Landscape Limnology: Lake morphology and process at the continental scale

By

Luke Adam Winslow

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Freshwater and Marine Sciences)

at the

UNIVERSITY OF WISCONSIN – MADISON

2014

Date of final oral examination: 8/20/2014

This dissertation is approved by the following members of the Final Oral Committee:

Emily Stanley, Professor, Zoology
Ankur Desai, Professor, Atmospheric and Oceanic Sciences
Michael Cardiff, Professor, Assistant Professor, Hydrogeology
Katherine McMahon, Professor, Civil and Environmental Engineering
Jordan Read, Civil Engineer, United States Geological Survey
Paul Hanson, Associate Research Professor, Limnology

TABLE OF CONTENTS

ABSTRACT

Many challenges facing humanity operate at the continental and global scales, pushing researchers to address environmental questions at such spatial scales. For issues of global change and biogeochemical cycles in inland waters, building a large-scale understanding of the extent and processes of inland waters is an important to predict and address present and future challenges.

In this work, I take the focus off any single lake and instead work at the lake population scale. This forced me to break new research ground, synthesize new datasets, and re-examine existing theories at a new scale. In Chapter 2, I synthesized and examined a new dataset of lakes in the U.S. and went beyond lake surface area. By improving our understanding of lake shoreline, I improved our understanding of how lakes are linked to their landscapes. In Chapter 3, I broke new ground by building a simple mathematical model of lake population morphology and process, giving researchers a simple and powerful tool for conceptualizing lake population morphology and process. In Chapter 4, I used a large-scale dataset and a fractal model of lakes to examine how different processes can result in different observed distributions of lakes. While theory points to a fractal world, we hypothesize that scale-dependent processes alter the underlying distributional form, altering the observed distribution of lakes. Lastly, with Chapter 5, I examined water temperature trends at a new scale and found that small lakes have a muted response to climate change than larger lakes.

More work is required to make global limnology accepted and recognized by the broad field of limnologists. To get there, new tools, datasets and broad-scale concepts are all required.

ACKNOWLEDGEMENTS

I'm not one for these types of things, so in rough chronological order…

In my life pre-college, I had important help. Without my parents, Jerry and Amanda Winslow, I would not be alive. Nor would I have the interests and skills I have today. My parents represent the perfect balance of practical application and academic pursuit needed in the environmental sciences. Without Brett, my brother, and my intense desire to compete with him as a child, I would not have the computer skills I have today.

Later, in my academic life, there were many that had important roles. Without Paul Hanson, my advisor and previously, boss, I would not be in grad school and would never have become interested in Linmology. Without Jordan Read, Kevin Rose, and Emily Read, I would have not have stuck to my convictions and pursued the ideas that interested me. Without many current and former graduate students of the UW and universities at large (Ryan Batt, Nicholas Preston, Brian Weidel, and many others) I would not understand the culture of academia today.

Lastly, I thank those in my personal life who have helped me in a variety of ways. Taylor Leach helps to remind me of my humanity every day. Jamin Dreyer has been critical in reminding me how it's all a joke. Many others each helped and provided inspiration in their own way. They include, but are not limited to, Dave Bunger, Hilary Dugan, Lucas Beversdorf, Hannes and Anne Naumann, Franzi and Marco Richter, and many many others.

CHAPTER 1 – INTRODUCTION

OVERVIEW

There is much to learn about the role of lakes in processes at continental and global scales. Limnological research has taken advantage of the well-defined boundaries of individual ecosystems, focusing on lakes surrounded by "large areas of inhospitable land" (Arnott et al. 2006). But today, many challenges facing humanity operate at the continental and global scale, challenging policymaker's decision making and pushing researchers to ask relevant questions crossing ecosystem boundaries and at large spatial scales. While large scale work is somewhat new to limnology, historically many disciplines have operated at continental and global scales. Climatology (Rohli and Vega 2011), oceanography (Talley et al. 2011), and demography (Cox 1970) are all examples. Without clear borders, such disciplines have necessarily worked at large spatial scales for decades. With regards to limnology, some have warned that the field has fallen behind others and not turned our broad knowledge of aquatic system function into answers to large-scale scientific and societally relevant questions (Jumars 1990; Reid and Beeton 1992). Recent work has begun showing how important the cumulative role lake ecosystems may play in global-scale systems (Ludwig et al. 1996; Cole et al. 2007; Lewis 2011). These studies have defined the emergence of global limnology, which strives to upscale our understanding of aquatic services and processes to continental and global scales (Downing 2009).

Importantly, a new perspective on data sharing is required to make the sort of cross-site, large-scale datasets available to a new generation of researchers. The work currently undertaken by the likes of the Global Lake Ecological Observatory Network (GLEON) and Cross Scale Interactions (CSI) limnology are a few, but not yet enough, examples of the kind of data sharing, collaborative efforts that are required. With data and tool sharing in mind, the tools and datasets used and developed throughout the course of this work are all published and freely available through sites referenced in each chapter. With resources available online, all figures and result in this manuscript could be replicated and reproduced.

Research into understanding aquatic system process at the global scale is entering an age of rapid growth. Internet-based communications and data sharing technologies have eased data sharing and collaboration across large spatial scales. With the right network of people today, large datasets could be collected from every corner of the globe in only a few days. Transit times and physically moving data are constraints of the past. In addition, large and spatially comprehensive datasets released by governments, research groups and non-governmental organizations are now freely available to any researcher with an internet connection. Just a few examples relevant to limnology include the USGS National Hydrography Dataset (nhd.usgs.gov), the EPA National Lakes Assessment dataset (water.epa.gov), and the work currently undertaken by GLEON and CSI Limnology groups. As global environmental challenges continue, advances in communications, computing and access to spatially-distributed datasets may help us tackle them head-on.

A critical underlying piece of information needed to understand lakes globally is a reliable, well understood measurement of the very existence and spatial extent of lakes in order to upscale process. While some work has been done to quantify the total number and surface area of lakes globally (Meybeck 1995; Kalff 2002), only recently have those numbers been updated based on new, globally cohesive datasets (Downing et al. 2006). This work has had a large impact, but some questions still remain over the abundance, size distribution, and total surface area of lakes globally. There is evidence that the distribution of lakes may not follow a Pareto distribution across all size classes, an underlying assumption used in current estimates of lake number and extent (Seekell and Pace 2011). In addition, new lake and reservoir datasets have suggested that both the overall abundance and surface area of lakes may have been overestimated (McDonald et al. 2012). Both of these studies show that the greatest uncertainty lies in quantifying the number and extent of small lakes and ponds. Understanding the differences presented in these studies is critical to upscaling process, especially with so much uncertainty in small lakes and ponds, which may have roles disproportionate to their surface area (Hanson et al. 2007; Downing 2010; Kastowski et al. 2011).

With only a few exceptions, studies of lake abundance at the global scale have focused on the estimate of total lake surface area and number, without addressing any other morphological characteristics, including volume and total perimeter that may be of use to better understand the role and potential importance of lakes globally. Robert Wetzel described in his 1990 Baldi memorial lecture what the distribution of world lakes might look like (Wetzel 1990). Since then, his description of the lake area versus abundance distribution has turned out to be strikingly accurate (Downing et al. 2006). But a lake is not defined *only* by its surface. Wetzel's description included the two other primary dimensions of lake morphology, the total volume (as described by mean depth) and the total shoreline length which he describes using the pelagic to littoral ratio. To-date, there has been little work examining the large-scale distribution of lake volume or perimeter, with the exception of a few regional studies which examined the distribution of lake volume or depth and related them to lake area and indices of the surrounding landscape topography (Sobek et al. 2007; Hollister and Milstead 2010; Hollister et al. 2011). These types of relationships may be critical in scaling lentic process to large scales when the volume of millions of lakes across the globe is not empirically known and must be estimated from available landscape and spatial information.

The three key morphological features in terms of the dimensionality of their units can be used as a simple conceptual model. Perimeter is 1-D, surface area, 2-D, and volume is 3-D. Surface area, the most commonly examined dimension, scales significantly with many processes, and processes are often reported in a per-unit area basis. For example, irradiative energy exchange, gas exchange, and primary productivity all scale significantly with the total surface area of a lake (Lewis 2011). Lake volume strongly affects the water residence time, relating to many in-lake processes. For example, volume, and by extension, depth has been show to affect nutrient dynamics, primary production, organic matter mineralization, and sedimentation (Hamilton and Jr 1990; Algesten et al. 2003; Jeppesen et al. 2005). Lastly, lake perimeter describes the amount of aquatic-terrestrial interface and thus, will scale with processes which transfer materials between aquatic and terrestrial ecosystems (Gasith and Hasler 1976; Preston et al. 2008) as well as the amount of littoral surface area, informing the balance between littoral and pelagic lake processes (Vadeboncoeur et al. 2008).

Understanding morphology is not the end-goal of global limnology. Morphology is a stepping-stone to better understanding processes in lakes that society is interested in predicting and understanding. One current issue of societal concern is the future of inland waters in a globally changing climate. Climate change impact on lakes may be translated through lake watersheds (Schindler 2009), changes to incoming hydrology (Mortsch and Quinn 1996), or may come in the form of direct changes to lake thermal structure and temperature (Livingstone 2003). Most work on the influence of climate change on lake temperature has focused on the surface waters of the largest lakes globally (Coats et al. 2006; Austin and Colman 2007; Hampton et al. 2008; Schneider and Hook 2010).

Individual lakes will not all have the same responses to climate change. Large lakes have been shown to respond strongly to a changing climate, with ice-albedo feedbacks potentially amplifying the climate signal (Austin and Colman 2007). The small lake response to climate change has not been well studied. Published work suggests that small lakes may respond differently to external climate forcing. For example, small lakes tend to have higher DOC concentrations (Hanson et al. 2007), resulting in different vertical distributions of thermal energy (Read and Rose 2013). Small lakes are often well sheltered from wind (Markfort et al. 2010) and tend to be shallower, on average, than larger lakes (Sobek et al. 2011). Lastly, small lakes have different drivers of turbulence and mixing than large lakes (Read et al. 2012).

IN THIS DISSERTATION

This dissertation is the product of several years' work to answer some questions that arise when trying to understand lake morphology and process at the continental scale. There are many unanswered questions and much progress yet to be made, but the field is now further along than it was a few years ago.

In **Chapter 2**, I undertook a large-scale analysis of the shoreline morphology of lakes. Past studies have already begun building a picture of lake morphology measured as lake surface area (Downing et al. 2006; McDonald et al. 2012). The goal of this chapter was to build an understanding of the large-scale morphological characteristics of lake populations, specifically lake perimeter. The chapter was published in the journal, Freshwater Biology in 2013 (Vol. 59, Issue 2).

In **Chapter 3**, I built a simple, mathematical model for lake populations based on the Pareto distribution that could be used to help quickly understand and conceptualize the distribution of lake morphological features as well as some processes across large population of lakes. The chapter is in press in the journal, Inland Waters (2014).

In **Chapter 4**, I simulate how scale-specific processes, such as evaporation or terrestrialization, could affect the shape of the lake size-abundance distribution. These simulations are based on an underpinning fractal landscape that has been hypothesized to be the underlying origin of lake population distributions. This chapter is in preparation to be submitted to JGR: Earth Surface.

In **Chapter 5**, I examine water temperature trends across a large population of lakes in Wisconsin for which observations are available. With this work, I find that low wind-driven buoyancy flux during the spring in small lakes reduces downward energy flux into the bottom waters, preventing small lake bottom temperatures from warming due to changes in springtime conditions. The chapter is currently in preparation for the journal Geophysical Research Letters.

References

Algesten, G., S. Sobek, A.-K. Bergström, A. Ågren, L. J. Tranvik, and M. Jansson. 2003. Role of lakes for organic carbon cycling in the boreal zone. Glob. Chang. Biol. **10**: 141–147.

Arnott, S. E., J. J. Magnuson, S. I. Dodson, and A. C. C. Colby. 2006. Lakes as Islands: Biodiversity, Invasion, and Extinction, p. 67–88. *In* J.J. Magnuson, T.K. Kratz, and B.J. Benson [eds.], Long-Term Dynamics of Lakes in the Landscape. Oxford Univ Press.

Austin, J. A., and S. M. Colman. 2007. Lake Superior summer water temperatures are increasing more rapidly than regional air temperatures: A positive ice-albedo feedback. Geophys. Res. Lett. **34**: L06604.

Coats, R., J. Perez-Losada, G. Schladow, R. Richards, and C. Goldman. 2006. The Warming of Lake Tahoe. Clim. Change **76**: 121–148.

Cole, J. J., Y. T. Prairie, N. F. Caraco, W. H. McDowell, L. J. Tranvik, R. G. Striegl, C. M. Duarte, P. Kortelainen, J. a. Downing, J. J. Middelburg, and J. Melack. 2007. Plumbing the Global Carbon Cycle: Integrating Inland Waters into the Terrestrial Carbon Budget. Ecosystems **10**: 172–185.

Cox, P. R. 1970. Demography, First Edit. R. Farley [ed.]. Cambridge University Press.

Downing, J. A. 2009. Global limnology: up-scaling aquatic services and processes to planet Earth. Verh. Internat. Verein. Limnol. **30**.

Downing, J. A. 2010. Emerging global role of small lakes and ponds: little things mean a lot. Limnetica **1**: 9–24.

Downing, J. A., Y. T. Prairie, J. J. Cole, C. M. Duarte, L. J. Tranvik, R. G. Striegl, W. H. McDowell, P. Kortelainen, N. F. Caraco, and J. M. Melack. 2006. The global abundance and size distribution of lakes, ponds, and impoundments. Limnol. Oceanogr. **51**: 2388–2397.

Gasith, a., and a. D. Hasler. 1976. Airborne litterfall as a source of organic matter in lakes. Limnol. Oceanogr. **21**: 253–258.

Hamilton, S., and W. L. Jr. 1990. Basin morphology in relation to chemical and ecological characteristics of lakes on the Orinoco River floodplain, Venezuela. Arch. fur Hydrobiol. Stuttgart **119**: 393–425.

Hampton, S. E., L. R. Izmest'Eva, M. V. Moore, S. L. Katz, B. Dennis, and E. a. Silow. 2008. Sixty years of environmental change in the world's largest freshwater lake - Lake Baikal, Siberia. Glob. Chang. Biol. **14**: 1947–1958.

Hanson, P. C., S. R. Carpenter, J. A. Cardille, M. T. Coe, and L. A. Winslow. 2007. Small lakes dominate a random sample of regional lake characteristics. Freshw. Biol. **52**: 814–822.

Hollister, J., and W. B. Milstead. 2010. Using GIS to estimate lake volume from limited data. Lake Reserv. Manag. **26**: 194–199.

Hollister, J. W., W. B. Milstead, and M. A. Urrutia. 2011. Predicting Maximum Lake Depth from Surrounding Topography G.J.-P. Schumann [ed.]. PLoS One **6**: e25764.

Jeppesen, E., M. Sondergaard, J. P. Jensen, K. E. Havens, O. Anneville, L. Carvalho, M. F. Coveney, R. Deneke, M. T. Dokulil, B. Foy, D. Gerdeaux, S. E. Hampton, S. Hilt, K. Kangur, J. Kohler, E. H. H. R. Lammens, T. L. Lauridsen, M. Manca, M. R. Miracle, B. Moss, P. Noges, G. Persson, G. Phillips, R. Portielje, S. Romo, C. L. Schelske, D. Straile, I. Tatrai, E. Willen, and M. Winder. 2005. Lake responses to reduced nutrient loading - an analysis of contemporary long-term data from 35 case studies. Freshw. Biol. **50**: 1747–1771.

Jumars, P. 1990. W(h)ither limnology? Limnol. Oceanogr. **35**: 1216–1218.

Kalff, J. 2002. Limnology: Inland Water Ecosystems, Prentice Hall.

Kastowski, M., M. Hinderer, and A. Vecsei. 2011. Long-term carbon burial in European lakes: Analysis and estimate. Global Biogeochem. Cycles **25**: 1–12.

Lewis, W. M. 2011. Global primary production of lakes: 19th Baldi Memorial Lecture. Inl. Waters **1**: 1–28.

Livingstone, D. 2003. Impact of secular climate change on the thermal structure of a large temperate central European lake. Clim. Change **57**: 205–225.

Ludwig, W., J. Probst, and S. Kempe. 1996. Predicting the oceanic input of organic carbon by continental erosion. Global Biogeochem. Cycles **10**: 23–41.

Markfort, C. D., A. L. S. Perez, J. W. Thill, D. a. Jaster, F. Porté-Agel, and H. G. Stefan. 2010. Wind sheltering of a lake by a tree canopy or bluff topography. Water Resour. Res. **46**: W03530.

McDonald, C. P., J. A. Rover, E. G. Stets, and R. G. Striegl. 2012. The regional abundance and size distribution of lakes and reservoirs in the United States and implications for estimates of global lake extent. Limnol. Oceanogr. **57**: 1–12.

Meybeck, M. 1995. Global Distribution of Lakes, p. 1–35. *In* Physics and Chemistry of Lakes.

Mortsch, L. D., and F. H. Quinn. 1996. Climate change scenarios for Great Lakes Basin ecosystem studies. Limnol. Oceanogr. **41**: 903–911.

Preston, N. D., S. R. Carpenter, J. J. Cole, and M. L. Pace. 2008. Airborne carbon deposition on a remote forested lake. Aquat. Sci. **70**: 213–224.

Read, J. S., D. P. Hamilton, A. R. Desai, K. C. Rose, S. MacIntyre, J. D. Lenters, R. L. Smyth, P. C. Hanson, J. J. Cole, P. a. Staehr, J. a. Rusak, D. C. Pierson, J. D. Brookes, A. Laas, and C. H. Wu. 2012. Lake-size dependency of wind shear and convection as controls on gas exchange. Geophys. Res. Lett. **39**: L09405.

Read, J. S., and K. C. Rose. 2013. Physical responses of small temperate lakes to variation in dissolved organic carbon concentrations. Limnol. Oceanogr. **58**: 921–931.

Reid, D. F., and A. M. Beeton. 1992. Large lakes of the world: A global science opportunity. GeoJournal **28**: 67–72.

Rohli, R. V., and A. J. Vega. 2011. Climatology, Jones & Bartlett Publishers.

Schindler, D. W. 2009. Lakes as sentinels and integrators for the effects of climate change on watersheds, airsheds, and landscapes. Limnol. Oceanogr. **54**: 2349–2358.

Schneider, P., and S. J. Hook. 2010. Space observations of inland water bodies show rapid surface warming since 1985. Geophys. Res. Lett. **37**, doi:10.1029/2010GL045059

Seekell, D. a., and M. L. Pace. 2011. Does the Pareto distribution adequately describe the size-distribution of lakes? Limnol. Oceanogr. **56**: 350–356.

Sobek, S., J. Nisell, and J. Fölster. 2011. Predicting the volume and depth of lakes from map-derived parameters. Inl. Waters **1**: 177–184.

Sobek, S., L. J. Tranvik, Y. T. Prairie, P. Kortelainen, and J. J. Cole. 2007. Patterns and regulation of dissolved organic carbon: An analysis of 7,500 widely distributed lakes. Limnol. Oceanogr. **52**: 1208–1219.

Talley, L. D., G. L. Pickard, W. J. Emery, and J. H. Swift. 2011. Descriptive Physical Oceanography: An Introduction (Google eBook), Academic Press.

Vadeboncoeur, Y., G. Peterson, M. J. Vander Zanden, and J. Kalff. 2008. Benthic algal production across lake size gradients: interactions among morphometry, nutrients, and light. Ecology **89**: 2542–52.

Wetzel, R. G. 1990. Land-water interfaces: metabolic and limnological regulators. Verh Intern. Verein Limnol **24**: 6–24.

CHAPTER 2 - LAKE SHORELINE IN THE CONTIGUOUS UNITED STATES: QUANTITY, DISTRIBUTION AND SENSITIVITY TO OBSERVATION RESOLUTION

*Luke A. Winslow[1], Jordan S. Read[2], Paul C. Hanson[1], Emily H. Stanley[1]

Publication Journal: Freshwater Biology

1. University of Wisconsin - Madison, Center for Limnology, 680 N. Park Street, Madison, Wisconsin 53706 USA
2. U.S. Geological Survey, Center for Integrated Data Analytics, 8505 Research Way, Middleton, WI 53562 USA

* Corresponding author. Email: lawinslow@wisc.edu

Keywords: lakes, littoral zone, global limnology, lake morphology, terrestrial-aquatic interface

Summary

1. Quantifying lake biogeochemical processing at broad spatial scales requires that we scale processes along with physical metrics. Past work has primarily scaled lentic processes using estimates of lake surface area. However, many processes important to lakes, such as material, energy and biological fluxes and biogeochemical cycling, scale with lake perimeter.

2. We estimate the total lake perimeter for the contiguous United States (U.S.) and examine the sensitivity of this estimate to measurement resolution. At the original mapping resolution, lakes in the contiguous U.S. have a total perimeter of over 1.8 million km.

3. The change in measured perimeter versus measurement resolution for the contiguous U.S. had a log-log slope (also known as the fractal dimension) of -0.21, generally less than previously reported estimates. With changing observation resolution, total measured perimeter was most sensitive to the inclusion or exclusion of small lakes, not shoreline complexity.

4. The total aquatic-terrestrial interface in lakes is less than one tenth that of streams and rivers, which collectively account for over 21 million km of shoreline in the contiguous U.S. This study further describes the distribution of lake perimeter and proposes a technique that can contribute to understanding continental-scale processes.

Introduction

Despite estimates that suggest lakes comprise only a small fraction of the Earth's surface relative to terrestrial and ocean environments (between 2 and 4% of total terrestrial surface area; Kalff 2002; Downing *et al.* 2006), they play a significant role in the global carbon cycle (Dean & Gorham 1998; Lewis 2011). However, quantifying aquatic processes at continental and global scales is inherently difficult because few fluxes into and out of lakes can be directly observed (Battin *et al.* 2009). For lakes, measurements from a relatively small number of systems have been upscaled to regions by using lake surface area as a scalar for inferring large-scale fluxes and process (Tranvik *et al.* 2009; Lewis 2011). While area is an appropriate scalar for many processes, other processes, such as those dependent upon connectivity to terrestrial systems, may benefit from scaling with lake perimeter.

Many important material fluxes and ecosystem processes occur at or near the perimeter of lakes. For example, terrestrial organic carbon subsidizes high rates of secondary production in littoral zones (Hershey *et al.* 2006; Francis & Schindler 2009) and may provide substantial organic substrates to pelagic consumers, such as zooplankton, especially in small lakes (Cole *et al.* 2011). Allochthonous subsidies that occur at the lake ecosystem interface, such as leaf litter input, have been recognized to scale with lake perimeter (Gasith & Hasler 1976; Preston *et al.* 2008).

Areas immediately adjacent to the perimeter, namely the littoral and riparian zones, are also important sites for biological activity and biogeochemical processing. Littoral zones are often hotspots of methane and nitrous oxide emissions (Wang *et al.* 2006; Bergström *et al.* 2007), as well as benthic primary productivity (Vadeboncoeur *et al.* 2008). In large lakes, 93% of species inhabit shallow, nearshore littoral zones (Vadeboncoeur, McIntyre & Vander Zanden

2011). The importance of nearshore littoral habitat is also apparent in links between decreased abundance of coarse woody debris and a decreased yellow perch density (Helmus & Sass 2008) and fish production capacity (Schindler, Geib & Williams 2000). Perhaps because of the importance of littoral habitat, shoreline development factor (SDF; Kalff 2001), a metric of shoreline complexity, has been a useful correlate of fish density and diversity (Drake & Pereira 2002; Scheuerell & Schindler 2004; Guy & Willis 2011). These examples are a partial representation of the work examining phenomena at the lakeshore, which has long been known to be important to ecosystem function (Wetzel 1990; Schindler & Scheuerell 2002; McClain *et al.* 2003; Strayer & Findlay 2010).

Recent work showing that aquatic-terrestrial boundaries are critical interfaces between ecosystems and that lakes are hotspots for material flux and processing in the landscape (Williamson *et al.* 2008; Karlsson *et al.* 2010; Buffam *et al.* 2011) suggests that focusing on the aquatic-terrestrial boundary may provide a useful approach to scaling. Advancing our understanding of the roles lakes play in the broader landscape requires that we better quantify littoral and perimeter extents and gain further understanding of the uncertainties of these estimates. Examining the spatially explicit distribution of total lake perimeter, in addition to lake area, could highlight different regional characteristics of lakes compared to examining lake area alone.

To quantify the fluxes and processes at the aquatic-terrestrial boundary at broad spatial scales, we first need to quantify the extent of interface and examine the sources and implications of uncertainties in those estimates. Part of the uncertainty in lake perimeter derives from the fractal nature of lake shoreline (Mandelbrot 1979; Kent & Wong 1982). The measured perimeter of a fractal shoreline does not converge with increasing measurement resolution. Under a fractal paradigm, a lake cannot be said to have a specific perimeter measurement that is absolute; rather, it has an observed perimeter that varies with measurement resolution. Despite this, observed perimeter has proven useful directly and as a component in SDF calculations in multiple studies (Drake & Pereira 2002; Scheuerell & Schindler 2004; Guy & Willis 2011). These studies suggest that perimeter observations, when measured at similar resolutions, are comparable and convey useful ecological information despite being sensitive to measurement resolution.

In this study, we explore the distribution and total length of lake perimeter for a geographically extensive data set with consistent resolution covering the contiguous United States (U.S.). We examine how lake perimeter can be quantified at large scales and how its spatial distribution differs from the spatial distribution of lake surface area. We use the U.S. Geological Survey's (USGS) National Hydrography Dataset to estimate the total length of the land-water interface contributed by lakes in the U.S. and then compare different perimeter estimation techniques to address the following questions: What is the total U.S. lake perimeter and how is it distributed across the U.S.? How does the extent of lake shoreline compare with that of rivers and streams? How sensitive is our estimate of perimeter to measurement resolution? What are the implications for our understanding of lake SDF? By answering these questions, we can begin to better understand the importance of lakes at the interface with their adjacent ecosystems and how that scales to broad spatial extents.

METHODS

The dataset examined was the USGS National Hydrography Dataset (retrieved January 2012, http://nhd.usgs.gov). This combines multiple surveys through time, using the resulting high resolution USGS topographical maps to delineate aquatic boundaries (Simley & Carswell 2009). While a variety of resolutions are available, only 1:24,000 was used in this analysis. The dataset covers the area of all 48 U.S. contiguous states, including Washington D.C. Because they are not completely contained within the U.S., we chose to exclude the Laurentian Great Lakes for this analysis.

We extracted all lake, reservoir and pond polygons from the GIS dataset. Because the dataset does not distinguish between artificial or natural lakes and ponds, no distinction was made for this analysis and all water bodies hereafter are collectively referred to as lakes. The elevation component of these polygons was removed using ArcGIS (ESRI ArcGIS v10.1). For simplicity, all island data were excluded from the results presented here, though we discuss the small resulting bias. Duplicate polygons were identified and removed using the permanent identifier field included with the data. All calculations were completed using the Mathworks Mapping Toolbox (v2010b; Mathworks, Natick, United States), which adds geographic information functionality to Matlab. A single-point location for each lake was defined as the centroid of the lake boundary polygon.

To examine perimeter while considering the issue of a fractal lake perimeter, we estimated perimeter using two complexity-insensitive and repeatable techniques. First, a theoretical minimum perimeter (P_{min}) was established by calculating the perimeter of a circle with the same area as each lake using:

$$Pmin = 2(\pi A)^{½} \quad (1)$$

where A is the total area of the individual lake (Kalff 2001). This represents the absolute minimum perimeter, on a flat plane, required to encompass a given area (spherical coordinates could reduce P_{min} further, though for the size scale of lakes examined here, the difference is negligible). Second, a resolution-specific perimeter was calculated using a simple yardstick method based on Mandelbrot (1979). With this method, a fixed-length line-segment was progressively "walked" along the polygon until reaching the start point (Fig. 1). This simulated a perimeter estimate at specific and adjustable mapping resolutions. The yardstick length was varied across a range of values (25 to 1600 m) to examine the sensitivity of the perimeter estimate to measurement resolution. To compare the sensitivity of perimeter estimate to previous studies, the slope of the relationship between log-perimeter and log-yardstick length was calculated using least-squares regression, based on Kent and Wong (1982). Lastly, we calculated the maximum observed perimeter of each polygon (P_{obs}) by using the full-resolution data and summing the lengths of each polygon segment. Shoreline development factor (SDF) was calculated using the equation:

$$SDF = \frac{P}{2(A\pi)^{½}} \quad (2)$$

where A is area and P is perimeter. The perimeter measurement technique used in SDF calculations (P_{obs} or the yardstick method) is indicated where discussed. Calculations of perimeters, area and centroid were fast for any single polygon, but because of the high number of lakes and high polygon resolution, calculations became computationally intensive. This and other geostatistical calculations were accelerated by use of a computer cluster. HTCondor software (Thain, Tannenbaum & Livny 2005) was used to distribute this task.

The extents of stream and river shorelines were calculated from the National Hydrography Dataset using the same technique as the fractal-naïve perimeter (P_{obs}). Only features classified with a type of Stream/River (USGS Feature #460) were included. This feature type includes intermittent, ephemeral and perennial streams and rivers, but excludes features such as underground streams and canals. The shoreline length of small streams, represented in the dataset by polyline objects, was estimated as the length of the lines doubled to account for both sides of the stream. The shoreline of larger rivers, which are stored in the dataset as polygons, was calculated directly as the observed perimeter of the polygons (P_{obs}) on the WGS84 datum.

Maps of lake abundance (number km^{-2}), area (percent cover) and shoreline density (m km^{-2}) were created by dividing the area of the U.S. into equal-area cells (cell size: 50 km^2). Each lake's attributes were assigned to a cell based on its unique single-point location and statistics were calculated for lake density and percent cover in each cell. Where a lake's area exceeded that of a single cell, the full lake shape was split into overlapping cells based on the amount of overlapping area. Cumulative distributions of lake number, area and perimeter as a function of lake area were used to evaluate general attributes of the entire U.S. lake population.

To aid future work in this area, we have released useful derived datasets on the web. While the data are freely available from the USGS, the National Hydrography Dataset's large size makes analyses challenging. To encourage additional research in this area, we have released the extracted perimeter data set. It is available at the data repository hosted by the North Temperate Lakes – Long-Term Ecological Research website at http://lter.limnology.wisc.edu.

RESULTS

Applying no size cutoff beyond the underlying sampling resolution (~0.07 hectares; McDonald *et al.* 2012), the dataset contained 5,800,000 lakes distributed over 8,000,000 km^2 of contiguous U.S. land area. Between quantification techniques, total perimeter estimates ranged from 1,200,000 (P_{min}) to 1,880,000 km (P_{obs}). Including the shoreline of islands increased P_{obs} by 3% to 1,940,000 km. Due to their small contribution, islands were excluded for subsequent analyses to simplify calculations. Applying minimum surface area cutoffs of 0.1 and 1 ha reduced P_{obs} to approximately 1,700,000 and 900,000 km respectively. Lakes below 1 ha (often used as the cutoff for 'small' lakes) accounted for only 7.8% of the total lake area but represented 23.6% of the total lake shoreline (P_{obs}). Although the Laurentian Great Lakes were not included in this study, their inclusion would represent a minor increase of between 1 and 2% to total lake shoreline of U.S. contiguous states.

Perennial rivers and streams had a combined shoreline length of 5,810,000 km. Ephemeral and intermittent streams and rivers had shoreline length totaling 15,560,000 km. Combined, all rivers and streams in U.S. contiguous states had over 21,370,000 km of shoreline.

Lakes were unevenly distributed across the contiguous U.S., with distinct differences in patterns for shoreline, abundance and surface area (Fig. 2). Lake abundances (number km^{-2}) were highest in the southeastern Great Plains, in the highlands between the lower Mississippi and Alabama/Tombigbee River valleys and in the Ohio River Valley region (Fig. 2b). The spatial pattern for lakeshore distribution appeared similar to that of abundance, but suggests an even greater extent of regions with high shoreline density (Fig. 2a). Classically identified lake districts (Upper Midwest, Northeast) did not have notably high lake abundances at the continental scale, but were shoreline-rich and area-rich (Fig. 2a; 2c). Topography and aridity had clear influences on lake distribution, as low gradient regions in the eastern U.S. were lake-rich while the

Colorado plateau and Great Basin regions in the west were comparably lake-poor. While state boundaries have little ecological or geological significance, they are the spatial units at which much of policy and management is enacted. To that end, we included a state-by-state breakdown of lakes and their characteristics (Table 1). Texas had the highest number of lakes (998,000), resulting primarily from its large size and, to a lesser degree, from its above average lake density (1.44 lakes km^{-2}). Of all states, Louisiana had the highest surface area covered by lakes (7.5%). Maryland and Nevada represented the extremes of lake density with 3.9 and 0.05 lakes km^{-2} respectively and Maryland's high density of lakes was also associated with the highest P_{obs} density at 1,100 m km^{-2}.

Small lakes made a disproportionally large contribution of shoreline length to the overall distribution, regardless of measurement technique. The size-distribution of P_{obs} and P_{min} (the two extremes) differed, with P_{obs} having slightly more skew towards large size classes compared to P_{min} (Fig. 3). The difference between the distributions of P_{obs} and P_{min} is a function of differences in average shoreline development factor (SDF) across lake size classes, because it is defined as the ratio between the two perimeters (SDF = P_{obs}/P_{min}). Lakes larger than 1 km^2 had a median SDF of 2.3 while lakes smaller than 1 km^2 had median SDF of 1.14. Lastly, the contribution of different lake size-classes to total lake area was skewed toward the larger lakes, though the distribution was somewhat even across size-classes larger than 0.01 km^2.

Total perimeter of the lake population decreased with increasing yardstick length (Fig. 4) with a linear slope between log-yardstick length and log-perimeter (also known as the fractal dimension) of -0.62. Much of the sensitivity of total perimeter to yardstick length was due to the loss of small lakes when the yardstick became too long to be contained even once in the lake polygon (Fig. 5). To exclude sensitivity due to the loss of small lakes, the population of larger lakes was examined. A sub-population of the dataset was chosen based on the coarsest yardstick resolution (3,200 m), so the lakes would be comparable across the full range of yardstick lengths. In the lake size bins in which no lakes were excluded due to yardstick length, sensitivity to resolution had an average log-perimeter, log-yardstick length slope of -0.2, less than the entire population's sensitivity (Fig. 6).

DISCUSSION

While the distribution of total lake surface in the U.S. is skewed towards larger lakes (McDonald *et al.* 2012), lake perimeter is strongly skewed towards small lakes, with lakes smaller than 1 hectare contributing almost 23% to P_{obs}. This indicates that processes that scale with total perimeter or occur primarily in the littoral or riparian zone may, at a lake-population level, be skewed strongly towards small lakes and ponds. This skew towards smaller lakes has further implications. Although previous work found that the fraction of total surface area contributed by lake size-classes increases with decreasing size (Downing *et al.* 2006), a more recent analysis of a dataset including a greater lake size range found that small lakes in the U.S. contribute a decreasing fraction to total surface area (McDonald *et al.* 2012). Regardless of the relative contribution of small versus large lakes to total area, both studies reported similar estimates of total area. In contrast, our estimates of perimeter have higher uncertainty because of the importance of small lakes in the perimeter distribution, and small lakes are more difficult to quantify at broad spatial scales.

Calculating P_{min} from the perimeter of an equal area circle provides a useful lower boundary for perimeter estimates. The absolute minimum perimeter possible for lakes in the U.S. contiguous states, P_{min} (1,200,000 km), was approximately 36% less than P_{obs}. Differences between P_{obs} and P_{min} are sensitive to the mapping resolution used and future estimates would probably increase when higher resolution mapping products are used. Despite sensitivity to resolution, most estimates of lake perimeter in previous studies (e.g., Kent & Wong 1982; Riera *et al.* 2000; Sharma & Byrne 2011) were measured using maps at similar or coarser resolutions than this study, meaning their perimeter estimates would fall between P_{obs} and P_{min} presented here.

The difference between the distributions of P_{min} and P_{obs} across lake area is indicative of a bias in observed shoreline development factor (SDF) as a function of lake size. This is more apparent if eqn. 2 is rewritten as:

$$SDF = \frac{P_{obs}}{P_{min}} \quad (3)$$

If average SDF were constant across lake sizes, P_{obs} and P_{min} would have similar distributions across lake size. Any difference in how P_{min} and P_{obs} are distributed across lake size must indicate a shift in average SDF. In lakes of the U.S. contiguous states, the contribution to total P_{obs} is higher in large lakes than would be predicted if there were a constant SDF across sizes. This means that on average, large lakes have higher observed SDFs than smaller lakes. This result is also apparent if the population is split and compared, as lakes smaller than and larger than 1 km^2 have median SDFs of 1.1 and 2.3, respectively.

While a positive trend in SDF across lake sizes has been observed previously (Hanson *et al.* 2007), it is unclear whether the trend represents a fundamental change in the morphology of lake shoreline or is simply a product of fixed mapping resolution applied to a large size range of lakes. The fixed mapping resolution could result in a smaller lake always presenting a lower observed SDF than a larger lake, even if their geometries were identical, scaled copies. For example, if a large lake's perimeter outline contained a small feature like a peninsula, and the same outline was scaled down but mapped at the same resolution, the peninsula feature would eventually become too small to be mapped and would no longer be part of the outline. This simplification would decrease the observed SDF without a change in actual shoreline complexity. To examine this potential issue in our analysis, we used a version of the yardstick perimeter measurement method. By pinning the characteristic measurement length to a constant proportion of P_{min}, we simulated a measurement resolution relative to the size of the lake. With this method, smaller features were neglected for larger lakes, virtually scaling the mapping resolution to the lake size. This technique changed the median SDF of lakes smaller than 1km^2 from 1.14 to 1.13 and lakes larger than 1km^2 from 2.3 to 1.5. While the median SDF for large lakes using the size-relative measurement resolution is still higher than smaller lakes (1.5 and 1.13, respectively), this increase in parity highlights a potential challenge when using SDF to compare lakes of large size difference. Counter intuitively, this issue is exacerbated by consistent measurement resolution. Further work is required to better understand the sensitivity of SDF across mapping resolutions and lake sizes, especially if SDF is to be used in studies comparing lakes with large ranges in size.

Compared to rivers and streams, lakes make up a small fraction of the total aquatic-terrestrial interface. Intermittent and perennial streams and rivers combined have an order of magnitude greater aquatic-terrestrial interface than lakes (21,370,000 km and 1,880,000 km of shoreline, respectively). The per-unit shoreline flux of material is often used to measure and upscale loading rates into aquatic systems (e.g., Conners & Naiman 1984; Preston *et al.* 2008). This suggests that systems with greater total interface (like streams versus lakes) may contribute more to collective flux into aquatic systems, and our estimate of total aquatic-terrestrial interface is important to accurately upscale fluxes. However, shoreline loading and methods for measuring these loads may differ between lakes and streams and not all fluxes may be sensitive to shoreline extent, confounding this simple scaling approach. For example, particulate organic carbon loading, one flux often published as a shoreline-specific rate, has reported rates generally of similar magnitude for streams and lakes (Gasith & Hasler 1976; Webster & Meyer 1997). While lake and stream fluxes are measured using in-water open-top traps, the lateral fluxes in streams often include a lateral transport trap not used in lakes. Whole-stream transport for streams is sometimes inferred from small-scale lateral transport traps scaled across many kilometers of shoreline. Particulate organic carbon loading is a clear example of a flux that is sensitive to shoreline length, since the dominant source of this material is often shoreline vegetation (Malanson & Kupfer 1993). On the other hand, if materials entering a lake are derived from the catchment as a whole (e.g., nutrient loading in well-drained agricultural basins; Fraterrigo & Downing 2008), then shoreline length may be a poor scalar for loading rates. Finally, it may also be the case that real differences in ecosystem morphometry affect shoreline flux. Convolution on the scale of a few metres may not be significant while larger convolution on the scale of 100s of metres may be enough to affect overall transport magnitudes. Understanding the differences among ecosystems in terrestrial-aquatic fluxes and the relevant scales of shoreline complexity may be an important element in improving our overall estimate of carbon fluxes between inland waters and their surrounding landscapes.

The spatial distributions of lake abundance, area and perimeter have distinct differences despite similarities in general trends. Our spatially explicit maps of abundance and percent lake cover, which display the USGS National Hydrography Dataset data at a resolution of 50 km^2 as opposed to coarse aggregations based on ecoregion (McDonald *et al.* 2012), highlight fine-scale spatial distributions. Smith et al. (2002) reported a remarkably high density of lakes in the southern and central Great Plains region. The authors argued that such high densities probably resulted from human activities based on the absence of past lake-creating geological processes and are manifestations of both agricultural land use and negative water balance in these areas. A similar spatial distribution is apparent in the abundance and shoreline densities we present. The significance of lentic habitat creation is even more striking when viewed in terms of shoreline distribution. Pond and reservoir construction has created a broad mid-continental region with extensive shoreline and littoral habitats that were virtually non-existent before European settlement (Smith *et al.* 2002). In the spatial distribution of percent lake cover, this high density is not as apparent because lakes in the southern and central Great Plains region are small and contribute disproportionally more to perimeter and abundance than to area. The spatial distributions of lake area, perimeter and abundance follow general trends while deviating notably in certain regions due to significant biases in average lake size and potentially in shoreline convolution. This is especially relevant when considering the spatial distribution of processes that may scale with lake area or perimeter (e.g., CO_2 efflux versus shoreline carbon loading), knowing now that the distributions of area and perimeter are not equivalently distributed across lake size classes. We leave further exploration of the spatial distribution to the reader by way of the publically available dataset (http://lter.limnology.wisc.edu).

Differences in lake area to perimeter ratio across lake size classes is driven by two processes, the increase in SDF with increasing lake size and the purely geometric, non-linear scaling of perimeter with changing area of P_{min} ($P_{min} = 2\pi^{0.5}A^{0.5}$). Geometric scaling of a circular lake would tend to increase the area to perimeter ratio with increasing lake size, while increasing SDF with increasing lake size would tend to decrease the area to perimeter ratio. Despite observing increasing SDF with increasing lake size, the geometric scaling across lake size dominates over differences in SDF. The mean area to perimeter (P_{obs}) ratio in larger lakes (10-100 km^2) was 448, significantly higher than 31, the mean in smaller lakes (size range of 0.01-0.1 km^2). This difference occurs despite larger lakes having generally more convoluted shorelines (higher SDF) that would tend to decrease their area to perimeter ratio. Despite lake size dominating the range seen in perimeter to area ratio, we do not claim that shoreline complexity is irrelevant when considering differences between lakes. While there has been some work relating shoreline complexity to a variety of lake process and characteristics, such as fish populations (Johnson *et al.* 1977) and terrestrial input of nutrients (Gasith & Hasler 1976), further links may be found in the large, geographically distributed lake datasets becoming increasingly available (e.g., the U.S. Environmental Protection Agency National Lakes Assessment).

A pattern of decreasing total perimeter with decreasing measurement resolution is consistent with the idea that lake shorelines are fractal and that estimates of length will increase with increasing mapping resolution, as represented by shorter yardstick length. Unlike previous studies that focused on a small number of larger lakes (Kent & Wong 1982), we applied the yardstick method to the entire population, yielding an overarching view of lake perimeter sensitivity to measurement resolution. Strikingly, the large increase in total shoreline with decreasing measurement coarseness is influenced more by the inclusion of small lakes than by shoreline complexity. Coarser measurement resolution excluded small lakes that were not big enough to be measured. The contribution of small lakes to total shoreline is eliminated with measurement resolutions of 800 and 3200 m.

Our interpretation of the relationship between overall perimeter and measurement resolution can be improved by removing small lakes from the data set. When all lakes are included, the relationship between total perimeter and measurement resolution has a slope of -0.62, which indicates that doubling measurement resolution increases total perimeter by about 50%. When only larger size bins (i.e., bins in which no lakes are lost across measurement resolution) are used, then the sensitivity due to shoreline convolution alone can be examined. In this case the average slope is -0.2, which indicates that a doubling of measurement resolution increases total perimeter by 15%. This slope estimate is not confounded by the loss of small lakes with coarser measurement resolution and suggests that perimeter estimates for a given lake, on average, may not be as sensitive to measurement resolution as previously reported. It is difficult to compare these results to other published work because the measurement sensitivity of shoreline is not well represented in the literature. One such study concerning lakes found slopes ranging from -0.1 to -0.16, though the relationships between observed perimeter and resolution also contained a second linear relationship with steeper slopes (-0.27 to -0.64; Kent & Wong 1982). At the large scale, we did not find such a secondary relationship, though we did not exhaustively examine individual lakes, and our data covered a different region. Our results are probably more comparable to large-scale measurements of ocean shoreline complexity. For example, the Hausdorff dimension of the coast of Great Britain is reported as 1.25 (Mandelbrot 1967). To convert it to a log-log slope of perimeter versus resolution, we subtract the Hausdorff dimension from one, giving us a slope of -0.25, similar to the collective slope of the lake shoreline of U.S. contiguous states.

Our study suggests that small lakes contribute more to total aquatic-terrestrial interface than large lakes. We highlight spatial patterns and a distribution of lake perimeter that differs from the distribution of lake surface area and lake population. These findings change the paradigm of the relative roles of small versus large lakes for processes that may scale with perimeter as opposed to surface area, such as terrestrial carbon input. The total aquatic-terrestrial interface in lakes is less than one tenth that of streams when using comparable mapping products, highlighting the importance of streams when describing processes occurring at or across this boundary. While the use of these estimates to upscale ecosystem processes was beyond the scope of this paper, combining measurements of shoreline with estimates of shoreline fluxes and process may contribute to improving estimates of aquatic processes at continental scales.

ACKNOWLEDGEMENTS

We thank those organizations involved in the collection, preparation and publishing of the data used, specifically the United States Geological Survey and affiliate organizations. We also thank the two anonymous reviewers who provided very useful comments. High-throughput computing resources were provided by the HTCondor project at the University of Wisconsin-Madison. This work was supported by the National Science Foundation grants DEB-0941510 (Global Lake Ecological Observatory Network), EF-1065818, DEB-0822700 (North Temperate Lakes LTER program) and the Gordon and Betty Moore Foundation, award 1182.

CITATIONS

Battin T.J., Luyssaert S., Kaplan L. a., Aufdenkampe A.K., Richter A. & Tranvik L.J. (2009) The boundless carbon cycle. *Nature Geoscience* **2**, 598–600.

Bergström I., Mäkelä S., Kankaala P. & Kortelainen P. (2007) Methane efflux from littoral vegetation stands of southern boreal lakes: An upscaled regional estimate. *Atmospheric Environment* **41**, 339–351.

Buffam I., Turner M.G., Desai A.R., Hanson P.C., Rusak J. a., Lottig N.R., *et al.* (2011) Integrating aquatic and terrestrial components to construct a complete carbon budget for a north temperate lake district. *Global Change Biology* **17**, 1193–1211.

Cole J.J., Carpenter S.R., Kitchell J., Pace M.L., Solomon C.T. & Weidel B. (2011) Strong evidence for terrestrial support of zooplankton in small lakes based on stable isotopes of carbon, nitrogen, and hydrogen. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 1975–80.

Conners M.E. & Naiman R.J. (1984) Particulate Allochthonous Inputs: Relationships with Stream Size in an Undisturbed Watershed. *Canadian Journal of Fisheries and Aquatic Sciences* **41**, 1473–1484.

Dean W.E.W. & Gorham E. (1998) Magnitude and significance of carbon burial in lakes, reservoirs, and peatlands. *Geology* **26**, 535–538.

Downing J.A., Prairie Y.T., Cole J.J., Duarte C.M., Tranvik L.J., Striegl R.G., *et al.* (2006) The global abundance and size distribution of lakes, ponds, and impoundments. *Limnology and Oceanography* **51**, 2388–2397.

Drake M.T. & Pereira D.L. (2002) Development of a Fish-Based Index of Biotic Integrity for Small Inland Lakes in Central Minnesota Development of a Fish-Based Index of Biotic Integrity for Small Inland Lakes in Central Minnesota. *North American Journal of Fisheries Management* **22**, 1105–1123.

Francis T.B. & Schindler D.E. (2009) Shoreline urbanization reduces terrestrial insect subsidies to fishes in North American lakes. *Oikos* **118**, 1872–1882.

Fraterrigo J.M. & Downing J. a. (2008) The Influence of Land Use on Lake Nutrients Varies with Watershed Transport Capacity. *Ecosystems* **11**, 1021–1034.

Gasith a. & Hasler a. D. (1976) Airborne litterfall as a source of organic matter in lakes. *Limnology and Oceanography* **21**, 253–258.

Guy C.S. & Willis D.W. (2011) Population Characteristics of Black Crappies in South Dakota Waters: A Case for Ecosystem-Specific Management. *North American Journal of Fisheries Management* **15**, 754–765.

Hanson P.C., Carpenter S.R., Cardille J.A., Coe M.T. & Winslow L.A. (2007) Small lakes dominate a random sample of regional lake characteristics. *Freshwater Biology* **52**, 814–822.

Helmus M.R. & Sass G.G. (2008) The rapid effects of a whole-lake reduction of coarse woody debris on fish and benthic macroinvertebrates. *Freshwater Biology* **53**, 1423–1433.

Hershey A., Beaty S., Fortino K. & Kelly S. (2006) Stable isotope signatures of benthic invertebrates in arctic lakes indicate limited coupling to pelagic production. *Limnology and Oceanograph* **51**, 177–188.

Johnson M., Leach J., Minns C. & Olver C. (1977) Limnological Characteristics of Ontario Lakes in Relation to Associations of Walleye (Stizostedion vitreum vitreum), Northern Pike (Esox lucius), Lake Trout (Salvelinus namaycush), and Smallmouth Bass (Micropterus dolomieui). *Journal of the Fisheries Research Board of Canada* **34**, 1592–1601.

Kalff J. (2001) Lake and Catchment Morphometry. In: *Limnology: Inland Water Ecosystems*. pp. 85–93. Prentice Hall, Upper Saddle River, NJ.

Kalff J. (2002) *Limnology: Inland Water Ecosystems*. Prentice Hall, New Jersey.

Karlsson J., Christensen T.R., Crill P., Förster J., Hammarlund D., Jackowicz-Korczynski M., *et al*. (2010) Quantifying the relative importance of lake emissions in the carbon budget of a subarctic catchment. *Journal of Geophysical Research* **115**, 2005–2010.

Kent C. & Wong J. (1982) Index of Littoral Zone Complexity and Its Measurement. *Canadian Journal of Fisheries and Aquatic Sciences* **39**, 847–853.

Lewis W.M. (2011) Global primary production of lakes: 19th Baldi Memorial Lecture. *Inland Waters* **1**, 1–28.

Malanson G.P. & Kupfer J.A. (1993) Simulated fate of leaf litter and large woody debris at a riparian cutbank. *Canadian Journal of Forest Research* **23**, 582–590.

Mandelbrot B. (1967) How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension. *Science* **156**, 636–638.

Mandelbrot B.B. (1979) Fractals: form, chance and dimension. *Fractals: form, chance and dimension., by Mandelbrot, BB. San Francisco (CA, USA): WH Freeman & Co., 16+ 365 p.* **1**.

McClain M.E., Boyer E.W., Dent C.L., Gergel S.E., Grimm N.B., Groffman P.M., *et al*. (2003) Biogeochemical Hot Spots and Hot Moments at the Interface of Terrestrial and Aquatic Ecosystems. *Ecosystems* **6**, 301–312.

McDonald C.P., Rover J.A., Stets E.G. & Striegl R.G. (2012) The regional abundance and size distribution of lakes and reservoirs in the United States and implications for estimates of global lake extent. *Limnology and Oceanography* **57**, 1–12.

Preston N.D., Carpenter S.R., Cole J.J. & Pace M.L. (2008) Airborne carbon deposition on a remote forested lake. *Aquatic Sciences* **70**, 213–224.

Riera J.L., Magnuson J.J., Kratz T.K. & Webster K.E. (2000) A geomorphic template for the analysis of lake districts applied to the Northern Highland Lake District, Wisconsin, U.S.A. *Freshwater Biology* **43**, 301–318.

Scheuerell M.D. & Schindler D.E. (2004) Changes in the Spatial Distribution of Fishes in Lakes Along a Residential Development Gradient. *Ecosystems* **7**, 98–106.

Schindler D., Geib S. & Williams M. (2000) Patterns of fish growth along a residential development gradient in north temperate lakes. *Ecosystems* **3**, 229–237.

Schindler D.E. & Scheuerell M.D. (2002) Habitat coupling in lake ecosystems. *Oikos* **98**, 177–189.

Sharma P. & Byrne S. (2011) Comparison of Titan's north polar lakes with terrestrial analogs. *Geophysical Research Letters* **38**, 1–7.

Simley J. & Carswell J. (2009) *The national map– hydrography: U.S. Geological Survey Fact Sheet 2009-3054*.

Smith S., Renwick W.H., Bartley J. & Buddemeier R. (2002) Distribution and significance of small, artificial water bodies across the United States landscape. *The Science of The Total Environment* **299**, 21–36.

Strayer D.L. & Findlay S.E.G. (2010) Ecology of freshwater shore zones. *Aquatic Sciences* **72**, 127–163.

Thain D., Tannenbaum T. & Livny M. (2005) Distributed computing in practice: the Condor experience. *Concurrency and Computation: Practice and Experience* **17**, 323–356.

Tranvik L.J., Downing J.A., Cotner J.B., Loiselle S.A., Striegl R.G., Ballatore T.J., *et al.* (2009) Lakes and reservoirs as regulators of carbon cycling and climate. *Limnology and Oceanography* **54**, 2298–2314.

Vadeboncoeur Y., McIntyre P.B. & Vander Zanden M.J. (2011) Borders of Biodiversity: Life at the Edge of the World's Large Lakes. *BioScience* **61**, 526–537.

Vadeboncoeur Y., Peterson G., Vander Zanden M.J. & Kalff J. (2008) Benthic algal production across lake size gradients: interactions among morphometry, nutrients, and light. *Ecology* **89**, 2542–52.

Wang W., Yin C., Wang Y. & Lu J. (2006) Littoral zones as the “hotspots” of nitrous oxide (N2O) emission in a hyper-eutrophic lake in China. *Atmospheric Environment* **40**, 5522–5527.

Webster J. & Meyer J. (1997) Organic matter budgets for streams: a synthesis. *Journal of the North American Benthological Society* **16**, 141–161.

Wetzel R.G. (1990) Land-water interfaces: metabolic and limnological regulators. *Verh Internat Verein Limnol* **24**, 6–24.

Williamson C.E., Dodds W., Kratz T.K. & Palmer M.A. (2008) Lakes and streams as sentinels of environmental change in terrestrial and atmospheric processes. *Frontiers in Ecology and the Environment* **6**, 247–254.

TABLES

TABLE 1: SUMMARY STATISTICS OF LAKE NUMBER, AREA, PERCENT WATER COVER, DENSITY AND TOTAL SHORELINE LENGTH FOR THE US AND INDIVIDUAL STATES.

