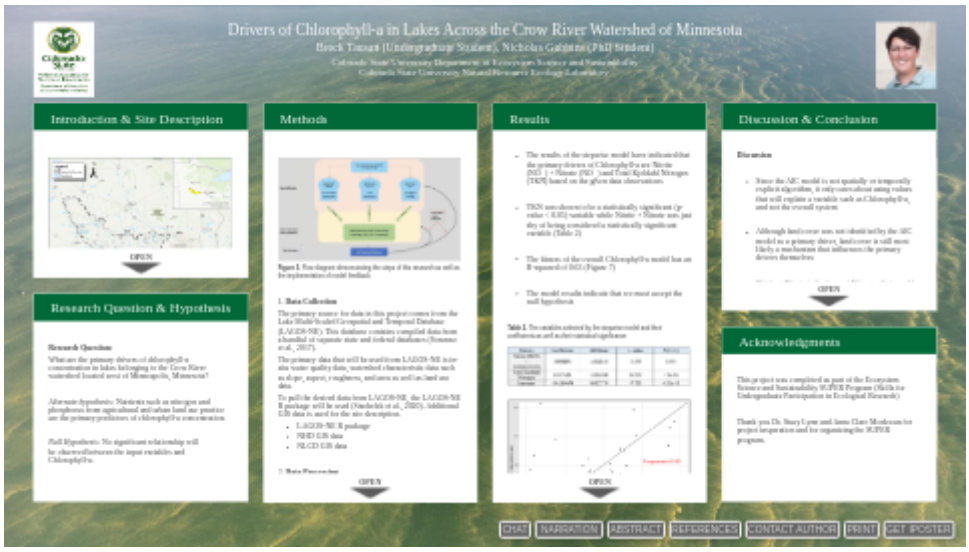
Drivers of Chlorophyll-a in Lakes Across the Crow River Watershed of Minnesota



Brock Tausan (Undergraduate Student), Nicholas Gubbins (PhD Student)

Colorado State University Department of Ecosystem Science and Sustainability
Colorado State University Natural Resource Ecology Laboratory

PRESENTED AT:

COLORADO STATE UNIVERSITY | CURC 2021 CELEBRATE UNDERGRADUATE RESEARCH AND CREATIVITY

INTRODUCTION & SITE DESCRIPTION

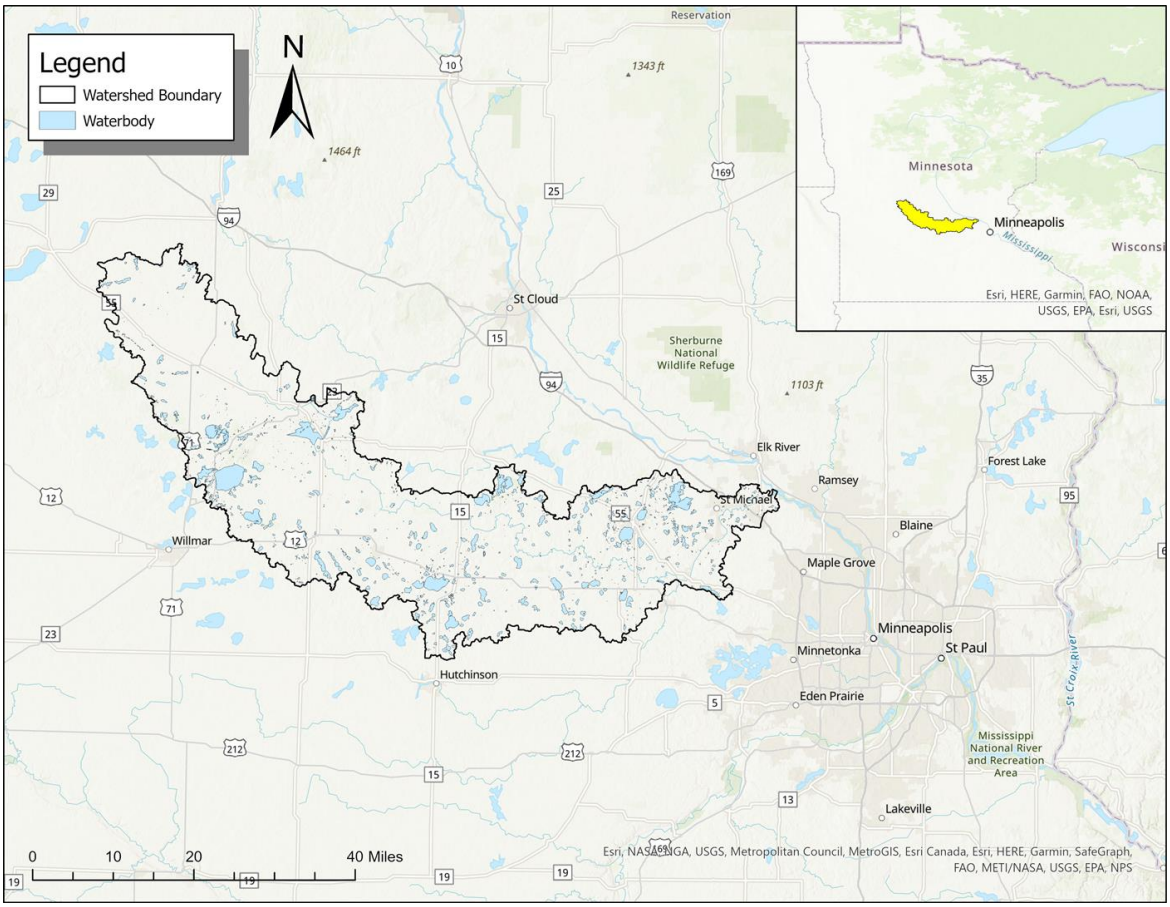


Figure 1. Crow River watershed boundary obtained from the USGS National Map Viewer's National Hydrography Dataset (NHD)

