Environmental Management (2008) 41:425–440
DOI 10.1007/s00267-007-9045-7

Grouping Lakes for Water Quality Assessment and Monitoring: The Roles of Regionalization and Spatial Scale

Kendra Spence Cheruvelil · Patricia A. Soranno · Mary T. Bremigan · Tyler Wagner · Sherry L. Martin

Published online: 14 December 2007

Abstract Regionalization frameworks cluster geographic data to create contiguous regions of similar climate, geology and hydrology by delineating land into discrete regions, such as ecoregions or watersheds, often at several spatial scales. Although most regionalization schemes were not originally designed for aquatic ecosystem classification or management, they are often used for such purposes, with surprisingly few explicit tests of the relative ability of different regionalization frameworks to group lakes for water quality monitoring and assessment. We examined which of 11 different lake grouping schemes at two spatial scales best captures the maximum amount of variation in water quality among regions for total nutrients, water clarity, chlorophyll, overall trophic state, and alkalinity in 479 lakes in Michigan (USA). We conducted analyses on two data sets: one that included all lakes and one that included only minimally disturbed lakes. Using hierarchical linear models that partitioned total variance into within-region and among-region components, we found that ecological drainage units and 8-digit hydrologic units most consistently captured among-region heterogeneity at their respective spatial scales using all lakes (variation among lake groups = 3% to 50% and 12% to 52%, respectively). However, regionalization schemes capture less among-region variance for minimally disturbed lakes. Diagnostics of spatial autocorrelation provided insight into the relative performance of regionalization frameworks but also demonstrated that region size is only partly responsible for capturing variation among lakes. These results suggest that regionalization schemes can provide useful frameworks for lake water quality assessment and monitoring but that we must identify the appropriate spatial scale for the questions being asked, the type of management applied, and the metrics being assessed.

Keywords Ecoregions · Hierarchical linear models · Lake classification · Landscape · Minimally disturbed lakes · Watersheds

Introduction

Many scientists and managers now view lakes as an integrated part of the landscape. This perspective exists in part because of our understanding that landscape features are hierarchically organized, such that broad-scale landscape and climatic features constrain the occurrence of local landscape features and in-lake processes (Frissell and others 1986, Tonn 1990, Poff 1997). By accounting for this hierarchy, we can better understand the factors that drive lake responses to disturbance, tease apart natural and human-induced variation among lakes, and define reference conditions, thereby setting more realistic expectations within and among lake types and spatial groupings.

Many state, federal, and international agencies have moved toward a regional approach for lake assessment and monitoring. One such approach is regionalization frameworks, which cluster geographic data to create contiguous regions. Although regionalization schemes do not account for local or site-specific variation among lakes, they may

K. S. Cheruvelil (✉)
Lyman Briggs College, Michigan State University,
35 East Holmes Hall, East Lansing, MI 48825, USA
e-mail: ksc@msu.edu

K. S. Cheruvelil · P. A. Soranno · M. T. Bremigan · T. Wagner · S. L. Martin
Department of Fisheries and Wildlife, Michigan State University, 13 Natural Resources Building, East Lansing, MI 48824, USA

capture broad-scale patterns in features such as climate, geology, hydrology, land use and land cover, soils, or vegetation that are important for understanding variation among lakes across large spatial scales. When coupled with a statistically valid sampling program, regionalization frameworks provide a logistically feasible framework for lake monitoring and assessment, especially across large geographic regions, and allow for extrapolation from sampled to unsampled lakes (Seelbach and others 2002). In fact, regionalization schemes are currently used as general frameworks to design sampling programs, assess water quality, determine the effects of acid rain and cultural eutrophication on biota, and implement monitoring strategies by many state, federal, and nongovernmental environmental agencies in the United States (*e.g.*, United States Environmental Protection Agency (USEPA), United States Geological Survey, The Nature Conservancy) and in European countries (Sandin and Johnson 2000, Santoul and others 2004, European Union Water Framework Directive 2000).

Although the majority of regionalization schemes were not originally designed for lake monitoring and assessment, these practices are very common. Using regionalization frameworks for such purposes assumes that characteristics of lakes within regions are more similar than characteristics of lakes across regions (Gallant and others 1989, Gerritsen and others 2000). Despite widespread use of regionalization schemes for lake assessment, there has been little effort to investigate this assumption by quantifying whether these frameworks partition significant amounts of regional lake variation (Jenerette and others 2002). Two possible approaches exist to address this issue. One approach is to use lake and stream monitoring data to create a new regionalization framework specifically designed to define water quality (as was done for streams by Robertson and others [2006] and for lakes by Omernik and others [1988, 1991]). The other approach is to use lake and stream monitoring data and their locations within a variety of different regionalization schemes to test their ability to partition regional lake variation. We took the latter approach because many existing regionalization frameworks are currently being used without having been tested or compared with each other to see which performs best for lake monitoring and assessment.

To date, most investigations that have taken this latter approach to assess the effectiveness of regionalization schemes and grouping schemes for aquatic ecosystem management have focused on streams (Newall and Magnuson 1999, Pan and others 2000). Furthermore, the emphasis of these investigations has often been on grouping aquatic ecosystems according to measures of community structure, such as species richness or diversity (*e.g.*, Newall and Magnuson 1999, Van Sickle and Hughes 2000, Moog and others 2004). Only a few such studies have examined the efficacy of regional frameworks to group similar lakes. For example, studies have found that littoral macroinvertebrate assemblages in Swedish lakes (Johnson 2000) and nutrients in Minnesota lakes (Heiskary and others 1987) correspond to ecoregions. We are aware of only one quantitative evaluation of the ecoregion concept for describing large-scale patterns in lake water chemistry and quality. Jenerette and others (2002) compared the classification ability of one ecoregion delineation to state political boundaries and land use clusters using 15 lake water chemistry and quality variables in 365 Northeastern United States lakes. They found that ecoregions were only 18% effective at classifying lakes. These results indicate that regionalization frameworks, such as ecoregions, may not be appropriate for lake monitoring and assessment and therefore require further testing.

Additional uncertainty exists about the spatial scale (or geographic extent) across which lake groups should be formed. Because landscape features are spatially organized, many regionalization schemes are delineated at multiple hierarchical spatial scales (Table 1). In addition, different regionalization frameworks result in regions of vastly different size. However, few studies have been conducted to determine the effect of spatial extent on the relative efficacy of different regionalization schemes. Finally, if the lake attributes being measured are not constrained spatially (*e.g.*, if lake properties vary independently over the landscape or if internal processes dominate independent of the landscape), then these frameworks will be ineffective at partitioning variance (Hawkins and Vinson 2000). These facts highlight the need for research such as ours that (1) determines the appropriate spatial scale for regionalization frameworks to be effective for lake monitoring and assessment and (2) quantifies how well regionalization schemes and other lake grouping schemes partition variability before they are implemented for ecosystem management (Johnson 2000, Omernik 2003, 2004). We addressed these issues by comparing many regionalization frameworks and by explicitly quantifying the variation among and within lake groups and the spatial autocorrelation of water quality variables with a large number of lakes across a large spatial extent.

We quantitatively compared multiple regionalization schemes and compared them with other lake grouping schemes as well as political boundaries to determine which approach maximizes among-lake group heterogeneity for a variety of water quality metrics. Using water quality data from 479 lakes in Michigan (USA), we quantified the amount of variation among and within regions (or lake groups). Because we wished to better understand patterns of lake water quality, which demands an understanding of

Table 1 Description of lake grouping schemes in this study[a]

