Transactions of the American Fisheries Society 141:414–425, 2012

ISSN: 0002-8487 print / 1548-8659 online
DOI: 10.1080/00028487.2012.667046

ARTICLE

A Landscape-Based Classification of Fish Assemblages in Sampled and Unsampled Lakes

Kevin E. Wehrly,* James E. Breck, Lizhu Wang, and Lidia Szabo-Kraft
Institute for Fisheries Research, Michigan Department of Natural Resources and University of Michigan, 212 Museums Annex, 1109 North University Avenue, Ann Arbor, Michigan 48109, USA

Abstract
We related fish species patterns and landscape-scale environmental data from 216 Michigan lakes to identify repeatable types of fish assemblages, identify environmental factors related to assemblage types, and classify fish assemblages in unsampled lakes. Multivariate regression tree modeling of fish species abundances identified six assemblage types that were explained by degree-days during the ice-free period, lake surface area, and mean lake surface temperature. Warmwater species dominated southern lakes, while coolwater and coldwater species had higher abundances in northern lakes. Coolwater species were present in large southern lakes, whereas warmwater species were excluded from northern lakes that had low mean surface temperatures or low degree-days. These results suggest that patterns of lake fish assemblages are shaped by differences in climate as well as lake-specific differences in surface temperature regimes and in vertical availability of coldwater and coolwater habitats. Because we related fish patterns to readily available landscape-level data, our approach can be used to characterize fish assemblages in all lakes across broad geographic extents.

Inland lakes exhibit a wide range of morphological, chemical, and biological characteristics. This diversity provides tremendous recreational opportunities and other ecosystem services (Wilson and Carpenter 1999), especially in lake-rich regions. The diversity and abundance of lakes also present a number of challenges to resource managers and policy makers charged with overseeing these waters. First, the abundance of lakes in many regions limits annual sampling to only a fraction of water bodies, yet the entire population of lakes must be managed. Second, the range of physicochemical and biological characteristics among lakes in a region complicates the assessment of ecological status because the expected condition of individual lakes is typically unknown (Søndergaard et al. 2005). Finally, the diversity of lake types makes it difficult to predict how an individual lake will respond to changes in habitat, landscape development, or regulations. Because of these challenges, a lake classification system that simplifies the myriad of lakes into a limited number of ecologically meaningful types is highly desirable (Tonn et al. 1983).

The development of a lake classification can provide several benefits, such as a simplified description of fishery resources that can improve communication and understanding among managers, policymakers, and the public. Lake classification can also be used as a sampling framework for monitoring that can help managers make more efficient use of limited personnel and improve precision of assessment statistics (Dolman 1990). In addition, lake classification can provide an expectation of ecological condition that can account for sources of natural variation among lakes and can be used to objectively assess the status of individual waters (Wang et al. 2010). Finally, lake classification can help guide management and policy decisions by providing information on how similar lakes will respond to management actions (Vehanen and Aspi 1996) and environmental perturbations (Mehner et al. 2005).

A number of lake classifications have been proposed based on a variety of physical, chemical, and biological variables (Leach and Herron 1992; Schupp 1992; Emmons et al. 1999; Gassner et al. 2005; Søndergaard et al. 2005). Tonn et al. (1983), however,

*Corresponding author: wehrlyk@michigan.gov
Received May 12, 2011; accepted October 10, 2011
Published online March 12, 2012

advocated for development of lake classifications that use fish assemblage structure, arguing that this approach would bring a much-needed ecological basis to fisheries management. In this classification approach, lakes are grouped together based on similarity in the fish species they support. Lake classes are then related to environmental variables so that factors influential in maintaining distinct fish assemblages can be identified and class membership of lakes lacking fish data can be determined. The current fish-based lake classifications (Johnson et al. 1977; Harvey 1978, 1981; Schneider 1981; Tonn et al. 1983; Marshall and Ryan 1987; Dolman 1990) primarily have used in-lake environmental variables, such as lake morphology and water chemistry, to determine class membership. A limitation of this approach is that only those lakes that have been the focus of physical and chemical surveys can be classified, leaving class membership of unsampled waters unknown. The development of a lake classification based upon spatially extensive data that could be obtained from geographical information systems (GIS) would provide a cost-effective method to classify every lake across broad regions. Spatially extensive information is readily available and can include landscape-scale data that characterize climate, geology, and land cover, as well as local information that describes lake area, fetch, and elevation. Spatially extensive data can also be derived from models that predict variables within a lake from landscape characteristics (Shuter et al. 1983; Hakanson 1996).

