

Drivers and spatial structure of abiotic and biotic properties of lakes, wetlands, and streams at the national scale

KATELYN KING ^{1,4} KENDRA SPENCE CHERUVELIL,^{1,2} AND AMINA POLLARD³

¹Department of Fisheries and Wildlife, Michigan State University, East Lansing, Michigan 48824 USA

²Lyman Briggs College, Michigan State University, East Lansing, Michigan 48824 USA

³U.S. Environmental Protection Agency Office of Water, Washington, D.C. 20004 USA

Citation: King K., K. S. Cheruvilil, and A. Pollard. 2019. Drivers and spatial structure of abiotic and biotic properties of lakes, wetlands, and streams at the national scale. *Ecological Applications* 00(00):e01957. 10.1002/eap.1957

Abstract. Broad-scale studies have improved our ability to make predictions about how freshwater biotic and abiotic properties will respond to changes in climate and land use intensification. Further, fine-scaled studies of lakes, wetlands, or streams have documented the important role of hydrologic connections for understanding many freshwater biotic and abiotic processes. However, lakes, wetlands, and streams are typically studied in isolation of one another at both fine and broad scales. Therefore, it is not known whether these three freshwater types (lakes, wetlands, and streams) respond similarly to ecosystem and watershed drivers nor how they may respond to future global stresses. In this study, we asked, do lake, wetland, and stream biotic and abiotic properties respond to similar ecosystem and watershed drivers and have similar spatial structure at the national scale? We answered this question with three U.S. conterminous data sets of freshwater ecosystems. We used random forest (RF) analysis to quantify the multi-scaled drivers related to variation in nutrients and biota in lakes, wetlands, and streams simultaneously; we used semivariogram analysis to quantify the spatial structure of biotic and abiotic properties and to infer possible mechanisms controlling the ecosystem properties of these freshwater types. We found that abiotic properties responded to similar drivers, had large ranges of spatial autocorrelation, and exhibited multi-scale spatial structure, regardless of freshwater type. However, the dominant drivers of variation in biotic properties depended on freshwater type and had smaller ranges of spatial autocorrelation. Our study is the first to document that drivers and spatial structure differ more between biotic and abiotic variables than across freshwater types, suggesting that some properties of freshwater ecosystems may respond similarly to future global changes.

Key words: aquatic vegetation; chlorophyll a; lakes; macroscale; National Aquatic Resource Surveys; nutrients; spatial autocorrelation; spatial scale; streams; wetlands.

INTRODUCTION

Ecologists have long understood the complex nature of ecosystems and the importance of spatial scale for understanding both ecosystem pattern and process (e.g., Allen and Starr 1982, Levin 1992, Peters et al. 2007). In fact, as the discipline has aged and technology has advanced, more ecological studies explicitly examine the role of scale for their species or ecosystem property of interest and take into account complex relationships among abiota and biota, as well as across ecosystems. The fields of landscape ecology, macroecology, and macrosystems ecology have further advanced the importance of studying phenomena at broad scales (i.e., regions to continents; Heffernan et al. 2014, McGill

2019), quantifying spatial structure of biotic and abiotic properties and their drivers (e.g., Baskett and Jordan 1995, Perry et al. 2002, Lapierre et al. 2018), considering the multi-scale drivers of biota and abiota (e.g., Whitaker et al. 2001, Stendera et al. 2012, Soranno et al. 2014), and recognizing the connections among and within terrestrial and aquatic ecosystems of various types (e.g., Peters et al. 2008, Rotjan and Idjadi 2013, Hotchkiss et al. 2018). However, it remains rare to conduct research at the macroscale that includes multiple ecosystem types and examines both biotic and abiotic properties.

Freshwater ecosystems present an ideal opportunity to do such macroscale, inter-ecosystem, and integrative research. Freshwaters collect and process material from the surrounding watershed, have well-defined boundaries, represent a widely distributed network, and are hierarchically organized (e.g., Frissell et al. 1986, Williamson et al. 2009). In addition, freshwater nutrients and biota reflect changes in their surrounding watershed

Manuscript received 28 September 2018; revised 21 May 2019; accepted 11 June 2019. Corresponding Editor: Ryan S. King.

⁴E-mail: katelyn8king@gmail.com

(e.g., Allan et al. 1997, Williamson et al. 2009), such as those caused by climate change and land use intensification, both of which operate at multiple spatial scales. To date, most freshwater research at the macroscale has focused on individual freshwater types (i.e., lakes, streams, or wetlands; Kling et al. 2000, Soranno et al. 2010, Stanley and del Giorgio 2018). However, when comparable properties of different freshwater types have been studied simultaneously at macro scales, surprising similarities and differences have emerged. For example, algal abundance per unit phosphorus across the United States was greater in lakes than rivers (Soballe and Kimmel 1987), likely due to differences in water residence time (the average duration water spends in the ecosystem). In contrast, macroinvertebrate communities in streams and lakes across Sweden were influenced more by ecosystem habitat properties (e.g., substrate, vegetation, and water chemistry) than by regional properties (e.g., elevation, latitude, and longitude; Johnson et al. 2004).

Interesting findings have also resulted from macro-scale studies of abiotic properties across freshwater types. For example, although lakes generally had lower nutrient concentrations than wetlands, water chemistry was positively related to modified land use regardless of freshwater type in the southern region of New Zealand (Galbraith and Burns 2007). In addition, lakes and streams exhibited similar dissolved organic carbon concentrations, likely due to inputs from wetlands in the upper midwestern United States (Lottig et al. 2011) and regional land use was a shared driver of stream and lake water chemistry in Sweden (Stendera and Johnson 2006). These few inter-ecosystem studies demonstrated important similarities and differences among freshwater types, facilitated the establishment of general frameworks that may cross ecosystem boundaries, and resulted in calls for research that includes a diversity of ecosystem properties and various ecosystem types (Chaloner and Wotton 2011, Rotjan and Idjadi 2013, Nöges et al. 2016).

Multi-scale drivers, predictor variables that operate and can be quantified at multiple scales (e.g., Tonn 1990, Allan et al. 1997, Poff 1997, Read et al. 2015), likely account for some of the macroscale study results of either biota or abiota across multiple freshwater types. Understanding how the spatial structure of these ecosystem properties compares to that of their drivers can provide insight about how spatial structure influences relationships among ecological properties (Lapierre et al. 2018) and across freshwater types. For example, a combination of broad-scale patterns in hydrology and land cover, as well as lake and catchment morphometry, generated an intermediate-scale spatial structure in lake total phosphorus across the northeastern United States (Lapierre et al. 2018). In streams and lakes, benthic macroinvertebrates exhibited a patchy spatial structure across several broad spatial extents (>1,000 km) in the United States and Sweden, likely

due to their small size and limited dispersal ability (Shurin et al. 2009). In contrast, the spatial structure of stream water chemistry across Maryland, USA was strongly related to broad-scale landscape variables, such as agriculture and geology type (Peterson et al. 2006). However, it is unknown to what extent these macroscale patterns are consistent across lakes, wetlands, and streams. Because streams have extensive networks and longitudinal flow, they may exhibit a broader range of spatial structure compared to lakes and wetlands (Shurin et al. 2009). Ecologists also know little about the multi-scale drivers or spatial structure of abiotic and biotic properties of all freshwaters (i.e., lakes, wetlands, and streams combined). Therefore, macroscale freshwater research that integrates across lakes, wetlands, and streams and studies both biotic and abiotic properties has the potential to greatly contribute to ecological understanding and prediction in the face of global threats (Chaloner and Wotton 2011, Rose et al. 2017, Stanley and del Giorgio 2018).

For the first time, we integrate data about all three freshwater types to ask do lake, wetland, and stream biotic and abiotic properties respond to similar ecosystem and watershed drivers and have similar spatial structure at the national scale? We studied thousands of water bodies at the scale of the conterminous United States using the U.S. Environmental Protection Agency's National Aquatic Research Surveys (NARS; *available online*).⁵ Owing to similar biogeochemical processing across freshwaters, we expected abiotic properties to respond to similar drivers at the macroscale, regardless of freshwater type. However, we expected biotic variables to respond to different external drivers across freshwater types because of intra-ecosystem species and community differences. We also expected the spatial structure of these abiotic and biotic variables to reflect the spatial structure of dominant drivers, which we predicted would vary by freshwater type due to different water residence times (streams < wetlands < lakes, generally; Kalff 2002). By studying both abiotic and biotic properties across lakes, wetlands, and streams at the macroscale, our results can increase understanding of broad-scale responses to changes in climate and land use intensification, as well as inform future national ecosystem assessments and landscape-scale management and conservation efforts.