		Perimeter		Area				
State Name	Total Lake (#)	Total (km)	Median (m)	Total (km2)	Median (m2)	Percent Water	Lake Density (# km^{-2})	Shoreline Density (m km^{-2})
TOTAL	5,819,000	1,883,000	155	132,995	1410	1.65%	0.72	234
Alabama	96,000	41,000	191	2,648	2080	1.95%	0.71	305
Arizona	37,000	10,000	154	596	1460	0.20%	0.13	37
Arkansas	176,000	54,000	135	2,827	1160	2.05%	1.28	393
California	91,000	42,000	183	6,050	1810	1.43%	0.22	99
Colorado	112,000	27,000	144	1,203	1250	0.45%	0.42	103
Connecticut	20,000	6,000	142	290	1140	2.02%	1.42	445
Delaware	9,000	3,000	176	114	1580	1.77%	1.51	544
Florida	205,000	93,000	236	7,368	2910	4.33%	1.2	550
Georgia	129,000	61,000	294	2,986	4610	1.94%	0.84	399
Idaho	35,000	12,000	137	1,915	1110	0.88%	0.17	59
Illinois	157,000	48,000	153	1,603	1350	1.07%	1.05	320
Indiana	95,000	25,000	143	819	1210	0.87%	1.01	275
Iowa	115,000	27,000	145	759	1280	0.52%	0.79	189
Kansas	231,000	51,000	149	1,495	1340	0.70%	1.09	242
Kentucky	190,000	38,000	125	1,161	1010	1.11%	1.82	367
Louisiana	178,000	92,000	192	10,130	1910	7.47%	1.31	684
Maine	31,000	23,000	181	4,015	1660	4.38%	0.35	261
Maryland	37,000	10,000	157	355	1410	3.69%	3.89	1119
Massachusetts	27,000	11,000	157	686	1390	2.51%	0.99	408
Michigan	45,000	36,000	342	5,403	5150	2.16%	0.18	144
Minnesota	127,000	77,000	176	11,181	1870	4.97%	0.57	343
Mississippi	253,000	68,000	154	2,255	1450	1.80%	2.02	543
Missouri	330,000	67,000	121	2,157	970	1.19%	1.83	373
Montana	126,000	51,000	209	3,849	2160	1.01%	0.33	134
Nebraska	117,000	39,000	201	1,272	2000	0.63%	0.58	198
Nevada	14,000	9,000	144	2,137	1220	0.75%	0.05	32
New Hampshire	16,000	7,000	166	785	1470	3.24%	0.66	322
New Jersey	33,000	13,000	189	509	1780	2.25%	1.48	602
New Mexico	71,000	18,000	157	894	1510	0.28%	0.23	59
New York	81,000	34,000	163	4,324	1490	3.06%	0.58	245
North Carolina	94,000	32,000	189	1,721	1900	1.23%	0.68	232
North Dakota	176,000	74,000	246	4,697	3330	2.57%	0.96	408
Ohio	99,000	25,000	149	842	1270	0.73%	0.86	220
Oklahoma	387,000	88,000	138	3,775	1200	2.09%	2.14	486
Oregon	59,000	21,000	134	2,598	1060	1.02%	0.23	83
Pennsylvania	93,000	22,000	140	827	1170	0.69%	0.78	190
Rhode Island	4,000	1,000	159	102	1420	2.56%	1.12	468
South Carolina	46,000	21,000	230	1,765	2680	2.13%	0.56	262
South Dakota	150,000	62,000	208	4,218	2220	2.11%	0.76	312
Tennessee	126,000	37,000	119	2,757	920	2.53%	1.16	342
Texas	998,000	243,000	139	10,909	1150	1.57%	1.44	350
Utah	42,000	17,000	130	6,895	1050	3.14%	0.19	81
Vermont	12,000	3,000	135	265	1080	1.07%	0.5	160
Virginia	78,000	24,000	161	936	1470	0.84%	0.71	225
Washington	45,000	19,000	174	2,120	1640	1.15%	0.25	105
West Virginia	29,000	6,000	121	167	900	0.27%	0.47	98
Wisconsin	83,000	42,000	185	4,300	1730	2.53%	0.49	251
Wyoming	91,000	28,000	163	2,027	1500	0.80%	0.36	111

FIGURE CAPTIONS

Figure 1: Yardstick method for calculating the perimeter of a fractal shoreline. The "yardstick" (black) as it is walked along the polygon (gray). This example is part of the Lake Mendota (Wisconsin) polygon and uses a yardstick length of 200m.

Figure 2: Maps of (a) areal distribution of lake shoreline (shoreline m km^{-2}), (b) lake density (# lakes km^{-2}) and (c) percent of area covered by lakes (% lake cover).

Figure 3: The relative contribution of each size class to total Area, P_{obs} and $P_{min.}$, based on non-extrapolated USGS data.

Figure 4: The sensitivity of total shoreline length to the yardstick length used.

Figure 5: The sensitivity of total perimeter to measurement resolution across all lake size classes. Yardstick length indicated by line color.

Figure 6: Comparison of sensitivity of the perimeter estimate to measurement resolution of only those size-bins that lose no lakes across all yardstick lengths (lakes 9.8 km^2 and larger). The legend indicates the lower bound of each lake-size bin.

Figure 1:

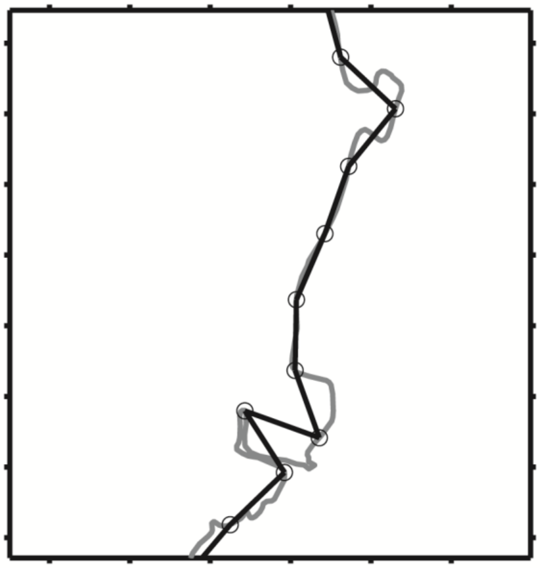


Figure 2:

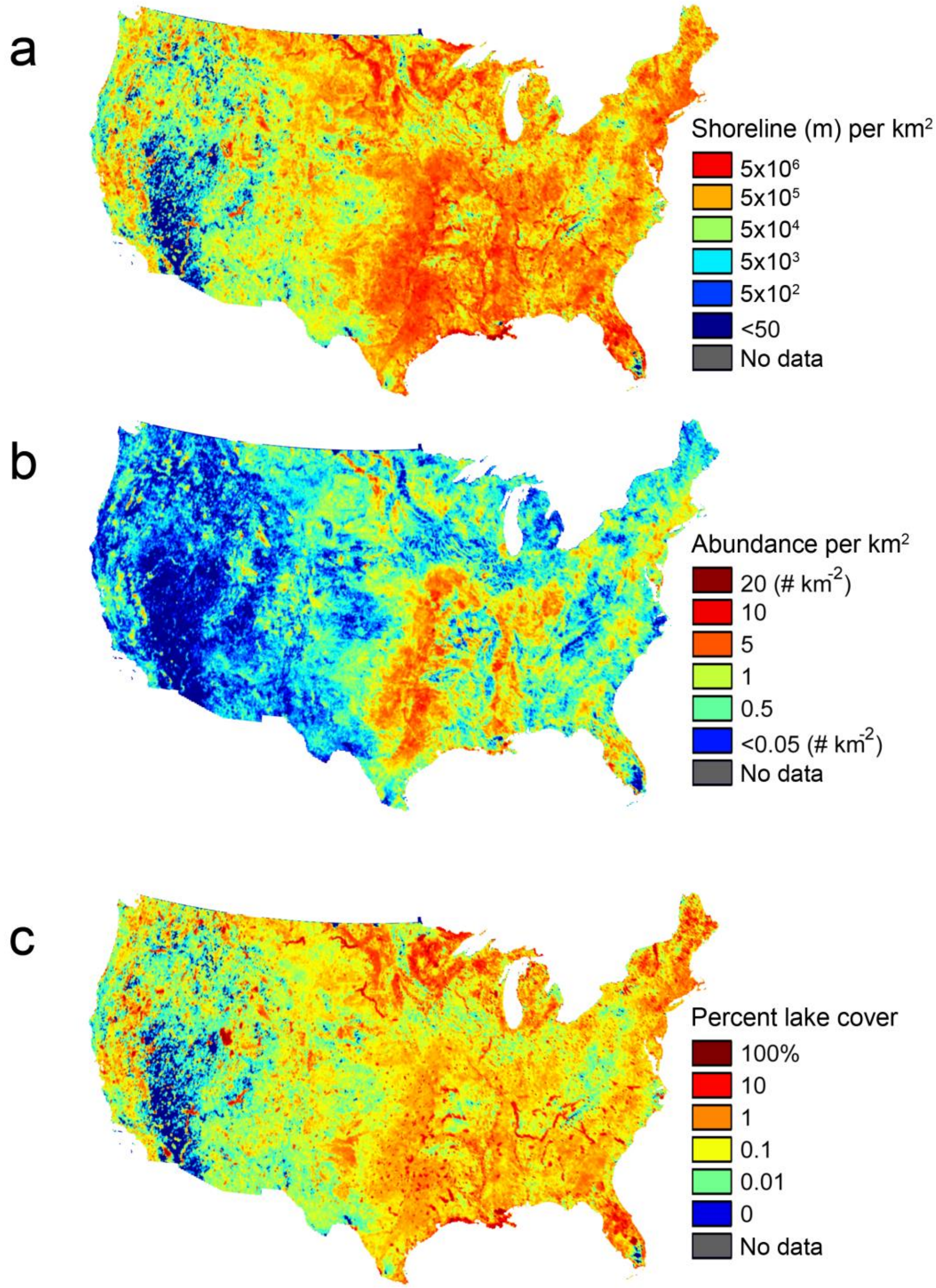

a
Shoreline (m) per km^2
5x10^6
5x10^5
5x10^4
5x10^3
5x10^2
<50
No data
b
Abundance per km^2
20 (# km^{-2})
10
5
1
0.5
<0.05 (# km^{-2})
No data
c
Percent lake cover
100%
10
1
0.1
0.01
0
No data

Figure 3:

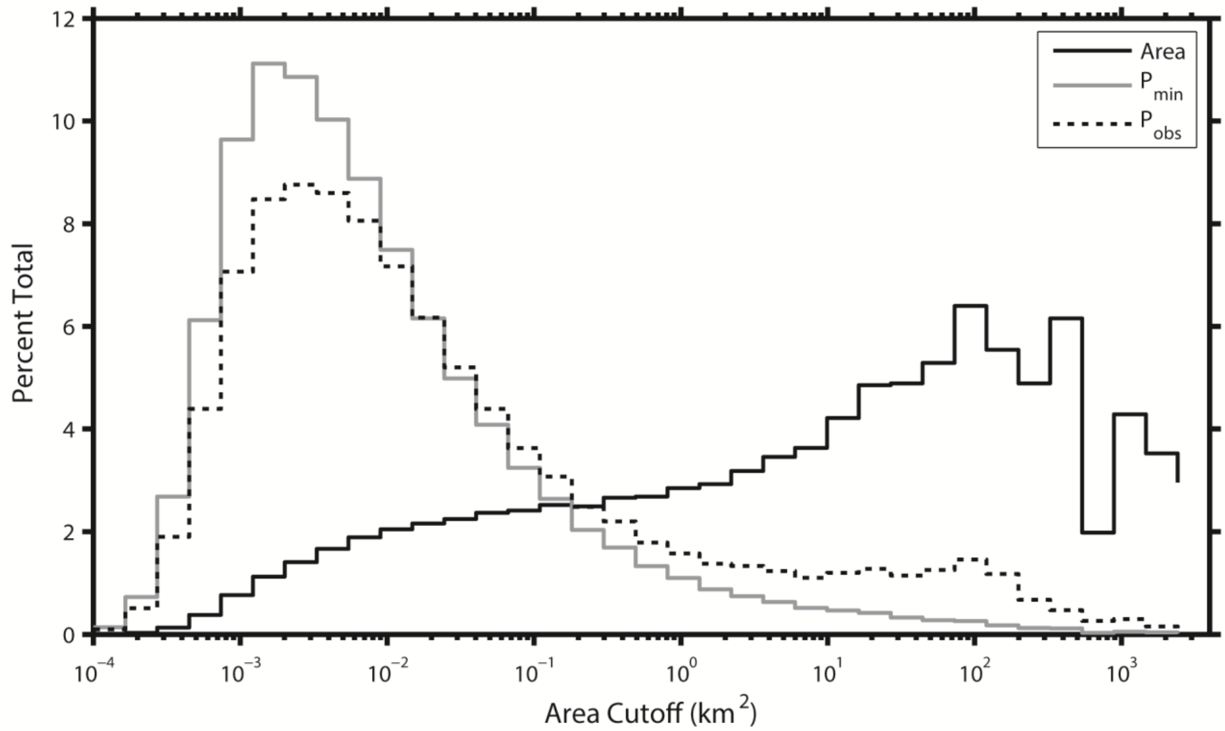
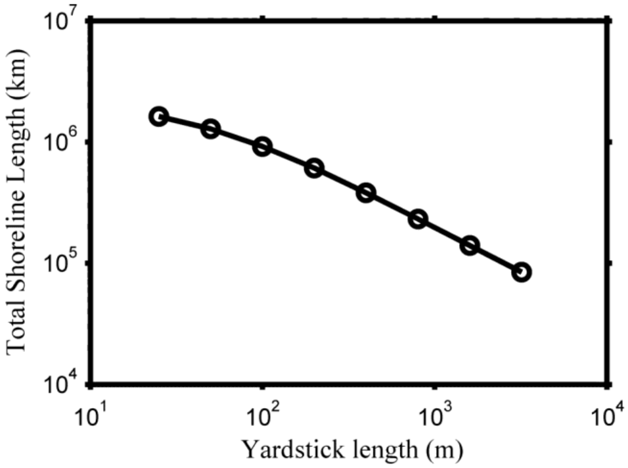
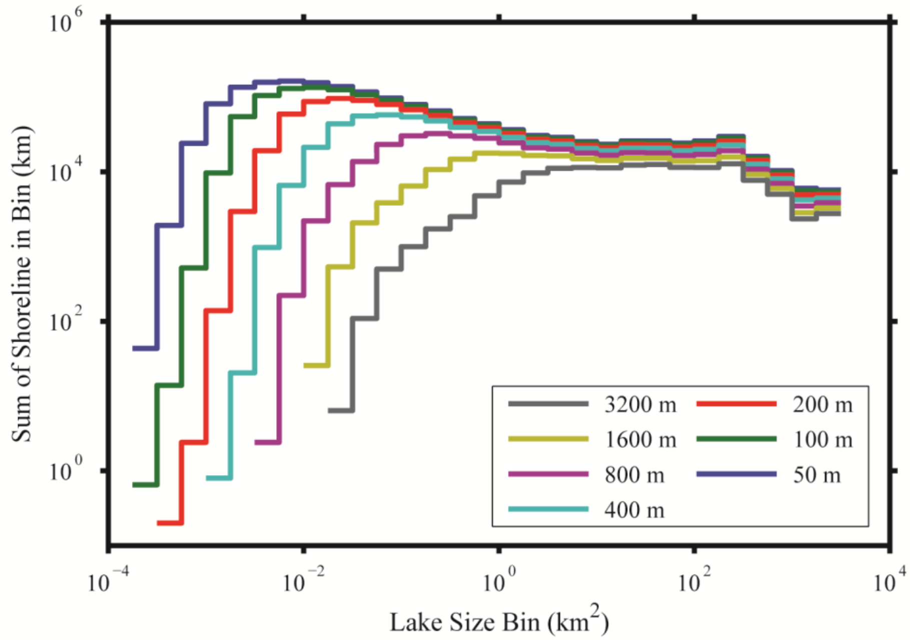

Area
P_{min}
P_{obs}
Percent Total
0
2
4
6
8
10
12
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
Area Cutoff (km^2)

Figure 4:



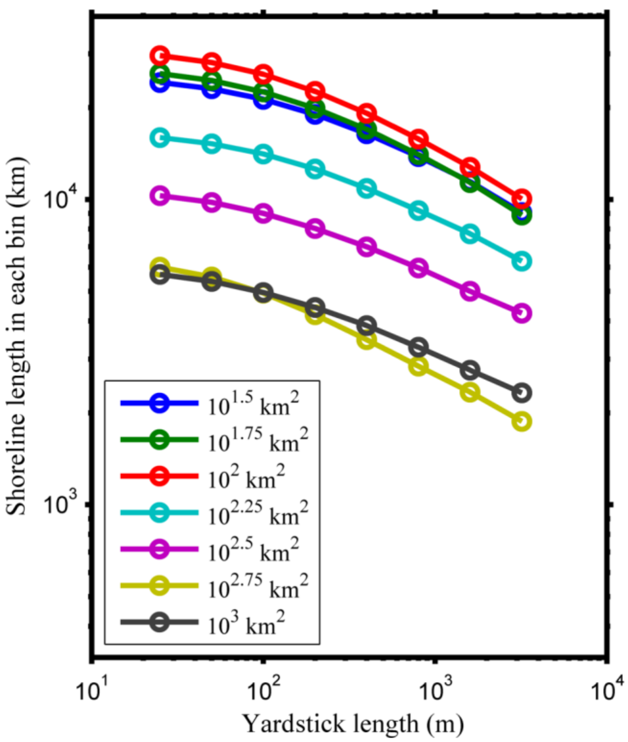
10^7
10^6
10^5
10^4
Total Shoreline Length (km)
10^1
10^2
10^3
10^4
Yardstick length (m)

Figure 5:



Sum of Shoreline in Bin (km)
Lake Size Bin (km²)
3200 m
1600 m
800 m
400 m
200 m
100 m
50 m

Figure 6:


Shoreline length in each bin (km)
10^4
10^3
$10^{1.5}$ km^2
$10^{1.75}$ km^2
10^2 km^2
$10^{2.25}$ km^2
$10^{2.5}$ km^2
$10^{2.75}$ km^2
10^3 km^2
10^1
10^2
10^3
10^4
Yardstick length (m)

CHAPTER 3 - A FRAMEWORK FOR COMBINING LAKE POPULATION MORPHOLOGY AND PROCESS TO BUILD LAKE POPULATION-SCALE MODELS

Luke A. Winslow[1*], Jordan S. Read[2], Paul C. Hanson[1], Emily H. Stanley[1]

Publication Journal: Inland Waters

1. University of Wisconsin - Madison, Center for Limnology, 680 N. Park Street, Madison, Wisconsin 53706 USA
2. U.S. Geological Survey, Center for Integrated Data Analytics, 8505 Research Way, Middleton, WI 53562 USA

* Corresponding Author: Email lawinslow@wisc.edu

Abstract

With lake abundances in the thousands to millions, creating an intuitive understanding of the distribution of morphology and processes in lakes is challenging. To improve researchers' understanding of large-scale lake processes, we developed a parsimonious mathematical model based on the Pareto distribution to describe the distribution of lake morphology (area, perimeter and volume). While debate continues over which mathematical representation best fits any one distribution of lake morphometric characteristics, we recognize the need for a simple, flexible model to advance understanding of how the interaction between morphometry and function dictates scaling across large populations of lakes. These models make clear the relative contribution of lakes to the total amount of lake surface area, volume, and perimeter. They also highlight the critical thresholds at which total perimeter, area, and volume would be evenly distributed across lake size-classes having Pareto slopes of 0.63, 1, and 1.12, respectively. These models of morphology can be used in combination with models of process to create overarching "lake population" level models of process. To illustrate this potential, we combine the model of surface area distribution with a model of carbon mass accumulation rate. We found that even if smaller lakes contribute relatively less to total surface area than larger lakes, the increasing carbon accumulation rate with decreasing lake size is strong enough to bias the distribution of carbon mass accumulation towards smaller lakes. This analytical framework provides a relatively simple approach to upscaling morphology and process that is easily generalizable to other ecosystem processes.

Key Words: Power-laws, Lake Morphology, Upscaling, Small Lakes, Macrosystems

INTRODUCTION

There is growing interest in better understanding the role of inland waters in carbon and nutrient cycles at broad scales (Bennett, Carpenter, and Caraco 2001; Cole et al. 2007; Harrison et al. 2008; Tranvik et al. 2009). To develop science incorporating lakes into large-scale cycles, some have argued for increased efforts in the field of global limnology, defined as “quantifying and understanding the role of continental waters in the functioning of the biosphere” (Downing 2009). A key challenge for this rapidly evolving research arena is discovering and understanding regular patterns in process rates across aquatic ecosystems in order to facilitate upscaling.

An emergent lesson from global limnology is that lake size matters. Lakes with surface areas that differ by orders of magnitude (which we describe here as lakes in different “size-classes”) sometimes have substantially different area-normalized process rates. Gas exchange (Read et al. 2012) and organic carbon burial (Downing et al. 2008; Kastowski, Hinderer, and Vecsei 2011) are two examples of processes with rates predicted in part by lake size. Such process with rates tied to lake area can be especially amenable for upscaling because unlike variables that cannot be remotely-sensed, the size and abundance are known with reasonable certainty for all but the smallest of lakes (McDonald et al. 2012). To upscale process estimates, models linking lake process with lake size are often combined with empirical lake size-abundance distributions. However, information on the number of small lakes is often missing. These gaps may represent a substantial source of error, as small water bodies can have particularly high rates for some processes, such as C storage or efflux (Downing et al. 2008; Read et al. 2012). To fill in such gaps, lake size-abundance models based on the Pareto distribution have been used to extrapolate unobserved small lake size-class abundances (e.g., Downing et al. 2008; Kastowski, Hinderer, and Vecsei 2011; Lewis 2011).

There are additional applications of lake size-abundance models beyond filling gaps in observation. Specifically, these models greatly simplify large and cumbersome datasets that contain information on many thousands (Lehner and Doll 2004) to millions (Downing et al. 2006) of lakes. When such a model approximates the full population to within a desired level of accuracy, the simplified mathematical form provides an easily manipulated representation, compared to a large empirical dataset, of the population and its key characteristics. At the first level, the mathematical form of the model can be modified to describe the relative contribution of key morphological characteristics (lake area, perimeter, and volume) of different lake size-classes. We refer to these models describing the relative contribution of different lake sizes to area, perimeter, and volume “morphology scaling relationships” (MSRs). These MSRs can provide a convenient and powerful mathematical representation of key components of the hydrosphere. MSRs may be combined directly with models of process to scale process to the full population of lakes. This concept is similar to the use of large-scale steady-state approximations of ocean dynamics to communicate key physical phenomenon in oceans (Brown et al. 1989). Such steady-state approximations are not directly used in modeling quantitative ocean process, but are useful in communicating and understanding important phenomenon (e.g., Ekman transport and the Sverdrup balance). Similar simplified models for lakes may offer a new and unique opportunity for understanding the collective behavior of these aquatic ecosystems at continental and global scales.

In this paper, we examine the applicability of the Pareto distribution as a simplified model of the lake size-abundance relationship within the continental United States. We show how additional models can be derived from the lake size-abundance distribution to describe, with minimal error, the distribution of morphology, as perimeter and volume, across almost the entire size range of lakes to create MSRs. Finally, we use a published model of carbon mass accumulation rate to show how MSRs can be combined with models of process to create simple models describing process for the whole population of U.S. lakes, though any large population of lakes could be used. We use these morphology and process models to answer the following questions. How do lake size and abundance combine to create MSRs of lake area, perimeter, and volume? What are the critical parameters in the lake size-abundance distribution that would modify our understanding of the relative contribution of small versus large lakes to total estimates of global aquatic morphology? At what rate does process need to scale with lake size in order to change our view regarding the contribution of small versus large lake size-classes to total continental-scale processes?

METHODS

Empirical Relationships

For an empirical lake population dataset, we used the United States Geological Survey's (USGS) National Hydrography Dataset (NHD; retrieved January 2013, http://nhd.usgs.gov). We used the data derived from high-resolution USGS topographical maps (1:24,000-scale) (Simley and Carswell 2009) and excluded Alaska and Hawaii due to the differing resolution of data available. While the dataset covers all 48 continental U.S. states, the Laurentian Great Lakes were not included in our analysis as their low number substantially reduces the applicability of population-level process simplifications and estimates.

From the NHD, we extracted all lake, reservoir, and pond polygons. State boundaries had overlapping coverage so we discarded identical polygons as duplicates. For all polygons, area and perimeter were calculated using the Mathworks Mapping Toolbox functionality (v2011a; http://mathworks.com). To avoid issues with missing or unobserved small lakes, only those with surface areas greater than 0.01 km^2 (1 hectare) were included here. For a more detailed analysis and geographical visualization of the NHD data, see (McDonald et al. 2012) and (Winslow et al. 2014).

Previous descriptions of lake size-abundance have used a Pareto distribution (Downing et al. 2006). The Pareto distribution

$$pdf\,(A) = \alpha x_m^{\alpha} A^{-(\alpha+1)} \quad (1)$$

is defined by two parameters, the scale parameter (x_m), and the shape parameter (α). x_m defines the minimum variable value of A, the values of the population of interest (in our case, A is lake area) and α is the exponent of the power law. To estimate α, we used the maximum likelihood estimator (MLE)

$$\hat{\alpha} = \frac{n}{\sum_{1}^{n} \ln \frac{X_i}{x_m}} \quad (2)$$

where n is the number of observations, $\hat{\alpha}$ is the maximum likelihood estimate for the population's α, x_m is the minimum value from the population of interest and X_i is the examined population variable, in this case, lake area (equation from Rytgaard 1990). Lastly, we used the NHD dataset to calculate a relationship between lake area and perimeter. The relationship was fitted using a non-linear, least-squares exponential fit of area versus perimeter to estimate the exponential relationship parameters.

Extending the Pareto distribution

We can use the Pareto probability density function (Eq. 1) to derive a number of functions that help highlight the relative contribution of different lake size-classes to the global distribution of lake area, volume, and perimeter, as well as biogeochemical processes that scale with these morphological parameters. When used as a size-abundance model, the Pareto probability density function gives the fraction of total lakes at a given size lake, which can be described roughly as:

$$pdf(A) \propto n(A) \quad (3)$$

where *n* is the number of lakes for a given lake area, *A*. The fraction of lakes in each size-class multiplied by the area of that class (*n***A*) gives us the relative total area contributed by that size-class, analogous to weighting the probability density function by lake area. Because of its relative ease of observation, lake area is the most commonly used parameter for large-scale lacustrine biogeochemistry estimates. To derive an equation for the lake area-distribution across lake size-classes, we multiply the size-abundance equation for the Pareto probability distribution by area and integrate:

$$A_{dens}(A) = \int \alpha x_m^{\alpha} A^{-(\alpha+1)} * A \; dA \quad (4)$$

The result gives an equation we can use to evaluate the relative contribution to total lake area (A_{dens}) across different lake size-classes. Because we want to focus on the relative contributions of different lake size-classes and emphasize simplicity, we combine all constant terms into a single term, *C*:

$$A_{dens} = CA^{1-\alpha} \quad (5)$$

Combining the constant terms into *C* greatly improves the presentation of relationship and helps highlight the important, area-dependent terms. For comparing process contribution of differently sized lakes, the absolute magnitude of the function is not important. Rather, how it scales across lake size-classes is the critical attribute.

With a few modifications, the same methods are applicable to lake perimeter. Because the distribution is formulated from the observed relationship between lake abundance and surface area, we first need a relationship between area and perimeter. We are unaware of any published empirical relationships between perimeter and surface area, so we make the most conservative assumption possible: that a circle with the given area adequately represents lake perimeter, which represents the lowest bound for perimeter of a given area. This relationship can easily be derived from the equations of a circle:

$$P = 2(\pi A)^{½} \quad (6)$$

where P is perimeter. To get the perimeter distribution across lake sizes, we combine the equation of a circle's perimeter (Eq. 6) with that of the Pareto distribution (Eq. 1) and again combine all constants into the term C. Integrating, we get:

$$P_{dens}(A) = CA^{½\text{-}\alpha} \quad (7)$$

where all constants are again subsumed into the C coefficient.

Lastly, while volume predicted from area alone results in relatively high uncertainty (±57% relative standard deviation of predicted versus observed volume (Sobek, Nisell, and Fölster 2011)), creating a similar model for volume is useful to contrast with area and perimeter distributions. To formulate the equation for volume distribution across lake sizes, we need to substitute in a published relationship between area and volume into Eq. 5. Because we are unaware of any published relationships for the U.S., we use a relationship based on lakes in Sweden ($V \sim A^{1.12}$) (Sobek, Nisell, and Fölster 2011). Substituting into Eq. 5, we get:

$$V_{dens}(A) = CA^{1.12-\alpha} \quad (8)$$

The derived MSRs for perimeter and volume were compared with observations by comparing the predicted fraction with the empirical distribution derived from the NHD continental U.S. dataset. The observations summed into decadal lake size bins (e.g., 1 to 10 km^2) and the model results and residuals were plotted for comparison.

RESULTS AND DISCUSSION

The Pareto distribution and simple scaling laws can help us understand how lakes of different sizes contribute to total perimeter, area, and volume. For example, while it is unclear whether small lakes contribute more than large lakes to the total surface area globally (Downing et al. 2006; McDonald et al. 2012), using the Pareto distribution and some simple calculus, we can illustrate how sensitive our inferences about total surface area are to our estimate of α. One can think of the distribution of lake surface area as a balance between decreasing area and increasing abundance with decreasing area. If lake abundance increases faster than area decreases, then total surface area in each size-class will increase (Figure 1). Conversely, if lake abundance does not increase quickly enough, then total surface area in each size-class will decrease, substantially altering our understanding of small versus large lake roles. Eq. 5 illustrates this tradeoff.

Using Eq. 5, it can be easily demonstrated how our perspective on the contribution of small versus large lakes can depend on the value of α. The critical threshold for area is when α equals 1, resulting in an exponent of zero. An exponent of zero makes the contribution to total area by each size-class unrelated to area. Knowing this critical threshold helps explain the different results found by past studies of the examined contribution of total lake surface area by lake size-class. Past work has calculated $\alpha = 1.06$ empirically from larger lakes of the globe (Downing et al. 2006), while others have found $\alpha = 0.92$ for a more complete size range of the continental U.S. lake population (McDonald et al. 2012). Despite the two estimates of α being of seemingly similar magnitude, they fall on opposite sides of the critical α cutoff. If α is less than 1, as reported by McDonald et al. and also found here ($\alpha = 0.92$, Figure 2), the exponent of area in Eq. 5 is positive and larger lakes make an increasing contribution to the total surface area of lakes. If α is greater than 1, the exponent is negative, resulting in a decreasing contribution to global surface area of lakes with increasing area. The deviation from linearity in the large lakes likely represents the edge effects of the continent, where a large lake has a higher likelihood of intersecting the continental edge and therefore not forming a lake (Goodchild 1988).

The perimeter MSR is described by Eq. 7, yielding a critical threshold for an even distribution of perimeter across all size-classes of $\alpha = 0.5$ when lake shape is simplified to a circle. An $\alpha = 0.5$ would produce a zero exponent for the area term and thus a constant perimeter density across all size-classes. In studies that have estimated values of α (Hamilton et al. 1992; Downing et al. 2006; Kastowski, Hinderer, and Vecsei 2011; McDonald et al. 2012) as well as here, α estimates are consistently >> 0.5, suggesting that the distribution of perimeter is skewed strongly towards small lakes. The contribution of any processes that scale proportionally to perimeter, such as particulate organic carbon (POC) import (Gasith and Hasler 1976), would be skewed towards small, rather than large lakes.

Using an empirical relationship between area and lake perimeter instead of a circular lake assumption can improve accuracy of the perimeter MSR. The non-linear least-squares exponential fit of area versus perimeter for continental U.S. lakes gave a slope of 0.63 (Figure 3). This result suggests large lakes have, on average, higher perimeters relative to their areas than would be predicted if lakes had a constant geometrically-scaled proportion of area to perimeter. Compared to a circular-lake assumption, steeper area to perimeter slope reduces the skew of perimeter towards small lakes, though the difference is not large enough to change the small-lake skew of perimeter. The U.S. NHD lakes-based MSR of perimeter would be:

$$P_{dens}(A) = CA^{0.63-\alpha} \quad (9)$$

Because there is no published relationship for lake area to volume for the contiguous U.S., we used the published relationship based on Swedish lakes for demonstration (Sobek, Nisell, and Fölster 2011). Using the published model for Swedish lakes, the critical threshold for an even distribution of volume across all size-classes would be α = 1.12, higher than previous α estimates. The α values reported for Eq. 8 for large collections of lakes are consistently < 1.12, which demonstrates that volume is likely skewed strongly towards larger lakes. To improve our mathematical model for the continuous U.S. and other beyond, future work examining area-volume relationships of other geographic regions is required.

For lake area and perimeter, the estimates made by the MSRs can be compared to a large-scale empirical lake distribution. Like any model, the desired accuracy is dependent on the scope and application. As in this case, when the geographic scope is very large (continental U.S.), even models that make predictions to within an order of magnitude may be useful. For each size-class, we compared the model results for contribution of each size-class to area and perimeter with the empirical results (Figure 4). For all size-classes between 0.01 km^2 and 1000 km^2, the estimates of area were within 4% of the empirical values. For perimeter, all bins were within 7% of the empirical measurements. Caution should be taken when extrapolating these models below lakes of 0.01 km^2 as it is unclear if the power-law model accurately describes abundance below that size-class (McDonald et al. 2012).

The distributions of area, perimeter, and volume strongly differ across size-classes and the understanding of how area is distributed is sensitive to the estimate of α (Figure 5). The differences between distributions of key lake morphological features suggest that some processes are likely to scale differently across lake size. This difference leads to shifts in the balance among processes in lakes of different size-class. Such shifts in process balance with lake area have been hypothesized to occur where processes such as primary and secondary production (scaling with surface area) interact with the lateral import and export of materials (scaling with perimeter) (Gasith and Hasler 1976; Vander Zanden and Gratton 2011). These differences in lake morphology distributions may have serious implications for how key processes in large-scale carbon cycling are distributed across lakes of different sizes. For example, if carbon import scales with perimeter while CO_2 evasion scales with area, is carbon import from the surrounding terrestrial ecosystem skewed towards small lakes while export to the atmosphere is skewed towards larger lakes?

The direct use of the Pareto distribution can be helpful in understanding how processes scale across lake sizes. By combining Eq. 5 and a relationship of process rate with lake area (an area-rate relationship), we can create a combined, overarching model with a simple form that estimates the relative contribution of each lake size-class to total process. Processes that have a power-function lake area-rate relationship are often represented in the literature as a log-transformed linear relationship. Such an area-rate relationship takes the form:

$$F = F_o A^{\beta} \quad (10)$$

Where F is the process rate with units dependent on the process being described, F_o is a linear scaling parameter (i.e., the intercept), A is lake area, and β is the parameter that scales the process rate with lake area. Care should be taken when fitting the β parameter. Non-linear fitting techniques, as opposed to log transformed linear regression which often distorts error, tend to be more robust and should be favored (Motulsky and Ransnas 1987). Because we want to examine how the process is distributed across lakes of different sizes, we focus on the exponent of lake area, β. This exponent, when negative, indicates an increasing process rate with decreasing lake area. Combining this area-rate relationship (Eq. 10) with the MSR for area (Eq. 5) yields a function that scales process across lake sizes:

$$F_{dens}(A) = CF_o A^{1-\alpha+\beta} \quad (11)$$

The key result of Eq. 11 comes from the exponent of area, which indicates the direction of areal skew in the process (towards or away from large lakes). To re-iterate the key point, a positive exponent would result in larger lakes making a larger relative contribution. A negative exponent would indicate that small lakes contribute more to the overall process magnitude. An exponent at or very near zero would suggest no scaling with area and thus an equal contribution of all lake size-classes.

What does our current understanding of the lake size-abundance distribution mean for the distribution of process across lake size-classes? The most recently published lake size-abundance slope parameter estimate is $\alpha = 0.92$ (McDonald et al. 2012). As discussed previously, $\alpha < 1$ means smaller lakes contribute less to total surface area than larger lakes. Despite this, process rates do not have to scale strongly with lake size to have an equal contribution across all lake sizes. For a given process to have an equal contribution across all lake sizes, the process scaling parameter (β) would only need to be -0.08, which would make the exponent of area zero (from Eq. 11: 1 - 0.92 + 0.08). To put this finding into perspective, $\beta = -0.08$ would imply that the process rate in a 0.01 km^2 lake is ~2 times higher than in a 100 km^2 lake. If $\beta < -0.08$, small lakes would contribute a larger fraction to the overall process than larger lakes.

A published example of such a process is carbon mass accumulation rate (CMAR) in European lakes (Kastowski, Hinderer, and Vecsei 2011). This work has found that organic carbon burial rates correlated with area, watershed slope, and percent cropland cover (Kastowski, Hinderer, and Vecsei 2011). The relationship was described by the equation:

$$\ln(CMAR) = 1.00 - 0.217 * \ln(lake\ \ area) + 0.194 * (slope) + 0.017 * (cropland\%) \qquad (12)$$

This equation shows that ln-CMAR varied with log-lake area with a slope of -0.217. If we use this β for CMAR and a published size-abundance slope (α = 0.92) in Eq. 11, we get a scaling equation for CMAR:

$$CMAR = CA^{-0.177} \quad (13)$$

The resulting equation (Eq. 13) has a negative area exponent, meaning that CMAR over the distribution of lakes is skewed towards smaller lakes (Figure 6). With α = 0.92, the further β is from -0.08, the more skewed towards large or small lakes the distribution of the process becomes. Using a relationship between lake area and carbon burial with β = -0.298 calculated from small eutrophic lakes in Iowa, U.S.A. (Downing et al. 2008), the distribution becomes even more strongly skewed toward smaller lakes (Figure 6). This skew suggests that despite being generally less studied than large lakes (Downing 2010), understanding small lakes is important to estimating the large-scale storage of carbon in the world's lakes. Unfortunately, carbon sedimentation in small lakes is highly variable, with commonly cited studies reporting maximum observed rates ranging over several orders of magnitude, from around 280 gC m^{-2} y^{-1} (Mulholland and Elwood 1982) to as high as 10,000 gC m^{-2} y^{-1} (Downing et al. 2008). The importance of small lakes, combined with high rate uncertainty and high abundance, will require novel research and ideas in the future to constrain large-scale carbon storage in small lakes.

CONCLUSIONS

This modeling framework represents a unique and simple approach to describe distributions of morphology and process across a population of lakes. While the quantification of morphology and process will always require detailed computational work using empirical data, this mathematical approach helps researchers form a more intuitive understanding of large populations of lakes while maintaining some quantitative aspects. With these equations and the calculated and published estimates of the relationship between area, perimeter, and volume, the critical thresholds for the Pareto slope where morphological characteristics would be evenly distributed across lake size-classes are 1 for area and 0.63 for perimeter. The volume relationship has the highest uncertainty and when better estimates are available, the equation for volume should be updated. Despite this uncertainty, using the available published relationship for demonstrative purposes (Sobek, Nisell, and Fölster 2011) results in a critical Pareto slope threshold of 1.12 for volume. Given a Pareto slope parameter of between 1.06 and 0.92, it is likely that larger lakes contribute relatively less to total perimeter and more to total volume (compared to the contribution from small lakes). The skew of the area distribution is dependent on the Pareto slope parameter, though the more recent estimate of 0.92, derived from a dataset spanning a large range of lake sizes, suggests smaller lakes contribute a decreasing fraction to total surface area. Despite the likely skew of surface area towards larger lakes, processes can offset the skew by having a process rate that increases with decreasing lake size, requiring a critical exponential slope with lake size of only -0.08 or less. The carbon mass accumulation is one such example, with published exponential slopes from -0.217 to -0.298, which results in a skew of carbon accumulation contribution towards smaller lakes. These simple mathematical tools will help bring quantitative, global limnological thinking to a broader group of students and researchers.

Acknowledgements

We would like to thank those who provided feedback throughout the manuscript preparation process, especially Emily Read and Kevin Rose. This work was supported by the National Science Foundation grants DEB-0941510 (Global Lake Ecological Observatory Network), EF-1065818, and and the Gordon and Betty Moore Foundation, award 1182. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

REFERENCES

Bennett EM, Carpenter SR, Caraco NF. 2001. Human impact on erodable phosphorus and eutrophication: a global perspective. Bioscience 51:227–234.

Brown E, Colling A, Park D, Phillips J, Rothery D, Wright J. 1989. Ocean Circulation. (Keynes M, editor.). Oxford: Pergamon Press.

Cole JJ, Prairie YT, Caraco NF, McDowell WH, Tranvik LJ, Striegl RG, Duarte CM, Kortelainen P, Downing J a., Middelburg JJ, et al. 2007. Plumbing the Global Carbon Cycle: Integrating Inland Waters into the Terrestrial Carbon Budget. Ecosystems 10:172–185.

Downing J a., Cole JJ, Middelburg JJ, Striegl RG, Duarte CM, Kortelainen P, Prairie YT, Laube K a. 2008. Sediment organic carbon burial in agriculturally eutrophic impoundments over the last century. Global Biogeochem. Cycles 22:1–10.

Downing JA, Prairie YT, Cole JJ, Duarte CM, Tranvik LJ, Striegl RG, McDowell WH, Kortelainen P, Caraco NF, Melack JM. 2006. The global abundance and size distribution of lakes, ponds, and impoundments. Limnol. Oceanogr. 51:2388–2397.

Downing JA. 2009. Global limnology: up-scaling aquatic services and processes to planet Earth. Verh. Internat. Verein. Limnol. 30.

Downing JA. 2010. Emerging global role of small lakes and ponds: little things mean a lot. Limnetica 1:9–24.

Gasith a., Hasler a. D. 1976. Airborne litterfall as a source of organic matter in lakes. Limnol. Oceanogr. 21:253–258.

Goodchild MF. 1988. Lakes on Fractal Surfaces: A Null Hypothesis for Lake-Rich Landscapes. Math. Geol. 20:615–630.

Hamilton SSK, Melack JJM, Goodchild M, Lewis WM. 1992. Estimation of the fractal dimension of terrain from lake size distributions. In: Lowland Floodplain Rivers: Geomorphological Perspectives. p. 145–163.

Harrison J a., Maranger RJ, Alexander RB, Giblin AE, Jacinthe P-A, Mayorga E, Seitzinger SP, Sobota DJ, Wollheim WM. 2008. The regional and global significance of nitrogen removal in lakes and reservoirs. Biogeochemistry 93:143–157.

Kastowski M, Hinderer M, Vecsei A. 2011. Long-term carbon burial in European lakes: Analysis and estimate. Global Biogeochem. Cycles 25:1–12.

Lehner B, Doll P. 2004. Development and validation of a global database of lakes, reservoirs and wetlands. J. Hydrol. 296:1–22.

Lewis WM. 2011. Global primary production of lakes: 19th Baldi Memorial Lecture. Inl. Waters 1:1–28.

McDonald CP, Rover JA, Stets EG, Striegl RG. 2012. The regional abundance and size distribution of lakes and reservoirs in the United States and implications for estimates of global lake extent. Limnol. Oceanogr. 57:1–12.

Motulsky H, Ransnas L. 1987. Fitting curves to data using nonlinear regression: a practical and nonmathematical review. FASEB J. 1:365–374.

Mulholland P, Elwood J. 1982. The role of lake and reservoir sediments as sinks in the perturbed global carbon cycle. Tellus 34:490–499.

Read JS, Hamilton DP, Desai AR, Rose KC, MacIntyre S, Lenters JD, Smyth RL, Hanson PC, Cole JJ, Staehr P a., et al. 2012. Lake-size dependency of wind shear and convection as controls on gas exchange. Geophys. Res. Lett. 39:L09405.

Rytgaard M. 1990. Estimation in the Pareto distribution. Astin Bull. 20.

Simley J, Carswell J. 2009. The national map– hydrography: U.S. Geological Survey Fact Sheet 2009-3054.

Sobek S, Nisell J, Fölster J. 2011. Predicting the volume and depth of lakes from map-derived parameters. Inl. Waters 1:177–184.

Tranvik LJ, Downing JA, Cotner JB, Loiselle SA, Striegl RG, Ballatore TJ, Dillon P, Finlay K, Fortino K, Knoll LB, et al. 2009. Lakes and reservoirs as regulators of carbon cycling and climate. Limnol. Oceanogr. 54:2298–2314.

Winslow LA, Read JS, Hanson PC, Stanley EH. 2014. Lake shoreline in the contiguous United States: quantity, distribution and sensitivity to observation resolution. Freshw. Biol. 59:213–223.

Vander Zanden J, Gratton C. 2011. Blowin' in the wind: reciprocal airborne carbon fluxes between lakes and land. Can. J. Fish. Aquat. Sci. 68:170–182.

FIGURE LEGENDS

Figure 1. Lake size and abundance tradeoff.

The tradeoff between decreasing size and increasing abundance in lake distributions can have two results, an overall decreasing contribution to total lake area with smaller size or an increasing contribution. (a) Shows the size-abundance distribution for two different published Pareto slope parameters, both showing increasing abundance with decreasing size. (b) Shows that while the size-abundance distributions themselves may not appear strikingly different, the result is two different views on the relative contribution of small versus large lakes.

Figure 2. Maximum likelihood size-abundance Pareto fit

NHD lakes size-abundance distribution with the maximum likelihood fit of the Pareto distribution.

Figure 3.Power-law lake area to perimeter relationship.

The power-law area to perimeter relationship of lakes in the continental U.S. A non-linear exponential fit is shown with slope 0.63, close to the theoretical relationship defined by the relationship between the area and perimeter of a circle (slope 0.5).

Figure 4. Comparison of lake morphology model to empirical dataset.

Comparison of the observed continental U.S. dataset with the Pareto-based model. The difference between observed and modeled for any bin of area and perimeter did not exceed 4% and 7% respectively.

Figure 5. Lake size-class contributions to total area, perimeter and volume.

Percent contribution of total for each lake size-class to total lake area, perimeter, and volume based on equations 7-9 for the range of published α values. (1.06 from Downing et al. (Downing et al. 2006) and 0.85 from Mcdonald et al. (McDonald et al. 2012)).

Figure 6. Percent contribution to total C Burial for each lake size-class.

Percent contribution to total C Burial for each lake size-class. Gray bars are calculated using β = -0.217 estimated from European lakes (Kastowski, Hinderer, and Vecsei 2011). Black bars were calculated using β = -0.298 estimated from eutrophic lakes in Iowa, U.S.A. (Downing et al. 2008).

Figures and Tables

Figure 1:

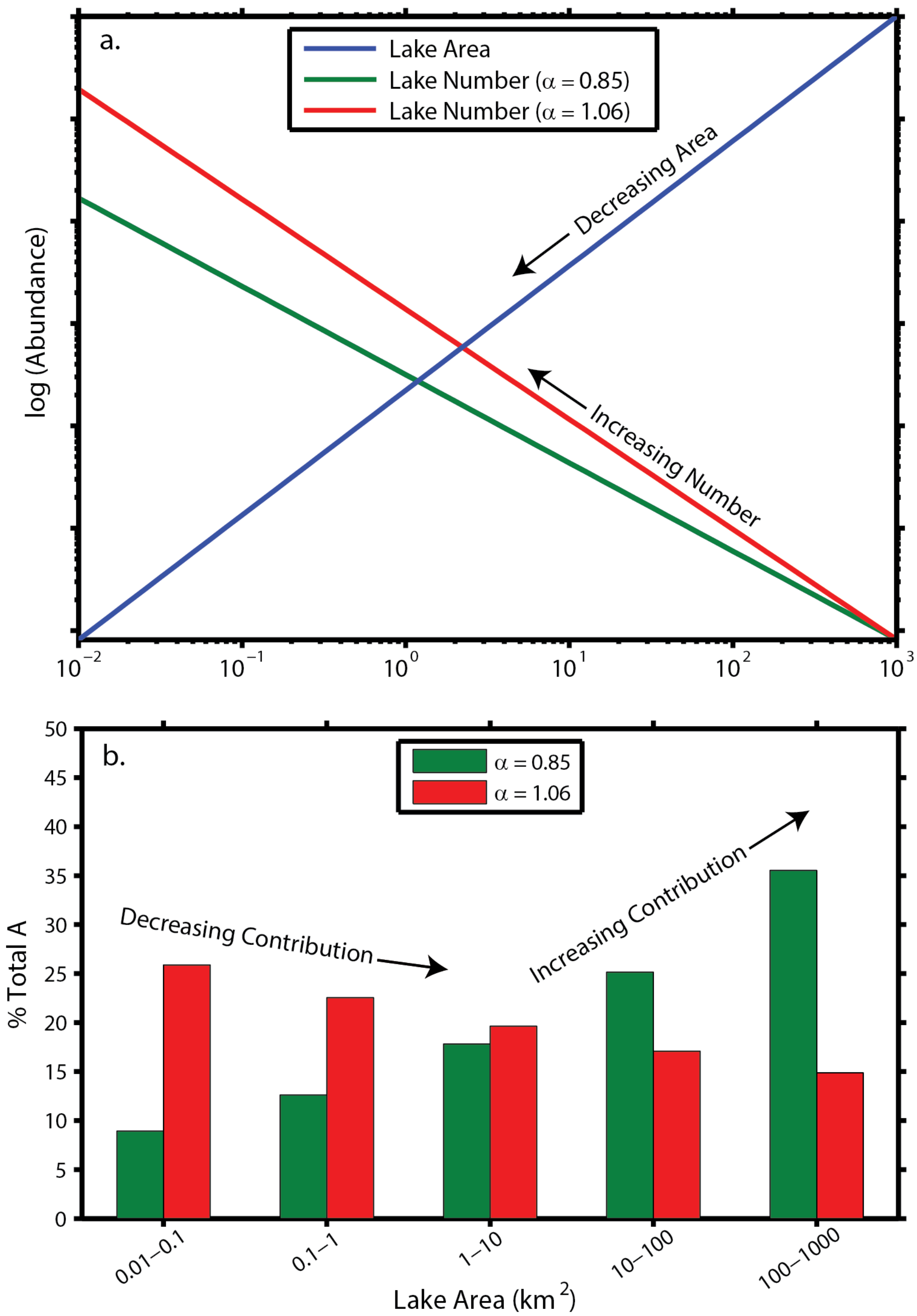
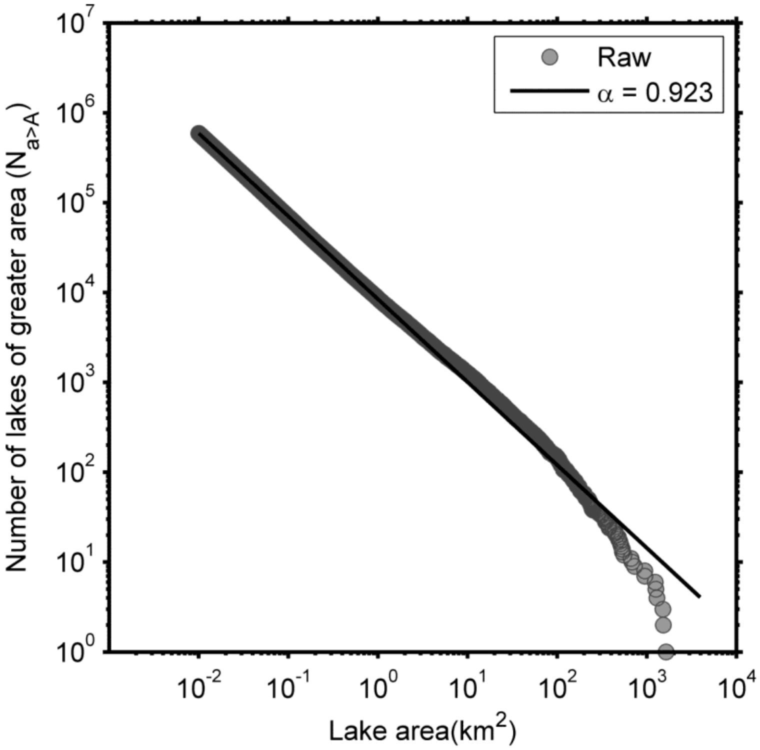


Figure 2:



Raw
α = 0.923
Number of lakes of greater area ($N_{a>A}$)
10^0
10^1
10^2
10^3
10^4
10^5
10^6
10^7
10^{-2}
10^{-1}
10^0
10^1
10^2
10^3
10^4
Lake area(km^2)

Figure 3:

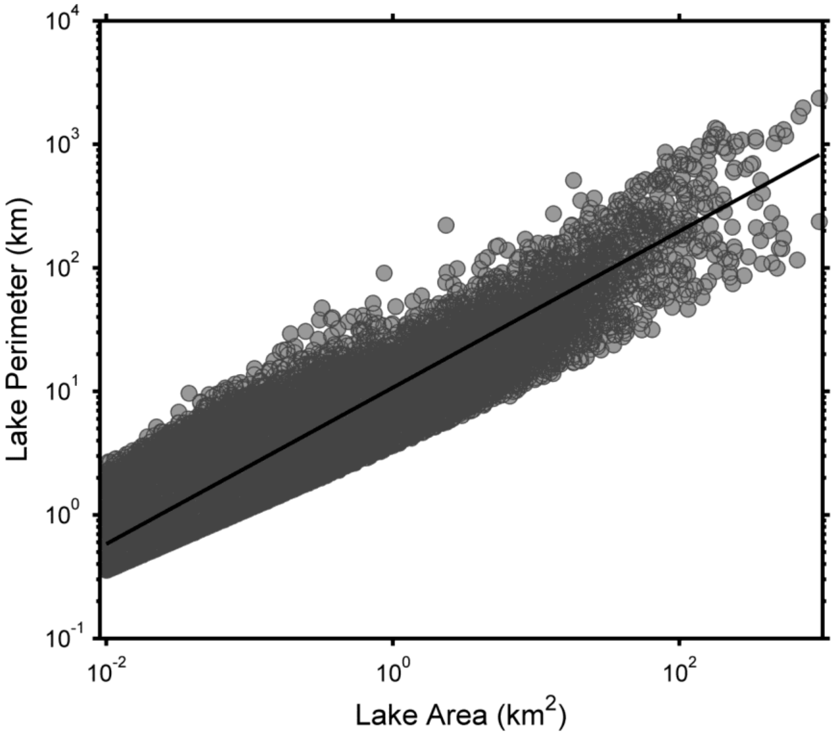

Lake Perimeter (km)
Lake Area (km^2)
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}
10^{4}
10^{-2}
10^{0}
10^{2}

Figure 4:

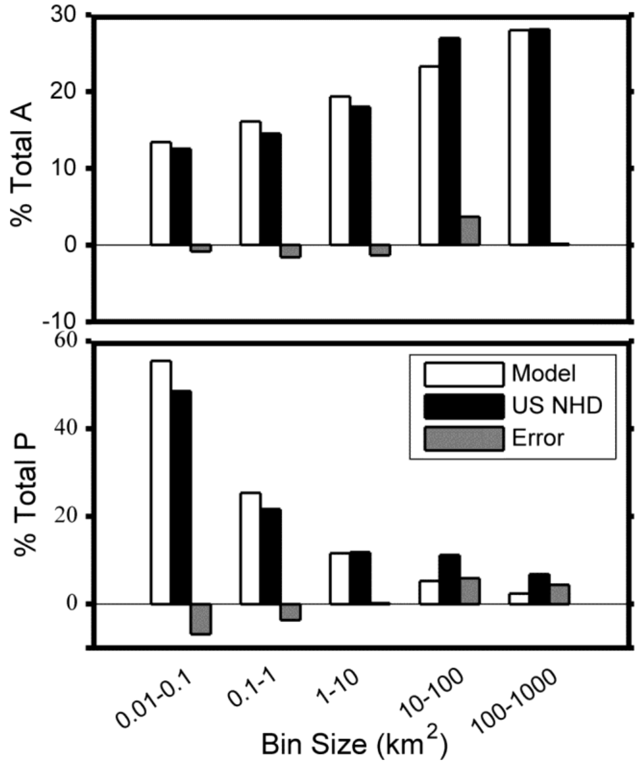

30
20
10
0
-10
% Total A
60
40
20
0
% Total P
Model
US NHD
Error
0.01-0.1
0.1-1
1-10
10-100
100-1000
Bin Size (km^2)

