Predicting Harmful Algal Blooms in Green Valley Lake Using a Machine Learning Model

Gu, Claire (School: Valley High School)

Cyanobacteria, or blue-green algae, are photosynthetic aquatic organisms that produce toxins. They can grow excessively to form large-scale blooms under favorable environmental conditions. These harmful algal blooms, known as CyanoHABs, can pose serious threats to human and animal health and aquatic ecosystems. In Iowa, lakes are susceptible to CyanoHABs due to excess nutrient runoffs from farm fields. Recently, CyanoHABs have been on the rise, but little research has been done to create models to predict CyanoHABs occurrences. The goal of this study was to develop a deep neural network (DNN) machine learning model to predict microcystin (a type of cyanobacterial toxin) and chlorophyll-a concentrations, two key indicators of CyanoHABs, using physiochemical water monitoring data collected from Green Valley Lake. The DNN model predicted results more accurately based on two error metrics than the traditional linear regression model developed for benchmarking. The study also identified nutrient conditions and seasonality as the two environmental variables that contributed most to CyanoHABs. Overall, the DNN model yielded good results using eight years of monitoring data. When more data becomes available in the future, this model can be used to effectively predict CyanoHABs occurrences in Iowa lakes to reduce the risk posed by them.

Awards Won: