SPATIAL PATTERNING OF LAKE NUTRIENTS AND MORPHOMETRY AT
MACROSCALES: IMPORTANCE OF REGIONAL FACTORS AND
AQUATIC-TERRESTRIAL LINKAGES

By

Joseph Jeremy Stachelek

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Fisheries and Wildlife – Doctor of Philosophy

2020

ABSTRACT

SPATIAL PATTERNING OF LAKE NUTRIENTS AND MORPHOMETRY AT
MACROSCALES: IMPORTANCE OF REGIONAL FACTORS AND
AQUATIC-TERRESTRIAL LINKAGES

By

Joseph Jeremy Stachelek

Lakes are classically viewed as discrete ecosystems bounded on all sides by land. However, a narrow focus on lakes as discrete units is incompatible with the scale of many management programs and ignores the placement of lakes relative to their larger ecological context. While it is clear that lakes are not isolated units but are instead embedded components of lake-river networks and have a broader landscape (i.e. regional) context, it is not always clear how this embedding is borne out quantitatively. For example, is the position of a lake in a multi-lake network a dominant predictor of nutrient retention? Or, how strongly does the arrangement of streams and near-stream land-use (i.e. aquatic-terrestrial linkages) affect lake nutrient concentrations? To date, we have been unable to quantitatively address these questions in a synthetic manner because the necessary data has not previously been available for many lakes over large geographic extents (i.e. the macroscale). As a result, prior research has mostly been conducted on single lakes or in some cases groups of nearby lakes in a single region. In each of the following chapters, I developed macroscale lake databases to examine hundreds to thousands of lakes across diverse local and regional settings and used these databases to investigate the roles of both local and regional processes and aquatic-terrestrial linkages in determining lake nutrient retention, nutrient concentrations, and basic lake morphometry.

In my first chapter, I show that throughout Northeastern and Midwest US, lakes with higher connectivity have lower nutrient retention but this "connectivity effect" is apparent at the scale of entire lake networks rather than more localized lake subwatersheds. My findings suggest that a broader whole-network perspective is likely to be more effective than a narrow lake-specific perspective in regulatory frameworks focused on eutrophication. In my second

chapter, I show that lake nutrient concentrations are related to a variety of agricultural activity measures beyond the percent of the watershed in agricultural land use. I show that when one measure in particular, the percentage of agricultural land use in near-stream areas, is elevated, this signals a high likelihood of elevated lake nutrient concentrations. I further show that lake total phosphorus concentrations have different relationships with measures of agricultural activity compared to lake total nitrogen concentrations. My findings suggest that differences in lake nutrient sensitivity to agricultural activity may affect the outcome of policies to enhance water quality depending on whether they focus on lake N or P. In my third and final chapter, I test the utility of geometric models for predicting lake depth for lakes with missing depth information. Using bathymetric data for 5,000 lakes, I show that one of the assumptions of such models, that slope proxies of the surrounding land are representative of true in-lake slope, is not supported by available evidence. This lack of relationship has implications for predictive accuracy and model bias in lakes of different types or shapes. My findings specifically suggest that geometric models are likely to overestimate the depth of lakes with concave cross-section shapes and those classified as reservoirs. Across all chapters, I use the frameworks of landscape limnology and macrosystems ecology to explore the relationships between watershed and lake characteristics. I show that such macroscale analyses, which explicitly consider hierarchy in the freshwater landscape, provide a nuanced understanding of controls on lake nutrients and morphometry across a broad array of lakes.

ACKNOWLEDGMENTS

It's cliche to say you've stood on the shoulders of giants but maybe it's just this type of elevated perspective that's needed to see how the pieces fit together - to see how lakes are related to each other - and to see how lakes are related to their ecological context. The shoulders I've stood on include of course my advisor, Pat Soranno, as well as the other members of our lab Kendra Cheruvelil, Katelyn King, Ian McCullough, Autumn Poisson, Nick Skaff, and Nicole Smith whose support has been much appreciated. I would also like to thank my committee members Drs Elise Zipkin and Dana Infante who provided excellent advice and direction on this Dissertation journey. As a group, I would like to thank the members of the Continental Limnology and CNH Lakes Projects whose diversity of perspectives provided much needed context for my research. Also, thank you to everyone from the numerous federal, state, tribal, and non-profit agencies, as well as university researchers and citizen scientist data collectors and curators whose data unpin my research. Finally, I would like to thank Sheena Stachelek and the rest of my family for their love and encouragement.

"You can't stop water. Water will find a way."

- Colin Saunders

The chapters in this Dissertation have been written as standalone papers. Each paper was written with at least one collaborator and myself as lead author. Two chapters have already been published in academic journals with a third chapter anticipated for submission. The citation for each chapter is as follows:

- Stachelek, J. and Soranno, P. A., 2019. Does freshwater connectivity influence phosphorus retention in lakes? *Limnology and Oceanography.* doi:10.1002/lno.11137

- Stachelek, J., Weng, W., Carey, C. C., Kemanian, A. R., Cobourn, K. M., Wagner, T., Weathers, K. C., and Soranno, P. A., 2020. Granular measures of agricultural land-use influence lake nitrogen and phosphorus differently at macroscales. *Ecological Applications.* doi:10.1002/eap.2187

- Stachelek, J., Hanly, P. J., and Soranno, P. A., In Prep. Geometric models overestimate lake depth due to imperfect slope measurement. *to be submitted to Water Resources Research*

In addition to the papers that make up my Dissertation, I was either lead author or co-author on papers that are not explicitly listed above but were instead precursors to Dissertation chapters:

- Stachelek, J., Ford, C., Kincaid, D., King, K., Miller, H., and Nagelkirk, R., 2018. The National Eutrophication Survey: Lake characteristics and historical nutrient concentrations. *Earth System Science Data*, page 6. doi:10.5194/essd-10-81-2018

- Hollister, J. and Stachelek, J., 2017. Lakemorpho: Calculating lake morphometry metrics in R. *F1000Research*, 6:1718. doi:10.12688/f1000research.12512.1

Each Dissertation chapter (and many of my co-authored papers) were written following the principles of "open science" where data (and often code or software) are archived and made publicly available. The publishing of these materials is meant to increase research transparency, facilitate reproducibility, and enable future research developments. The software and data packages produced as part of this open science effort include:

- Stachelek, J. and Oliver, S, 2019. *LAGOSNE: Interface to the Lake Multi-Scaled Geospatial and Temporal Database.* https://github.com/cont-limno/LAGOSNE

- Stachelek, J., 2019a. *Gssurgo: Python Toolbox Enabling an Open Source gSSURGO Workflow.* https://github.com/jsta/gssurgo

- Stachelek, J., 2019c. Freshwater connectivity and stream morphology metrics for Northeast and Midwestern US lakes. doi:10.5281/zenodo.2554212

- Stachelek, J., 2018a. *nhdR: R Tools for Working with the National Hydrography Dataset.* https://github.com/jsta/nhdR

- Stachelek J, Ford C, Kincaid, D, King K, Miller H, and Nagelkirk R, 2017. The National Eutrophication Survey: Lake characteristics and historical nutrient concentrations. doi:10.5063/F10G3H3Z

I also authored papers that were not directly related to my Dissertation but were instead a result of my participation with two collaborative research groups, the Continental Limnology project (https://lagoslakes.org/projects/continental-limnology) and the CNH Lakes project (https://www.cnhlakes.frec.vt.edu/). My participation included attending monthly conference calls, annual project meetings, and collaborating on additional research projects outside of my Dissertation.

The Continental Limnology project consisted of approximately 25 people with expertise in ecology, freshwater science, statistics, and machine learning. The goal of this project was to understand and predict nutrient patterns for all continental US lakes to inform estimates of lake contributions to continental and global nutrient cycles. The papers I co-authored while on this project include:

- Wagner, T., Lottig, N. R., Bartley, M. L., Hanks, E. M., Schliep, E. M., Wikle, N. B., King, K. B. S., McCullough, I., Stachelek, J., Cheruvelil, K. S., Filstrup, C. T., Lapierre, J. F., Liu, B., Soranno, P. A., Tan, P.-N., Wang, Q., Webster, K., and Zhou, J., 2019. Increasing accuracy of lake nutrient predictions in thousands of lakes by leveraging water clarity data. *Limnology and Oceanography Letters.* doi:10.1002/lol2.10134

- Collins, S. M., Yuan, S., Tan, P. N., Oliver, S. K., Lapierre, J. F., Cheruvelil, K. S., Fergus, C. E., Skaff, N. K., Stachelek, J., Wagner, T., and Soranno, P. A., 2019. Winter

Precipitation and Summer Temperature Predict Lake Water Quality at Macroscales. *Water Resources Research*, 55(4):2708–2721. doi:10.1029/2018WR023088

- McCullough, I. M., Cheruvelil, K. S., Lapierre, J.-F., Lottig, N. R., Moritz, M. A., Stachelek, J., and Soranno, P. A., 2019a. Do lakes feel the burn? Ecological consequences of increasing exposure of lakes to fire in the continental United States. *Global Change Biology*, 25(9):2841–2854. doi:10.1111/gcb.14732

- McCullough, I. M., King, K. B. S., Stachelek, J., Diaz, J., Soranno, P. A., and Cheruvelil, K. S., 2019b. Applying the patch-matrix model to lakes: A connectivity-based conservation framework. *Landscape Ecology*. doi:10.1007/s10980-019-00915-7

- Qian, S. S., Stow, C. A., Nojavan A., F., Stachelek, J., Cha, Y., Alameddine, I., and Soranno, P., 2019. The implications of Simpson's paradox for cross-scale inference among lakes. *Water Research*, 163:114855. doi:10.1016/j.watres.2019.114855

- Soranno, P. A. and et, a., 2017. LAGOS-NE: A multi-scaled geospatial and temporal database of lake ecological context and water quality for thousands of US lakes. *GigaScience*, 6(12):1–22. doi:10.1093/gigascience/gix101

The CNH Lakes project consisted of approximately 16 people with expertise in freshwater ecology, hydrology, agriculture, economics, and social science. The goal of this project was to understand and predict feedbacks between human decision making, agriculture, hydrology, lake ecology, and lake advocacy associations. The papers I co-authored while on this project include:

- Ward, N. K., Fitchett, L., Hart, J. A., Shu, L., Stachelek, J., Weng, W., Zhang, Y., Dugan, H., Hetherington, A., Boyle, K., Carey, C. C., Cobourn, K. M., Hanson, P. C., Kemanian, A. R., Sorice, M. G., and Weathers, K. C., 2018. Integrating fast and slow processes is essential for simulating human–freshwater interactions. *Ambio*. doi:10.1007/s13280-018-1136-6

- Cobourn, K. M., Carey, C. C., Boyle, K. J., Duffy, C., Dugan, H. A., Farrell, K. J., Fitchett, L., Hanson, P. C., Hart, J. A., Henson, V. R., Hetherington, A. L., Kemanian, A. R., Rudstam, L. G., Shu, L., Soranno, P. A., Sorice, M. G., Stachelek, J., Ward, N. K., Weathers, K. C., Weng, W., and Zhang, Y., 2018. From concept to practice to policy: Modeling coupled natural and human systems in lake catchments. *Ecosphere*, 9 (5):e02209. doi:10.1002/ecs2.2209

- Hanson, P. C., Stillman, A. B., Jia, X., Karpatne, A., Dugan, H. A., Carey, C. C., Stachelek, J., Ward, N. K., Zhang, Y., Read, J. S., and Kumar, V., 2020. Predicting lake surface water phosphorus dynamics using process-guided machine learning. *Ecological Modelling*, 430:109136. doi:10.1016/j.ecolmodel.2020.109136

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Lakes are classically viewed as discrete ecosystems bounded on all sides by land (Golley, 1993). This framing has served as a useful organizing principle particularly with regard to the study of specific waterbodies (O'Neill, 2001). However, a narrow focus on lakes as discrete units is incompatible with the scale of many management programs and ignores the links between lakes, their aquatic-terrestrial linkages, and their larger ecological context (Likens and Bormann, 1974; Soranno et al., 2010). In recognizing these limitations, a specific focus of modern limnology has been on further integrating freshwater and terrestrial environments connected to and surrounding lakes (i.e. their ecological context). Frameworks developed as part of these efforts include landscape limnology (Soranno et al., 2010) and macrosystems ecology (Heffernan et al., 2014) that call for explicit consideration of hierarchy in the freshwater landscape and interactions across time and space among hierarchy levels.

Despite these efforts, several limiting factors have prevented a more full consideration of lakes relative to their ecological context. Foremost of these has been limited data availability. It is common for studies to consider the ecological context of lakes relative to their watersheds for a single lake or several lakes within a single geographic region. However, it is far less common for studies to consider many lakes across multiple regions (i.e. at macroscales) where distances span hundreds to thousands of kilometers (but see USEPA (1975), Landers et al. (1988), and USEPA (2016)). A consequence of the mostly narrow focus on relatively few lakes is that we lack the perspective necessary to formulate general relationships between lakes and their ecological context that hold beyond a single lake or region. As macroscale lake datasets become available, (such as Soranno and et (2017), Read et al. (2017), and Williams and Labou (2017)), we now have the ability to test our intuition and the generality of fine-scale relationships identified in prior studies against hundreds to thousands of lakes at the macroscale.

A further limiting factor preventing a fuller consideration of lake ecological context

has been lack of data analysis techniques that both account for and leverage spatial variability at broad scales. Such techniques are necessary to account for multiple interacting drivers of the lake ecological status (Allan, 2004). A key technique, which I use throughout the following chapters, is to build models with regionally varying coefficients that capture different relationships between lake characteristics and landscape predictors depending on geographic region. This technique avoids some of the risk of drawing imprecise or incorrect conclusions due to lumping together lakes with fundamentally different responses to a given predictor variable (Qian et al. 2019). Furthermore, this technique allows for a richer post-hoc examination of model results in light of unobserved or unaccounted regional variation in underlying lake drivers.

Given the limiting factors identified above, my aim in this Dissertation was to use the landscape limnology and macrosystems ecology frameworks to more rigorously test conventional understanding of lakes gained from studies at limited spatial extents to a more diverse set of lakes representative of continental-scale gradients through scaling and extrapolation. I developed each chapter following these general steps 1) Identify an aspect of lake ecology that has not been explored at broad spatial extents, 2) Develop a conceptual model linking lake characteristics to different components of the freshwater landscape taking into account variation across spatial scales, and 3) Formulate statistical or machine learning models to test this conceptual model. Underpinning each conceptual model were questions such as: What is the role of local and regional processes and aquatic-terrestrial linkages in determining lake nutrient retention, nutrient concentrations, or basic lake morphometry? and How do relationships between lake nutrient retention, nutrient concentrations and morphology differ when computed over different extents and with different levels of detail?

CHAPTER 1

DOES FRESHWATER CONNECTIVITY INFLUENCE PHOSPHORUS RETENTION IN LAKES?

1.1 Abstract

Lake water residence time and depth are known to be strong predictors of phosphorus (P) retention. However, there is substantial variation in P retention among lakes with the same depth and residence time. One potential explanatory factor for this variation is differences in freshwater connectivity of lakes (i.e. the type and amount of freshwater connections to a lake), which can influence watershed P trapping or the particulate load fraction of P delivered to lakes via stream connections. To examine the relationship between P retention and connectivity, we quantified several different measures of connectivity including those that reflect downstream transport of material (sediment, water, and nutrients) within lake-stream networks (lake-stream-based metrics) as well as those that reflect transport of material from hillslope and riparian areas adjacent to watershed stream networks (stream-based metrics). Because it is not always clear what spatial extent is appropriate for determining functional differences in connectivity among lakes, we compared connectivity metrics at two important spatial extents: the lake subwatershed extent and the lake watershed extent. We found that variation in P retention among lakes was more strongly associated with connectivity metrics measured at the broader lake watershed extent rather than metrics measured at the finer lake subwatershed extent. Our results suggest that both connectivity between lakes and streams as well as connectivity of lakes and their terrestrial watersheds influence P retention.

1.2   Introduction

Lake phosphorus (P) retention is an important characteristic of lakes that can be used to predict P concentrations and to evaluate lake sensitivity to nutrient loading and eutrophication (Alexander et al., 2008; Milstead et al., 2013). Specifically, P retention is an integrated measure of internal P losses including permanent sedimentation, biological uptake, and other processes that remove P from the water column (Chapra, 2008). P retention has been well studied in lakes because P determines lake trophic status and downstream watershed yields (i.e. export to terminal lakes and coastal estuaries). Although previous studies have shown that lake P retention is related to water residence time (Vollenweider, 1975), large uncertainties exist around this relationship (Brett and Benjamin, 2008; Milstead et al., 2013). These uncertainties can be large in lakes with intermediate water residence times, particularly compared to lakes with either extremely short or extremely long water residence times (Figure S1.1). For example, lakes with very long water residence times (on the order of a decade or longer) have complete or near complete P retention, while lakes with very short water residence times (on the order of days) have almost no P retention (Brett and Benjamin, 2008). The reason for the substantial uncertainty in lake P retention between these two extremes may be that predicting retention solely on the basis of water residence time does not capture many of the other factors and processes that affect P retention (Figure 1.1).

One well-studied factor that has been shown to influence lake P retention is lake depth through its influence on internal processing of P loads (Søndergaard et al., 2013; Vollenweider, 1975). The mechanism for such an influence is that lake depth controls thermal stratification and material resuspension from the benthos. As a result, shallow lakes have a tendency to mix throughout the summer causing redistribution of sedimented phosphorus throughout the mixed zone (Fee et al., 1996). This mixing and redistribution of sedimented P often leads to tighter benthic-pelagic coupling and increased P recycling (Cha et al., 2013). Thus, depth is one example of a lake characteristic that is likely to influence P retention in concert with water residence time (Cheng et al., 2010).

4
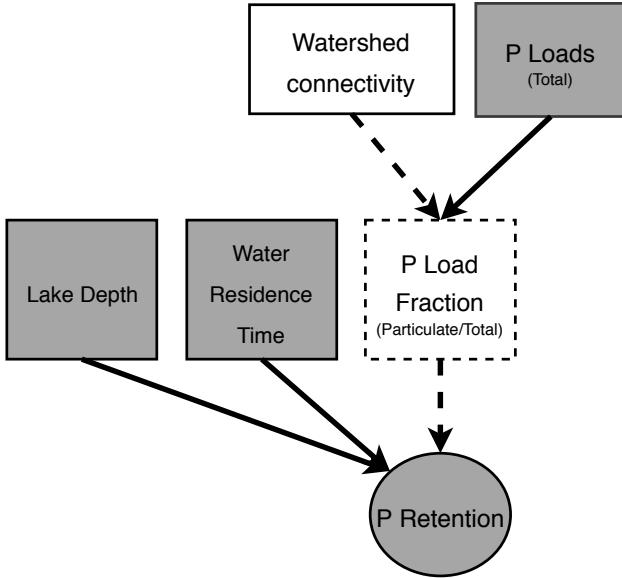
Figure 1.1: Major lake and watershed factors affecting lake P retention. Shaded symbols indicate factors typically considered in P retention models whereas open symbols indicate additional factors specific to the present study. Dashed lines indicate inferred relationships, which cannot be tested with available data, but are otherwise discussed herein.

A primary tool for evaluating the influence of specific lake characteristics, like lake depth, on P retention is statistical phosphorus retention modelling. Although the form of such models is variable, many models estimate P retention as a function of water residence time and a parameter k that represents in-lake P decay (Chapra, 2008; Vollenweider, 1975). Many studies treat k as a constant global parameter (i.e., it has the same value for all lakes), which may be valid only for studies that consider small numbers of lakes in a limited geographic region, with similar characteristics. Few studies have modeled different k values based on lake or watershed characteristics, despite the many differences among lakes that likely influence their ability to process P (Cheng et al., 2010).

For example, there is evidence that the relative proportion of particulate versus dissolved P loads (hereafter, particulate load fraction), influences P retention in stream and wetland ecosystems (Jarvie et al., 2011; Kronvang, 1992; Russell et al., 1998; Vanni et al., 2001). However, evidence for particulate load fraction controls on P retention in lakes remains limited because of difficulties in tracking the fate of particulate loads after entering a lake

(Brett and Benjamin, 2008; Dillon and Molot, 1996). Therefore, there is potential to further study particulate load Ffraction using proxies that may be closely related to it, such as the relative amounts of point and nonpoint nutrient sources to a lake subwatershed.

In general, point source inputs to lakes are associated with increased dissolved P loads (Kronvang, 1992; Russell et al., 1998), whereas non-point source inputs to lakes, such as those in lake watersheds with high agricultural land use cover, typically have higher particulate P loads (Carpenter et al., 1998; Sharpley et al., 1994). Exceptions to these generalizations have been observed in areas where increased dissolved P loading occurs not as a result of point source nutrient inputs but rather as a result of non-point source runoff due to P saturated soils or legacy P release (Bennett et al., 2001; Powers et al., 2015). Despite the fact that particulate loads are usually related to non-point source inputs, the fraction of non-point source inputs that are in particulate form is highly variable and may depend on intra-annual flow variations as well as watershed erosion characteristics (Jarvie et al., 2011). Furthermore, quantifying non-point source inputs at broad spatial scales is not commonly done due to logistical and sampling constraints (Guy et al., 1994). As a result, studies often infer the relative amounts of particulate and dissolved loading from other available proxy data such as land-use cover (Ellison and Brett, 2006; Djodjic and Markensten, 2018).

Apart from land-use cover, another potential proxy for particulate load fraction entering lakes is the type and amount of freshwater connections to a lake, which we argue is easier to measure than other proxies, and could help to improve estimates of lake P retention, especially in lakes for which we lack P loading data (Figure 1.1). Although freshwater connectivity may be easy to measure from a logistical standpoint, there are still many ways to measure connectivity that likely represent different mechanisms of water and material flow (Figure 1.2). We broadly define and study two types of freshwater connectivity that correspond to either stream-based metrics or lake-stream-based metrics. First, lake-stream-based metrics measure the connections between a lake, other upstream lakes, and streams in their watershed. This type of connectivity can be quantified by measuring the closest

6

Low   High

Lake-stream-based Metrics

A. Closest Lake Distance: Network distance to the closest upstream lake.

B. Upstream Lake Area: Total area of upstream lakes limited to lakes that are greater than or equal to 4ha.

Stream-based Metrics

C . Average Link Length: Sum of the total length of stream reaches between junctions divided by the total number of reaches.

D. Stream Density: the length of all streams (minus artificial lines through lakes) divided by the area of watershed.

E. Stream Order Ratio: The total number of first order stream reaches divided by the total number of higher order stream reaches. In this example the lower value lake has a stream order ratio of 2 whereas the higher value lake has a stream order ratio of 3.

Figure 1.2: Connectivity metric definitions along with simplified examples of high and low value lakes that might arise from a binary classification. Both lake-stream-based and stream-based metrics are associated with restrictions on in-stream transport whereas stream-based metrics are associated with differences in transport of P from terrestrial runoff to streams. We use the term "first order stream" to describe a headwater stream.

distance to an upstream stream-connected lake (Figure 1.2A), or by measuring the total upstream lake area (Figure 1.2B). Second, stream-based metrics measure the connections between inflowing streams and their surrounding land whereby increasing stream connections lead to a greater abundance of land-water interfaces and greater transport of material from hillslope and riparian areas adjacent to watershed stream networks (Figure 1.2C-E).

7
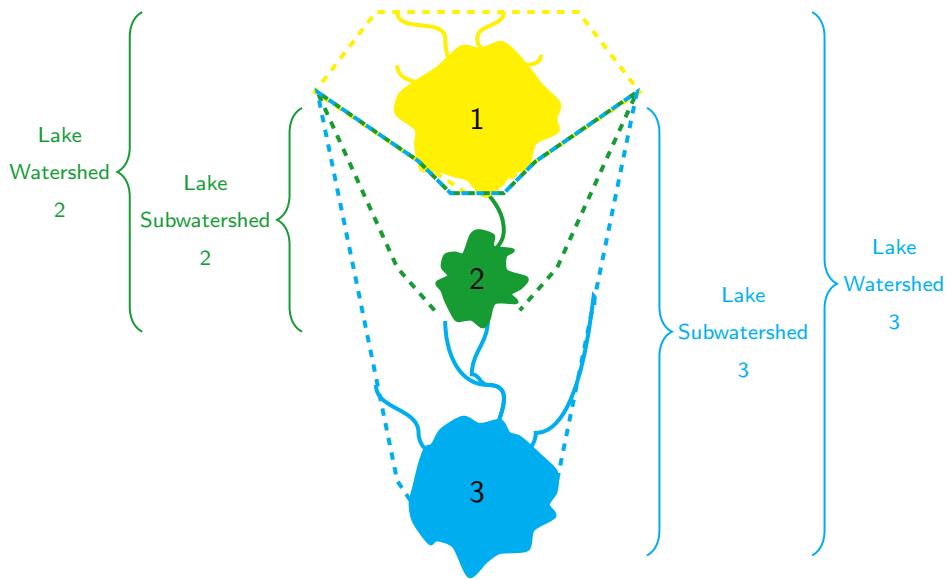
Figure 1.3: Diagram showing the lake subwatershed and lake watershed of three lakes. Here the lake subwatershed of lake 3 encompasses the lake subwatershed of lake 2 because it is smaller than 10 ha but it does not encompass the lake subwatershed of lake 1 because it has an area of at least 10 ha. In contrast to the lake subwatershed boundaries, the lake watershed boundaries extend to the headwaters of the lake chain.

In addition to variation among connectivity metrics and connectivity metric types, it is also not always clear which spatial extent is appropriate for determining functional differences in connectivity among lakes (Soranno et al., 2015). Such information is needed to inform the design of regulatory frameworks balancing controls on cumulative nutrient transport along stream networks and controls on localized nutrient transport (Alexander et al., 2008; Carpenter et al., 1998; Withers and Jarvie, 2008). To test the importance of different spatial extents, we examined connectivity metrics measured for both the lake subwatershed extent and the lake watershed extent (Figure 1.3). Here, the lake subwatershed extent includes the elements of the immediate watershed in the direct drainage of a lake whereas the lake watershed extent includes all of the elements in the entirety of the lake-stream network up to and including headwater streams (Figure 1.3).

We propose that measures of freshwater connectivity are related to P retention in the following ways. First, some connectivity metrics reflect the proximity of terrestrial

watershed areas to the stream network (Covino, 2017). For example, lakes with a high watershed stream density should have increased particulate matter loading from terrestrial hillslope and riparian areas adjacent to watershed stream networks because particulate matter export is limited by overland distance (Gomi et al., 2002). Second, connectivity metrics can reflect the configuration of lakes within lake-stream networks. For example, lakes with upstream lakes in close proximity may receive P loads that have previously undergone in-lake processing whereby labile fractions have already been trapped in upstream lakes (Cardille et al., 2007). In contrast, lakes with more distant upstream lakes are more likely to receive the more labile fractions from terrestrial runoff that serve to increase P retention as opposed to receiving the more recalcitrant fractions that are resistant to biological uptake and are thus not retained. Although some connectivity metrics have an intuitive relation to P retention, it is not clear which specific measures of freshwater connectivity are important for transport of particulate matter. Therefore, our study is designed to examine and compare which measures of connectivity are more related to lake P retention.

We addressed the above knowledge gaps by quantifying and comparing a range of freshwater connectivity measures at multiple spatial extents. Taken together, our suite of connectivity metrics reflect both freshwater connectivity in the direct drainage of lakes (i.e. the lake subwatershed extent) and freshwater connectivity of the entirety of the stream network up to and including watershed headwater areas (i.e. the lake watershed extent, Figure 1.3). Our motivation for measuring connectivity so many ways is that it is easy to measure connectivity of lakes with small watersheds situated at the beginning of lake chains but it is much more challenging to identify the type and extent of connectivity in larger, more complex lake networks. Another reason we examined relationships between P retention, multiple measures of connectivity, and multiple spatial extents is that many commonly used connectivity metrics merely reflect watershed size (spatial extent) rather than types of material transport or particulate load fraction (Leibowitz et al., 2018). Although lakes in larger watersheds have both a greater potential area from which to source particulate runoff

and total phosphorus export from the watershed, we expect that delivery of sediment-bound phosphorus is dependent on connectivity-mediated trapping in the upstream watershed (T-Prairie and Kalff, 1986).