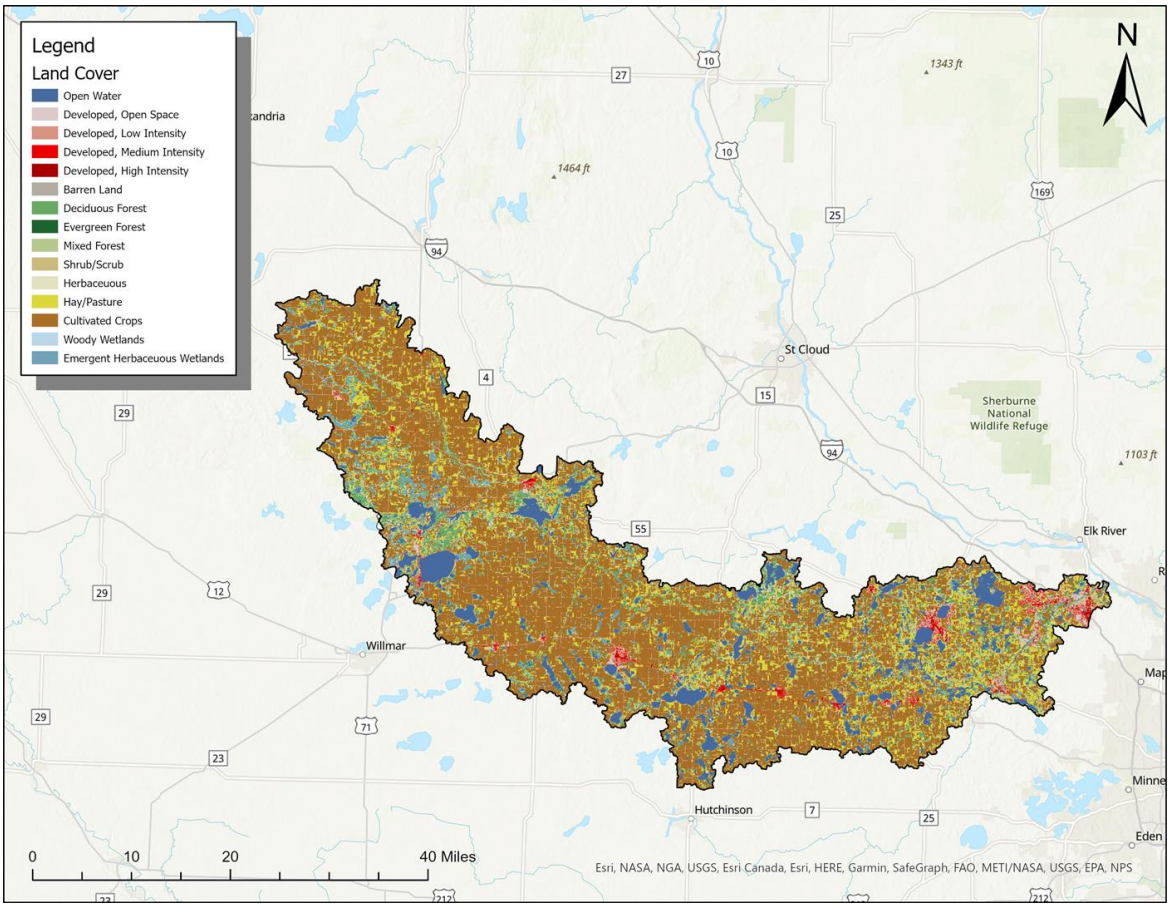


Figure 2. Crow River watershed land use and land cover data obtained from the NRCS geospatial data gateway's National Landcover Data Set (NLCD)

Large scale land use and land cover change such as agriculture and urban development can cause an increase in the amount of nutrients that reach lakes and streams (Carpenter et al., 2011).

Nutrients such as nitrogen and phosphorus, which are widely used ingredients in fertilizers used in agriculture, can cause increased algal growth in lakes (Anderson et al., 2002).

Algal blooms are a naturally occurring phenomenon in all bodies of water, however the effect of different anthropogenic land use practices has become an area of concern in the scientific community for several reasons including their adverse effects on fish and wildlife as well as humans (Zhang et al., 2009).

Chlorophyll-a, which is used in oxygenic photosynthesis by plants such as algae can be used as an indicator of for algal blooms (Carpenter et al., 2011).

- The Crow River watershed is located west of Minneapolis, Minnesota and is 949,107 acres (Figure 1).
- The primary LULC is agriculture with scattered development, occurring more frequently in the eastern portion of the watershed (Figure 2).
- Many of the lakes in this watershed do not meet State water quality standards because of summer algal blooms (North Fork Crow River, 2011)

RESEARCH QUESTION & HYPOTHESIS

Research Question:

What are the primary drivers of chlorophyll-a concentration in lakes belonging to the Crow River watershed located west of Minneapolis, Minnesota?

Alternate hypothesis: Nutrients such as nitrogen and phosphorus from agricultural and urban land use practice are the primary predictors of chlorophyll-a concentration.

Null Hypothesis: No significant relationship will be observed between the input variables and Chlorophyll-a.

