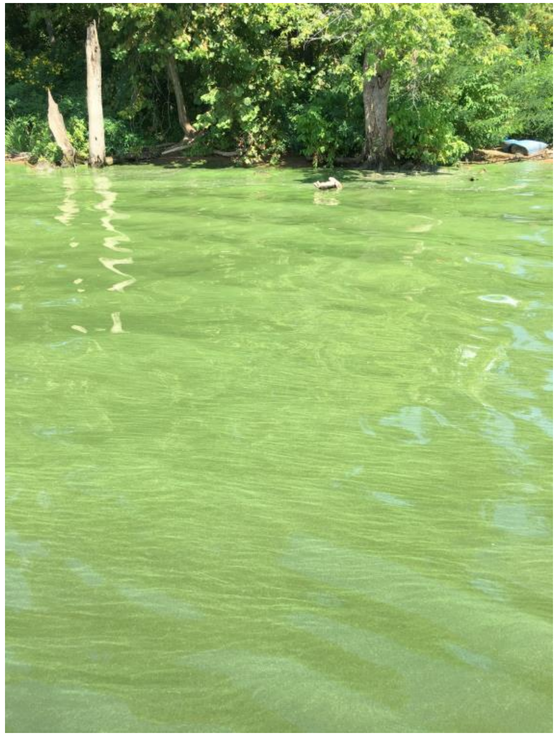
Frank Borsuk, EPA Region 3 Office in Wheeling WV received a phone call from ORSANCO in August 2015 that the lockmaster at the Pike Island Lock and Dam near Wheeling, WV reported a ‘green paint’ spill in the Ohio River. The ‘green paint’ was actually the start of the 2015 Harmful Algal Bloom that affected over 700 miles of the 900-mile Ohio River.

Microcystis aeruginosa – dominant species

Process

- Consider data sources, availability, and usability
- Match data availability with risk characterization objectives, i.e., provide a risk characterization for entire river’s length at anytime
- Determine statistical modeling approach
- Work with river managers to develop information organizational framework and build software application for end use
- This was a collaborative process across multiple agencies and areas of expertise that worked closely to develop the risk models and the visuals provided in the Shiny app