We asked two questions in this study: 1) Which measures of freshwater connectivity influence lake phosphorus retention? 2) What spatial extent of connectivity most strongly influences P retention? To answer these questions, we fit statistical P retention models in a Bayesian hierarchical framework following Cheng et al. (2010) where 2 separate values of the k processing parameter were estimated for lakes with either high or low values of each connectivity metric. Using this approach, higher k values for a specific lake connectivity class indicates more extensive in-lake processing and higher P retention. We applied this model to a dataset of 129 lakes across a wide range of hydrologic, geologic, and climatic settings. We fit separate models using each combination of connectivity metric and spatial extent in an effort to determine whether P retention is more strongly controlled at the lake subwatershed extent or the lake watershed extent.

1.3 Methods

1.3.1 Dataset description

We used data on P retention, maximum depth, and water residence time from 129 lakes in the National Eutrophication Survey (USEPA, 1975; Stachelek et al., 2018). Mean annual P loading, P discharge, and P retention values in the National Eutrophication Survey (NES) dataset were calculated based on monthly sampling for P in tributary and outlet discharge points as well as any municipal waste discharges from 1972 to 1975. Here, P retention is a unitless value representing the fraction of incoming P loads. Sampling frequency for water discharge and residence time varied among lakes but details of these variations were not provided in the source dataset (USEPA, 1975; Stachelek et al., 2018). Estimates of water discharge and residence time in the NES dataset represent normalized mean flow estimates expected to occur during a period of average precipitation and hydrology (USEPA, 1975).

	Min	Median	Max	IQR
Total Phosphorus (ug/L)	4	43	1380	20 - 111
Chlorophyll (ug/L)	1	12	381	6 - 21
Secchi Depth (m)	0.2	1.5	19.3	0.9 - 2.4
P Loading (kg/yr)	204	6041	418485	2035 - 24370
P Retention	0.06	0.46	0.99	0.24 - 0.59
Residence Time (yr)	0.03	0.63	17.4	0.2 - 1.8
Lake Area (km^2)	0.25	6.54	453.26	2.93 - 21.88
Maximum Depth (m)	1.1	12.9	96.3	9.2 - 21.3
Agricultural Landuse (%)	0.11	53.27	94.74	17.68 - 74.05
Lake Subwatershed Area (km^2)	5	88	4018	21 - 410
Lake Watershed Area (km^2)	120	143	18641	54 - 550

Table 1.1: Minimum, median, maximum, and interquartile range of selected characteristics of the study lakes (N = 129).

Water residence time for our study lakes ranged from 1 week to 17 years with an interquartile range of 3 months to 1.8 years while P retention ranged from 0.06 to 0.99 with an interquartile range of 0.24 to 0.59 (Table 1.1).

We supplemented the NES dataset with boundaries for lake subwatersheds, as well as estimates of stream density, upstream lake area, upstream lake connection(s), baseflow (an index of groundwater inputs), land-use cover, and other water quality measurements from the LAGOS-NE dataset (Table 1.1; Soranno and et, 2017). Our study lakes encompassed a wide range of land-use cover types and nutrient levels (Table 1.1). Although, lake subwatersheds were variable with respect to agricultural land use cover, we did not observe a strong relationship with lake P retention (Figure S1.2). On average, the water quality (total phosphorus, chlorophyll concentration, and Secchi depth) of the lakes in our study are similar to other US lakes as measured by the stratified random sampling design of the National Lakes Assessment (NLA) lake population (USEPA, 2016). However, our lakes are substantially larger and deeper than most NLA lakes (Figure S1.3).

We restricted the lakes in the study to those located within the footprint of LAGOS-NE which includes lakes located in 17 northeastern and midwestern US states (Soranno and et, 2017). We excluded lakes from our analysis if they had a surface area of greater than 1000

km^2 or a surface area of less than 0.1 km^2. We also excluded lakes if they had a maximum depth of greater than 70 m, lacked upstream surface water connections, or had one of the North American Great Lakes in its upstream watershed. A total of 129 out of 236 NES lakes met each of these selection criteria.

1.3.2   Connectivity metrics and spatial extents

In addition to data from the NES and LAGOS-NE datasets, we calculated several connectivity metrics that we expected would be related to lake P retention (Figure 1.2). Some of these metrics were stream-based with the goal of capturing aspects of the configuration of each lakes' upstream surface water network (Figure 1.2C-D). In particular, we chose metrics that would quantitatively approximate network complexity under the assumption that highly complex networks are also low connectivity networks. This assumption is supported by the findings of stream network simulations where increased network complexity leads to increased network resistance and ultimately decreases in network connectivity (Rodriguez-Iturbe and Rinaldo, 2001). In addition to stream-based metrics, we calculated lake-stream-based metrics that we expected would reflect the likelihood of P trapping in upstream lakes prior to arriving at a given focal lake via tributary flow (Figure 1.2A, 2B). A simple metric that captures this likelihood is the presence (or absence) of an upstream lake (greater than 4 ha) which we define as "lake connection" (Fergus et al., 2017). In addition to lake connection, we calculated related metrics such as total upstream lake area and the network distance to the closest upstream lake.

   To examine the importance of spatial extent relative to our connectivity metrics, we calculated each metric at multiple extents (Figure 1.3). First, we calculated connectivity metrics at the scale of individual lake subwatersheds. We defined a lake subwatershed as the area draining into a particular lake exclusive of any upstream areas that drain into a lake greater than or equal to 10 ha (0.1 km^2). Next, we calculated connectivity metrics at the scale of entire upstream lake networks (lake watershed extent). We defined a lake watershed

as the area draining into any part of the upstream network irrespective of the presence or absence of upstream lakes (Figure 1.3).

All connectivity metrics were calculated using the high-resolution National Hydrography Dataset (NHD) as a primary input (USGS, 2018). Average link length was calculated as the total stream length in a given watershed divided by the number of stream reaches after dissolving (removing) any network points that do not occur at a stream junction. Stream density was calculated as the length of all streams in the watershed (minus artificial lines through lakes) expressed in units of meters per hectare. Upstream lake area was calculated as the sum of the lake area in the upstream watershed expressed in square meters. Stream order ratio was defined as the number of headwater (first-order) streams in the upstream watershed of the focal lake divided by the total number of higher order (> 1) streams (La Barbera and Rosso, 1989). Closest distance to an upstream lake was defined as the shortest path-distance (rather than the straight-line distance) to a lake upstream from the focal lake.

We calculated all connectivity metrics and lake watershed extents using the `streamnet` and `nhdR` R packages respectively (Stachelek, 2018b,a). The algorithms in the `streamnet` package use the `sf` R package (Pebesma, 2018) as well as the `v.net` and `v.stream.order` modules (Jasiewicz and Metz, 2011) included in GRASS GIS (GRASS Development Team, 2017). All processed connectivity data and code are available at doi:10.5281/zenodo.2554212.

1.3.3   Modelling lake P retention

We modelled lake total phosphorus retention (hereafter, P retention) using the Vollenweider equation that models P retention as a function of water residence time and a parameter (k) conceptually representing in-lake P decay (Chapra, 2008; Vollenweider, 1975). Although there are several variants of this basic equation, we selected a 2-parameter form (equation 1.1) that has been shown to have good performance in multiple cross-sectional studies (Brett and Benjamin, 2008; Cheng et al., 2010):

$$R_i = 1 - \frac{1}{1 + k\tau_i^x} \tag{1.1}$$

where R_i is P retention as a fraction of P inputs, τ is water residence time, k is a unitless parameter representing in-lake P decay, and x is a unitless parameter representing P export via hydrologic flushing. Here, higher values of k are associated with greater integrated P losses from sedimentation and biological uptake resulting in greater P retention. Note that Eq 1.1 does not include a recycling term. Therefore, our results represent net P retention (as opposed to gross P retention) under a steady state assumption where lakes are at equilibrium with respect to recycling (Vollenweider, 1975). Note that although some forms of the Vollenweider equation use P loading as a predictor variable, it does not appear in Eq 1.1. The reason for this is two-fold. First, estimates of P loading are more difficult to obtain than estimates of water residence time and our aim was to develop a model than can be widely applied to lakes for which we lack detailed data on P loading. Second, loading based model forms have been shown to be mathematically equivalent to water residence time based model forms (Brett and Benjamin, 2008).

We used the model described by Eq 1.1 to compare lake P retention in lakes with different connectivity by fitting the model in two ways. First, we modelled the overall relationship between P retention and water residence time for all lakes in our dataset (global model). Second, we fit hierarchical versions of Eq. 1.1 where k was modelled separately (k_j) as a function of a binary sub-population indicator g_i denoting membership in one of two lake classes formed on the basis of specific connectivity metrics (or lake depth) and specific spatial extents:

$$R_i = 1 - \frac{1}{1 + k_j\tau_i^x} \tag{1.2}$$

$$k_j = g_i \tag{1.3}$$

where greater differences in k between the two groups indicate greater support for a connectivity effect on P retention. Prior to model fitting, we examined the bivariate relationships between each connectivity metric, water residence time, P loading, and other factors related to P retention using Pearson's correlation coefficients. The purpose of this exercise was twofold, to determine the potential for collinearity among any of the variables in Eq 1.2 and to identify any relationships between P retention, water residence time, and other watershed and lake characteristics that were not included in our model. As only one connectivity metric was used to define g for each model we did not use the results of this exercise to exclude variables from further investigation. We quantified the relative support for an effect of each connectivity metric on P retention in more detail by calculating the difference in the median value of the P decay parameter k between groups (i.e. Δk). We used these median k values along with median estimates of x to determine how differences in k translated to differences in P retention (Eq 1.2). We judged significance by whether or not differences in group-wise P retention were greater than the measurement precision of P retention (> 0.01).

We fit all models in a Bayesian framework using the non-linear extension to the `brms` package to interface with the Stan statistical program (Bürkner et al., 2017; Stan Development Team, 2017). In both models, we set a semi-informative prior on k and x of N(1.3, 0.1) and N(0.45, 0.1) respectively. These priors were based on the confidence intervals presented in Brett and Benjamin (2008) and qualitatively matched those used by Cheng et al. (2010). We used the default settings of `brms` and `rstan` to generate posterior estimates using four chains of 4,000 iterations each with no thinning and initial parameter values drawn from a uniform distribution bounded between -2 and 2. We also used the `brms` package for model evaluation by computing a Bayesian R^2 following the method of Gelman et al. (2017).

1.3.4 Lake connectivity classes

Our lake connectivity classes were formed by dividing the lake dataset into two classes for each connectivity metric based on the bivariate relationship between each metric and P retention. Prior studies have used a similar binary splitting approach to examine the effect of various exogenous factors on lake P retention (Cheng et al., 2010; Shimoda and Arhonditsis, 2015). We determined splitting criteria for each metric from the results of a random forest procedure incorporating conditional inference trees (Hothorn and Zeileis, 2015). This procedure creates binary splits of the independent variables (i.e. each of the connectivity metrics), which are recursively repeated to find the split that maximizes association with the dependent variable (P retention). The advantage of the ctree technique over the more typical classification-regression tree (CART) technique is that tree growth stopping rules are pre-specified (Hothorn et al., 2006). As a result, some of subjectivity associated with post-hoc tree pruning is avoided.

1.4 Results

1.4.1 Interactions between connectivity, hydrology, and P loading

We examined the bivariate relationships between connectivity, water residence time, P loading and other factors related to P retention to determine the potential for strong relationships among any of the variables which were not accounted for by our model structure (Eq 1.2, Figure 1.4). We found evidence for some relationships among these variables, but none that suggest either redundancy among connectivity metrics or that otherwise limit our ability to infer relationships with P retention (Figure 1.4). For example, the Pearson correlation (r) between water residence time and stream density, which is implicitly accounted for in our model structure, was 0.29 ($p < 0.05$). In contrast, the correlation between water residence time and upstream lake area, which is unaccounted for in our model structure, was only 0.10 ($p > 0.05$). Note that interactions between connectivity metrics and water residence time are

Figure 1.4 (correlation matrix):

	Average link length	Closest lake distance	Stream density	Max depth	Baseflow	Stream order ratio	Lake Area	Point Source P	SWS area	P Retention	Upstream lake area	P Loading	WS area
Closest lake distance													
Stream density	0.26	0.27											
Max depth													
Baseflow	0.20	0.20	0.52	0.23									
Stream order ratio	0.26		0.26		0.19								
Lake Area				0.46									
Point Source P							0.13						
SWS area		0.33	0.30		0.25	0.19	0.33	0.15					
P Retention					0.27	0.25		0.28					
Upstream lake area						0.21	0.25	0.43		0.32			
P Loading							0.24		0.76		0.52		
WS area						0.36	0.18	0.36		0.77	0.62		
Water Residence Time	0.25	0.26	0.29	0.47	0.22	0.26	0.30			0.44			

Figure 1.4: Correlation among connectivity metrics and selected lake characteristics. Cell shading and color reflects the correlation coefficient value. Only coefficients accompanied by a significant p value < 0.05 are shown as text. Here, WS is the lake watershed extent whereas SWS is the lake subwatershed extent. Connectivity metrics are defined in Figure 1.2.

accounted for in our hierarchical model structure because Eq. 1.2 uses a global coefficient x for water residence time that is estimated separately from the hierarchical and connectivity-dependent P decay coefficient k. This is conceptually similar to fitting the k-connectivity-P retention relationship to the residuals of the water residence time to P retention relationship. Perhaps surprisingly, we did not observe strong correlations between P loading and water residence time (r = 0.13, p > 0.05) or between lake depth and most connectivity metrics (r < 0.17). This suggests that our P retention results are not confounded by variations in lake depth, by interactions between connectivity and P loading, or by interactions between water residence time and P loading.

A secondary purpose of examining the bivariate relationships between connectivity, water residence time, P loading and other factors related to P retention was to determine if

our connectivity metrics were related to each other such that they provide similar information. With the exception of stream density and baseflow (r = -0.52, p < 0.05), correlations among connectivity metrics were low and of a similar magnitude as the correlations between connectivity metrics and water residence time (0.10 < r < 0.29). The strongest correlation between any connectivity metric or lake characteristic was upstream lake area and lake watershed area (r = 0.77, p < 0.05).

We observed notably strong correlations between P loading and several lake characteristics not accounted for in Eq 1.2. These include the correlation between P loading and lake watershed area (r = 0.62, p < 0.05) as well as the correlation between P loading and lake subwatershed area (r = 0.76, p < 0.05). In addition, the correlation between P loading and upstream lake area was quite strong (r = 0.52, p < 0.05). Despite strong correlations between P loading, watershed area, and upstream lake area, we did not observe a strong correlation between P loading and P retention (r = 0.15, p > 0.05). Notably, this correlation was much weaker than the correlation between water residence time and P retention (r = 0.44, p > 0.05). Our observation that the correlation between P retention and P loading was not appreciably stronger than correlations between P retention and our connectivity metrics suggests that our estimates of connectivity metric effects on P retention are not confounded by interactions between water residence time and P loading.

Although we did not observe strong correlations among connectivity metrics, we found that lakes with similar connectivity metric values, were spatially clustered (Figure 1.5). In particular, we found that lakes were concentrated in either the southern or northern portions of our study area depending on their connectivity metric value (Figure 1.5). This observation is consistent with the findings of Fergus et al. (2017) that lakes in the northern portion of our study area have distinct freshwater connectivity as compared to lakes in the southern portion of our study area.

A. SWS

Average link length

Closest lake distance

Stream density

Baseflow

Stream order ratio

B. WS

Average link length

Closest lake distance

Stream density

Upstream lake area

Baseflow

Stream order ratio
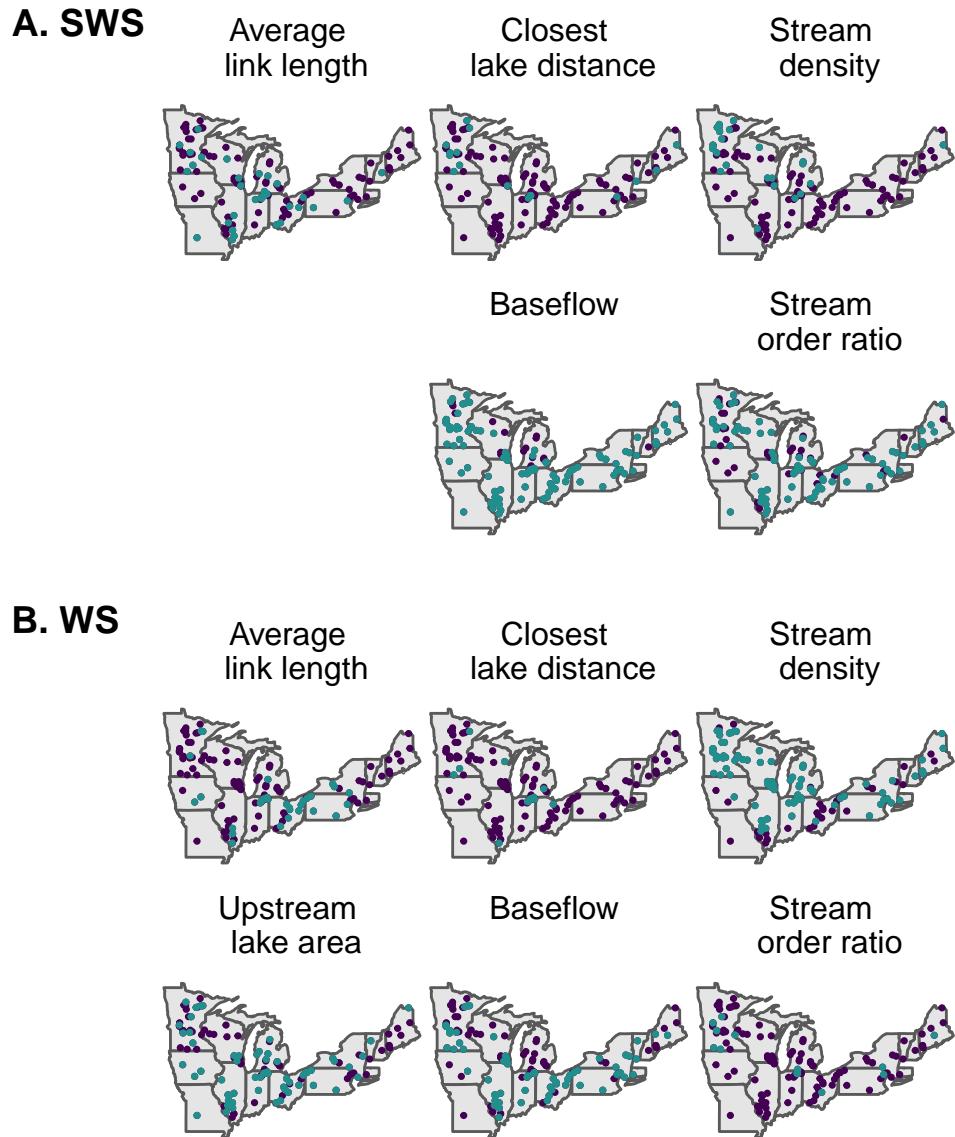
Figure 1.5: Locations of lakes of with different connectivity metric values at the (A) lake subwatershed and (B) lake watershed extent. Low valued lakes are represented by lighter (green) symbols, and high values lakes are represented by darker (purple) symbols. WS is the lake watershed extent. SWS is the lake subwatershed extent. Connectivity metrics are defined in Figure 1.2.
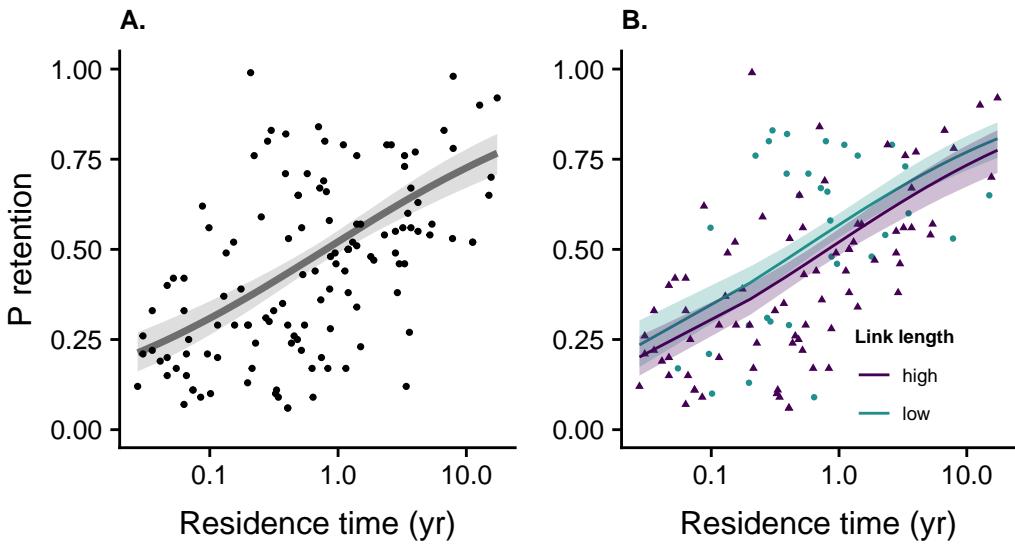
Figure 1.6: Residence time (yr) versus P retention for the NES dataset and the global model fit to the data ($R^2 = 0.34$, n = 129) where the solid line and shaded interval represents the median and central 95% interval estimates respectively (A). As above except that solid lines and shaded interval estimates represent hierarchical model fits to the data ($R^2 = 0.41$, n = 129) based on watershed average link length, which had the strongest association of any connectivity metric relative to P retention (B). Equations for these lines at median water residence time for low and high link length lakes are: Rp = 1 - (1 / (1 + 1.32 $\tau^{0.4}$)) and Rp = 1 - (1 / (1 + 1.08 $\tau^{0.4}$)) respectively.

1.4.2 Effects of connectivity on P retention

We found that freshwater connectivity metrics were associated with lake P retention (Figure 1.6, 1.7). Across all connectivity metrics except stream order ratio, we found that the P decay coefficient k and thus P retention was associated with whether a lake had a high or low value of each connectivity metric. These findings matched our expectations in several ways. Most notably, lakes with shorter average link lengths had higher P retention relative to lakes with longer average link lengths. In addition, lakes with less upstream lake area had higher P retention than lakes with more upstream lake area (Figure 1.6, 1.7).

We found that some connectivity metrics were more strongly related to P retention than others. For example, the model R^2 for network average link length was higher ($R^2 = 0.41$) than the global model ($R^2 = 0.34$). For other connectivity metrics, such as network stream order ratio, goodness-of-fit ($R^2 = 0.36$) was very similar to the global model ($R^2 = $
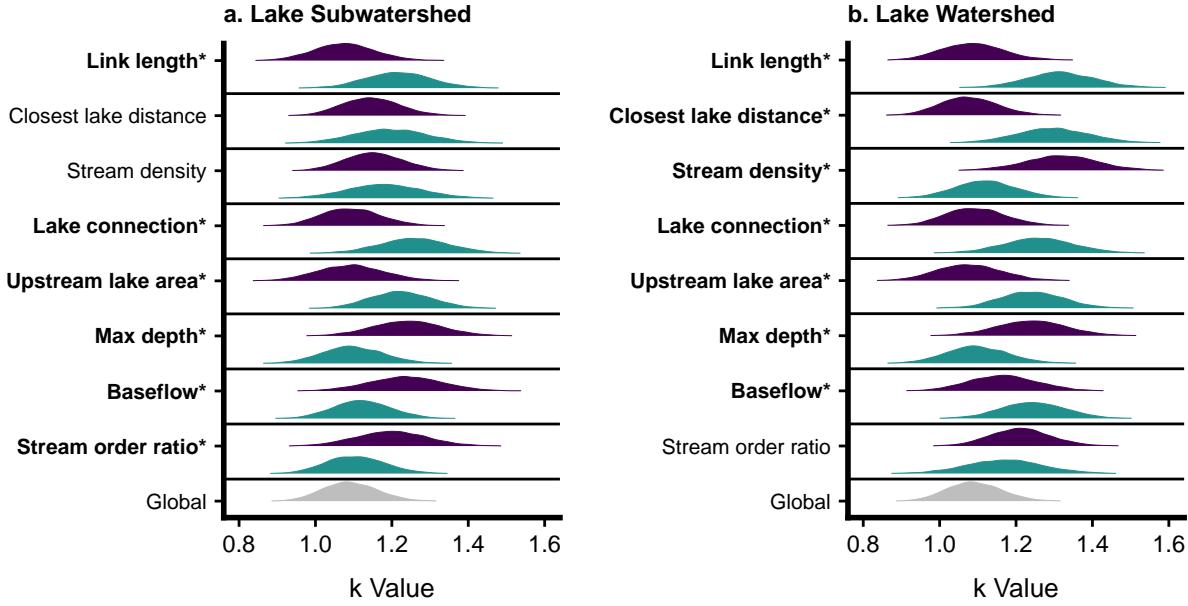
Figure 1.7: Distribution of the k parameter from Eq. 1.1 in lakes of differing connectivity, depth, or baseflow at the (A) lake subwatershed and (B) lake watershed extents. Lighter (green) lines indicate lakes with lower connectivity metrics values while darker (purple) lines indicate lakes with higher connectivity metrics values. For lake connection, lighter colored lines indicate lakes without upstream lakes. Connectivity metrics are defined in Figure 2. Labels associated with models where differences in k translated to significant differences in P retention are bolded and starred.

0.34). Model fit for other connectivity metrics was in between these two extremes ($0.36 < R^2 < 0.41$). Overall, we found that all hierarchical models had at least a marginally better fit to the water residence time versus P retention relationship than a global model which does not account for connectivity (Figure 1.6). Although we found a discernable effect of connectivity metrics on lake P retention, the somewhat modest improvements in model fit may be due to the fact that water residence time remains a dominant effect even after accounting for freshwater connectivity.

1.4.3 Comparison across connectivity metrics and spatial extent

Differences in P retention among lakes with different connectivity metric values was reflected in differences among connectivity class-specific values of the P decay parameter k (Figure

21

					Low	High
Metric	Units	Scale	Delta k	Split Value	N	N
Average link length	m	WS	0.23*	2380	33	96
Closest lake distance	m	WS	0.22*	3774	26	103
Stream density	-	WS	0.2*	13.84	95	34
Lake connection	-	focal	0.17*	-	27	102
Upstream lake area	ha	WS	0.16*	154	62	67
Max depth	m	focal	0.15*	19.81	87	42
Average link length	m	SWS	0.14*	2177	36	93
Upstream lake area	ha	SWS	0.13*	279	69	60
Baseflow	-	SWS	0.12*	63.76	113	16
Stream order ratio	-	SWS	0.1*	0.67	87	42
Baseflow	-	WS	0.08*	53.43	65	64
Closest lake distance	m	SWS	0.05	3274	16	113
Stream order ratio	-	WS	0.04	0.4	17	112
Stream density	-	SWS	0.03	4.43	24	105

Table 1.2: Connectivity class split values and samples sizes for connectivity metrics and lake depth ranked according to the difference in median k (P decay parameter, Δk) values. Differences in Δk that translate to differences in P retention greater than measurement precision are marked with an asterisk. Here, WS is the lake watershed extent whereas SWS is the lake subwatershed extent. Lakes with metric values above or equal to the split value were assigned to a separate connectivity class relative to lakes below the split value. N is sample size. Connectivity metrics are defined in Figure 1.2.