Spatial scale	Regional				Subregional			
Group/use	Name	No. groups (area)	Average no. (range) all lakes (n = 479)	Average no. (range) MD lakes (n = 167)	Name	No. groups (area)	Average no. (range) all lakes (n = 416)	Average no. (range) MD (n = 78)
Contiguous geographically dependent								
Ecoregions (Omernik 1987)/assessment of regional patterns and trends in the extent and quality of environmental resources and relations with natural and human-related factors	*Omernik L-IIIs*[b]	5, 3, 3 (29,670)	160 (34–224)	56 (10–129)	N/A	N/A	N/A	N/A
Ecoregions (Bailey and others 1994)/ provide information for development of resources and conservation of the environment	*Bailey sections*	4, 4, 3 (36,880)	120 (9–228)	56 (15–90)	N/A	N/A	N/A	N/A
Regional landscape ecosystems (Albert 1995)/integrated resource management and planning, biologic conservation, and comparison of differences in composition, occurrence, interactions, and productivity of plants and animals among ecosystems	*RLEs*	4, 4, 4 (36,833)	120 (40–180)	42 (16–63)	*RLE subsections RLE sub-subsections*	22, 17, 6 (6,697) 45, 23, 7 (3,274)	25 (4–71) 18 (4–48)	13 (4–38) 11 (4–38)
Watersheds or hydrologic units (Seaber and others 1987)/delineation of Great Lake and river basins for reporting on various aspects of water resources	*GLBs*	4, 4, 4 (36,857)	120 (31–321)	42 (7–98)	*HUC-8s*	57, 28, 10 (2,545)	15 (4–50)	8 (4–14)
Freshwater ecoregions (Abell and others 2000)/assessment of freshwater diversity and conservation of biota	*FW Ecos*	4, 2, 2 (36,906)	240 (29–450)	84 (5–162)	N/A	N/A	N/A	N/A
Ecological drainage units (Higgins and others 2005)/classify freshwater ecosystems to capture representative freshwater biodiversity for regional conservation planning	*EDUs*	10, 7, 7 (14,759)	68 (29–157)	24 (4–56)	N/A	N/A	N/A	N/A
Counties/arbitrary political boundaries	N/A	N/A	N/A	N/A	*CO*	83, 46, 9 (1,779)	9 (4–31)	7 (4–21)
Noncontiguous geographically dependent								
Hydrologic landscape regions (Winter 2001, Wolock and others 2004)/identify areas with similar hydrologic characteristics to select representative study areas	*HLRs*	10, 5, 5 (14,490)	96 (15–245)	33 (11–72)	N/A	N/A	N/A	N/A
Human land use groups (this study)/ exploratory; to examine whether MI lakes are grouped by human land use	*Human LUs*	N/A, 4, 4	120 (117–122)	42 (39–45)	*Human LUs*	N/A, 26, 8	16 (11–24)	10 (8–14)

Table 1 continued

Spatial scale	Regional				Subregional			
Group/use	Name	No. groups (area)	Average no. (range) all lakes (n = 479)	Average no. (range) MD lakes (n = 167)	Name	No. groups (area)	Average no. (range) all lakes (n = 416)	Average no. (range) MD (n = 78)
Geographically independent								
Random lake groups (this study)/ exploratory; null model representing the lower bound or the minimum amount of variation among lake groups	*LB*	N/A, 4, 4	120 (119–120)	42 (41–42)	*Random*	N/A, 26, 8	16 (16–16)	10 (9–10)
Lake water quality groups (this study)/ exploratory; upper bound representing the maximum amount of variation among lake groups	*UB*	N/A, 4, 4	118 (92–139)	42 (33–46)	*Quality*	N/A, 26, 8	16 (5–39)	10 (6–11)

Area = Average group area (km^2) for each framework

MD = Minimally disturbed. MD lakes are natural lakes with < 25% human land use (agricultural and urban) within 500 m surrounding the lake.

[a] Ranges correspond to the ranges across regions, except in the case of lake water quality groups, for which the ranges correspond to ranges across response metrics. Lake groups within each scheme were included in analyses if n >4 lakes. Number of groups indicates the total number of lake groups in Michigan, the number in this study using all lakes, and the number in this study using only minimally disturbed lakes, respectively.

[b] Omernik L-IV ecoregions are not available for the state of Michigan

reference conditions and the degree of alteration from that state, we conducted all analyses twice: once with all lakes and once with a subset of those lakes identified as being minimally disturbed by human activities (Stoddard and others 2006). Finally, because the question of regionalization frameworks is inherently spatial, we examined how the spatial extent of each regionalization scheme influenced our results and whether spatial autocorrelation influenced the choice of "best" regionalization framework. The results of our study provide information critical for more informed lake water quality assessment and monitoring. We will begin by describing the regionalization schemes compared in our study.

Existing Regionalization Frameworks and Other Approaches to Grouping Lakes

Many different contiguous geographic regions have been defined in the United States. Perhaps most widely known are ecoregions, which are defined as units of land that are homogenous with respect to multiple landscape characteristics (Omernik & Bailey 1997), but others exist, such as watersheds (defined as the topographic area that drains water into a water body) and political boundaries, (*e.g.*, states or counties). Many different ecoregion frameworks exist both nationally and internationally. In North America, each ecoregion delineation has been developed independently and for different purposes, and each emphasizes somewhat different sets of landscape features (Table 1). For example, Omernik's Level III ecoregion sections (Omernik L-IIIs) are based on land use and land cover, soils, land surface form, and potential natural vegetation and were developed to examine patterns and trends in environmental resources (Omernik 1987). Bailey's ecoregion sections (Bailey sections), which are based on climate, potential vegetation, natural land cover, and terrain, and regional landscape ecosystem sections (RLEs), which are based on climate and physiography, were both developed for resource and ecosystem management, planning, and biologic conservation (Bailey and others 1994, Albert 1995, Bailey 2005). In contrast, freshwater ecoregions (FW Ecos) are based on potential freshwater species assemblages and were developed for freshwater biodiversity conservation (Abell and others 2000).

Watershed and political boundaries are also contiguous geographically dependent regions used as regionalization schemes for aquatic assessment and monitoring. We might expect that the lakes within a river watershed experience similar hydrologic and landscape features. For the state of Michigan, the Great Lake basins (GLBs) are watersheds delineated at a relatively large spatial scale (GLB average area = 36,857 km^2). Nested within these basins are the most common watershed delineation used in the United States, eight-digit hydrologic units (HUC-8; also referred to as "subbasins"), which are based on delineations of major river hydrologic features developed by the United States Geological Survey (Seaber and others 1987). However, because HUC-8s may or may not completely overlap with a river's true topographic watershed (Omernik 2003) and are based on rivers, we do not know whether they represent a valuable and accessible framework for grouping lakes. Although one might assume that political boundaries have little to do with the ecology of aquatic ecosystems, they have been, and to some degree continue to be, used for lake management purposes. For example, land use planning is often conducted at the county level and water quality is managed at the state level.

One regionalization framework represents a combined approach of ecoregions and watersheds. Ecological drainage units (EDUs) were developed by agglomerating HUC-8s using landscape features, such as climate and landform, to form contiguous, geographically dependent regions (Higgins and others 2005). These regions were developed to classify freshwater ecosystems to capture representative freshwater biodiversity for regional conservation planning.

Compared with contiguous regionalization schemes, the landscape can also be classified geographically into noncontiguous lake groups. In other words, lakes can be assigned to the same "landscape group" but not be physically located next to each other. For example, hydrologic landscape regions (HLRs) are based on hydrologic, climatic, and geologic data and result in noncontiguous, hydrologically similar land areas (Winter 2001, Wolock and others 2004). Finally, lakes can be grouped by assigning each lake group membership according to a particular landscape feature (e.g., major land use and land cover classes that occur throughout any region).

Approach: Data and Methods

Study Lakes and Water Quality Data

The 479 study lakes are located in Michigan. Water quality data were collected by the Michigan Department of Environmental Quality from public lakes >20 ha from 1974 to 1984 using standard sampling and laboratory procedures. All data are from the stratified summer season (July, August, and September); all data were collected from the epilimnion; each lake is represented in the database once. For lakes that were sampled more than once in a summer, we randomly chose one sample date; for lakes that were sampled more than once during the decade, we chose the most recent year. We examined five response variables that were available for a large number of lakes across the state, that exhibited wide response ranges, and that indicated water clarity (*i.e.*, Secchi disk depth; "Secchi" hereafter), risk of acidification (alkalinity), and trophic status (total phosphorus, total nitrogen, and chlorophyll *a*) (Table 2).

For statistical analyses, nonnormally distributed response variables were natural log transformed to accommodate the assumption of normality and homogeneity of variance. We also created a composite response variable for overall lake trophic state using principal component analysis (PCA) on the correlation matrix (Systat 9.0; SPSS, Chicago, IL). We used the "broken-stick" criterion to determine how many axes to retain (Jackson 1993, Legendre & Legendre 1998), and ± 0.63 was used as the loadings cut-off (Tabachnik & Fidell 1989). We identified one axis that included Secchi, chlorophyll *a*, total phosphorus, and total nitrogen and accounted for approximately 60% of the original variation (Table 3). Note that for a few lakes, ≥1 response variables were not sampled, resulting in smaller sample sizes for trophic state variable analyses compared with analyses of each individual response variable (Table 3).

Table 2 Abbreviation, minimum, mean, and maximum for the five water chemistry and water clarity response variables using the all lake and minimally disturbed lake data sets at the regional and subregional spatial scales[a]

Response variable (units)	Abbr	Regional			Subregional		
		Min	Mean	Max	Min	Mean	Max
Secchi disk depth (m)	Secchi	0.5 (0.5)	3.1 (3.1)	9.1 (9.1)	0.5 (0.5)	3.4 (3.3)	7.8 (7.0)
Chlorophyll *a* (μg/L)	Chl *a*	0.1 (0.2)	5.7 (5.9)	66 (66)	0.1 (0.2)	5.3 (6.7)	54 (54)
Total nitrogen (μg/L)	TN	66 (66)	586 (581)	2756 (1717)	111 (111)	485 (448)	1430 (1430)
Total phosphorus (μg/L)	TP	3 (3)	18 (18)	155 (155)	1 (1)	12 (12)	70 (46)
Alkalinity (mg/L $CaCO_3$)	Alk	0 (0)	97 (96)	225 (225)	1 (1)	70 (47)	186 (186)

n = 464 to 478 and 405 to 415 lakes and *n* = 145 to 169 and 68 to 78 lakes (minimally disturbed) at the regional and subregional spatial scales, respectively, depending on the response variable.