METHODS

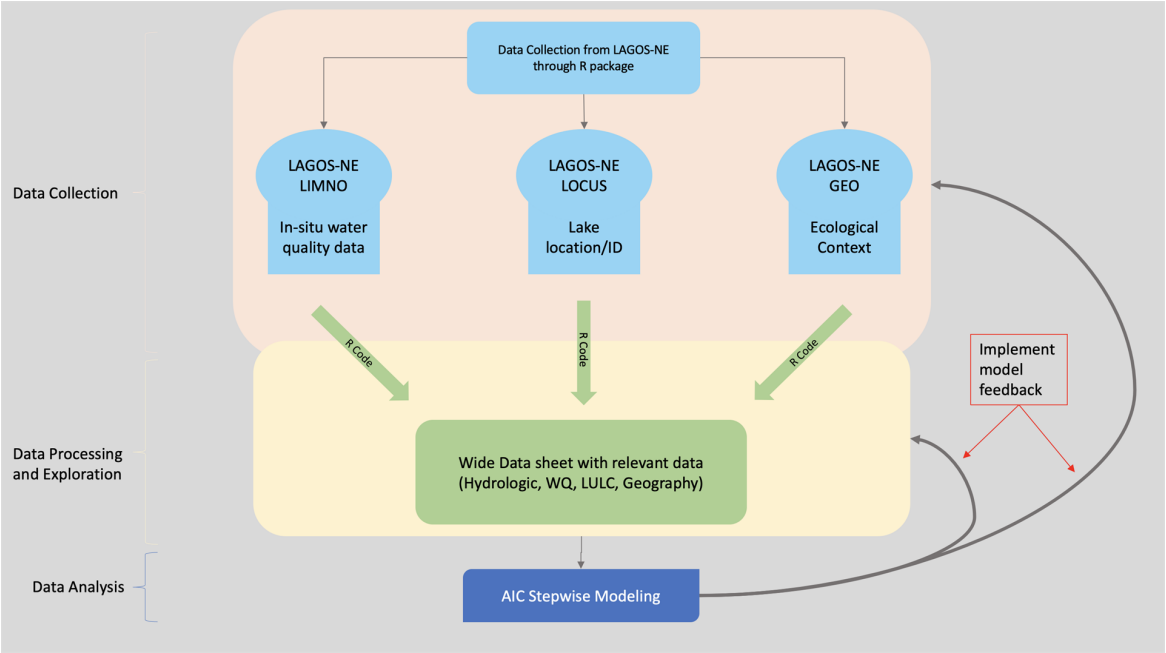


Figure 3. Flow diagram demonstrating the steps of this research as well as the implementation of model feedback

1. **Data Collection**

The primary source for data in this project comes from the Lake Multi-Scaled Geospatial and Temporal Database (LAGOS-NE). This database contains compiled data from a handful of separate state and federal databases (Soranno et al., 2017).

The primary data that will be used from LAGOS-NE is in-situ water quality data, watershed characteristic data such as slope, aspect, roughness, and area as well as land use data.

To pull the desired data from LAGOS-NE, the LAGOS-NE R package will be used (Stachelek et al., 2020). Additional GIS data is used for the site description.

- LAGOS-NE R package
- NHD GIS data
- NLCD GIS data

2. **Data Processing**

There is no data entry needed in this project, however there is lots of data processing and exploration required.

- *Exploratory data analysis*
  - Using different methods in R, the data must be cleaned and filtered to the variables of interest
  - Comparing frequent in-situ chl-a measurements to hydrologic and land use data will require that the water quality data be harmonized.
  - Chlorophyll-a data must be compared to other variables so that any covariance can be verified before feeding them to the stepwise model.

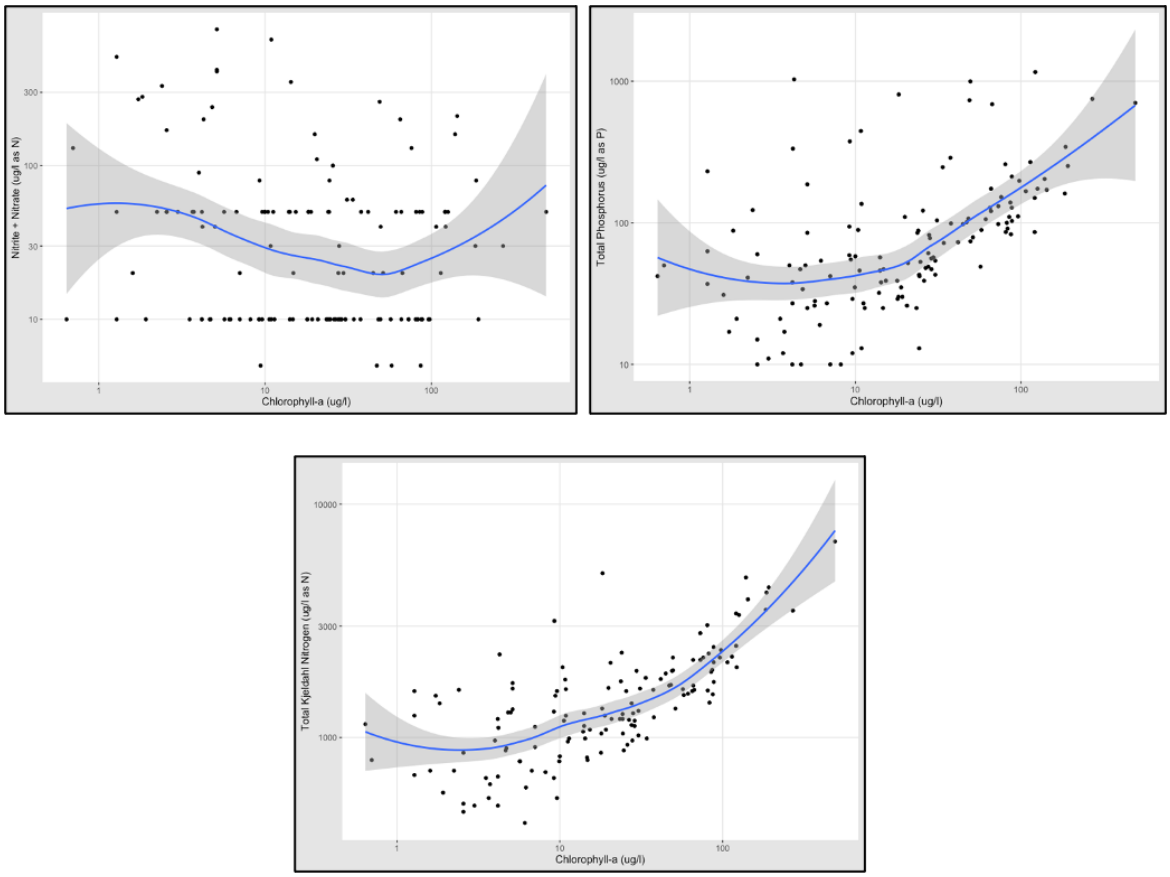


Figure 4. Covariance of Chlorophyll-a and Nitrite + Nitrate, Total Phosphorus, and Total Kjeldahl Nitrogen from the data subset

Table 1. Variables used in the AIC model from LAGOS-NE

Variable	Type	Description
Chlorophyll-a (µg/l)	Water Quality	Indicator of algae
Nitrite (NO2) + Nitrate (NO3) (µg/l)	Nutrient	Plant available form of Nitrogen
Total Kjeldahl Nitrogen (µg/l)	Nutrient	The total concentration of organic nitrogen and ammonia
Total Phosphorus (µg/l)	Nutrient	A measure of all forms of phosphorus present in a sample
2011 500-meter buffered NLCD (%)	Land Use and Land Cover	A 500-meter buffer of land use and land cover type surrounding all lakes >=4 ha from the year 2011
Mean Terrain Ruggedness Index (TRI) (m)	Land Use and Land Cover	The absolute difference in meters between the elevation of the focal cell and its immediate neighbors (10m scale)
Mean Slope (m)	Land Use and Land Cover	The slope at each cell is the slope with respect to its immediate neighbors (10m scale)
Mean Annual Runoff (in/yr)	Hydrologic	Mean annual runoff in the zone from 1951 to 1980
30-year mean annual precipitation (mm/yr)	Climate	30-yr long-term annual mean precipitation for zone
30-year mean annual temperature (deg C)	Climate	30-yr long-term annual mean temperature for zone
Lake area (ha)	Hydrologic	Total surface area of each lake
Lake Perimeter (km)	Hydrologic	Total perimeter of each lake

- *GIS data*
  - NRCS geospatial data gateway - 2016 Nationlal Land Cover Data Set (NLCD)
  - USGS National Map Veiwer - National Hydrography Dataset (NHD)

3. **Data Analysis**

- AIC stepwise modeling
  - The data was analyzed with a general linear model created using AIC stepwise model creation.
  - The linear model was then fed into a backward stepwise model, where the AIC algorithm starts with all possible variables and removes them one at a time in different combination until the maximum model performance is reached (Yamashita et al., 2007).

RESULTS

- The results of the stepwise model have indicated that the primary drivers of Chlorophyll-a are Nitrite (NO^-_2) + Nitrate (NO^-_3) and Total Kjeldahl Nitrogen (TKN) based on the given data observations

- TKN was shown to be a statistically significant (p-value < 0.05) variable while Nitrite + Nitrate was just shy of being considered a statistically significant variable (Table 2)

- The fitness of the overall Chlorophyll-a model has an R-squared of 0.63 (Figure 7)

- The model results indicate that we must accept the null hypothesis

Table 2. The variables selected by the stepwise model and their coefficients as well as their statistical significance

Drivers	Coefficients	Std. Error	t – value	Pr(>| t |)
Nitrite (NO2) + Nitrate (NO3)	-0.004294	0.002833	-1.515	0.131
Total Kjeldahl Nitrogen	0.047629	0.002484	19.172	< 2e-16
Intercept	-34.009674	4.403778	-7.723	4.32e-13

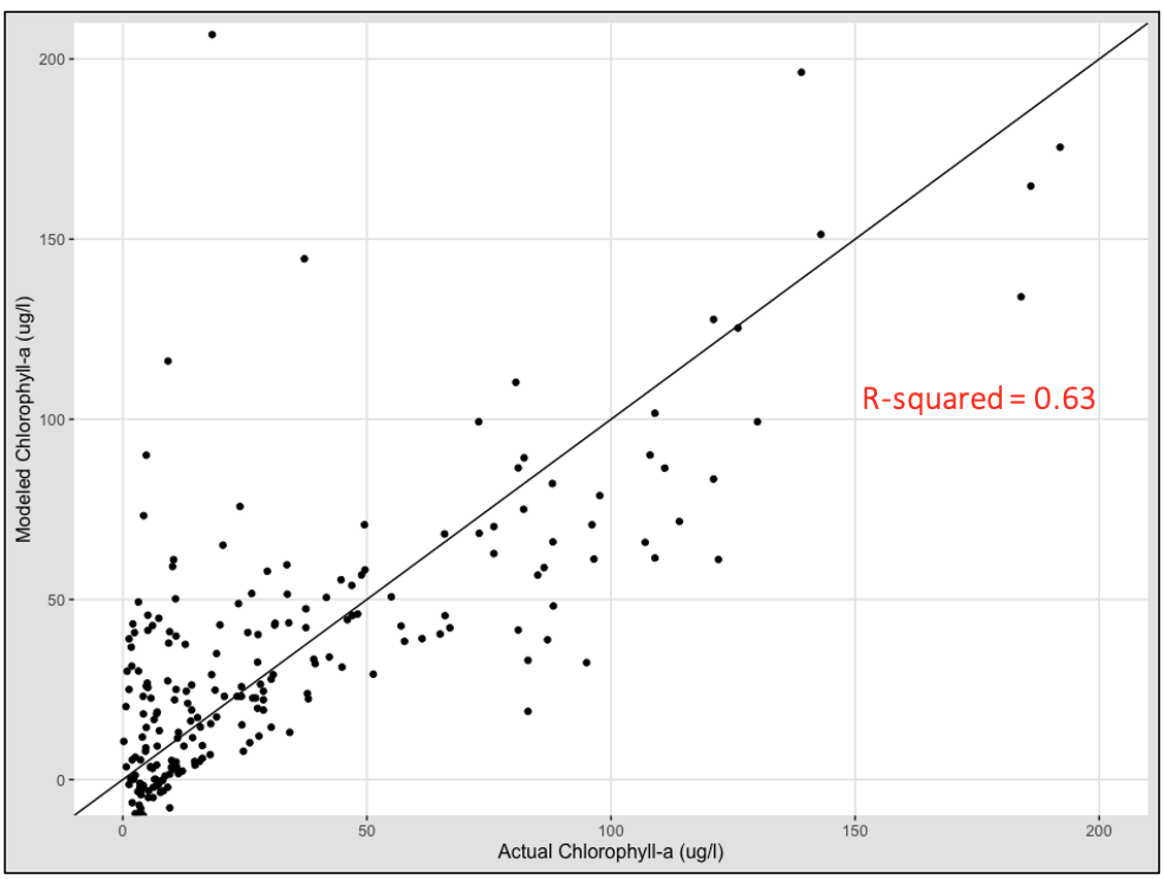


Figure 7. Modeled Chlorophyll-a using the coefficients from Table 2 vs. actual Chlorophyll-a values from the LAGOS-NE subset