Although spatially extensive data have been used to classify stream fish assemblages (Brenden et al. 2008), we are unaware of other studies that have developed similar classifications for lake fish assemblages. Lake fish assemblages are structured by variables operating at both local and regional scales (Jackson and Harvey 1989; Tonn 1990; Hinch et al. 1991; Jackson et al. 2001; van Zyll de Jong et al. 2005; Bertolo and Magnan 2006; Wang et al. 2010). Consequently, spatially extensive data may prove useful in predicting lake fish assemblage structure, especially when lakes are viewed across broad regions.

The goal of this study was to develop a fisheries classification that could be used to classify every lake across a state or multistate region. Our specific objectives were to (1) identify lake types based on fish species assemblages, (2) identify environmental variables from spatially extensive data sets that could be used to characterize these lake types, and (3) determine class membership of over 6,500 Michigan lakes with a surface area of 4 ha or greater.

METHODS

Environmental data.—For this analysis, we used readily available spatial data to quantify the characteristics of all 4-ha or larger lakes in Michigan. To begin, we selected polygons representing natural and manmade lakes from the 1:24,000-scale National Hydrography Dataset by using GIS (ESRI 2002). We included manmade lakes in our analysis because many lakes in Michigan are maintained by lake-level control structures and contain fish assemblages that are similar to those of natural lakes. Catchment boundaries were delineated for all lakes by using GIS algorithms to identify runoff directions based on a 30-m-resolution digital elevation model and to restrict the outmost catchment boundaries by using a 12-digit hydrological unit or aggregated hydrological units that were developed by the Michigan Department of Environmental Quality.

We calculated several measures of lake network position, morphometry, connectivity, and lake thermal regime (Table 1) because of their demonstrated importance in structuring lake fish assemblages. Measures of network position included (1) lake order, which was calculated as the stream order (Strahler 1957) of the largest tributary flowing into each lake; (2) the total

TABLE 1. Summary of characteristics for the 216 lakes used to classify fish assemblages in Michigan. Total phosphorus and chlorophyll-*a* values are epilimnetic concentrations measured during summer stratification. All variables except total phosphorus, chlorophyll *a*, and mean depth were considered in the classification analysis.

Variable	Minimum	Median	Mean	Maximum
Total phosphorus (μg/L)	0.0	14.0	17.8	120.0
Chlorophyll *a* (μg/L)	0.0	3.0	4.9	70.8
Mean depth (m)	0.5	4.5	5.1	22.7
Lake order	0	2	5.2	99
Number of lakes upstream	0	0	10.4	270
Number of lakes downstream	0	1	2.2	17
Catchment area (ha)	1	1,073	13,778	592,226
Catchment area : lake area ratio	0.04	8.4	55.7	1,362.3
Lake elevation (m)	176.7	262.7	283.6	520.7
Catchment slope	0.0	1.3	1.6	5.4
Surface area (ha)	4	94	303	4,374
Shoreline development index	1.1	1.9	2.2	8.8
Mean temperature (°C)	13.4	15.8	15.7	17.5
Degree-days	2,982	3,979	3,962	5,015

