Scaling relationships between lake surface area and catchment area

Jonathan A. Walter[1], Rachel Fleck[2], Michael L. Pace[1], Grace M. Wilkinson[2]

[1]Department of Environmental Sciences, University of Virginia
[2]Department of Ecology, Evolution, and Organismal Biology, Iowa State University

Corresponding Author:
Jonathan A. Walter
Department of Environmental Sciences
University of Virginia
291 McCormick Rd
Charlottesville, VA 22904
jaw3es@virginia.edu

Key Words: watershed, power law, drainage ratio

Abstract

Scaling relationships, including power laws, provide quantitative predictions used in basic and applied sciences. We investigated scaling relationships between catchment area and lake surface area, the ratio of which has important implications for terrestrial-aquatic linkages. Synthesizing evidence from 9 datasets from three continents, we show that there is an approximately linear relationship between lake surface area and catchment area, and that reservoirs and other human-made lakes tend to have larger catchments than natural lakes. Using the example of DOC export from forested catchments, we illustrate how the relationships observed in this study can be used to provide first-order estimates of ecosystem processes coupling lakes and their catchments.

Introduction

Across the environmental and biological sciences, many well-known scaling relationships reflect fundamental processes and provide quantitative predictions useful in research, conservation, and natural resource management. For example, in numerous types of measurements the variance of a set increases as its mean also increases according to a power-law relationship, known in ecology as Taylor's law (Taylor 1961). A multiplicity of biotic and abiotic variables, from the abundances of insect pests to human demography to Tornado outbreaks, conform to Taylor's law (Taylor 1961, Downing 1979, Xu et al. 2014, Bohk et al. 2015, Lagrue et al. 2015, Tippett and Cohen 2016, Reuman et al. 2017, Zhao et al. 2019), and knowledge of this scaling relationship has wide-ranging usefulness, from design of sampling protocols (Downing 1979, Taylor 2018), to scaling process measurements up in time and space, to quantifying portfolio effects (Doak et al. 1998, Hallett et al. 2014). Similar power-law relationships may hold for features of aquatic ecosystems. Two examples are the relationships between lake size and lake abundance (Downing et al. 2006, Seekell and Pace 2011) and stream network lengths and basin area (Shreve 1974, Robert and Roy 1990). There are many more small lakes than large lakes, and the logarithm of abundance declines approximately linearly with the logarithm of surface area, at least over some intervals (Downing et al. 2006, Seekell and Pace 2011). Basin area tends to increase with stream network length according to a power-law (Shreve 1974, Robert and Roy 1990).

Here, we consider another possible power-law relationship, between lake surface area and catchment area. The ratio of catchment area to surface area, also known as the drainage ratio, is a commonly used indicator of the influence of catchment processes on lake dynamics and is related to factors including nitrogen removal (Harrison et al. 2009), carbon inputs (Sobek et al.

2007, Xenopoulos et al. 2013), carbon burial (Downing et al. 2008), and lake color (Rasmussen et al. 1989). However, surprisingly little has been reported about how catchment area scales with lake surface area. Reservoirs tend to have larger catchments than natural lakes, even relative to lake size (Hayes et al. 2017), and drainage ratios may differ between regions representing different geophysical settings (Kortelainen 1993). However, whether there is a power-law scaling relationship between lake area and catchment area, and whether this relationship is consistent across geographic regions and lake types (e.g., natural, reservoir) is unclear.

This study synthesizes evidence from 6542 lakes across 9 datasets from North America, Europe, and New Zealand to examine: 1) does catchment area scale with lake surface area? 2) do relationships between catchment area and lake area differ geographically or according to lake type? We find approximately linear (i.e. power-law) relationships between lake surface areas and catchment areas on logarithmic scales. These relationships change according to lake type, with reservoirs and other human-made lakes having consistently larger catchments relative to lake surface area. These relationships can be used to scale up local estimates of important ecosystem processes and to derive first-order approximations of important quantities, which we illustrate using dissolved organic carbon (DOC) export from forested watersheds as an example.

Methods

We analyzed using linear regression the scaling relationship between catchment area and lake surface area in nine datasets representing different regions and areal extents: the Adirondacks (New York, USA), the Eastern Lake Survey (USA), Iowa, Ohio, Finland, Florida, New Zealand, the Western Lake Survey (USA) and the USA National Lakes Assessment (NLA). Where possible, given sample sizes and provided information, we considered whether these relationships differed by region and/or lake type attributes by asking whether model fit was improved by including these factors as random effects in linear mixed effects regression. Except when particular groups had too few members for robust analysis, we used lake type designations as provided in the original datasets. The first eight datasets were used primarily to assess the variety and consistency of scaling relationships across diverse ecological and geomorphic settings. Due to its sampling design, geographic scope, and detailed lake type information, we gave greater weight to the NLA data, and used this survey in particular to compare the relative strength of variability attributable to geography versus lake type. A summary of lake type designations for each dataset is provided in Table 1.

Datasets

The Adirondacks data were generated by the Adirondacks Lake Survey (Kretser et al. 1989, Baker et al. 1990) and includes data on 1468 lakes in Adirondack Park, a mountainous ≈24000 km^2 region of northeast New York state. The dataset contains a representative sample of ≈50% of lakes in the region, ranging in surface area from 0.2 to 203 ha. Lake type information was not available.

The Iowa data come from the Iowa Ambient Lakes Monitoring Program which monitors water quality of "significant publicly-owned" lakes in the state of Iowa, or lakes and impoundments greater than 4 ha and capable of supporting a sport fish stock of at least 225 kg per hectare. The dataset includes 123 lakes located over the ≈150000 km^2 state. We used lake type information in the dataset to designate Iowa lakes as human-made (borrow pits, impoundments) or natural.

The Ohio data were obtained from Hayes and Vanni (2018), and include 86 lakes over the ≈115000 km^2 state, ranging from 7 to 6500 ha in surface area. Lakes types in this dataset included tributary reservoirs, canal reservoirs, permanently filled quarries, and natural lakes. We considered differences between reservoirs (tributary and canal) and other lakes; permanently filled quarries and natural lakes were not separately identified in the dataset.

The Finnish data were obtained from the Finnish Environment Institute and are described in Forsius et al. (1990) and Kortelainen (1993). These data include 976 lakes over the ≈340000 km^2 area of Finland. The lakes span 1 to 750 ha in surface area, and are predominantly drainage (70%) or headwater lakes (17%) (Kortelainen 1993). Lake type information was not provided.

The Florida data were obtained from (Xiong and Hoyer 2019) and contain observations of 87 lakes in Florida, USA. Lake surface areas ranged from <1 to >7000 ha. Lake type information was not provided.

The New Zealand data were obtained by compiling records available online from Land Air Water Aotearoa at <https://www.lawa.org.nz> and from the Waikato Regional Government at <https://www.waikatoregion.govt.nz/assets/PageFiles/19300/2011-05.pdf>. These data cover 136 lakes, from 0.7 to 61500 ha in surface area, distributed over the ≈270000 km^2 area of New Zealand. Lakes in this dataset are predominantly natural, and other lake types were not sufficiently common for analysis of lake origin effects.

The Eastern Lake Survey (ELS) and Western Lake Survey (WLS) datasets were obtained from the US Environmental Protection Agency. The surveys, conducted in 1984 and 1985, respectively, were part of the National Acid Precipitation Assessment Program. The ELS features data on 1669 lakes in the northeast, southeast, and upper Midwest regions of the USA. The WLS features data on 752 lakes over the remaining contiguous US. The ELS contained some lake type attributes, but this was ignored in our analyses because of the large fraction of lakes with missing information and the predominance of natural lakes among those for which lake type was recorded. The WLS identified lakes by hydrologic classes: closed, drainage, reservoir, and seepage, which we used to analyze lake type effects. Although these hydrologic lake types are somewhat different from the lake types for other datasets, which focused on natural versus human origin and modification, the hydrologic lake types in the WLS allowed us to address another potential source of variation in catchment area to surface area relationships.

The NLA data were taken from the 2012 survey (U.S. Environmental Protection Agency 2016). These data represent 1230 lakes over the contiguous United States (≈8000000 km^2). The NLA includes lakes >1 ha in surface area, and lakes are selected for inclusion following a stratified random sampling procedure that ensures that lakes from major ecological zones are represented. The NLA distinguished five types of lakes: natural, enhanced natural, man-made, man-made abandoned, and reservoir. Enhanced natural lakes were originally open water shallow lakes that had been enhanced with flow diversions to form a larger/deeper lake. Man-made lakes were created by humans with current use for purposes of water storage, fishing, and recreation, but not irrigation or power generation; this category also includes lakes historically used for other purposes, such as mining pits and mill ponds. Man-made abandoned lakes have been abandoned from their intended function for many decades, and for our analysis were combined with the man-made class. Reservoirs are man-made impoundments used for drinking water, hydropower, flood control, and/or active irrigation. The NLA also grouped lakes by Level-1 North American Ecoregions. Some regions were represented by few lakes and were, for our analysis, merged into adjacent regions. We merged regions 11 (Mediterranean California), 12 (Southern Semi-Arid Highlands), and 13 (Temperate Sierras) into region 10 (North American

Deserts); region 15 (Tropical Wet Forests) intro region 8 (Eastern Temperate Forests); and region 7 (Marine West Coast Forest) into region 6 (Northwestern Forested Mountains).

Analyses

To analyze the scaling relationships between catchment area and lake surface area, we first took the base-10 logarithm of both quantities. The log transformation stabilizes variances and allows us to assess whether the relationship between lake surface and catchment area is a power-law. The power-law relationship $C = aL^b$, where C is catchment area and L is lake surface area, is linearized by taking the logarithm of both sides to $\log_{10}(C) = \log_{10}(a) + b\log_{10}(L)$. We used linear models to estimate $\log_{10}(a)$ and b, and thereby quantify the relationship between catchment area and lake surface area. Where possible, given sample sizes and information provided in the selected datasets, we included random effects to examine whether the scaling relationship between catchment area and lake surface area varies by lake type or region. Except when particular groups had too few members for robust analysis, we used lake type designations as provided in the original datasets; exceptions are described above under the *Datasets* sub-heading. We considered all models with random effects of lake origin or region on the intercept, the slope, or both, and used Akaike's Information Criterion with correction for small sample sizes (AICc) to select the best model. For the Iowa, Ohio, and WLS datasets, we considered random effects of lake type. For the NLA data, we considered random effects of region and lake origin.

To determine whether the relationship between catchment area and lake surface area for each dataset was well-described by a power law, we considered the linearity and homoscedasticity of each relationship (Zhao et al. 2019), and whether the estimate of b (the slope of the linear model, or the exponent on the natural scale) differs from 1. To determine whether the linearity criterion was met, we compared the fit of the linear model, or the selected linear mixed-effects model, to an equivalent model including a quadratic term, e.g. $\log_{10}(C) = \log_{10}(a) + b_1\log(L) + b_2\log(L)^2$. Linearity was rejected if a likelihood ratio test indicated the quadratic model to perform better at the $\alpha = 0.01$ significance level (Zhao et al. 2019). To determine whether the homoscedasticity criterion was met, we regressed absolute valued residuals from the power-law model against $\log_{10}(L)$. If the relationship between residuals and $\log_{10}(L)$ was statistically significant at the $\alpha = 0.01$ level, homoscedasticity was rejected. The $\alpha = 0.01$ significance level was used to evaluate consistency with a power-law relationship because, as our goal was to explore an approximate empirical relationship in data with error and potential sampling biases, we were willing to accept that relationships between lake surface area and catchment area approximately followed a power law unless there were considerable deviations from these assumptions. We used the 95% confidence interval on the estimate of the linear regression coefficient to determine if it differed from 1.

We analyzed all datasets separately because of differences between datasets in how lakes were selected for inclusion in each dataset, and in whether and how lake types were represented. There are also potential differences between datasets in methods for determining lake area and catchment area, the precise nature of which we were not able to be determined from available documentation. Additionally, we note that model-II regression would in some ways have been appropriate because it assumes that there is measurement error in both the predictor and response variables, but we are unaware of an established methodology for including random effects in these procedures. Furthermore, the independent variable, lake area, is likely measured with high precision and accuracy while the dependent variable, catchment area, typically has higher uncertainty. In this context the analysis is arguably closer to model I regression. Analyses were conducted in R version 3.5.1 (R Core Team 2018) using the 'lmerTest' package (Kuznetsova et al. 2017).