DISCUSSION & CONCLUSION

Discussion

- Since the general linear model is not spatially or temporally explicit algorithm, it only cares about using values that will explain a variable such as Chlorophyll-a, and not the overall system.
- Although AIC model selection did not identifiy land cover as a primary driver, land cover is still most likely a mechanism that influences the primary drivers themselves.
- Nitrite + Nitrate is the form of Nitrogen that would be most available to algae, which is an interesting point because this would suggest there should be a higher statistical significance of Nitrite + Nitrate in the model results.
- Total Kjeldahl Nitrogen is a less bioavailable form of Nitrogen, however just because it is not bioavailable upon measurement does not mean that it never will be (Applequist, 2012).

Conclusion

Total Kjeldahl Nitrogen and Nitrite + Nitrate were identified as the primary drivers of Chlorophyll-a in this study. The overall model and variable selection lacked enough statistical significance to reject the null hypothesis, but this research project has given more than just results of a model.

- This project has shed light on the importance of quality and quantity of data.
- There are not very many data sets like LAGOS-NE that allow such streamlined access to so much data, but even then, sometimes there is not enough to truly model the system.
- A more proper use of LAGOS-NE for future projects would be to use it for a project with better scope.

ACKNOWLEDGMENTS

This project was completed as part of the Ecosystem Science and Sustainability SUPER Program (Skills for Undergraduate Participation in Ecological Research)

Thank you Dr. Stacy Lynn and Anna Clare Monlezun for project insperation and for organizing the SUPER program.

ABSTRACT

As more land is transformed from its natural state, the hydrology, ecological context, and water quality of landscapes across the world are also changed (Soranno et al., 2017). A major consequence that has arisen from this mass land transformation is the occurrence of algal blooms. Algal blooms can be identified by chlorophyll-a concentrations, but chlorophyll-a is only a product of other nutrients and hydrologic phenomena of the area (Chen et al., 2011). Nitrogen and phosphorus are limiting inputs of algal blooms, making them very important to algal bloom dynamics. (Carpenter et al., 2011). With major advances in the overall availability of data needed to assess these processes, it is now possible to conduct a study that covers a large hydrologic area that has hundreds of lakes with verified in-situ data. Here we show the primary drivers of chlorophyll-a in 31 lakes located within the Crow River watershed (HUC-8) located west of Minneapolis, Minnesota. Using the Lake Multi-Scaled Geospatial and Temporal Database (LAGOS-NE), and Akaike Information Criterion (AIC) stepwise modeling, a watershed specific equation can be generated using coefficients from the model. The AIC model identified Nitrite + Nitrate and Total Kjeldahl Nitrogen (TKN) to be the best predictor variables out of the 28 used. Our study will demonstrate the power of LAGOS-NE in dissecting important region-specific water quality issues.

REFERENCES

Anderson, D. M., Glibert, P. M., & Burkholder, J. M. (2002). Harmful algal blooms and eutrophication: Nutrient sources, composition, and consequences. Estuaries, 25(4), 704–726. https://doi.org/10.1007/BF02804901

Applequist, B. (2012). A brief introduction to Kjeldahl Nitrogen Determination. Labconco. https://www.labconco.com/articles/a-brief-introduction-to-kjeldahl-nitrogen-determ

Carpenter, S. R., Stanley, E. H., & Vander Zanden, M. J. (2011). State of the World's Freshwater Ecosystems: Physical, Chemical, and Biological Changes. Annual Review of Environment and Resources, 36(1), 75–99. https://doi.org/10.1146/annurev-environ-021810-094524

Chen, N., Wang, S., Zhang, X., & Yang, S. (2020). A risk assessment method for remote sensing of cyanobacterial blooms in inland waters. Science of The Total Environment, 740, 140012. https://doi.org/10.1016/j.scitotenv.2020.140012