Figure 5:

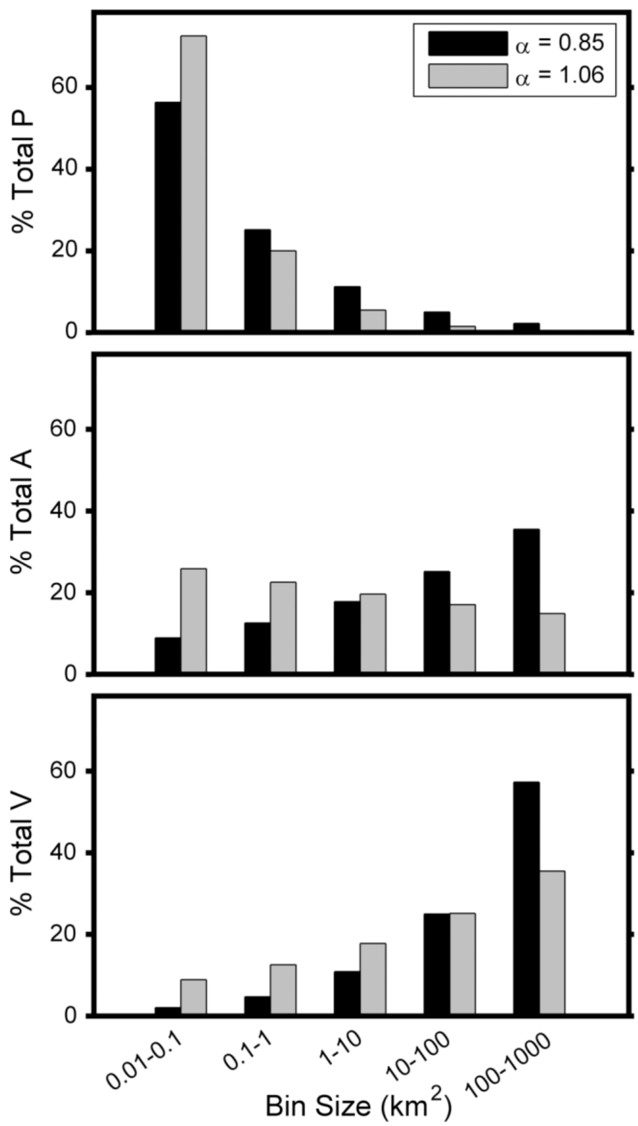
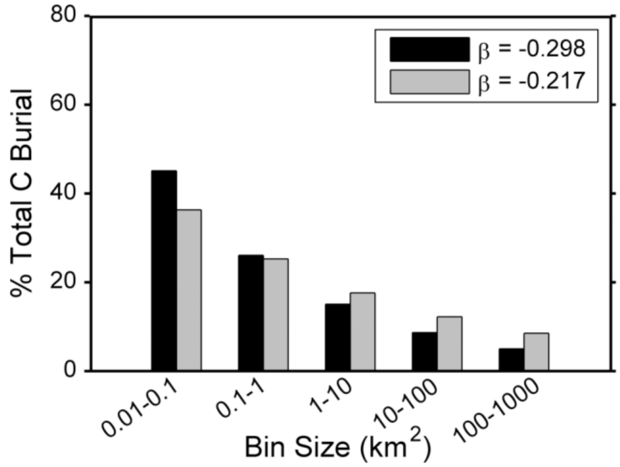

α = 0.85
α = 1.06
% Total P
% Total A
% Total V
0
20
40
60
0.01-0.1
0.1-1
1-10
10-100
100-1000
Bin Size (km^2)

Figure 6:


β = -0.298
β = -0.217
% Total C Burial
0
20
40
60
80
0.01-0.1
0.1-1
1-10
10-100
100-1000
Bin Size (km2)

CHAPTER 4 - LAKES ON A FRACTAL LANDSCAPE: ALTERNATIVE MECHANISMS EXPLAIN DEVIATIONS FROM THE PARETO DISTRIBUTION FOR LAKES

Winslow LA, M Cardiff, PC Hanson, TH Leach, EH Stanley

Target Journal: Journal of Geophysical Research: Earth Surface

KEY POINTS

1. Fractal landscapes can reproduce the lake size-abundance distribution and the power-law relationships of lake area vs. volume, and perimeter, though the area-volume and area-perimeter exponents could not be concurrently matched on a single landscape.
2. Large lake abundance tends to fall below power-law predictions due to large lakes having a high likelihood of intersecting the landscape boundary.
3. Simple, scale-specific models improve similarity of lakes on a fractal landscape with the U.S. lake distribution.

ABSTRACT

The processes that create and modify terrestrial lakes are well known. Glaciation, landslides, and thermokarst processes, among others, are known to create lakes on earth. Conversely, sedimentation, terrestrialization, and water seepage have been identified as a few processes that can eliminate existing lakes. To-date, little work has examined how these processes affect the observed lake size-abundance distribution. Here we use a fractal model to generate populations of lakes and compare the resulting morphological characteristics with those of observed lake populations. We apply three scale-specific processes that eliminate lakes as alternative models for explaining the observed lake distribution shape. We found that lakes on a fractal surface have generally similar characteristics and patterns of area, perimeter and volume when compared to observed lakes. However, the specific exponents for the relationships of area vs. perimeter and volume tended to be steeper for fractal surfaces. The difference in fitted exponents can be explained by the loss of fine-scale features in observed lakes when compared with fractal surfaces and lakes. These relationship exponents varied across fractal terrains generated with different underlying roughness coefficient. However, no single roughness coefficient could be chosen to concurrently match the lake area-volume and area-perimeter relationships. On the fractal surfaces, due to edge effects, large lake abundance tended to be deficient when compared to the small lake abundance. Lastly, of the three scale-specific processes applied to the distribution, terrestrialization fit the observed data the best.

INDEX TERMS AND KEYWORDS

Lakes, power-laws, geomorphology, global limnology, fractal surfaces

1. INTRODUCTION

Near the front cover of almost every Limnology textbook comes a chapter on the "Origin of Lakes" that describes the well-known and well-vetted local scale processes such as glaciation, tectonics and landslides that result in lake creation. Not addressed in most texts is how these processes result in the observed shape of the distribution of lakes at landscape to global geographic scales. Empirical studies have characterized the general nature of the size distribution and abundance of lakes at the continental [*McDonald et al.*, 2012] and global scales [*Lehner and Doll*, 2004; *Downing et al.*, 2006]. This work often reports distributions that are power-law (Pareitan) in form, despite noted deviations from a pure power law. How the processes modify that underlying power-law distribution at the landscape-level remain largely unexplored.

Understanding the mechanisms that generate observed lake populations is useful for several reasons. Such understanding could improve estimates of lake extent and distribution into areas lacking high resolution hydrologic maps. Currently, quantifying the role of lakes in global processes is challenged by the uncertainty in total lake abundance and area [*Lewis*, 2011]. Better understanding of the mechanisms that impact the lake distribution could also help improve predictions of climate change impacts on lakes. As distributions of precipitation, runoff and evaporation change in response to climate, the spatial distribution of lakes will likely change. Lastly, identifying mechanisms that create and remove lakes from the landscape could help us understand the natural fate of lakes as biologic and geologic processes proceed through time. The currently observed lake distribution is the result of a continuously active process of lake formation and destruction. Landscapes recently renewed through processes, such as glaciation or uplift, may have different lake distributions than older landscapes [e.g., *Englund et al.*, 2013], especially if the formation or destruction mechanisms act preferentially on smaller or larger lakes.

While studies have focused on describing lake distributions, there is much to be learned about connecting distributions to physical and biological processes. *Downing et al.* [2006] examined the size-abundance distribution of lakes in a variety of locations and identified a population well described by a Pareto distribution. They used this Pareto distribution assumption to extrapolate small lake abundance from a database of the globe's largest lakes to estimate the total population and surface area of lakes, but that work did not explore mechanistic underpinning for a Pareto lake distribution. In subsequent work, *Seekell and Pace* [2011] showed that if a Pareto distribution did not accurately describe the lake size-abundance distribution, then extrapolating from only the largest lakes could result in highly erroneous estimates of the number of small lakes globally. This log-normal hypothesis was refuted by work using a continental-scale, high-resolution lake dataset, which showed that while the distribution seems to deviate from a Pareto distribution, it does not deviate as strongly as would be predicted if the distribution were log-normal [*McDonald et al.*, 2012; *Winslow et al.*, 2014]. Nonetheless, an analysis of this continental dataset called into question the original assumption of a pure Paretian lake distribution and reported a reduced estimate of the total global abundance and area of lakes. Finally, recent work by *Seekell et al.* [2013] suggests that the lake distribution has an underlying Pareitan distribution when the population examined is restricted to lakes existing at the mean landscape elevation. However, this conclusion is based on a small population (~100 lakes) and the authors did not apply a hypothesis test designed specifically for power-laws [*Clauset et al.*, 2009].

Most work at the lake population level has focused on the relationship between area and abundance, with relatively little attention paid to volume and perimeter. While early work conceptually introduced the relationships between different aspects of lake-population morphology [*Wetzel*, 1990], it lacked empirical grounding. Lake volume and perimeter has been examined separately at large-scales [*Sobek et al.*, 2011; *Winslow et al.*, 2014], but how the three dimensions of lake morphology relate across the full landscape remains unexplored.

The fractal nature of landscape geomorphology may provide a means for resolving differences in interpretation of the observed lake distribution and could help link different dimensions of lake morphology. Many landforms and surface processes that shape the earth's surface generally obey fractal statistics by presenting power-law distributions [*Mandelbrot*, 1983; *Turcotte*, 1992]. This observed landscape with fractal characteristics is the template on which terrestrial lakes form. Given a fractal surface, a near-perfect Pareto distribution of lakes can be generated by flooding the surface's depressions [*Goodchild*, 1988]. Such a model produces realistic power-law distributions of lake size-abundance (Figure 1), though it lacks deviations seen in the small and large lake tails of the observed distribution (Figure 2).

One way to generate deviations from a power-law distribution is to apply a scale-dependent process. A scale-dependent process would either act to alter the observed population in a specific size range while leaving the rest of the population unaltered, or act across the full population disproportionally to lake area. How scale-dependent processes affect the shape of the lake size-abundance distribution has not been directly examined in the literature. However, published evidence hints at scale-dependent processes that may lead to the size-specific loss of lake abundance or area. Lake abundance may be reduced in a size-selective way by sedimentation, which would tend to fill small, shallow lakes more quickly than larger lakes [*Englund et al.*, 2013]. Power-law distributions in non-lake topographic depressions suggest that in some regions, there is a potential for a Pareto distribution of small lakes, but drying or draining prevents their perennial existence [*Le and Kumar*, 2014]. Biological processes, such as the encroachment of peat mats in northern boreal zone lakes (terrestrialization), could alter existing lakes and their apparent area, or even eliminate them altogether [*Roach et al.*, 2011]. Real landscapes have finite area that constrain geomorphological features at their boundaries. Depressions that would form large lakes have a higher likelihood of intersecting the landscape edge and not being formed [*Goodchild*, 1988]. For continental-scale distributions, the continental edge may represent such a boundary. Furthermore, the distribution of U.S. lakes may not represent a single landscape with a single overarching distribution; rather it might be better described as a mosaic of landscapes with different underlying characteristics. To our knowledge, these mechanisms have not been incorporated into an analytical framework that accurately recreates the observed lake size distribution.

To investigate the underlying mechanisms that explain the observed distribution of lakes, we ask three questions. Can lakes on a fractally generated landscape re-create the observed distribution of lakes in the U.S., and how do the fractal lakes' perimeter and volume distributions compare to the observed distributions? How might a suite of scale-specific mechanisms be added to fractal scaling to explain deviations from the Pareto distribution in the observed data? Can different interpretations of the lake size-abundance distribution be resolved by a process-modified fractal lake distribution? Answering these questions can help rectify some of the conflicting results reported and bring greater understanding to landscape-level patterns in lake morphology and distribution.

2. METHODS

2.1.1 FRACTAL LAKES

The fractal landscape was generated using a diamond-square surface generating algorithm. While there are numerous implementations of this algorithm, we used a freely available implementation for Matlab [*Viswanathan*, 2001]. The algorithm requires a numerical roughness coefficient that represents a moderator of the random process. When this index increases in size, it produces terrains with higher elevation ranges (i.e. a "rougher" landscape). For the diamond-square algorithm, the landscape width needs to be a $2^n + 1$, where n is an integer (width and height were equal). From this point forward, we omit the "+1" when describing the landscape size. We generated landscapes in several ways to examine different questions. One, to create a single, large landscape with a single, large lake population, we generated a landscape with width 2^{12}. Larger landscapes were constrained by computation time and memory availability. Two, to examine how the lake distribution was affected by edge effects and variability between randomly generated landscapes, we generated 100 random landscapes with width 2^{10}. Lastly, on a landscape of width 2^{10}, we varied the roughness coefficient from 0.01 (mountainous) to 1 (very flat) to examine its impact on resulting fractal lake morphology. Because the dimensions of the generated landscapes lack real units, a convenient way to interpret the landscape is based on the size range of lakes which can be generated. The largest landscape, 2^{12}, resulted in lakes which spanned 5 orders of magnitude in size. For contrast, the U.S. lake distribution spans about 7 orders of magnitude, 9 if the Laruentian Great Lakes are included, so the largest generated landscape is roughly similar to the extent of a small U.S. state. For brevity, we call lakes that reside in these fractal landscapes, "fractal lakes".

With the generated landscapes, we applied the algorithm of *Goodchild* [1988], designed to mimic the "filling" of all potential basins to the point of overflowing. The algorithm applied three rules; 1) the edge of the landscape was considered an infinite sink, meaning that no edge cell could be part of a lake, 2) fluid could travel across adjacent grid cells of equal or less elevation and 3) all local minima were filled until the edge of the landscape was reached. Lakes were considered full when further increase in water level intersected an outflow cell. No virtual rivers or streams were recorded, though the algorithm could be modified to include them. All connected flooded grid cells were considered a single lake. Each lake was then analyzed to determine surface area, volume and perimeter. All units reported for fractal lakes are arbitrary and herein referred to as virtual units (VU), which represents the width of a single cell on the fractal landscape. For example, the smallest lakes have an area of one and perimeter of four.

2.1.2 Scale-dependent processes

We considered three different scale-dependent models applied to the fractal lake distributions. Each of these models had a single free parameter that defined its magnitude. The parameter values were selected to show a range of observable impacts on the population. However, caution should be used in the interpretation of the parameter values, as they were chosen for effect on the fractal lake populations and not derived from literature values. For models that showed interesting results, we compared the model parameter values with estimated values from the literature. The first and simplest model mimicked the effects of sedimentation or evaporation applied evenly to all lakes on the landscape and acted to remove lakes from the landscape proportionally to their volume [*Brooks and Hayashi*, 2002]. The second scale-dependent model applied to the lake distributions simulated terrestrialization. In some regions, lakes experience edge encroachment through the growth of a surrounding peat mat [*Roach et al.*, 2011]. To mimic this process, our algorithm removed lake pixels based on a depth threshold (from 1 to 50% of the maximum simulated lake depth), thereby simulating the growth of moss, beginning in shallower areas and progressing into deeper areas. The last scale-specific model was a simple groundwater seepage model. To simulate the water table, we created a smoothed version of the overlying landscape using a 2-dimensional moving average window width area 200 VU^2. This virtual groundwater surface was shifted downward by a variable amount to simulate areas with different groundwater depths. Where lakes currently existed and the moving average intersected the terrain surface, the lake was retained. In areas where the terrain surface did not intersect the lake bottom, the lake was considered "dry" and removed from the population [modeled after *Kratz et al.*, 1997].

2.2.1 Lake observation data

For power-law analyses and to compare the fractally derived lake distributions to existing lakes, we used the National Hydrography Dataset available from the U.S. Geological Survey. We excluded lakes in Hawaii and Alaska and used only lakes in the contiguous U.S., excluding the Laurentian Great Lakes (except where noted). To compare the U.S. lakes distribution to the fractal lake distribution, a minimum size cutoff of 0.01 km^2 was used. It was difficult to identify lakes below that size based on the original mapping resolution, and data below that cut-off is considered unreliable [*McDonald et al.*, 2012]. For other analyses, different cutoffs were used and noted with the results. Further information on this dataset and collection methods can be found in *Winslow et al.*, [2014] and *McDonald et al.*, [2012].

Bathymetric data were collected from the state natural resource agencies of Wisconsin, Minnesota and Nebraska. Combined, there were 1319 unique lakes. The original sources vary, with some data based on modern acoustic-sounding techniques and others derived from hand-drawn survey maps. When not already in a digital format, data were digitized, georeferenced, and converted to a raster map of depths with a spatial resolution of 5 meters. From these raster maps, lake volume and area was directly calculated using R (v2.15.2) and the raster package [*Hijmans*, 2014].

2.3 ANALYSES

To test if a distribution could have been generated from an underlying power-law, we used a semi-parametric approach, which improves significantly over using R^2 values in evaluating the fit of data to a power-law [*Clauset et al.*, 2009]. A hundred fractal populations were generated (width of each landscape: 10^{10} VU) to examine the variability in lake generation of the random landscape-generating process. Each resulting population was fit and tested using the *Clauset et al.*, [2009] fitting and power-law hypothesis testing methods.

To test the large-lake, landscape-boundary hypothesis that predicts many smaller fractal surfaces combined would be large-lake deficient due to boundary effects, we combined 100 smaller fractal lake populations and fit the full population in two contrasting ways. One, the slope parameter was fit to the full population of lakes to see what Pareto slope parameter would be estimated for the whole. Two, the full population was combined, but only lakes with areas greater than 10^3 VU were fit. This mimicked the difference between fitting the full size range of lakes at a large scale [*McDonald et al.*, 2012] and fitting only the largest lakes of a population [*Downing et al.*, 2006].

To compare the shapes of the observed U.S. lakes population, we normalized the distributions to eliminate differences in the range of lake sizes and number of lakes. This was necessary to normalize units between the fractal lake distribution and the observed distribution. To determine which scale-dependent model best described the existing distribution, we used the two-sample

Kolmogorov–Smirnov (KS) test. Before calculation, we log-transformed the lake area values and then normalized the observed and modeled lake populations using the following equation:

$$X_{norm} = \frac{(X - median(X))}{mean(X - mean(X))} \text{ [eq. 1]}$$

where X is the set of lake areas. The denominator is the mean absolute deviation of the set of lake areas. This is similar to the standard normal transform, but less sensitive to outliers and non-normal distributions. From this test, we used the KS statistic as an index of observed distribution similarity, where smaller numbers indicate higher similarity.

3. RESULTS

3.1 FRACTAL LAKES

The single, large fractal landscape **(**width 2^{12}, roughness 0.7) generated a surface with 434,000 lakes. The lakes covered 6,236,000 VU^2 in total, 37% of the total landscape area. Lakes on the fractal surface ranged in area from 1 to 626,000 VU^2. The generated lake distribution passed the semi-parametric power-law test ($p<0.05$) and had a Pareto slope parameter of 0.95. The regression slope of lake area to perimeter was 0.66 and the slope of lake area to volume was 1.27. As-generated, the fractal surface elevations had a standard deviation of 14.2 VU.

The lake populations from 100 randomly generated fractal surfaces (width 2^{10}, roughness 0.7) helped to describe the expected variability of key parameters due purely to random differences across landscapes, These 100 lake populations had an average Pareto slope of 1.008 with a range from 0.85 to 1.13. The relationships of area with volume and perimeter were highly constrained, having medians of 0.66 ±0.002 and 1.27± 0.005 respectively. Landscape elevation standard deviation showed large variability, ranging from 10.7 to 28.7 with a median of 15.9. Of the 100 randomly generated lake populations on fractal surfaces (width 2^{10}, roughness 0.7), 9 lake populations were rejected by the semi-parametric power-law test ($p<0.05$) applied to lake areas.

It was difficult to determine how landscape elevation standard deviation varied with the roughness coefficient as standard deviation varied greatly across landscapes with the same underlying landscape generation parameters. To examine how standard deviation changed with roughness coefficient, an additional 100 fractal landscapes were generated with a different roughness coefficient (0.4). These landscapes had a median standard deviation of 20.4 and ranged from 16.2 to 29.5. Other morphological aspects had lower variability and matched the results of our cross-roughness analysis based on single generated landscapes.

3.2 Comparing fractal and observed lake populations

For the observed U.S. distribution, the regression slopes of area vs abundance, perimeter and volume were similar to the fractal landscapes (Figure 3). The size-abundance Pareto slope for lakes in the contiguous U.S. is 0.92 (0.85 with no size cutoff [*McDonald et al.*, 2012]). The contiguous U.S. relationship of area to perimeter had an exponent of 0.58, less than the values from the fractal lakes (0.66). The empirical relationship between area and volume deviated furthest from the relationship found on the fractal landscape (1.15 and 1.27 respectively), though this is the empirical relationship based on data from only three of the states in the contiguous U.S. (Wisconsin, Minnesota, Nebraska). Because of the limited spatial extent of available bathymetric data, this result may not be representative of the lakes in the contiguous U.S. It is unclear how our results would change if a U.S.-scale dataset were available.

The area to volume and area to perimeter relationship exponents depend on the fractal landscape roughness (Figure 4). The perimeter exponent decreased while the volume exponent increased with an increasing roughness coefficient. Because a large abundance of very small lakes may be influencing the observed relationships, the model was also fit to only lakes larger than 100 VU^2. Those exponents showed the same trend with roughness, though the estimated values were more sensitive to changes in the landscape roughness coefficient.

While most fractal landscape lake distributions were statistically indistinguishable from a power-law, the populations often had a visible negative deviation in the large-tail of the distribution, suggesting that large lakes are missing from the populations. When the lakes from 100 individual landscapes are combined, mimicking a fragmented landscape, the missing large lakes become apparent (Figure 5). When only the large tail of the distributions is fit (> 10^4 VU^2), the Pareto slope is steeper than when fit to the whole population (1.3 vs 1.01 respectively, Figure 5).

The three scale-dependent models varied in their effects on the lake size-abundance distribution (Figure 6). The simplest effect was from the evaporation/sedimentation model (Figure 6b). When lakes were removed from the distribution based on their volume, the small lake abundances quickly drop to zero which can be seen as the flattening of the cumulative distribution. Large lakes are not affected until the magnitude of the evaporation/sedimentation process is very large. Terrestrialization (Figure 6a) had an effect that could be seen across the lake size-abundance distribution, though it has the strongest effect on small lakes. Little to moderate encroachment tended to fill in small lakes, reducing their abundance. Unlike the evaporation/sedimentation process, small lakes were not eliminated from the distribution.

Across all three scale-dependent models, terrestrialization created the lake size-abundance most similar to the observed lake population (Table 1). With a depth cutoff of 2.8 VU, terrestrialization had the lowest reported KS value of 0.14, a substantially better fit than the fractal population without any scale-specific modification (KS: 0.28).

4. DISCUSSION

4.1 FRACTAL AND OBSERVED LAKE MORPHOLOGY

While fractal landscapes should result in power-law lake distributions [*Goodchild*, 1988], some of our landscapes had lake populations that differed significantly from a power law. From 100 the randomly generated landscapes, it was surprising that 9 lake populations were rejected by the power-law test ($p < 0.05$). Because this technique has not previously been applied to lake distributions, it is unclear if such a false negative rate is to be expected, or if the hypothesis test is too conservative when applied in this situation. For our landscapes, the rejected populations may be a result of true biases of the population from a power-law due to reduced large lakes abundances on a surface with boundaries (discussed below), or may suggest the test is overly conservative for large numbers of observations.

The Pareto slope of the area distribution is an important feature of the lake population that defines the relative contribution of large and small lakes to the total population area, perimeter and volume. A slope of 1.0 is the critical point where the relative contribution by large and small lakes to total lake area is balanced. Slopes above and below 1.0 describe populations skewed towards large or small lakes [*Winslow et al.*, n.d.]. Empirical studies have reported regionally differing slopes, varying from 0.56 to 1.10 [*McDonald et al.*, 2012]. The 100 fractal landscapes with the same roughness coefficient, differing only due to random processes, generated lake populations with Pareto slopes that varied from 0.85 to 1.13. This suggests that some slope variability can be attributed to random variation alone, not to changes to the underlying generation mechanism. However, the variability range from observed datasets is higher than for the fractal terrains, suggesting underlying differences between regional lake distributions that go beyond random variability. This is not surprising considering the large climactic and geological differences across the observed datasets examined [*McDonald et al.*, 2012]. The steeper overall average Pareto slope for fractal lakes (fractal: 1.01 observed: 0.92) suggests the contiguous U.S. has relatively fewer small lakes than would be predicted with a fractal lake model alone. Scale-dependent processes preferentially affecting small lakes may explain the observed discrepancy from a fractal lake distribution.

Lake area vs. volume and perimeter also showed relationships comparable to those estimated from observed lake populations (Figure 3). Large lakes tended to have disproportionately longer perimeters and larger volumes than small lakes, comparable to results from the U.S. and Sweden [*Sobek et al.*, 2011; *Winslow et al.*, 2014]. The relationship exponents of area vs perimeter and volume were steeper for fractal lakes. Because a steeper slope in these morphology relationships is driven by a high abundance of fine-scale features (e.g., shoreline convolution and lakes of small volume), the exponent differences between the fractal and observed relationships indicate the fractal population has more fine-scale features. The steeper area to perimeter slope (fractal: 0.66 observed: 0.58) can only be explained by longer perimeters in the fractal population's large lakes as small lakes have a bottom limit to perimeter. This suggests the large fractal lakes have finer-scale shoreline features than observed lakes. The steeper slope of area to volume (fractal: 1.27 observed: 1.15) may be a result of shallower lakes on the fractal landscape than would be unlikely to exist in reality (mean depth c. 0.01 VU, roughly equivalent to 1 cm). These lakes can be seen at the small end of the area to volume regression (green line, figure 3a). Overall, fractal lakes display more fine-scale features than the observed lake population, an important difference potentially pointing to key processes driving observed lake morphology. Scale-specific processes that that act to eliminate and modify morphology at the population scale can help explain the differences between observed and fractal lakes.

The results from the fractal surfaces helps inform why landscape roughness broadly predicts the relationship between lake area and volume, but fails to improve individual lake volume estimates. Reported results have shown that when predicting lake volume, only lake area is a significant predictor, though broadly, the area-volume relationship exponent is higher in areas of rougher terrain [*Sobek et al.*, 2011, *S. Oliver*, personal communication, July 2014]. On our fractal surfaces, the underlying roughness coefficient is a very good predictor of the lake area-volume relationship exponent, with higher values for rougher terrain due to smaller lakes being disproportionately deeper. This is similar to the significantly higher lake area-volume coefficient found in a mountainous region by *Sobek et al.*, [2011]. Despite the good relationship of roughness coefficient with lake area-volume exponent, the standard deviation of landscape elevation is highly variable across individual landscapes generated using the same roughness coefficient. Therefore, at a fine scale, the observed landscape roughness is a poor predictor of the relationship between lake area and volume. This suggests that for terrestrial lake populations, immediately surrounding landscape characteristics may not improve models of lake volume, but at large-scales, roughness will help predict the relationship exponent between lake area and volume. To better understand how broad-scale landscape roughness predicts the lake area-volume relationship, more comprehensive and geographically-distributed datasets of lake volume are needed.