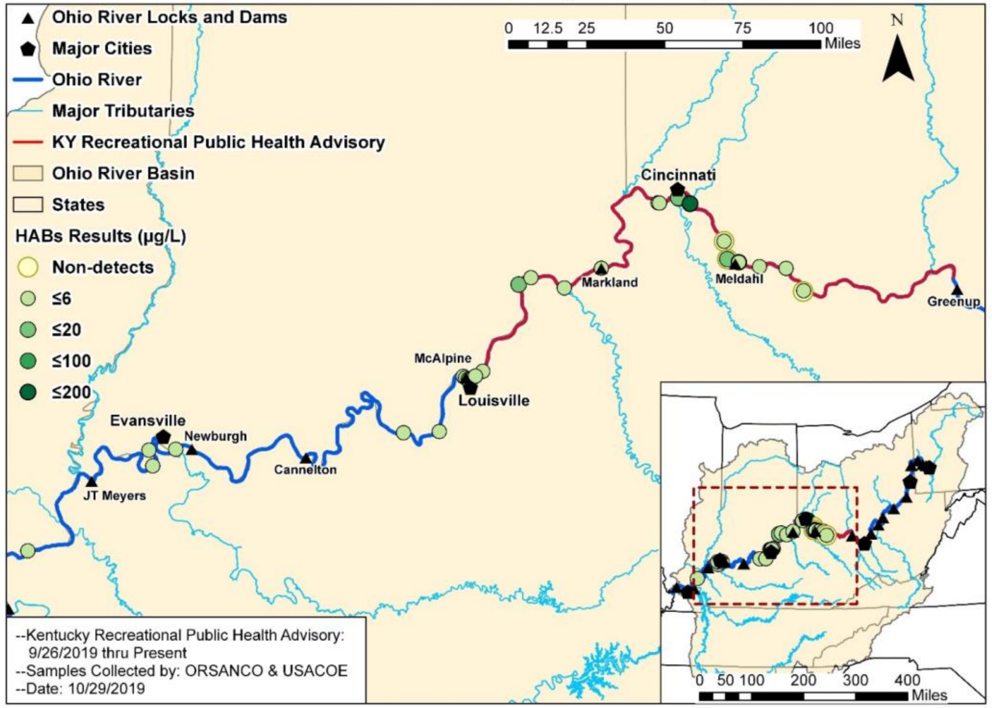
Ohio River HABs 09/12/2019



Russel DWTP Intake

Greenup L&D

Ohio River HABs 10/15/2019-10/24/2019

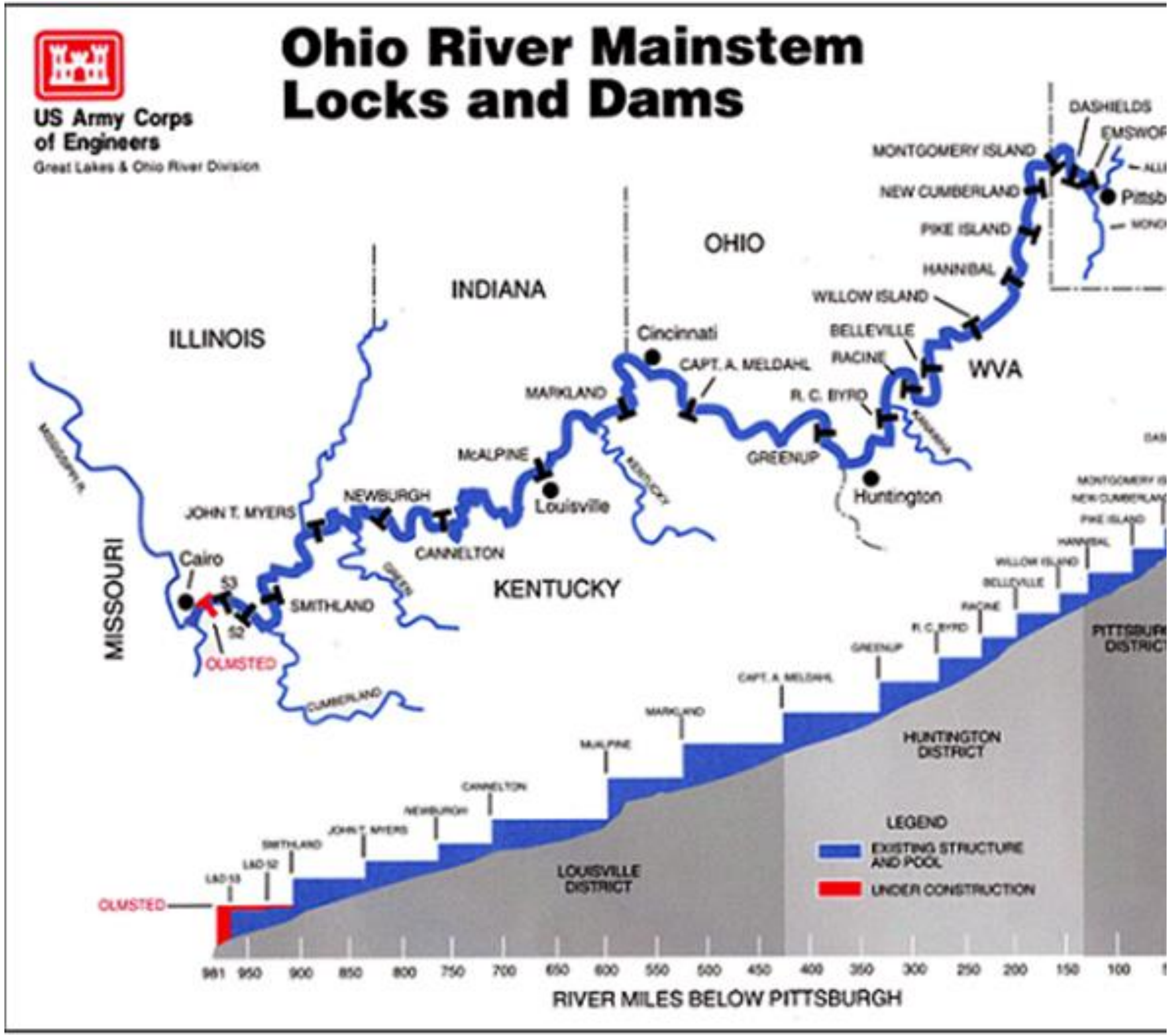


2019 Blooms

- First observed in Greenup Pool on 9/12/2019
- Covered 300 miles of river from Greenup to Louisville, but more spatial variability within pools compared to 2015
- Multiple species of *Microcystis* and high toxin concentrations

Water Data Availability

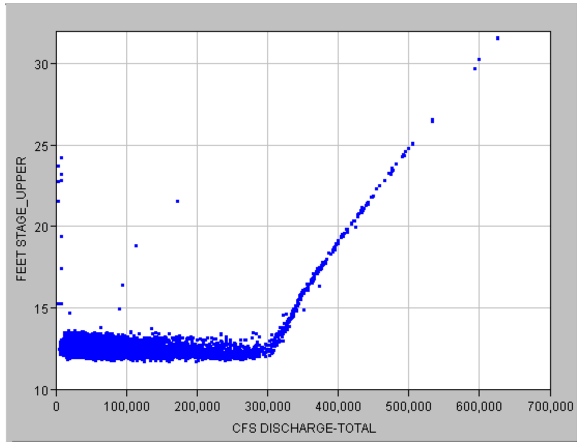
- Despite a lot of data, much of it did not lend itself to supporting the modeling objectives
- We needed multiple years, multiple sites, and high and on-going temporal resolution to support a real-time modeling tool