number of lakes in the tributary catchment of each lake; (3) the total number of lakes downstream between each lake and the Great Lakes (all rivers in Michigan flow into the Great Lakes); (4) catchment area; (5) catchment area : lake area ratio; and (6) lake elevation and catchment slopes, which were calculated based on the digital elevation model. Catchments were defined as the land area where surface water drains directly into lakes. In the case of lakes with tributary inputs, catchments included land area where surface water drains into rivers and then into lakes. Measures of lake morphometry included shoreline development index ($D = L/[2 \times \{\pi A\}^{0.5}]$, where L = perimeter and A = lake area). Hydrologic connectivity of all lakes was identified based on both perennial and intermittent connecting streams. Inline lakes were defined as having both inflows and outflows, headwater lakes were defined as having only outflows, and disconnected lakes were defined as having no inflows or outflows. Seasonal thermal regime of each lake was modeled as a truncated sine function by using equations from Shuter et al. (1983) for calculating maximum surface water temperature, duration of the ice-free period, and mean water temperature during the ice-free period. These temperature variables were modeled for all Michigan lakes as functions of mean annual air temperature, lake fetch, and depth (J. E. Breck, unpublished data). When depth was not available for a lake, an equation from Shuter et al. (1983) was used to calculate summer thermocline depth from fetch. We calculated degree-days (from a base of 0°C) as the product of the duration of the ice-free period and mean water temperature during the ice-free period. Water temperature during ice cover, which was used in the calculation of mean water temperature during the ice-free period, was modeled as a constant that depended on fetch (Kalff 2002). Temperature models developed for Michigan lakes had a root mean square error of 1.8°C, and temperature estimates were assumed to accurately reflect the differences in thermal regimes among lakes.

Fish assemblage data.—Fish data in this study were collected as a part of the Inland Lakes Status and Trends Program of the Michigan Department of Natural Resources (MDNR). Lakes with a surface area of 4 ha or greater were selected by means of a stratified random design that used fisheries management unit and lake size as strata. This sampling framework is used by MDNR to collect information from lakes that are representative of the population of lakes in the region. Fish sampling occurred from May to July by using standardized sampling methods. To minimize the effects of seasonal differences in catch rates, a sampling window of approximately 4 weeks began in southern Michigan when lake temperatures warmed to 14°C and progressed northward as lakes reached this temperature threshold. Information on sampling and gears is detailed by Wehrly et al. (2012). Fish in each lake were sampled by using a combination of three types of gear: trap nets (2.5- and 3.8-cm stretch mesh) in the littoral zone, boat electrofishing in the littoral zone, and experimental gill nets in deeper waters. Gill nets were 38 m in length and were made up of five 7.6-m panels of 3.8-, 5.1-, 6.4-, 7.6-, and 10.2-cm stretch mesh. Trap nets were set in less than 3 m of water with the lead perpendicular to shore. Boat electrofishing was conducted at night with two netters on board. Gill nets were set overnight on the bottom in random locations at depths greater than 3 m. Minimum netting effort per lake varied across the range of lake sizes (4 ha to >2,000 ha); trap-netting effort ranged from three nets for three nights to eight nets for four nights, and gillnetting effort ranged from two nets for two nights to six nets for four nights. Minimum sampling effort for electrofishing involved three 10-min transects across all lake sizes. Data were available for 216 lakes sampled from 2002 to 2009 (Figure 1). Lakes were sampled only once during the period of study. These lakes represented a broad range of productivity, morphology, temperature, and hydrology (Table 1).