Abbr = Abbreviation

Max = Maximum

Min = Minimum

Table 3 Component loadings from PCA of the four water quality response variables using all lakes and only minimally disturbed lakes[a]

Component	Regional		Subregional	
	All lakes (n = 394)	MD lakes (n = 142)	All lakes (n = 342)	MD lakes (n = 68)
Secchi disk depth	−0.82	−0.78	−0.84	−0.85
Chlorophyll *a*	0.69	0.59	0.71	0.63
Total nitrogen	0.8	0.76	0.8	0.81
Total phosphorus	0.86	0.86	0.87	0.88
Percent total variance explained	63.5	56.5	65.0	63.4

[a] All variables are considered to dominate the first axis because component loadings are ≥ ±0.63 (Tabachnik & Fidell 1989)

Regionalization Frameworks and Other Approaches to Group Lakes

We grouped lakes using 11 grouping schemes that stem from three broadly different approaches related to whether or not the groups were geographically dependent and spatially contiguous (Table 1). All grouping schemes used ≥1 global information system (GIS) coverage to assign lakes to categories using the sources listed in Table 1 (except for water quality and random lake groups; see later).

Contiguous Geographically Dependent Regionalization Schemes

We used six published ecologically based regionalization frameworks with existing GIS coverages: three types of ecoregions, RLEs, watersheds, and EDUs (Table 1 and citations therein). All six of these regionalization schemes were previously delineated at the coarser "regional" spatial scale (Fig. 1, views i through vi), but just two (RLEs and watersheds; views viii–x) were previously delineated at the finer "subregional" spatial scale for the state of Michigan (Fig. 1, views vii through xi). We also grouped lakes by the arbitrary political boundaries of counties, which we classified as "subregional" in spatial scale (Fig. 1, view xi).

Noncontiguous Geographically Dependent Groups

We used two such lake grouping schemes. First, lakes were placed into HLRs using a GIS coverage. Based on the number of HLRs in Michigan, we classify this grouping scheme as "regional" in spatial scale. Second, for the human land use lake groups (Human LU), we used Michigan GIS land use/cover data from the Michigan Resource Information Service (Michigan Resource Information System 2000) to create a grouping scheme for each spatial scale. This database has a resolution of 2.5 ha and was created from aerial photographs of the state taken from 1978 to 1985 and classified using the Anderson Classification scheme (Anderson and others 1976). We based the number of groups and the number of lakes within Human LUs on the median number of groups and lakes for the published regionalization frameworks at each spatial scale. Group membership was based on proportion of human land use (sum of all urban and agricultural land use) within a 500 m buffer around each lake. For example, at the regional scale using all lakes, this method resulted in four groups with proportions of human land use in the buffer of 0 to 0.31, 0.33 to 0.52, 0.53 to 0.70, and 0.71 to 1.00.

Geographically Independent Groups

We used two grouping schemes to generate upper (UBs) and lower boundaries (LBs) for the amount of variation that occurred among lake groups in our data set. For the UB, the lake water quality groups were determined similarly to the Human LUs but according to values for each of the six response variables (*i.e.*, lakes were sorted by each response variable and assigned to groups that maximized response variation among those groups). For the LB, we developed lake groups randomly so that we could compare all other lake grouping schemes with a null model. We delineated these two groupings at both the regional and subregional spatial scales (Fig. 1), and we based the number of groups and the number of lakes within groups on the median number of groups and lakes for the published regionalization schemes at each spatial scale.

Statistical Methods

Analyses were conducted for each lake grouping scheme at each available spatial scale (regional and subregional). For all grouping schemes, an individual group was included if there were at least four lakes in the group (*i.e.*, individual regions or subregions with <4 lakes were removed from analyses), resulting in different data sets at each spatial

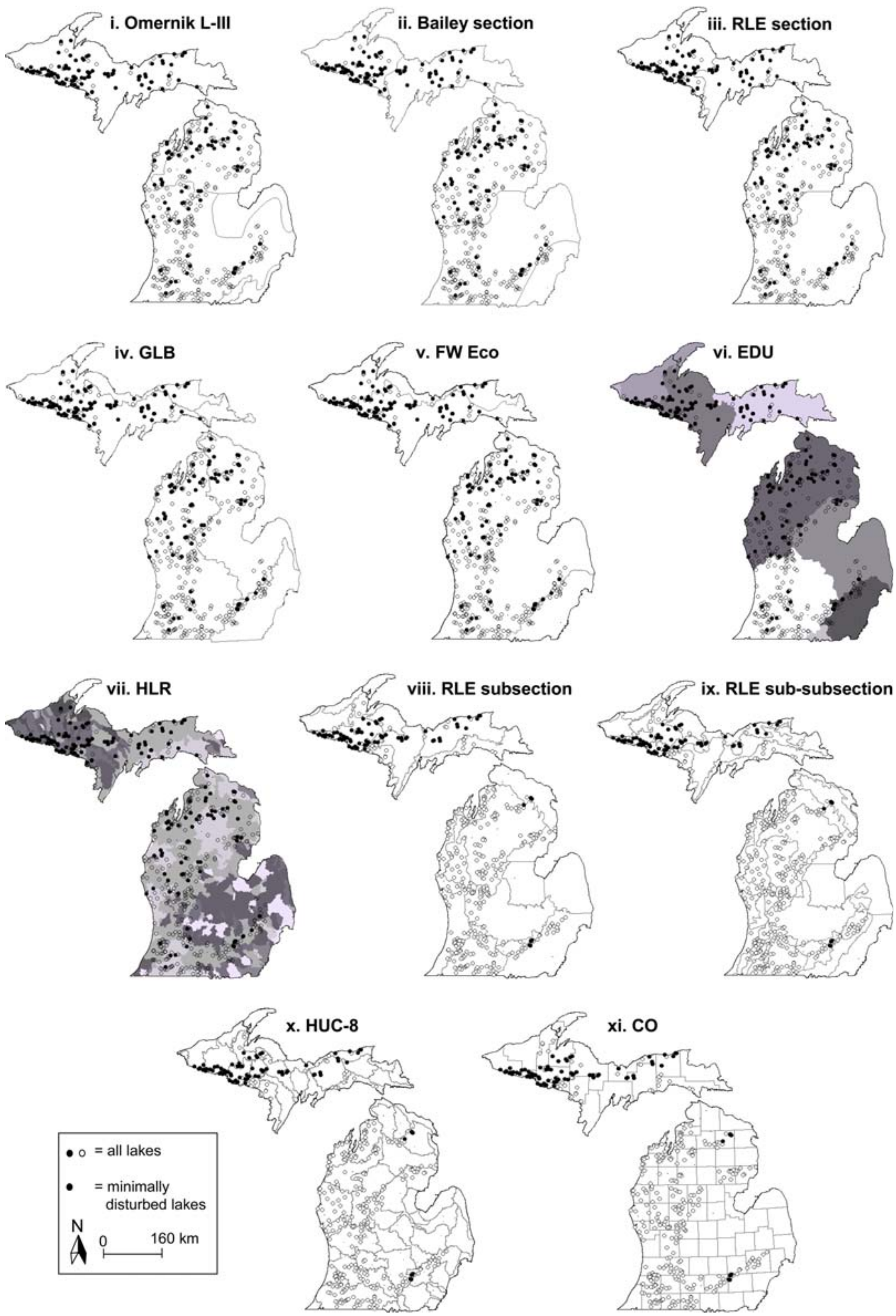

Fig. 1 Study lakes with contiguous geographically dependent regionalization schemes (i through vi) and noncontiguous geographically dependent regionalization schemes (vii) at the regional spatial scale and contiguous geographically dependent regionalization schemes (viii to xi) at the subregional spatial scale. $n = 479$ (all lakes) and 167 (minimally disturbed) at the regional scale and 416 (all lakes) and 78 (minimally disturbed) at the subregional spatial scale

scale (Tables 1 and 2). All analyses were conducted twice at each spatial scale: once using all lakes and once using only minimally disturbed lakes. We defined minimally disturbed lakes as natural lakes without dams that have

$\leq$25% human land use (agricultural plus urban land use/ cover) within the 500 m buffer surrounding the lake. When using all lakes, the lakes were widely distributed across the state of Michigan at both the regional and subregional spatial scales (Fig. 1, closed and open circles). When minimally disturbed lakes were used, the regional distribution was similar to that for all lakes; at the subregional spatial scale, most of the lakes were located in Michigan's Upper Peninsula (Fig. 1, closed circles only). To compare the minimally disturbed lake data set with the all lake data set, we conducted one-way analyses of variance (ANOVA) (Systat 9.0; SPSS). These tests demonstrated that alkalinity and total nutrients were lower in the minimally disturbed lake data set than in the all lake data set at both spatial scales ($p < 0.01$). Secchi, chlorophyll *a*, and trophic state were similar in the two data sets at both spatial scales ($p >$ 0.06).