- HAB-related satellite imaging data was considered unreliable for the Ohio River
- When we started only one bloom had been 'documented'
- We ended up using flows from the lock and dam tailwaters and some free channel gauging sites



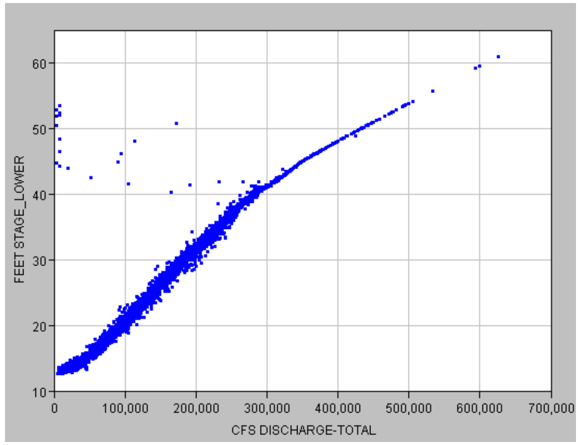
Estimating Discharge and establishing sites

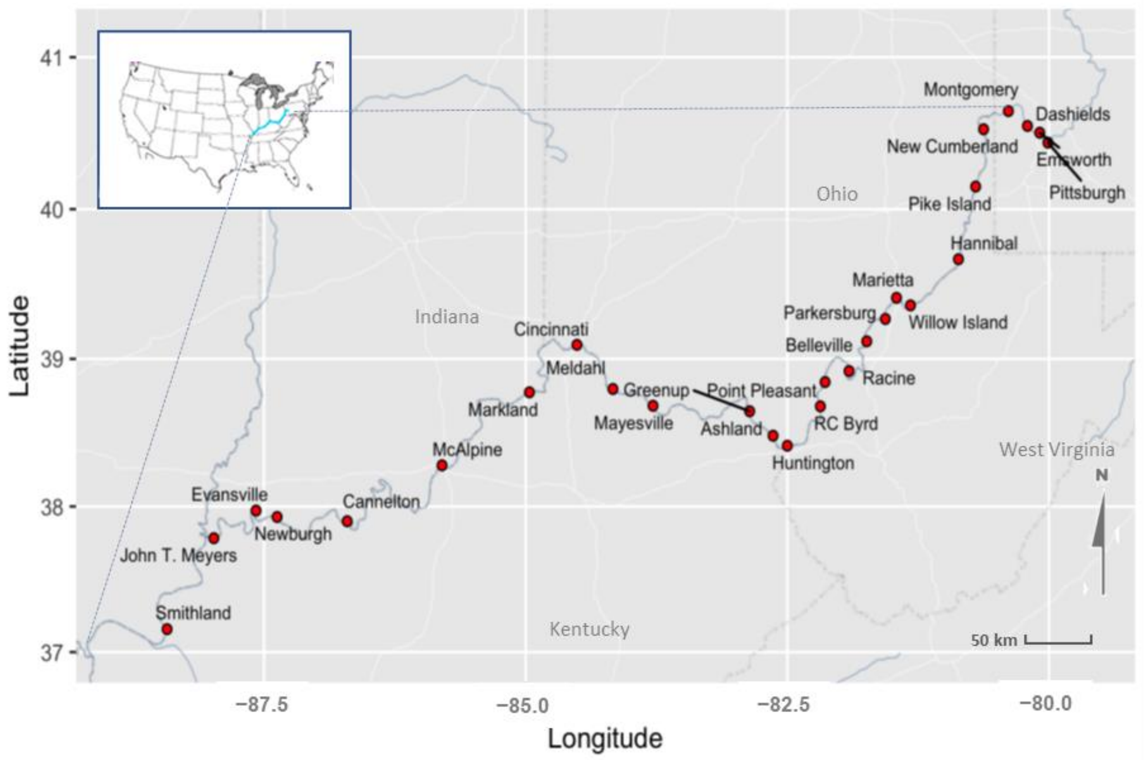
- Modeling goal: Use historical data for flow on the Ohio River to characterize hydrologic conditions that coincided with HABs in the late summers of 2015 and 2019
- Both high and, especially, low flow conditions are critical, but not all gaged sites can be used to accurately estimate discharge over the entire range of flows