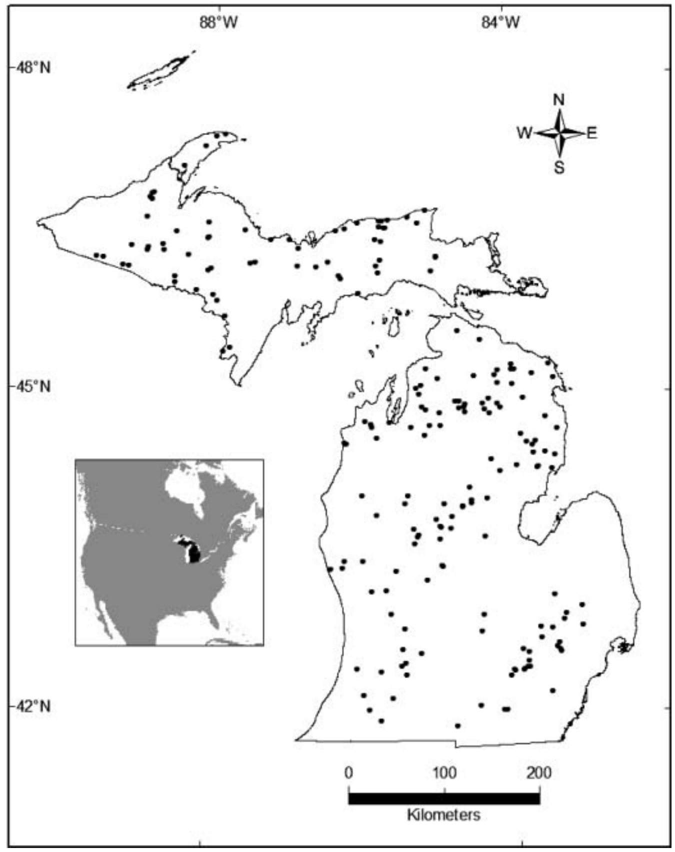

FIGURE 1. Locations of 216 Michigan lakes that were used to develop a lake classification based on fish assemblages.

In total, 85 fish species were collected across the study lakes. Based on preliminary analyses, we determined that identifying fish assemblages and predictor variables was not influenced by stocked populations of walleyes *Sander vitreus* and coldwater species or by species that were less common (<50% occurrence). For our final classification analysis, we included stocked populations and excluded species that were captured in less than 5% of the lakes. We chose this cutoff because all major game fishes were found in more than 5% of the study lakes and because it simplified our data analyses and interpretation by eliminating 42 species, the majority (78%) of which occurred in fewer than six lakes.

Coldwater fishes were found throughout the state but were poorly sampled by all gear types. As a result, in our database no single coldwater species was well represented across the study area. To better characterize lakes containing coldwater habitat, we computed a coldwater species metric by pooling the number of individuals caught in each gear type for the following nine species: brook trout *Salvelinus fontinalis*, brown trout *Salmo trutta*, burbot *Lota lota*, cisco *Coregonus artedi*, lake trout *Salvelinus namaycush*, lake whitefish *Coregonus clupeaformis*, mottled sculpin *Cottus bairdii*, rainbow trout *Oncorhynchus mykiss*, and rainbow smelt *Osmerus mordax*. Given the difficulty in capturing coldwater species, we assumed that our coldwater metric was an indicator of coldwater habitat but probably underestimated the relative abundance of coldwater species. In our analyses, the distribution of the coldwater metric was treated similar to the distribution of an individual species, and hereafter we refer to the coldwater metric as "coldwater species."

For each species and gear type, catch per unit effort (CPUE) was computed as the total number of captured individuals of each species divided by the amount of fishing effort. Effort for trap nets and gill nets was the number of lifts; effort for electrofishing was the number of minutes for which electroshock was applied. Data for CPUE were rescaled from 0 to 1 by using the Hellinger transformation (Legendre and Gallagher 2001). The Hellinger transformation maintains the Euclidean distance or similarity in species abundance among lakes, which is critical for multivariate analyses of community data (Legendre and Gallagher 2001).

Data analysis.—We evaluated whether species assemblages differed by sampling gear by using a multiresponse permutation procedure (MRPP) in PC-ORD software (McCune and Mefford 2006). The MRPP is a nonparametric test that evaluates two or more groups of observations and is robust to departures from parametric assumptions (McCune et al. 2002). We performed the MRPP on a rank-transformed distance matrix by using the Sorensen coefficient of similarity distance, and items were weighted by $n/\Sigma(n)$.