1.7). The connectivity metric that had the most effect on k was average link length (Table 1.2). For instance, hierarchical models fit with the average link length metric had a median effect size of 0.23 and 0.05 for k and P retention respectively (Table 1.2), which means that for this metric, lakes with shorter average link lengths retained 4.7 to 4.9% more P than lakes with longer average link lengths. The influence of lake-stream-based connectivity metrics on lake P retention was similar to stream-based connectivity metrics (Figure 1.7). This suggests that both lake-stream-based connectivity between lakes and streams as well as stream-based connectivity of lakes and their terrestrial watersheds influence P retention (Figure 1.2).

We found that connectivity metrics measured at the lake watershed extent were more strongly associated with P retention than metrics measured at the lake subwatershed extent (Table 1.2, Figure 1.7). Specifically, the metrics that had the strongest association

with P retention such as average link length ($\Delta k = 0.23$), closest lake distance ($\Delta k = 0.22$), and stream density ($\Delta k = 0.20$) also had a stronger association at the lake watershed extent rather than at the lake subwatershed extent (Table 1.2, Figure 1.7). An exception to this pattern was observed with the baseflow connectivity metric where although greater differences at the lake subwatershed extent were more strongly associated with P retention, the sign of the effect was variable depending on measurement extent (i.e. the value of the P decay parameter was positively related to connectivity metric values at the subwatershed extent but was negatively related at the watershed extent). For several connectivity metrics, we judged that differences in P decay among lakes with either low or high connectivity metric values were not significant because they translated to differences in P retention that were less than measurement precision (Table 1.2).

1.5 Discussion

Although prior studies have found that lake P retention is related to water residence time (Brett and Benjamin, 2008; Cheng et al., 2010; Vollenweider, 1975), there is substantial variation around this relationship, particularly at intermediate water residence times. We found that some of this remaining variation could be explained using a hierarchical modelling framework that accounts for differences in freshwater connectivity among lakes. Although we found that the magnitude of this effect depends on the specific connectivity metric, our results are consistent with the findings of previous studies showing that connectivity metrics are associated with differences in lake carbon input fluxes and differences in lake nitrogen output fluxes (Cardille et al., 2007; Schmadel et al., 2018). We also found important differences in the association between connectivity metrics and P retention at different measurement extents. Specifically, we found that P retention was more strongly associated with connectivity measured at the broader lake watershed extent rather than connectivity measured at the finer lake subwatershed extent.

1.5.1 Connectivity and P retention

We found that differences in P retention among lake connectivity classes was influenced by specific connectivity metrics including average link length, closest lake distance, and stream density (Table 1.2, Figure 1.7). Indeed, these metrics were more strongly associated with P retention than covariates typically used in statistical P retention models (e.g. lake depth, Figure 1.1). Note that we were able to examine the influence of specific metrics as separate effects apart from water residence time because our model structure treats them as hierarchical coefficients on the P decay parameter k. As a result, although water residence time remains the dominant effect on lake P retention, we were able to estimate the specific influence of each metric in a way that is not possible using integrated water residence time and connectivity metrics such as watershed transport capacity (Fraterrigo and Downing, 2008).

Several of the connectivity metrics that were most strongly associated with P retention were weakly correlated with watershed size (Figure 1.4). The weak nature of these correlations are consistent with the mixed results of prior studies linking watershed size to lake processes. For example, Zimmer and McGlynn (2018) found that carbon export was related to watershed size, but Smith (2003) found that the nitrogen flux was not strongly related to watershed size. Taken together, our results and the results of prior studies suggest that watershed size and lake depth alone may not always reflect functional differences in potential material transport. One consequence of a correlation between connectivity metrics and watershed size is that metrics derived from watershed size, such as catchment to lake area ratio, are also likely to be associated with connectivity. Catchment to lake area ratio in particular has been previously used as an approximation or proxy of water residence time (Rasmussen et al., 1989; Sobek et al., 2007). Although our results differ from those of Soranno et al. (2015) who found that the presence of an upstream lake connection was not strongly associated with catchment to lake area ratio, our results suggest that catchment to lake area ratio likely incorporates some connectivity information and caution is needed before using it

as a proxy for water residence time.

Another lake characteristic that was strongly related to watershed size was P loading. In particular, positive correlations between P loading and watershed size are consistent with the idea that lakes in larger watersheds receive greater P loading (T-Prairie and Kalff, 1986). A related observation is that the correlation between P loading and upstream lake area was positive. This can be explained by the fact that larger watersheds often have greater numbers of lakes nested within them (Zhang et al., 2012). Another notable result is that we observed a weaker correlation between P loading and watershed area relative to the correlation between P loading and subwatershed area. One explanation for this result is that land-use cover at the finer lake subwatershed extent has more of an influence on P loading than land-use cover at the broader lake watershed extent (Soranno et al., 2015).

1.5.2 Relative importance of connectivity spatial extent

One of the challenges in quantifying the effect of freshwater connectivity on lake P retention is that it is not always known what spatial extent of the watershed is functionally connected to a lake. We found that connectivity at the broader lake watershed extent rather than connectivity at the finer lake subwatershed extent was more strongly associated with differences in lake P retention. In addition, we found that individual characteristics such as lake depth (or more discrete measures of connectivity such as the presence of an upstream lake) were more strongly associated with P retention than any of the other connectivity metrics measured at the lake subwatershed extent (Table 1.2). These findings contrast with Soranno et al. (2015) who examined the association between lake nutrient concentrations (as opposed to retention) and land-use measured at varying spatial extents and found that measurements at the finer lake subwatershed extent rather than measurements at the broader lake watershed extent were more strongly associated with lake nutrient concentrations. One explanation for the difference between our results and those of Soranno et al. (2015) is that connectivity metrics may reflect long-range watershed processes to a greater degree than land-use cover. An

alternative explanation is that controls of lake P retention may differ compared to controls on lake P concentration.

1.5.3 How connectivity metrics may influence P retention

Prior studies at regional extents have shown that P retention in streams and rivers is largely determined by the fate of the particulate load fraction (Cushing et al., 1993; Kronvang, 1992; Russell et al., 1998; Vanni et al., 2001). For instance, the findings of Jarvie et al. (2011) show that riverine P loads can be controlled by nonpoint-source P delivery of particulate P. Therefore, it stands to reason that P retention in lakes may also be largely determined by the fate of the particulate load fraction. However, in the context of our broad-scale study it is difficult to examine this relationship because although we have estimates of nonpoint versus point source loading we do not have explicit estimates of particulate load fraction for large numbers of lakes. Unfortunately, we cannot use nonpoint source loading as a direct proxy for particulate load fraction because the two quantities do not have a consistent relationship. For example, Russell et al. (1998) report that the particulate phosphorus fraction of nonpoint source loads can be anywhere between 62 - 90% while Djodjic and Markensten (2018) report that this fraction can be anywhere between 33 and 80%. This may explain why prior broad scale studies that estimate lake P retention have not attempted to estimate separate effects of particulate versus dissolved loading (Alexander et al., 2008; Brett and Benjamin, 2008).

We developed a conceptual model that places particulate load fraction in context with other processes affecting P retention (Figure 1.1). We expected that both land-stream and stream-lake connectivity influences how much particulate P is transported into lakes, which in turn affects their P retention. This expectation is supported by the findings of Cushing et al. (1993) as well as Guy et al. (1994) showing that particulate P can be transported beyond the direct drainage from stream-adjacent hillslopes.

Despite our inability to test such differential transport processes across many lake watersheds at broad scales, we note that such processes are indirectly supported by our finding

that connectivity metrics are associated with lake P retention. For example, differential transport of particulate matter whereby barriers to flux and differences in drainage path configuration only become apparent beyond fine spatial extents may explain why we observed stronger association of P retention with metrics measured at the broader lake watershed extent rather than metrics measured at the finer scale lake subwatershed extent. Another finding consistent with differential transport of particulate P is our observation that average link length, which approximates stream network structure and along-stream transport potential (La Barbera and Rosso, 1989), was one of the more strongly associated metrics with lake P retention. Finally, it is notable that our stream density metric was influential at the lake watershed extent but not at the lake subwatershed extent. Given that the stream density metric captures the average distance or drainage potential between any streams in the network and their adjacent hillslopes, floodplains, and wetlands (see Leibowitz et al., 2018), this suggests that differences in terrestrial runoff of particulate matter from hillslope and riparian areas are likely to be important for P retention.

Taken together, our findings are consistent with the idea that both connectivity between lakes and streams as well as connectivity of lakes and their terrestrial watersheds affect lake P retention. This conclusion matches that of prior studies showing that aquatic transport of phosphorus and nitrogen at the sub-continental scale is strongly controlled by processes affecting along-stream flux such as reservoir trapping (Alexander et al., 2008; Schmadel et al., 2018).

1.6 Conclusion

We provide evidence that freshwater connectivity has an effect on lake P retention and that connectivity metrics measured at the broader lake watershed extent more strongly captures functional differences in the effect of connectivity on P retention among lakes compared to connectivity metrics measured at the finer lake subwatershed extents. Furthermore, our results suggest that lake P retention is related to both connectivity of lakes and streams as

well as connectivity of lakes and their terrestrial watersheds. Taken together, our findings suggest that a broader network perspective would be useful for the design of regulatory frameworks and the development of best management practices focused on eutrophication, given the importance of lake P retention in determining the trophic state of lakes. Specifically, our findings highlight the need to consider cumulative network effects of P transport in addition to localized transport mechanisms.

1.7   Supplemental information

It is difficult to examine the uncertainties surrounding the relationship between water residence time and lake P retention because empirical measures of these quantities do not exist for many lakes. One of the few comprehensive sources of these data is modelling studies which produce largely unverified estimates of water residence time and P retention (Milstead et al., 2013). Such model output data may not be appropriate for statistical P retention modelling because P retention and water residence time values are not independently estimated. As an alternative, we performed a more qualitative analysis fitting separate smoothed functions to lakes of differing connectivity (see appendix source code for details). Specifically, we used the connectivity data provided in LAGOS-NE (Soranno and et, 2017) and described in detail by (Fergus et al., 2017) where lakes are grouped based on whether they have upstream lake connections (DR_LakeStream), or upstream stream connections (DR_Stream), are headwater lakes, or are isolated lakes.
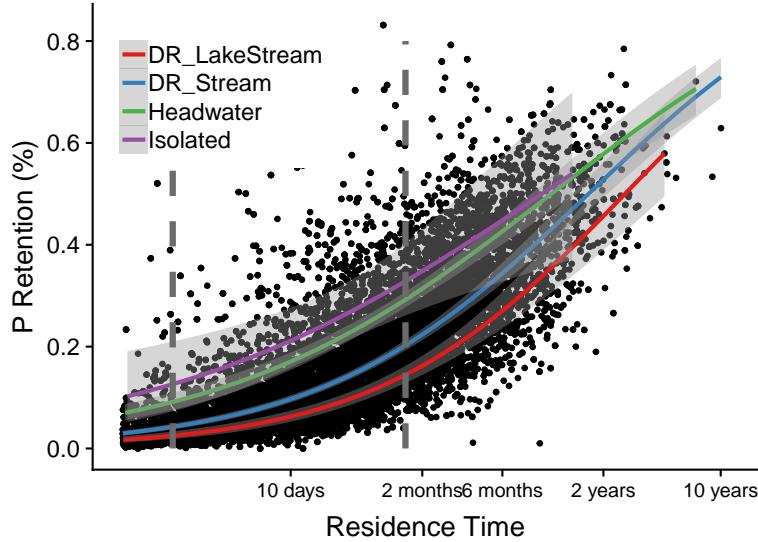
Figure S1.1: Relationship between residence time and phosphorus retention for inland lakes with surface area greater than 4 ha in Northeastern USA (New England). Vertical dashed lines denote lakes with intermediate residence times (within the interquartile range). Best fit lines are colored based on whether they have upstream lake connections (DR_LakeStream), or upstream stream connections (DR_Stream), are headwater lakes, or are isolated lakes. Data from (Milstead et al., 2013).

Our study lakes encompassed a range of land-use cover types and nutrient levels. Here, we examined whether a strong relationship of lake P retention with agricultural land use cover could be obscuring relationships with connectivity metrics. We used P retention data from the National Eutrophication Survey (Stachelek et al., 2018) and land cover data from the 1992 National Land Cover Database clipped to lake watersheds and available in the LAGOS-NE dataset (Soranno and et, 2017).

Although our study considered many lakes spread across a broad spatial extent, we wanted to assess whether their basic characteristics were representative of lakes in general. We compared water quality (total phosphorus, chlorophyll concentration, and Secchi depth) of the lakes in our study with other US lakes as measured by the stratified random sampling design of the National Lakes Assessment (NLA) lake population (USEPA, 2016). Our lakes are similar in most respects except that they are substantially larger and deeper than most NLA lakes.

29
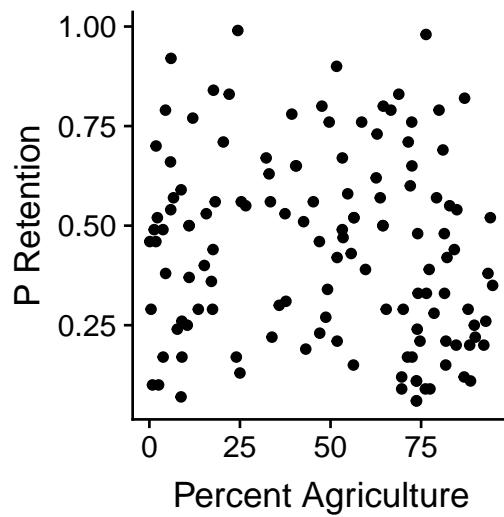
Figure S1.2: P retention versus percent agriculture in lake watersheds.
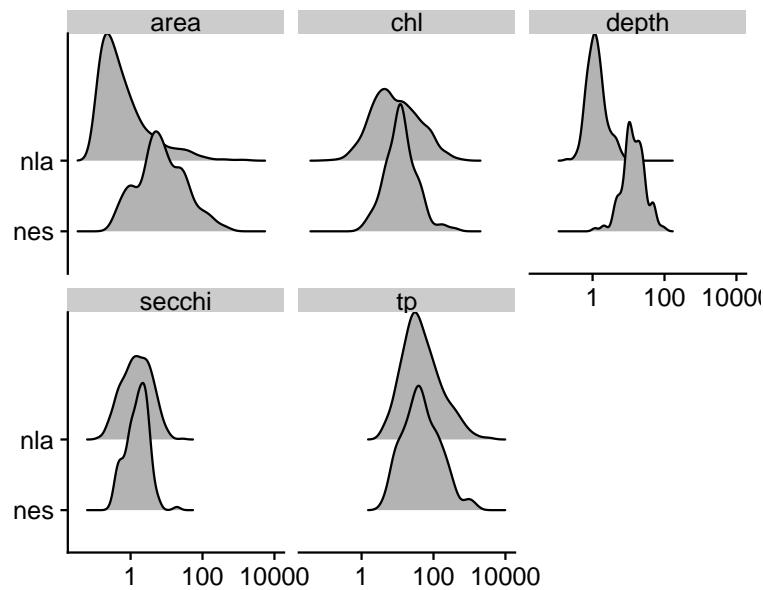
Figure S1.3: Comparison of selected lake characteristics among the stratified random sampling design of the National Lakes Assessment (nla) and the haphazard sampling of the National Eutrophication Survey (nes) lakes analyzed in the present study.

GRANULAR MEASURES OF AGRICULTURAL LAND-USE INFLUENCE LAKE NITROGEN AND PHOSPHORUS DIFFERENTLY AT MACROSCALES

Stachelek, J., Weng, W., Carey, C. C., Kemanian, A. R., Cobourn, K. M., Wagner, T., Weathers, K. C., and Soranno, P. A., 2020. Granular measures of agricultural land-use influence lake nitrogen and phosphorus differently at macroscales. *Ecological Applications.* doi:10.1002/eap.2187

2.1   Abstract

Agricultural land-use is typically associated with high stream nutrient concentrations and increased nutrient loading to lakes. For lakes, evidence for these associations mostly comes from studies on individual lakes or watersheds that relate concentrations of nitrogen (N) or phosphorus (P) to aggregate measures of agricultural land-use, such as the proportion of land used for agriculture in a lake's watershed. However, at macroscales (i.e., in hundreds to thousands of lakes across large spatial extents), there is high variability around such relationships and it is unclear whether considering more granular (or detailed) agricultural data, such as fertilizer application, planting of specific crops, or the extent of near-stream cropping, would improve prediction and inform understanding of lake nutrient drivers. Furthermore, it is unclear whether lake N and P would have different relationships to such measures and whether these relationships would vary by region, since regional variation has been observed in prior studies using aggregate measures of agriculture. To address these knowledge gaps, we examined relationships between granular measures of agricultural activity and lake total phosphorus (TP) and total nitrogen (TN) concentrations in 928 lakes and their watersheds in the Northeastern and Midwest U.S. using a Bayesian hierarchical modelling approach. We found that both lake TN and TP concentrations were related to these measures

of agriculture, especially near-stream agriculture. The relationships between measures of agriculture and lake TN concentrations were more regionally variable than those for TP. Conversely, TP concentrations were more strongly related to lake-specific measures like depth and watershed hydrology relative to TN. Our finding that lake TN and TP concentrations have different relationships with granular measures of agricultural activity has implications for the design of effective and efficient policy approaches to maintain and improve water quality.

2.2   Introduction

Freshwaters are vulnerable to eutrophication in areas of high agricultural land-use and land cover because agricultural activities are associated with high nutrient runoff and loading to groundwater and streams (Allan, 2004; Arbuckle and Downing, 2001; Daniel et al., 2010; Taranu and Gregory-Eaves, 2008). High runoff and loading in these areas is a result of high rates of nutrient input combined with hydrologic modifications that decrease the travel time of these inputs from the land surface to lakes (Blann et al., 2009). Surprisingly, while some studies have found strong relationships between agricultural land-use and land cover (hereafter referred to as "land-use" or LULC) and lake nutrient concentrations (Arbuckle and Downing, 2001; Downing et al., 2001; Taranu and Gregory-Eaves, 2008), others have found more mixed results (Jones et al., 2008), particularly in studies that include many lakes located in multiple regions (Soranno et al., 2015). Further examples of such macroscale studies (see Heffernan et al. 2014) in which lakes are spread across many regions at distances spanning hundreds to thousands of kilometers include Collins et al. (2017) and Read et al. (2015), who found that the strength of agriculture and lake nutrient relationships varied depending on geographic region and lake characteristics.

Mixed results from prior studies may be due to two difficulties in quantifying lake nutrient and agricultural land-use relationships. First, the pathways of nutrients from fields to streams and ultimately to lakes are complex and indirect (Cherry et al., 2008; Heathwaite

et al., 2003; King et al., 2005). For example, in order for nitrogen applied as fertilizer to reach a lake, it must be transported to streams or groundwater in excess of microbial denitrification, plant use, and microbial uptake. Then, it must travel, again, often undergoing repeated chemical transformation as it passes through riparian buffers and along stream networks before finally entering the lake (Maranger et al., 2018). Each step of the journey represents an opportunity for those nutrients to be sequestered or removed. Further, overall hydrologic transport is influenced by soil type (Naiman et al., 2010), and topography. Thus, the nutrients that ultimately enter a lake are a function of filtering by the landscape as well as geochemical transformation processes that are difficult to capture at broad scales (Canham et al., 2012; Maranger et al., 2018).

Second, much of our evidence for a connection between agricultural land-use and increased nutrient concentrations comes from studies focusing on a single watershed or on several watersheds within a single geographic region (Capel, 2018; Daniel et al., 2010; Hayes et al., 2015; Renwick et al., 2008). These studies tend to focus on very detailed measures of agricultural activity such as tillage and other practices, nutrient amendments, and their spatial arrangement. However, studies at broader spatial scales (i.e., the macroscale, Heffernan et al. 2014, which aim to provide a more general view of relationships between lake nutrient concentrations and agriculture, tend to focus only on relatively coarse measures of agricultural activity (Filstrup et al., 2018; Read et al., 2015; Schmadel et al., 2019). As a result, it is unclear to what extent the mixed results from prior broad-scale studies might be due to the limited use of detailed measures of agricultural activity or simply from regional variation.

There are two ways in which detailed (i.e., granular) measures of agricultural activity may be substituted for their coarse (i.e., aggregate) counterparts. The first is by using granular measures that are recorded in the same locations as their aggregate equivalents but are more descriptive. For example, in broad-scale studies, the proportion of land used for agriculture in a lake watershed is sometimes replaced by separate representations of the land used for pasture and the land used for row-crops (Abell et al., 2011; Collins et al., 2017). The second

is by using granular measures that have the same description as their aggregate equivalents, but that are measured in more specific locations. For example, a small number of lake studies have compared the proportion of land used for agriculture in near-stream buffers versus the watershed as a whole (Gémesi et al., 2011; Soranno et al., 2015). Although the term granular can be used in a general sense to describe any detailed agricultural measure, we define the term more narrowly as only those that have a specific aggregate counterpart (Table 2.1).

Prior use of granular data in broad-scale studies of lake water quality has been limited. For this reason, findings from broad-scale studies may be less useful in applied management settings because coarse (i.e. aggregate) land-use and land cover change metrics have become less widely used policy instruments (Morefield et al., 2016). Instead, recent policy interventions go beyond aggregate measures of agricultural activity to target more specific measures such as implementation of specific farming practices, no-till agriculture, and construction of riparian buffer strips (Capel, 2018; NRC, 2010; Yang et al., 2005). Thus, broad-scale studies could be made more relevant for informing policy interventions if they used covariates that have a similar level of granularity to those used in fine-scale studies. For example, implementation of no-till agriculture policies may be better informed by covariates at the granular crop level rather than solely by aggregate covariates like land used for agriculture (Figure S2.1).

One likely reason that broad-scale studies have rarely used granular agricultural data is that until recently, such data have not been available with corresponding lake nutrient concentration data over large geographic extents. Although a few examples exist of studies connecting granular agricultural data to either stream nitrogen or phosphorus concentrations at the macroscale (Boyer et al., 2002; Bellmore et al., 2018; Metson et al., 2017), most studies have focused on either nitrogen (N) or phosphorus (P) but not on both at the same time (Alexander et al., 2008). Crucially, we are not aware of prior work examining relationships between a multiple granular measures of agricultural activity on either lake N or P concentrations at macroscales.

Variable	Granularity	Median	Q25	Q75
Ag (percent)	Aggregate	42.0	25.00	63.0
Pasture (percent)	Granular	14.0	7.20	24.0
Corn (percent)	-	6.9	2.20	17.0
Soybeans (percent)	-	4.5	0.86	14.0
Buffer Ag (percent)	-	25.0	11.00	48.0
Buffer natural (percent)	-	41.0	23.00	59.0
Fertilizer N (kg/ha)	-	55.0	32.00	91.0
Fertilizer P (kg/ha)	-	9.6	6.40	16.0
Manure N (kg/ha)	-	27.0	17.00	45.0
Manure P (kg/ha)	-	7.0	4.80	12.0
Forest (percent)	Other	25.0	12.00	46.0
Wetlands (percent)	-	2.8	0.52	8.2
N deposition (kg/ha)	-	5.7	4.60	7.0
Precipitation (mm/yr)	-	910.0	830.00	1000.0
Baseflow	-	49.0	33.00	62.0
Wetland potential (percent)	-	15.0	5.30	26.0
Soil organic carbon (g C/m2)	-	4000.0	2900.00	5300.0
Clay (percent)	-	10.0	5.00	17.0
Max depth (m)	-	9.4	6.10	14.0
Watershed-lake ratio	-	15.0	6.00	34.0

Table 2.1: Medians followed by first and third quantiles of predictor variables for 928 lakes. Also shown is whether each predictor is defined as an aggregate or granular measure of agriculture or as a non-agriculture (other) predictor. Dashed entries for the granularity category indicate an identical categorization as the preceding predictor.

There are several plausible expectations, which are based on the findings of broad-scale stream studies as well as the findings of fine-scale lake studies, for the type of relationships between such measures and lake nutrient concentrations that may emerge at macroscales. First, we expect that increased nutrient inputs to the land surface as fertilizer and manure will increase lake nutrient concentrations (Bellmore et al., 2018; Renwick et al., 2008). Second, we expect that lakes with watersheds that have high soil clay content and high soil organic carbon content will have higher lake nutrient concentrations. This will occur because clay soils tend to have higher rates of surface runoff and have higher organic matter content relative to sandy soils, despite the tendency for higher organic matter content to increase water storage and reduce surface runoff (Capel, 2018). Finally, we expect that lakes with

stream networks characterized by extensive near-stream agriculture will have higher nutrient concentrations because there will be less interception of agricultural runoff (Naiman et al., 2010).

Although we can formulate potential expectations for relationships between agricultural activities and lake nutrients at the macroscale by building on the findings of prior studies, several key uncertainties remain. The first key uncertainty is the extent to which lake and watershed characteristics, such as watershed hydrology and soil type, affect relationships between granular measures of agricultural activity and lake nutrient concentrations at the macroscale. For example, macroscale studies have found that lake P concentrations are strongly dependent on lake depth (Collins et al., 2017), but the degree to which granular agricultural data provide additional explanatory power is unknown. Similarly, Abell et al. (2011) found that watershed to lake area ratio (i.e. lake water residence time) was positively related to lake N concentrations after controlling for aggregate measures of agricultural land-use, but it is unknown whether this mediation effect would also affect relationships with more granular measures of agriculture.

The second key uncertainty is the extent to which relationships between granular measures of agricultural activity and lake nutrient concentrations vary regionally. For example, previous macroscale studies on lake nutrient concentrations have found that relationships between lake chlorophyll and nutrient concentrations vary regionally according to hydrologic subregions (Wagner et al., 2011; Qian et al., 2019). Models in which separate relationships (e.g. slopes) are estimated for different regions, such as those used by Wagner et al. (2011), can be used to test differences among regions in the sensitivity to nutrient predictors. Wagner et al. (2011) found that the slope of the chlorophyll to P relationship was notably higher in several of their study regions and found that lakes with high pasture land-use in their watershed were more sensitive to changes in P concentrations (larger, positive slope estimate). They suggest that elevated sensitivity in high pasture regions is due to dual N and P nutrient enrichment associated with this land cover type. Given the findings of this and other studies

in stream ecosystems (Alexander et al., 2008), it follows that other relationships in lakes may vary regionally, but whether or not this includes relationships with granular measures of agricultural activity remains unknown.

We addressed the knowledge gaps described above by asking two questions using data from approximately 900 lakes in the Northeastern and Midwestern US: 1) How do granular measures of agricultural activity relate to lake N and P concentrations? And, 2) How do relationships between agricultural activities and lake nutrients vary regionally among hydrologic and climatic regions? To answer these questions, we fit statistical models of lake nutrient concentrations as a function of granular measures of agricultural activity such as the proportion of watershed area land-use of specific crops, and near-stream land-use, as well as fertilizer and manure applications. We also included a variety of lake and watershed characteristics as predictors to account for the influence of other drivers. Finally, our models included a hierarchical component where relationships between watershed land-use and lake nutrient concentrations were allowed to vary among hydrologic subregions to examine potential regional variation in these relationships.

2.3   Methods

2.3.1   Data description

We analyzed data on lake total nitrogen (TN) and total phosphorus (TP) concentrations from the LAGOS-NE data product (Soranno and Cheruvelil, 2017a). The LAGOS-NE nutrient data are limited to surface (or "epilimnetic") samples and were derived from federal, state, tribal, and non-profit agencies, as well as university researchers and citizen scientist data collections (Soranno and et, 2017). We collapsed lake nutrient data to long-term median values computed using all data from the summer stratified period (i.e. 15 June through 15 September) available for each lake between 2000 and 2010. For the lakes that met our selection criteria (see below) the median number of samples per lake was nine (range = 3-94) and the median number of years sampled was four (range = 2-10).