(a) Meldahl L&D Upper Pool



(b) Meldahl L&D Tailwater

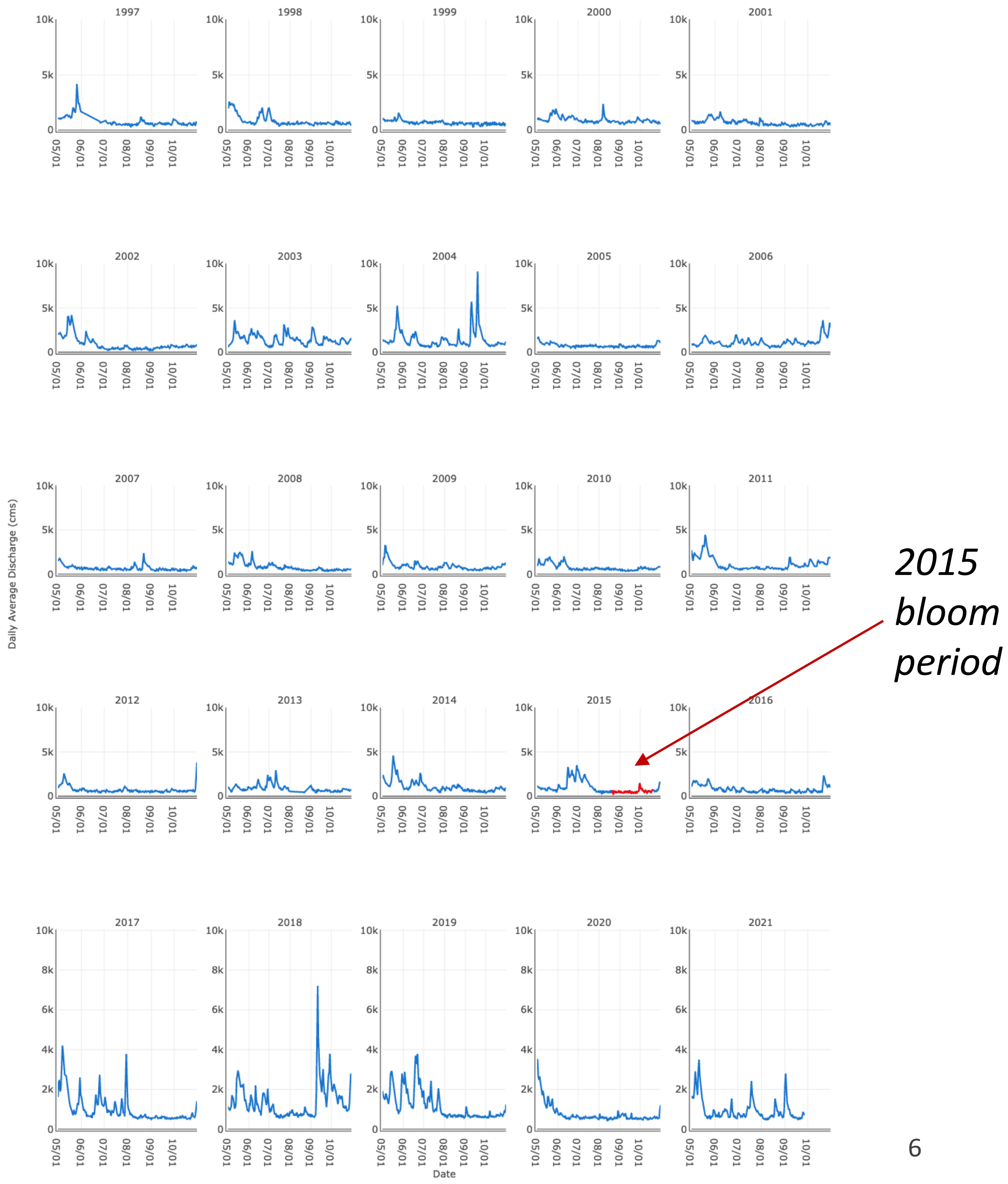


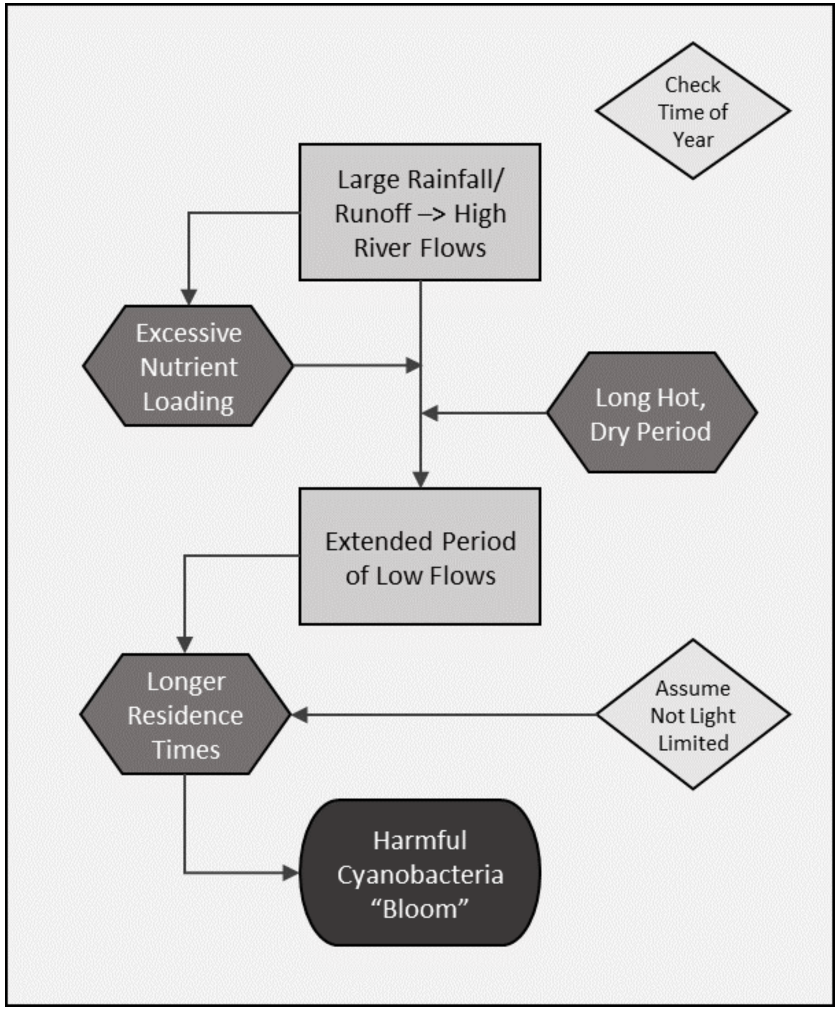


- *25 sites, spanning the entire river's length used for model development*

Exploratory Data Analysis

- Average daily discharge data for the Pike Island site plotted for 1995 through 2021, beginning of May to the end of October each year. Bloom first reported at Pike Island in 2015 (points in red signify bloom period).
- Note the unique flow dynamic in 2015
  - Period of high flows in late June/early July
  - Followed by long period of uninterrupted low flows





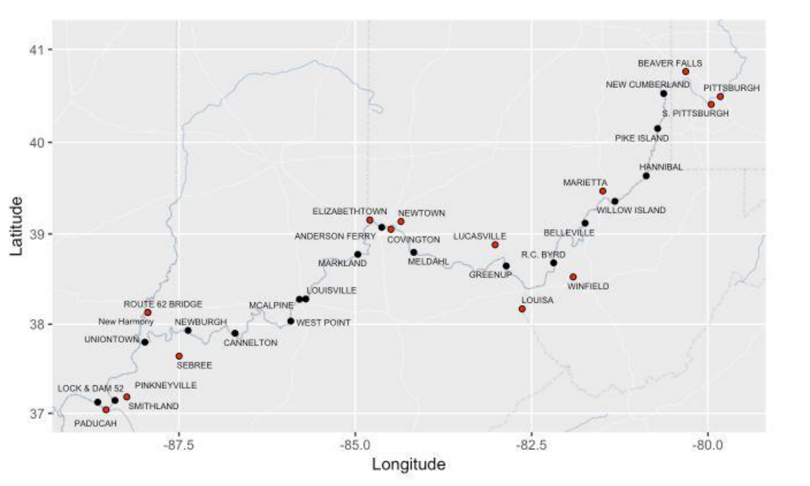
Guiding hypothetical cause and effects model linking cyanoHABs to preceding river flow conditions

<u>Hypothesis</u>

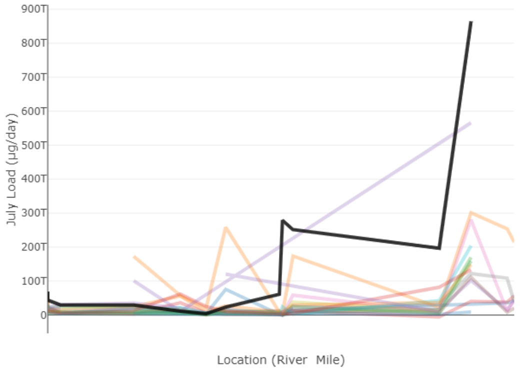
A period of high flows delivers excess nutrients to the river, which are effectively detained behind dams as an uninterrupted period of low flows ensues. As residence time of river's pools increases, they become weakly stratified, i.e., more lake-like, and this increases nutrient availability, favoring cyanobacteria proliferation.

July nutrient loads across years from monitored tributaries and main stem sites: 2002-2017

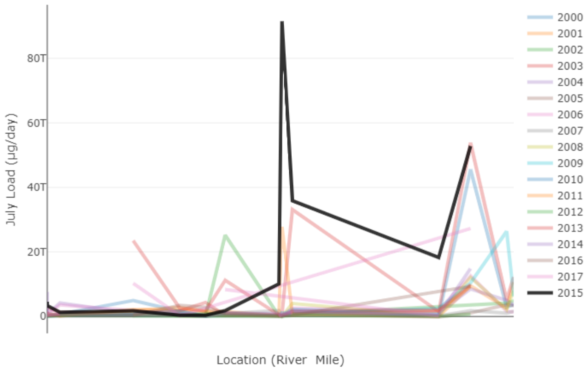
Supporting Evidence from Nutrient grab sampling site locations