4.2 Scale-dependent processes

The lack of large lakes given the number of small lakes in the fractal lake populations is a result of edge effects. Due to their size, depressions that would form large lakes have a high probability of intersecting the landscape boundary and flow off the edge thus forming a small lake, or not forming a lake at all [*Goodchild*, 1988]. Combining 100 lake populations generated on smaller landscapes into one larger population highlights the outcome of this process (Figure 5). Globally, land effectively has boundaries at the edges of all continents. Large rivers may also form additional boundaries where, without erosion, large lakes may have formed. These boundary effects combine to reduce the number of large lakes otherwise predicted by small lake abundance and a power-law distribution. Because of their low abundance in the distribution, large lakes have relatively little impact on Pareto distribution fits spanning the entire size-range. Fits estimated from a large-biased population will be adversely affected by this boundary bias.

The three scale-dependent models resulted in lake distributions with different shapes and key features. The evaporation/sedimentation model has generally the simplest result, completely eliminating small lakes preferentially from the lake distribution. This tendency to eliminate small lakes from the population may be somewhat unrealistic as sedimentation rates can be highly variable across different lakes [*Downing et al.*, 2008]. A variable sedimentation rate would be difficult to parameterize on the fractal landscape, but would likely result in a less complete elimination of small lakes. Groundwater seepage had a gradually changing effect across small to medium sized lakes, and had little effect on the largest lakes. Small to medium sized lakes were strongly impacted as they had higher densities in higher landscape positions. The largest lakes had lower elevations, meaning even with a strongly depressed virtual water table, they remained unchanged. The resulting populations had the least amount of curvature in the distribution, acting more to change the slope while maintaining a linear shaped distribution. This has an interesting parallel with the U.S. lake distribution, which does not have a slope of 1.0 [*Winslow et al.*, n.d.], as would generally be predicted by an unaltered fractal lake distribution. Instead, in the central portion of the distribution, the population displays a near-linear relationship with a slope of 0.92 (Figure 2). Lastly, terrestrialization created the most curved distributions. This curvature is a result of the encroachment growth eliminating small lakes, but replacing them with lower abundance medium sized lakes turned into small lakes.

While we did not link the fractal landscape units directly to observational units, it is useful to ground-truth the magnitude of the best fitting scale-dependent process found here. To do that, we look at the results of the terrestrialization model fit and compare the relative magnitude of terrestrialization depth and median lake depth with published literature values. The cutoff value that resulted in the best terrestrialization model was 2.8 VU. This means that all fractal lake area with a depth of less than 2.8 VU was considered filled and no longer counted as lake. On the affected fractal landscape, the median simulated lake depth was 1.4 VU, and so depth of terrestrialization was greater than the median lake depth. In a lake district with a high amount of terrestrialization, the Northern Highland Lake District (NHLD), mean observed lake depth is 3.9 meters [*Hanson et al.*, 2007], while estimated mean region-wide peat depth is 2.1 meters [*Buffam et al.*, 2010]. Furthermore, the fractal lake median depth of 1.4 VU is based on the original, unmodified fractal lakes. A more comparable metric would be the median depths for lakes remaining after the terrestrialization process has been applied (7.4 VU). So an average peat depth below, but near the median lake depth is plausible. Furthermore, the landscape roughness of the NHLD and the fractal terrain used for the models is similar, with elevation standard deviations of 18.5 VU and 16.7 meters respectively. This suggests that the vertical processes and morphology described on the fractal surface (volume and terrestrialization related to depth) may represent realistic scales.

4.3 RESOLVING DIFFERENT REPORTED GLOBAL LAKE ESTIMATES

The underlying reasons for differences in reported global lake abundances are based on the landscape boundary effect on the large-lake end of the distribution. When the Pareto distribution is fit to large global lakes over 10km^2, the slope is 1.06 [*Downing et al.*, 2006] but when the Pareto slope is fit to a larger area range of lake in the U.S., the value is 0.85 [*McDonald et al.*, 2012]. This flattening of the distribution substantially reduces the number of small lakes that would be predicted from a given number of large lakes, meaning that the estimate made by *Downing et al.*, [2006] based on extrapolating large lakes is almost certainly an over-estimate of abundance and surface area. More accurate estimates of global processes can likely be made using the global values estimated by *McDonald et al.*, [2012]. Furthermore, process estimates that are based on a pure Pareto-distribution assumption [*Winslow et al.*, n.d.; *Kastowski et al.*, 2011] may underestimate the contribution of large lakes.

Differences in the relationships of key lake morphology, such as area vs volume, may be a result of different topographical roughness. Additionally, difficulty in estimating the underlying roughness coefficient from the standard deviations of landscape elevation may explain why such measures of roughness seem to improve individual lake volume predictions very little [*Sobek et al.*, 2011]. Despite this, the low variability in relationships between lake area, perimeter and volume on landscapes with a single roughness coefficient suggest that underlying roughness is an important factor in controlling the relationships between lake area, perimeter, and volume. While information may not improve models predicting the volume of any single lake, large-scale, regional estimates of roughness may improve models of lake volume at such broad scales. It is unclear if lake-specific estimates of volume can be substantially improved. The curation of a continental-scale database of bathymetry would improve this and other large-scale limnological analyses.

5. CONCLUSIONS

We found that lakes on a fractal surface have generally similar characteristics and patterns of area, perimeter and volume when compared to observed lakes. However, the specific exponents for the relationships of area vs. perimeter and volume tended to be steeper for fractal surfaces. The difference in fitted exponents can be explained by the loss of fine-scale features in observed lakes when compared with fractal surfaces and lakes. These relationship exponents varied across fractal terrains generated with different underlying roughness parameters. However, no single roughness parameter could be chosen to concurrently match the lake area-volume and area-perimeter relationships. The landscape roughness coefficient predicted the relationship between lake area and volume, though standard deviation of landscape elevation was too variable to be used as an observable proxy for landscape roughness.

Examining lakes on fractal surfaces helps resolve differences in reported results among large-scale lake abundance studies and provides insight into how different small-lake scale-dependent processes modify the shape of the size-abundance distribution (Figure 7). Due to edge effects, large lake abundance tended to be deficient when compared to the small lake abundance. This deviation would explain reported overestimates in small lake abundance based on large-lake population extrapolation. Small lake deviations from a pure fractal distribution can be in-part explained by scale-specific processes applied to the distribution, with terrestrialization fitting the observed data best. Further work is required to understand how these processes, among others, apply to different regions and lake populations to create the currently observed lake distribution.

ACKNOWLEDGEMENTS

U.S. lake data are available from the U.S. Geological Survey at nhd.usgs.gov. Lake volume data derived from digitized bathymetry organized and available at www.bathybase.org. Funding for this work was provided by NSF DEB-0941510 grant and an Anna Grant Birge Fellowship. Special thanks to Jordan Read and Steph Januchowski-Hartley for their input at various stages of the project.

REFERENCES

Brooks, R. T., and M. Hayashi (2002), Depth-area-volume and hydroperiod relationships of ephemeral (vernal) forest pools in southern New England, *Wetlands*, *22*(2), 247–255, doi:10.1672/0277-5212(2002)022[0247:DAVAHR]2.0.CO;2.

Buffam, I., S. R. Carpenter, W. Yeck, P. C. Hanson, and M. G. Turner (2010), Filling holes in regional carbon budgets: Predicting peat depth in a north temperate lake district, *J. Geophys. Res.*, *115*(G1), G01005, doi:10.1029/2009JG001034.

Clauset, A., C. R. Shalizi, and M. E. J. Newman (2009), Power-Law Distributions in Empirical Data, *Soc. Ind. Appl. Math.*, *51*(4), 661–703, doi:10.1137/070710111.

Downing, J. A., Y. T. Prairie, J. J. Cole, C. M. Duarte, L. J. Tranvik, R. G. Striegl, W. H. McDowell, P. Kortelainen, N. F. Caraco, and J. M. Melack (2006), The global abundance and size distribution of lakes, ponds, and impoundments, *Limnol. Oceanogr.*, *51*(5), 2388–2397, doi:10.4319/lo.2006.51.5.2388.

Downing, J. a., J. J. Cole, J. J. Middelburg, R. G. Striegl, C. M. Duarte, P. Kortelainen, Y. T. Prairie, and K. a. Laube (2008), Sediment organic carbon burial in agriculturally eutrophic impoundments over the last century, *Global Biogeochem. Cycles*, *22*(1), 1–10, doi:10.1029/2006GB002854.

Englund, G., H. Eriksson, and M. B. Nilsson (2013), The birth and death of lakes on young landscapes, *Geophys. Res. Lett.*, *40*(7), 1340–1344, doi:10.1002/grl.50281.

Goodchild, M. F. (1988), Lakes on Fractal Surfaces: A Null Hypothesis for Lake-Rich Landscapes, *Math. Geol.*, *20*(6), 615–630.

Hanson, P. C., S. R. Carpenter, J. A. Cardille, M. T. Coe, and L. A. Winslow (2007), Small lakes dominate a random sample of regional lake characteristics, *Freshw. Biol.*, *52*(5), 814–822, doi:10.1111/j.1365-2427.2007.01730.x.

Hijmans, R. J. (2014), raster: Geographic data analysis and modeling,

Kastowski, M., M. Hinderer, and A. Vecsei (2011), Long-term carbon burial in European lakes: Analysis and estimate, *Global Biogeochem. Cycles*, *25*(3), 1–12, doi:10.1029/2010GB003874.

Kratz, T., K. E. Webster, C. Bowser, J. Maguson, B. Benson, and J. J. Magnuson (1997), The influence of landscape position on lakes in northern Wisconsin, *Freshw. Biol.*, *37*(1), 209–217, doi:10.1046/j.1365-2427.1997.00149.x.

Le, P. V. V, and P. Kumar (2014), Power-law scaling of topographic depressions and their hydrologic connectivity, *Geophys. Res. Lett.*, doi:10.1002/2013GL059114.

Lehner, B., and P. Doll (2004), Development and validation of a global database of lakes, reservoirs and wetlands, *J. Hydrol.*, *296*(1-4), 1–22, doi:10.1016/j.jhydrol.2004.03.028.

Lewis, W. M. (2011), Global primary production of lakes: 19th Baldi Memorial Lecture, *Inl. Waters*, *1*(1), 1–28, doi:10.5268/IW-1.1.384.

Mandelbrot, B. B. (1983), *The fractal geometry of nature*, WH freeman.

McDonald, C. P., J. A. Rover, E. G. Stets, and R. G. Striegl (2012), The regional abundance and size distribution of lakes and reservoirs in the United States and implications for estimates of global lake extent, *Limnol. Oceanogr.*, *57*(3), 1–12, doi:10.4319/lo.2012.57.3.0000.

Roach, J., B. Griffith, D. Verbyla, and J. Jones (2011), Mechanisms influencing changes in lake area in Alaskan boreal forest, *Glob. Chang. Biol.*, *17*(8), 2567–2583, doi:10.1111/j.1365-2486.2011.02446.x.

Seekell, D., M. Pace, L. J. Tranvik, and C. Verpoorter (2013), A fractal • based approach to lake size distributions, *Geophys. Res. Lett.*, *40*, 1–5, doi:10.1002/grl.50139.1.

Seekell, D. a., and M. L. Pace (2011), Does the Pareto distribution adequately describe the size-distribution of lakes?, *Limnol. Oceanogr.*, *56*(1), 350–356, doi:10.4319/lo.2011.56.1.0350.

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

Turcotte, D. L. (1992), *Fractals and chaos in geology and geophysics*, Cambridge University Press.

Viswanathan, A. (2001), Surface Defined by a Plasma Fractal, *Mathworks File Exch.* Available from: http://www.mathworks.com/matlabcentral/fileexchange/702-surface-defined-by-a-plasma-fractal (Accessed 15 February 2014)

Wetzel, R. G. (1990), Land-water interfaces: metabolic and limnological regulators, *Verh Intern. Verein Limnol*, *24*(1), 6–24.

Winslow, L. A., J. S. Read, P. C. Hanson, and E. H. Stanley (2014), Lake shoreline in the contiguous United States: quantity, distribution and sensitivity to observation resolution, *Freshw. Biol.*, *59*(2), 213–223, doi:10.1111/fwb.12258.

Winslow, L. A., J. S. Read, P. C. Hanson, and E. H. Stanley (n.d.), Does lake size matter? Combining morphology and process modelling to examine the contribution of lake size classes to population-scale processes, *Inl. Waters*.

TABLES

Table 1: KS statistics for the three scale-dependent models and different model parameters examined.

Process	Model Parameter	KS Value
Drying	0.1	0.18
Drying	0.4	0.20
Drying	1.6	0.17
Drying	2.4	0.25
Groundwater	-13	0.18
Groundwater	-9	0.19
Groundwater	-5	0.18
Groundwater	-1	0.20
Moss	0.4	0.34
Moss	1.2	0.16
Moss	2.8	0.14
Moss	3.2	0.15

FIGURE CAPTIONS

Figure 1: Theoretical lake distribution based on filling depressions on a fractally generated landscape. The lakes shown on the landscape (A) and the size-abundance distribution (B).

Figure 2: Size abundance distribution of contiguous U.S. lakes, excluding the Laurentian Great Lakes, showing deviation from a pure power-law distribution.

Figure 3: Comparison of how key morphological features scale across the population of observed (blue) and fractal (green) lakes. (a) The volume to area relationships, (b) the perimeter to area relationships and (c) the lake abundance distribution. The line positions on the x axis are arbitrary for fractal lakes and offset to prevent overlap.

Figure 4: Relationship between landscape roughness and the exponent for (a) area to perimeter relationship and (b) area to volume relationship for lakes on fractal landscapes. The dashed lines show the fitted exponent estimates for U.S. Lakes.

Figure 5: Power-law fits to the size abundance distribution of the combined population of 100 random fractal landscapes (size: 2^{10}, roughness: 0.7). The three lines show the difference between Power-law fits based on the whole population versus fits to only the largest lakes.

Figure 6: The different effects of scale-dependent processes on the lake size-abundance distribution for lakes generated on a fractal landscape. For each process, multiple relative magnitudes are plotted.

Figure 7: The regions of the lake size-abundance distribution that deviate from a pure Pareto distribution and the relevant scale-dependent processes that are discussed here

Figure 1:

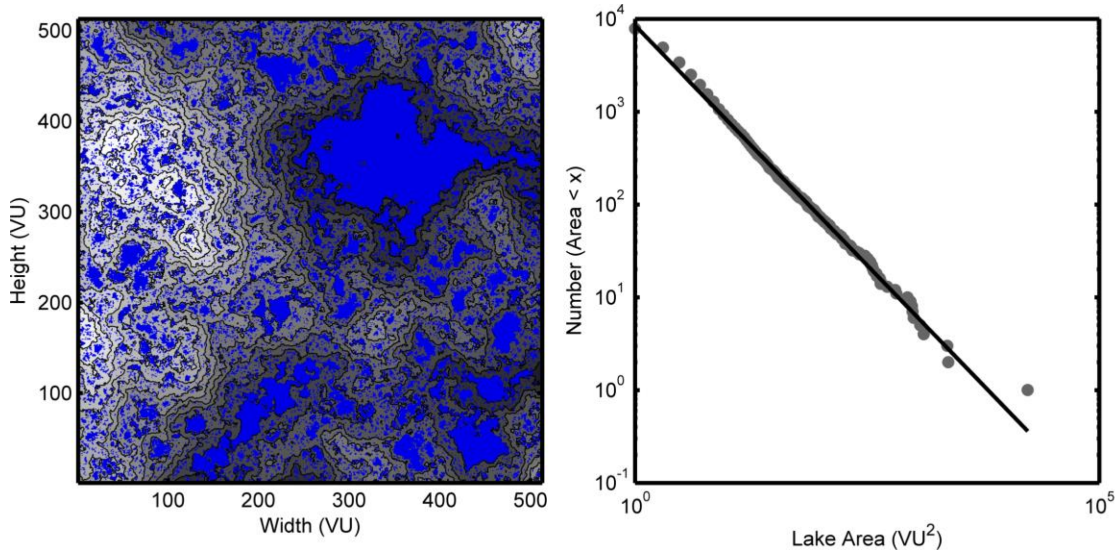
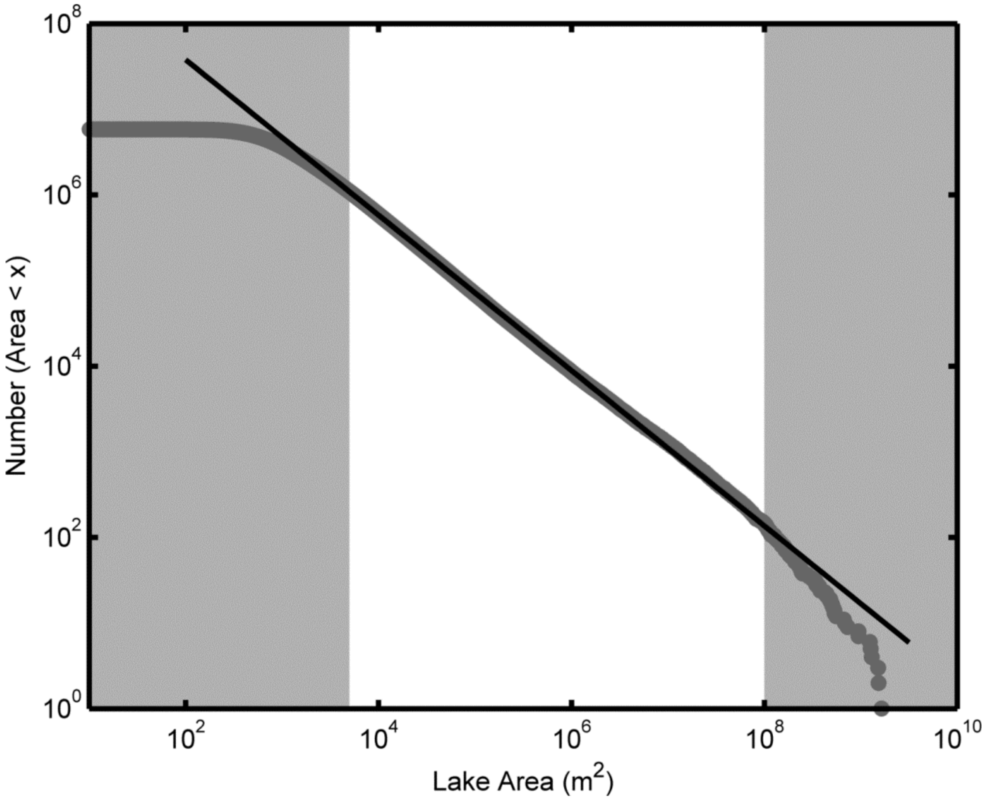

500
400
300
200
100
Height (VU)
100
200
300
400
500
Width (VU)
10^4
10^3
10^2
10^1
10^0
10^{-1}
Number (Area < x)
10^0
10^5
Lake Area (VU^2)

Figure 2:



10^8
10^6
10^4
10^2
10^0
Number (Area < x)
10^2
10^4
10^6
10^8
10^10
Lake Area (m^2)

Figure 3:

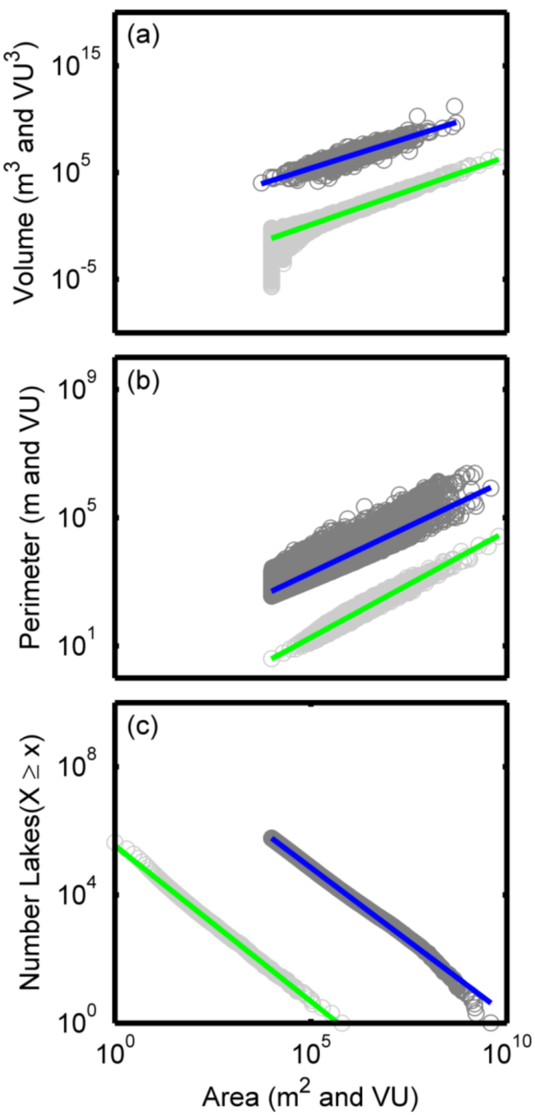

(a)
Volume (m^3 and VU^3)
10^{15}
10^5
10^{-5}
(b)
Perimeter (m and VU)
10^9
10^5
10^1
(c)
Number Lakes(X ≥ x)
10^8
10^4
10^0
10^0
10^5
10^{10}
Area (m^2 and VU)

Figure 4:

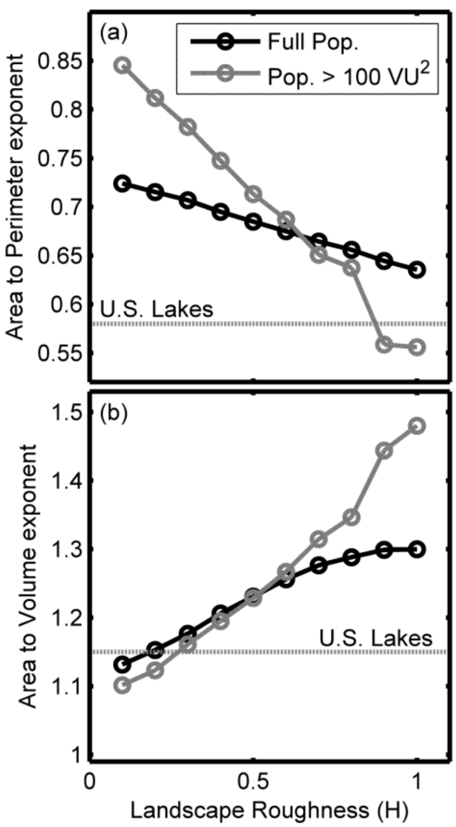

(a)
Full Pop.
Pop. > 100 VU^2
Area to Perimeter exponent
0.85
0.8
0.75
0.7
0.65
0.6
0.55
U.S. Lakes
(b)
Area to Volume exponent
1.5
1.4
1.3
1.2
1.1
1
U.S. Lakes
0
0.5
1
Landscape Roughness (H)

Figure 5:

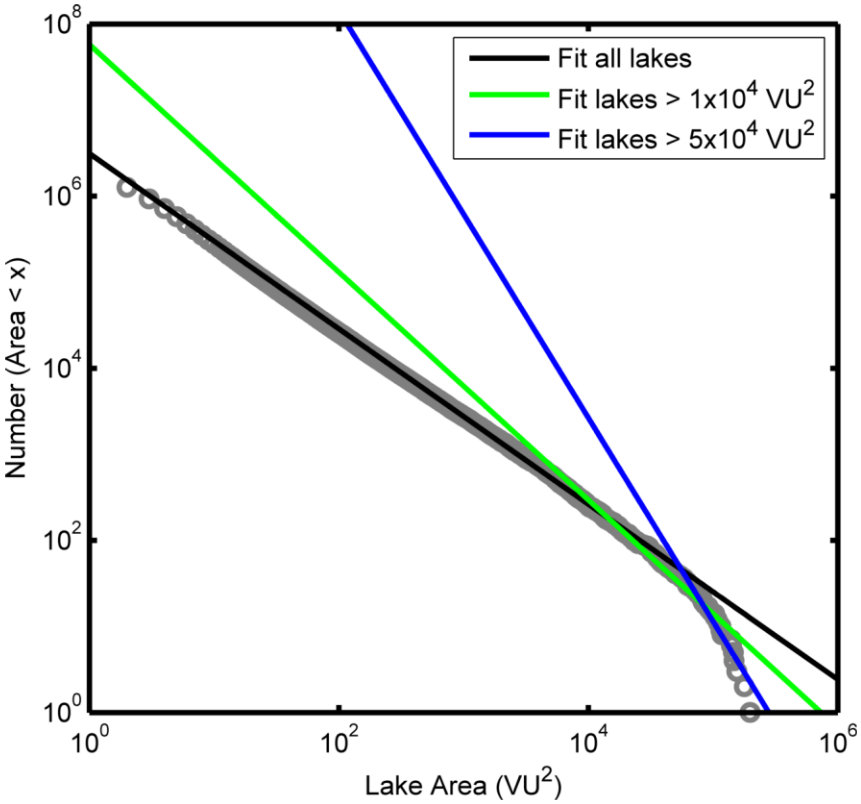

Fit all lakes
Fit lakes > 1x10^4 VU^2
Fit lakes > 5x10^4 VU^2
Number (Area < x)
Lake Area (VU^2)
10^0
10^2
10^4
10^6
10^8

Figure 6:

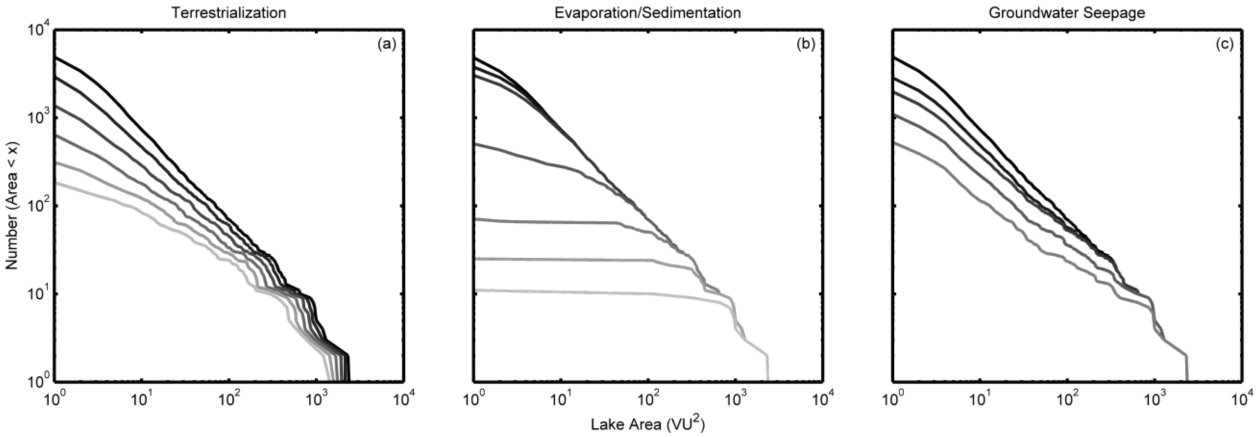
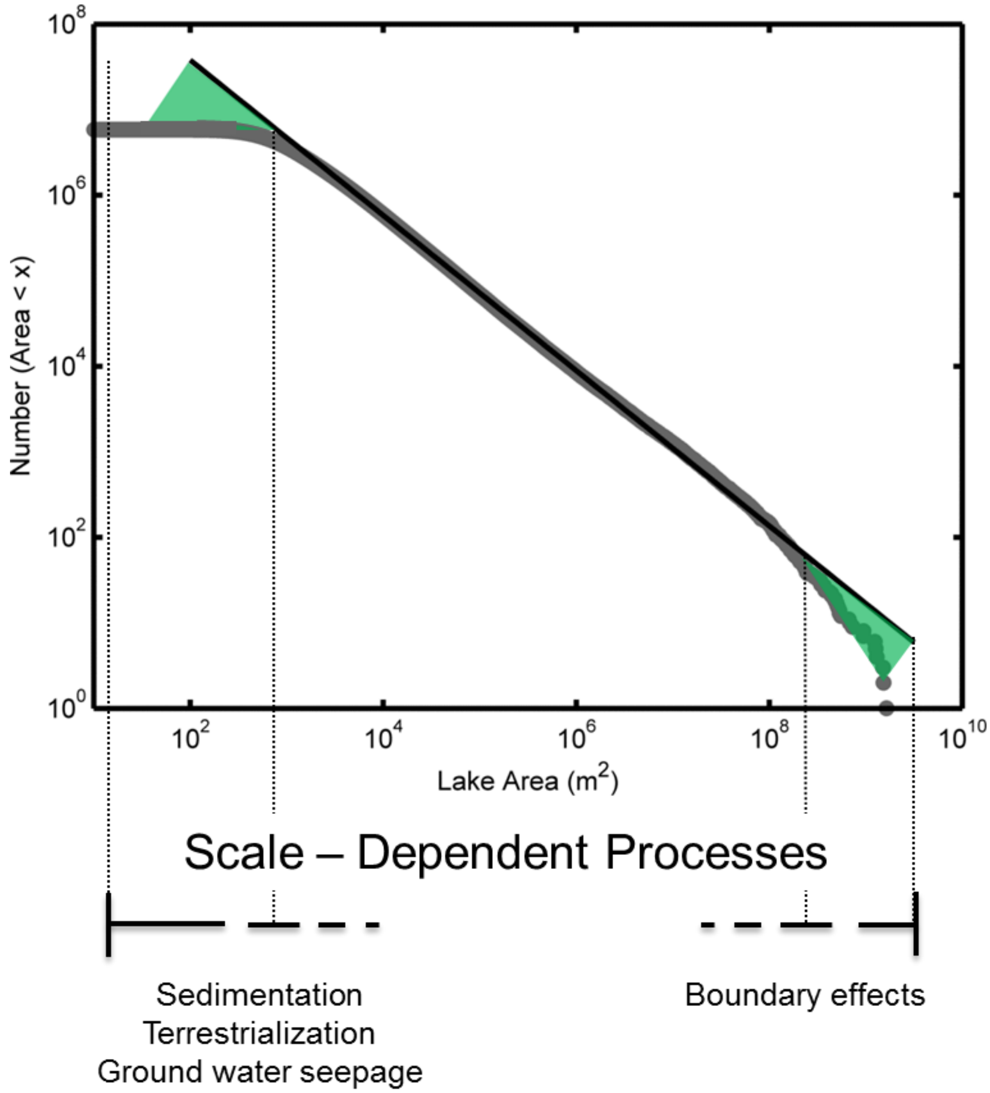

Terrestrialization
Evaporation/Sedimentation
Groundwater Seepage
(a)
(b)
(c)
Number (Area < x)
Lake Area (VU^2)
10^0
10^1
10^2
10^3
10^4

Figure 7:


10^8
10^6
10^4
10^2
10^0
Number (Area < x)
10^2
10^4
10^6
10^8
10^10
Lake Area (m^2)
Scale – Dependent Processes
Sedimentation
Terrestrialization
Ground water seepage
Boundary effects

CHAPTER 5 - SMALL LAKES SHOW MUTED CLIMATE CHANGE SIGNAL IN DEEP-WATER TEMPERATURES

*Luke A. Winslow[1], Jordan S. Read[2], Gretchen J.A. Hansen[3], Paul C. Hanson[1]

Target Journal: Geophysical Research Letters

1) Center for Limnology, University of Wisconsin - Madison, Madison, WI, U.S.A.