METHODS

Study sites

We used data from the U.S. Environmental Protection Agency's National Aquatic Research Surveys that include sample sites in lakes, streams, and wetlands: the 2012 National Lakes Assessment (NLA), the 2008–2009 National Rivers and Streams Assessment (NRSA), and

⁵ www.epa.gov/national-aquatic-resource-surveys

the 2011 National Wetland Condition Assessment (NWCA; Fig. 1a). Field crews sampled each freshwater type during a summer index period, typically May–September, during respective NARS years. For each freshwater type, the EPA selected sample locations using a statistically representative sample design that was stratified by EPA ecoregion (Fig. 1b), as well as either lake area, stream Strahler order (i.e., the position of a stream in the river network), or wetland vegetation category (e.g., intertidal emergent, forested palustrine; USEPA 2016a,b,c). Each NARS also a priori hand-selected “least disturbed” reference sites using screening

criteria that varied by state or ecoregion, which were also included in our study.

NARS sample lakes were chosen from a target population that included freshwater lakes, ponds, and reservoirs (hereafter referred to as lakes) ≥ 1 ha, at least 1 m deep, with a minimum 0.1 ha of open water, and a minimum retention time of one week ($n = 1,130$; USEPA 2016a). Sample streams were chosen from perennial streams and rivers, hereafter referred to as streams ($n = 2,123$; USEPA 2016b). Sample wetlands were chosen from both marine and freshwater wetlands that contained rooted vegetation; if open water was present, 90%

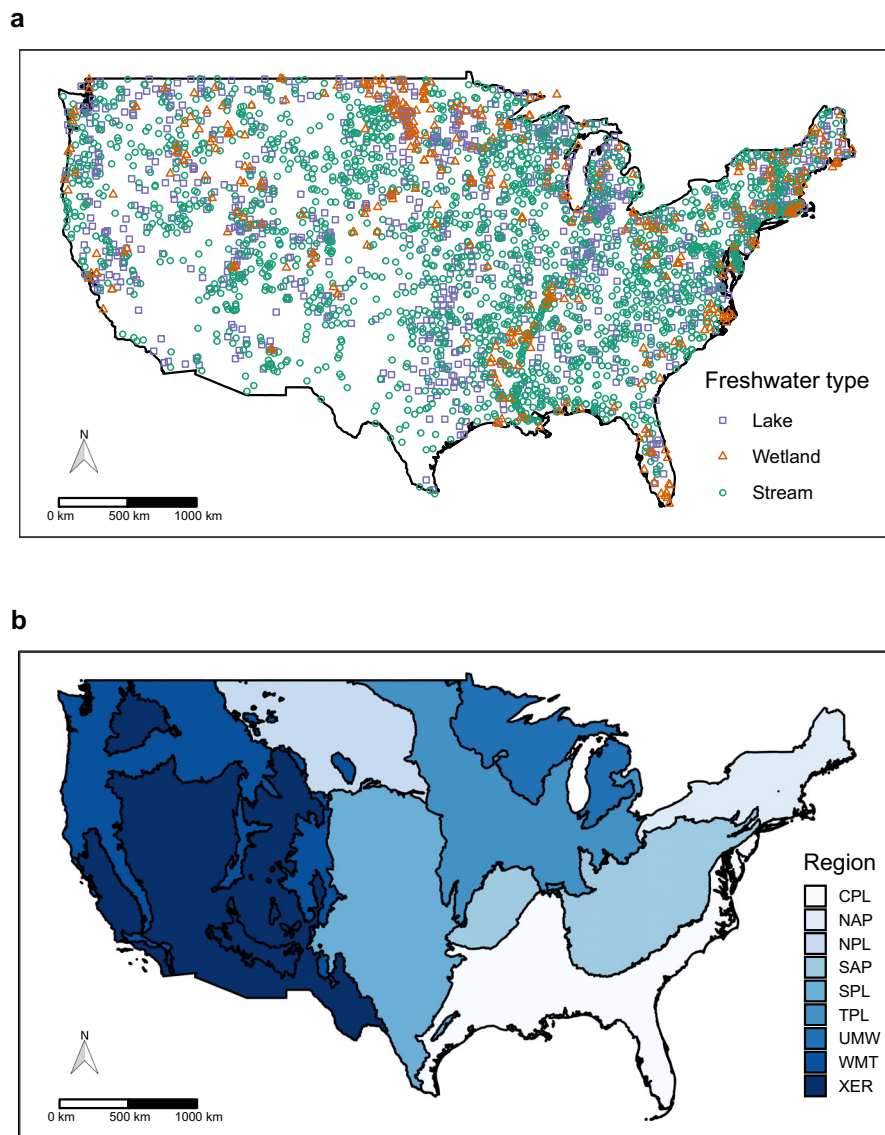


FIG. 1. (a) Freshwater sample sites for the (a) National Lakes Assessment in blue ($n = 1,130$), the National Wetland Condition Assessment in orange ($n = 400$), and National Rivers and Streams Assessment in green ($n = 2,123$). (b) U.S. EPA ecoregions (agglomerated Omernik III regions; Omernik 1987): CPL, Coastal Plains; NAP, Northern Appalachians; NPL, Northern Plains; SAP, Southern Appalachians; SPL, Southern Plains; TPL, Temperate Plains; UMW, Upper Midwest; WMT, Western Mountains; and XER, Xeric.

TABLE 1. Descriptive statistics for the lake, stream, and wetland response variables included in analyses.

Response variable (units)	Lakes		Streams		Wetlands	
	<i>n</i>	Mean \pm SD	<i>n</i>	Mean \pm SD	<i>n</i>	Mean \pm SD
Total phosphorus ($\mu\text{g/L}$)	1,130	116.5 \pm 278.1	2,116	156.1 \pm 543.2	400	455.3 \pm 1004.6
Total nitrogen ($\mu\text{g/L}$)	1,130	1141.8 \pm 2579.6	2,116	1216.7 \pm 2392	390	2078.6 \pm 4748.7
Chlorophyll <i>a</i> ($\mu\text{g/L}$)	1,034	27.6 \pm 57.3	2,010	12.3 \pm 62.9	393	35.6 \pm 136.5
Aquatic vegetation (% cover)	1,115	28.7 \pm 26.7	2,113	7.7 \pm 15.0	400	29.3 \pm 33.3

of wetland area was <1 m deep (USEPA 2016c). We used only freshwater wetlands ($n = 400$) in this study to compare ecological properties with lakes and streams. Most response variables and driver variables were acquired from each NARS raw data file (USEPA 2016d,e,f). However, lake watershed land use metrics and stream chlorophyll *a* concentration were obtained directly from EPA personnel. The resulting database covers a broad geographical extent that includes nine ecoregions across the conterminous U.S. (Fig. 1b), three freshwater types, and two abiotic and two biotic properties. All data and code used for analyses are available in an online repository (King 2018, 2019).

Abiotic and biotic freshwater properties

Detailed sampling methods for ecosystem properties, total phosphorus, total nitrogen, and chlorophyll *a* concentrations, as well as percent aquatic vegetation cover (all measures of abundance) can be found in EPA field and lab manuals (e.g., USEPA 2007, 2008, 2011a,b,c, 2012). Briefly, samples for total phosphorus (TP), total nitrogen (TN), and chlorophyll *a* (CHL) were collected using an integrated collection tube from the top 2 m of the water column at an open water location in lakes, grab samples from mid-channel at a depth of 0.5 m in streams, and grab samples from open water >15 cm in wetlands. Samples were then packed on ice and immediately sent for lab analysis to measure nutrient and algal concentrations. Field crews recorded observations of percent cover of aquatic vegetation for each sample site, including emergent, submersed, and floating vegetation. Lake aquatic vegetation cover was estimated based on plants collected with a double-sided rake, stream plant cover estimates were recorded by visual inspection or by sounding with a pole, and wetland plant cover estimates were determined by visual inspection (USEPA 2007, 2011a,b). Descriptions of the ecosystem properties (i.e., response variables in analyses) are in Table 1, and spatial distributions of these variables are in Fig. 2.