Chen, X., Yang, X., Dong, X., & Liu, Q. (2011). Nutrient dynamics linked to hydrological condition and anthropogenic nutrient loading in Chaohu Lake (southeast China). Hydrobiologia, 661(1), 223–234. https://doi.org/10.1007/s10750-010-0526-y

North Fork Crow River. (2011, May 23). Minnesota Pollution Control Agency. https://www.pca.state.mn.us/water/watersheds/north-fork-crow-river

R: The R Project for Statistical Computing. (n.d.-a). Retrieved April 9, 2021, from https://www.r-project.org/

Soranno, P. A., Bacon, L. C., Beauchene, M., Bednar, K. E., Bissell, E. G., Boudreau, C. K., Boyer, M. G., Bremigan, M. T., Carpenter, S. R., Carr, J. W., Cheruvelil, K. S., Christel, S. T., Claucherty, M., Collins, S. M., Conroy, J. D., Downing, J. A., Dukett, J., Fergus, C. E., Filstrup, C. T., … Yuan, S. (2017). LAGOS-NE: A multi-scaled geospatial and temporal database of lake ecological context and water quality for thousands of US lakes. GigaScience, 6(gix101). https://doi.org/10.1093/gigascience/gix101

Soranno, P. A., Bissell, E. G., Cheruvelil, K. S., Christel, S. T., Collins, S. M., Fergus, C. E., Filstrup, C. T., Lapierre, J.-F., Lottig, N. R., Oliver, S. K., Scott, C. E., Smith, N. J., Stopyak, S., Yuan, S., Bremigan, M. T., Downing, J. A., Gries, C., Henry, E. N., Skaff, N. K., … Webster, K. E. (2015). Building a multi-scaled geospatial temporal ecology database from disparate data sources: Fostering open science and data reuse. GigaScience, 4(1), 28. https://doi.org/10.1186/s13742-015-0067-4

Soranno, P. A., & Cheruvelil, K. S. (2017a). LAGOS-NE-GEO v1.05: A module for LAGOS-NE, a multi-scaled geospatial and temporal database of lake ecological context and water quality for thousands of U.S. Lakes: 1925-2013 [Data set]. Environmental Data Initiative. https://doi.org/10.6073/PASTA/B88943D10C6C5C480D5230C8890B74A8

Soranno, P. A., & Cheruvelil, K. S. (2017b). LAGOS-NE-LOCUS v1.01: A module for LAGOS-NE, a multi-scaled geospatial and temporal database of lake ecological context and water quality for thousands of U.S. Lakes: 1925-2013 [Data set]. Environmental Data Initiative. https://doi.org/10.6073/PASTA/940B25D022C695B440E1BDBC49FBB77B

Soranno, P. A., Lottig, N. R., Delany, A. D., & Cheruvelil, K. S. (2019). LAGOS-NE-LIMNO v1.087.3: A module for LAGOS-NE, a multi-scaled geospatial and temporal database of lake ecological context and water quality for thousands of U.S. Lakes: 1925-2013 [Data set]. Environmental Data Initiative. https://doi.org/10.6073/PASTA/08C6F9311929F4874B01BCC64EB3B2D7

Stachelek, J., Oliver, S., & Masrour, F. (2020). LAGOSNE: Interface to the Lake Multi-Scaled Geospatial and Temporal Database (2.0.2) [Computer software]. https://CRAN.R-project.org/package=LAGOSNE

Yamashita, T., Yamashita, K., & Kamimura, R. (2007). A Stepwise AIC Method for Variable Selection in Linear Regression. Communications in Statistics - Theory and Methods, 36(13), 2395–2403. https://doi.org/10.1080/03610920701215639

Zhang, Y., Wang, Y., Wang, Y., & Xi, H. (2009). Investigating the Impacts of Landuse-landcover (LULC) Change in the Pearl River Delta Region on Water Quality in the Pearl River Estuary and Hong Kong's Coast. Remote Sensing, 1(4), 1055–1064. https://doi.org/10.3390/rs1041055