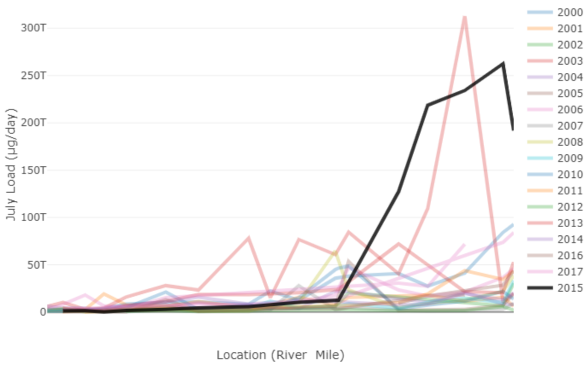
TN – Tributary Sites



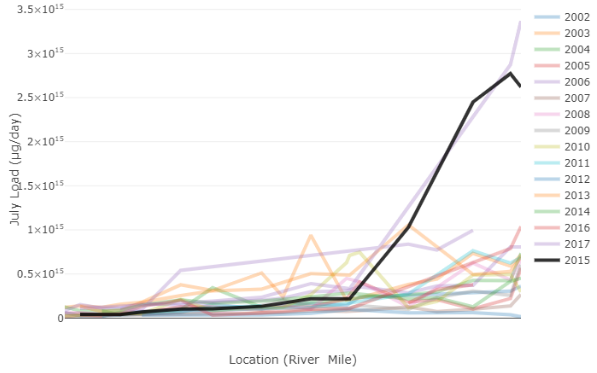
TN – Main Stem Sites



TP – Tributary Sites



TP – Main Stem Sites

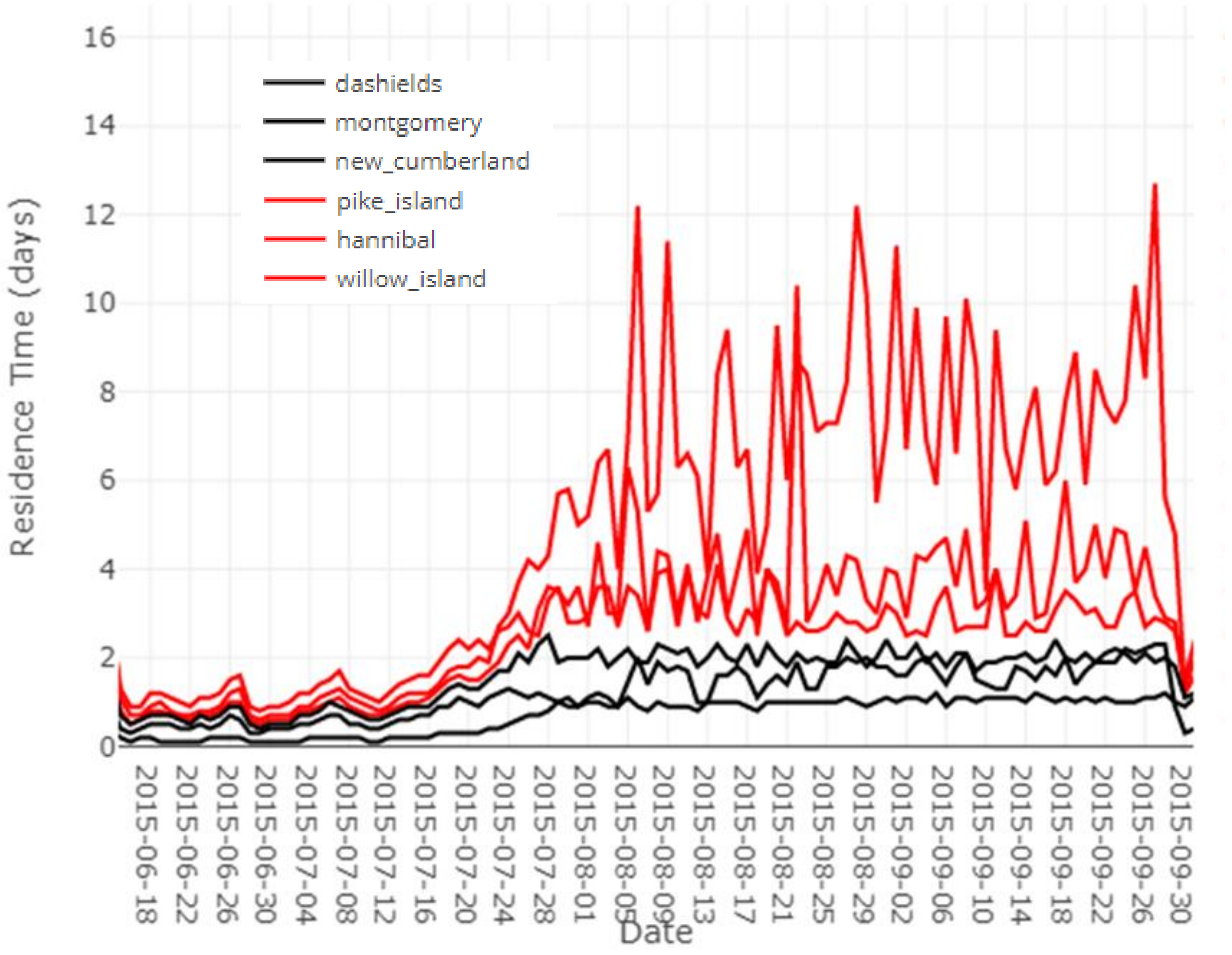


2015 HAB year in bolded black

- *Nutrient loading to and in the river for July 2015 was one of the highest years in the record, especially for mid and lower river sites*

Supporting Evidence from estimated residence times

- Residence time (or decrease in flushing rate) for the sites that bloomed (red) in 2015 compared with those that did not (black) in the upper river