Drivers of freshwater properties

The EPA characterized ecosystem and watershed contexts of each waterbody using watersheds from the NHDPlus v2 for lakes and streams and used concentric buffers around the center point of each wetland (200, 500, 1,000 m; USEPA 2016d,e,f). We used 1,000-m

wetland buffers as a proxy for wetland watersheds because they were most similar in area to the lake watersheds (median area = 3.14 and 10.6 km², respectively). Soranno et al. (2015) found that when comparing lake watersheds to 1,500 m buffers around lakes, they were equally effective in quantifying the effect of land use/cover on lake nutrients. Previous studies at the national scale have similarly used wetland buffers to characterize the landscape (e.g., Moon et al. 2017, Stapanian et al. 2018). The EPA quantified a variety of driver variables at the watershed (and wetland buffer) scale: NLCD 2006 land use/cover, NADP 2007–2011 nitrogen deposition, 2000/2010 U.S. Census human population density, 2010 TIGER road density, and 2006 NED elevation (Table 2) (USEPA 2016d,e,f). The EPA field crews estimated riparian zone vegetation cover and we aggregated the three vegetation layers (canopy, mid-layer, ground cover) and types (herbaceous, woody, trees) (USEPA 2007, 2011a,b). Descriptions of predictor variables can be found in Table 2.

Because climate is an important driver of freshwater nutrients and biota at broad scales (Allan et al. 2005, Wrona et al. 2006, O'Reilly et al. 2015), we added climate variables to our compiled NARS data set. We extracted 800-m resolution 30-yr normal values for mean annual precipitation, mean maximum temperature, and mean minimum temperature for the point location of each sample site (1981–2010; data *available online*).⁶ Winter precipitation was calculated from monthly averages of 4-km resolution data for the December–February prior to the sample period for each NARS (i.e., streams sampled during 2008 have winter precipitation data from Dec 2007–Feb 2008). Summer precipitation and temperatures were calculated using monthly averages of the 5-month sampling period of May–September during the same year of each NARS.

Analysis

We used random forest (RF) analysis to determine the important drivers of nutrients and biota in lakes, wetlands, and streams simultaneously (i.e., in a single combined model). RF is a machine-learning approach that grows a regression tree for each bootstrap sample of the original data set using a random subset of predictor

⁶<http://prism.oregonstate.edu>

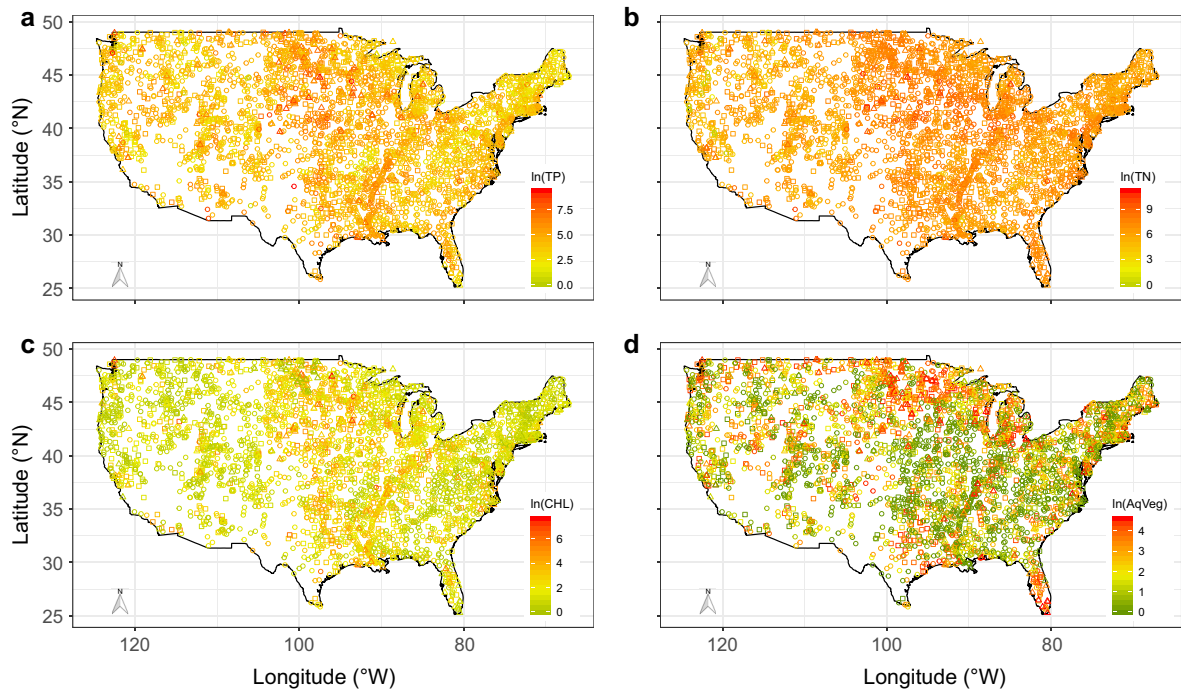


FIG. 2. Maps of the spatial distribution of the natural log (ln) of ecosystem properties, (a) total phosphorus (TP) (b) total nitrogen (TN), (c) chlorophyll *a* (CHL), and (d) percent aquatic vegetation cover (AqVeg) across lakes (squares), wetlands (triangles), and streams (circles).

TABLE 2. Descriptive statistics for predictor variables calculated at either the ecosystem or watershed scale.

Predictor variable (units)	Abbreviation	Median	5th	95th
Ecosystem scale				
Depth (m) [†]	DEPTH	1.0	0.13	14.6
Riparian vegetation (%)	Rveg	30.0	9.1	69.6
Mean annual precipitation, 30-yr norm (mm)	PrecipNorm	965.5	290.7	1511.0
Summer precipitation average May–Sep of sample period (mm)	PrecipSummer	83.1	13.0	164.1
Winter precipitation average Dec–Feb of sample period (mm)	PrecipWinter	60.1	8.2	151.6
Minimum annual temperature, 30-yr norm (°C)	TMIN	3.9	−2.6	12.9
Maximum annual temperature, 30-yr norm (°C)	TMAX	16.2	10.0	25.6
Summer temperature average May–Sep of sample period (°C)	Tsummer	19.3	12.0	26.6
Watershed or 1,000-m buffer scale				
Mean elevation (m)	ELEVMEAN	408.5	30.6	2504.1
Forest (%)	FOREST_PCT	33.9	0	88.3
Agriculture (%)	AG_PCT	8.5	0	78.3
Wetland (%)	WETLAND_PCT	1.4	0	31.7
Urban (%)	URBAN_PCT	3.8	0	26.1
Grassland (%)	SHRUB_GRASS_PCT	9.0	0	84.7
Mean nitrogen deposition (kg/ha)	NDEP	3.5	0.9	5.9
Population density (people/km ²)	POPDEN	12.7	0	702.0
Road density (km/km ²)	ROADDEN	1.3	0.1	4.0

Notes: For wetlands, 1,000-m buffers were used as a proxy for unavailable watersheds. Statistics shown are calculated across all three freshwater types. Predictor variables without statistics are freshwater type (lake, wetland, stream) and ecoregion (AGGR_ECO9_2015). 5th and 95th refer to percentiles.

[†] Maximum (lakes and wetlands) or mean thalweg (deepest part of the channel along reach) (stream) depth.

variables for each split that partitions the observations (Breiman 2001). Each tree model then predicts observations not used in the model development (out-of-bag samples) using both original and randomly permuted

data, resulting in mean square error estimates. Predictor importance can be determined by the mean decrease in accuracy (also known as “permutation accuracy importance measure”) whereby variable importance is the

difference in prediction accuracy before and after permutation (Strobl et al. 2007). Ecologists increasingly use RFs to determine which predictor variables are important for classifying ecosystem properties (e.g., Cutler et al. 2007, Hollister et al. 2016).

