Monitoring Cyanobacteria Blooms using Remote Sensing and Classification Trees

Katherine V. Hooker[1], Thayer C. Hallidayschult[1], Jie Wang[2], Jessica E. Beyer[1], Dani L. Glidewell[1], Xiangming Xiao[2],and K. David Hambright[1]

[1]Plankton Ecology and Limnology Lab University of Oklahoma Department of Biology, [2]Department of Microbiology and Plant Biology University of Oklahoma

INTRODUCTION:

Seasonally recurrent harmful algal blooms (HABs), especially of cyanobacteria, are a global problem. Due to the human health risks posed from ingestion and exposure to cyanotoxins, rigorous monitoring for HABs is necessary, but current monitoring approaches are expensive and often inadequate for reliably assessing human health risk from exposure to HAB toxins. Remote sensing is considered capable of offering an inexpensive, real-time tool for monitoring cyanobacteria in lakes and reservoirs, but systems to date have generally been based on complex multiple linear regression algorithms that are difficult to implement by the average lake managing authority. We hypothesized that simpler, yet reliable and robust, classification tress could be constructed using remote sensing data to produce a more practical solution for lake management.

METHODS:

DATA:

- The study system was Lake Texoma, a 303(d) listed reservoir on the border of Texas and Oklahoma (Fig. 1).
- We sampled five sites from 2006-2014 and five additional sites from 2012 -2014.
- We determined chlorophyll-a and phycocyanin concentrations through fluorometry as proxies for algal and cyanobacterial biomass.
- We extracted reflectance values from Landsat 7 images within an eight-day window of the sampling time and a 10% cloud-cover threshold.