Our first source for granular agricultural information was the 2010 Cropland Data Layer (CDL, USDA-NASS 2019), which we used to generate lake watershed and stream buffer estimates of the proportion of land cover due to specific crops such as corn and soybeans. We used the R package `cdlTools` to access CDL products (Chen and Lisic, 2018). We used the 2010 CDL data because this was the first year of high resolution (30 x 30m) CDL data collection that matched our lake nutrient data collection window. Because the CDL contains detailed coverage for dozens of crops, including rare crops with little to no coverage, we re-categorized CDL data based on Lark et al. (2015) to a more limited set of categories (Table S2.1).

As a more granular representation of watershed agricultural land use, we measured the proportion of land cover in agricultural and "natural" land-uses in 100-m buffers of lake shorelines and streams for each lake watershed. We chose a 100-m buffer width because this width is inclusive of nearly all "riparian buffers" (Mayer et al., 2007). Here agricultural land-use is defined in the aggregate sense as the proportion of land used for agriculture in a lake's watershed whereas "natural" land-uses include the sum of all non-agriculture and non-developed CDL classes. We identified the streams associated with each lake using the stream network extraction tool in the `nhdR` interface (Stachelek, 2019b) to the National Hydrography Network (USGS, 2019). We compiled soil characteristics for each lake watershed using the Gridded Soil Survey Geographic Database (gSSURGO, Soil Survey Staff 2016) where we used the python package `gssurgo` (Stachelek, 2019a) to access gSSURGO products. One such product was wetland potential, defined as the percentage of the soil grid that meets the criteria for hydric soils formed under conditions of saturation, flooding, or ponding long enough to develop anaerobic conditions but not so long as to be classified as a permanent waterbody (Soil Survey Staff, 2016). Finally, we compiled mean annual nutrient inputs via fertilizer and manure to lake watersheds from 1982-2001 using county level data provided by Ruddy et al. (2006). We spatially aligned these county level estimates to lake watershed polygons provided by LAGOS-NE (Soranno and Cheruvelil, 2017b) using the area-weighted

interpolation functions provided by the `sf` R package (Pebesma, 2018). All of our data and data processing code are available at doi:10.5281/zenodo.3754916.

2.3.2 Location information

We restricted the lakes included in the study to those located within the footprint of LAGOS-NE which includes lakes located in 17 Northeastern and Midwest U.S. states (Soranno and et, 2017). We excluded lakes from our analysis if they had a surface area $> 400 \ km^2$ or a maximum depth > 35 m. These removals resulted in exclusion of approximately 40 lakes which we regarded as outliers that would likely have undue influence on model results because such large and deep lakes are likely to respond differently to enrichment as a result of enhanced stratification. To ensure an adequate comparison between aggregate and granular measures of agriculture, we further limited lakes in our study to those with at least 10% of the total watershed area devoted to agricultural land-use and those that were sampled at least three times between 2000 and 2010. A total of 928 lakes met these selection criteria (Figure 2.1). Because we focused on lakes in agricultural watersheds, more than 35 percent of lakes in our study are considered eutrophic to hypereutrophic using chlorophyll a as a diagnostic versus only 15 percent for all lakes from Soranno and et (2017) located within our study extent (Figure S2.3). For the regional terms in our models, we used hydrologic regions at the subbasin (i.e. HUC4) level because this level was small enough to give a sense of overall spatial variation but large enough to encompass sufficient numbers of lakes to estimate within region variance.

2.3.3 Model overview

We evaluated the effects of agricultural activities on lake TN and TP concentrations using a Bayesian hierarchical modelling approach. A list of lake and watershed covariates with their summary statistics is available in Table 2.1. In the first part of our model evaluation, we compared models that each had only a single measure of watershed land-use along with all
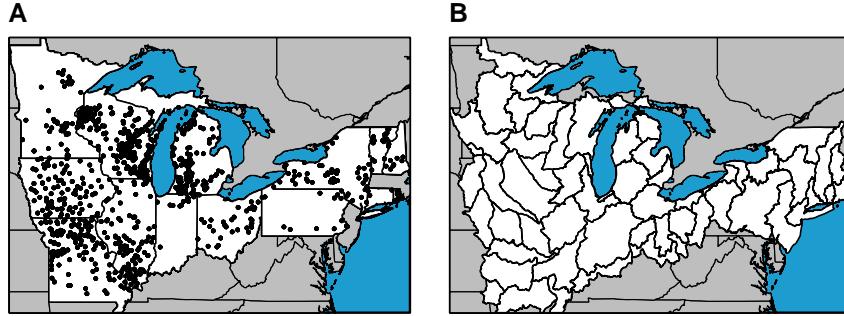
Figure 2.1: A) Map of lake locations and B) hydrologic (HUC4) regions.

remaining non-watershed land-use predictors. These non-watershed land-use predictors were included in every fitted model and were defined as measures of near-stream land-use, soil characteristics, and lake and watershed characteristics. We took this approach of evaluating one watershed land-use measure at a time for two reasons: 1) because measures of watershed land-use were highly correlated, and 2) because it allowed us to more rigorously test our expectation that granular measures of agriculture provide additional explanatory power beyond that offered by more typical aggregate measures of agriculture. In the second part of our model evaluation, we selected the top-ranked watershed land-use model according to our selection criteria for lake N and P (see below) for further inspection of their standardized coefficient values (See Table 2.1). Our models were of the form:

$$y_i = N(\alpha_{j(i)} + \beta_1 * X_{1i}...\beta_n * X_{ni} + \gamma_{1j(i)} * W_{1i} + \gamma_{mji} * W_{mi})$$
$$\begin{pmatrix} \alpha_j \\ \gamma_{1j} \end{pmatrix} \sim MVN \left(\begin{pmatrix} \mu_\alpha \\ \mu_{\gamma 1} \end{pmatrix}, \Sigma \right)$$
(2.1)

where y_i is either TN or TP concentration for lake i, and β are "global" (i.e. fixed effect) coefficients. This set of global coefficients included estimates for watershed soil characteristics, near-stream land-use, as well as fertilizer and manure inputs for each lake (X_i). Whereas β coefficients were estimated as fixed effects, γ coefficients and α intercepts were estimated as varying (i.e. random) slopes and intercepts respectively among m hydrologic regions j on watershed land-use (W_{ij}). Region specific intercepts α and γ slopes were assumed to come from a multivariate normal distribution (MVN) where μ_α and $\mu_{\gamma 1}$ are their respective grand

mean (i.e., population level) estimates. We tested a variety of watershed land-use types for W_{ij}, including both granular and more aggregate measures of agricultural activity. Our only aggregate measure of agricultural activity was agricultural land-use whereas we used several granular measures of agricultural activity. These included both detailed measures of watershed land-use such as corn and soybean cover as well as more detailed measures of nutrient inputs and near-stream land-use (Table 2.1). We included measures of both N and P inputs in both N and P models because of the possibility for stoichiometric interaction. Given our regularization scheme (i.e. our use of "horseshoe priors" described below) there was little reason to exclude P inputs from N models (and vice versa) because unless P inputs are strongly related to lake TN concentration, their coefficient will be forced close to zero.

All models had the same set of fixed effect coefficients while each individual model used a single different watershed land-use variable as a random effect. This modelling strategy is supported by our view that watershed land-use is an indirect proxy (sensu Burcher et al. 2007; Hayes et al. 2015; King et al. 2005) for other unquantifiable agricultural activities. Therefore, we expect the makeup of specific activities represented by this indirect measure to vary regionally. This contrasts with other predictor variables e like lake depth, where we have little evidence from prior studies that its effect on lake nutrient concentrations is spatially variable. Prior to model fitting, we examined the bivariate relationships between lake nutrient concentrations and all predictor variables using Pearson's correlation coefficients (Figure A5) to determine the overall structure of the predictor dataset. We did not use the results of this exercise for building our model, performing model selection, or variable selection. For qualitative analysis of our model results, we classified predictor variables following Collins et al. (2017) into categories based on the dominant mechanism affecting lake nutrient concentrations, which includes nutrient inputs, nutrient transport, spatial configuration of land-use in stream buffers (Spatial config.), lake characteristics, and watershed land-use (LULC).

We fit all models in a Bayesian framework using the `brms` R package interface to the Stan statistical program (Bürkner et al., 2017; Stan Development Team, 2017). We used

horseshoe shrinkage priors on all fixed effect coefficients to evaluate variable importance (Carvalho et al., 2010). We considered a response variable sensitive to a given predictor if the predictor's 95% credible interval did not overlap 0. We standardized all predictor variables by subtracting the mean of each variable and dividing by their respective standard deviation so that model coefficients could be compared on roughly the same scale. As a result, the relative sensitivity of a response variable to a particular predictor is related to the relative magnitude of its coefficient estimate. We evaluated model fit of each watershed land-use variable (i.e., each model having one regionally varying coefficient) in two ways. First, we computed a Bayesian R^2 following the method of Gelman et al. (2017) and second, we computed differences in expected log predictive density (ELPD) using the leave-one-out cross e validation routines provided by the `loo` R package and implemented in `brms` (Vehtari et al., 2017). The `loo` package uses leave-one-out cross validation to estimate overall model error by computing the average error of models iteratively trained on all the data except for a single point. ELPD values have a similar interpretation to information criterion measures such as Akaike information criterion (AIC) or Watanabe-Akaike information criterion (WAIC) except that values are on a different scale (Gelman et al., 2013). Typically, models are considered to be different if they are separated by an AIC value of greater than 2 (Anderson and Burnham, 2002), which is equivalent to an ELPD value of -1 (Gelman et al., 2013). We report differences in ELPD among models using the notation ΔELPD. We selected the model with the lowest absolute value ELPD as the "top-ranked" model for detailed reporting and discussion as this signifies the model with the lowest leave one out cross-validation error for N and P respectively.

We used the default settings of `brms` to generate posterior estimates using four chains of 4,000 iterations each with no thinning and discarding the first 1,000 iterations. We examined model fits to ensure that all models had acceptable convergence of MCMC chains and had approximately normal model residuals. We further tested for spatial correlation among model residuals using the `spind` R package (Carl et al., 2018). All of our code for

model fitting and evaluation is available at doi:10.5281/zenodo.3754916.

2.4 Results

2.4.1 Effects of agriculture on lake nitrogen and phosphorus

Lake characteristics (e.g., maximum depth and watershed to lake area ratio) along with measures of nutrient transport (e.g., baseflow) and near-stream agriculture were significant predictors in all lake N and P models that we fit. When we compared among different models for each nutrient individually, we found that those with agricultural watershed land-use (in the aggregate sense) were top-ranked (i.e., had the lowest absolute value leave-one-out cross validation score) for models of both TN and TP concentrations (Figure 2.2, Table 2.2). Although we observed no difference in the specific watershed land-use predictor used in each top-ranked model, we found differences in the extent to which each top-ranked model was substantially different from lower ranked models. For example, in the case of P, all models had nearly identical R^2 (0.63) and the difference between the top-ranked model and second ranked model was modest (ΔELPD = 0.41). For N models, however, agricultural and corn land-use models had higher R^2 (0.58) compared to other models and the difference between the top-ranked and second ranked model was more substantial (ΔELPD = 2.58, Figure 2.2, Table 2.2).

When we looked more closely at each top-ranked model, we found similarities in the types of predictors that contributed significantly to the top-ranked N and P models (Figure 2.3). First, both N and P models included measures of lake characteristics such as maximum depth and watershed to lake area ratio as significant predictors, with the sign of these associated coefficients matching our conventional understanding, in which shallower lakes and lakes with greater hydrologic loads have higher TN and TP concentrations. Second, both models included measures of nutrient transport such as baseflow and precipitation, in which lakes with a "flashier" hydrology (i.e., having lower baseflow) where incoming water is primarily from surface runoff rather than from groundwater, had higher TN and TP
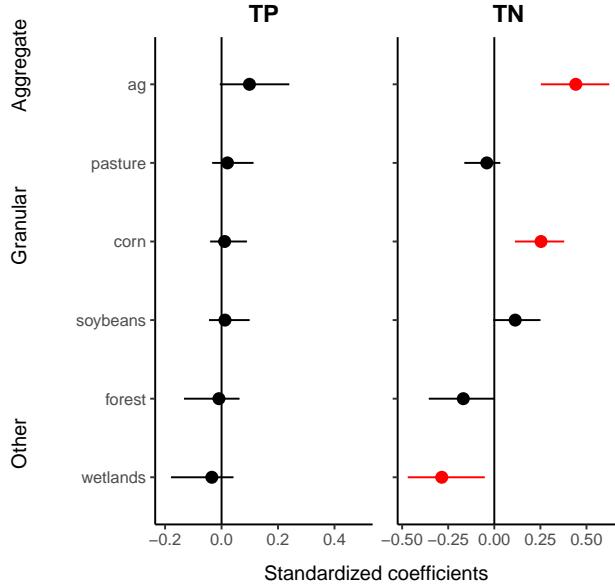
Figure 2.2: Population level slope estimates (μ_γ) for the effect of watershed land-use cover on lake TN and TP from six candidate models. Values shown are posterior medians (filled circles) and 95% credible intervals (solid lines). Also shown is a comparison to a zero effect (solid vertical line). Values that do not overlap zero are shaded in red. Coefficient estimates are reported relative to standardized predictor variables centered at zero with unit variance.

response	term	R^2	LOO-ELPD
tp	ag	0.63	0.00
tp	wetlands	0.63	-0.41
tp	corn	0.63	-0.59
tp	pasture	0.63	-0.75
tp	forest	0.63	-0.76
tp	soybeans	0.63	-1.43
tn	ag	0.58	0.00
tn	corn	0.58	-2.58
tn	wetlands	0.54	-16.41
tn	soybeans	0.53	-20.88
tn	pasture	0.53	-21.01
tn	forest	0.53	-22.37

Table 2.2: Diagnostics for each model listed by regionally varying coefficient. Table is sorted by decreasing R^2 and expected log predictive density (ELPD). ELPD has a similar interpretation to information criterion measures like AIC. Typically models are considered to be different if they are separated by an Akaike information criterion (AIC) value of greater than 2, which is equivalent to an ELPD value of -1.
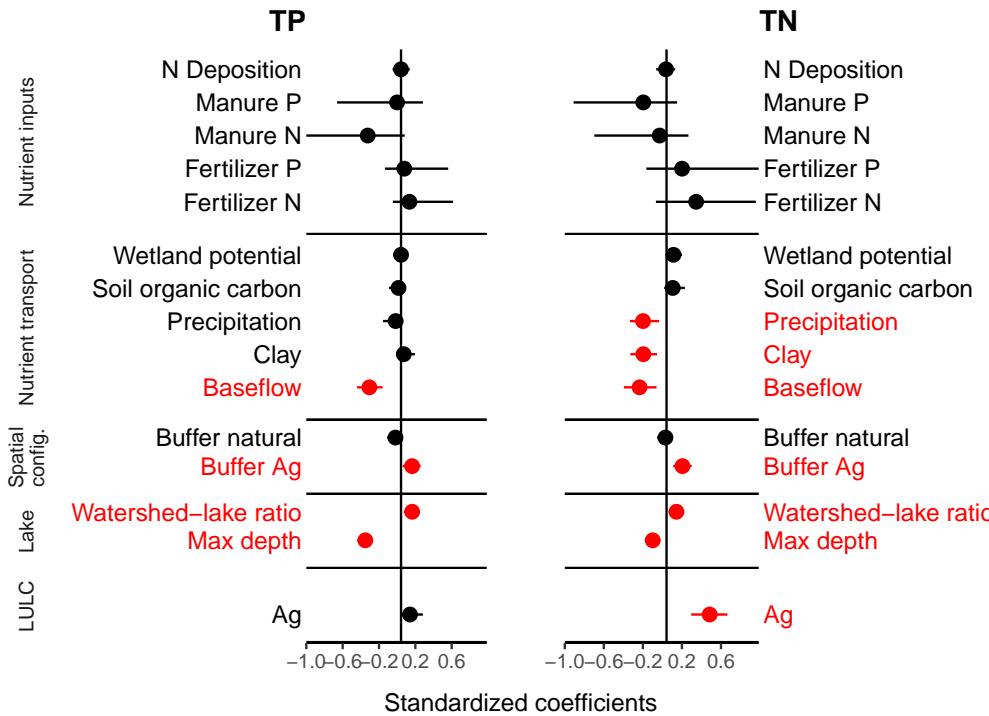
Figure 2.3: Global (fixed effect) coefficient values (β, for all non-LULC predictors) and population level estimates for the effect of watershed land-use (μ_γ, for LULC) on lake TN and TP for each respective top-ranked model. Note that the values for LULC here are identical to their corresponding values in Figure 2. Values shown are posterior medians (filled circles) and 95% credible intervals (solid lines). Also shown is a comparison to a zero effect (solid vertical line). Values that do not overlap zero are shaded in red. Horizontal bars separate coefficients in distinct predictor categories. Coefficient estimates are reported relative to standardized predictor variables centered at zero with unit variance and correspond with β (and μ_γ for LULC) from Equation 2.1.

concentrations. Finally, both models indicated that high near-stream agriculture (i.e., a high proportion of the area adjacent to the stream network was in agricultural land-use) was associated with lakes having higher TN and TP concentrations. Where N and P models differed was in the effect of soil clay content, in which soils with low clay content were associated with high lake N but had no significant relationship with P.

Although top-ranked N and P models shared some similarities in the type of predictors that contributed significantly to each model, the coefficients of these models differed in magnitude, and thus the top-ranked models varied in their sensitivity to specific predictors (Figure 2.3). For example, the top-ranked P model was more sensitive to lake

characteristics, whereas the top-ranked N model was more sensitive to watershed land-use. Quantitatively, lake TP concentrations were more sensitive to maximum depth (β_{depth}: median = -0.39, SD = 0.04) compared to lake TN concentrations (β_{depth}: median = -0.14, SD = 0.04); whereas lake TN concentrations were more sensitive to watershed agriculture land-use (β_{ag}: median = 0.44, SD = 0.11) compared to lake TP concentrations (β_{ag}: median = 0.10, SD = 0.08). Finally, although we found that near-stream agriculture was associated with both higher TN and TP concentrations (i.e., a source-effect of near-stream agriculture), there was not a significant difference in the magnitude (i.e., sensitivity) of this coefficient between N ($\beta_{bufferag}$: median = 0.16, SD = 0.06) and P ($\beta_{bufferag}$: median = 0.12, SD = 0.06) models.

No predictors in the nutrient input category appeared to be strongly related to either TN or TP concentrations. One explanation may be that these variables co-varied with watershed and near-stream land-use variables (Figure A5). In an attempt to further investigate this possibility, we fit alternative models excluding all land-use predictors. The results show that, at least in the case of N, removing these predictors caused model variance to be apportioned from watershed and buffer land-use predictors to N input and P fertilizer predictors (Figure A4). However, this non-land-use N model had a relatively poor fit ($R^2 = 0.40$) compared to the top-ranked model that included land-use as a predictor ($R^2 = 0.58$).

2.4.2 Regional variation in agriculture sensitivity

Both TN and TP concentrations were sensitive to measures of watershed land-use as well as near-stream agriculture (Figure 2.2, Table 2.2). Despite these similarities, we found differences in both the magnitude of these effects and the extent to which we observed regional variation in watershed land-use sensitivity for N and P models (Figure 2.4). For P, there was little evidence that sensitivity to watershed land-use was regionally variable. More specifically, the credible intervals for the slope of each individual region overlapped the global slope estimate (Figure 2.4). In contrast, for N we found evidence for regional variation in sensitivity
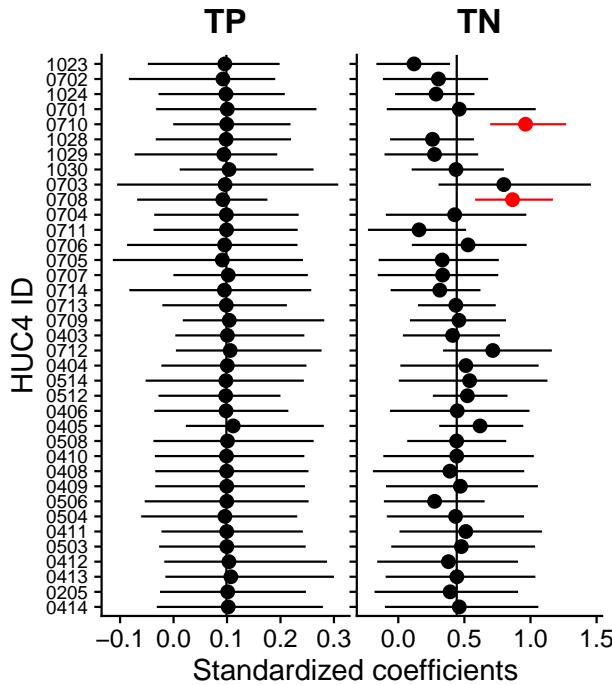
Figure 2.4: Effect of watershed land-use (γ_j) for individual regions in the top-ranked lake N and P models. Values shown are posterior medians (filled circles) and 95% credible intervals (solid lines) for individual hydrologic units (HUC4s) ordered from top to bottom according to longitude (west to east). Also shown is a comparison to a zero effect (solid vertical line). Values that are different from the population level effect are shaded in red.

to watershed land-use. For this nutrient, the credible intervals for 2 of the 37 regions did not overlap the global estimate (Figure 2.4). These regions were found in parts of Iowa, Minnesota, and Illinois and appear to be more sensitive to watershed land-use – i.e., lake N increases at a faster rate per unit increase in agricultural land-use compared to other regions (Figure 2.5). The median soil clay content of watersheds in these regions was higher than the median across watersheds in all other regions (Figure 2.6). Furthermore, these two regions had a unique combination of both high soil clay content and extensive tile drainage (Figure S2.6).
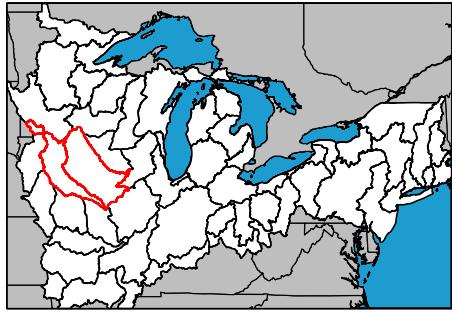
Figure 2.5: Location of hydrologic regions sensitive to watershed land-use cover corresponding to highlighted credible intervals in Figure 2.4
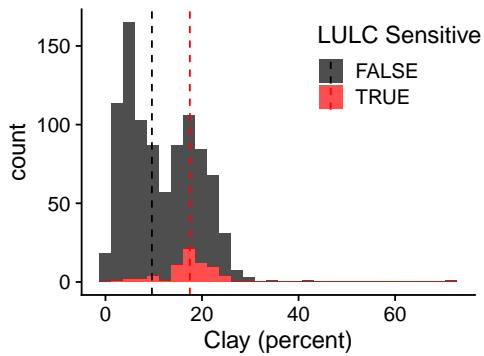
Figure 2.6: Histograms showing the distribution of soil clay content for watersheds in regions sensitive to watershed land-use (see highlighted credible intervals in Figure 2.4) relative to watersheds all other regions. Medians for each group are shown as vertical dashed lines.

2.5 Discussion

2.5.1 Effects of granular measures of agriculture on lake nitrogen and phosphorus

There is substantial unexplained variation around simple linear relationships between aggregate representations of watershed land-use in agriculture and both lake N and P (Figure S2.2). Our study was designed to examine these relationships in greater detail by testing whether more granular measures of agriculture could help explain some of this uncertainty for both N and P, and whether there were regional differences in these relationships. In sum, all models for TN and TP concentrations included at least one granular measure of agriculture, but there were also important differences between N and P related to the type of measures

that were important to each. For example, we found little benefit of increased granularity of description (i.e., where measures are recorded in all the same locations as their aggregate equivalents) but consistent benefit of granular representation of the spatial configuration of land-use in near-stream buffers.

2.5.2 Spatial configuration

The result showing that lake TN and TP concentrations are sensitive to the spatial configuration of land-use in near-stream buffers is consistent with our expectation and prior research. Specifically, our finding that model coefficients on near-stream agriculture (i.e., agricultural land-use in stream buffers) in all N and P models were significant and positive indicates a nutrient-delivery effect of stream buffer agriculture and suggests that the spatial configuration of agriculture with respect to stream buffers has a detectable influence on both lake N and P at macroscales. This is consistent with prior studies conducted over more limited spatial extents which examined relationships between lake or stream nutrient concentrations and agricultural land-use in stream buffers (Diebel et al., 2009; Baker et al., 2006; Gémesi et al., 2011; Soranno et al., 2015). In contrast to near-stream agriculture, we found that near-stream "natural" land-use was not significant in either N or P models. While we cannot definitively answer why, it may be related to the role that natural buffers play in nutrient cycling for N and P (Alexander et al., 2008; Canham et al., 2012). In the case of N, natural buffers may reduce stream loading by facilitating denitrification, whereas, in the case of P, natural buffers may trap particulate bound material without necessarily removing it (Mayer et al., 2007; Naiman et al., 2010). An alternative possibility is that we are observing a scale effect whereby the delivery effect of near-stream agriculture is spatially consistent whereas the trapping and removal effects of N and P by natural buffers is more spatially variable (Pärn et al., 2012). Finally, we may not have observed a protective effect of natural buffers because natural land-use in buffers is too coarse of a proxy for "riparian buffers" composed of forest or herbaceous vegetation (Mayer et al., 2007).

2.5.3 Crop type

We found that for N models, land-use of specific crops was significant, although it was not found in the top-ranked model. Specifically, we found that watershed land-use in corn production was a significant predictor in the second-ranked N model (Table 2.2). In the case of P by contrast, neither aggregate nor granular measures of watershed land-use were significant in any models (Figure 2.2). This can be explained by our finding that all watershed land-use metrics had weak relationships with lake P, especially relative to the strong relationships we observed between P and other factors like lake depth, hydrology, and near-stream agricultural land-use. Here, our finding is consistent with prior studies in stream ecosystems showing that the influence of hydrology exceeds that of agricultural land-use or anthropogenic P inputs (Metson et al., 2017).

2.5.4 Nutrient inputs

None of the nutrient input variables were significant for either N or P models. On the surface, this would seem to contradict the findings of prior research such as that of Bellmore et al. (2018), who found that stream TN concentrations were more strongly controlled by N input predictors relative to measures of either watershed or near-stream land-use. However, our finding has several alternative explanations. First, differences between our lake study and the Bellmore et al. (2018) stream study may simply point to differences in the controls on stream nitrogen concentrations relative to lake nitrogen concentrations (Allan, 2004; Canham et al., 2012). Second, our finding that N and P were insensitive to nutrient input variables may be a result of shared variance between nutrient input variables and watershed land-use. This makes sense given that agricultural land-uses and corn land-use in particular are associated with high rates of fertilizer and manure application (Powers, 2007). To test this explanation, we formulated N models without a watershed land-use term and observed that model variance was transferred from watershed land-use to the nutrient input terms total N input and P fertilizer (Figure S2.4). While the positive coefficient on total N input has a straightforward

interpretation, the negative coefficient on P fertilizer is unclear. Rather than evidence of a true inverse relationship between P fertilizer and N (or evidence for stoichiometric interaction sensu Paerl et al. 2016, we interpret the negative coefficient on P fertilizer as evidence of either model misspecification (i.e., model fit was poor compared to the model with land-use because N was very sensitive to land-use) or as evidence of multicollinearity among nutrient input variables (Figure S2.5). Thus, we think that it is likely that the removal of one of the key drivers of N (being land-use) caused model variance to be transferred to nutrient input predictors generally such that the specific highlighting of P fertilizer is a result of noise rather than a true relationship.