We ran 500 trees for each natural log-transformed response variable and investigated the unscaled mean decrease in accuracy measure using the randomForest package (Liaw and Wiener 2002) in R 3.5.0 (R Core Team 2018). When we plotted the change in error vs. the number of trees, the error flattened at about 100 trees (Appendix S1: Fig. S1); however, we selected 500 trees to ensure unbiased out-of-bag estimates (Breiman 2001). For each response variable, we included a mix of ecosystem and watershed context variables (i.e., driver variables at different spatial scales), freshwater type, and ecoregion membership as potential predictor variables (Table 2). We identified a different number of important predictor variables for each response variable and have bolded them in Fig. 3. We identified the top predictors by repeating the analysis with several random seeds, as the top-ranked variables should remain unchanged from each run (Strobl et al. 2007), and using inflection points in variable importance plots. To determine whether our results were influenced by the relatively few wetlands sampled by the EPA, we ran RF analysis on a subset of data that included just 400 each of randomly selected lakes and streams and compared the results to those found when using the whole data set (Appendix S1: Fig. S2).

For each biotic and abiotic response variable, we determined the broad-scale spatial structure for lakes, wetlands, and streams individually. We then compared the spatial structure of the response variables to the spatial structure of the top drivers from the RF analysis to determine how well abiotic and biotic variables reflect the scale of the dominant drivers. We used semivariogram analysis to quantify spatial structure because it provides a dissimilarity measure where $\gamma(\mathbf{h})$ is the variance between point pairs for a particular Euclidean lag distance \mathbf{h} (Legendre and Fortin 1989). Euclidean relationships reveal lateral connectivity between a waterbody and the landscape, indicate landscape influences, and show spatial dependence at broad scales (McGuire et al. 2014). To determine broad-scale spatial structure with semivariograms, we used a bin size of 20 km (relatively small scale) for each variable that maintained at least 50 point-pairs per bin (Rossi et al. 1992, Turner et al. 2001).

We report the range (Appendix S2), which is the distance at which spatial autocorrelation ends and the measure of variance asymptotes, as well as semivariogram shape as either single-scale or nested structures (one or multiple ranges, respectively; McGuire et al. 2014). Although standard practice for semivariance analysis is to calculate semivariance up to one-half of the largest Euclidean distance between point pairs (Rossi et al. 1992), we report semivariogram results up to 2,500 km

(0.6 the maximum distance) because this distance was useful for visualization purposes and the resulting statistics were the same for both distances. Spherical or exponential models were fit to the semivariograms to estimate ranges of spatial autocorrelation using the gstat package (Pebesma 2004) and visual adjustment. For nested semivariograms, we determined the second range of spatial autocorrelation visually, as the goal of our analysis was not for prediction but to compare semivariogram shapes, which does not require fitting a model (sensu McGuire et al. 2014).

RESULTS

Drivers of abiotic and biotic properties across lakes, wetlands, and streams

Random forest models, one for each ecosystem property across all three freshwater types, show that freshwater type (lake, wetland, stream) was not one of the most important predictor variables for the two abiotic properties (TP and TN). These properties were best explained by watershed percent forest cover, followed by watershed agriculture and ecoregion membership (Fig. 3a, b). Watershed mean elevation was also a top driver of TN (Fig. 3b). In contrast, freshwater type was the most important predictor variable of the two biotic properties (CHL and aquatic vegetation; Fig. 3c, d). Both biotic properties were also driven by depth and CHL was driven by watershed percent forest cover (Fig. 3c, d). The overall variation explained by models was greater for abiotic properties ($R^2 = 0.62$ for TN and $R^2 = 0.50$ for TP) than for biotic properties (CHL $R^2 = 0.44$ and aquatic vegetation $R^2 = 0.41$).

Supplemental analysis of equal numbers of ecosystems across freshwater types generally corroborated these findings. However, equal sample size models pointed to summer temperature as an additional driver of CHL and both depth and freshwater type as additional drivers of TP (Appendix S1). The overall variation explained by models was similar regardless of the number of ecosystems sampled (Appendix S1).

Spatial structure of biotic and abiotic properties and drivers across lakes, wetlands, and streams

Spatial structures were more different between biotic and abiotic variables than among freshwater types. For example, TP and CHL exhibited similar nested spatial structure (multi-scale), regardless of freshwater type (Fig. 4), and the first ranges of abiotic properties were generally larger than those of biotic properties across all freshwater types (except aquatic vegetation in wetlands and CHL in streams; Fig. 4; Appendix S2). Ecosystem properties that exhibited a nested spatial structure (Table 3) showed a second range at approximately 1,500–2,000 km (Appendix S2), regardless of freshwater type.

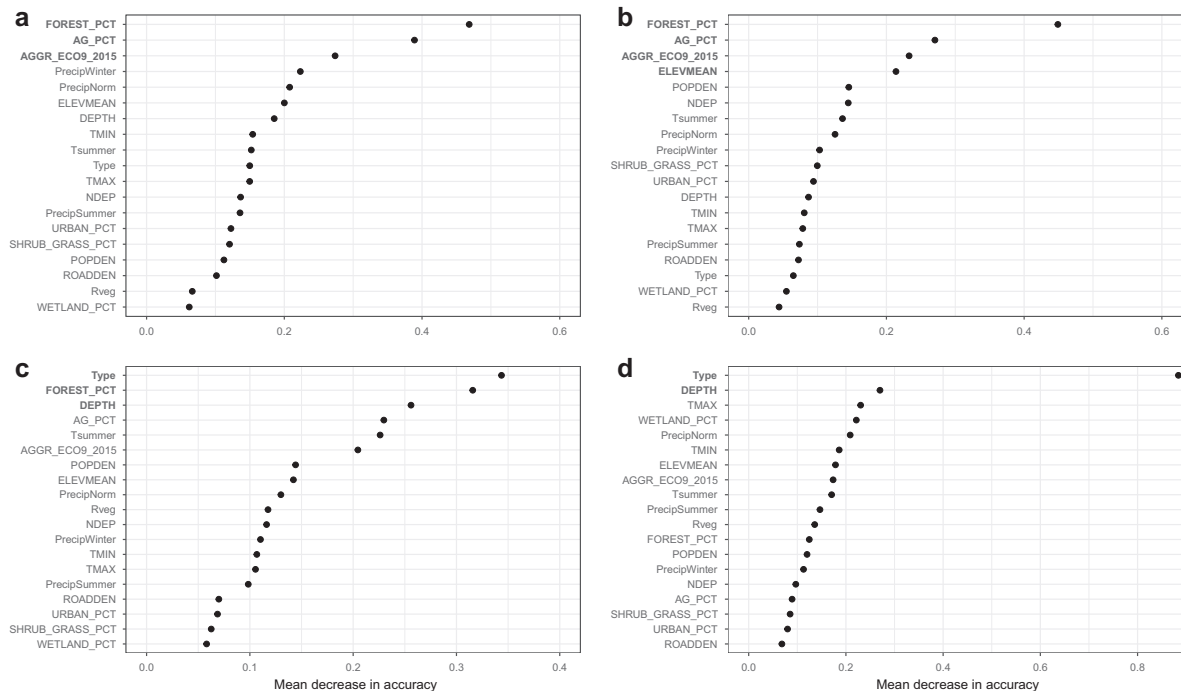


Fig. 3. Variable importance plots from random forests of (a) total phosphorus ($R^2 = 0.50$), (b) total nitrogen ($R^2 = 0.62$), (c) chlorophyll *a* ($R^2 = 0.44$), and (d) percent aquatic vegetation cover ($R^2 = 0.41$). More important predictor variables have higher values of mean decrease in accuracy, which is the unscaled difference between the observed minus the randomly permuted out-of-bag estimates. Important predictor variables are bold. See Table 2 for abbreviations of predictor variables.

Despite these similarities in regional spatial structure across freshwater types, we also detected some interesting differences across types. For example, aquatic vegetation in lakes and wetlands had single-scale structures (albeit wetlands had a very large range), whereas in streams there was a nested spatial structure (Fig. 4). For TP and TN, the first range of spatial autocorrelation was largest for lakes (757 and 1,195 km, respectively), followed by wetlands (605 and 789 km, respectively), and then streams (271 and 389 km, respectively; Appendix S2). Finally, TN in lakes displayed a nested spatial structure, wetlands a single spatial structure, and streams a nested spatial structure with no second plateau (Fig. 4, Table 3), which suggests a broad-scale gradient and no end to spatial autocorrelation of TN across the study area.

We found that the spatial structure of the top drivers reflected the structure of the associated ecosystem property (Fig. 5). For example, the spatial structures of watershed percent forest and agriculture are multi-scaled and look similar to the structure of TP and CHL (Figs. 4 and 5). Depth shows single-scale structure with a small range of spatial autocorrelation, which is similar to the small ranges found for aquatic vegetation in lakes and streams and CHL in lakes and wetlands (Figs. 4 and 5). There was no plateau for elevation, which resembles the broad-scale TN gradient (Figs. 4 and 5). These results indicate that landscape context drivers acting at

multiple spatial scales are likely shaping patterns of ecosystem properties across freshwater types at the macroscale.