Results

For all nine datasets, there were strong relationships between lake area and catchment area (Figures 1 and 2). Most of the lake area to catchment area relationships were best fit by linear, as opposed to quadratic relationships, and most relationships had homoscedastic residuals. Among lakes meeting the linearity and homoscedasticity criteria, two datasets exhibited super-linear relationships ($b > 1$), one dataset had b not statistically different from 1, and three datasets exhibited sub-linear relationships ($b < 1$).

For the Adirondacks, there was a quadratic relationship between $\log_{10}$(lake surface area) and $\log_{10}$(catchment area) (Fig. 1a). The quadratic relationship gave a substantially better fit to the data than the linear relationship (ΔAICc = -25.5; likelihood ratio test $F = 27.7$, $df = 1$, $p < 0.0001$). Lakes in the ELS also exhibited a quadratic relationship between $\log_{10}$(lake surface area) and $\log_{10}$(catchment area) (Fig. 1b), but the relationship was less strongly nonlinear than for the Adirondacks (ΔAICc = -7.3; likelihood ratio test $F = 9.32$, $df = 1$, $p = 0.002$). Finnish lakes exhibited a linear relationship (Figure 1c) between $\log_{10}$(lake surface area) and $\log_{10}$(catchment area), but heteroscedasticity was detected (absolute valued residuals increase with $\log_{10}$(lake surface area); $b = 0.082$, $p < 0.0001$).

For datasets where relationships were linear and homoscedastic (Florida, Iowa, New Zealand, Ohio, WLS, NLA), most estimates of b were near one, but in some cases significantly greater than one, and in other cases, less than one. In addition, where data were available, accounting for different lake origins was important, as human-made water bodies had higher estimates of $\log_{10}(a)$ (y-intercepts on the log-log scale). In every dataset where the data allowed testing its effect (Iowa, Ohio, WLS, and NLA), the best model by AICc also included random effects of lake type on the linear regression intercept. For Florida lakes (Fig 1d), the estimate of *b* was indistinguishable from 1 ($b = 0.972 \pm \text{SE} = 0.065$). For Iowa lakes (Fig. 1e), human-made lakes—predominantly reservoirs—had larger catchments than natural lakes, but *b* was the same across lake origins (Fig. 1e). The relationship was super-linear, i.e., $b > 1$ ($b = 1.147 \pm \text{SE} = 0.069$). New Zealand lakes had a sub-linear ($b = 0.860 \pm \text{SE} = 0.050$), power-law relationship between catchment area and lake surface area (Fig. 1f). For the Ohio lakes, reservoirs had larger catchments than other lakes, including natural lakes, and the slope was the same across lake origins (Fig. 1g). The slope estimate was less than 1 ($b = 0.773 \pm \text{SE} = 0.094$). The WLS data showed a power-law relationship with effects of lake origin on catchment area, but *b* did not differ by lake origin (Fig. 1h). Reservoirs had the largest catchments, and seepage lakes the smallest. The scaling relationship was sublinear ($b = 0.802 \pm \text{SE} = 0.032$). For the NLA data, reservoirs and other man-made lakes had the largest watersheds, but *b* did not differ among lake origins. The scaling relationship was super-linear ($b = 1.050 \pm \text{SE} = 0.024$). The model accounting for lake type alone outperformed models allowing variation among ecoregions in the scaling relationship.

Discussion

There is generally a linear scaling relationship between lake surface area and catchment area. Lake type can modify this relationship such that a reservoir or impoundment of a given size will typically have a larger catchment than a natural lake of comparable surface area, but the rate at which catchment area increases as a function of lake surface area does not appear to change according to lake origin, at least within a given data set. Although we observed some variation between datasets in the shape of this scaling relationship and its parameters, differences among lake types appeared to be more important than differences among regions in the scaling between lake surface area and catchment area. Knowledge of these relationships is useful for scaling the total magnitudes of fluxes of matter and energy between landscapes and lakes, for example to extrapolate loading rates known for one system to another, differently-sized system.