We used multivariate regression trees (MRTs) to classify lakes based on their fish assemblages and to identify environmental predictors of each class. Multivariate regression trees were well-suited for this analysis because they provide a quantitative method for splitting data into groups, defining groups based on environmental variables, and predicting group membership of lakes that are described only by environmental data. We also chose MRTs because they can accommodate both continuous and categorical data, are robust to collinearity, and can handle nonlinear relationships between variables and high-order interactions (De'ath and Fabricus 2000; De'ath 2002). The MRT analysis is a form of constrained clustering that groups lakes together by repeatedly splitting the data set to minimize dissimilarity in species assemblages among sites within a group. Each split of the data is defined by a rule determined from the environmental variable that minimizes the sum of squared Euclidean distances within each group. Once an initial split is made, each of the resulting groups is split again and all of the environmental variables are evaluated to determine rules for subsequent splits. The output from this analysis is a branched diagram or tree with each split represented by two nodes. Unsplit or terminal nodes are referred to as "leaves" and represent groups, or in this case, lake types. The splitting process is repeated until it produces a complex tree that fits the data very well but has poor predictive ability. A final, more simplified tree must be selected that balances model fit and predictive ability. In this study, MRT analysis was used to identify lake classes from the Hellinger-transformed species abundance matrix and to identify predictors of these classes from the 12 spatially extensive environmental variables (one categorical variable describing hydrologic connectivity and 11 continuous variables listed in Table 1). We used a cross-validation test ($n = 10$) to select the most complex tree within one SE of the best predictive tree (Breiman et al. 1984). We compared our MRT solution with the solutions obtained from unconstrained clustering to determine whether there was unexplained variance that could be attributed to other environmental factors not considered in this study (De'ath 2002). Unconstrained cluster solutions were calculated for two to six clusters by using complete linkage hierarchical clustering. Clusters were refined by using k-means clustering to minimize the within-cluster sums of squares. All MRT analyses were performed with R version 2.9.1 (R Development Core Team 2010) and the mvpart library (De'ath 2002).

Indicator species analysis (ISA) was used to determine fish species characteristic of each lake class resulting from the MRT (De'ath 2002). Indicator species scores were calculated as the product of a species' frequency of occurrence and relative abundance within each lake class (Dufrêne and Legendre 1997). Species with high ISA scores within a lake class were considered good indicators of that class. Monte Carlo randomizations ($n = 1{,}000$) were used to determine whether ISA scores were statistically significant (McCune et al. 2002). Indicator species scores and randomization tests were calculated by using PC-ORD (McCune and Mefford 2006).

We used leave-out-one cross validation (Efron and Gong 1983) to evaluate the ability of the MRT model to assign an unclassified lake to its appropriate class. We chose this method because it provides unbiased estimates of model performance (Efron and Gong 1983; Dolman 1990; Olden et al. 2002) and because further subsetting of our data would result in relatively few observations (<30) for most classes. Because MRT can be sensitive to the number of observations (Hastie et al. 2001), we included all observations in model development to produce the most robust lake classification possible. To calculate unbiased error rates, we left out one observation, fitted an MRT model to the remaining observations, and then used the resultant MRT model to classify the excluded observation. Overall error rate was calculated as the percentage of lakes that were correctly classified into their original species association.

To determine class membership for all 4-ha and larger lakes in Michigan, we predicted fish assemblages for over 6,500 unsampled lakes based on the environmental rules defined in our MRT model. We then characterized the landscape pattern of different lake types by summarizing the numbers and surface areas of lakes within each class and mapping their locations.

RESULTS

Forty-three fish species occurred in at least 5% of the study lakes (Table 2). Differences in catch composition were evident across gear types. The highest number of species was captured by electrofishing (43 species) followed by trap nets (40 species) and gill nets (27 species). Electrofishing tended to better capture small-bodied species (e.g., minnows and darters) and smaller size-groups of other species (e.g., black basses *Micropterus* spp. and sunfishes). Trap nets tended to capture larger black basses and sunfishes as well as benthic species (e.g., suckers and catfishes). Gill-net catches were very different from the other gear types' catches and tended to best represent coldwater species, white suckers, yellow perch, and larger predators, such as northern pike and walleyes. Results from the MRPP analysis indicated that assemblage structure across the three gear types was significantly different (test statistic $t = -159.5$, $P < 0.001$). The differences in assemblage structure between electrofishing and trap-netting ($t = -87.2$, $P < 0.001$) was relatively low compared with differences between gillnetting and electrofishing ($t = -124.0$, $P < 0.001$) or between gillnetting and trap-netting ($t = -124.8$, $P < 0.001$). Based on these differences in catch composition among sampling methods, we combined catch data from the three sampling gear types of each lake to provide a more complete picture of fish assemblage structure. Hellinger-transformed CPUE data for each species were averaged across gear types to compute a single measure of relative abundance for each species at each site.