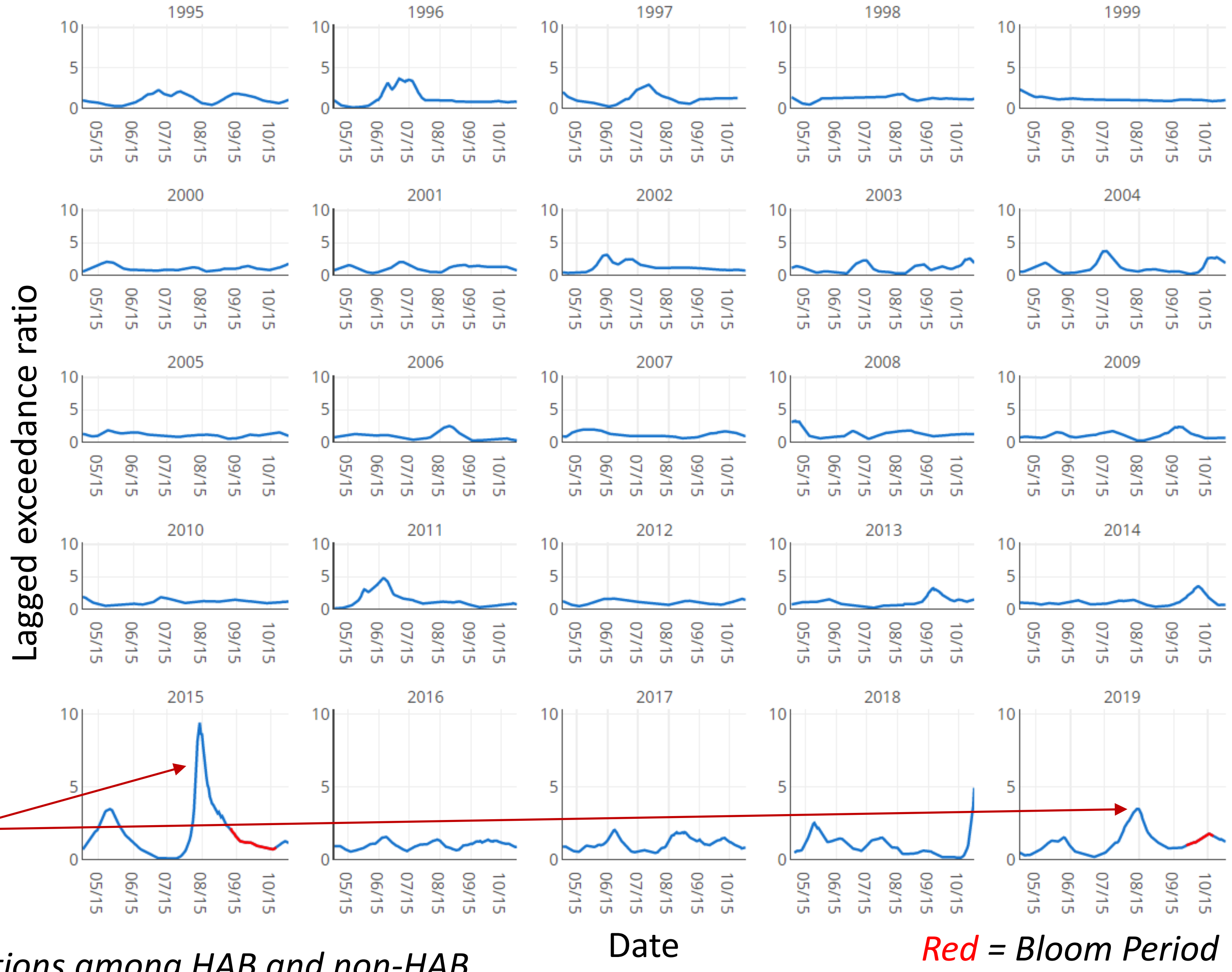


Characterizing the flow conditions that produce HABs

- Use lag terms that characterize hypothesized prerequisite flow conditions for HABs:
    - 21-55-day lagged exceedance captures high flow period
    - 19-day lagged exceedance captures low flow period
    - “best fit” lag terms determined using AIC
    - The ratio of the two is used for modeling

Exceedance ratio uniquely forecasting bloom period (in red)

Markland lock and dam site



- *The degree of similarity in flow conditions among HAB and non-HAB years at a site characterizes the relative risk in terms of a probability*

Two HAB Prediction Models:

Occurrence Model

$y_i \sim Bernoulli(p_i)$

logit(p_s)= α_{0s}+ β_0+(α_{1s}+β_1)*maxratio*+ β_2*inc15*+ β_3*meanrt*+ β_4*maxratio x meanrt*

$\alpha_{0s} \sim Normal(0,\sigma_0)$
$\alpha_{1s} \sim Normal(0,\sigma_1)$
$\beta_0 \sim Cauchy(0,2.5)$
$\beta_1 \sim Cauchy(0,2.5)$
$\beta_2 \sim Cauchy(0,2.5)$
$\beta_3 \sim Cauchy(0,2.5)$
$\beta_4 \sim Cauchy(0,2.5)$
$\sigma_0 \sim half\text{–}Cauchy(0,2.5)$
$\sigma_1 \sim half\text{–}Cauchy(0,2.5)$

Model Inputs:
- Location
- The current day's lagged exceedance ratio
- Number of days ratio has been increasing

Model Output:
Probability that current year will be a bloom year, with a 95% posterior interval

Persistence Model

logit(p_s)= α_{0s}+ β_0+(α_{1s}+β_1)*maxratio*+ β_2*ndays*+ β_3*ndays*^{2}+ β_4*I*+ β_5*I x maxratio* + β_6*meanrt*+ β_7*maxratio x meanrt*

b0j~Normal(0,σ0)
b1j~Normal(0,σ1)
β0~Cauchy(0,2.5)
β1~Cauchy(0,2.5)
β2~Cauchy(0,2.5)
β3~Cauchy(0,2.5)
β4~Cauchy(0,2.5)
β5~Cauchy(0,2.5)
β6~Cauchy(0,2.5)
β7~Cauchy(0,2.5)
σ0~half–Cauchy(0,2.5)
σ1~half–Cauchy(0,2.5)

Model Inputs:
- Location
- Maximum YTD ratio
- The # of days after maxratio occurred (ndays)
- I = binary threshold indicator for increasing flows:
  - threshold is passed if the 0-19d lagged average exceedance dropped by at least 15 in the previous 19d (indicating increased flows)

Model Output:
Probability that the bloom persists on the current day, with a 95% credible interval

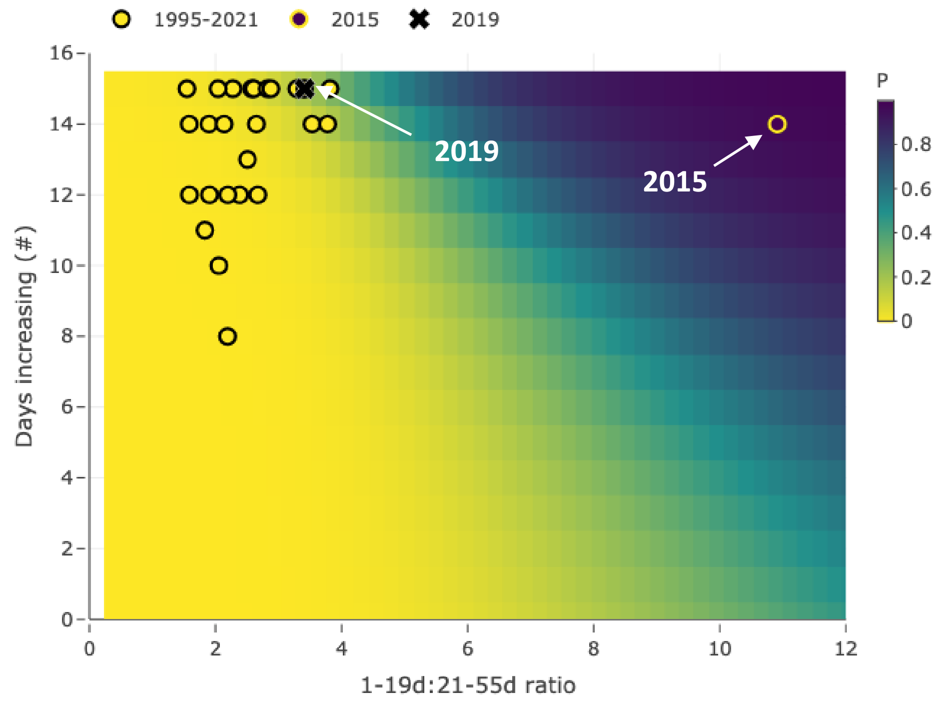
- *The limited number of events constrained formal model validation*
- *However, as a measure of performance, leave-one-out cross validation returned low misclassification rates*