DISCUSSION

For the first time, we integrated data from three freshwater types to ask do lake, wetland, and stream biotic and abiotic properties respond to similar ecosystem and watershed drivers and have similar spatial structure at the national scale? We found that drivers and spatial structure differed more between biotic and abiotic response variables than among freshwater types (lakes, wetlands, and streams). The similar response of nutrients to drivers acting at multiple spatial scales (i.e., watershed land use/cover and ecoregion membership), regardless of freshwater type, suggests that classifying an ecosystem as lentic or lotic may not be necessary for nutrient monitoring at the national scale. In contrast, classifying by freshwater types is likely important for monitoring, managing, and predicting biotic properties at such broad scales.

The intent of this study was not to best understand the variables driving differences among ecosystems *within* a freshwater type; rather, it was to look for similarities and differences in drivers and spatial structure *across* freshwater types. Therefore, we included in models only response and predictor variables that were

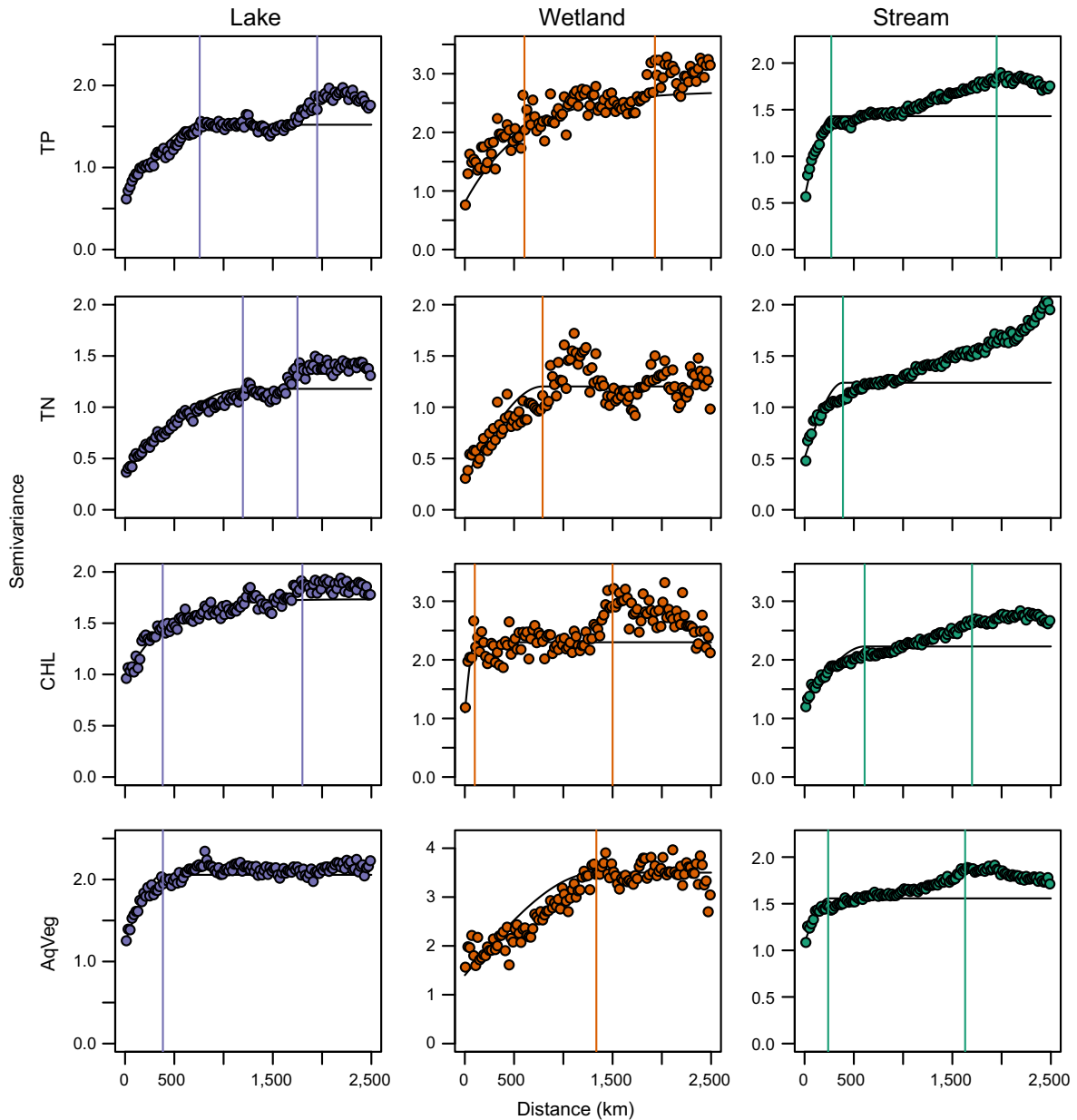


FIG. 4. Semivariograms for lakes (blue; left), wetlands (orange; middle), and streams (green; right) for response variables going top to bottom: total phosphorus, total nitrogen, chlorophyll *a*, and aquatic vegetation (AqVeg). The vertical lines on each plot represent the estimated range or ranges (single or nested spatial structure, respectively) of spatial autocorrelation. Semivariance is a measure of average dissimilarity between point pairs of a certain distance.

consistent across freshwater types. These decisions may have implications for the interpretation of our results; we did not include some high-interest ecosystem response variables nor drivers that likely contribute to some of the unexplained variation in ecosystem properties. For example, future macroscale research across freshwater types may gain important insights by considering community composition in addition to biotic abundance metrics. Because our goal was also to better understand *macroscale* patterns rather than *local*

patterns in ecosystem properties, we used relatively course-grain data and techniques appropriate for regional analyses. Although these facts preclude us from capturing fine-scale patterns, our research demonstrates the potential utility of such macroscale, inter-ecosystem, and integrative research, and suggests that nutrient concentrations of lakes, wetlands, and streams may respond similarly to future global change at macroscales.

Macroscale understanding of the spatial patterns and drivers of wetlands requires more scientific attention for

TABLE 3. Summary results from RF models and semi-variogram analysis across freshwater types and response variables.

Response variable	Freshwater type	Variable type	Spatial structure	Top driver
Total phosphorus	lake	abiotic	nested	% forest
Total phosphorus	wetland	abiotic	nested	% forest
Total phosphorus	stream	abiotic	nested	% forest
Total nitrogen	lake	abiotic	nested	% forest
Total nitrogen	wetland	abiotic	single	% forest
Total nitrogen	stream	abiotic	nested	% forest
Chlorophyll <i>a</i>	lake	biotic	nested	freshwater type
Chlorophyll <i>a</i>	wetland	biotic	nested	freshwater type
Chlorophyll <i>a</i>	stream	biotic	nested	freshwater type
Aquatic vegetation	lake	biotic	single	freshwater type
Aquatic vegetation	wetland	biotic	single	freshwater type
Aquatic vegetation	stream	biotic	nested	freshwater type

Note: Spatial structure indicates a single or nested (i.e., multi-scale) spatial structure; top driver indicates the predictor variable with the highest importance measure.

these ecosystems. At the national scale, only 400 freshwater wetland sites were sampled, as compared to thousands of lakes and streams. In addition, wetland watersheds have not been delineated at the national scale (or for the NARS wetlands). Because EPA land use/cover metrics were quantified using 1,000-m buffers around a center point, watershed areas do not vary and an uncertain and highly variable part of these estimates inevitably contains the focal wetland. Although previous studies at the national scale have used wetland buffers to characterize the landscape (e.g., Moon et al. 2017,

Stapanian et al. 2018), the use of buffers may have affected our study results and their interpretation. Thus, we have identified a gap in macroscale characterization and study of freshwater wetlands that warrants further attention.

Macroscale drivers of biota across lakes, wetlands, and streams

Scientists have long known that lakes, streams, and wetlands are not isolated from the surrounding landscape, that agricultural lands contribute nutrients to freshwaters via surface runoff and atmospheric deposition, and that riparian forests and wetlands can serve as sinks for those nutrients (e.g., Peterjohn and Correll 1984, Frissell et al. 1986, Johnston 1991, Matson et al. 1997). However, we demonstrated that concentrations of TP and TN had similar drivers regardless of freshwater type. Therefore, future land use intensification is likely to similarly negatively impact nutrients in all three freshwater types.