The final MRT (i.e., within 1 SE of the best predictive tree) was made up of six leaves that were explained by lake thermal regime and surface area (Figure 2). The MRT analysis created the first split based on whether degree-days were relatively high ($\geq$3,916 degree-days) or low (<3,916 degree-days). In lakes having more degree-days, species assemblage differences occurred at a surface area threshold of 177 ha, thus forming lake classes 1 and 2 (Figure 2). Bluegills, largemouth bass, and other warmwater species were significantly associated with lakes smaller than 177 ha (Table 3). The yellow perch was the only coolwater species that was common in lake class 1, but this species tended to occur in relatively low abundance. Lakes having more degree-days and a surface area of 177 ha or greater were also characterized by a predominance of warmwater species, but indicator species tended to be larger-bodied fishes (e.g., bowfin, longnose gar, and channel catfish) and smaller-bodied fishes (e.g., brook silverside, sand shiner, and logperch; Table 3). Yellow perch and larger-bodied coolwater species (e.g., smallmouth bass and walleye) were common in lake class 2 but tended to occur in relatively low abundance. Lakes having fewer than 3,916 degree-days were split based on a mean surface temperature of 14.7°C (Figure 2). Fish assemblages in lakes with temperatures less than 14.7°C were dominated by coolwater centrarchids (smallmouth bass and rock bass) and percids (walleye and yellow perch) and by coldwater species. Warmwater species (e.g., largemouth bass and bluegill) rarely occurred in lakes belonging to lake class 3. Lakes with a mean surface temperature of 14.7°C or higher were characterized by a mixture of coolwater and warmwater species and were subsequently split based on a surface area of 53 ha (Figure 2). Although no significant indicator species were found in lakes smaller than 53 ha, warmwater species (e.g., largemouth bass and bluegill) and coolwater species (e.g., yellow perch and white sucker) were typically associated with lake class 4 (Table 3). Lakes with a surface area of 53 ha or greater were split based on degree-days, thus forming lake classes 5 and 6 (Figure 2). Fish assemblages in lakes having 3,486 degree-days or more were dominated by a group of species that were tolerant of low-oxygen conditions, including the brown bullhead, yellow bullhead, bluntnose minnow, northern pike, and yellow perch. A mix of coolwater and warmwater-tolerant centrarchids as well as white suckers, golden shiners, and walleyes were also associated with lakes in class 5. White suckers were significantly associated with lakes having less than 3,486 degree-days, and other coolwater species (e.g., walleye, yellow perch, and northern pike) were typically found in lake class 6 (Table 3). Warmwater species (e.g., bluegill and largemouth bass) were common in lakes belonging to class 6, but those species occurred in relatively low abundance.

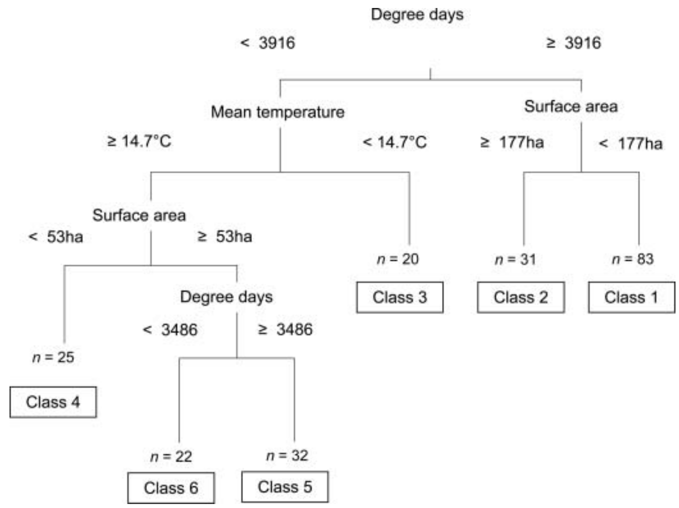

FIGURE 2. Multivariate regression tree output, showing the six lake classes based on fish assemblages and abiotic thresholds (n = number of lakes in each class).