In the context of assessing power-law relationships, identifying whether the relationship is sub- or super-linear, i.e., b is less than or greater than 1, is key. In the absence of evidence, linearity is often the default assumption, but severe mispredictions can occur if the increase in catchment area as lake surface area increases occurs more slowly ($b < 1$) or more rapidly ($b > 1$) than linearly. For example, if we take the y-intercept $\log(a) = 1.2$ (a typical value for our study) and estimate the catchment area of a 100 ha lake with $b = 0.8$ versus $b = 1.2$, the estimated catchment area ranges 630 to 3980 ha, a difference of >3000 ha. Although some individual datasets gave estimates of b that were statistically distinguishable from 1, all were qualitatively close to 1, so even if the true relationship is subtly nonlinear, other sources of variation and error likely have a greater effect on predictions based on this relationship than a slight misspecification of the form of the relationship.

Estimates of the $\log(a)$, the y-intercept on the linearized scale, when converted back to the natural scale, correspond to the average ratio of catchment area to lake surface area. Estimates of $\log(a)$ were typically between 1 and 2, meaning that catchment areas were, on average, 10 to 100 times greater than lake surface area. The catchments of natural lakes were generally smaller, relative to lake surface area, than artificial lakes, which had much larger catchments. This is consistent with analysis of a U.S. and global dataset and a subsetted NLA database, where catchment areas of artificial lakes were 3 to 4 times greater than the catchment areas of natural lakes (Thornton et al. 1980, Harrison et al. 2009, Hayes et al. 2017).

A utility of the relationships is illustrated by a simple example. An approximate average export of dissolve organic carbon (DOC) from a forested catchment is 40 kg C ha^{-1} (Aikenhead et al. 2000; Canham et al. 2004). Using the relationships for Finland, New Zealand, and WLS, we can estimate of catchment areas for a 100 ha lake. Using the DOC export, we can then calculate areal loading rates as ~ 1000, ~ 700, and ~ 450 kg C ha^{-1} for Finland, New Zealand, and WLS, respectively. This calculation suggests a forested Finnish lake likely has a substantially higher DOC loading rate than a similar New Zealand lake and that a lake in the Western U.S receives less than half the DOC input relative to a similar Finnish lake. Similar to these calculations, we could estimate a range of DOC loading for lakes of the same size using the prediction intervals of the regressions. While such calculations must be used cautiously—for example, they do not account for differences in processing of material from the catchment *en route* to the lake—they can provide the basis for first order estimates and hypotheses for further investigation that is particularly useful in expanding limnological investigations to regional and global scales. Because lake surface areas are commonly measured, but catchment areas and, especially, export and loading rates, are more difficult to quantify, calculations such as this could be summed over many lakes to estimate regional and global-scale values.

Only two datasets, Adirondacks and ELS, better fit a quadratic than linear regression model for lake surface area and catchment area on the log-log scale (Fig. 1). It is possible that the $\log_{10}$(catchment area) to $\log_{10}$(lake surface area) relationship is truly nonlinear in some settings. That this pattern was shared between both eastern USA datasets could suggest an unknown geomorphic explanation. However, we suspect that attributes of these datasets mask what is truly a linear relationship with lake type effects, as was seen in the Iowa, Ohio, WLS, and NLA datasets (Figs. 1, 2). Because different lake origins have different size distributions, and different lake types also have different y-intercepts (on the log-log scale), overrepresentation of particular lake types at different ranges of lake size could make the overall shape of the relationship appear nonlinear. While the Adirondacks data did not contain lake type information, the study does include water bodies with “reservoir” in the name. We suspect that some of the larger catchments relative to lake area (positive residuals in Figure 1) may be reservoirs but we did not have the data to test this conjecture. This same argument may explain the heteroscedasticity in the Finland relationship.

In summary, catchment area scales to lake area typically in linear form on a log-log scale, with slopes near one. Although we found evidence of some regional variability in the catchment area to lake area scaling relationship, within a region, catchment area-lake area slopes are consistent but intercepts differ in relation to lake type, with reservoirs and other man-made water bodies tending to have larger drainage ratios. Given that the geographic scope and stratified random sampling procedure used to select lakes for inclusion in the dataset ensured wide representation of lake characteristics and geomorphic settings, the relationships for the NLA dataset likely generalize best to lakes outside the specific regions investigated here. Catchment area-lake area relationships provide a means to calculate inputs and impacts of catchment processes on lakes and how these vary.