Comparisons of Lake Grouping Schemes

To determine which regionalization framework maximizes within-group homogeneity, we used hierarchical linear models (HLMs). This is a multivariate approach specifically used to analyze hierarchical or nested data, such as lakes within regions (Singer 1998, Raudenbush & Bryk 2002). HLMs accommodate unbalanced sampling designs, missing data points, catagorical data, and small sample sizes. Because HLMs recognize the hierarchical nature of the data, they account for the nonindependence of lakes (our unit of analysis) within regions and separate the total variance into components at each level (*i.e.*, within and among regions). This variance partitioning allows for a quantitative test of which regionalization scheme best maximizes among-region variance (and thus maximizes within-region homogeneity; see later). All models were performed using the SAS MIXED procedure (SAS, Cary, NC). Although we describe here the approach above for lakes within "regions," the approach works for lakes within any sort of grouping scheme, which allows comparisons between contiguous regionalization frameworks, such as ecoregions, and noncontiguous lake grouping schemes, such as water quality groups.

To quantify the among-lake group percent variation, an unconditional HLM (a one-way ANOVA with random effects and no predictor variables; Raudenbush & Bryk 2002) was performed for each of the response variables, each of the grouping schemes, and the two different data sets (all lakes and minimally disturbed lakes) resulting in a total of 204 models. These models partition the variance into local (lake) and regional (lake group) variance components. An example of an unconditional model with alkalinity as the response variable is as follows:

$$\mathrm{Y_{ij}} = \gamma_{00} + \mathrm{r_{ij}} + \mathrm{u_{0j}}, \qquad (1)$$

where Y_{ij} is alkalinity for lake i in lake group j; γ_{00} is the intercept representing the grand mean alkalinity for all lake groups; r_{ij} is the local (lake) error for lake $_i$ in lake group $_j$, where $r_{ij} \sim N(0, \sigma^2)$ and σ^2 represent the within-lake group variability in alkalinity; u_{0j} is the regional error for lake group j, where $u_{0j} \sim N(0, \tau_{00})$ and τ_{00} represent the among-lake group variability in alkalinity.

We tested for the significance of τ_{00}, we present significant results at alpha levels of 0.05 and 0.10, and we calculated the intraclass correlation coefficient to measure the proportion of variance in alkalinity that is among lake groups ($\hat{\rho}$) as:

$$\hat{\rho} = \hat{\tau}_{00}/(\hat{\tau}_{00} + \hat{\sigma}^2).$$

We determined the best grouping scheme using two lines of evidence from these unconditional HLMs. We defined the best grouping scheme as the one with: (1) the largest among-lake group percent variation ($\hat{\rho}$) and (2) the lowest corrected Akaike Information Criterion (AIC_C) value, a model selection criterion that takes into account model fit and model complexity and is corrected for small sample sizes (Burnham & Anderson 2002, Johnson & Omland 2004). For comparing among models, a smaller AIC_C indicates a better-fit model, and a difference of seven indicates a significantly better model (Legendre & Legendre 1998).

Finally, because regionalization schemes are spatial frameworks, spatial autocorrelation of the response variables is likely (Legendre and others 2002). In fact, regionalization frameworks assume that lakes that are closer to one another are more similar than lakes further apart. Therefore, we examined whether the best regionalization scheme is simply the one that breaks the landscape up into the most and smallest regions to maximize the spatial autocorrelation structure of the response variables. To explore this idea, we took three approaches. First, we regressed among-group percent variation against the average area of regions and the number of regions. Second, we performed regression tree analysis (RTA) to quantify the nonlinear relationship between the average number or area of a region and the among-group percent variation for each response variable using Systat 11.0 (SPSS). RTA is a recursive data partitioning algorithm that splits the response variable into two subsets based on the value of the predictor variable (average region area) that maximizes the reduction in total residual sum of squares from the parent group to the two daughter groups (Breiman and others 1984). We examined the "proportional reduction in error," which is a goodness-of-fit statistic and similar to an r^2 value. Third, we fit spherical semivariogram models to total nutrients, Secchi, chlorophyll *a*, and alkalinity using

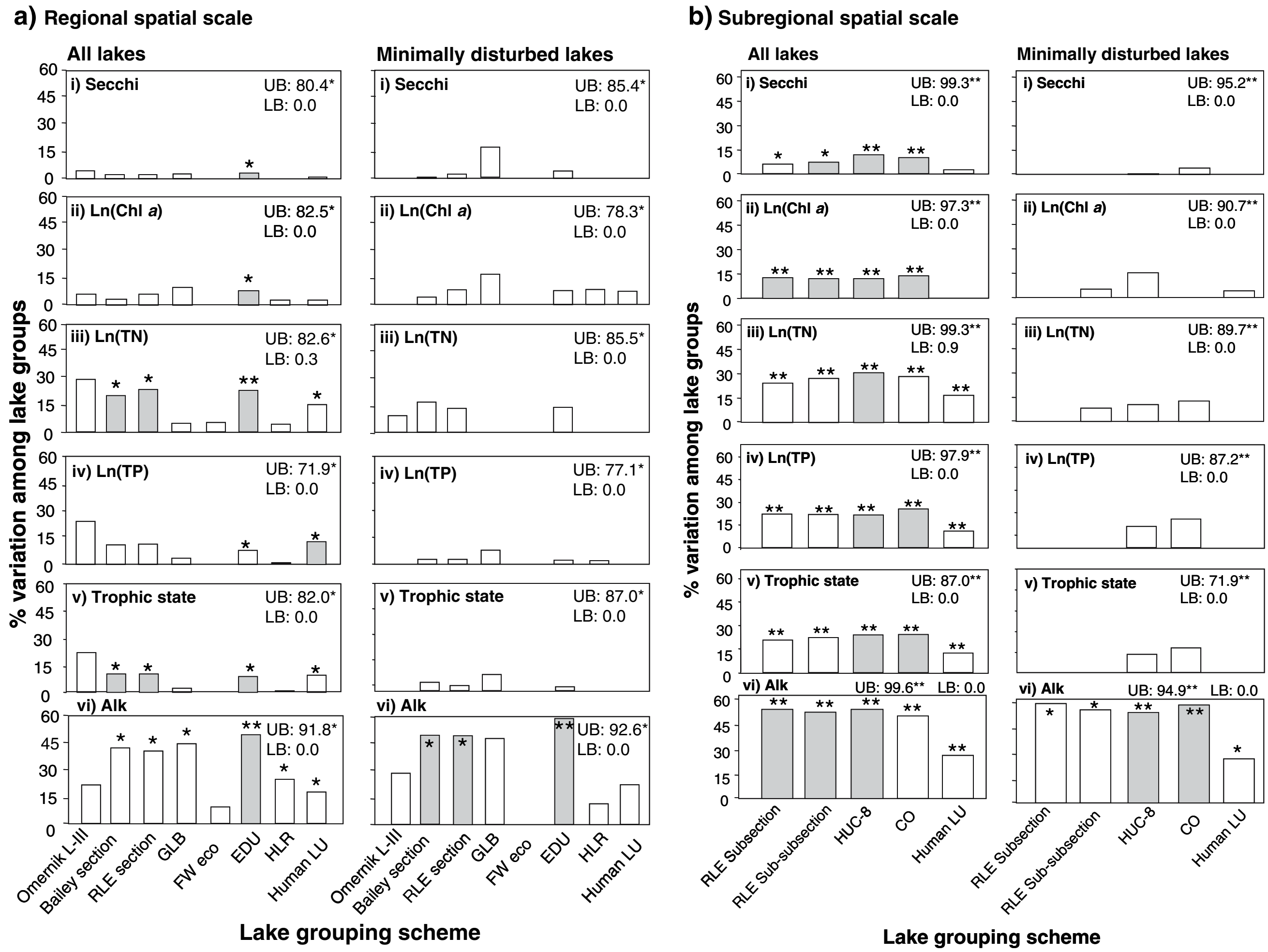

Fig. 2 Percent variation among lake groups using all lakes (left panels) and only minimally disturbed lakes (right panels) at (A) the regional and (B) the subregional spatial scales for (i) Secchi, (ii) chlorophyll *a*, (iii) total nitrogen, (iv) total phosphorus, (v) trophic state (composite variable defined by PCA on the four-trophic state variables in i through iv; see Table 3 for loadings), and (vi) alkalinity. *Percent variation is significantly different from zero at $0.05 < p < 0.10$; **Percent variation is significantly different from 0 at $p < 0.05$. Shaded bars indicate the best lake grouping scheme(s) based on AIC_C values. LBs and UBs of percent variation are based on random and quality lake groups, respectively. See Table 1 for lake grouping scheme names and descriptions. See Tables 1 and 3 for the number of lakes included in each analysis.

the all lake data set to quantify the spatial autocorrelation for each response variable. The parameters for the spherical semivariogram models were estimated using maximum likelihood in the programming environment R, using the likfit function (geoR library).