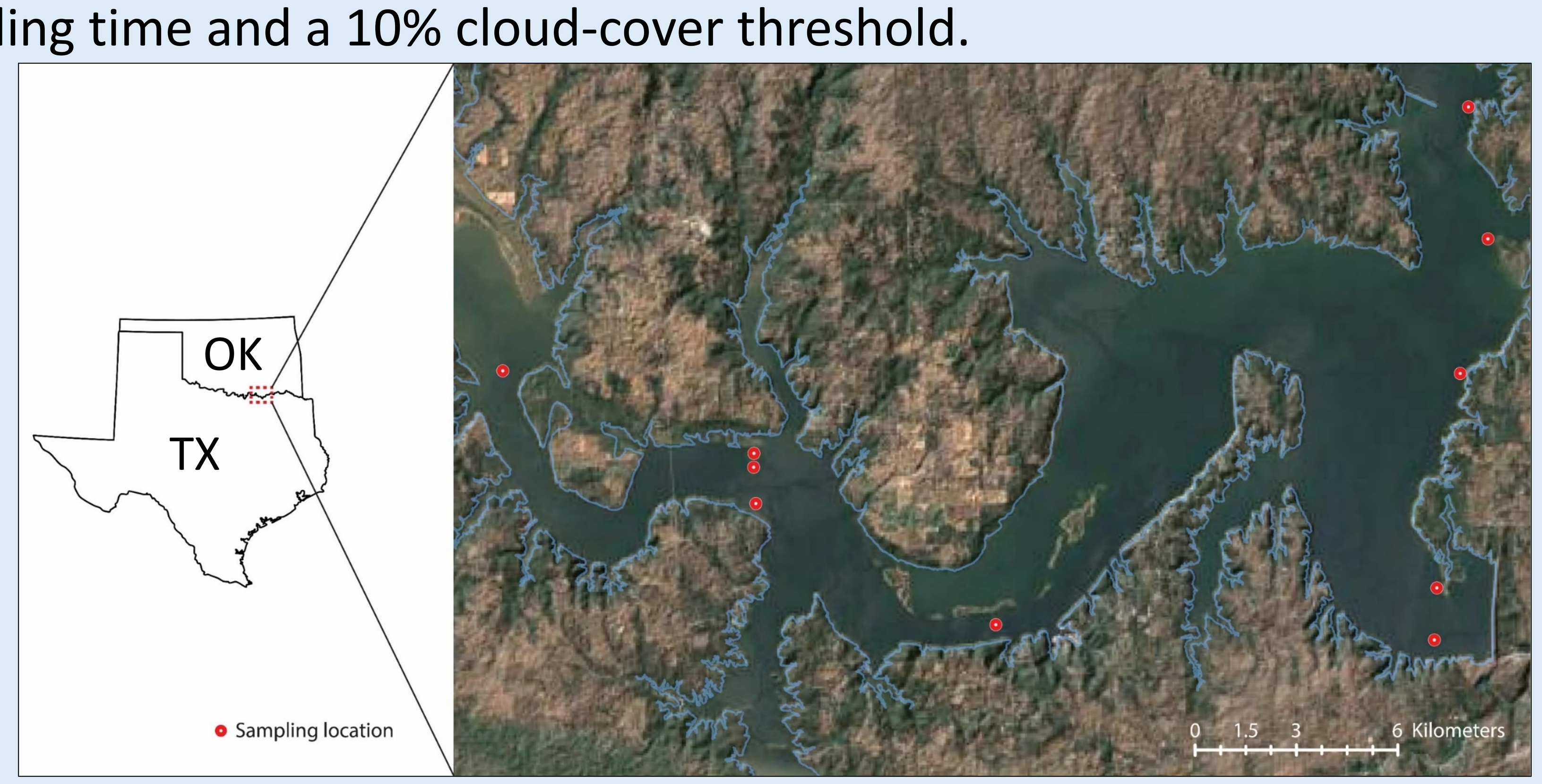

Fig. 1. Map of ten sampling locations on Lake Texoma, an impaired reservoir located on the border of Texas and Oklahoma. (Base map image credit: USGS)

MODEL CONSTRUCTION:

- Classification trees were created in R (R Core Team 2016).
- Predictor variables: Green:Red, Green:Blue, and Red:Blue reflectance ratios, and season
- Response variables:
  - Chlorophyll-a (50 µg/L threshold based on the World Health Organization's (WHO) guidelines (WHO 1999))
  - Phycocyanin (30 µg/L threshold based on the values of phycocyanin which where similar to 50 µg/L chlorophyll- a in Lake Texoma)

RESULTS:

Chlorophyll-a classification tree

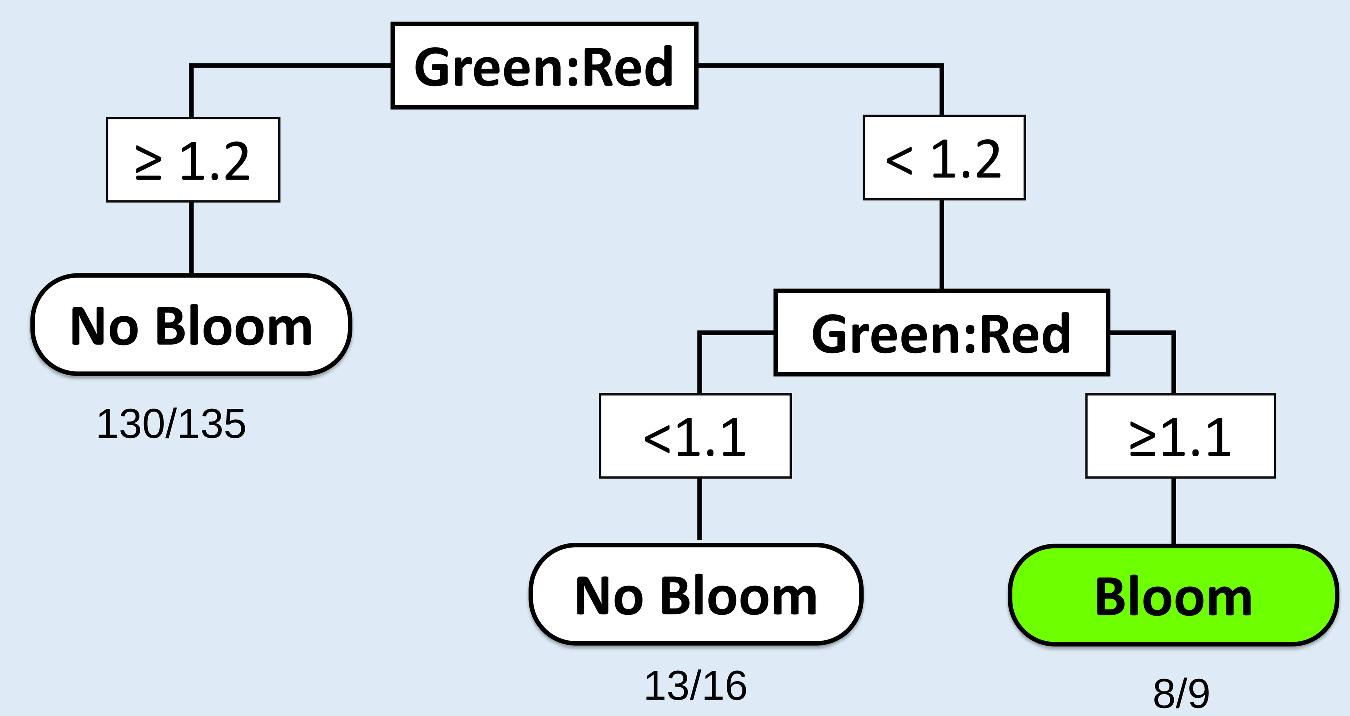

Fig. 2. Classification tree for chlorophyll-a built using chlorophyll-a data from Lake Texoma (n = 160). The threshold was set at 50 µg/L chlorophyll-a.