Acknowledgements

This work was supported by NSF Grants OAC-1839024 and OAC-1839011. JAW was also supported by a NatureNet Science Fellowship.

Data availability statement

A compilation of the ELS, Iowa, New Zealand, Ohio, and NLA data used in this study is publicly archived at https://doi.org/10.6073/pasta/3a45bd3bb328fb202ff24b7d54b83cba. The Adirondacks, Finland, and Florida datasets were used with permission from the original authors of the datasets, but without permission to distribute them. Contacts for those datasets should be to the authors of the original papers cited in the text.

References

Baker, J. P., S. A. Gherini, S. W. Christensen, C. T. Driscoll, J. Gallagher, R. K. Munson, R. K. Newton, K. H. Reckhow, and C. L. Shofield. 1990. Adirondack lakes survey: An interpretive analysis of fish communities and water chemistry, 1984-87. Ray Brook, N.Y.

Bohk, C., R. Rau, and J. E. Cohen. 2015. Taylor's power law in human mortality. Demographic Research 33:589–610.

Doak, D. F., D. Bigger, E. K. Harding, M. A. Marvier, R. E. O'Malley, and D. Thomson. 1998. The statistical inevitability of stability-diversity relationships in community ecology. The American Naturalist 151:264–276.

Downing, J. A. 1979. Aggregation, transformation and the design of benthos sampling programs. Canadian Journal of Fisheries and Aquatic Sciences 36:1454–1463.

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 Biogeochemical Cycles 22:1–10.

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, J. M. Melack, and J. J. Middelburg. 2006. The global abundance and size distribution of lakes, ponds, and impoundments. Limnology and Oceanography 51:2388–2397.

Forsius, M., V. Malin, I. Makinen, J. Kamari, P. Kortelainen, and M. Verta. 1990. Finnish lake acidification survey: survey design and random selection of lakes. Envirometrics 1:73–88.

Hallett, L. M., J. S. Hsu, E. E. Cleland, S. L. Collins, T. L. Dickson, E. C. Farrer, L. A. Gherardi, K. L. Gross, R. J. Hobbs, L. Turnbull, and K. N. Suding. 2014. Biotic mechanisms of community stability shift along a precipitation gradient. Ecology 95:1693–1700.

Harrison, J. A., R. J. Maranger, R. B. Alexander, A. E. Giblin, P. A. Jacinthe, E. Mayorga, S. P. Seitzinger, D. J. Sobota, and W. M. Wollheim. 2009. The regional and global significance of nitrogen removal in lakes and reservoirs. Biogeochemistry 93:143–157.

Hayes, N. M., J. R. Corman, B. R. Deemer, K. E. Strock, and N. R. Razavi. 2017. Key differences between lakes and reservoirs modify climate signals: A case for a new conceptual model. Limnology and Oceanography Letters 2:47–62.

Hayes, N. M., and M. J. Vanni. 2018. Microcystin concentrations can be predicted with phytoplankton biomass and watershed morphology. Inland Waters 8:273–283.

Kortelainen, P. 1993. Total organic carbon in Finnish lakes and its relationship to catchment characteristics. Canadian Journal of Fisheries and Aquatic Sciences 50:1477–1483.

Kretser, W. A., J. Gallagher, and J. Nicholette. 1989. Adirondack lakes study 1984-87: an evaluation of fish communities and water chemistry. Ray Brook, N.Y., USA.

Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2017. lmerTest package: tests in linear mixed effects models. Journal of Statistical Software 82:1–26.

Lagrue, C., R. Poulin, and J. E. Cohen. 2015. Parasitism alters three power laws of scaling in a metazoan community: Taylor’s law, density-mass allometry, and variance-mass allometry. Proceedings of the National Academy of the United States of America 112.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Rasmussen, J. B., L. Godbout, and M. Schallenberg. 1989. The humic content of lake water and watershed and lake morphometry. Limnology and Oceanography 34:1336–1343.

Reuman, D. C., L. Zhao, L. W. Sheppard, P. C. Reid, and J. E. Cohen. 2017. Synchrony affects Taylor’s law in theory and data. Proceeding of the National Academy of Sciences

114:6788–6793.

Robert, A., and A. G. Roy. 1990. On the fractal interpretation of the mainstream length-drainage area relationship. Water Resources Research 26:839–842.