Results

Comparisons of Lake Grouping Schemes: The Regional Spatial Scale

For all response variables and when using both the all lake and the minimally disturbed lake data sets, the random lake groups resulted in no significant variation among groups ($p > 0.30$), and the water quality groups resulted in 71.9% to 97.6% variation among groups ($p = 0.07$ to 0.08) (Fig. 2A). These two grouping schemes therefore serve as our LBs and UBs, respectively, to which we compared the ability of all other lake grouping schemes to group similar lakes.

For the all lake data set at the regional spatial scale (Fig. 2A, left panel), EDUs most consistently grouped similar lakes for water quality. The percent variation among EDUs was significantly different from zero for all six response variables ($p = 0.03$ to 0.09) and ranged from 3.2% to 50.0%, depending on the variable (Fig. 2A, left panel). The variation among EDUs was highest for alkalinity (percent variation 50.0) and lowest for Secchi and total phosphorus (percent variations 3.2 and 4.0, respectively). The lowest AIC_C was also associated with EDU for five of the six variables (Fig. 2A, left panel). Although EDUs most consistently maximized within-lake group homogeneity across all response variables, Human LUs best partitioned variation for total phosphorus (Fig. 2A, left panel, box iv). For total nitrogen and trophic status, Bailey and RLE sections resulted in a significant amount of among-lake group percent variation and AIC_C values that were not significantly different from the EDU AIC_C values (Fig. 2A, left panel, boxes iii and v). For alkalinity, only Omernik L-IIIs and FW Ecos were not better than random lake groups at grouping similar lakes. However, EDUs resulted in the largest among-group percent variation, and the AIC_C was significantly better than for all other schemes (Fig. 2A, left panel, box vi). Therefore, across all six response variables using all lakes at the regional scale, EDUs most consistently maximized within-lake group homogeneity relative to the other lake grouping schemes. This result indicates that EDUs best meet the underlying assumption of these regionalization frameworks: Attributes of lakes within EDUs are more similar than attributes of lakes across EDUs.

For the minimally disturbed lakes at the regional spatial scale (Fig. 2A, right panel), we found that no lake grouping scheme resulted in a significant amount of among-lake group variation for five of the six response variables. However, for alkalinity, there was a significant amount of among-group variation as well as similar AIC_C values for Bailey's sections, RLE sections and EDUs, with EDUs resulting in the largest percent variation among its groups (Fig. 2A, right panel, box vi). Therefore, regionalization schemes appear to be more effective at capturing among-lake variation for all lakes than for minimally disturbed lakes.

Comparisons of Lake Grouping Schemes: The Subregional Spatial Scale

Similar to the results obtained at the regional spatial scale, when using both the all lake and the minimally disturbed lake data sets, our random lake groups resulted in no variation among groups ($p > 0.30$), and our water quality lake groups resulted in 71.9% to 99.7% variation among groups ($p = 0.002$ to 0.08) for all response variables (Fig. 2B).

For the all lake data set at the subregional spatial scale (Fig. 2B, left panel), HUC-8s best grouped lakes according to water quality, but we found a significant amount of among-lake group variation for all six response variables and all lake grouping schemes (with the exception of Human LUs for Secchi and chlorophyll *a*) (Fig. 2B, left panel). The variation among lake groups was highest for alkalinity and lowest for Secchi and chlorophyll *a*. HUC-8s resulted in the lowest AIC_C for total nitrogen, and HUC-8s and counties resulted in similarly low AIC_Cs for total phosphorus and the composite trophic state variable. For Secchi and chlorophyll *a*, HUC-8s, counties, and at least one of the finer-scaled RLEs resulted in similar AIC_C values and percent among-lake group variation. For alkalinity, HUC-8s and both of the finer-scaled RLEs resulted in similar AIC_C values and percent among-lake group variation. Therefore, considering all six response variables using all lakes at the subregional scale, HUC-8s most consistently captured the most variation among lake groups (Fig. 2B, left panel), indicating that this lake grouping scheme best meets the underlying assumption of regionalization frameworks.

For the minimally disturbed lakes at the subregional spatial scale (Fig. 2B, right panel), we found a significant amount of among-lake group variation only for alkalinity. All lake grouping schemes resulted in a significant amount of percent among-lake group variation for alkalinity, with HUC-8s and counties resulting in the lowest AIC_C values (Fig. 2A, left panel, box vi). Therefore, similar to our results at the regional scale, regionalization schemes do not appear to be effective at capturing among-lake variation for many water quality metrics for minimally disturbed lakes.

Comparisons of Regional and Subregional Spatial Scales

At the regional scale, the ranks of regionalization frameworks in capturing among-lake group variation for all lakes and water quality metrics (from most to least) resulted in (1) EDUs, (2) Bailey and RLE sections, (3) GLBs and Human LUs, and (4) HLRs. There was no significant among-lake group variation for any metric using Omernik L-IIIs or FW Ecos. At the subregional spatial scale, the rankings were (1) HUC-8s and counties, (2) RLE subsubsections, (3) RLE subsections, and (4) Human LUs. Between HUC-8s and counties, the AIC_C values support HUC-8s as the scheme that better partitions variation among regions.

We used the best frameworks, EDUs and HUC-8s, at the regional and subregional spatial scales, respectively, to compare the amount of variation among lake groups at these two spatial scales. Using all lakes, HUC-8s resulted in lower AIC_C values and greater among-lake group variation (7.9% on average) for all six response variables than did EDUs. Overall, this is a rather small difference in the ability of these two regionalization schemes and spatial scales to group similar lakes, especially given the large differences in the number and average size of EDUs compared with HUC-8s (Table 1).

Finally, we explored the spatial autocorrelation of our lakes to examine whether the best regionalization framework is simply the one that breaks the landscape up into the most and smallest regions, therefore capturing any spatial autocorrelation present in the data. With the exception of alkalinity, the percent variation among regions for all response variables was significantly negatively related to the average area of those regions (r^2 = 0.60 to 0.97, $p \leq$ 0.013; Fig. 3) and significantly positively related to the number of regions included in analyses (r^2 = 0.54 to 0.97, p < 0.02; data not shown; plots very similar to Fig. 3). However, there were clear exceptions to the overall trend. For example, two regionalization schemes performed better than would be predicted based on average area alone for total nitrogen (Bailey and RLE sections). Using RTA, we found a threshold in the percent variation among regions at an average region area of 29,670 km^2, which is the average size of Omernik L-IIIs (Fig. 3), above which there is no to little percent variation explained among groups. Similarly, there was little to no percent variation explained among groups for regionalization frameworks with <7 regions, which is the number of EDUs (data not shown because plots are similar to those in Fig. 3). Finally, depending on the response variable, semivariograms identified the range of spatial autocorrelation in Michigan as 167 to 491 km (27,889 to 241,081 km^2; Fig. 4). Collectively, these results indicate that (1) regionalization schemes with smaller and more regions will better group similar Michigan lakes than those with larger and fewer regions, (2) regionalization frameworks with large or few regions (>30,000 km^2 or <7 regions, respectively) are not likely to group similar Michigan lakes, and (3) some regionalization schemes better group similar Michigan lakes than others, regardless of expectations based on size or number of regions. In other words, we need to consider not just size and number of regions but also how those regions are delineated as well as the existing ranges of variation in landscape and lake features.