2) U.S. Geological Survey, Center for Integrated Data Analytics, Middleton, WI, U.S.A.

3) Wisconsin Department of Natural Resources, Science Services, Madison, WI, U.S.A.

* Corresponding author. Email: lawinslow@wisc.edu

Running Head: Small lake temperature trends

ABSTRACT (<150)

[1] Water temperature observations were collected from 142 lakes across Wisconsin, U.S.A. to examine variation in temperature of lakes exposed to similar regional climate. Whole lake water temperatures increased across the state from 1990 to 2012, with an average trend of 0.055 °C yr^{-1} ± 0.01 °C yr^{-1}. In large (> 0.5km^2) lakes, the positive temperature trend was similar across all depths. In small lakes (<0.5 km^2), the warming trend was restricted to shallow waters, with no significant temperature trend observed in water >0.5 times the maximum lake depth. The differing response of small versus large lakes is likely a result of wind sheltering reducing turbulent mixing magnitude in small lakes. These results demonstrate that numerically dominant small lakes respond differently to climate change than large lakes, suggesting that current predictions of impacts to lakes from climate change may require modification.

1. INTRODUCTION

Most research addressing the impacts of global warming on lake temperature has focused on relatively large lakes, while neglecting to test whether small lakes show similar responses. New techniques for estimating lake surface temperature using satellite infrared imagery are constrained by image resolution, and as a result, a recent global analysis by Schneider and Hook [2010] focused solely on warming trends in lakes larger than 500 km^2. Long-term records of in-situ temperature measurements are also more common for large lakes and have revealed warming patterns in the largest lakes in the world [*O'Reilly et al.*, 2003; *Austin and Colman*, 2007; *Hampton et al.*, 2008]. Examples of warming analyses in medium-sized lakes include lakes with areas of 65km^2 [*Livingstone*, 2003], 13.3 km^2 [*Carvalho and Kirika*, 2003], and 0.77 km^2 [*Tanentzap et al.*, 2008]. Comparatively fewer long-term temperature data collection efforts exist for smaller inland lakes, especially lakes smaller than 0.5 km^2. Exceptions to this pattern include long-term monitoring programs for lake districts that include small lakes [*Mannio and Vuorenmaa*, 1995; *Schindler et al.*, 1996; *Magnuson et al.*, 2006]. Because the majority of lakes in the world are small [*Downing et al.*, 2006; *Winslow et al.*, 2014], analyses of climate change impacts restricted to large lakes may not capture temperature trends representative of the majority of the world's lakes.

The physical and chemical attributes that tend to differentiate small lakes from large lakes are also important controls on water temperatures and may influence warming trends. Small lakes tend to have higher DOC concentrations [*Hanson et al.*, 2007], resulting in different vertical distributions of thermal energy due to faster extinction of light in the water column [*Read and Rose*, 2013]. Small lakes are often well sheltered from wind [*Markfort et al.*, 2010] and consequently, have less wind-driven turbulent mixing that large lakes [*Read et al.*, 2012]. Lastly, lake size is correlated with several geometric properties of lakes, as small lakes typically have less complex morphometry [*Winslow et al.*, 2014] and shallower depths [*Sobek et al.*, 2011]. These factors may interact in complex ways to change how energy is gained, stored and lost in lakes. As a result, small lakes may express different thermal responses to climate warming than larger lakes.

Small lake research has lagged behind the study of large lakes for much of the past century [*Downing*, 2010], resulting in a knowledge deficit about small lakes and their response to long-term change. To bring a greater understanding of the impacts climate change may have on smaller, infrequently monitored lakes, we examine water temperature trends in a large population of lakes across the state of Wisconsin, U.S.A. We address the following questions: How have water temperatures changed across a population of lakes in Wisconsin? Do temperature trends differ with lake size, depth or water clarity? By answering these questions, we can better understand how climate change impacts the large and diverse population of the world's lakes.

2. Data

The water temperature data used in this study were from the Wisconsin Department of Natural Resources (WDNR) and the North Temperate Lakes Long-term Ecological Research (NTL-LTER) program. The WDNR maintains a record of lake monitoring data including observations collected by both WDNR staff and citizen volunteers from around the state. The sampling frequency and temporal extent of the WDNR data varies between lakes, though the majority of data were collected between 1990 and 2013. NTL-LTER data are split geographically across Wisconsin, covering 7 lakes in the north and 4 lakes in the south, with bi-weekly and monthly sampling throughout the open water and ice-covered seasons respectively. The northern NTL-LTER lakes data span 1984-2013 while the southern lakes have data from 1994-2013. Combined, the dataset contained 889 lakes, which was later reduced to 142 lakes based on a set of criteria discussed below.

Depth and clarity data were collected from a variety of sources, and lake surface area and elevation were calculated from WDNR hydrography geospatial dataset. Maximum depth (Z_{max}) was provided by the WDNR and NTL-LTER. When missing from lakes in the WDNR database, Z_{max} was manually digitized from historical bathymetric maps. Satellite-derived Secchi depth (an index of water clarity) estimates were provided by the WDNR (lakesat.org) and were based on Landsat imagery [methods of *Torbick et al.*, 2013].

3. Trend Analysis

3.1 Sen's Slope Estimator

We examined both the overall temperature trend across 142 lakes and trends across gradients of observation depth, season, and lake characteristics. Because the data set is large and observations were made by a variety of people, we chose a trend analysis technique that is relatively insensitive to outliers to mitigate the influence of possible errors in the data that may arise from transcription or observation errors. The Theil-Sen slope estimator (referred to here as Sen's slope) is a non-parametric technique for estimating trends that is robust to outliers and non-normality based on Kendall's tau rank correlation. Sen's slope is defined as the median of a set of slopes that join all permutations of observations [*Sen*, 1968]. We refer to this set of slopes (*B*) as a set of "paired-sample slopes".

In its original form, the Sen's slope estimator cannot be used on single timeseries containing a seasonal component. For example, if the original Sen's slope were applied to water temperature observations, some of the paired-sample slopes would be the result of comparing temperatures from January with temperatures from July. To account for this, the seasonal Sen's slope estimator was developed [*Hirsch et al.*, 1982]. In the seasonal version, observations are grouped by a seasonal component (often month) and paired-sample slopes are calculated only between observations from the same season. The equation for the seasonal paired-sample slope (b_{ijk}) is:

$$b_{ijk} = \frac{y_{jk} - y_{ik}}{t_{jk} - t_{ik}}$$

where y is the observed value, t is time, i is the start index, j is the end index, k is the seasonal identifier index. The overall slope is then defined as the median of all b_{ijk} paired-sample slopes.

To examine the trends across all observed lakes, we extended the seasonal estimator to create a seasonal, cross-site estimator. We extended the seasonal Sen's slope estimator by adding a location identifier so that only temperatures collected at the same location and season were compared. The resulting equation for the cross-site, seasonal paired-sample slope (b_{ijks}) is:

$$b_{ijks} = \frac{y_{jks} - y_{iks}}{t_{jks} - t_{iks}}$$

where *s* is the location identifier index.

To analyze the trends across Wisconsin lakes, we used the cross-site seasonal Sen's slope estimator. While the seasonal identifier is often delineated based on the sampling month, this can cause issues in months where the season is rapidly changing. For example, in Wisconsin, lake water temperatures often change rapidly in September, meaning that early-September observations will typically be much warmer than late-September observations. To avoid pairing such temperatures for assessment of annual temperature trends, we defined each sample's season as the week of the year, an approach that results in fewer paired-sample slopes but reduces the issues of calculating slopes across longer periods of rapid change. Sample location was defined as a specific lake at a specific depth rounded to the nearest whole meter. The resulting paired-sample slopes were each linked with attributes (specifically: start and end year, season and location) which allowed us to aggregate the samples along key gradients.

3.2 Air temperature trend analysis

We estimated the statewide average air temperature trend for the state of Wisconsin for comparison with water temperature trends. We retrieved statewide yearly average air temperatures from the Wisconsin State Climatology Office originally calculated by the National Climatic Data Center as of May 2014. The trend in the statewide annual mean air temperature was calculated using the original Sen's slope analysis [*Sen*, 1968] for observations from 1990 to

2013.

3.2 Water temperature trend analysis

The complete dataset contained 352,000 temperature observations from 889 lakes, but these observations were not evenly distributed across lakes or years. The total number of lakes was reduced to 142 as follows. The cross-site, seasonal Sen's slope results could also be biased towards lakes and time periods with more observations. While there were observations from all years from 1970 to 2012, the majority were collected after 1990. To prevent bias caused by uneven sampling across years, only data after 1990 were included. So that each lake contributed the same number of paired-sample slopes, 800 paired-sample slopes were sub-sampled randomly without replacement. All lakes with insufficient observations were dropped. A cutoff of 800 was selected to maximize the overall number of slopes in the analysis. A larger cutoff reduced the lake count more than the larger sub-sample increased the overall number of slopes. Similarly, a smaller cutoff reduced the number of slopes contributed by each lake more than it increased the number of lakes included. Overall, 113,600 paired-sample slopes and 142 lakes remained.

To compare depth-specific temperature trends in lakes with different maximum depths, a relative depth for each sample was calculated by normalizing the sample's observed depth to the max observed depth of the lake, resulting in a relative depth metric spanning from zero to one where zero is the surface and one is the maximum depth of sampling. "deep water" temperatures were defined as relative depths > 0.5 while "shallow water" temperatures were defined as relative depths <= 0.5.

To examine the overall sensitivity of our estimates to a specific set of sampled lakes and time periods, we randomly sampled from the remaining 142 lakes 1000 times and ran the cross-site

Sen's slope analysis on the random subset. For each random subset, we dropped 30% of the lakes and randomly sampled without replacement 800 paired-sample slopes. This eliminated bias due to different total number of observations between lakes and normalized the contribution of each lake on the final result. With the 1000 random subsets we performed two analyses. To examine how trends change across lake area, the median slope and median lake area was calculated across a 30-lake window applied to lakes sorted by area. A Loess regression was calculated across the full set of median slopes and areas. To contrast large and small lakes, a cutoff of 0.5 km^2 was selected based on a visual inspection of the relationship between lake surface area and temperature trend. To compare how trends differ across depth, median trends were calculated for shallow and bottom waters in large and small lakes for each subset. 95% confidence intervals were defined as the 95th percentile ranges of the randomized results.

5. Results

Wisconsin lakes are warming. Across all lakes and all depths, temperature increased at a median rate of 0.055 ± 0.01 °C yr^{-1}. Larger lakes warmed faster than smaller lakes (Figure 1). Across all depths, water temperature in large lakes (>0.5 km^2) increased by a median of 0.06 °C yr^{-1} while that of small lakes (<0.5 km^2) increased by a median of 0.02 °C yr^{-1}. Between 1990 and 2013, statewide annual average air temperature increased at an estimated rate of 0.067 °C yr^{-1}.

Depth-specific temperature trends depended on lake size. The warming rates for large lakes were consistent across all depths, but the deeper waters of small lakes were not warming at a rate significantly different from zero (Figure 2). In lakes larger than 0.5 km^2, the median trends for shallow and deep waters were 0.056 °C yr^{-1} and 0.05 °C yr^{-1} respectively and were not significantly different ($p > 0.05$). In small lakes, the median trends for shallow and deep waters were 0.03 °C yr^{-1} and 0 °C yr^{-1} respectively and were significantly different ($p < 0.01$). Large and small lakes exhibited similar ranges of water clarity, max depth, elevation, latitude and longitude (Table 1).

6. Discussion

Our analysis reveals important size-dependent differences in the lake thermal responses of lakes to climate change even when subjected to similar regional climate drivers. Many investigators have examined the impact of climate change on water temperatures in lakes, but this study is the first to examine temperature trends across the lake depth profile for a large number of lakes spanning a wide range of surface areas. In particular, we found that small lakes (< 0.5 km^2 in surface area) respond differently to climate forcing than large lakes. Large lake warming was consistent throughout the water column and small lakes were warming only in the shallow waters. The different behavior of these small lakes has relevance to the full distribution of U.S. lakes as they represent 99% of lakes by number and 30% of lake surface area [*Winslow et al.*, 2014]. These differences in full-depth warming patterns are important for predicting the impacts of climate change on both biogeochemical cycles [*Marotta et al.*, 2014] and the distribution and available habitat for fishes [*Ficke et al.*, 2007].

Deep and shallow water temperatures are controlled by climate at different times of year. Although shallow waters respond to climate throughout the open water season, the temperatures of bottom waters in stratified lakes are controlled by springtime conditions preceding stratification onset [*Kalff*, 2002]. In large lakes, warmer springtime shallow temperatures are mixed downwards by periodic mixing events, but in small lakes the warming climate signal was not transferred to the bottom waters, as they likely lacked sufficient turbulent mixing fluxes [*Read et al.*, 2012].

Lakes with areas of approximately 0.5 km^2 are in a size range of rapid transition in wind sheltering. Wind sheltering is exponentially related with lake size and changes quickly in lakes with areas between 0.1 and 1 km^2 [See figure 4 in *Hondzo and Stefan*, 1993]. Lakes in forested regions, such as many of the lakes in Wisconsin, can have higher wind sheltering than would be predicted by fetch alone [*Markfort et al.*, 2010]. For lakes in the size range where wind sheltering changes rapidly, bottom water temperatures may be especially sensitive to changes in drivers of wind sheltering. For example, *Tanentzap et al.*, [2008] found that increasing canopy height and decreasing clarity resulted in cooling bottom waters despite regional climate warming. The area of the lake examined (0.77 km^2) was within the highly sensitive 0.1-1 km^2 range. While we report a transition of 0.5 km^2 between warming lakes and lakes with a muted climate signal, there is likely a range of areas across which multiple factors, including fetch, depth, and water clarity, control the coupling between the climate signal and water temperature. Larger datasets, including lakes with high diversity in sheltering, depth, and water clarity would be necessary to improve our understanding of how different lake characteristics moderate the transmission of warming climate signals into the bottom waters of small lakes.

The strength of stratification for this population of lakes increased during the study period. To evaluate stratification, we examined how density trends diverged in surface and bottom waters by calculating density from temperature [*Read et al.*, 2011], assuming zero salinity, and applying the same cross-site, seasonal Sen's slope analysis. The stratification trend for large lakes was due to the nonlinearity of the water temperature and density relationship. Differences between warming rates in shallow and deeper depths were not significant ($p > 0.05$; Fig 2), but the nonlinear relationship between temperature and thermal expansion of water resulted in a significant trend in stratification strength. Stratification trends in small lakes were driven by the combined effects of density nonlinearity and significant differences in the cross-depth temperature trend ($p < 0.01$). Regardless of lake size, these results further support hypotheses that lake stratification strength will increase in response to climate change [*Livingstone*, 2003; *Hadley et al.*, 2013].

While these cross-lake patterns are likely applicable to lakes in other temperate regions lakes, there may be exceptions. Small lakes that are very shallow or very clear may allow for sufficient turbulence to mix the warming climate signal to bottom waters [*Folkard et al.*, 2007]. Similarly, large lakes of extreme depth or strong haline stratification may lack sufficient mixing to connect bottom waters to the surface warming signal [*Hamblin et al.*, 1999]. The results shared here are exclusively from temperate lakes subjected to winter ice cover. Lakes with different properties or lakes in different regions may show different warming patterns.

Our lake population shows a temperature trend similar to other studies [*Livingstone*, 2003; *Schneider and Hook*, 2010]. When trends are aggregated across multiple lakes, then average reported temperature trends were 0.45 ± 0.011 °C yr^{-1} [*Schneider and Hook*, 2010]. Whereas, temperatures for some individual lakes have been reported to increase more rapidly [*Austin and Colman*, 2007; *Schneider et al.*, 2009]. Spatial climate variability may explain differences in reported lake trends. Comparing our results with those from Lake Michigan, a lake in closer proximity and similar latitude to Wisconsin, yields similar reported trends (0.078 ± 0.036 and 0.051 ± 0.034 °C yr^{-1}) [*Austin and Colman*, 2007].

7. Conclusion

The outcomes of this study identify the importance of lake size and the concomitant influence of fetch-dependent turbulent processes on temperature trends. While our results are in general agreement with previous findings for surface waters [*Schneider and Hook*, 2010], we demonstrate how temperature trends in small lakes diverge from expectations based on studies of larger lakes. A warming signal was not detected in the deeper waters of small lakes, likely because small lakes have lower turbulent mixing and transfer less of the climate signal to deeper waters. Large lakes, with greater magnitude of turbulent mixing, showed consistent warming across all depths. The difference in response to climate change by lakes of different sizes must be considered when estimating future impact of climate change on a global lake population dominated by small lakes.

8. Acknowledgements

We would like acknowledge helpful feedback from Samantha Oliver and Ryan Batt. Funding for this research was provided by the U.S. National Science Foundation (NSF) grants DEB-0941510 (Global Lake Ecological Observatory Network) and DEB-0822700 (North Temperate Lakes Long-term Ecological Research program, NTL-LTER), US Geological Survey Award #10909172, and the Wisconsin Department of Natural Resources (WDNR). Special thanks to the WDNR teams and citizen volunteers who collected and curated the dataset presented here. All temperature data are publically available on the WDNR (http://dnr.wi.gov/topic/surfacewater/swims) and NTL-LTER (http://lter.limnology.wisc.edu/) websites.

9. References

Austin, J. A., and S. M. Colman (2007), Lake Superior summer water temperatures are increasing more rapidly than regional air temperatures: A positive ice-albedo feedback, *Geophys. Res. Lett.*, *34*(6), L06604, doi:10.1029/2006GL029021.

Carvalho, L., and A. Kirika (2003), Changes in shallow lake functioning: response to climate change and nutrient reduction, *Hydrobiologia*, *506-509*(1-3), 789–796, doi:10.1023/B:HYDR.0000008600.84544.0a.

Downing, J. A. (2010), Emerging global role of small lakes and ponds: little things mean a lot, *Limnetica*, *1*(29), 9–24.

Downing, J. A., Y. T. Prairie, J. J. Cole, C. M. Duarte, L. J. Tranvik, R. G. Striegl, W. H. McDowell, P. Kortelainen, N. F. Caraco, and J. M. Melack (2006), The global abundance and size distribution of lakes, ponds, and impoundments, *Limnol. Oceanogr.*, *51*(5), 2388–2397, doi:10.4319/lo.2006.51.5.2388.

Ficke, A. D., C. a. Myrick, and L. J. Hansen (2007), Potential impacts of global climate change on freshwater fisheries, *Rev. Fish Biol. Fish.*, *17*(4), 581–613, doi:10.1007/s11160-007-9059-5.

Folkard, A. M., A. J. Sherborne, and M. J. Coates (2007), Turbulence and stratification in Priest Pot, a productive pond in a sheltered environment, *Limnology*, *8*(2), 113–120, doi:10.1007/s10201-007-0207-3.

Hadley, K. R., a. M. Paterson, E. a. Stainsby, N. Michelutti, H. Yao, J. a. Rusak, R. Ingram, C. McConnell, and J. P. Smol (2013), Climate warming alters thermal stability but not stratification phenology in a small north-temperate lake, *Hydrol. Process.*, n/a–n/a, doi:10.1002/hyp.10120.

Hamblin, P. F., C. L. Stevens, and G. A. Lawrence (1999), Simulation of vertical transport in mining pit lake, *J. Hydraul. Eng.*, *125*(10), 1029–1038.

Hampton, S. E., L. R. Izmest'Eva, M. V. Moore, S. L. Katz, B. Dennis, and E. a. Silow (2008), Sixty years of environmental change in the world's largest freshwater lake - Lake Baikal, Siberia, *Glob. Chang. Biol.*, *14*(8), 1947–1958, doi:10.1111/j.1365-2486.2008.01616.x.

Hanson, P. C., S. R. Carpenter, J. A. Cardille, M. T. Coe, and L. A. Winslow (2007), Small lakes dominate a random sample of regional lake characteristics, *Freshw. Biol.*, *52*(5), 814–822, doi:10.1111/j.1365-2427.2007.01730.x.

Hirsch, R. M., J. R. Slack, and R. a. Smith (1982), Techniques of trend analysis for monthly water quality data, *Water Resour. Res.*, *18*(1), 107–121, doi:10.1029/WR018i001p00107.

Hondzo, M., and H. Stefan (1993), Lake water temperature simulation model, *J. Hydraul. Eng.*, *119*(11), 1251–1273.

Kalff, J. (2002), Temperature Cycles, Lake Stratification, and Heat Budgets, in *Limnology: Inland Water Ecosystems*, pp. 154–178, Prentice Hall, New Jersey.

Livingstone, D. (2003), Impact of secular climate change on the thermal structure of a large temperate central European lake, *Clim. Change*, *57*, 205–225.

Magnuson, J. J., T. K. Kratz, and B. J. Benson (2006), *Long-term Dynamics of Lakes in the Landscape: Long-term Ecological Research on North Temperate Lakes*, edited by J. J. Magnuson, T. K. Kratz, and B. J. Benson, Oxford University Press, New York.

Mannio, J., and J. Vuorenmaa (1995), Regional Monitoring of lake acidification in Finland, *Water. Air. Soil Pollut.*, *85*, 571–576.

Markfort, C. D., A. L. S. Perez, J. W. Thill, D. a. Jaster, F. Porté-Agel, and H. G. Stefan (2010), Wind sheltering of a lake by a tree canopy or bluff topography, *Water Resour. Res.*, *46*(3), W03530, doi:10.1029/2009WR007759.

Marotta, H., L. Pinho, and C. Gudasz (2014), Greenhouse gas production in low-latitude lake sediments responds strongly to warming, *Nat. Clim. Chang.*, *4*(May), 11–14, doi:10.1038/NCLIMATE2222.

O'Reilly, C. M., S. R. Alin, P.-D. Plisnier, A. S. Cohen, and B. a McKee (2003), Climate change decreases aquatic ecosystem productivity of Lake Tanganyika, Africa., *Nature*, *424*(6950), 766–8, doi:10.1038/nature01833.

Read, J. S., and K. C. Rose (2013), Physical responses of small temperate lakes to variation in dissolved organic carbon concentrations, *Limnol. Oceanogr.*, *58*(3), 921–931, doi:10.4319/lo.2013.58.3.0921.

Read, J. S., D. P. Hamilton, I. D. Jones, K. Muraoka, L. A. Winslow, R. Kroiss, C. H. Wu, and E. Gaiser (2011), Derivation of lake mixing and stratification indices from high-resolution lake buoy data, *Environ. Model. Softw.*, *26*(11), 1325–1339, doi:10.1016/j.envsoft.2011.05.006.

Read, J. S. et al. (2012), Lake-size dependency of wind shear and convection as controls on gas exchange, *Geophys. Res. Lett.*, *39*(9), L09405, doi:10.1029/2012GL051886.

Schindler, D. W., S. E. Bayley, B. R. Parker, K. G. Beaty, D. R. Cruikshank, E. J. Fee, E. U. Schindler, and M. P. Stainton (1996), The effects of climatic warming on the properties of boreal lakes and streams at the Experimental Lakes Area, northwestern Ontario, *Limnol. Ocean.*, *41*(5), 1004–1017, doi:10.4319/lo.1996.41.5.1004.

Schneider, P., and S. J. Hook (2010), Space observations of inland water bodies show rapid surface warming since 1985, *Geophys. Res. Lett.*, *37*, doi:10.1029/2010GL045059.

Schneider, P., S. J. Hook, R. G. Radocinski, G. K. Corlett, G. C. Hulley, S. G. Schladow, and T. E. Steissberg (2009), Satellite observations indicate rapid warming trend for lakes in California and Nevada, *Geophys. Res. Lett.*, *36*(22), L22402, doi:10.1029/2009GL040846.

Sen, P. (1968), Estimates of the regression coefficient based on Kendall’s tau, *J. Am. Stat. Assoc.*, *63*(324), 1379–1389.

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

Tanentzap, A., N. Yan, and B. Keller (2008), Cooling lakes while the world warms: Effects of forest regrowth and increased dissolved organic matter on the thermal regime of a temperate, urban lake, *Limnol. Oceanogr.*, *53*(1), 404–410.

Torbick, N., S. Hession, S. Hagen, N. Wiangwang, B. Becker, and J. Qui (2013), Mapping inland lake water quality across the Lower Peninsula of Michigan using Landsat TM imagery, *Int. J. Remote Sens.*, *34*(21), 7607–7624, doi:10.1080/01431161.2013.822602.