2.5.5   Nutrient transport

An important category of predictors in our models was nutrient transport. For example, we found that baseflow was a significant predictor in all N but especially all P models. This is consistent with prior research at macroscales showing the sensitivity of lake nutrients to this metric (Collins et al., 2017). In contrast to baseflow, we found that watershed soil clay content was a significant predictor in all N models but not in any P models. Furthermore, we found that the coefficient on soil clay content was negative, which was contrary to our expectation, and prior research. It seems to suggest a negative correlation between clay content and N. However, upon closer inspection, we did not find evidence for such a relationship. Instead, we found evidence for a non-linear relationship, which may explain the negative coefficient on clay, whereby clay and TN were positively correlated over most of the range of watershed clay content but were negatively correlated in watersheds with very high soil clay content ($>$ 20%; Figure S2.7). This non-linear relationship may be an artifact of the specific non-random sampling of our lakes whereby the lakes with extremely high soil clay content watersheds happen to have extremely long water residence times leading to extensive removal of N loads due to denitrification (Groffman et al., 2009). Overall, granular measures of agriculture were significant in both N and P models. However, the contribution of such measures relative

to other non-agricultural predictors was greater for N models (Figure 2.2, Figure 2.3). One reason why granular measures of agricultural activity had a greater effect on N models may be that variation in N is less effectively captured solely by lake and watershed characteristics owing in part to the more complex nature of transformations in the nitrogen cycle. This finding is consistent with that of Collins et al. (2017) and Wagner and Schliep (2018) who found that lake depth coefficients were of a much higher magnitude for P relative to N. This is expected since depth strongly controls internal P loading (i.e., recycling), which is a dominant control on lake phosphorus dynamics (Søndergaard et al., 2013).

2.5.6 Regional variation in agriculture–nutrient relationships

The macroscale nature of our study motivated our second research question examining how relationships between agricultural activities and lake nutrients vary regionally. This is because recent research has shown that analyzing macroscale lake datasets without considering the possibility of regionally varying relationships runs the risk of drawing imprecise or incorrect conclusions because it can lump together lakes with fundamentally different responses to a given predictor variable (Qian et al., 2019). Additionally, we looked at this question because prior studies have shown regional variation in the relationship of lake nutrient concentrations to aggregate measures of agriculture (Wagner and Schliep, 2018).

In our study, we found mixed evidence for regional variation in relationships with watershed land-use, depending on the lake nutrient response variable. For N, we found evidence for regional variation whereby lakes in two of the 37 regions were more sensitive to changes in agricultural land-use relative to other regions. The reasons for this elevated sensitivity are unclear, but one possible reason may be that watersheds in these more sensitive regions had higher median soil clay content than the median soil clay content of watersheds in less sensitive regions (Figure 2.6). Higher soil clay content in particular may ultimately control the nitrogen content of field runoff because it is associated with more direct (i.e., tile) drainage. For example, maps produced by Capel (2018) suggest that our more sensitive regions

correspond roughly with areas where field exports of nutrients are likely to bypass trapping by riparian buffers. As evidence of this association, data from Nakagaki and Wieczorek (2016) indicate that watersheds in these two sensitive regions had a unique combination of high soil clay content and extensive tile drainage (Figure S2.6).

For P, we found no evidence for regional variation in its relationship with watershed land-use. This finding, that watershed land-use can be modelled as a global (fixed) effect, is consistent with that of Taranu and Gregory-Eaves (2008), who found no statistically significant differences in region-specific relationships between lake P and agricultural land-use. However, it is inconsistent with that of Wagner et al. (2011), who found regionally variable relationships between lake P and agricultural land-use using a multilevel modelling framework. One reason that both our analysis and the Taranu and Gregory-Eaves (2008) study did not observe regional variation in the watershed land-use versus P relationship may be that P is so strongly controlled by lake depth that there is little additional explanatory power offered by including a watershed land-use term. In addition, if lake P is controlled primarily by lake depth, which we can assume does not change with time, then our results may explain the finding of Oliver (2017) that lake P trends are spatially consistent whereas lake N trends have distinct regional variability.

2.5.7 Management implications

Knowledge of differences in the drivers of lake N and P can support the design of effective and efficient policy approaches to maintain or improve water quality. For instance, our finding of regional variation in the relationship between lake TN concentrations and watershed land-use in agriculture suggests spatial targeting of best management practices (BMPs) to specific regions known to be highly sensitive (Holmes et al., 2016). In addition, our finding of strong relationships between lake TP concentrations and lake characteristic predictors contrasts with the strong relationships we observed between watershed land-use and lake TN. Given that watershed land-use, rather than lake characteristics, is a more feasible management target,

our results suggest that the cost-effectiveness of BMPs could differ depending on whether the goal is to protect against excess N, P, or both (Paerl et al., 2016). For P, our analysis suggests that nutrient control policies are likely to be especially effective in shallow lakes and lakes with low baseflow (i.e., those with flashier hydrology). Conversely, phosphorus control in deeper lakes and reservoirs with long residence times will likely require recovery efforts in addition to prevention efforts due to the long time-scales of stored (i.e., legacy) P (Powers et al., 2015). In contrast to P, our results show that lake TN concentrations are more sensitive to watershed land-use. This suggests that policies to enhance the use of BMPs to reduce N inputs to lakes are likely to require a greater degree of stakeholder involvement, possibly through consideration of tradeoffs between land retirement and working lands programs (Capel, 2018; NRC, 2010).

2.5.8   Future research priorities

Due to a lack of temporally resolved data, our study focused on spatial patterns in sensitivity of lake TN and TP concentrations to measures of agricultural activity and did not examine the possibility that such relationships could be temporally variable. A consequence of this was that we assumed that relationships between lake N and P relative to agricultural drivers did not change over our data collection window (2000 - 2010). However, there are several instances where time may be important, and these would likely be fruitful areas for future research. For example, Lark et al. (2015) showed marked conversion of conservation reserve program (CRP) lands to cropland throughout the footprint of our study from 2008 to 2012. Such changes in land-use could make it more difficult to quantify lake sensitivity to agriculture if relationships vary through time especially if relationships are subject to threshold effects (Renwick et al., 2008). A more subtle illustration of when time may be important is when field-scale nutrient export is highly dependent on episodic hydrology. For example, a number of previous studies have shown that field-scale nutrient export of N is greatest when a series of dry years is followed by a wet spring (Motew et al., 2017; Strickling and Obenour, 2018). It might make

sense then to organize modeling around whether lake watersheds are generally subject to slowflow, fastflow, or drainflow (Capel, 2018) nutrient transport rather than solely taking a spatial regionalization approach (i.e., using HUCs). Barriers to more detailed temporal approaches include greater demands of spatio-temporally resolved data products. Overall, our results point to hydrology predictors like baseflow as an instance where we only have spatially coarse information and development of more granular estimates of watershed hydrology, possibly using the output of hydrology models, would likely improve future research efforts.

2.6   Conclusion

We show that granular measures of agricultural activity are related to both lake N and P and that these relationships are regionally variable for lake N. Taken together, our results suggest that lake TP concentrations are more strongly driven by lake characteristics; whereas, lake TN concentrations are more strongly driven by watershed land-use. A consequence of our finding is that lake TP concentrations are largely predictable from lake-specific measures such as near-stream land-use, lake depth, and transport metrics like baseflow; whereas, accurate predictions of lake N likely requires not only lake specific information (including granular measures of agriculture) but also consideration of regional context due to complex regional variation of soil characteristics. Such differences in lake nutrient model sensitivity to measures of agricultural activity may affect the outcome of policies to enhance water quality depending on whether they focus on lake N or P.

2.7 Supplemental information
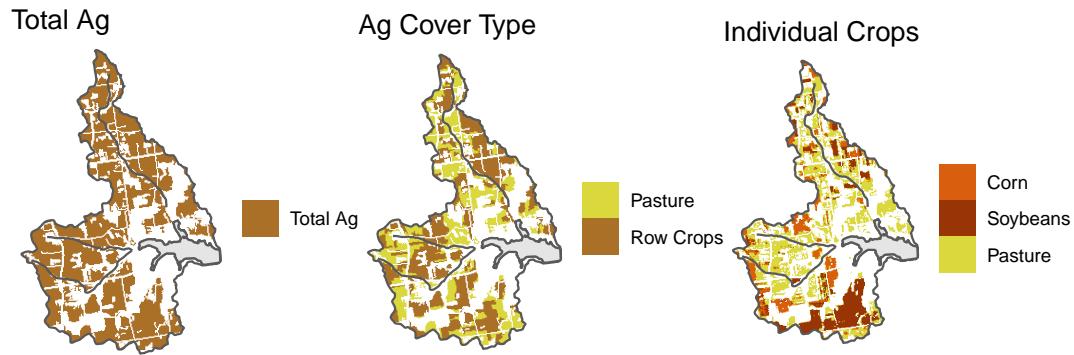
Figure S2.1: Example of increasing granularity for total Ag to Ag versus pasture, to pasture versus specific crops. For illustration, only corn, soybeans, and pasture are shown rather than all CDL land-use categories.

Category	Description
Corn	Corn
Corn	Sweet corn
Corn	Pop or orn corn
Corn	Non irrigated corn
Forest	Forest
Forest	Deciduous forest
Forest	Evergreen forest
Forest	Mixed forest
Pasture	Grass pasture
Soybeans	Soybeans
Soybeans	Non irrigated soybeans
Wetlands	Wetlands
Wetlands	Woody wetlands
Wetlands	Herbaceous wetlands

Table S2.1: Category definitions from the 2010 CDL. See code supplement for listing of variables classified as 'ag' (doi:10.5281/zenodo.3754916).
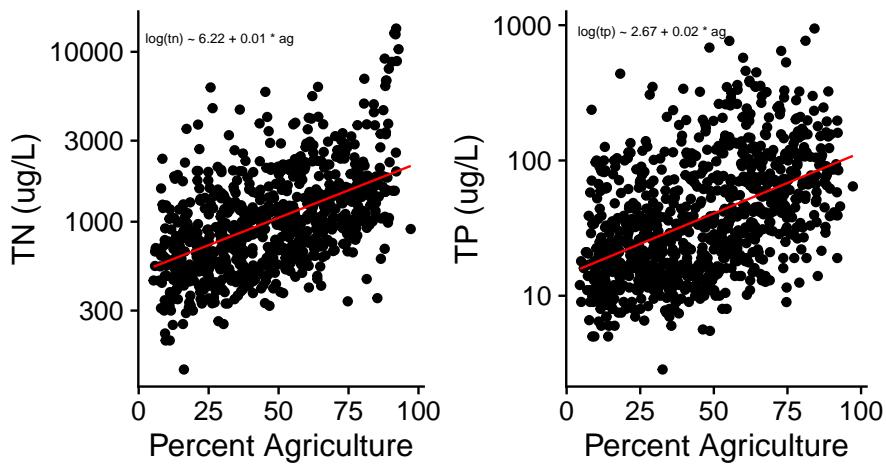
Figure S2.2: Lake nutrient concentrations plotted against percent watershed agriculture.
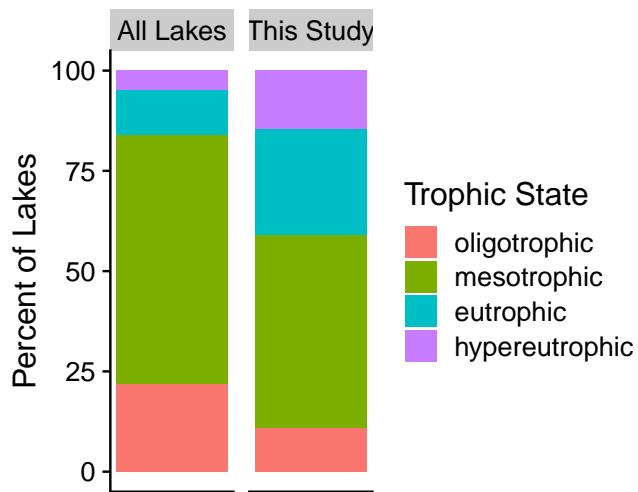
Figure S2.3: Lake trophic state in our study lakes versus all lakes from Soranno and et (2017) located within our study extent. Trophic state based on the chlorophyll criteria from Carlson and Simpson (1996).
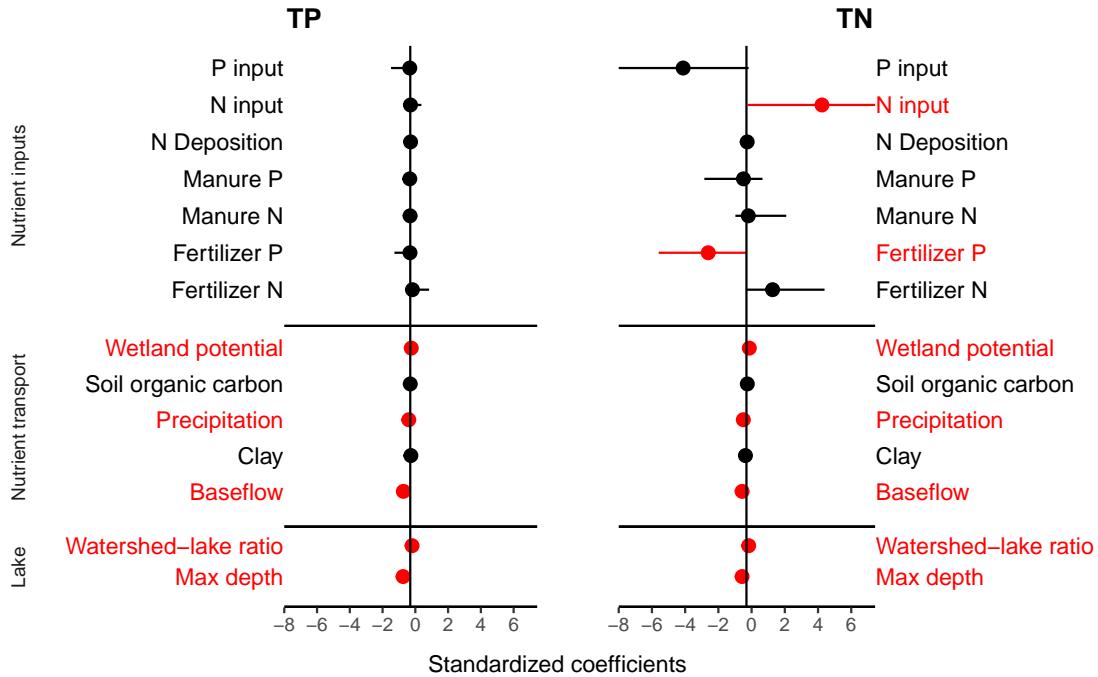
Figure S2.4: Global (fixed effect) coefficient values and credible intervals for top-ranked lake TP and TN models when land-use predictors are excluded. Values shown are posterior medians (filled circles) and 95% credible intervals (solid lines). Also shown is a comparison to a zero effect (solid vertical line). Values that do not overlap zero are shaded in red. Horizontal bars separate coefficients in distinct predictor categories. Coefficient estimates are reported relative to standardized predictor variables centered at zero with unit variance and correspond with β from Equation 2.1.
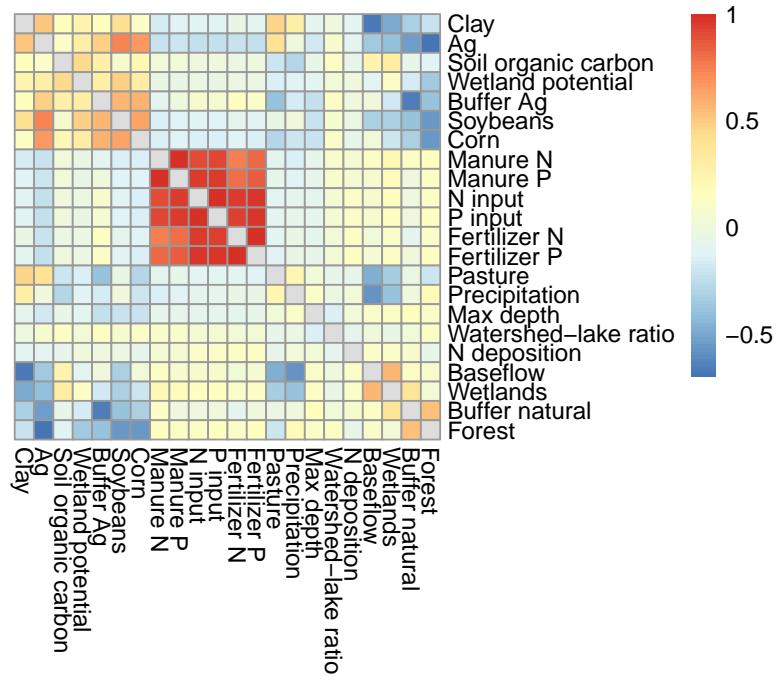
Figure S2.5: Heatmap showing Pearson correlation coefficients among predictor variables. Grey cells denote correlation matrix diagonals.
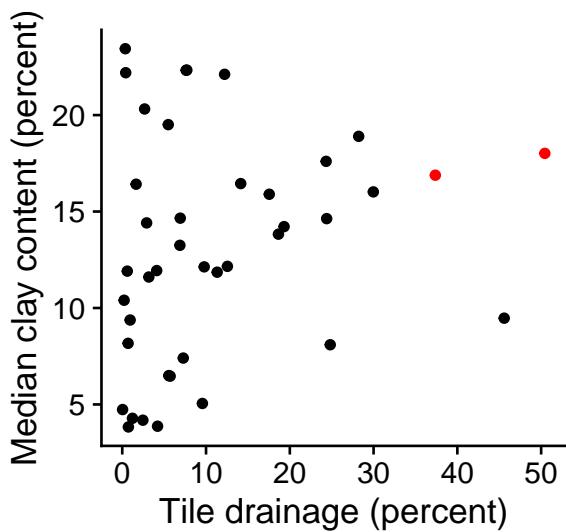
Figure S2.6: Scatterplot showing the median clay content of watersheds in our hydrologic regions plotted against percent tile drainage from Nakagaki and Wieczorek (2016). The regions that are highly sensitive to agricultural land-use from Figure 4 are highlighted in red.
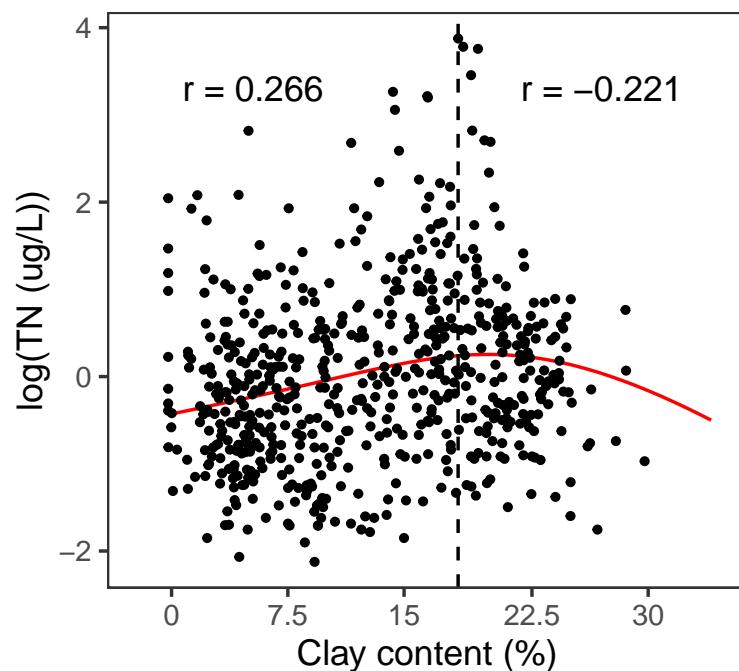
Figure S2.7: Scatterplot showing the non-linear relationship between watershed clay content and lake TN concentration. Vertical dashed line shows transition between a positive and a negative correlation (r) between the two variables. Solid red line shows the fit of a generalized additive model from the `mgcv` R package.

GEOMETRIC MODELS OVERESTIMATE LAKE DEPTH DUE TO IMPERFECT SLOPE MEASUREMENT

Stachelek, J., Hanly, P. J., and Soranno, P. A., In Prep. Geometric models overestimate lake depth due to imperfect slope measurement. *to be submitted to Water Resources Research*

3.1 Abstract

Lake depth is a critical characteristic that influences many important ecological processes in lakes. Unfortunately, lake depth measurements are labor-intensive to gather and are only available for a small fraction of lakes globally. Therefore, scientists have tried to predict lake depth from characteristics that are easily obtained for all lakes such as lake surface area or the slope of the land surrounding a lake. One approach for predicting lake depth simulates lake basins using a geometric model where nearshore land slope is assumed to be a representative proxy for in-lake slope and the distance to the center of the lake is assumed to be a representative proxy for the distance to the deepest point of the lake. However, these assumptions have rarely been tested in a broad range of lakes. We used bathymetry data from 5,000 lakes and reservoirs to test these assumptions and to examine whether differences in lake type or shape influences depth prediction error. We found that nearshore land slope was not a representative proxy of in-lake slope and using it for prediction increases prediction error substantially relative to models using true in-lake slope. We also found that models using nearshore land slope as a proxy systematically overpredicts lake depth in concave lakes (i.e.~bowl-shaped; up to 18% of lakes in the study population) and reservoir lakes (up to 30% of lakes in the study population), suggesting caution in using geometric models for depth prediction in unsampled lakes.

3.2 Introduction

Lake depth is an important factor controlling lake physics, chemistry, and biota. Deeper lakes generally have higher water clarity and less complete mixing compared to shallow lakes (Fee et al., 1996; Read et al., 2014). These differences are reflected in variation among lakes in terms of biological productivity (Qin et al., 2020) and rates of greenhouse gas production (Li et al., 2020). However, because measured depth data is only available for a small fraction (~25%) of all lakes, our ability to understand and predict depth-dependent processes is limited. The importance of lake depth, coupled with its limited availability, has led to numerous attempts to predict depth using measures available for all lakes such as lake surface area or the nearshore slope of the land surrounding a lake (Heathcote et al., 2015; Oliver et al., 2016; Sobek, 2011). Given the limited prediction accuracy of prior attempts ($\pm$ 6-7 m), studies have explored strategies such as employing more diverse covariates (Oliver et al., 2016), varying lake buffer sizes (Heathcote et al., 2015), or estimating hidden groupings (e.g. fitting different models for distinct size classes) among lakes (Sobek, 2011).

One approach for predicting lake depth involves using a geometric model that assumes lake basins correspond to an idealized shape such as a cone, bowl, or an elliptic sinusoid (Hollister et al., 2011; Neumann, 1959; Yigzaw et al., 2018). All such geometric models for lake depth prediction involve implicit assumptions about the terms of geometric formulae. In the simplest case, where lakes basins are treated as cones (Eq 3.1, Figure 3.1), two assumptions are required to make depth predictions for all lakes: 1) that nearshore land slope is a representative proxy for in-lake slope and 2) that the distance to the center of the lake is a representative proxy for the distance to the deepest point of the lake (Figure 3.1). This cone model imposes the following fixed (i.e. geometric) relationship between slope and horizontal distance:

$$depth_{geometric} = tan(slope) * distance \tag{3.1}$$
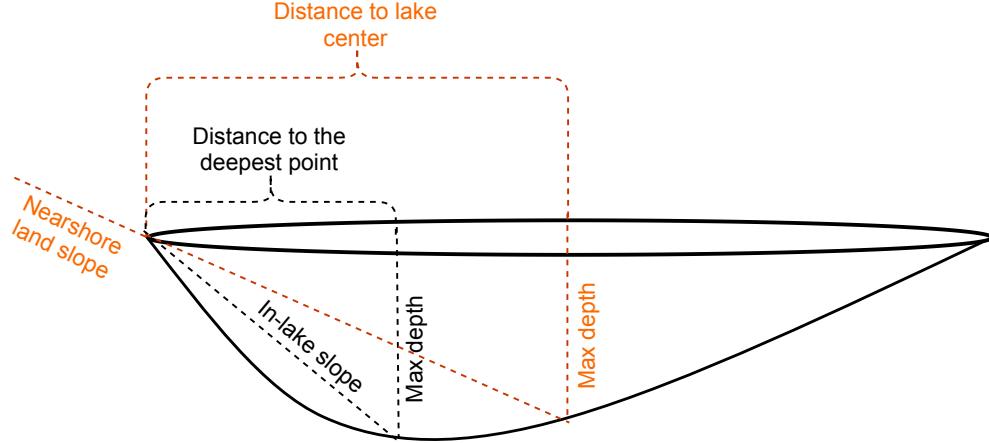
Figure 3.1: Diagram showing the relations between true (black) and proxy (orange) metrics of lake geometry. Geometric depth calculated via Equation 3.1 requires a single distance and slope metric.

where the product of slope and horizontal distance yields an exact geometric depth estimate ($depth_{geometric}$).

The assumptions of the cone model (as well as other geometric models) can be tested by comparing proxy measures of lake geometry against corresponding "true" (i.e. in-lake) values derived from bathymetric maps and by evaluating how lake cross-section shapes differ from that of an idealized cone (Johansson et al., 2007). For instance, lake cross-section shapes have been shown to vary from narrow "convex" forms to outstretched "concave" forms (Hakanson, 1977). Because tests of geometric model assumptions require bathymetric map data, which is only available for about 15% of all lakes, existing evidence may not be applicable to all lakes. The few studies that have tested these assumptions have mostly been limited to individual studies of very large (> 500 ha) lakes or studies on small numbers (< 100) of lakes (Johansson et al., 2007). Studies focused specifically on reservoirs (as opposed to the more typical case where reservoirs and natural lakes are combined), have been even more restricted to that of extremely large lakes > 1000 ha (Lehner et al., 2011; Messager et al., 2016).

As a result of this limited testing, we lack knowledge on both the predictive performance of geometric models, the effect of proxies on depth prediction, and whether

depth predictions are more sensitive to measurement errors in the horizontal dimension (i.e. distance to the deepest point of the lake) or measurement errors in the vertical dimension (i.e. in-lake slope). Additionally it is unclear whether model prediction error is related to differences in lake type such those with different cross-section shapes or those classified as reservoirs versus natural lakes.

Given the knowledge gaps identified above, we asked three research questions: 1) How representative is nearshore land slope of in-lake slope; and how representative is the distance to the center of a lake compared to the distance to the deepest point of a lake? 2) How does the use of proxies for lake geometry affect lake depth prediction error? 3) How does lake cross-section shape (i.e. concave versus convex) and lake type (i.e. natural lake vs reservoir) affect depth prediction error? To answer these questions, we extracted maximum depth (hereafter referred to as "observed maximum depth"), in-lake slope, cross-section shape (i.e., concave versus convex), and distance to the deepest point, of approximately 5,000 lakes in the Northeastern and Midwestern US from bathymetric map data and supplemented this data with classification estimates of whether lakes are reservoirs or natural lakes. First, we examined whether measures of lake geometry (in-lake slope and distance to the deepest point of lakes) were related to geometry proxies (nearshore land slope and distance to lake centers). Next, we computed geometric depth estimates (Equation 3.1) and prediction "offsets" to these estimates using the random forest algorithm (Equation 3.2). Covariates used in offset modeling included a variety of lake, watershed, and hydrologic subbasin measures that are available for all lakes (Table 3.1). We examined differences in overall prediction error corresponding to different inputs to geometric depth as well as differences in the relative distribution of prediction error in lakes with different characteristics.