Discussion

We have shown that regionalization frameworks can partition significant amounts of variation (3% to 60%) in lake water chemistry and clarity variables and that regionalization schemes with smaller and more regions will better group similar Michigan lakes than those with larger and fewer regions. The amount of variation that was accounted for depended on water quality metrics, spatial scale, and whether all lakes or only minimally disturbed lakes were considered. For example, when using all lakes, the variation among regions was highest for alkalinity and lowest for Secchi and total phosphorus. Therefore, it is critical to identify the appropriate regionalization framework for the questions being asked, the management application, the needs of the users, and the metrics being assessed (*e.g.*, water quality, biology). In other words, it may not be

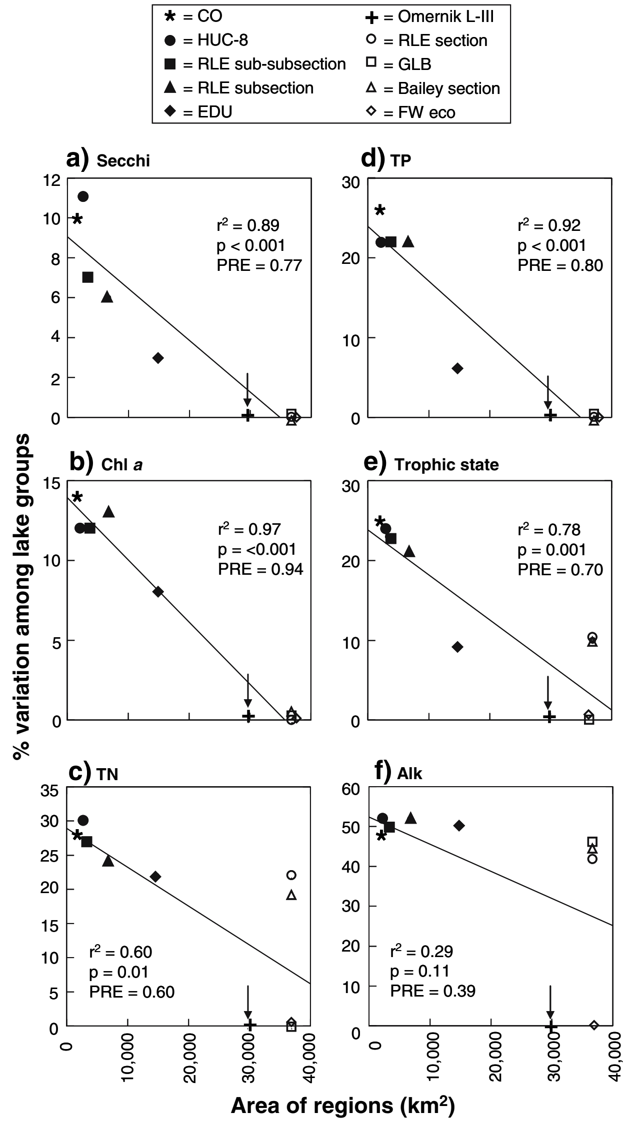

Fig. 3 Percent variation among contiguous lake groups using all lakes versus the area of regions for (a) Secchi, (b) chlorophyll *a*, (c) total nitrogen, (d) total phosphorus, (e) trophic state, and (f) alkalinity. PRE = proportional reduction in error from the regression tree analysis. The arrow indicates the threshold or split in the data for each variable. See Table 1 for lake grouping scheme names, descriptions, and number of regions. See Table 2 for water chemistry abbreviations. See Tables 1 and 3 for the number of lakes included in each analysis.

realistic to use a single regionalization scheme to meet a variety of lake management objectives.

Although many of the regionalization frameworks resulted in a significant amount of variation among regions, EDUs and HUC-8s captured the most variation among lake groups at their respective spatial scales. To better understand why these two regionalization schemes capture the most variation among lake groups, we explored the basis of these two spatial frameworks. Although it has been argued that HUCs are not always true topographic watersheds

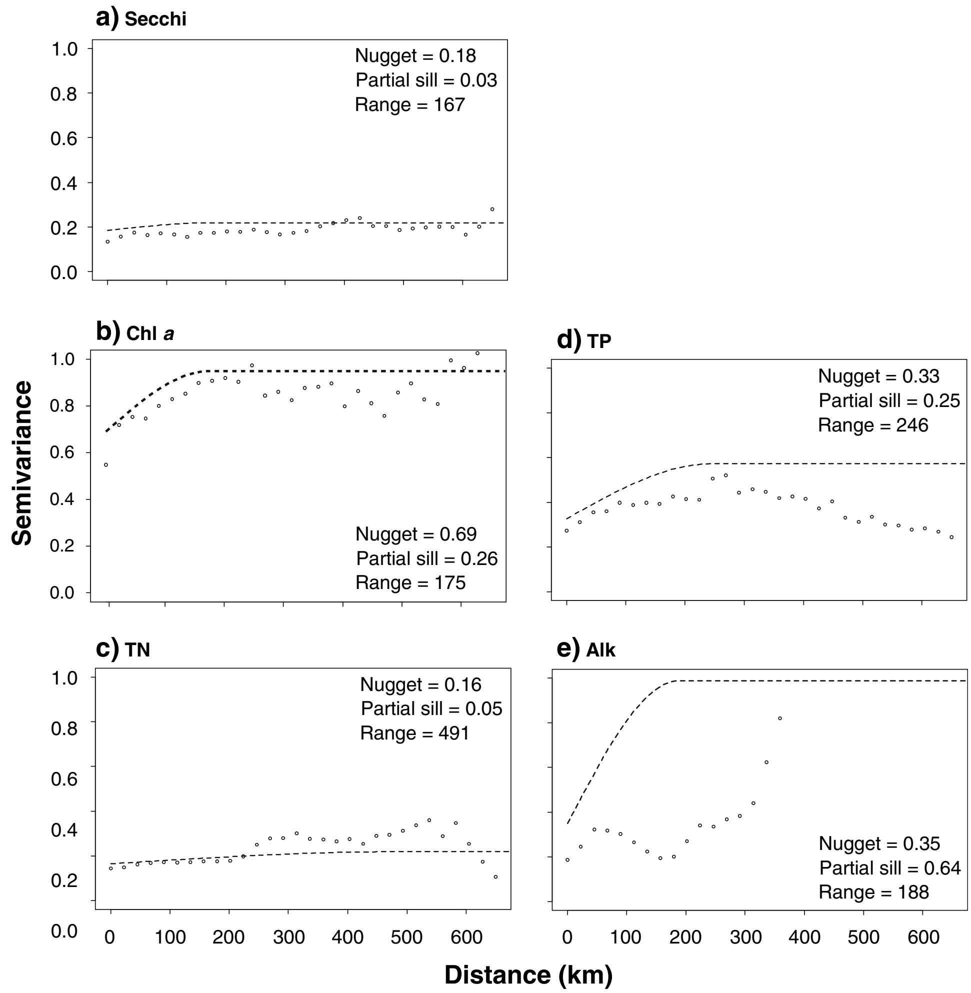

Fig. 4 Model fit and maximum likelihood parameter estimates for spherical semivariogram models using the all lake data set for (a) Secchi, (b) chlorophyll *a*, (c) total nitrogen, (d) total phosphorus, and (e) alkalinity. Nugget = estimated dissimilarity as the geographic separation approaches zero; the range indicates the distance at which the asymptote (sill) is reached and observations are independent. Partial sill = the difference between the sill and the nugget; this is part of the semivariance caused by autocorrelation. All variables were natural log transformed before analyses

(Omernik 2003, Griffith and others 1999), the results of our study suggest that HUC-8s may be sufficient approximations to topographic watersheds in Michigan for them to group lakes for water quality. A potential mechanism explaining why HUCs capture variation among lakes is that lakes within the same HUC-8 can share surface water and groundwater features through hydrologic connections. It is interesting that EDUs are actually agglomerations of HUC-8s based on abiotic landscape features, such as dominant patterns of surficial geology, drainage density, and climate (Higgins and others 2005), which is how most terrestrial ecoregions are delineated. Therefore, EDUs represent a hybrid approach between ecoregion and watershed delineation, and they are based on attributes that influence both regional and subregional patterns. Our results suggest that it is not a matter of whether ecoregions or watersheds are better at grouping lakes. Rather, both regionalization frameworks capture some variation among lakes.

We also found that although EDUs and HUC-8s are delineated at different spatial scales, they grouped lakes similarly for water quality. There were relatively small differences in the magnitude of percent variation among lake groups, and the spatial scale that maximized among-lake group percent variation depended on the lake variable. This result may be especially important for lake water quality monitoring and assessment. Depending on the question of interest, it may be more feasible for managers and policy makers to use the coarser-scale EDUs rather than the finer-scaled HUC-8s to design sampling programs, to assess lake integrity, and to set standards. For example, Michigan's Department of Environmental Quality is currently developing lake nutrient standards. Based on the results of this study and the logistics involved in developing and implementing statewide standards, we have suggested they not use a regionalization scheme for setting lake total phosphorus standards because the percent variation among regions was low for all regionalization frameworks, whereas we recommended they use EDUs as their regionalization scheme for setting lake total nitrogen standards (P. Soranno and K. S. Cheruvelil, unpublished data, 2007).

Landscape analyses such as these are challenging in that land use and natural hydrogeomorphic features are intertwined in complex ways that cannot always be distinguished (Robertson and others 2006). We have tried to separate these effects by using minimally disturbed lakes to better understand natural patterns of lake water quality metrics and to examine how different regionalization frameworks might be driven by human land use patterns. Although our minimally disturbed data set had fewer lakes and a smaller geographic range (*i.e.*, many of the minimally disturbed lakes are located in Michigan's Upper Peninsula) than the all lake data set, the ranges and means of the response variables were surprisingly similar between the two. We are able to draw two major conclusions from the results of the minimally disturbed lakes: (1) EDUs and HUC-8s most consistently capture the most among-lake group variation and (2) some lake water quality metrics may not coincide well with regionalization boundaries after removing the effects of human land use. This second conclusion is especially important when you consider that some regionalization schemes include human land use as a variable used to delineate the regions, which has important effects on water quality.