The nested spatial structure of TP and TN across all freshwater types (except TN in wetlands) (Fig. 4) supports the idea that drivers operating at multiple scales act on ecosystem properties, regardless of freshwater type. Previous studies of individual freshwater types have shown that regional variables such as precipitation, elevation, and regional land use/cover drive abiotic properties such as dissolved organic carbon, TP, and alkalinity (e.g., Stendera and Johnson 2006, Sobek et al. 2007, Cheruvilil et al. 2013). Another study demonstrated that although the spatial structure of the ecosystem properties generally reflects the top drivers, multiple drivers and complex interactions between scales can alter the structure of the ecosystem properties (Lapierre et al. 2018). The fact that we found similar spatial structures for TP and TN and their main predictor variables (Fig. 4 compared with Fig. 5) provides further support that land use/cover are drivers of nutrients and demonstrates the importance of understanding multi-scaled

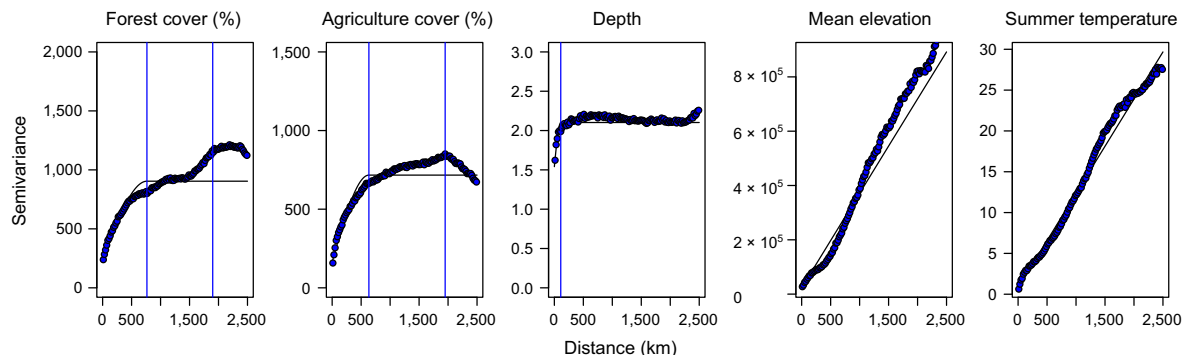


FIG. 5. Semivariograms for a subset of predictor variables identified as important using random forests: watershed percent forest, watershed percent agriculture, depth, mean elevation, and summer temperature. The vertical lines on each plot represent the estimated range or ranges (single or nested spatial structure, respectively) of spatial autocorrelation. Semivariance is a measure of average dissimilarity between point pairs of a certain distance.

spatial structure in order to interpret ecosystem relationships.

We also found several cases where drivers and spatial structure varied by freshwater types, three of which we describe here. First, when accounting for sample size differences, freshwater type emerged as a driver of TP concentrations, which were higher in wetlands than in lakes and streams in the central United States (Fig. 2). This result is supported by studies that show wetlands perform as a TP in agricultural regions (Johnston 1991, Fergus et al. 2011). Second, the spatial structure of TN varied across freshwater types (i.e., nested for lakes, single for wetlands, and nested with no second plateau for streams). Regional differences in various types of TN input (e.g., fertilizer, atmospheric deposition) at the national scale may alter the effects of drivers on TN concentrations (Bellmore et al. 2018). Therefore, differences in nutrient inputs or processing at the ecosystem scale (Saunders and Kalff 2001) can interact with landscape context variables (e.g., land cover, ecoregion, precipitation) and cause differences in spatial structure among freshwater types. Third, contrary to our expectations, we found that streams had the shortest range of spatial autocorrelation for TP and TN as compared to lakes and wetlands. This result may be due to the longitudinal form and flow of a stream, the fact that streams run through landscapes of different land use/cover, and the patchy physical structure of streams (i.e., riffles and pools), all of which may result in more local heterogeneity for streams (Peterson et al. 2006, McGuire et al. 2014) than for lakes or wetlands. Therefore, we advocate the use of hydrologic distance (rather than Euclidean), which is restricted to the stream network, to better represent spatial autocorrelation of nutrients within stream networks (Peterson et al. 2006, McGuire et al. 2014) for future study.

Macroscale drivers of biota across lakes, wetlands, and streams

We found that the drivers of biotic properties depended on freshwater type, likely due to large differences in ecosystem form and function among lakes, wetlands, and streams. For example, nutrient processing differs between lentic and lotic ecosystems, whereby the algal abundance per unit phosphorus increases with residence time (streams < lakes) (Soballe and Kimmel 1987). Aquatic vegetation studied across ponds, rivers, and streams found ponds to be most diverse due to freshwater-type-specific properties such as permanence, depth, and flow (Williams et al. 2003). Interestingly, we found that CHL was driven by a mix of predictor variables that were important for both aquatic vegetation and freshwater nutrients. Although CHL is a measure of algal open-water biomass (i.e., it is a biotic property) and it is used as an indicator of primary productivity, this property is determined from water column samples and is often highly correlated with TP and TN (Dillon and Rigler

1974, Smith 1982). Therefore, it follows that our models of CHL had similar drivers to both nutrients (e.g., percent forest) and aquatic vegetation (e.g., depth, type).

As with the abiotic response variables, we found that predictor variables at multiple spatial scales were important for shaping biotic ecosystem properties. For example, the nested spatial structure of CHL across all freshwater types (Fig. 4) implies that this property is likely shaped by drivers acting at both relatively local and regional scales (i.e., 382 km and 1,800 km for lakes, respectively). Although this spatial structure most closely resembles that of watershed percent forest (landscape scale, multi-scale), the shorter first range of spatial autocorrelation may be indicative of the importance of ecosystem depth for primary production (ecosystem scale, single-scale; Fig. 5). In addition, summer temperature (broad-scale spatial structure; Fig. 5) can interact with landscape and ecosystem scale drivers to influence CHL concentrations (e.g., Flanagan et al. 2003, Wagner et al. 2011).

The presence of a multi-scaled spatial structure for aquatic vegetation depended on freshwater type (Fig. 4). The spatial structure was single-scaled for lakes and wetlands, likely influenced by depth (ecosystem scale, single-scale), which further supports the idea that processes driving biotic properties vary by freshwater type. Unfortunately, we are unable to infer further drivers shaping these patterns across freshwater types because freshwater type was the only other driver that explained aquatic vegetation percent cover. Although our study is an important first step in the integrated research of lakes, wetlands, and streams that shows local and regional drivers likely shape CHL across freshwater types, more research is needed across freshwater types to understand regional drivers of aquatic vegetation.

CONCLUSION

Broad-scale studies of lakes *or* wetlands *or* streams have improved scientists' ability to make predictions about how freshwater biotic and abiotic properties may respond to changes in climate and land use intensification. However, by studying lakes *and* wetlands *and* streams, we elucidated drivers and spatial structure of abiotic and biotic properties that were previously unknown. For example, we found that drivers and spatial structure of nutrients occur at multiple spatial scales and that undisturbed watersheds (those with high percent forest cover) drive ecosystem properties, regardless of freshwater type. Freshwater type was an important driver of biotic variables, which exhibited both single and nested spatial structures. This is the first time that these broad-scale patterns have been quantified across freshwater types and our results lend support to the idea that both abiota and biota have multi-scale spatial structure, which can inform landscape-scale management and conservation efforts. Our research also demonstrates that national and multi-national ecosystem assessments

(e.g., EPA NARS, Canada lake-PULSE; European Water Framework Directive; EC 2000) can provide a valuable source of data to extend the study of multiple ecosystem types to build macroscale knowledge and inform land use management and policy at macroscales.