Given that, by definition, the distance proxy (distance to the center of the lake) must always be greater or equal to the true distance value (distance to the deepest point of the lake), we expect that the use of this proxy will lead to overestimation of lake depth (Figure 3.1). Furthermore we expect to see greater overestimation error in reservoirs as compared
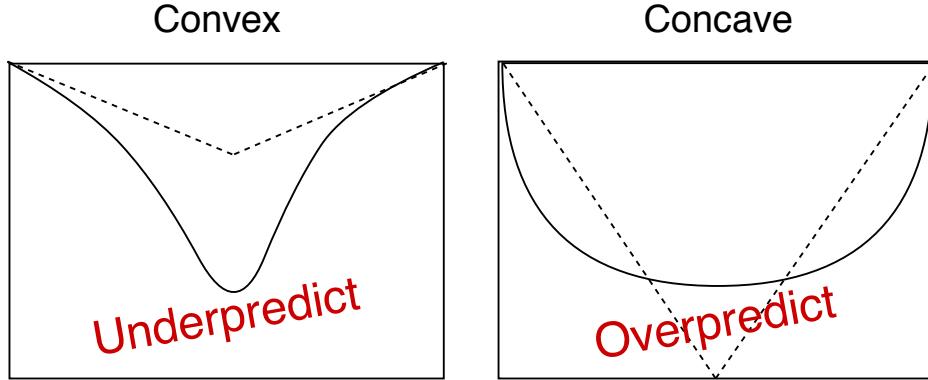
Figure 3.2: Diagram showing our expectation that slope-based models of lake depth will under predict true depth in convex lakes (left) and over predict true depth in concave lakes (right). Dashed lines represent extrapolated nearshore land slope while solid lines represent the lake bottom.

to natural lakes because many reservoirs are known to be drowned river valleys where the deepest point is close to the edge at the end of the reservoir (i.e. next to the dam) rather than in the center of the reservoir (Lanza and Silvey, 1985). In a similar fashion, we expect to see overestimation error associated with using a nearshore land slope proxy in lakes with differing cross-section shape (Figure 3.2) such that the depth of U-shaped (i.e. concave) lakes will be overpredicted whereas the depth of V-shaped (i.e. convex) lakes will be underpredicted. Finally, we expect that overall depth prediction will be strongly related to lake area and hydrologic subbasin variables as these measures have been influential in prior studies (Oliver et al., 2016).

By testing these expectations, we can establish whether barriers to increased depth prediction accuracy lie in lack of correspondence between true and proxy measures of lake geometry or in hidden grouping (such as lake cross-section shape or reservoir status). This information could help direct future research efforts to focus on particular dimensions of lake geometry (i.e. horizontal versus vertical) or to stratify model predictions based on specific lake types and cross-section shapes. Ultimately, achieving increased depth prediction accuracy would allow for more precise estimates of depth-dependent biotic and chemical processes across broad spatial extents.

3.3 Methods

3.3.1 Data description

We compiled bathymetry data on approximately 5,000 lakes in the Northeastern and Midwestern US from nine official state databases (Figure S3.1). The original data came in a variety of formats including pre-interpolated rasters (Minnesota), contour lines (Nebraska, Michigan, Massachusetts, Kansas, Iowa), contour polygons (New Hampshire, Connecticut), or point depth soundings (Maine). For the Minnesota data, we simply clipped the raster for each lake to its outline. For data from the remaining states, we processed each lake by converting its original representation to a point layer (if necessary), rasterizing these points, and creating an interpolated bathymetry "surface" using a simple moving window average in the `raster` R package (Hijmans, 2019). The size of the moving window was adjusted iteratively to ensure that each bathymetry raster contained no missing data.

All lake bathymetry was specifically calculated relative to high-resolution (1:24,000 scale) NHD (USGS, 2019) waterbodies such that source data and bathymetry surface outputs were clipped to the area of each lake polygon. We restricted the lakes in our study to those with an area of at least 4 ha and a maximum depth of at least 0.3 m (1 ft). The purpose of these restrictions was to ensure that lakes had enough contours (or points, or polygons) to generate adequately smooth interpolations with which to calculate in-lake geometry metrics.

We used our generated bathymetry surfaces to find the location of the deepest point in the lake and we resolved ties by choosing the deepest point that was closest to the center of the lake. We calculated the center of the lake not as its centroid but rather by finding the point farthest from the lake shoreline (i.e. its "visual distance to lake center"). For these calculations, we used the `polylabelr` R package (Larsson, 2019), which interfaces with the Mapbox pole of inaccessibility algorithm (Agafonkin, 2019). We calculated in-lake slope as maximum lake depth divided by the distance to the deepest point and we calculated nearshore land slope for each lake by computing the slope within a 100-m buffer using data

Table 3.1: Summary of lake characteristics for the present study (and for lakes in the contiguous United States from Stachelek et al. (In prep)). Predictor variables for computing random forest offsets (Eq 3.2) are printed in bold face. Dashes (-) indicate an identical sample size among this study and that of the contiguous United States.

Variable	Median	Q25	Q75	n
Max depth (m)	8.2 (7)	4.6 (3.7)	14 (12)	4850 (17700)
Elevation (m)	300 (340)	180 (210)	400 (460)	4850 (17700)
Area (ha)	55 (33)	21 (11)	140 (100)	4850 (17700)
Island area (ha)	0 (0)	0 (0)	0.18 (0.076)	4850 (17700)
Perimeter (m)	4400 (3500)	2500 (1800)	8100 (7300)	4850 (17700)
Shoreline development	1.7 (1.7)	1.4 (1.4)	2.1 (2.2)	4850 (17700)
Watershed-lake ratio	7.8 (10)	3.9 (4.4)	17 (28)	4850 (17700)
Distance to deepest point (m)	180 (-)	110 (-)	290 (-)	4850 (-)
Distance to lake center (m)	240 (-)	160 (-)	380 (-)	4850 (-)
In-lake slope (m/m)	0.046 (-)	0.024 (-)	0.079 (-)	4850 (-)
Nearshore land slope (m/m)	0.077 (-)	0.051 (-)	0.11 (-)	4850 (-)

from a high resolution digital elevation model (~15x15m grain) accessed using the `elevatr` R package (Hollister and Shah, 2017) and computed using the terrain function in the `raster` R package (Hijmans, 2019).

We categorized lakes based on their cross-section shape and reservoir class. For cross-section shape, we categorized lakes as either convex or concave following the method of Hakanson (1977) by computing normalized lake depth-area relationships (i.e. hypsographic curves) and assigning class membership based on whether a lake's curve falls above or below that of a simple straight-sided cone (Figure S3.2). We further classified lakes using data from Polus et al. (In prep), which uses the output of a machine learning algorithm to assign a probability to each lake as to whether it is a reservoir or a natural lake. For our purposes, we determined a lake to be a reservoir if the classification probability was 0.75 or greater. The Polus et al. (In prep) data product defines reservoirs as any permanent waterbody that has a water control structure likely to significantly impact flow or pool water, beyond simply controlling water level. It makes no distinction between different dam types or dam heights.

Covariates used in random forest modeling (Table 3.1, Equation 3.2) for lake elevation, area, island area, perimeter, shoreline development, watershed to lake area ratio,

and hydrologic subbasin (i.e. HUC4s), were obtained from the LAGOS-US LOCUS database (Smith et al., In prep). One such measure, that of shoreline development, is a measure of lake perimeter shape defined as:

$$shoreline_{devel} = perimeter/(2 * \sqrt{(\pi * waterarea * 10000)}) \qquad (3.2)$$

where sinuous lakes have larger values of shoreline development and circular lakes have smaller values of shoreline development. Watershed to lake area ratio is an approximation of water residence time and is defined as watershed area divided by lake area.

3.3.2 Proxy evaluation

We conducted a qualitative assessment of whether or not proxy measures of lake geometry are representative of their true values by visual inspection (i.e. plotting each proxy measure against its corresponding true value) and by computing coefficients of determination (R^2). We further tested proxy measures by examining their effect on lake depth prediction error. Our approach involved several steps. In the first step, we computed a geometric estimate of lake depth using only geometry information ($depth_{geometric}$, Equation 3.1). In the second step, we fit a random forest model to predict observed (i.e. true) depth as a function of geometric depth along with several covariates available for all lakes (Table 3.1). The purpose of this random forest "offset" modeling was to more rigorously test our expectations regarding prediction error among different formulations of $depth_{geometric}$ and among different lake types. Each of these steps were executed iteratively for each combination of true and proxy values of slope and distance (Table 3.2).

3.3.3 Model description

3.3.3.1 Geometric model

We used a geometric model of lakes where basins are treated as cones with a fixed relationship between slope and distance (Equation 3.1). Note that Equation 3.1 is a geometric formula and has no intercept or "coefficients" and it produces a perfect depth estimate given true values of slope and distance. To use this model to predict the depth of all lakes, there is a necessary assumption that proxy slope and distance measures, which are available for all lakes, are representative of true slope and distance (Figure 3.1).

3.3.3.2 Random forest models

Nearly all prior studies predicting lake depth using geometric models include a statistical or machine learning model "layer" or "offset" to boost predictive accuracy. For our purposes, this offset modeling enabled us to test our expectations that prediction error would be different among different formulations of $depth_{geometric}$ and among different lake types. It also facilitated direct comparison against prior models of lake depth including those that are non-geometric. We generated an "offset" to geometric depth (sensu Hollister et al., 2011) using the random forest algorithm and the `ranger` R package (Wright and Ziegler, 2017) to predict observed maximum depth as a function of covariates including geometric maximum depth (from Equation 3.1) along with the lake elevation, area, perimeter, and ratio/index measures listed in Table 1:

$$depth_{observed} \sim depth_{geometric} + covariates \qquad (3.3)$$

Neither cross-section shape nor reservoir class was used as a covariate in any random forest models. We used the random forest algorithm because it makes no assumptions about the distribution of model residuals, allows for non-linearity, and is insensitive to interactions (i.e. multicollinearity) among covariates (Prasad et al., 2006).

3.3.3.3 Model comparisons

We tested model sensitivity to slope and distance proxies by generating multiple "geometric maximum depth" estimates from 3 different model runs using each of the possible metric combinations for Equation 3.1 (true slope - proxy distance, proxy slope - true distance, proxy slope - proxy distance). Prior to entry into Equation 3.1, we standardized proxy distances to have the same numeric range as their true counterpart. The purpose of this standardization was to prevent lakes with extremely long proxy distances from having an outsized impact on model evaluation metrics.

3.3.3.4 Model evaluations

We evaluated model fit and prediction error using root-mean-square error (RMSE) and coefficient of determination (R^2) metrics on a holdout set containing 25% of all lakes. We evaluated the residuals (residual = observed - predicted) of each model relative to lake cross-section shape and reservoir classes to determine whether depth is consistently over or under predicted for some lake types relative to others.

3.4 Results

Lakes belonging to each cross-section shape and reservoir class were not evenly distributed across our study area (Figure S3.1). For example, concave lakes were nearly absent from Michigan whereas Maine lakes had an overabundance of lakes categorized as neither concave nor convex. Lakes in the southern portions of our study area tended to be classified as reservoirs whereas lakes in the northern portions of our study area were a more even mix between reservoirs and natural lakes (Figure S3.1). Approximately 18%, 80%, and 2% of lakes were classified as having a concave, convex, or neither shape respectively whereas approximately 30% and 70% of lakes were classified as being a reservoir or a natural lake.

Although proxy distance to lake center was often much larger in magnitude compared to the true distance to the deepest point of lakes', they were strongly related ($R^2 = 0.8$).
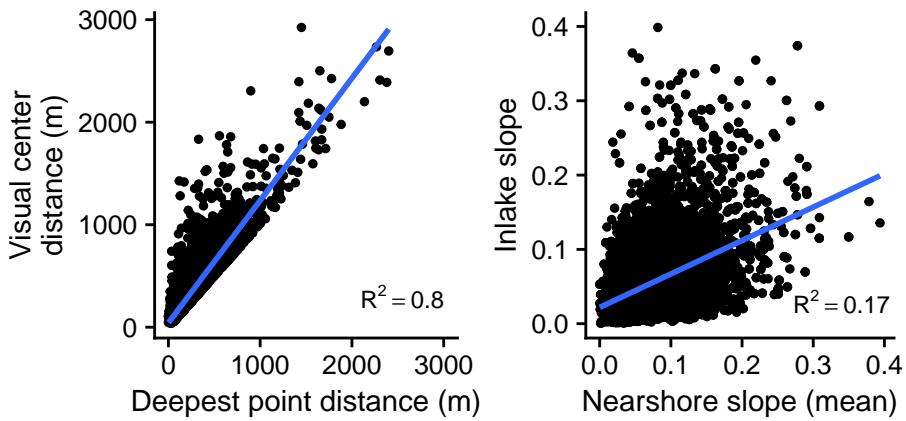
Figure 3.3: Comparison among proxy and true values of lake geometry for A) distance to deepest point versus distance distance to lake center and B) nearshore land slope versus in-lake slope. A best-fit line and coefficient of determination is shown to illustrate representativeness.

In contrast, proxy nearshore land slope and true in-lake slope were more weakly related ($R^2 = 0.17$). For slope measures, most lakes had lower magnitude (i.e. shallower) nearshore land slope compared to true in-lake slope (Figure 3.3). Taken together, these results suggest that proxy distance to the center of lakes is representative of true distance to the deepest point of lakes while proxy nearshore land slope is not representative of true in-lake slope. In addition to overall differences between slope and distance measures, we found differences in these relationships among lake shape classes. For example, in-lake slope and distance to the deepest point of the lake metrics were consistently larger in magnitude for convex lakes as compared to concave lakes (Figure S3.3). However, there were not similar differences among slope and distance metrics for natural lakes versus reservoirs (Figure S3.3).

The use of proxy nearshore land slope had a larger effect on model fit and prediction error than the use of proxy distance to lake center (Table 3.2). More specifically, the true slope - proxy distance model had a better fit ($R^2 = 0.73$) and lower prediction error (RMSE = 4.23m) compared to the proxy slope - true distance model ($R^2 = 0.26$, RMSE = 6.87m). Furthermore, analysis of model residuals showed overestimation of lake depth for concave lakes when models included a proxy slope measure (Figure 3.4). We observed similar but smaller overestimation depending on if a lake was classified as a reservoir rather than a

Table 3.2: Model fit and predictive accuracy metrics (RMSE = root mean square error, R^2 = coefficient of determination) for all combinations of true (in-lake slope, distance to the deepest point of the lake) and proxy (nearshore land slope, distance to lake center) metrics.

slope	distance	RMSE	R^2
true	true	-	-
true	proxy	4.2 m	0.73
proxy	true	6.9 m	0.26
proxy	proxy	6.6 m	0.31
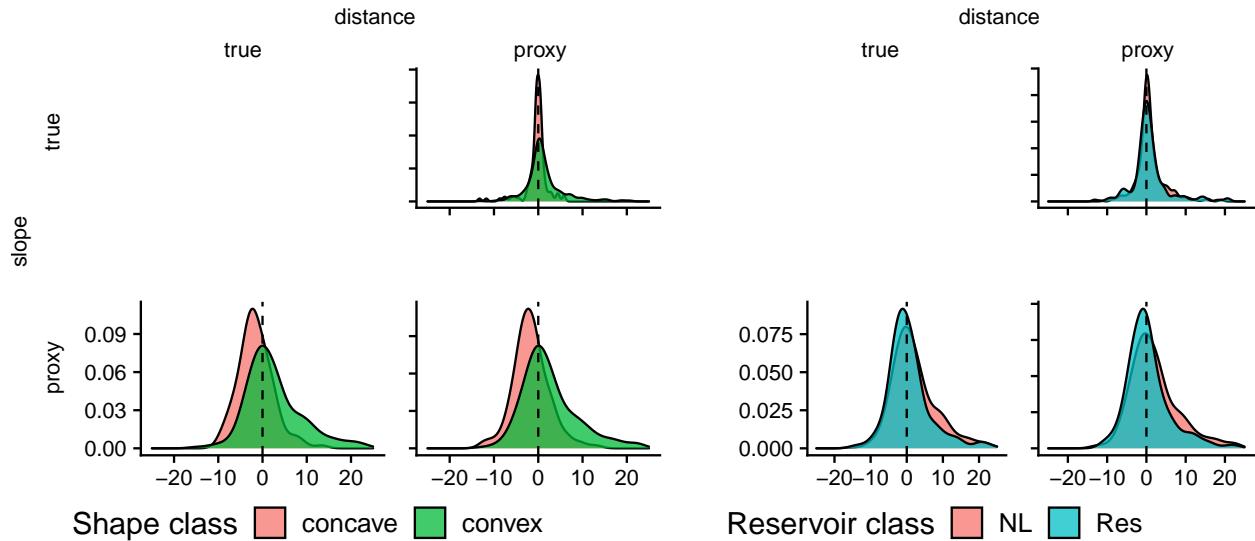
Figure 3.4: Depth model residuals (residual = observed - predicted) in meters by cross-section shape and reservoir class indicating overprediction of concave and reservoir lakes.

natural lake (Figure 3.4).

The most important covariates in offset models were those relating to spatial location, lake area, and perimeter (Figure S3.4). Conversely, watershed metrics and lake elevation had little contribution to random forest model fit (Figure S3.4). The spatial location (i.e. HUC4) covariate was notably less importance in the true slope model compared to the two proxy slope models. Model importance calculations indicated that omitting a geometric max depth measure results in a 130%, 60%, or 50% increase in mean square error depending on the formulation of geometric max depth in Eq 3.1 (Figure S3.4).

3.5 Discussion

Our tests of geometric lake depth models show that specific proxy measures of lake geometry are not representative of true measures of lake geometry across a broad array of lakes. Using a cone model example, we show that nearshore land slope is not representative of in-lake slope. Furthermore, our results indicate that the use of nearshore land slope for prediction results in increased error and systematic overestimation of depth in concave and reservoir lakes. Although our analysis was limited to lakes with available bathymetry data, these lakes did not have characteristics that differed from that of the overall lake population (Figure S3.5, S3.7). This lack of difference suggests that our results are likely to be broadly applicable to all lakes although there is a possibility that there is some hidden bias not explored for in our analyses.

3.5.1 Representativeness of proxy measures of lake geometry

In comparing among lake geometry measures, our analysis suggests that proxy distance to lake center is representative of true distance to the deepest point of the lakes but that proxy nearshore land slope is not representative of true in-lake slope. A simple indication of this non representativeness is that proxy nearshore land slope was often (in $> 74\%$ of cases) steeper than true in-lake slope. This finding is consistent with Heathcote et al. (2015) whos results suggest that in-lake slopes are shallower compared to the surrounding land. The shallow nature of in-lake slopes is likley a function of erosion and sediment transport processes (Håkanson, 1981; Johansson et al., 2007).

One surprising finding with respect to the relationship between true and proxy geometry measures when examined by lake class was the fact that there was no greater difference between proxy and true distances in reservoirs compared to natural lakes. This is contrary to the idea that most reservoirs are drowned river valleys where the deepest point is close to the edge at the end of the reservoir (i.e. next to the dam) rather than

in the center of the reservoir (Lanza and Silvey, 1985). One possible explanation is that the reservoir classification data from Polus et al. (In prep) uses a more general definition of a reservoir (i.e. any permanent waterbody that has a water control structure likely to significantly impact flow or pool water) compared to that of conventional classifications that are tied to specific dam types or dam heights. Another possible explanation is that conventional reservoir classifications are conceptually biased towards more southern areas with few natural lakes (Figure S3.1).

We found other differences among lake geometry measures according to lake cross-section shape. One finding was that convex lakes, when compared to concave lakes, had longer distances to lake centers relative to corresponding distances to the deepest point of lakes. In addition, convex lakes often had steeper in-lake slopes relative to nearshore land slopes as compared to concave lakes. These differences are reflected in the class-wise differences in shoreline development (e.g. shorelines were more sinuous in convex lakes compared to concave lakes, Figure S3.6). It was notable that convex lakes were deeper than concave lakes despite having similar distributions of lake surface area (Figure S3.6). Given the similarity in lake surface area, the underlying cause of these differences is unknown but one possibility is that geometry is tied to the circumstances of lake formation whereby concave lakes were formed as a result of more intense glacial scouring compared to convex lakes (Gorham, 1958). While there is some evidence in support of this idea, namely that there is a geographic hotspot of concave lakes associated with the glaciated "prairie pothole region" (see Hayashi and van der Kamp, 2000), the overall geographic distribution of lake cross-section shapes does not support this idea. Instead of a concentrated area of concave lakes in formerly glaciated regions, there appears to be a fairly even mix of concave and convex lakes distributed amongst the northern (i.e. glaciated) and southern (non-glaciated) portions of our study area (Figure S3.1).

3.5.2 Effects of proxy measures of lake geometry depth prediction error

Models using only proxy variables had prediction error rates (RMSE = 6.6m) of a similar magnitude as that of prior studies (RMSE = 6 - 7.3m) predicting lake depth at broad geographic extents (Hollister et al., 2011; Oliver et al., 2016; Messager et al., 2016). When only a single proxy measure was used there was a difference in model sensitivity depending on if it was a horizontal distance measure or a vertical slope measure. In the case of a true slope and proxy distance combination, models were more accurate ($\pm$ 4.2m) than even the most accurate of prior studies (Hollister et al., 2011; Oliver et al., 2016; Messager et al., 2016). Conversely, models using a proxy slope and true distance combination had prediction error rates ($\pm$ 6.9m) of a similar magnitude as that of the baseline proxy-proxy model ($\pm$ 6.6m). The greater sensitivity of depth predictions to proxy slope measures relative to proxy distance measures may be explained by the fact that proxy slope measures were a more imperfect representation of true in-lake slopes relative to proxy versus true distances. In addition, these results help explain the relatively poor predictive performance of prior non-geometric lake depth models given that they rely heavily on lake area as a predictor (Messager et al., 2016; Oliver et al., 2016; Sobek, 2011) and both horizontal distance measures and vertical slope measures appear to be decoupled from lake area (Figure S3.6).

3.5.3 Effects of lake shape and lake type on depth prediction error

As expected, we found that the maximum depth of concave lakes was systematically overpredicted by a simple geometric model using proxy nearshore land slope. However, contrary to our expectation, we did not observe underprediction of depth in convex lakes. The reason we did not observe underprediction of the depth of convex lakes is likely because geometric depth itself was always greater than observed maximum depth owing to the fact that proxy distance is constrained to be greater than true distance. Although the magnitude of overestimation is likely related to class imbalance in our dataset (i.e. there was a greater number of convex lakes), these results suggest that broad scale estimates of lake depth are overestimated

particularly when those estimates encompass large numbers of lakes with diverse cross-section shapes.

3.5.4 Future research

One of our models (true slope, proxy distance) was more accurate than even the most accurate of prior studies. However, parameterization of this model requires data on bathymetry which is not available for all lakes. We propose that the error rate of this model ($\pm$ 4.2m) be used as an out-of-sample prediction benchmark for future studies such that they should attempt to match it but not expect to exceed it. Because this model requires bathymetry data, this suggests that it may not be possible with current data and models to produce depth predictions for all lakes with error rates below about 6m. To approach our benchmark using data available for all lakes, future studies could explore alternative modeling approaches such as ordinal modeling, which would capture whether or not a lake crosses some important depth threshold, but would not seek to predict a specific depth value. These studies could also explore emerging data types to boost prediction accuracy such as "topobathymetric" products that integrate both topographic and bathymetric data in a seamless fashion rather than treating them as separate entities. This would allow for more robust tests of the representativeness of geometric model inputs. Unfortunately, topobathymetric products are rare, have mostly been limited nearshore marine environments, and as such are not widely available for inland waters (Danielson et al., 2016).

Finally, our findings indicate that geometry measures differ according to lake cross-section shape. This makes it an attractive target for inclusion in depth prediction models. Unfortunately, identifying a lake's cross-section shape requires bathymetry data which is unavailable for most lakes. However, given the conceptual links between cross-section shape, glaciation, and sedimentation (Johansson et al., 2007) it may be advantageous for future studies to compile data on sedimentation to determine if this data can be used to predict cross-section shape and boost depth prediction accuracy.

3.6   Conclusion

To our knowledge, the present study is the largest and most comprehensive test to date of geometric models of lake depth. Using bathymetry data on approximately 5,000 lakes, we show that proxy slope measures are not representative of true in-lake slope and this leads to inaccuracies in predicting the depth of concave and reservoir lakes. These innaccuracies suggest that caution is warranted in using geometric models for depth prediction in unsampled lakes. Despite these apparent biases, overall prediction accuracy was equivalent to that of prior depth prediction studies ($\pm$ 6-7m). Only our models using a true measure of in-lake slope, which is limited in availability to lakes with bathymetry data and where we already know lake depth, had greater accuracy than that of prior studies ($\pm$ 4.2m). Lack of improved prediction accuracy (short of including data that is unavailable for most lakes) suggests that improved prediction may require new types of data or novel analysis techniques.