Seekell, D. A., and M. L. Pace. 2011. Does the Pareto distribution adequately describe the size-distribution of lakes? Limnology and Oceanography 56:350–356.

Shreve, R. L. 1974. Variation of mainstream length with basin area in river networks. Water Resources Research 10:1167–1177.

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. Limnology and Oceanography 52:1208–1219.

Taylor, L. R. 1961. Aggregation, variance and the mean. Nature 189:732–735.

Taylor, R. A. J. 2018. Spatial distribution, sampling efficiency and Taylor's power law. Ecological Entomology 43:215–225.

Thornton, K. W., R. H. Kennedy, J. H. Carroll, W. W. Walker, R. C. Gunkel, and S. Ashby. 1980. Reservoir sedimentation and water quality-an heuristic model. Pages 654–661 *in* H. G. Stefan, editor. Proceedings of the symposium on surface water impoundments. American Society of Civil Engineering.

Tippett, M. K., and J. E. Cohen. 2016. Tornado outbreak variability follows Taylor's power law of fluctuation scaling and increases dramatically with severity. Nature Communications 7:1–7.

U.S. Environmental Protection Agency. 2016. National Aquatic Resource Surveys. https://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys.

Xenopoulos, M. A., D. M. Lodge, J. Frentress, T. A. Kreps, S. D. Bridgham, E. Grossman, and C. J. Jackson. 2013. Regional comarparisons of watershed determinants of dissolved organic carbon in temperate lakes from the Upper Great Lakes region and selected regions globally. Limnology and Oceanography 48:2321–2334.

Xiong, C., and M. V. Hoyer. 2019. Influence of land use and rainfall variability on nutrient concentrations in Florida Lakes. Lake and Reservoir Management 35:25–37.

Xu, M., W. S. F. Schuster, and J. E. Cohen. 2014. Robustness of Taylor's law under spatial hierarchical groupings of forest tree samples. Population Ecology 57:93–103.

Zhao, L., L. W. Sheppard, P. C. Reid, J. A. Walter, and D. C. Reuman. 2019. Proximate determinants of Taylor's law slopes. Journal of Animal Ecology 88:484–494.

Tables

Table 1: Datasets used in this analysis and associated lake origin designations. Lake origins correspond to random effects considered in our analysis.

Dataset	Lake type designations
Adirondacks	None
Eastern Lake Survey	None
Finland	None
Florida	None
Iowa	**Human-made:** borrow pits, impoundments **Natural:** naturally-occurring lakes
New Zealand	None
Ohio	**Other:** natural lakes and permanently-filled quarries **Reservoir:** tributary and canal reservoirs
Western Lake Survey	**Closed:** lakes having no outlet **Drainage:** lakes having an inlet and an outlet **Reservoir:** human-made impoundments used for drinking water, hydropower, and/or active irrigation **Seepage:** lakes having no inlet or outlet
National Lakes Assessment (2012)	**Man-made:** created by humans, includes abandoned man-made lakes, but not reservoirs **Natural:** naturally occurring, unmodified lakes **Enhanced natural:** naturally occurring, but enhanced with flow diversions e.g. to make a larger or deeper lake. **Reservoir:** human-made impoundments used for drinking water, hydropower, and/or active irrigation

Figures:

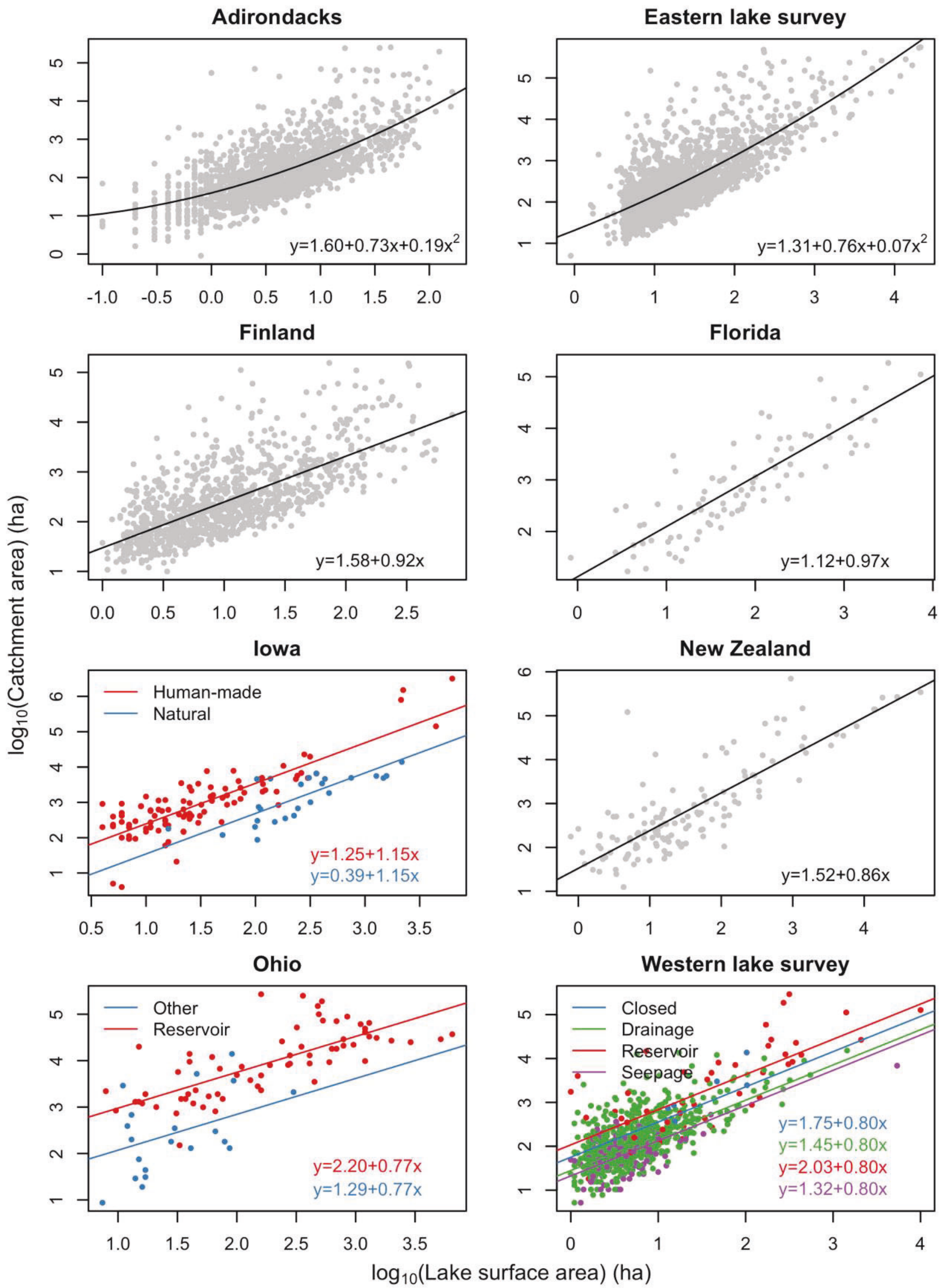

Fig. 1: Relationships between lake surface area and catchment area for 8 selected datasets.

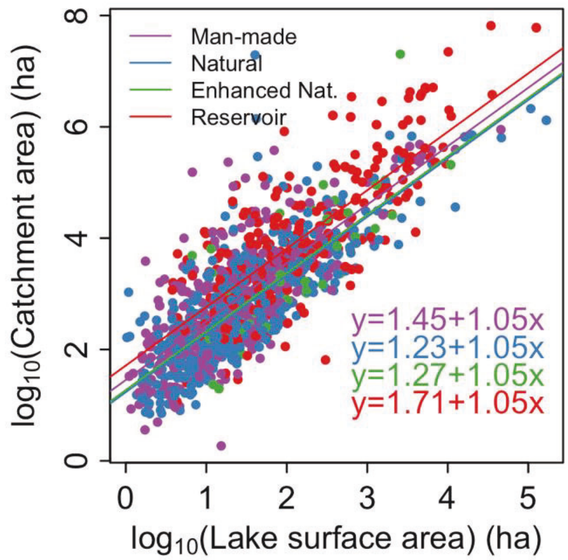

Fig. 2: Relationship between lake surface area and catchment area for lakes sampled in the 2012 National Lakes Assessment survey.