A Bayesian Multilevel Model for Microcystin Prediction in Lakes of the Continental United States

The frequency of cyanobacteria blooms in North American lakes is increasing. A major concern with rising cyanobacteria blooms is microcystin, a common cyanobacterial hepatotoxin. To explore the conditions that promote high microcystin concentrations, we analyzed the US EPA National Lake Assessment (NLA) dataset collected in the summer of 2007. The NLA dataset is reported for nine eco-regions. .

We used the results of random forest modeling as a means of variable selection from which we developed a Bayesian multilevel model of microcystin concentrations. Model parameters under a multilevel modeling framework are eco-region specific, but they are also assumed to be exchangeable across eco-regions for broad continental scaling. The exchangeability assumption ensures that both the common patterns and eco-region specific features will be reflected in the model. Furthermore, the method incorporates appropriate estimates of uncertainty.

Our preliminary results show associations between microcystin and turbidity, total nutrients, and N:P ratios. Upon release of a comparable 2012 NLA dataset, we will apply Bayesian updating. The results will help develop management strategies to alleviate microcystin impacts and improve lake quality.