ACKNOWLEDGMENTS

We thank Emi Fergus for her feedback throughout the research project; Patricia Soranno, Ian McCullough, Joe Stachelek, Nicole Smith, and Gregg Serenbetz for providing feedback on early drafts; Gregg Serenbetz, Amanda Nahlik, and Richard Mitchell for answering questions about the surveys, data, and methods; and reviewers for their constructive feedback. This research was partially supported by U.S. National Science Foundation Macrosystems Biology program (EF #1638679), MSU College of Agriculture and Natural Resources Summer Fellowship, and the Robert C. Ball and Betty A. Ball Fisheries and Wildlife Fellowship. The NRSA 2008-2009, NWCA 2011, and NLA 2012 data were a result of the collective efforts of dedicated field crews, laboratory staff, data management and quality control staff, analysts, and many others from EPA, states, tribes, federal agencies, universities, and other organizations. The views expressed in this article are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

LITERATURE CITED

- Allan, J. D., D. L. Erickson, and J. Fay. 1997. The influence of catchment land use on stream integrity across multiple spatial scales. *Freshwater Biology* 37:149–161.
- Allan, J. D., M. A. Palmer, and N. L. Poff. 2005. Climate change and freshwater ecosystems. Pages 272–290 in T. E. Lovejoy and L. Hannah, editors. *Climate change and biodiversity*. Yale University Press, New Haven, Connecticut, USA.
- Allen, T. F. H., and T. B. Starr. 1982. *Hierarchy perspectives for ecological complexity*. University of Chicago Press, Chicago, Illinois, USA.
- Baskett, E., and G. A. Jordan. 1995. Characterizing spatial structure of forest landscapes. *Canadian Journal of Forest Research* 25:1830–1849.
- Bellmore, R. A., J. E. Compton, J. R. Brooks, E. W. Fox, R. A. Hill, D. J. Sobota, D. J. Thornbrugh, and M. H. Weber. 2018. Nitrogen inputs drive nitrogen concentrations in U.S. streams and rivers during summer low flow conditions. *Science of the Total Environment* 639:1349–1359.
- Breiman, L. 2001. Random forests. *Machine Learning* 45:5–32.
- Chaloner, D. T., and R. S. Wotton. 2011. Overview: the links that bind aquatic ecosystems. *Journal of the North American Benthological Society* 30:751–761.
- Cheruvilil, K. S., P. A. Soranno, K. E. Webster, and M. T. Bremigan. 2013. Multi-scaled drivers of ecosystem state: Quantifying the importance of the regional spatial scale. *Ecological Applications* 23:1603–1618.
- Cutler, D. R., T. C. Jr Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler. 2007. Random forests for classification in ecology. *Ecology* 88:2783–2792.
- Dillon, P. J., and F. H. Rigler. 1974. The phosphorus–chlorophyll relationship in lakes. *Limnology and Oceanography* 28:792–795.
- European Community (EC). 2000. Directive 2000/60/EC of the European parliament and of the council of 23 October 2000 establishing a framework for community action in the field of water policy. *Official Journal of the European Communities* L327:1–72.
- Fergus, C. E., P. A. Soranno, K. S. Cheruvilil, and M. T. Bremigan. 2011. Multiscale landscape and wetland drivers of lake total phosphorus and water color. *Limnology and Oceanography* 56:2127–2146.
- Flanagan, K. M., E. McCauley, F. Wrona, and T. Prowse. 2003. Climate change: the potential for latitudinal effects on algal biomass in aquatic ecosystem. *Canadian Journal of Fisheries and Aquatic Sciences* 60:635–639.
- Frissell, C. A., W. J. Liss, C. E. Warren, and M. D. Hurley. 1986. A hierarchical framework for stream habitat classification: viewing streams in a watershed context. *Environmental Management* 10:199–214.
- Galbraith, L. M., and C. W. Burns. 2007. Linking land-use, water body type and water quality in southern New Zealand. *Landscape Ecology* 22:231–241.
- Heffernan, J. B., et al. 2014. Macrosystems ecology: Understanding ecological patterns and processes at continental scales. *Frontiers in Ecology and the Environment* 12:5–14.
- Hollister, J. W., W. B. Milstead, and B. J. Kreakie. 2016. Modeling lake trophic state: a random forest approach. *Ecosphere* 7:1–14.
- Hotchkiss, E. R., S. Sadro, and P. C. Hanson. 2018. Toward a more integrative perspective on carbon metabolism across lentic and lotic inland waters. *Limnology and Oceanography Letters* 3:57–63.
- Johnson, R. K., W. Goedkoop, and L. Sandin. 2004. Spatial scale and ecological relationships between the macroinvertebrate communities of stony habitats of streams and lakes. *Freshwater Biology* 49:1179–1194.
- Johnston, C. A. 1991. Sediment and nutrient retention by freshwater wetlands: Effects on surface water quality. *Critical Reviews in Environmental Science and Technology* 21:491–565.
- Kalff, J. 2002. *Limnology*. Prentice-Hall, Upper Saddle River, New Jersey, USA.
- King, K. 2018. Lake, wetland, and stream biotic and abiotic properties from the National Aquatic Resource Surveys. Knowledge Network for Biocomplexity. <https://doi.org/10.5063/f13j3b5d>
- King, K. B. S. 2019. Lakes, wetlands, and streams at the national scale (GitHub repository) (Version 1.0). <https://doi.org/10.5281/zenodo.3246537>
- Kling, G. W., G. W. Kipphut, M. M. Miller, and W. J. O'Brien. 2000. Integration of lakes and streams in a landscape perspective: the importance of material processing on spatial patterns and temporal coherence. *Freshwater Biology* 43:477–497.
- Lapierre, J. F., S. M. Collins, D. A. Seekell, K. Cheruvilil, P.-N. Tan, N. K. Skaff, Z. E. Taranu, C. E. Fergus, and P. A. Soranno. 2018. Similarity in spatial structure constrains ecosystem relationships: Building a macroscale understanding of lakes. *Global Ecology and Biogeography* 27:1251–1263.
- Legendre, P., and M. J. Fortin. 1989. Spatial pattern and ecological analysis. *Vegetatio* 80:107–138.
- Levin, S. A. 1992. The problem of pattern and scale in ecology. *Ecology* 73:1943–1967.
- Liaw, A., and M. Wiener. 2002. Classification and regression by randomForest. *R News* 2:18–22.
- Lottig, N., E. H. Stanley, P. C. Hanson, and T. K. Kratz. 2011. Comparison of regional stream and lake chemistry: Differences, similarities, and potential drivers. *Limnology and Oceanography* 56:1551–1562.
- Matson, P. A., W. J. Parton, A. G. Power, and M. J. Swift. 1997. Agricultural intensification and ecosystem properties. *Science* 379:285–357.
- McGill, B. J. 2019. The what, how and why of doing macroecology. *Global Ecology and Biogeography* 28:6–17. <https://doi.org/10.1111/geb.128555>