Our interpretation of the land use results make sense in that variables, such as total nutrients may be less driven by natural landscape features, such as climate and geology, than by local lake features (*e.g.*, lake depth) and human activities (*e.g.*, agriculture and urban land use). In contrast, after minimizing the effects of human disturbance by examining only minimally disturbed lakes, regionalization boundaries should still coincide well with more conservative lake attributes such as dissolved ions or alkalinity that are tightly linked with the landscape and less sensitive to human disturbance, which is what we indeed found. This latter conclusion is supported by a recent study of Kansas lakes and reservoirs. Dodds and others (2006) found a significant EPA level III ecoregion effect on chlorophyll, Secchi, and total nutrients when using the entire data set, but no ecoregion effect was detected when only reference lakes and reservoirs were analyzed. These results highlight the importance of carefully defining the question of interest and the appropriate data set for analysis before deciding on whether and which regionalization framework to include as part of lake assessment. For example, current water quality conditions are probably strongly influenced by human land use and other human impacts. Therefore, regionalization schemes based on human land use might partition patterns of eutrophication. Alternatively, understanding the natural potential of lakes (*e.g.*, reference condition) to establish a management goal (*e.g.*, nutrient criteria) might require a framework developed using only natural features or minimally disturbed lakes.

To our knowledge, our study is just the second to quantitatively compare the ability of different spatial frameworks to group lakes according to water quality. Previously, Jenerette and others (2002) found that land use groups were able to correctly classify lakes better than Omernik level IIIs or state boundaries using composite water chemistry and quality variables from 365 Northeastern United States lakes, and that ecoregions were only 18% effective at classifying lakes. We can best compare their results with ours at the regional spatial scale when using all lakes and the composite trophic state variable, and on doing so our results support some of their findings. For example, we found that Omernik L-IIIs (1987) did not result in a significant amount of variation among lake groups but that Human LUs resulted in a significant amount of variation among groups. However, contrary to the study by Jenerette and others (2002), human land use was not the best way to group similar lakes in our study. The fact that their and our studies come to different conclusions regarding the relative importance of land use for grouping lakes may be caused by differences in how the land use lake groups were calculated in the two studies, differences inherent in land use patterns, or differences in the degree to which land use influenced water quality parameters between the two study areas.

We also cannot directly compare the results of our test of political boundaries partitioning variance with the Jenerette et al study because our study area was a single large state and our political boundaries were counties; their study area was multiple, smaller states, and their political boundaries were states (Jenerette and others 2002). However, we were surprised to find that county boundaries at the subregional spatial scale partitioned variance for multiple response variables. When using all lakes and examining the combined evidence of AIC_C values and the among-lake group percent variation, counties maximized among-lake group heterogeneity as well as other lake grouping schemes for Secchi, chlorophyll *a*, total phosphorus, and trophic state (Fig. 2B, left panel). In addition to counties having small areas, it is possible that we are detecting some sort of human activity signature resulting from county-level land use planning patterns. For example, drains in Michigan are managed at the county level and may have an effect on lake Secchi, chlorophyll *a*, total phosphorus, and overall lake trophic state.

Many researchers have pointed out how important it is to account for the spatial structure of data (*e.g.*, Legendre and others 2002, Meot and others 1998), ways to do so (*e.g.*, Bell and others 1993, Gotway & Young 2002), and the difficulties inherent in quantifying the spatial structure of ecological data (Meisel & Turner 1998). When conducting studies such as this one, it is important to recognize that a regionalization framework is spatial, and as such it assumes spatial autocorrelation of data. In fact, in many instances, spatial aggregation is necessary to create meaningful units of analysis (Gerritson and others 2000). Our results from linear regressions, regression tree analyses, and semivariograms support the assumption that lakes are spatially autocorrelated, with the distance ranging from 167 to 491 km, depending on the response variable. This conclusion means that, in general, regionalization schemes with smaller and more regions will better group similar lakes than those with larger and fewer regions. However, our study quantifies these spatial relationships for the state of Michigan and identifies interesting thresholds in region size (approximately 30,000 km^2) and number of regions (n = 7). Finally, we found instances in which particular regionalization frameworks grouped similar Michigan lakes better or worse than would be expected based solely on average region size or number of regions. These results indicate that (1) we should consider how regions are delineated, in addition to the spatial scale of delineation, when deciding on a regionalization scheme for lake assessment and monitoring and (2) coarse regionalization frameworks are not likely to group similar lakes for similar study areas in terms of size, geography, and ecology. However, if we were to examine the nation as whole, the quantification of spatial autocorrelation and the best regionalization scheme will differ because of differences in spatial extent, landscape and lake features. Therefore, we need more studies across a range of geographic areas and extents to address these questions of spatial correlation and its influence on spatial patterns among lakes.

When we use regionalization as a framework within which lakes are grouped for the purposes of research, management, or regulation, we must recognize the inherent nested data structure of lakes within regions. Hierarchical linear modeling provides a necessary and useful tool for such purposes because it is a statistical modeling approach that specifically takes into account the nested structure of data and separates the total variance into components at each level (within- and among-lake groups) (Wagner and others 2006). In our study, just three regionalization schemes were nested hierarchically: (1) RLEs (sub-subsections and subsections are nested within sections), (2) "watersheds" (HUC-8s are nested within GLBs), and (3) HUC-8s (which are nested within EDUs). Although HLM can accommodate such nesting, it was beyond the scope of this study to analyze these potential three-level hierarchical relations for those regionalization frameworks. However, our results suggest that some hierarchical regionalization schemes are more likely to be relevant (*e.g.*, HUC-8s and EDUs) compared with others (*e.g.*, HUC-8s and GLBs).

Summary and Conclusions

We found that EDUs and HUC-8s most consistently (and similarly) maximize within-lake group homogeneity for a variety of water quality metrics using an all lake data set. However, there was a significant amount of among-group variation only for alkalinity after removing the apparent effects of human activities by using minimally disturbed lakes. These results will inform our ability to effectively group lakes according to water quality to facilitate the appropriate identification of reference conditions and allow for the more precise assessment of water quality, aquatic community status, and human impacts.

Our results also point to four important research areas that must be considered in the future to improve our understanding of the landscape ecology of lakes and to improve the management of large groups of lakes (*e.g.*, state and national lake assessments). First, similar to the formation of EDUs, we must develop regionalization frameworks specifically for aquatic ecosystem management that are based on natural environmental features that are known to drive lake variability. The fact that EDUs were particularly effective at grouping lakes suggests that more work could be done in this area. An important second step is to explain variation among lakes by incorporating local (or site-specific) and regional predictor variables into hierarchical models that include a regionalization scheme. Here, we identified the amount of variation that occurs at the regional and local scales. The next step is to identify how much of that variation can be explained by both local and regional landscape variables (*e.g.*, Cheruvelil 2004, Wagner and others 2007). Third, given the importance of spatial scale for aquatic ecosystems, we must further increase our spatial scale to multiple states or the entire nation to explore the effects of spatial extent, different landscape features, and lake properties on the ability of regionalization to group similar lakes. Finally, given the importance of comparing an all lake data set with a minimally disturbed lake data set, we will likely need to explore better or alternate ways to define minimally disturbed lakes. A landscape perspective provides us with the opportunity to adapt and refine the ways in which we are assessing and monitoring lake water quality.

Acknowledgments This work was supported in part by grants from the Michigan Department of Environmental Quality and the USEPA National Lakes Assessment Planning Project. We thank our anonymous reviewers, Joe Pavelitis, Jan Stevenson, Jen Tank, Peter Vaux, and Katherine Webster, for their constructive criticism and Dan Hayes and Ken Frank for their statistical guidance. Thanks to Ralph Bednarz, Mike Belligan, Jim Breck, Lidia Szabo Kraft, and Kevin Wehrly for their contributions to and creation of the landscape and lake GIS databases. Special thanks to the personnel at The Nature Conservancy, World Wildlife Fund, and the United States Geological Survey, especially Michele DePhilip, Jonathan Higgins, and David Wolock for help in acquiring GIS layers of regionalization frameworks. In addition, Laura Barrett, Stephen Bowman, Tyler Rosa, Remy Brim, Cassie Meier, Dave Meyers, and Sarah Wills aided with compiling and conducting quality control on the databases.

References

Abell RA, Olson DM, Dinerstein E, Hurley PT, Diggs JT, Eichbaum W, Walters S, Wettengel W, Allnutt T, et al. (2000) Freshwater ecoregions of North America: a conservation assessment. Island Press, Washington, DC

Albert, DA (1995) Regional landscape ecosystems of Michigan, Minnesota, and Wisconsin: a working map and classification. United States Department of Agriculture Forest Service North Central Forest Experiment Station General Technical Report No. NC-178

Anderson JR, Harvey EH, Roach JT, Whitman RE (1976) A land use sensor and land cover classification system for use with remote sensor data geological survey. Professional Paper 964. United States Government Printing Office, Washington, DC

Bailey, RG, Avers PE, King T, McNab WH (1994) Ecoregions and subregions of the United States (map; scale 1:7,500,000) (supplementary table of map unit descriptions compiled and edited by McNab WH, Bailey RG). United States Department of Agriculture Forest Service, Washington, DC

Bailey RG (2005) Identifying ecoregion boundaries. Environmental Management 34(S1):S14–S26

Bell G, Lechowicz MJ, Appenzeller A, Chandler M, Deblois E, Jackson L, Mackenzie B, Preziosi R, Schallenberg M, et al. (1993) The spatial structure of the physical-environment. Oecologia 96:114–121

Breiman L, Friedman JH, Olshen R, Stone CJ (1984) Classification and regression trees. Wadsworth International Group, Belmont, CA

Brown RS, Marshall K (1996) Ecosystem management in state governments. Ecological Applications 6:721–723

Burnham KP, Anderson DR (2002) Model selection and multimodel inference. Springer-Verlag, New York, NY

Cheruvelil KS (2004) Examining lakes at multiple spatial scales: predicting fish growth, macrophyte cover and lake physio-chemical variables. Doctoral thesis, Department of Fisheries and Wildlife, Michigan State University, East Lansing, MI

Dodds WK, Carney E, Angelo RT (2006) Determining ecoregional reference conditions for nutrients, Secchi depth and chlorophyll *a* in Kansas lakes and reservoirs. Lake and Reservoir Management 22(2):151–159

European Union Water Framework Directive (2000) Available at: http://www.europa.eu.int/comm/environment/water/water-framework/index_en.html. Accessed: 4 Nov 2007

Frissell CA, Liss WJ, Warren CE, Hurley MD (1986) A hierarchical framework for stream habitat classification: viewing streams in a watershed context. Environmental Management 10:199–214

Gallant AL, Whittier TR, Larsen DP, Omernik JM, Hughes RM (1989) Regionalization as a tool form managing environmental resources. Report EPA/600/3-89/060. United States Environmental Protection Agency, Corvallis, OR

Gerritsen J, Barbour MT, King K (2000) Apples, oranges, and ecoregions: on determining pattern in aquatic assemblages. Journal of the North American Benthological Society 19:487–496

Gotway CA, Young LJ (2002) Combining incomplete spatial data. Journal of the American Statistical Association 97:632–648

Griffith GE, Omernik JM, Woods AJ (1999) Ecoregions, watersheds, basins, and HUCs: how state and federal agencies frame water quality. Journal of Soil and Water Conservation 54:666–677

Hawkins CP, Vinson MR (2000) Weak correspondence between landscape classification and stream invertebrate assemblages: implications for bioassessment. Journal of the North American Benthological Society 19:501–517

Heiskary SA, Wilson CB, Larsen DP (1987) Analysis of regional patterns in lake water quality: using ecoregions for lake management in Minnesota. Lake and Reservoir Management 3:337–344

Higgins JV, Bryer MT, Khoury ML, Fitzhugh TW (2005) A freshwater classification approach for biodiversity conservation planning. Conservation Biology 19:432–445

Jackson DA (1993) Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. Ecology 74:2204–2214

Jenerette GD, Lee J, Waller DW, Carlson RE (2002) Multivariate analysis of the ecoregion delineation for aquatic systems. Environmental Management 29:67–75

Johnson RK (2000) Spatial congruence between ecoregions and littoral macroinvertebrate assemblages. Journal of the North American Benthological Society 19:75–486

Johnson JB, Omland KS (2004) Model selection in ecology and evolution. Trends in Ecology and Evolution 19:101–108

Kalff J (2002) Limnology. Prentice Hall, Upper Saddle River, NJ

Legendre P, Dale TMR, Fortin MJ, Gurevitch J, Hohn M, Myers D (2002) The consequences of spatial structure for the design and analysis of ecological field surveys. Ecography 25:601–615

Legendre P, Legendre L (1998) Numerical ecology, second edition. Elsevier Science BV, Amsterdam, The Netherlands

Meisel JE, Turner MG (1998) Scale detection in real and artificial landscapes using semivariance analysis. Landscape Ecology 13:347–362

Meot A, Legendre P, Borcard D (1998) Partialling out the spatial component of ecological variation: questions and propositions in the linear modelling framework. Environmental and Ecological Statistics 5:1–27

Michigan Resource Information System (2000) Michigan Department of Natural Resources. Available at: http://www.ciesin.org/datasets/mirislcover/miriscov.html.

Moog O, Schmidt-Kloiber A, Ofenbock T, Gerritsen J (2004) Does the ecoregion approach support the typological demands of the EU Water Framework Directive? Hydrobiologia 516:21–3

Newall PR, Magnuson JJ (1999) The importance of ecoregion versus drainage area on fish distributions in the St. Croix River and its Wisconsin tributaries. Environmental Biology of Fishes 55:245–254

Omernik JM (2004) Perspectives on the nature and definition of ecological regions. Environmental Management 34(S1):27–38

Omernik JM (2003) The misuse of hydrologic unit maps for extrapolation, reporting, and ecosystem management. Journal of the American Water Resources Association 39:563–573

Omernik JM (1987) Ecoregions of the conterminous United States. Annals of the Association of American Geographers 77:118–125

Omernik JM, Bailey RG (1997) Distinguishing between watersheds and ecoregions. Journal of the American Water Resources Association 33:935–949

Omernik JM, Larsen DP, Rohm CM, Clarke SE (1988) Summer total phosphorus in lakes: A map of Minnesota, Wisconsin, and Michigan, USA. Environmental Management 12(6):815–825

Omernik JM, Rohm CM, Lillie RA, Mesner N (1991) Usefulness of natural regions for lake management: analysis of variation among lakes in northwestern Wisconsin, USA. Environmental Management 15(2):281–293

Pan Y, Stevenson RJ, Hill BH, Herlihy AT (2000) Ecoregions and benthic diatom assemblages in the Mid-Atlantic Highlands streams, USA. Journal of the North American Benthological Society 19:518–540

Poff NL (1997) Landscape filters and species traits: toward mechanistic understanding and prediction in stream ecology. Journal of the North American Benthological Society 16:391–409

Raudenbush SW, Bryk AS (2002) Hierarchical linear models, second edition. Sage Publications, Thousand Oaks, CA

Robertson DM, Saad DA, Heisey DM (2006) A regional classification scheme for estimating reference water quality in streams using land-use-adjusted spatial regression-tree analysis. Environmental Management 37(2):209–229

Sandin L, Johnson RK (2000) Ecoregions and benthic macroinvertebrate assemblages of Swedish streams. Journal of the North American Benthological Society 19:462–474

Santoul F, Soulard A, Figuerola J, Céréghino R, Mastrorillo S (2004) Environmental factors influencing local fish species richness and differences between hydroecoregions in south-western France. International Review of Hydrobiology 89:79–87

Seaber PR, Kapinos FP, Knapp GL (1987) Hydrologic unit map. United States Geological Survey Water-Supply Paper 2294

Seelbach PW, Wiley MJ, Soranno PA, Bremigan MT (2002) Aquatic conservation planning: using landscape maps to predict ecological reference conditions for specific waters. In: Gutzwiller KJ (ed.), Applying landscape ecology in biological conservation. Springer-Verlag, New York, NY pp 454–478

Singer JD (1998) Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. Journal of Educational and Behavioral Statistics 24:323–355

Stoddard JL, Larsen DP, Hawkins CP, Johnson RK, Norris RH (2006) Setting expectations for the ecological condition of streams: the concept of reference conditions. Ecological Applications 16(4):1267–1276

Tabachnik BG, Fidell LS (1989) Using multivariate statistics, second edition. Harper and Row, New York, NY

Tonn WM (1990) Effects of climate change on fish. Transactions of the American Fisheries Society 119:337–352

Van Sickle J, Hughes RM (2000) Classification strengths of ecoregions, catchments, and geographic clusters for aquatic vertebrates in Oregon. Journal of the North American Benthological Society 19:370–384

Wagner T, Hayes DB, Bremigan MT (2006) Accounting for multilevel data structures in fisheries data using mixed models. Fisheries 31(4):180–187

Wagner T, Bremigan MT, Cheruvelil KS, Soranno PA, Nate NA, Breck JE (2007) A multilevel modeling approach to assessing regional and local landscape features for lake classification and assessment of fish growth rates. Environmental Monitoring and Assessment 130:437–454

Winter TC (2001) The concept of hydrologic landscapes. Journal of the American Water Resources Association 37:335–349

Wolock DM, Winter TC, McMahon G (2004) Delineation and evaluation of hydrologic-landscape regions in the United States using geographic information system tools and multivariate statistical analyses. Environmental Management 34(S1):S71–S88