** Mean residence time for a site's 19d lag for the 2015 event is an explanatory variable, but is fixed for a location, so does not need to be input in real time*

Model output visualization

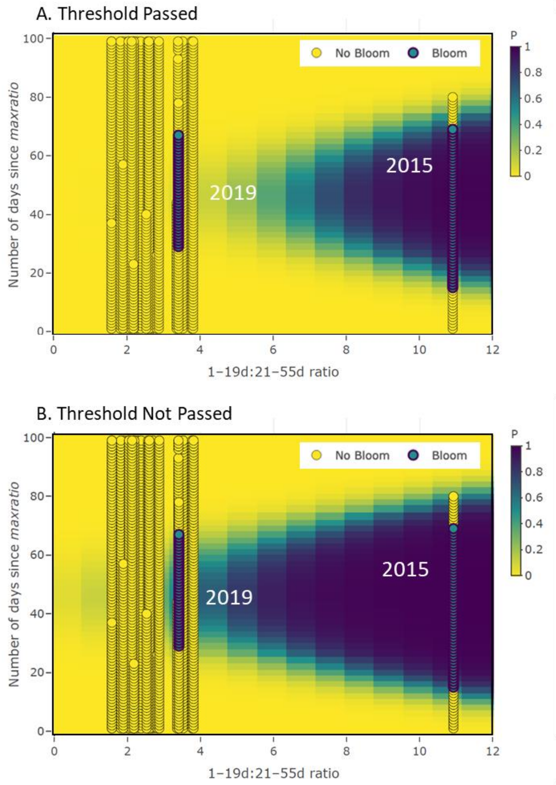
- Predicted probability (risk) with uncertainty of a bloom occurring and persisting
- Caveats:
  - Can only characterize the risk of new blooms in terms of 2015 and 2019 conditions
  - Models were fit despite some uncertainty about start and end dates
  - With more HABs, these models may or may not be better performing; this depends on how critical the specified flow conditions actually are to blooms

Occurrence



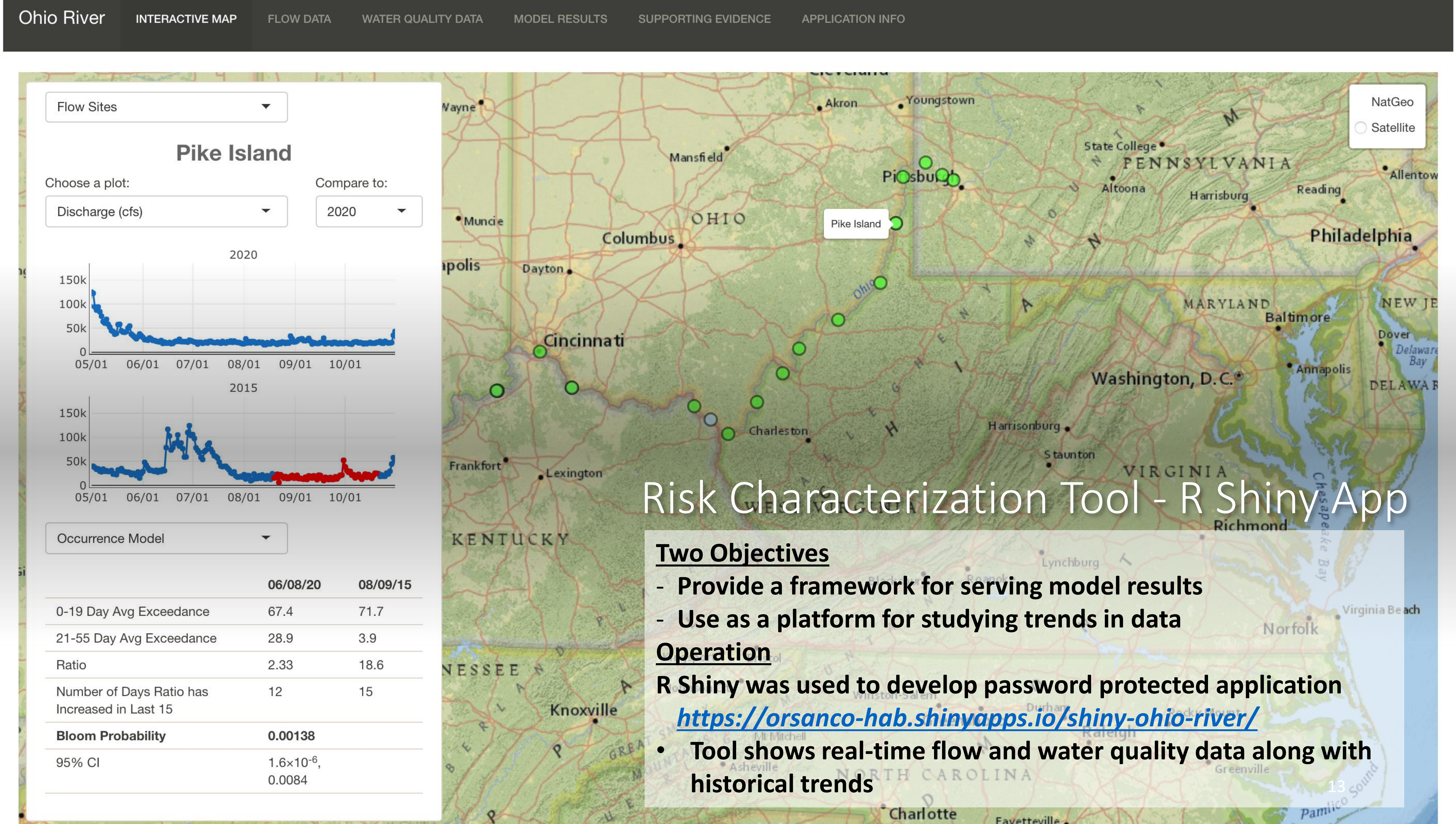
Example visualization of occurrence model results for the Greenup site

Persistence



Persistence model results for Greenup site. Yearly maxratio vs. the number of days since the maxratio occurred. Panels show how the persistence risk decreases when the flow threshold is passed (top)

Models are meant to be updated with new data each year. Input variables may need to change as more is learned about HABs on the river.



	06/08/20	08/09/15
0-19 Day Avg Exceedance	67.4	71.7
21-55 Day Avg Exceedance	28.9	3.9
Ratio	2.33	18.6
Number of Days Ratio has Increased in Last 15	12	15
Bloom Probability	**0.00138**	
95% CI	1.6×10^{-6}, 0.0084	

Risk Characterization Tool - R Shiny App

Two Objectives

- **Provide a framework for serving model results**
- **Use as a platform for studying trends in data**

Operation

R Shiny was used to develop password protected application

https://orsanco-hab.shinyapps.io/shiny-ohio-river/

- **Tool shows real-time flow and water quality data along with historical trends**

Water Quality Data Visualization Objectives

Eight sites are reporting continuous water quality data on the Ohio River (established between 2015 and 2019)

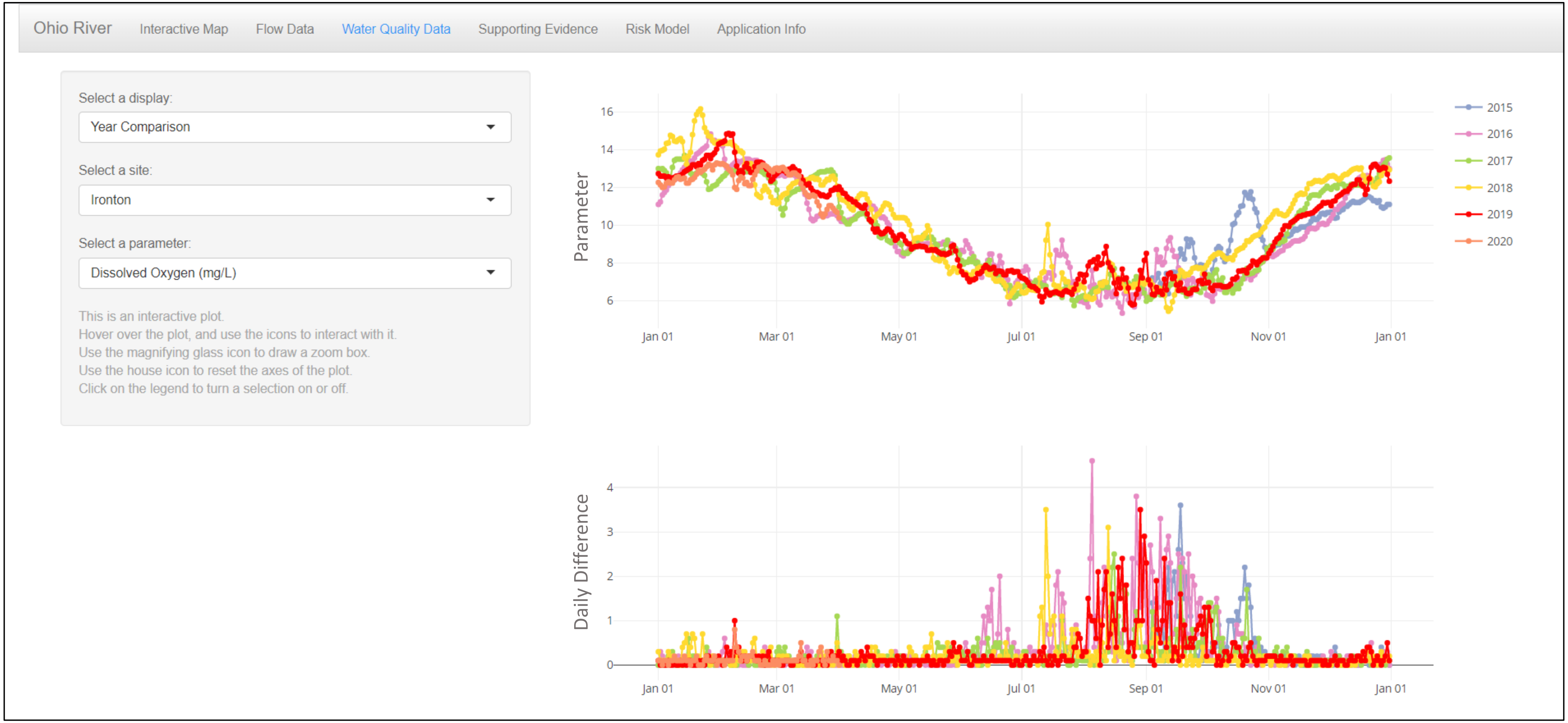
Configure the risk characterization tool so that the WQ information can be visualized in real time

Objectives for visualizing data in context of HABs modeling and active HABs management

- Compare data across sites
- Compare data across years
- Compare two variables at a time
- Report data as indicators of algal activity)

Current utility allows for data exploration in real time, which is critical for real-time HABs risk managment

Screen capture of water quality data page of Shiny app: Year comparison display



Lessons Learned

- For regulated big rivers, flow dynamics characterizations appear to be a powerful tool for developing HAB predictors. It is a ‘master’ variable, if you will. This is very different compared to lakes.
- Much risk management of HABs relies on satellite imagery and monitoring buoys that can be strategically positioned to capture water quality dynamics. Currently these tools are not as useful on rivers, so we must come-up with new approaches, like this one.
- Even if we had more WQ data it likely would not have improved our model’s prediction uncertainties, at least in the near term.
- Instead, we recommend, as we have done here, using the WQ data in a multi-prong approach. This aspect is not discussed often in the scientific literature.
- The Risk Characterization Tool is a unique application that combines data from multiple agencies (USACE, USGS, NWS, and ORSANCO). We hope that this effort will promote cross agency collaboration going forward.

Future Work

- The timing and magnitude of the preceding high flow period appears critical. We hypothesize this is related to nutrient availability as river flows decrease and the residence time of pools increase. It would be worthwhile to test this experimentally.
- We are working currently to use the National Weather Services ensemble hydrological forecasts to drive our prediction models. This will allow us to forecast bloom occurrence/persistence in a 1-to-2-month window preceding the period when we are likely to see blooms.
- We have very little control over climate change aspects fostering big river blooms. We can pay more attention to reducing nutrient loadings and the timing of those loads (i.e., if our hypothesis is correct). However, this is also going to take time. Therefore, supporting regulatory agencies like ORSANCO in assessing, sampling, and monitoring blooms, as we have done here, is critical to minimizing risk over the near future.

Acknowledgements

- Frank Borsuk, Regina Poeske Region 3; Carole Braverman, Meghan Hemken and Wendy Drake, Region 5; Robert Moyer and Erich Emery, USACE; Jason Heath and Richard Harrison, ORSANCO; Anna Springsteen, Neptune and Company; Jeff Vogt, Greater Cincinnati Water Works