Winslow, L. A., J. S. Read, P. C. Hanson, and E. H. Stanley (2014), Lake shoreline in the contiguous United States: quantity, distribution and sensitivity to observation resolution, *Freshw. Biol.*, *59*(2), 213–223, doi:10.1111/fwb.12258.

FIGURES AND TABLES

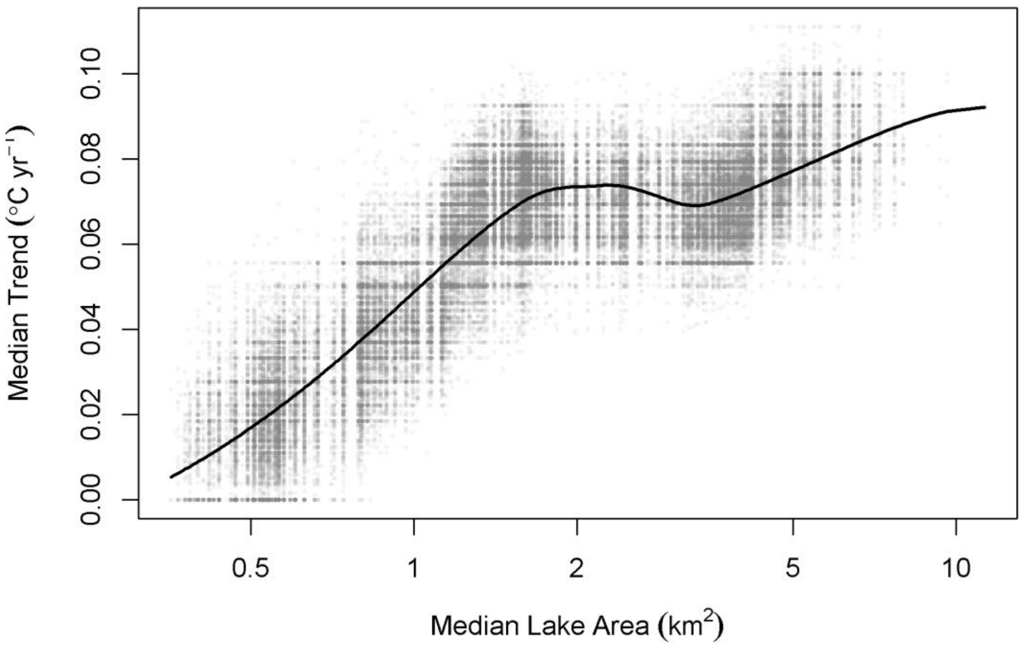


Figure 1: Median temperature trends across different lake sizes. Black line is a loess regression through all random subsets. The dataset was randomly sampled 1000 times to eliminate potential bias introduced by any single lake or dataset. Gray points are median temperature slope and median lake area for a moving window of 30 lakes for each random sample. X-axis is log scaled.

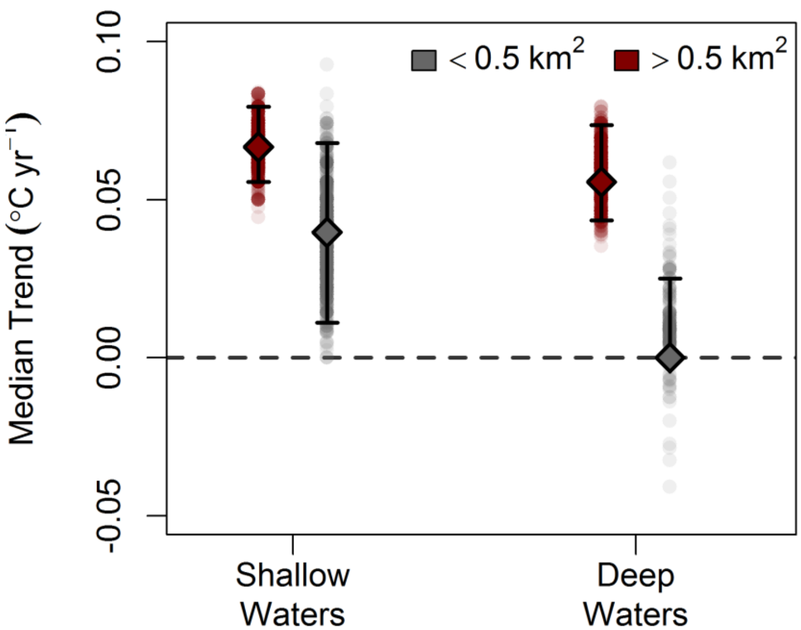


Figure 2: Median temperature trends for shallow (relative depth < 0.5) and deep (relative depth > 0.5) waters split into large and small lake populations (0.5 km^2 cutoff). The dataset was randomly sampled 1000 times to eliminate bias introduced by any single lake or dataset. All random subsample medians shown (circles) with overall median (diamond) and 95th percentile limits indicated (lines).

Table 1: Key characteristics of lakes represented in dataset. Format shown is *median [minimum, maximum]*.

Lake Characteristics	Large Lakes (> 0.5 km^2)	Small Lakes (< 0.5 km^2)
Number (#)	118	24
Area (km^2)	2.1 [0.51, 533]	0.32 [0.006, 0.49]
Z_{max} (m)	15 [2.4, 72]	11 [2.5, 20]
Secchi Depth (m)	2.6 [0.5, 8.2]	3.3 [0.8, 5.6]
Elevation (m)	360 [180, 527]	295 [248, 537]
Latitude (°N)	45.5 [42.6, 46.5]	44.7 [42.8, 46]
Longitude (°E)	89.6 [87.2, 92.6]	89.2 [88, 92.6]

APPENDIX A - SMALL LAKES SHOW MUTED CLIMATE CHANGE SIGNAL IN DEEP-WATER TEMPERATURES

by Luke A. Winslow, Jordan S. Read, Gretchen J.A. Hansen, Paul C. Hanson

A.1 Temperature trends across seasons

While the focus of this manuscript is on a broad-scale examination of water temperature trends in lakes, the variability of individual lake responses and how temperature trends may differ seasonally are important to consider. This variability is of interest from a mechanistic perspective, and also in how it may affect the performance of the cross-site, seasonal Sen's slope method. Here, we present further exploration into the seasonal and cross-site patterns of temperature trends in Wisconsin lakes.

A.1.1 Seasonal Trends

As seasonal Sen's slope can be affected by seasonal differences in trends [*Hirsch et al.*, 1982], we examined our dataset to see if seasonality in trends could introduce bias into the overall results. We plotted trends grouped by week of the year (Figure S1 and found no discernable seasonal pattern in trend. The large uncertainties and extreme values in the early and late year trends are likely due to the small number of observations in these seasons and should be ignored. Throughout the rest of the year, there is no clear season or month with clearly different trends that would affect the overall results presented. Further work is required to understand the specific drivers of the temperature trends and why they are so seasonally consistent.

A.1.2 Monthly air and water temperature changes correlated

Inter-annual climate variability offers a natural gradient temperature to help understand how these lakes respond to climate variability. To determine how inter-annual variability in air temperatures within specific seasons correlated with inter-annual variability in water temperatures within seasons and across different depths, we built a month-specific, inter-annual variability comparison.

Because our data are span many different lakes, we could not simply correlate annual water temperature with air temperature and instead built the analysis around our cross-site seasonal Sen's slope method. We tested how the inter-annual variability month-specific water temperatures correlated with the inter-annual variability of month-specific water temperatures (e.g., how differences in June water temperatures correlated with differences in June air temperatures). To do this, we used monthly mean air temperatures for the state of Wisconsin and calculated paired-sample slopes for each month for all permutations of years. For the lakes, we used our calculated paired-sample slopes, split them into shallow water and deep water slopes using a relative depth cutoff of 0.5 and further split them into large and small lakes by an area cutoff of 0.3 km^2. For each group, the month-specific, inter-annual water temperature slopes across all lakes was joined with the inter-annual air temperature slopes. This allowed us to examine how inter-annual variability in air temperature for any month of the year predicts the inter-annual variability in water temperature for any month. For example, inter-annual variability in air temperature in March was correlated with variability in water temperature for months March to December to see if water temperature variability in later months is predicted by March air temperatures. For each pair of months, the Pearson's correlation coefficient was calculated and plotted on a grid with color indicating the strength of correlation. The correlation direction was ignored.

The inter-annual air and water temperature analysis highlighted several important points. First, surface water temperatures across examined large and small lakes are generally driven by immediate, month-to-month climate (Figure S2a-b). The strong correlations along the diagonal show that surface water variability is driven by the same month's climate variability while previous months have little explanatory power. This is in contrast to findings from Lake

Superior, showing strong correlation between springtime ice conditions and summertime temperatures [*Austin and Colman*, 2007]. Our results suggest that the surface water temperatures in these lakes have little thermal inertia and respond relatively quickly to changing meteorological drivers.

While surface waters show little memory, bottom water temperatures in large lakes show substantial memory of springtime conditions (Figure S2c). Air temperature variability in the months of both March and April shows considerable influence on bottom-water temperatures. This influence extends through the stratified period. Conversely, air temperature variability has little influence on bottom water temperatures during the commonly stratified period (May to August). This confirms long-understood concepts of physical limnology (Some classic references). Lastly, small-lake bottom water temperatures show little correlation with climate variability of any month (Figure S2d). The lack of connectedness to climate helps explain the lack of climate signal shown in the manuscript's analysis.

A.2 Cross-lake Trends

The analysis presented in this manuscript aggregates observations from a large number of lakes to focus on broad-scale, cross-lake pattern and does not focus on patterns in any individual lake. Despite the broad-scale focus, showing individual-lake trends and presenting detail on a small number of well-studied lakes helps present the pattern seen at the broad scale.

Figure S3 shows the depth-specific trends for three lakes with almost three decades of consistently sampled and calibrated water temperature data from the North Temperate Lakes Long Term Ecological Research group. In the two large-lake examples (Figure S3a-b), the median temperature change is consistent across depths, with an overall median value of 0.046 °C

yr^{-1} for Trout Lake (a) and 0.05 °C yr^{-1} for Sparkling Lake (b). For the lake under 0.5 km^2, Crystal Lake (c), the trend was similar in the surface waters, changing towards a trend of 0 between depths of 9 to 12 meters. Overall median trend of Crystal lake is 0.018 °C yr^{-1}, though this value is the combination of the steeper trends in shallow waters and near-zero trends in deep waters. Splitting the lake into shallow and deep waters, the median trend for shallow depths (< 9m) and for deep depths (> 12m) are 0.05 and 0 °C yr^{-1} respectively. The deviation in trend in Sparkling lake (Figure S3b) between depths of 7 to 10 meters corresponds with the generally observed thermocline depth and may indicate a deepening trend in upper mixed layer of the lake, though we did not explore it in detail.

Calculating lake-specific trends for all 142 lakes is challenged by the heterogeneous nature of the dataset. While some of the lakes are very consistently sampled throughout the season, others have observations for only one or two times in the year. These lakes with low numbers of observation yielded highly variable median trends. Figure S4 shows a map of the lakes in the analysis and uses color to indicate median overall temperature trend. Lake specific trends were highly variable (25% and 75% quantiles: -0.027 and 0.187) and included lakes with both positive and negative trends. The 39 lakes with large numbers of observations (large circles) showed more constrained lake-specific trends (25% and 75% quantiles: 0.006 and 0.1 °C yr^{-1}) than all 142 lakes, though the median trend was the same for both groups. This large variability in lake-specific trends shows how the cross-site seasonal Sen's slope method allowed us to use more of the dataset for the overall trend analysis than would have been possible if trends were calculated at the individual lake scale.

A.3 References

Austin, J. A., and S. M. Colman (2007), Lake Superior summer water temperatures are increasing more rapidly than regional air temperatures: A positive ice-albedo feedback, *Geophys. Res. Lett.*, *34*(6), L06604, doi:10.1029/2006GL029021.

Hirsch, R. M., J. R. Slack, and R. a. Smith (1982), Techniques of trend analysis for monthly water quality data, *Water Resour. Res.*, *18*(1), 107–121, doi:10.1029/WR018i001p00107.

A.4 Figures

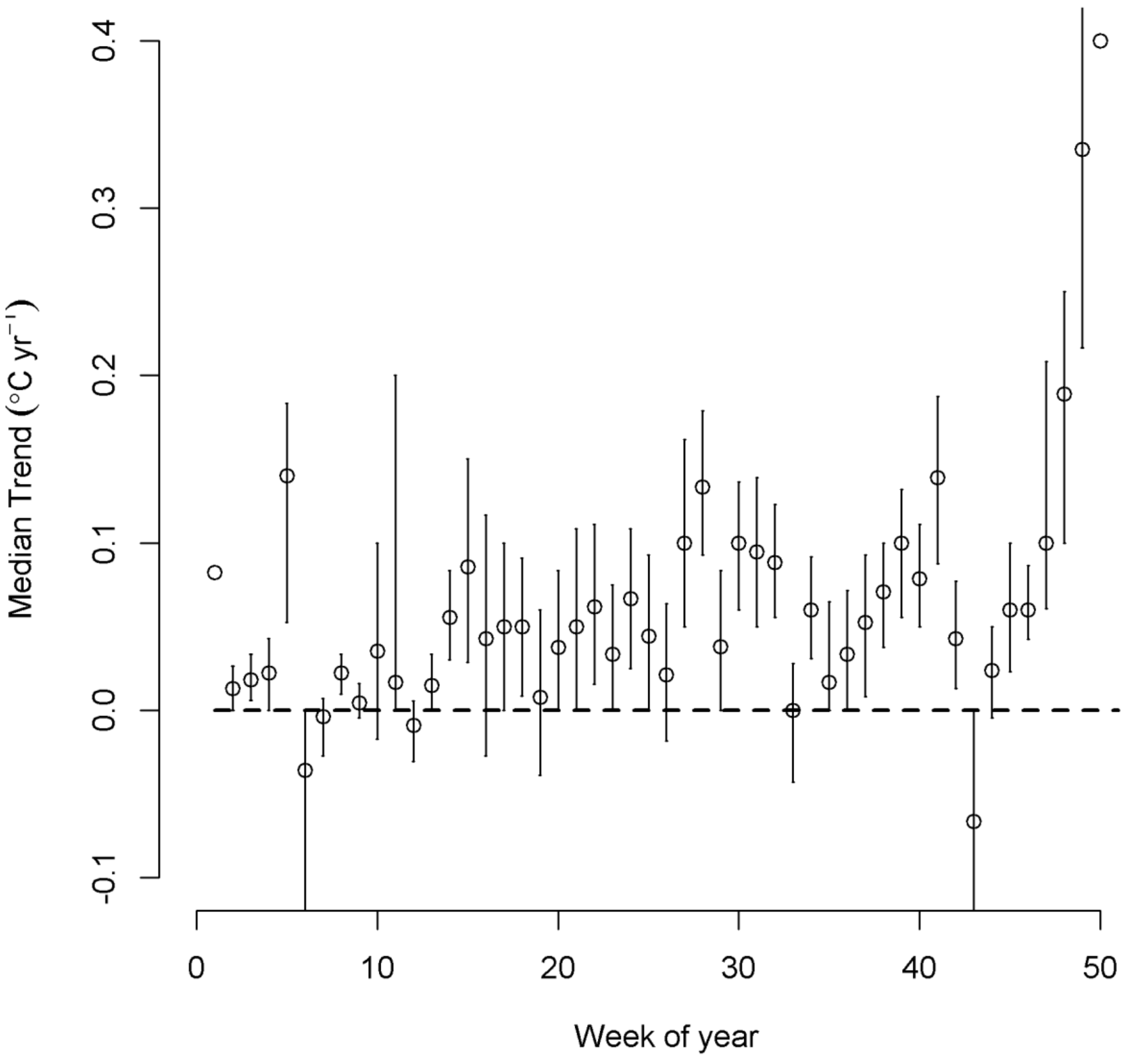


Figure S1: Median paired-value slopes grouped by the week of the year. Lines show 95% confidence intervals calculated using the method described by [*Hirsch et al.*, 1982].**The winter and late fall edges of the year have fewer observations and show higher variability in the medians.**

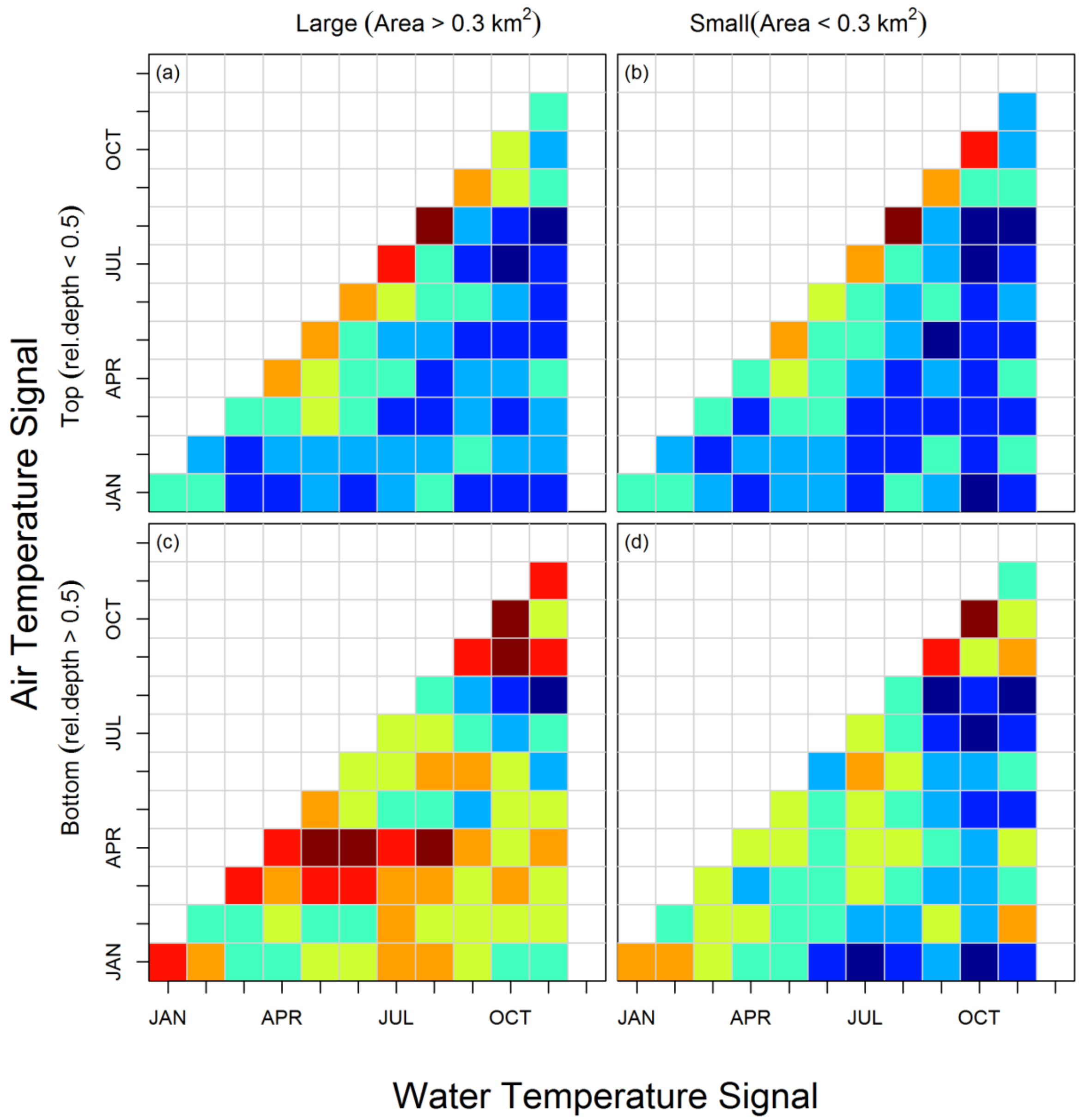


Figure S2: Relationship between inter-annual variability in water temperature (x-axis) and state-wide air temp (y-axis). (a-b) Show shallow water temperatures while (c-d) show bottom water temperatures. Darker red indicates strong correlation while yellow/blue indicates weak or no correlation. Direction of correlation is not shown.

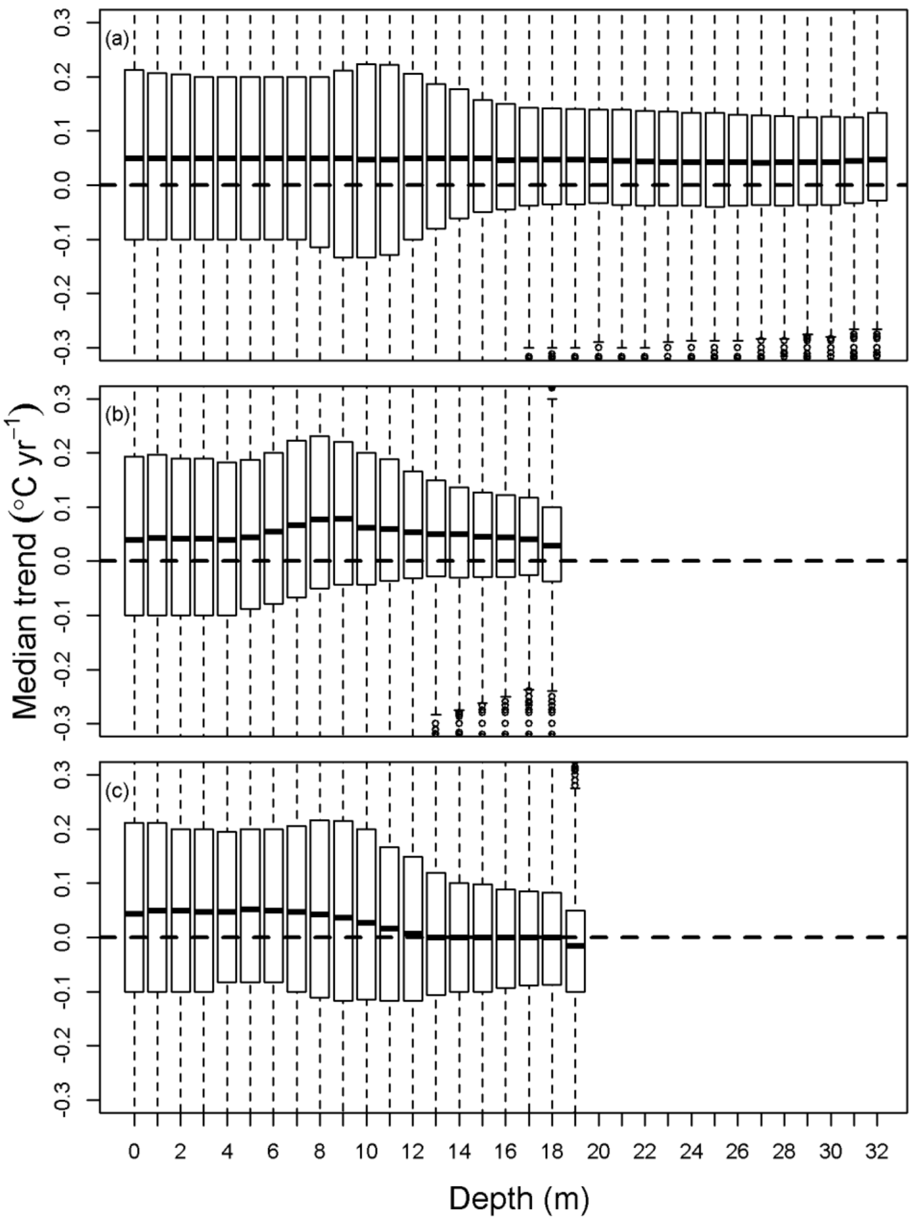


Figure S3: Median paired-sample slope trends grouped by relative depth for two exemplar lakes. (a) Shows trends from Trout Lake (area: 15.6 km^2, average Secchi: 4.7 m), (b) Sparkling Lake (area: 0.6 km^2, average Secchi: 6.2 m) and (c) Crystal Lake (area: 0.4 km^2, average Secchi: 7.5 m). All lakes are from the North Temperate Lakes, Long-Term Ecological Research site, are in close proximity (within 20 km) and have routinely sampled and calibrated observations spanning 30 years from 1985 to 2014.

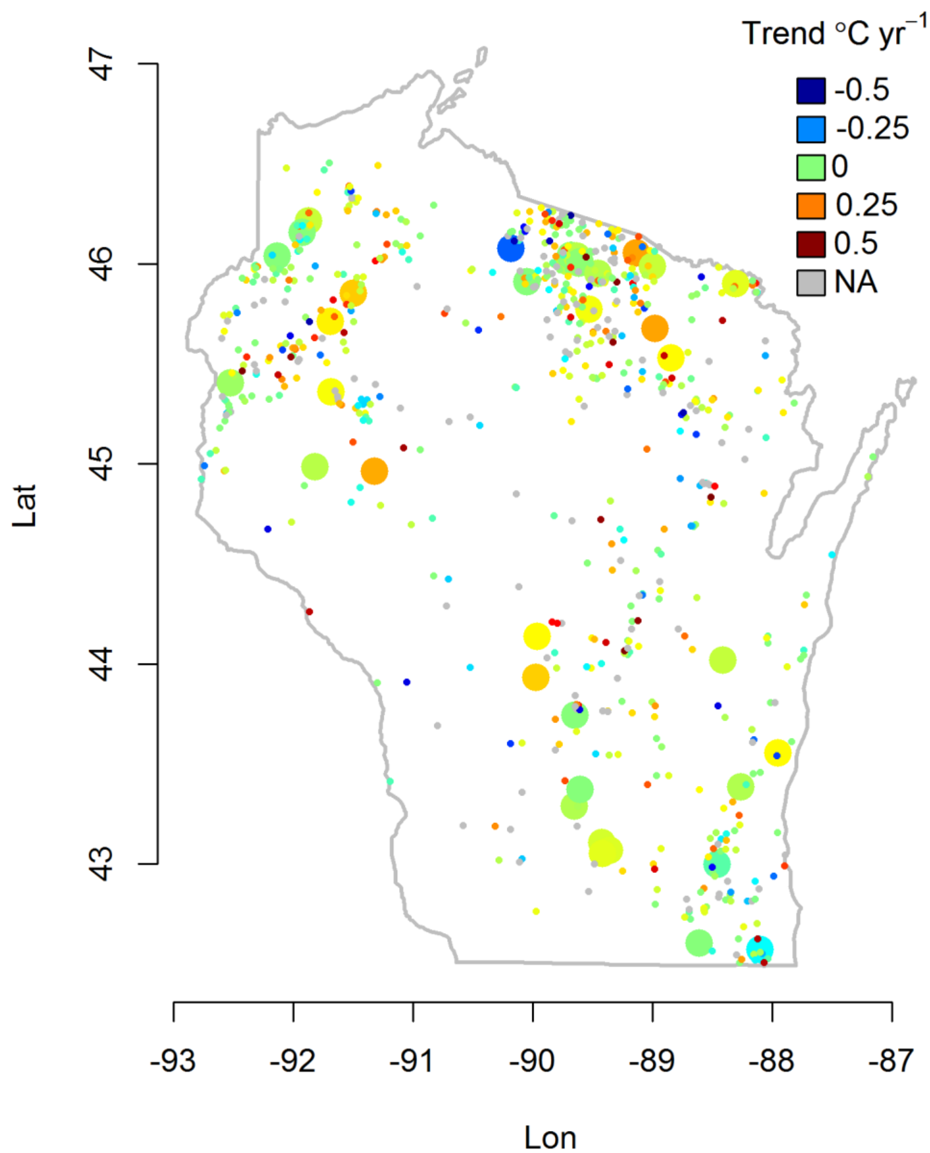


Figure S4: Map of lake-specific temperature trends. Small points are lakes with less than 3,000 paired-sample slopes. Large points are lakes with greater than 3,000 paired-sample slopes. Grey points are lakes with median trends that fell outside of ± 0.5 °C yr-1.