3.7   Supplemental information
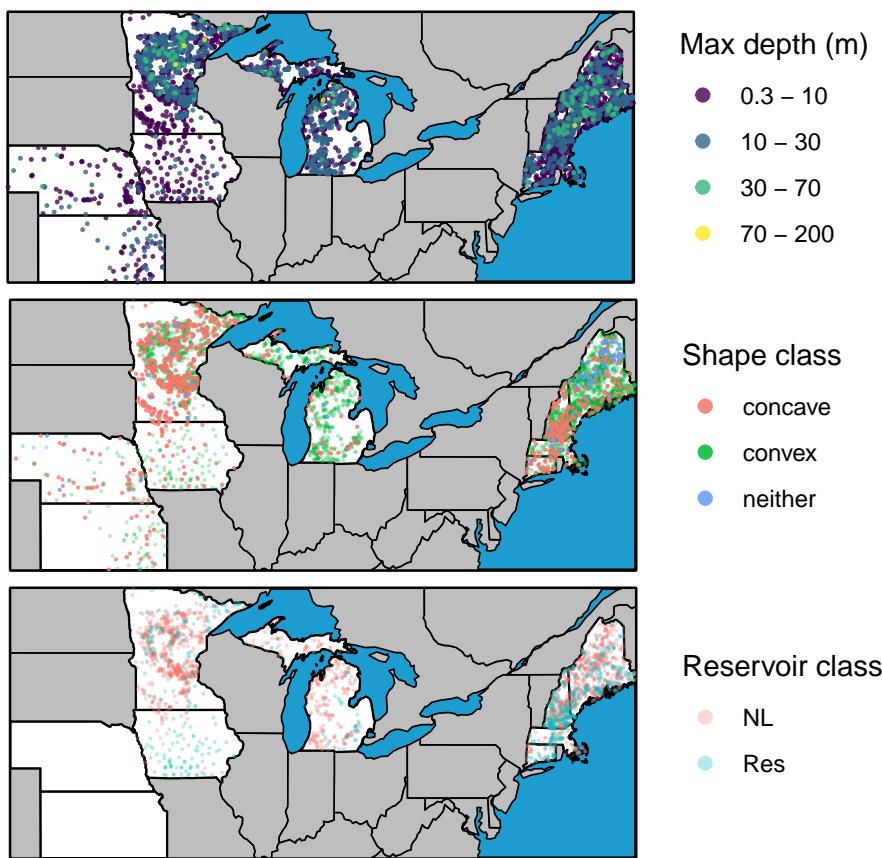
Figure S3.1: Map of study lakes showing A) lake maximum depth measurements, B) cross-section shape class, and C) reservoir classification.
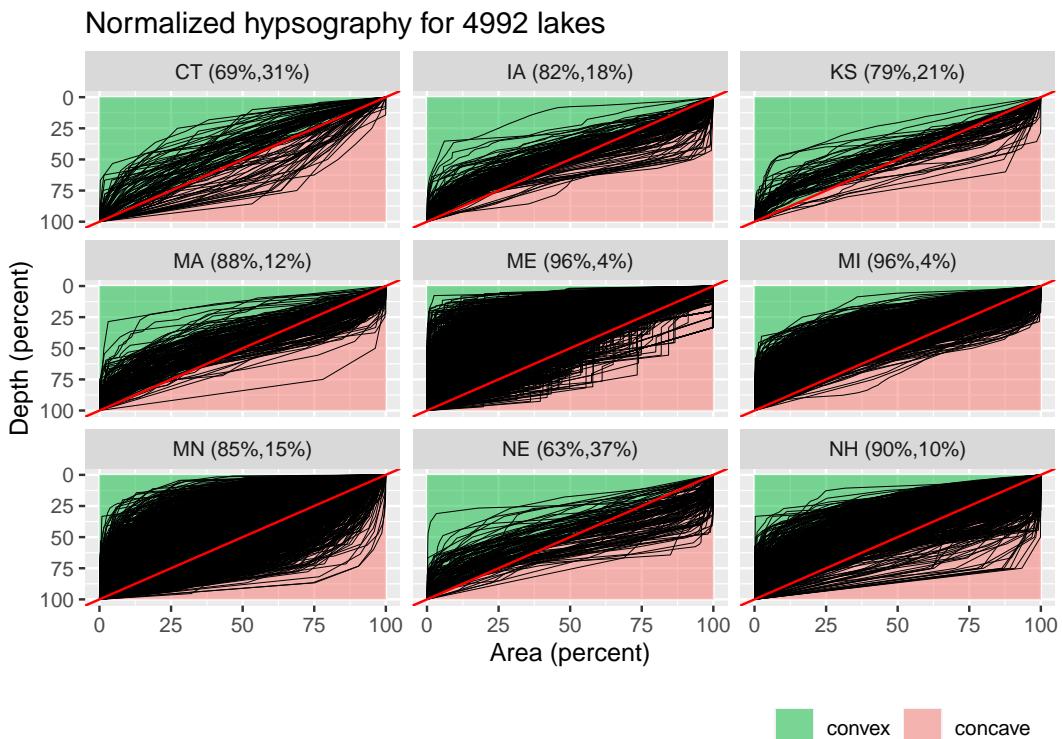
Figure S3.2: Hypsography classification by state. Numbers on panel labels indicate the percentage of lakes in each state with a convex versus a concave cross-section shape.
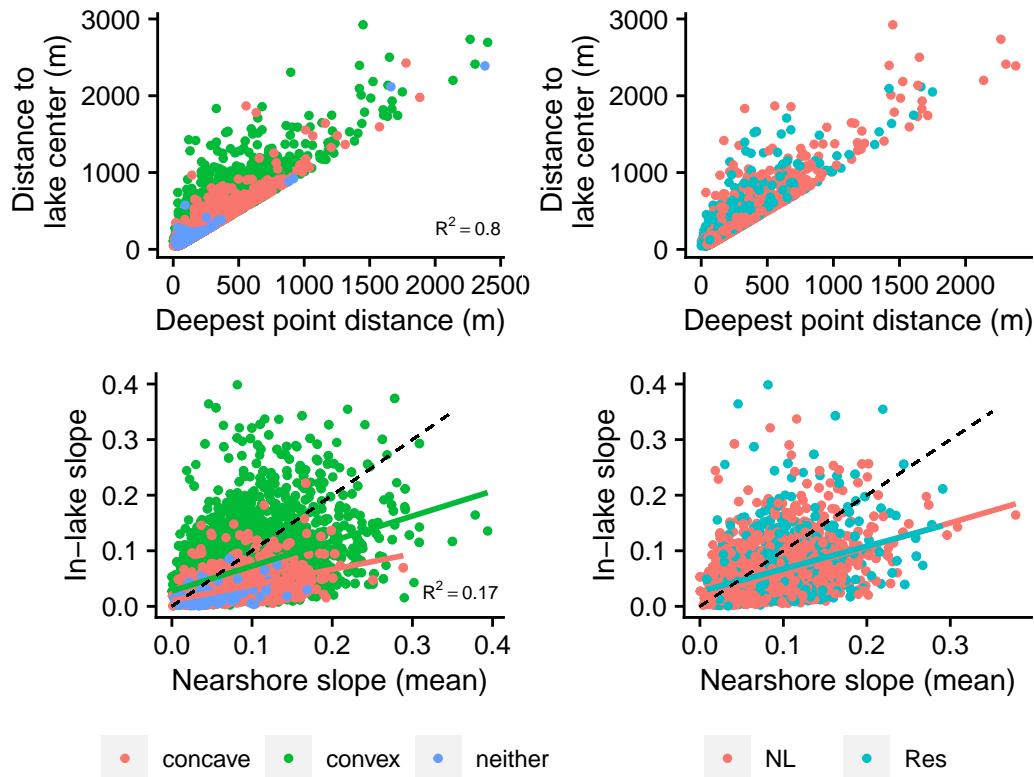
Figure S3.3: Comparison among lake shape and reservoir classes for A-B) distance to deepest point versus distance to lake visual center and C-D) nearshore land slope versus in-lake slope. A dashed 1:1 line is shown for comparison. Cross-section shape and reservoir class plots are not identical because not all lakes had a reservoir classification exceeding a 0.75 probability confidence level.
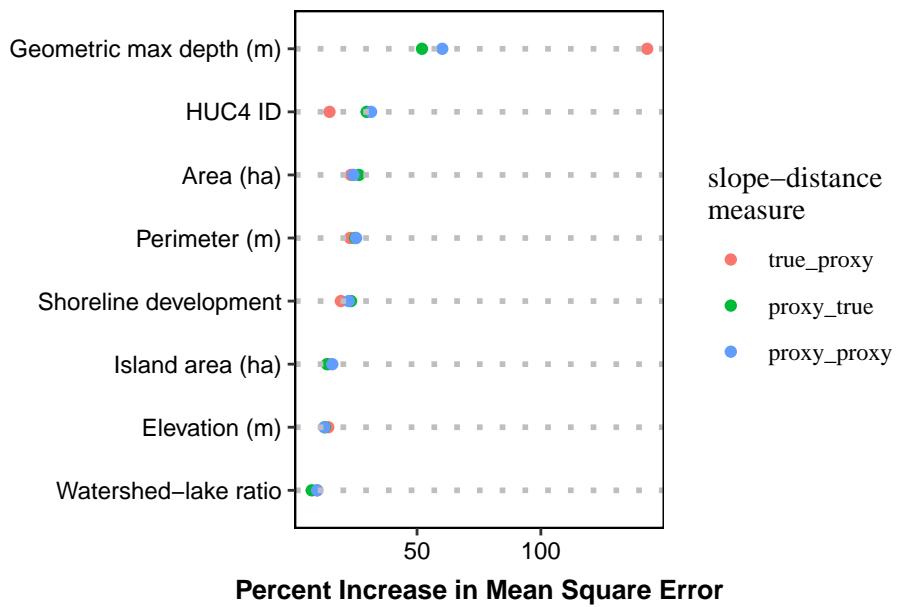
Figure S3.4: Importance plot for random forest variables showing increase in mean square error. Higher values indicate greater importance to model predictions. See Equation 1 for a definition of geometric max depth. HUC4 ID is a 'dummy' variable of geographic (hydrologic subbasin) location.
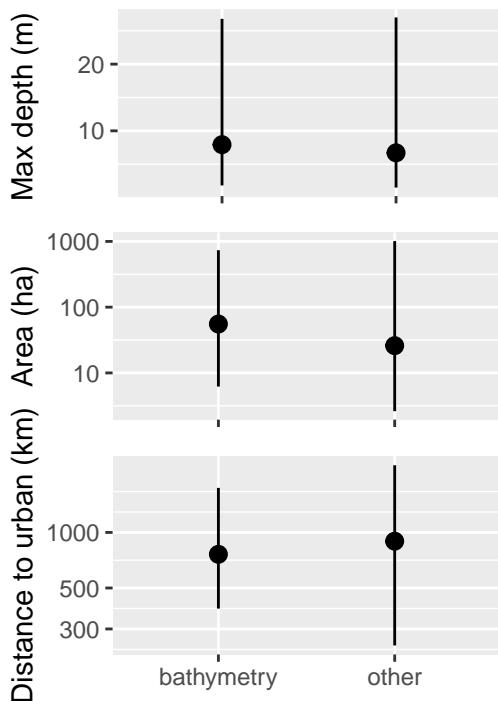
Figure S3.5: Comparison between characteristics of lakes with bathymetry data against lakes with depth from other sources in the LAGOSUS-Depth product (Stachelek et al., In prep). The distance to urban area metric is calculated using data from the 2018 US Census Urban and Rural Classification.
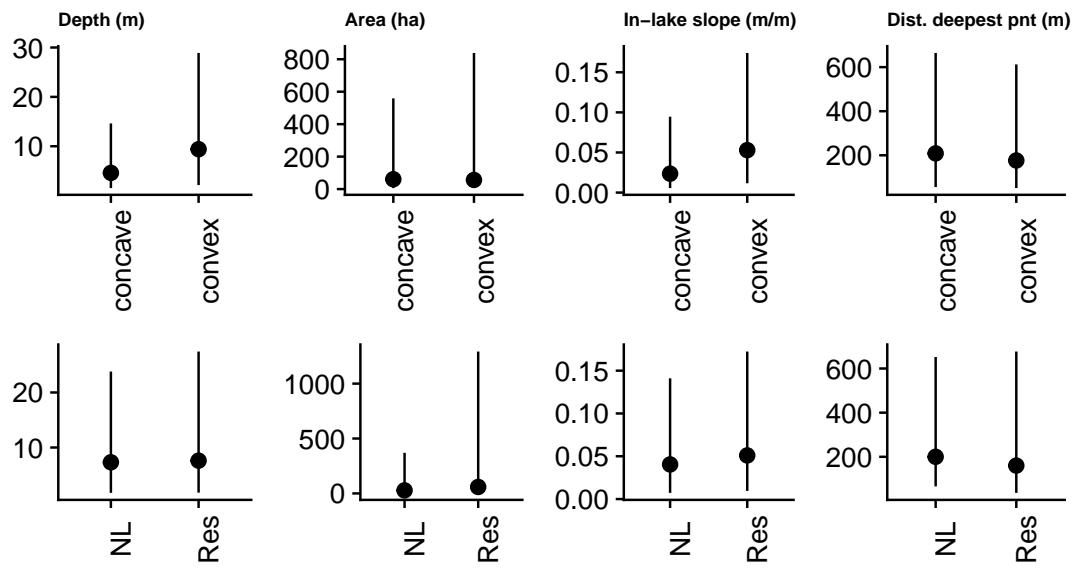
Figure S3.6: Lake characteristics by categorical variables.
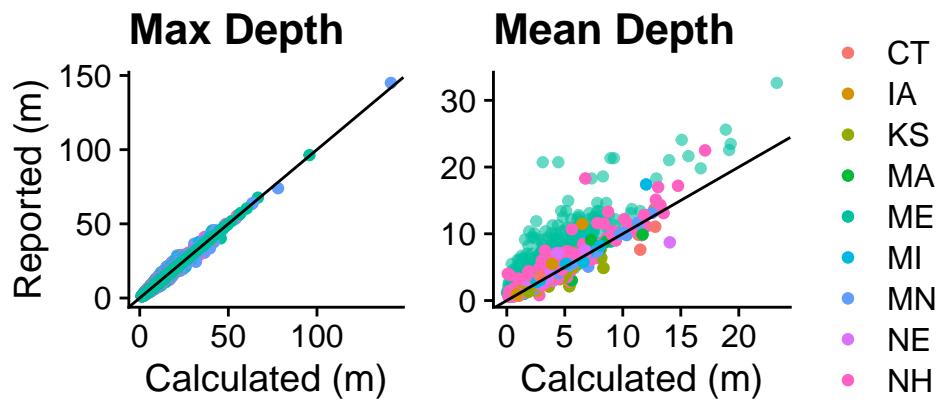
Figure S3.7: Comparison between reported depth and depth extracted from bathymetry surfaces by US State where reported depths come from the LAGOSUS-Depth product (Stachelek et al., In prep). For this figure, no reported depth values originated from the same source as its corresponding bathymetry-derived value.

CONCLUSION

Research Frontiers

Each of the preceding chapters left some analyses unexplored that represent future research frontiers in macroscale lake research. These unexplored analyses include the development of new techniques and assembly of new types of data. In Chapter 1, I used lake connectivity metrics to quantify the effect of connectivity among lakes and streams as well as connectivity of lakes and their terrestrial watersheds on lake P retention. However, "connectivity" is a somewhat nebulous idea that is difficult to quantify with specific metrics. For instance, I found that not all metrics can be mapped onto a "high connectivity" versus "low connectivity" gradient. A potential alternative to this messy mapping of connectivity onto discrete metrics is to compare the drainage pattern of a given stream network against an optimal drainage pattern which has the highest possible connectivity for a given watershed. A recent tool, the `OCNet` R package, may facilitate exactly such calculations as it can calculate the optimal distributions of upstream and downstream lengths, contributing area, and the space-filling attributes of specific watersheds (Carraro et al. 2020).

In Chapter 2, I examined relationships between lake nutrient concentrations and measures of agricultural activity quantified at varying levels of spatial and process-level detail. This effort was limited in the types of agricultural measures that were considered owing to the spatial extent of the study, challenges in data integration of diverse data types, and complicated model interpretation and building. For instance, it would have been great to have included other aspects of agriculture like animal feeding operations. Unfortunately, the limited spatial resolution of this data (only to the county level) makes it difficult to integrate with crop and nutrient input data. In addition, I was unable to explore temporal variation in lake nutrient-agriculture relationships which may have contributed to overall model uncertainty. Unfortunately, few lakes have adequate nutrient time series data to drive

such a model.

In Chapter 3, I tested the assumptions of geometric models of lake depth using bathymetry data. I show that such models are highly sensitive to imperfect proxies of lake slope. Even with measures of true slope derived from bathymetry, model prediction accuracy is still barely accurate enough to distinguish between "shallow" and non-shallow lakes. A fruitful area of future research would be to test different modeling techniques such as ordinal modeling (i.e. shallow versus not shallow) and new topobathymetric data products that integrate topographic and bathymetric data into a single product. This may boost predictive power and avoid changes in resolution and changes in scale that likely introduce uncertainty in model predictions.

Future Directions

Throughout the process of writing this Dissertation I noticed several commonalities amongst each chapter. One commonality was simply the large number of modeling choices that underpin each result. My hope is that the most important of these choices is reflected in the text and the remaining ones are sufficiently documented in the code supplements accompanying each chapter. In some cases, I was fortunate to have the opportunity to explore the sensitivity of my model results to each choice. However, the sheer number of these choices made it intractable to test them all. Choices such as what type of regionalization to use or how and when to aggregate data from one spatial scale to another were not always straightforward. Future exploration of these choices would likely go a long way in strengthening our understanding of spatial patterning of lake characteristics at macroscales and generating exciting new research questions.

A second commonality I found was that models incorporating spatial variation are often superior in terms of predictive performance and model fit relative to non-spatial models. However, I only really interrogated one level of spatial variation, which was variation within and among hydrologic subbasins. The question of why two neighboring lakes might

differ in terms of trophic status despite having nearly identical land-use, nutrient loading, morphometry, and climate remains largely unresolved by my analyses. Answering this question may require the application of new techniques considering spatial autocorrelation within stream networks or new data types derived from the output of fine-scale hydrologic models. Exploring these is likely to be critical for resolving the types of local-scale spatial variation which is so important for the management of freshwater ecosystems.

REFERENCES

REFERENCES

Abell, J. M., Özkundakci, D., Hamilton, D. P., and Miller, S. D., 2011. Relationships between land use and nitrogen and phosphorus in New Zealand lakes. *Marine and Freshwater Research*, 62(2):162. doi:10.1071/MF10180.

Agafonkin, V., 2019. *A JS Library for Finding Optimal Label Position inside a Polygon*. https://github.com/mapbox/polylabel.

Alexander, R. B., Smith, R. A., Schwarz, G. E., Boyer, E. W., Nolan, J. V., and Brakebill, J. W., 2008. Differences in Phosphorus and Nitrogen Delivery to The Gulf of Mexico from the Mississippi River Basin. *Environmental Science & Technology*, 42(3):822–830. doi:10.1021/es0716103.

Allan, J. D., 2004. Landscapes and Riverscapes: The Influence of Land Use on Stream Ecosystems. *Annual Review of Ecology, Evolution, and Systematics*, 35(1):257–284. doi:10.1146/annurev.ecolsys.35.120202.110122.

Anderson, D. R. and Burnham, K. P., 2002. Avoiding Pitfalls When Using Information-Theoretic Methods. *The Journal of Wildlife Management*, 66(3):912. doi:10.2307/3803155.

Arbuckle, J. H. and Downing, G. A., 2001. The influence of watershed land use on lake N:P in a predominantly agricultural landscape. *Limnol. Oceanogr*, 34(2):267–285.

Baker, M. E., Weller, D. E., and Jordan, T. E., 2006. Improved methods for quantifying potential nutrient interception by riparian buffers. *Landscape Ecology*, 21(8):1327–1345. doi:10.1007/s10980-006-0020-0.

Bellmore, R., Compton, J., Brooks, J., Fox, E., Hill, R., Sobota, D., Thornbrugh, D., and Weber, M., 2018. Nitrogen inputs drive nitrogen concentrations in U.S. streams and rivers during summer low flow conditions. *Science of The Total Environment*, 639:1349–1359. doi:10.1016/j.scitotenv.2018.05.008.

Bennett, E. M., Carpenter, S. R., and Caraco, N. F., 2001. Human Impact on Erodable Phosphorus and Eutrophication: A Global Perspective. *BioScience*, 51(3):227. doi:10.1641/0006-3568(2001)051[0227:HIOEPA]2.0.CO;2.

Blann, K. L., Anderson, J. L., Sands, G. R., and Vondracek, B., 2009. Effects of Agricultural Drainage on Aquatic Ecosystems: A Review. *Critical Reviews in Environmental Science and Technology*, 39(11):909–1001. doi:10.1080/10643380801977966.

Boyer, E. W., Goodale, C. L., Jaworski, N. A., and Howarth, R. W. Anthropogenic nitrogen sources and relationships to riverine nitrogen export in the northeastern U.S.A. In Boyer, E. W. and Howarth, R. W., editors, *The Nitrogen Cycle at Regional to Global Scales*, pages 137–169. Springer Netherlands, Dordrecht, 2002.

Brett, M. T. and Benjamin, M. M., 2008. A review and reassessment of lake phosphorus retention and the nutrient loading concept. *Freshwater Biology*, 0(0). doi:10.1111/j.1365-2427.2007.01862.x.

Burcher, C. L., Valett, H. M., and Benfield, E. F., 2007. The land-cover cascade: Relationships coupling land and water. *Ecology*, 88(1):228–242. doi:10.1890/0012-9658(2007)88[228:TLCRCL]2.0.CO;2.

Bürkner, P.-C. et al., 2017. Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1):1–28.

Canham, C. D., Pace, M. L., Weathers, K. C., McNeil, E. W., Bedford, B. L., Murphy, L., and Quinn, S., 2012. Nitrogen deposition and lake nitrogen concentrations: A regional analysis of terrestrial controls and aquatic linkages. *Ecosphere*, 3(7):art66. doi:10.1890/ES12-00090.1.

Capel. A Conceptual Framework for Effectively Anticipating Water-Quality Changes Resulting From Changes in Agricultural Activities. Scientific Investigations Report, 2018.

Capel, P. D. Agriculture—A River Runs Through It—The Connections Between Agriculture and Water Quality. Circular, 2018.

Cardille, J. A., Carpenter, S. R., Coe, M. T., Foley, J. A., Hanson, P. C., Turner, M. G., and Vano, J. A., 2007. Carbon and water cycling in lake-rich landscapes: Landscape connections, lake hydrology, and biogeochemistry. *Journal of Geophysical Research*, 112 (G2). doi:10.1029/2006JG000200.

Carl, G., Levin, S., and Kühn, I., 2018. Spind: An R Package to Account for Spatial Autocorrelation in the Analysis of Lattice Data. *Biodiversity Data Journal*, 6:e20760. doi:10.3897/BDJ.6.e20760.

Carlson, R. E. and Simpson, J., 1996. A coordinator's guide to volunteer lake monitoring methods. *North American Lake Management Society*, 96:305.

Carpenter, S. R., Caraco, N. F., Correll, D. L., Howarth, R. W., Sharpley, A. N., and Smith, V. H., 1998. Nonpoint Pollution of Surface Waters with Phosphorus and Nitrogen. *Ecological Applications*, 8(3):559. doi:10.2307/2641247.

Carvalho, C. M., Polson, N. G., and Scott, J. G., 2010. The horseshoe estimator for sparse signals. *Biometrika*, 97(2):465–480. doi:10.1093/biomet/asq017.

Cha, Y., Stow, C. A., and Bernhardt, E. S., 2013. Impacts of dreissenid mussel invasions on chlorophyll and total phosphorus in 25 lakes in the USA: *Dreissenid invasion impacts on Chl and TP*. *Freshwater Biology*, 58(1):192–206. doi:10.1111/fwb.12050.

Chapra, S. C., 2008. *Surface Water-Quality Modeling*. Waveland press.

Chen, L. and Lisic, J., 2018. *Tools to Download and Work with USDA Cropscape Data*. https://cran.r-project.org/package=cdlTools.

Cheng, V., Arhonditsis, G. B., and Brett, M. T., 2010. A revaluation of lake-phosphorus loading models using a Bayesian hierarchical framework. *Ecological Research*, 25(1):59–76. doi:10.1007/s11284-009-0630-5.

Cherry, K., Shepherd, M., Withers, P., and Mooney, S., 2008. Assessing the effectiveness of actions to mitigate nutrient loss from agriculture: A review of methods. *Science of The Total Environment*, 406(1-2):1–23. doi:10.1016/j.scitotenv.2008.07.015.

Cobourn, K. M., Carey, C. C., Boyle, K. J., Duffy, C., Dugan, H. A., Farrell, K. J., Fitchett, L., Hanson, P. C., Hart, J. A., Henson, V. R., Hetherington, A. L., Kemanian, A. R., Rudstam, L. G., Shu, L., Soranno, P. A., Sorice, M. G., Stachelek, J., Ward, N. K., Weathers, K. C., Weng, W., and Zhang, Y., 2018. From concept to practice to policy: Modeling coupled natural and human systems in lake catchments. *Ecosphere*, 9(5):e02209. doi:10.1002/ecs2.2209.

Collins, S. M., Yuan, S., Tan, P. N., Oliver, S. K., Lapierre, J. F., Cheruvelil, K. S., Fergus, C. E., Skaff, N. K., Stachelek, J., Wagner, T., and Soranno, P. A., 2019. Winter Precipitation and Summer Temperature Predict Lake Water Quality at Macroscales. *Water Resources Research*, 55(4):2708–2721. doi:10.1029/2018WR023088.

Collins, S. M., Oliver, S. K., Lapierre, J.-F., Stanley, E. H., Jones, J. R., Wagner, T., and Soranno, P. A., 2017. Lake nutrient stoichiometry is less predictable than nutrient concentrations at regional and sub-continental scales. *Ecological Applications*, 27(5):1529–1540. doi:10.1002/eap.1545.

Covino, T., 2017. Hydrologic connectivity as a framework for understanding biogeochemical flux through watersheds and along fluvial networks. *Geomorphology*, 277:133–144. doi:10.1016/j.geomorph.2016.09.030.

Cushing, C. E., Minshall, G. W., and Newbold, J. D., 1993. Transport dynamics of fine particulate organic matter in two Idaho streams. *Limnology and Oceanography*, 38(6): 1101–1115. doi:10.4319/lo.1993.38.6.1101.

Daniel, F. B., Griffith, M. B., and Troyer, M. E., 2010. Influences of Spatial Scale and Soil Permeability on Relationships Between Land Cover and Baseflow Stream Nutrient Concentrations. *Environmental Management*, 45(2):336–350. doi:10.1007/s00267-009-9401-x.

Danielson, J. J., Poppenga, S. K., Brock, J. C., Evans, G. A., Tyler, D. J., Gesch, D. B., Thatcher, C. A., and Barras, J. A., 2016. Topobathymetric Elevation Model Development using a New Methodology: Coastal National Elevation Database. *Journal of Coastal Research*, 76:75–89. doi:10.2112/SI76-008.

Diebel, M. W., Maxted, J. T., Robertson, D. M., Han, S., and Vander Zanden, M. J., 2009. Landscape Planning for Agricultural Nonpoint Source Pollution Reduction III: Assessing Phosphorus and Sediment Reduction Potential. *Environmental Management*, 43(1):69–83. doi:10.1007/s00267-008-9139-x.

Dillon, P. J. and Molot, L. A., 1996. Long-term Phosphorus Budgets And An Examination Of A Steady-state Mass Balance Model For Central Ontario Lakes. *Water Research*, 30 (10):8.

Djodjic, F. and Markensten, H., 2018. From single fields to river basins: Identification of critical source areas for erosion and phosphorus losses at high resolution. *Ambio*. doi:10.1007/s13280-018-1134-8.

Downing, J. A., Watson, S. B., and McCauley, E., 2001. Predicting Cyanobacteria dominance in lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, 58(10):1905–1908. doi:10.1139/cjfas-58-10-1905.

Ellison, M. E. and Brett, M. T., 2006. Particulate phosphorus bioavailability as a function of stream flow and land cover. *Water Research*, 40(6):1258–1268. doi:10.1016/j.watres.2006.01.016.

Fee, E. J., Hecky, R. E., Kasian, S. E. M., and Cruikshank, D. R., 1996. Effects of lake size, water clarity, and climatic variability on mixing depths in Canadian Shield lakes. *Limnology and Oceanography*, 41(5):912–920.

Fergus, C. E., Lapierre, J.-F., Oliver, S. K., Skaff, N. K., Cheruvelil, K. S., Webster, K., Scott, C., and Soranno, P., 2017. The freshwater landscape: Lake, wetland, and stream abundance and connectivity at macroscales. *Ecosphere*, 8(8):e01911. doi:10.1002/ecs2.1911.

Filstrup, C. T., Wagner, T., Oliver, S. K., Stow, C. A., Webster, K. E., Stanley, E. H., and Downing, J. A., 2018. Evidence for regional nitrogen stress on chlorophyll *a* in lakes across large landscape and climate gradients: N stress on lake phytoplankton. *Limnology and Oceanography*, 63(S1):S324–S339. doi:10.1002/lno.10742.

Fraterrigo, J. M. and Downing, J. A., 2008. The Influence of Land Use on Lake Nutrients Varies with Watershed Transport Capacity. *Ecosystems*, 11(7):1021–1034. doi:10.1007/s10021-008-9176-6.

Gelman, A., Hwang, J., and Vehtari, A., 2013. Understanding predictive information criteria for Bayesian models. *arXiv:1307.5928 [stat]*.

Gelman, A., Goodrich, B., Gabry, J., and Ali, I., 2017. R-squared for Bayesian regression models∗. page 7.

Gémesi, Z., Downing, J. A., Cruse, R. M., and Anderson, P. F., 2011. Effects of Watershed Configuration and Composition on Downstream Lake Water Quality. *Journal of Environment Quality*, 40(2):517. doi:10.2134/jeq2010.0133.

Golley, F. B., 1993. *A History of the Ecosystem Concept in Ecology: More than the Sum of the Parts*. Yale University Press.

Gomi, T., Sidle, R. C., and Richardson, J. S., 2002. Understanding Processes and Downstream Linkages of Headwater Systems. *BioScience*, 52(10):905. doi:10.1641/0006-3568(2002)052[0905:UPADLO]2.0.CO;2.

Gorham, 1958. The Physical Limnology of Northern Britain: An Epitome of the Bathymetrical Survey of the Scottish Freshwater Lochs, 1897.–1909. *Limnology and Oceanography*, page 11.

GRASS Development Team, 2017. *Geographic Resources Analysis Support System (GRASS GIS) Software, Version 7.4*. http://grass.osgeo.org.