The final MRT explained 25% of the variation in species assemblage structure. A comparison of the MRT solution with

TABLE 2. Percent occurrence of fish species captured by individual gear types (and for all gear types combined) in the 216 Michigan study lakes. Only fish species occurring in at least 5% of lakes are shown; see Methods for a list of species that were included in the coldwater species metric.

		Gear type			
Species	Code	Electrofishing	Gill net	Trap net	Combined
Lepisosteidae					
Spotted gar *Lepisosteus oculatus*	SPG	1.9	0.9	5.6	5.6
Longnose gar *Lepisosteus osseus*	LNG	10.2	5.1	16.7	19.9
Amiidae					
Bowfin *Amia calva*	BOW	19.9	6.0	38.9	38.9
Cyprinidae					
Spotfin shiner *Cyprinella spiloptera*	SFS	5.1		0.5	5.6
Common carp *Cyprinus carpio*	CAR	15.7	7.9	25.0	28.7
Common shiner *Luxilus cornutus*	CSH	15.7		2.3	17.6
Golden shiner *Notemigonus crysoleucas*	GOS	30.6	9.3	25.5	43.5
Blackchin shiner *Notropis heterodon*	BCS	6.0		1.4	6.5
Blacknose shiner *Notropis heterolepis*	BNS	11.1		0.5	11.6
Spottail shiner *Notropis hudsonius*	STS	17.1		2.3	17.1
Sand shiner *Notropis stramineus*	SAS	14.4		0.9	15.3
Mimic shiner *Notropis volucellus*	MIS	7.9		0.9	8.3
Bluntnose minnow *Pimephales notatus*	BNM	47.2		3.7	48.6
Creek chub *Semotilus atromaculatus*	CRC	2.8		2.8	5.6
Catostomidae					
White sucker *Catostomus commersonii*	CWS	39.4	53.2	60.2	69.4
Lake chubsucker *Erimyzon sucetta*	LCS	7.4	0.9	6.5	9.7
Golden redhorse *Moxostoma erythrurum*	GOR	4.2	3.2	3.2	6.5
Shorthead redhorse *Moxostoma macrolepidotum*	SHR	2.3	2.8	5.1	5.1
Ictaluridae					
Black bullhead *Ameiurus melas*	BLB	6.9	6.0	22.2	25.0
Yellow bullhead *Ameiurus natalis*	YLB	30.6	20.8	50.0	53.2
Brown bullhead *Ameiurus nebulosus*	BRB	38.4	21.3	69.0	71.3
Channel catfish *Ictalurus punctatus*	CCF	2.3	6.5	10.2	12.0
Esocidae					
Northern pike *Esox lucius*	NOP	37.0	81.0	81.0	87.5
Muskellunge *Esox masquinongy*	MUS	2.3	3.2	7.9	8.8
Redfin pickerel *Esox americanus*	GRP	13.9		1.9	14.4
Umbridae					
Central mudminnow *Umbra limi*	MUD	17.1			17.1
Fundulidae					
Banded killifish *Fundulus diaphanus*	BKF	5.6			5.6
Atherinopsidae					
Brook silverside *Labidesthes sicculus*	BSS	19.0			19.0
Centrarchidae					
Rock bass *Ambloplites rupestris*	RKB	67.1	45.4	78.2	78.7
Green sunfish *Lepomis cyanellus*	GSF	15.3	2.3	15.3	25.5
Pumpkinseed *Lepomis gibbosus*	PSF	82.4	25.9	90.3	92.1
Warmouth *Lepomis gulosus*	WAR	21.8	13.9	23.6	26.4
Bluegill *Lepomis macrochirus*	BLG	82.4	50.9	87.0	88.9
Redear sunfish *Lepomis microlophus*	RSF	7.4	2.3	10.6	10.6
Smallmouth bass *Micropterus dolomieu*	SMB	44.9	20.4	47.7	52.8
Largemouth bass *Micropterus salmoides*	LMB	84.7	49.5	81.5	87.5
Black crappie *Pomoxis nigromaculatus*	BCR	34.7	42.1	71.8	73.1
Percidae					
Iowa darter *Etheostoma exile*	IOD	15.7		0.9	15.7
Johnny darter *Etheostoma nigrum*	JOD	12.5		0.5	12.5
Yellow perch *Perca flavescens*	YEP	94.4	78.2	66.2	98.6
Logperch *Percina caprodes*	LOG	20.4		0.5	20.8
Walleye *Sander vitreus*	WAE	38.4	45.8	57.4	63.0
Coldwater species	CW	4.2	20.8	13.0	26.9