- McGuire, K. J., C. E. Torgersen, G. E. Likens, D. C. Buso, W. H. Lowe, and S. W. Bailey. 2014. Network analysis reveals multiscale controls on streamwater chemistry. *Proceedings of the National Academy of Sciences USA* 111:70370–77035.
- Moon, J. B., T. H. DeWitt, M. N. Errend, R. J. F. Bruins, M. E. Kentula, S. J. Chamberlain, M. S. Fennessy, and K. J. Naithani. 2017. Model application niche analysis: assessing the transferability and generalizability of ecological models. *Ecosphere* 8:e01974.
- Nöges, P., C. Argillier, A. Borja, J. M. Garmendia, J. Hanganu, V. Kodeš, F. Pletterbauer, A. Sagouis, and S. Birk. 2016. Quantified biotic and abiotic responses to multiple stress in freshwater, marine and ground waters. *Science of the Total Environment* 540:43–52.
- Omernik, J. M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* 77:118–125.
- O'Reilly, C. M., et al. 2015. Rapid and highly variable warming of lake surface waters around the globe. *Geophysical Research Letters* 42:10773–10781.
- Pebesma, E. J. 2004. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* 30:683–691.
- Perry, J. N., A. M. Liebhold, M. S. Rosenberg, J. Dungan, M. Miriti, A. Jakomulska, and S. Citron-Pousty. 2002. Illustrations and guidelines for selecting statistical methods for quantifying spatial pattern in ecological data. *Ecography* 25:578–600.
- Peterjohn, W. T., and D. L. Correll. 1984. Nutrient dynamics in an agricultural watershed: observations on the role of a riparian forest. *Ecology* 65:1466–1474.
- Peters, D. P. C., B. T. Bestelmeyer, and M. G. Turner. 2007. Cross-scale interactions and changing pattern-process relationships: consequences for system dynamics. *Ecosystems* 10:790–796.
- Peters, D. P. C., P. M. Groffman, K. J. Nadelhoffer, N. B. Grimm, S. L. Collins, W. K. Michener, and M. A. Huston. 2008. Living in an increasingly connected world: A framework for continental-scale environmental science. *Frontiers in Ecology and the Environment* 6:229–237.
- Peterson, E. E., A. A. Merton, D. M. Theobald, and N. S. Urquhart. 2006. Patterns of spatial autocorrelation in stream water. *Environmental Monitoring and Assessment* 121:571–596.
- Poff, L. N. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. *Journal of the North American Benthological Society* 16:391–409.
- R Core Team. 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Read, E. K., et al. 2015. The importance of lake-specific characteristics for water quality across the continental United States. *Ecological Applications* 25:943–955.
- Rose, K. C., R. A. Graves, W. D. Hansen, B. J. Harvey, J. Qiu, S. A. Wood, C. Ziter, and M. G. Turner. 2017. Historical foundations and future directions in macrosystems ecology. *Ecology Letters* 20:147–157.
- Rossi, R. E., D. J. Mulla, A. G. Journel, and E. H. Franz. 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. *Ecological Monographs* 62:277–314.
- Rotjan, R. D., and J. Idjadi. 2013. Surf and Turf: Toward better synthesis by cross-system understanding. *Oikos* 122:285–287.
- Saunders, D. L., and J. Kalff. 2001. Nitrogen retention in wetlands, lakes and rivers. *Hydrobiologia* 443:205–212.
- Shurin, J. B., K. Cottenie, and H. Hillebrand. 2009. Spatial autocorrelation and dispersal limitation in freshwater organisms. *Oecologia* 159:151–159.
- Smith, V. H. 1982. The nitrogen and phosphorus dependence of algal biomass in lakes: An empirical and theoretical analysis. *Limnology and Oceanography* 27:1101–1111.
- Soballe, D. M., and B. L. Kimmel. 1987. A large-scale comparison of factors influencing phytoplankton abundance in rivers, lakes, and impoundments. *Ecology* 68:1943–1954.
- 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.
- Soranno, P. A., K. S. Cheruvilil, K. E. Webster, M. T. Bremigan, T. Wagner, and C. A. Stow. 2010. Using landscape limnology to classify freshwater ecosystems for multi ecosystem management and conservation. *BioScience* 60:440–454.
- Soranno, P. A., et al. 2014. Cross-scale interactions: Quantifying multi-scaled cause-effect relationships in macrosystems. *Frontiers in Ecology and the Environment* 12:65–73.
- Soranno, P. A., K. S. Cheruvilil, T. Wagner, K. E. Webster, and M. T. Bremigan. 2015. Effects of land use on lake nutrients: the importance of scale, hydrologic connectivity, and region. *PLoS ONE* 10:e0135454.
- Stanley, E. H., and P. A. del Giorgio. 2018. Toward an integrative, whole network approach to C cycling of inland waters. *Limnology and Oceanography Letters* 3:39–40.
- Stapanian, M. A., B. Gara, and W. Schumacher. 2018. Surrounding land cover types as predictors of palustrine wetland vegetation quality in conterminous USA. *Science of the Total Environment* 619:366–375.
- Stendera, S., and R. K. Johnson. 2006. Multiscale drivers of water chemistry of boreal lakes and streams. *Environmental Management* 38:760–770.
- Stendera, S., R. Adrian, N. Bonada, M. Cañedo-Argüelles, B. Hugueny, K. Januschke, F. Pletterbauer, and D. Hering. 2012. Drivers and stressors of freshwater biodiversity patterns across different ecosystems and scales: a review. *Hydrobiologia* 696:1–28.
- Strobl, C., A.-L. Boulesteix, A. Zeileis, and T. Hothorn. 2007. Bias in Random Forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics* 8:25.
- Tonn, W. M. 1990. Climate change and fish communities: a conceptual framework. *Transactions of the American Fisheries Society* 119:337–352.
- Turner, M., R. Gardner, and R. O'Neill. 2001. *Landscape ecology in theory and practice pattern and process*. Springer-Verlag, New York, New York, USA.
- USEPA. 2007. National rivers and streams assessment: field operations manual. EPA-841-B-07-009. U.S. Environmental Protection Agency, Washington, D.C., USA.
- USEPA. 2008. National Rivers and Streams Assessment: Laboratory Methods Manual. EPA-841-B-07-010. U.S. Environmental Protection Agency, Washington, D.C.
- USEPA. 2011a. 2012 National lakes assessment. Field operations manual. EPA 841-B-11-003. U.S. Environmental Protection Agency, Washington, D.C., USA.
- USEPA. 2011b. National wetland condition assessment: field operations manual. EPA-843-R-10-001. U.S. Environmental Protection Agency, Washington, D.C., USA.
- USEPA. 2011c. National wetland condition assessment: laboratory operations manual. EPA-843-R-10-002. U.S. Environmental Protection Agency, Washington, D.C., USA.
- USEPA. 2012. 2012 National Lakes Assessment. Laboratory Operations Manual. EPA-841-B-11-004. U.S. Environmental Protection Agency, Washington, D.C.
- USEPA. 2016a. National lakes assessment 2012: a collaborative survey of lakes in the United States. EPA 841-R-16-113. U.S. Environmental Protection Agency, Washington, D.C., USA.

- USEPA. 2016*b*. National rivers and streams assessment 2008–2009: a collaborative survey. EPA/841/R-16/007. U.S. Environmental Protection Agency, Washington, D.C., USA.
- USEPA. 2016*c*. National wetland condition assessment 2011: a collaborative survey of the nation's wetlands. EPA 843-R-15-005. U.S. Environmental Protection Agency, Washington, D.C., USA.
- USEPA. 2016*d*. National Lakes Assessment 2012 (nla12_key-variables_data.csv, nla2012_waterchem_wide.csv, nla2012_wide_phabmet_10202016.csv, nla2012_wide_siteinfo_08232016.csv, and corresponding metadata files). Available from U.S. EPA website: <http://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys>
- USEPA. 2016*e*. National Rivers and Streams Assessment 2008–2009 (chem.csv, phabmed.csv, siteinfo_0.csv, land.csv, bentcond.csv, and corresponding metadata files). Available from U.S. EPA website: <http://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys>
- USEPA. 2016*f*. National Wetland Condition Assessment 2011 (nwca2011_siteinfo.csv, nwca2011_waterchem.csv, nwca2011_cond_stress.csv, nwca2011_chla.csv, nwca2011_salinity.csv, nwca2011_vegtype_grndsurf.csv, nwca2011_tree.csv, nwca2011_landscapechar.csv, nwca2011_aawaterchar_edited.csv, and corresponding metadata files). Available from U.S. EPA website: <http://www.epa.gov/national-aquatic-resource-surveys/data-national-aquatic-resource-surveys>
- Wagner, T., P. Soranno, K. Webster, and C. Cheruvilil. 2011. Landscape drivers of regional variation in the relationship between total phosphorus and chlorophyll in lakes. *Freshwater Biology* 56:1811–1824.
- Whittaker, R., K. J. Willis, and R. Field. 2001. Scale and species richness: towards a general, hierarchical theory of species diversity. *Journal of Biogeography* 28:452–470.
- Williams, P., M. Whitfield, J. Biggs, S. Bray, G. Fox, P. Nicolet, and D. Sear. 2003. Comparative biodiversity of rivers, streams, ditches and ponds in an agricultural landscape in Southern England. *Biological Conservation* 115:329–341.
- Williamson, C. E., J. E. Saros, W. F. Vincent, and J. P. Smol. 2009. Lakes and reservoirs as sentinels, integrators, and regulators of climate change. *Limnology and Oceanography* 54:2273–2282.
- Wrona, F. J., T. D. Prowse, J. D. Reist, J. E. Hobbie, L. M. J. Lévesque, and W. F. Vincent. 2006. Climate change effects on aquatic biota, ecosystem structure and function. *Ambio* 35:359–369.

SUPPORTING INFORMATION

Additional supporting information may be found online at: <http://onlinelibrary.wiley.com/doi/10.1002/eap.1957/full>

DATA AVAILABILITY

Data are available from the Knowledge Network for Biocomplexity: <https://doi.org/10.5063/f13j3b5d>. Code is available on Zenodo: <https://doi.org/10.5281/zenodo.3246537>.