Groffman, P. M., Butterbach-Bahl, K., Fulweiler, R. W., Gold, A. J., Morse, J. L., Stander, E. K., Tague, C., Tonitto, C., and Vidon, P., 2009. Challenges to incorporating spatially and temporally explicit phenomena (hotspots and hot moments) in denitrification models. *Biogeochemistry*, 93(1-2):49–77. doi:10.1007/s10533-008-9277-5.

Guy, M., Taylor, W. D., and Carter, J. C. H., 1994. Decline in Total Phosphorus in the Surface Waters of Lakes during Summer Stratification, and its Relationship to Size Distribution of Particles and Sedimentation. *Canadian Journal of Fisheries and Aquatic Sciences*, 51(6): 1330–1337. doi:10.1139/f94-132.

Hakanson, L., 1977. On Lake Form, Lake Volume and Lake Hypsographic Survey. *Geografiska Annaler. Series A, Physical Geography*, page 31.

Håkanson, L., 1981. On lake bottom dynamics—the energy–topography factor. *Canadian Journal of Earth Sciences*, 18(5):899–909. doi:10.1139/e81-086.

Hanson, P. C., Stillman, A. B., Jia, X., Karpatne, A., Dugan, H. A., Carey, C. C., Stachelek, J., Ward, N. K., Zhang, Y., Read, J. S., and Kumar, V., 2020. Predicting lake surface water phosphorus dynamics using process-guided machine learning. *Ecological Modelling*, 430:109136. doi:10.1016/j.ecolmodel.2020.109136.

Hayashi, M. and van der Kamp, G., 2000. Simple equations to represent the volume–area–depth relations of shallow wetlands in small topographic depressions. *Journal of Hydrology*, 237(1-2):74–85. doi:10.1016/S0022-1694(00)00300-0.

Hayes, N. M., Vanni, M. J., Horgan, M. J., and Renwick, W. H., 2015. Climate and land use interactively affect lake phytoplankton nutrient limitation status. *Ecology*, 96(2):392–402. doi:10.1890/13-1840.1.

Heathcote, A. J., del Giorgio, P. A., Prairie, Y. T., and Brickman, D., 2015. Predicting bathymetric features of lakes from the topography of their surrounding landscape. *Canadian Journal of Fisheries and Aquatic Sciences*, 72(5):643–650. doi:10.1139/cjfas-2014-0392.

Heathwaite, A., Fraser, A., Johnes, P., Hutchins, M., Lord, E., and Butterfield, D., 2003. The Phosphorus Indicators Tool: A simple model of diffuse P loss from agricultural land to water. *Soil Use and Management*, 19(1):1–11. doi:10.1111/j.1475-2743.2003.tb00273.x.

Heffernan, J. B., Soranno, P. A., Angilletta, M. J., Buckley, L. B., Gruner, D. S., Keitt, T. H., Kellner, J. R., Kominoski, J. S., Rocha, A. V., Xiao, J., Harms, T. K., Goring, S. J., Koenig, L. E., McDowell, W. H., Powell, H., Richardson, A. D., Stow, C. A., Vargas, R., and Weathers, K. C., 2014. Macrosystems ecology: Understanding ecological patterns and processes at continental scales. *Frontiers in Ecology and the Environment*, 12(1):5–14. doi:10.1890/130017.

Hijmans, R. J., 2019. *Raster: Geographic Data Analysis and Modeling.* https://CRAN.R-project.org/package=raster.

Hollister, Milstead, W. B., and Urrutia, M. A., 2011. Predicting maximum lake depth from surrounding topography. *PloS one*, 6(9):e25764.

Hollister, J. and Shah, T., 2017. *Elevatr: Access Elevation Data from Various APIs.* http://github.com/usepa/elevatr.

Hollister, J. and Stachelek, J., 2017. Lakemorpho: Calculating lake morphometry metrics in R. *F1000Research*, 6:1718. doi:10.12688/f1000research.12512.1.

Holmes, R., Armanini, D. G., and Yates, A. G., 2016. Effects of Best Management Practice on Ecological Condition: Does Location Matter? *Environmental Management*, 57(5):1062–1076. doi:10.1007/s00267-016-0662-x.

Hothorn, T. and Zeileis, A., 2015. Partykit: A modular toolkit for recursive partytioning in R. *The Journal of Machine Learning Research*, 16(1):3905–3909.

Hothorn, T., Hornik, K., and Zeileis, A., 2006. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical statistics*, 15(3):651–674.

Jarvie, H. P., Neal, C., Withers, P. J., Baker, D. B., Richards, R. P., and Sharpley, A. N., 2011. Quantifying Phosphorus Retention and Release in Rivers and Watersheds Using Extended End-Member Mixing Analysis (E-EMMA). *Journal of Environment Quality*, 40(2):492. doi:10.2134/jeq2010.0298.

Jasiewicz, J. and Metz, M., 2011. A new GRASS GIS toolkit for Hortonian analysis of drainage networks. *Computers & Geosciences*, 37(8):1162–1173. doi:10.1016/j.cageo.2011.03.003.

Johansson, H., Brolin, A. A., and Håkanson, L., 2007. New Approaches to the Modelling of Lake Basin Morphometry. *Environmental Modeling & Assessment*, 12(3):213–228. doi:10.1007/s10666-006-9069-z.

Jones, J. R., knowlton, M. F., and Obrecht, D. V., 2008. Role of land cover and hydrology in determining nutrients in mid-continent reservoirs: Implications for nutrient criteria and management. *Lake and Reservoir Management*, 24(1):1–9. doi:10.1080/07438140809354045.

King, R. S., Baker, M. E., Whigham, D. F., Weller, D. E., Jordan, T. E., Kazyak, P. F., and Hurd, M. K., 2005. Spatial considerations for linking watershed land cover to ecological indicators in streams. *Ecological Applications*, 15(1):137–153. doi:10.1890/04-0481.

Kronvang, B., 1992. The export of particulate matter, particulate phosphorus and dissolved phosphorus from two agricultural river basins: Implications on estimating the non-point phosphorus load. *Water Research*, 26(10):1347–1358. doi:10.1016/0043-1354(92)90129-R.

La Barbera, P. and Rosso, R., 1989. On the fractal dimension of stream networks. *Water Resources Research*, 25(4):735–741. doi:10.1029/WR025i004p00735.

Landers, D. H., Overton, W. S., Linthurst, R. A., and Brakke, D. F., 1988. Eastern Lake Survey, regional estimates of lake chemistry. *Environmental science & technology*, 22(2): 128–135.

Lanza, G. R. and Silvey, J. Interactions of reservoir microbiota: Eutrophication—related environmental problems. In *Microbial Processes in Reservoirs*, pages 99–119. Springer, 1985.

Lark, T. J., Meghan Salmon, J., and Gibbs, H. K., 2015. Cropland expansion outpaces agricultural and biofuel policies in the United States. *Environmental Research Letters*, 10 (4):044003. doi:10.1088/1748-9326/10/4/044003.

Larsson, J., 2019. *Polylabelr: Find the Pole of Inaccessibility (Visual Center) of a Polygon.* https://github.com/jolars/polylabelr.

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rödel, R., Sindorf, N., and Wisser, D., 2011. High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management. *Frontiers in Ecology and the Environment*, 9(9): 494–502. doi:10.1890/100125.

Leibowitz, S. G., Wigington, P. J., Schofield, K. A., Alexander, L. C., Vanderhoof, M. K., and Golden, H. E., 2018. Connectivity of Streams and Wetlands to Downstream Waters: An Integrated Systems Framework. *JAWRA Journal of the American Water Resources Association*.

Li, M., Peng, C., Zhu, Q., Zhou, X., Yang, G., Song, X., and Zhang, K., 2020. The significant contribution of lake depth in regulating global lake diffusive methane emissions. *Water Research*, page 115465. doi:10.1016/j.watres.2020.115465.

Likens, G. E. and Bormann, F. H., 1974. Linkages between Terrestrial and Aquatic Ecosystems. *BioScience*, 24(8):447–456. doi:10.2307/1296852.

Maranger, R., Jones, S. E., and Cotner, J. B., 2018. Stoichiometry of carbon, nitrogen, and phosphorus through the freshwater pipe: Stoichiometry of carbon, nitrogen, and phosphorus. *Limnology and Oceanography Letters*. doi:10.1002/lol2.10080.

Mayer, P. M., Reynolds, S. K., McCutchen, M. D., and Canfield, T. J., 2007. Meta-Analysis of Nitrogen Removal in Riparian Buffers. *Journal of Environmental Quality*, 36(4):1172–1180. doi:10.2134/jeq2006.0462.

McCullough, I. M., Cheruvelil, K. S., Lapierre, J.-F., Lottig, N. R., Moritz, M. A., Stachelek, J., and Soranno, P. A., 2019a. Do lakes feel the burn? Ecological consequences of increasing exposure of lakes to fire in the continental United States. *Global Change Biology*, 25(9): 2841–2854. doi:10.1111/gcb.14732.

McCullough, I. M., King, K. B. S., Stachelek, J., Diaz, J., Soranno, P. A., and Cheruvelil, K. S., 2019b. Applying the patch-matrix model to lakes: A connectivity-based conservation framework. *Landscape Ecology*. doi:10.1007/s10980-019-00915-7.

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., and Schmitt, O., 2016. Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications*, 7:13603. doi:10.1038/ncomms13603.

Metson, G. S., Lin, J., Harrison, J. A., and Compton, J. E., 2017. Linking terrestrial phosphorus inputs to riverine export across the United States. *Water Research*, 124: 177–191. doi:10.1016/j.watres.2017.07.037.

Milstead, W. B., Hollister, J. W., Moore, R. B., and Walker, H. A., 2013. Estimating Summer Nutrient Concentrations in Northeastern Lakes from SPARROW Load Predictions and Modeled Lake Depth and Volume. *PLoS ONE*, 8(11):e81457. doi:10.1371/journal.pone.0081457.

Morefield, P. E., LeDuc, S. D., Clark, C. M., and Iovanna, R., 2016. Grasslands, wetlands, and agriculture: The fate of land expiring from the Conservation Reserve Program in the Midwestern United States. *Environmental Research Letters*, 11(9):094005. doi:10.1088/1748-9326/11/9/094005.

Motew, M., Chen, X., Booth, E. G., Carpenter, S. R., Pinkas, P., Zipper, S. C., Loheide, S. P., Donner, S. D., Tsuruta, K., Vadas, P. A., and Kucharik, C. J., 2017. The Influence of Legacy P on Lake Water Quality in a Midwestern Agricultural Watershed. *Ecosystems*, 20(8):1468–1482. doi:10.1007/s10021-017-0125-0.

Naiman, R. J., Decamps, H., and McClain, M. E., 2010. *Riparia: Ecology, Conservation, and Management of Streamside Communities*. Elsevier.

Nakagaki, N. and Wieczorek, M., 2016. Estimates of subsurface tile drainage extent for 12 Midwest states, 2012. *USGS Data Release*, 10. doi:10.5066/F7W37TDP.

Neumann, J., 1959. Maximum depth and average depth of lakes. *Journal of the Fisheries Board of Canada*, 16(6):923–927.

NRC, 2010. *Toward Sustainable Agricultural Systems in the 21st Century*. National Academies Press, Washington, D.C. ISBN 978-0-309-14896-2. doi:10.17226/12832. http://www.nap.edu/catalog/12832.

Oliver, S., 2017. Unexpected stasis in a changing world: Lake nutrient and chlorophyll trends since 1990. *Global Change Biology*.

Oliver, S. K., Soranno, P. A., Fergus, C. E., Wagner, T., Winslow, L. A., Scott, C. E., Webster, K. E., Downing, J. A., and Stanley, E. H., 2016. Prediction of lake depth across a 17-state region in the United States. *Inland Waters*, 6(3):314–324.

O'Neill, R. V., 2001. IS IT TIME TO BURY THE ECOSYSTEM CONCEPT? (WITH FULL MILITARY HONORS, OF COURSE!) [1]. *Ecology*, 82(12):3275–3284. doi:10.1890/0012-9658(2001)082[3275:IITTBT]2.0.CO;2.

Paerl, H. W., Scott, J. T., McCarthy, M. J., Newell, S. E., Gardner, W. S., Havens, K. E., Hoffman, D. K., Wilhelm, S. W., and Wurtsbaugh, W. A., 2016. It Takes Two to Tango: When and Where Dual Nutrient (N & P) Reductions Are Needed to Protect Lakes

and Downstream Ecosystems. *Environmental Science & Technology*, 50(20):10805–10813. doi:10.1021/acs.est.6b02575.

Pärn, J., Pinay, G., and Mander, Ü., 2012. Indicators of nutrients transport from agricultural catchments under temperate climate: A review. *Ecological Indicators*, 22:4–15. doi:10.1016/j.ecolind.2011.10.002.

Pebesma, E., 2018. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*.

Polus, S., Danila, L., Wang, Q., Tan, P.-N., Zhou, J., Cheruvelil, K., and Soranno, P., In prep. LAGOS-US: RSVR v1.0: Module of the classification of lakes that identifies the probability of a lake being a natural lake or a reservoir for lakes in the conterminous U.S. greater than or equal to 4 ha.

Powers, S. M., Tank, J. L., and Robertson, D. M., 2015. Control of nitrogen and phosphorus transport by reservoirs in agricultural landscapes. *Biogeochemistry*, 124(1-3):417–439. doi:10.1007/s10533-015-0106-3.

Powers, S. E., 2007. Nutrient loads to surface water from row crop production. *The International Journal of Life Cycle Assessment*, 12(6):399–407. doi:10.1065/lca2007.02.307.

Prasad, A. M., Iverson, L. R., and Liaw, A., 2006. Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction. *Ecosystems*, 9(2): 181–199. doi:10.1007/s10021-005-0054-1.

Qian, S. S., Stow, C. A., Nojavan A., F., Stachelek, J., Cha, Y., Alameddine, I., and Soranno, P., 2019. The implications of Simpson's paradox for cross-scale inference among lakes. *Water Research*, 163:114855. doi:10.1016/j.watres.2019.114855.

Qin, B., Zhou, J., Elser, J. J., Gardner, W. S., Deng, J., and Brookes, J. D., 2020. Water Depth Underpins the Relative Roles and Fates of Nitrogen and Phosphorus in Lakes. *Environmental Science & Technology*, 54(6):3191–3198. doi:10.1021/acs.est.9b05858.

Rasmussen, J. B., Godbout, L., and Schallenberg, M., 1989. The humic content of lake water and its relationship to watershed and lake morphometry. *Limnology and Oceanography*, 34 (7):1336–1343. doi:10.4319/lo.1989.34.7.1336.

Read, E. K., Patil, V. P., Oliver, S. K., Hetherington, A. L., Brentrup, J. A., Zwart, J. A., Winters, K. M., Corman, J. R., Nodine, E. R., Woolway, R. I., et al., 2015. The importance of lake-specific characteristics for water quality across the continental United States. *Ecological Applications*, 25(4):943–955.

Read, E. K., Carr, L., De Cicco, L., Dugan, H. A., Hanson, P. C., Hart, J. A., Kreft, J., Read, J. S., and Winslow, L. A., 2017. Water quality data for national-scale aquatic research: The Water Quality Portal: DATA FOR NATIONAL-SCALE AQUATIC RESEARCH. *Water Resources Research*, 53(2):1735–1745. doi:10.1002/2016WR019993.

Read, J. S., Winslow, L. A., Hansen, G. J., Van Den Hoek, J., Hanson, P. C., Bruce, L. C., and Markfort, C. D., 2014. Simulating 2368 temperate lakes reveals weak coherence in stratification phenology. *Ecological Modelling*, 291:142–150. doi:10.1016/j.ecolmodel.2014.07.029.

Renwick, W. H., Vanni, M. J., Zhang, Q., and Patton, J., 2008. Water Quality Trends and Changing Agricultural Practices in a Midwest U.S. Watershed, 1994–2006. *Journal of Environment Quality*, 37(5):1862. doi:10.2134/jeq2007.0401.

Rodriguez-Iturbe, I. and Rinaldo, A., 2001. *Fractal River Basins: Chance and Self-Organization*. Cambridge University Press.

Ruddy, B. C., Lorenz, D. L., and Mueller, D. K., 2006. County-level estimates of nutrient inputs to the land surface of the conterminous United States, 1982-2001.

Russell, M. A., Walling, D. E., Webb, B. W., and Bearne, R., 1998. The composition of nutrient fluxes from contrasting UK river basins. *Hydrological Processes*, 12(9):1461–1482. doi:10.1002/(SICI)1099-1085(199807)12:9<1461::AID-HYP650>3.0.CO;2-6.

Schmadel, N. M., Harvey, J. W., Alexander, R. B., Schwarz, G. E., Moore, R. B., Eng, K., Gomez-Velez, J. D., Boyer, E. W., and Scott, D., 2018. Thresholds of lake and reservoir connectivity in river networks control nitrogen removal. *Nature Communications*, 9(1). doi:10.1038/s41467-018-05156-x.

Schmadel, N. M., Harvey, J. W., Schwarz, G. E., Alexander, R. B., Gomez-Velez, J. D., Scott, D., and Ator, S. W., 2019. Small ponds in headwater catchments are a dominant influence on regional nutrient and sediment budgets. *Geophysical Research Letters*. doi:10.1029/2019GL083937.

Sharpley, A. N., Chapra, S., Wedepohl, R., Sims, J., Daniel, T. C., and Reddy, K., 1994. Managing agricultural phosphorus for protection of surface waters: Issues and options. *Journal of environmental quality*, 23(3):437–451.

Shimoda, Y. and Arhonditsis, G. B., 2015. Integrating hierarchical Bayes with phosphorus loading modelling. *Ecological Informatics*, 29:77–91. doi:10.1016/j.ecoinf.2015.07.005.

Smith, N., Webster, K., Rodriguez, L., Cheruvelil, K., and Soranno, P.A., In prep. LAGOS-US: LOCUS v1.0: Module of location, identifiers, and physical characteristics of lakes and their watersheds in the conterminous U.S.

Smith, V. H., 2003. Eutrophication of freshwater and coastal marine ecosystems a global problem. *Environmental Science and Pollution Research*, 10(2):126–139.

Sobek, S., 2011. Predicting the depth and volume of lakes from map-derived parameters. *Inland Waters*, 1(3):177–184. doi:10.5268/IW-1.3.426.

Sobek, S., Tranvik, L. J., Prairie, Y. T., Kortelainen, P., and Cole, J. J., 2007. Patterns and regulation of dissolved organic carbon: An analysis of 7,500 widely distributed lakes. *Limnology and Oceanography*, 52(3):1208–1219. doi:10.4319/lo.2007.52.3.1208.

Soil Survey Staff. Gridded soil survey geographic (gSSURGO) database for the conterminous united states. Technical report, United States Department of Agriculture, Natural Resources Conservation Service, 2016. https://gdg.sc.egov.usda.gov/.

Søndergaard, M., Bjerring, R., and Jeppesen, E., 2013. Persistent internal phosphorus loading during summer in shallow eutrophic lakes. *Hydrobiologia*, 710(1):95–107. doi:10.1007/s10750-012-1091-3.

Soranno, P. and Cheruvelil, K., 2017a. LAGOS-NE-LIMNO v1.087.1: A module for LAGOS-NE, a multi-scaled geospatial and temporal database of lake ecological context and water quality for thousands of US Lakes: 1925–2013. *Environmental Data Initiative*, 10.

Soranno, P. and Cheruvelil, K., 2017b. LAGOS-NE-GIS v1.0: 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. *Environmental Data Initiative.*

Soranno, P. A. and et, a., 2017. LAGOS-NE: A multi-scaled geospatial and temporal database of lake ecological context and water quality for thousands of US lakes. *GigaScience*, 6(12): 1–22. doi:10.1093/gigascience/gix101.

Soranno, P. A., Cheruvelil, K. S., Webster, K. E., Bremigan, M. T., Wagner, T., and Stow, C. A., 2010. Using Landscape Limnology to Classify Freshwater Ecosystems for Multi-ecosystem Management and Conservation. *BioScience*, 60(6):440–454. doi:10.1525/bio.2010.60.6.8.

Soranno, P. A., Cheruvelil, K. S., Wagner, T., Webster, K. E., and Bremigan, M. T., 2015. Effects of Land Use on Lake Nutrients: The Importance of Scale, Hydrologic Connectivity, and Region. *PLOS ONE*, 10(8):e0135454. doi:10.1371/journal.pone.0135454.

Stachelek, J., Rodriguez, L., Namovich, J., Diaz, J., Hawkins, A., Shoffner, A., McCullough, I., King, K., Egedy, L., Lottig, N., Polus, S., Cheruvelil, K., and Soranno, P.A., In prep. LAGOS-US: DEPTH v0.1: Module of lake depths of lakes in the conterminous U.S.

Stachelek, J., 2018a. *nhdR: R Tools for Working with the National Hydrography Dataset.* https://github.com/jsta/nhdR.

Stachelek, J., 2018b. Streamnet: Morphology Analysis of Stream Networks. doi:10.5281/zenodo.2549736.

Stachelek, J., 2019a. *Gssurgo: Python Toolbox Enabling an Open Source gSSURGO Workflow.* https://github.com/jsta/gssurgo.

Stachelek, J., 2019b. *nhdR: Tools for Working with the National Hydrography Dataset.* https://github.com/jsta/nhdR.

Stachelek, J., 2019c. Freshwater connectivity and stream morphology metrics for Northeast and Midwestern US lakes. doi:10.5281/zenodo.2554212.

Stachelek, J. and Oliver, S, 2019. *LAGOSNE: Interface to the Lake Multi-Scaled Geospatial and Temporal Database.* https://github.com/cont-limno/LAGOSNE.

Stachelek, J. and Soranno, P. A., 2019. Does freshwater connectivity influence phosphorus retention in lakes? *Limnology and Oceanography*. doi:10.1002/lno.11137.

Stachelek, J., Ford, C., Kincaid, D., King, K., Miller, H., and Nagelkirk, R., 2018. The National Eutrophication Survey: Lake characteristics and historical nutrient concentrations. *Earth System Science Data*, page 6. doi:10.5194/essd-10-81-2018.

Stachelek, J., Weng, W., Carey, C. C., Kemanian, A. R., Cobourn, K. M., Wagner, T., Weathers, K. C., and Soranno, P. A., 2020. Granular measures of agricultural land-use influence lake nitrogen and phosphorus differently at macroscales. *Ecological Applications*. doi:10.1002/eap.2187.

Stachelek, J., Hanly, P. J., and Soranno, P. A., In Prep. Geometric models overestimate lake depth due to imperfect slope measurement. *to be submitted to Water Resources Research*.

Stachelek J, Ford C, Kincaid, D, King K, Miller H, and Nagelkirk R, 2017. The National Eutrophication Survey: Lake characteristics and historical nutrient concentrations. doi:10.5063/F10G3H3Z.

Stan Development Team, 2017. *The Stan Core Library*. http://mc-stan.org/.

Strickling, H. L. and Obenour, D. R., 2018. Leveraging Spatial and Temporal Variability to Probabilistically Characterize Nutrient Sources and Export Rates in a Developing Watershed. *Water Resources Research*, 54(7):5143–5162. doi:10.1029/2017WR022220.

T-Prairie, Y. and Kalff, J., 1986. Effect of catchment size on phosphorus export. *JAWRA Journal of the American Water Resources Association*, 22(3):465–470.

Taranu, Z. E. and Gregory-Eaves, I., 2008. Quantifying Relationships Among Phosphorus, Agriculture, and Lake Depth at an Inter-Regional Scale. *Ecosystems*, 11(5):715–725. doi:10.1007/s10021-008-9153-0.

USDA-NASS, 2019. *National Agricultural Statistics Service Cropland Data Layer*. https://nassgeodata.gmu.edu/CropScape/.

USEPA. National eutrophication survey methods 1973–1976 (working paper no. 175). Technical report, United States Environmental Protection Agency, 1975.

USEPA. National Aquatic Resource Surveys. National Lakes Assessment 2012. Technical report, 2016.

USGS. National hydrography Dataset. Technical report, 2018.

USGS. National hydrography Dataset. Technical report, 2019. https://nhd.usgs.gov/.

Vanni, M. J., Renwick, W. H., Headworth, J. L., Auch, J. D., and Schaus, M. H., 2001. Dissolved and particulate nutrient flux from three adjacent agricultural watersheds: A five-year study. *Biogeochemistry*, 54(1):85–114.

Vehtari, A., Gelman, A., and Gabry, J., 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and computing*, 27(5):1413–1432.

Vollenweider, R. A., 1975. Input-output models. *Aquatic Sciences-Research across boundaries*, 37(1):53–84.

Wagner, T. and Schliep, E. M., 2018. Combining nutrient, productivity, and landscape-based regressions improves predictions of lake nutrients and provides insight into nutrient coupling at macroscales: Joint nutrient-productivity model. *Limnology and Oceanography*. doi:10.1002/lno.10944.

Wagner, T., Soranno, P. A., Webster, K. E., and Cheruvelil, K. S., 2011. Landscape drivers of regional variation in the relationship between total phosphorus and chlorophyll in lakes: Relationship between total phosphorus and chlorophyll. *Freshwater Biology*, 56(9): 1811–1824. doi:10.1111/j.1365-2427.2011.02621.x.

Wagner, T., Lottig, N. R., Bartley, M. L., Hanks, E. M., Schliep, E. M., Wikle, N. B., King, K. B. S., McCullough, I., Stachelek, J., Cheruvelil, K. S., Filstrup, C. T., Lapierre, J. F., Liu, B., Soranno, P. A., Tan, P.-N., Wang, Q., Webster, K., and Zhou, J., 2019. Increasing accuracy of lake nutrient predictions in thousands of lakes by leveraging water clarity data. *Limnology and Oceanography Letters*. doi:10.1002/lol2.10134.

Ward, N. K., Fitchett, L., Hart, J. A., Shu, L., Stachelek, J., Weng, W., Zhang, Y., Dugan, H., Hetherington, A., Boyle, K., Carey, C. C., Cobourn, K. M., Hanson, P. C., Kemanian, A. R., Sorice, M. G., and Weathers, K. C., 2018. Integrating fast and slow processes is essential for simulating human–freshwater interactions. *Ambio*. doi:10.1007/s13280-018-1136-6.

Williams, J. and Labou, S. G., 2017. A database of georeferenced nutrient chemistry data for mountain lakes of the Western United States. *Scientific Data*, 4:170069.

Withers, P. and Jarvie, H., 2008. Delivery and cycling of phosphorus in rivers: A review. *Science of The Total Environment*, 400(1-3):379–395. doi:10.1016/j.scitotenv.2008.08.002.

Wright, M. N. and Ziegler, A., 2017. ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, 77(1):1–17. doi:10.18637/jss.v077.i01.

Yang, W., Sheng, C., and Voroney, P., 2005. Spatial Targeting of Conservation Tillage to Improve Water Quality and Carbon Retention Benefits. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie*, 53(4):477–500. doi:10.1111/j.1744-7976.2005.00031.x.

Yigzaw, W., Li, H.-Y., Demissie, Y., Hejazi, M. I., Leung, L. R., Voisin, N., and Payn, R., 2018. A New Global Storage-Area-Depth Data Set for Modeling Reservoirs in Land Surface and Earth System Models. *Water Resources Research*, 54(12). doi:10.1029/2017WR022040.

Zhang, T., Soranno, P. A., Cheruvelil, K. S., Kramer, D. B., Bremigan, M. T., and Ligmann-Zielinska, A., 2012. Evaluating the effects of upstream lakes and wetlands on lake phosphorus

concentrations using a spatially-explicit model. *Landscape Ecology*, 27(7):1015–1030. doi:10.1007/s10980-012-9762-z.

Zimmer, M. A. and McGlynn, B. L., 2018. Lateral, Vertical, and Longitudinal Source Area Connectivity Drive Runoff and Carbon Export Across Watershed Scales. *Water Resources Research*, 54(3):1576–1598. doi:10.1002/2017WR021718.