**Project Name**: The influence of abiotic controls on algal blooms in freshwater lakes of Ontario, Canada.

**Team name & members**: Academia blooms (Meg Britt, Eric Massa, Kristen Hayward, Meghan Hamp)

BACKGROUND

Cyanobacteria are prokaryotic organisms that are ubiquitous in aquatic environments (Paerl and Otten 2013). Certain cyanobacteria can produce a range of bioactive secondary metabolites (Smith et al. 2008) that are capable of reducing water quality or even harming and killing pets and wildlife (Stewart and Falconer 2008). The unchecked growth of certain harmful species in aquatic ecosystems can have negative implications for humans and wildlife. A variety of abiotic and biotic factors contribute to the prevalence and population growth of cyanobacteria in water bodies. Physical characteristics of the waterbody such as temperature and stratification can be important indicators of potential cyanobacterial growth (Barros et al., 2019). In recent years, algal blooms have been increasing in their occurrence and a number of environmental factors have been contributing to this trend.

Traditional algae bloom prediction methods typically rely on satellite imagery, wind predictions, and modelling based on previous observations. These methods are most often applied to marine coastlines and are unable to make predictions on scales finer than 30km2. Therefore, traditional prediction methods are inappropriate for lakes such as Three Mile Lake (8km2). A recent study by Wang et al. (2018) found a predictable pattern between chlorophyll fluorescence and spikes in algal cell density in Lake Ehrai, Yunnan Province, China.  Wang et al. found that their results allowed them to reliably predict algal blooms 40 days before they occurred; a novel feat that has profound implications for ecosystem conservation and human health.

RESEARCH OBJECTIVE

**The aim of our study is to determine whether chlorophyll fluorescence, and/or other abiotic parameters, can be used as as early** **predictors of harmful algal blooms in Canadian freshwater systems.**

**Hypothesis 1:** Algal blooms, indicated by water chlorophyll a content, can be predicted based on relationships that exist between abiotic environmental conditions and chlorophyll a fluorescence.

**Hypothesis 2:** The rate of change observed for abiotic environmental conditions can be used to predict algal blooms, as indicated by chlorophyll a concentrations in water.

DATA DESCRIPTION

Three Mile Lake is an 8.8 km2 lake with a shoreline perimeter of 31.2 km, surrounded by lakefront cottages in Muskoka, Ontario. It is composed of two basins, with maximum depths of 4 and 12 m. Both basins have a nutrient content that is notably higher than the rest of the freshwater in the region.

Data consist of 10-minute averages of continuously measured environmental variables, stored in a 100 MB .csv file. Each data point is indexed by the date and time at which it was recorded, with approximately 25000 measurements recorded in the ice-free season of 2018. Water temperature was measured every 25 cm from 10 cm below surface level to 10 m below surface level. Measurements of conductivity (milliSiemens/meter) and pH (dimensionless), Chlorophyll a (mg/L), and relative phycocyanin fluorescence (percentage) at 1 m below surface level were also recorded under the same conditions.

PIPELINE

Our bioinformatic pipeline will select and clean up the appropriate data columns for use in (1) addressing our hypotheses with linear model selection, and (2) visualization of these results.

Our pipeline will be composed of 4 modules:

1.     **Data exploration and subsetting: Meg Britt will use Bash/Unix to** explore thedata/evaluate preliminary trends, subset the data (only include fluorescence, conductivity, pH, surface temperature, radiance, chlorophyll a), and exclude any rows with NAs. (wget, cat, awk, grep, etc) *\*input = raw data, output = subsetted data with no NAs*

**2.**     **Data modification and outlier checks: Eric will use relevant packages in R to:**

a.     Plot histograms of each variable to approximately assess normality (*ggplot2*)

b.     Determine appropriate transformations for non-normal variables using symbox() function (*car*)

c.     Transform appropriate variables using mutate() function and piping (*dplyr*)

d.     Perform outlier checks using influenceIndexPlot() (*car*)

*\*input = data from (1), output = transformed data with no outliers*

**3.**     **Data analysis and assumption checks: Kristen will use relevant packages in R to:**

a.     Perform backwards model selection using Anova() and anova(), starting with the full model (chlorophyll a ~ fluorescence + conductivity + pH + surface temperature + radiance)

b.     Perform AICc model selection (*MuMIn*)

c.     Calculate standardized partial regression coefficients to determine the relative strength of each predictor

d.     Assess assumptions of normality and homoscedasticity with autoplot() (*cowplot(*)

*\*input = data from (2), output = minimum adequate model with p-values and standardized regression coefficients for each significant predictor*

4.     **Data visualization – Meghan Hamp will use ggplot2 in R to** plot how well the selected model fits our data, as well as create added variable plots*. \*input = data from (2) & (3), output = formal composite scatterplots/added variable plots with regressions*

**Team coordination**: We will use a github folder in Kristen’s repository with subfolders for the raw data (untouched) and each of the modules from the pipeline. Each of these folders will include annotated scripts, the input data, and the output data. Each team member will have access to these folders and will be able to easily find the input data for their module.

BROADER SIGNIFICANCE

Throughout North America, tourism companies, municipalities, and property owners have experienced the economic, environmental and health impacts of harmful algal blooms. Ontario cottages face similar issues, but an in depth understanding of the abiotic factors that could be use to predict potential blooms is lacking. Algal blooms can have dramatic effects on nearby regions as blooms can produce dangerous toxins and foul taste and odor compounds. These toxins can concentrate in aquatic animal and result in harm to pets and humans. Taste and odor issues on the other hand, can result in distrust in the safety of local drinking water and negatively affect tourism and cottage industries. As such, local governments and community organizations have a significant incentive to predict, identify, and manage the risk of algal blooms.

Three Mile Lake in Muskoka region has a history of algal blooms, including in 2018. The recent exposure of community members of the Muskoka area to these algal blooms provides an important opportunity to improve our understanding of the relevant abiotic factors that contribute to the prevalence of blue green algae in our freshwater systems.

PREDICTIONS

We know from the dataset that there was a notable algal bloom in Three Mile Lake during the sampling period. We predict that the spike in chlorophyll a concentration (mg/L) during the bloom, a proxy for algal biomass, will be preceded by a steep increase of an abiotic factor (Figure 1). Based on the results of Wang et al. the most likely predicting variable of an increase in algal biomass is fluorescence (%).

APPENDIX

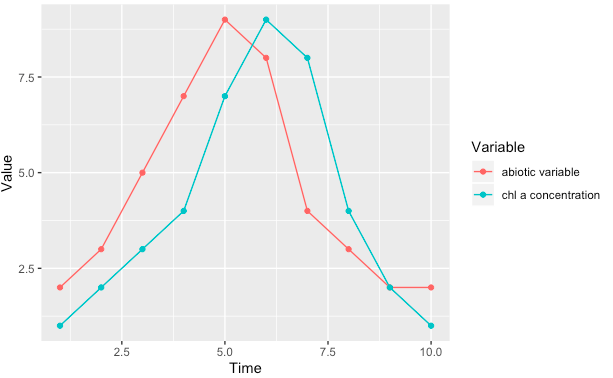


Figure 1. Prediction of how the chlorophyll a concentration will be influenced by on eor more predictor variables (fluorescence (%), temperature, conductivity, pH, surface temperature, radiance).

Table 1: Description of variables used in analysis with units. All variables above are recorded as 10-minute averages of continuously-measured variables. Additional variables were recorded in the raw data but were excluded from analysis. Due to limitations in the project scope, we eliminated X of the Y water temperature variables (measured every 25 cm starting at 10 cm below surface level to 10.1 m below surface level), restricting ourselves to only the water temperature at 0.1 m below surface level. Other measurements that we have eliminated from analysis include atmospheric conditions, dissolved oxygen in water, and phycocyanin pigment data, again due to limits in project scope.

|  |  |  |
| --- | --- | --- |
| Variable name | Units | Description |
| Water temperature | ° C | Water temperature 0.1 m below surface level. |
| Conductivity | milliSieverts | A measure of total dissolved solids in water (particularly ions) at ambient temperature. |
| Longwave radiation | Watts/m2 | A proxy for photosynthetically available radiation, reflecting the sunlight that makes it to the lake’s surface. |
| Fluorescence | % | Relative chlorophyll fluorescence. |
| pH | unitless | Water pH, indicating the free hydrogen ion content of the water. |
| Chlorophyll a | mg/L | A proxy for algal cell presence in the water. |

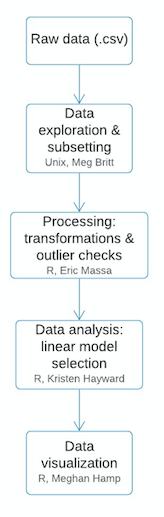


Figure 2. Flowchart illustrating the outline to our overall pipline for the project and the team member in charge of each step of the pipeline.

REFERENCES

Barros, M., Wilson, A., Leitao, J., Pereira, S., Buley, R., Fernandez-Figueroa, E., & Capelo-Neto, J. (2019). Environmental factors associated with toxic cyanobacterial blooms across 20 drinking water reservoirs in a semi-arid region of Brazil. *Harmful Algae*, 86:128-137.

Darki, B., & Krakhmalnyi, A. (2019). Biotic and abiotic factors affecting the population dynamics of Certium hirundinella, Peridinium cinctum, and Peridiniopsis elpatiewskyi. Diversity, 11(137):1-13.

Paerl, H.W., & Otten, T.G. (2013). Harmful cyanobacterial blooms: causes, consequences and

controls. *Microbial Ecology*. 65: 995–1010.

Smith, J.L., Boyer, G.L., Zimba, P.V. (2008). A review of cyanobacterial odorous and bioactive metabolites: impacts and management alternatives in aquaculture. *Aquaculture*, 280:5–20.

Stewart, I., & Falconer, I.R. (2008). Cyanobacteria and cyanobacterial toxins. In: Walsh, P.J.,

Smith, S.L., Fleming, L.E., Solo-Gabriele, H.M., Gerwick, W.H. (Eds.), Oceans and

Human Health: Risks and Remedies from the Seas. Academic Press: 271–296.

Wang, H., Zhu, R., Zhang, J., Ni, L., Shen, H., & Xie, P. (2018). A Novel and Convenient Method for Early Warning of Algal Cell Density by Chlorophyll Fluorescence Parameters and Its Application in a Highland Lake. *Frontiers in Plant Science*, *9*. doi: 10.3389/fpls.2018.00869