Phycocyanin classification tree

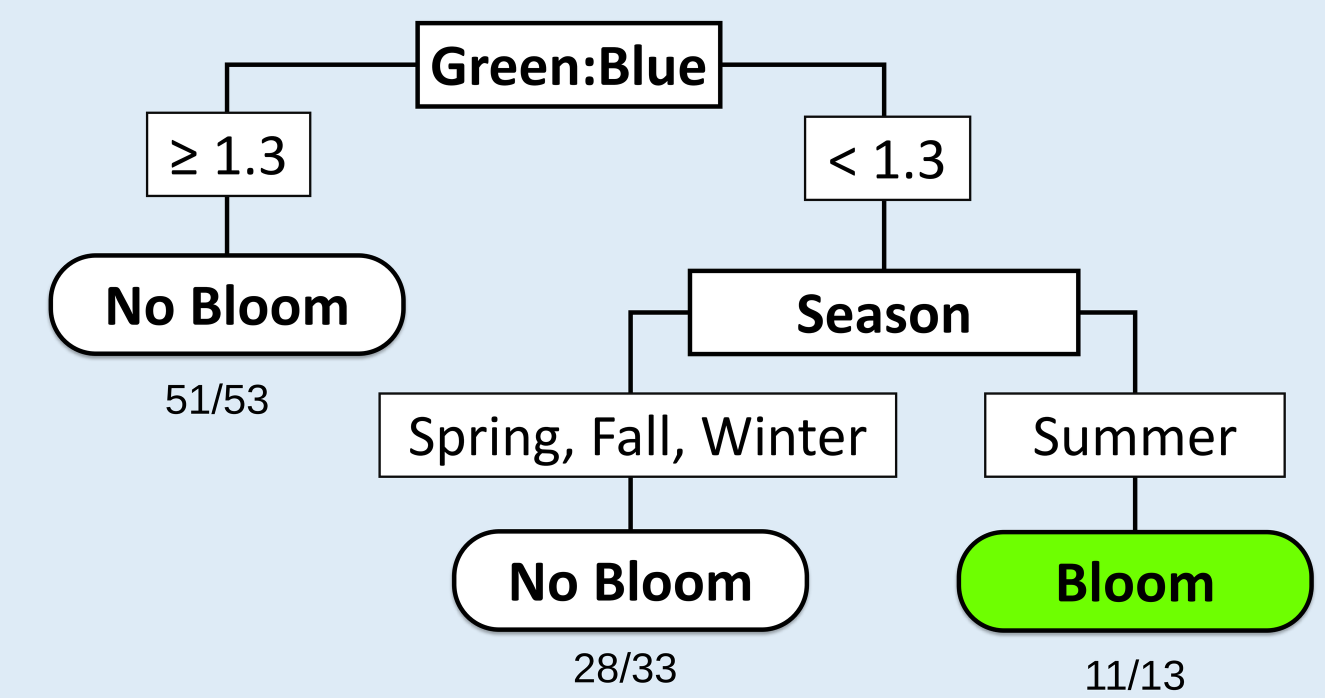

Fig. 3. Classification tree for phycocyanin built using phycocyanin data from Lake Texoma (n = 99). The threshold was set at 30 µg/L phycocyanin.

- Chlorophyll-a tree had nine recorded blooms and a misclassification rate of 5.6%
- Phycocyanin tree had 13 recorded blooms and a misclassification rate of 7.2%
- False positives occurred when the model predicted a bloom when there was not a bloom (Table 1);
- False negatives occurred when the model predicted there was not a bloom when there was a bloom

	Chl-a	Pcy
% Correct	94.4	91
% False Positive	0.6	2.0
% False Negative	5.0	7.0

Table 1. The percentage of times the tree correctly identified a bloom or no bloom, and the breakdown of the misclassification rate into false positives and false negatives.

CONCLUSIONS AND FUTURE GOALS:

- Classification trees provide a simpler decision mechanism than multiple linear regression algorithms for health risk management; using remote sensing of chlorophyll and phycocyanin, a lake manager could readily predict the presence of algal and cyanobacterial blooms in Lake Texoma with 94 and 91% accuracy.
- Future plans include testing and validating these models using additional lakes in the south central United States
- Our goal is to create a web portal for lake managers and lake patrons to access real-time remotely-sensed information on harmful algal blooms in reservoirs around the south central US

ACKNOWLEDGEMENTS: We would like to thank the Plankton Ecology and Limnology Lab at OU for Lake Texoma water sampling data as well as the Earth Observation and Modeling Facility for help with remote sensing images and analysis. In addition, we want to acknowledge our many collaborators(US Army Cops of Engineers, USGS, OWRB, and DEQ) and funding from the Oklahoma Water Resources Research Institute.

CITATIONS:

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

World Health Organization. "Toxic Cyanobacteria in Water: A guide to their public health consequences, monitoring and management." *London: E & FN Spon* (1999)