TABLE 3. Characteristic fish species of the six lake classes identified in multivariate regression tree analysis (species codes are defined in Table 2). Assemblage membership was based on above-average indicator scores within each lake class. Species are listed in order of prominence in each lake class; species listed in the same row had identical indicator species scores. Species with significant indicator scores are shown in bold italics.

Lake class					
1	2	3	4	5	6
WAR	***BOW***	***SMB***	LMB	***BRB***, YLB	***CWS***
YLB	***LNG***	***WAE***	YEP	***BNM***	WAE
BLG	***CAR***	***RKB***	BLG	NOP, PSF	YEP
LMB	***CCF***	CWS	PSF	RKB	NOP
BCR	BLG, LMB	***YEP***	CWS	LMB, YEP	PSF
PSF	BCR, ***BSS***, RKB	***CW***	CW	BLG	BCR, BRB, CSH, RKB
BOW	***SAS***, ***LOG***	***CSH***, NOP	NOP, SMB	CWS	SMB, GOS
NOP	BRB, PSF, YLB	PSF	GSF	BCR	BLG
BSS, CAR, ***GRP***, GSF, ***SPG***	NOP	STS	GOS	GOS, WAE	BNM
BRB	SMB, YEP	BCR	BNM, RKB	SMB	LMB, LOG
LCS	WAE		BRB, MUD, WAE		
YEP					

the solutions derived from unconstrained clustering showed that unconstrained clustering accounted for an average of 16% more of the variation in assemblage structure across the two to six clusters evaluated. These results indicate the existence of additional variation in assemblage structure that is not accounted for by the environmental variables considered in our study.

Lake classes (Figure 3) did not differ significantly in summer total phosphorus concentration (Kruskal–Wallis test: $\chi^2 = 7.6$, $P = 0.18$) but differed in mean depth ($\chi^2 = 24.3$, $P < 0.01$), surface area ($\chi^2 = 135.7$, $P < 0.01$), mean temperature ($\chi^2 = 113.9$, $P < 0.01$), and degree-days ($\chi^2 = 171.1$, $P < 0.01$). Lakes in classes 2 and 3 tended to be larger and deeper than lakes in other classes. Lakes belonging to lake classes 1 and 4 tended to be relatively small, and lakes in classes 5 and 6 were intermediate in surface area. Despite this pattern in surface area, lakes in class 1 had a relatively high mean depth and lakes in class 6 had the lowest mean depth of any category. Mean depth was similar between lake classes 4 and 5. Mean temperature was highest in class 1 lakes, lowest in class 3 lakes, and similar among the remaining lake classes. Degree-days in lake classes 1 and 2 were much higher than those in other lake classes, thereby reflecting the first split in the regression tree. Degree-days in lake class 5 were intermediate, and degree-days were relatively similar among the remaining lake types.

Overall predictive ability of the MRT model was 68% based on leave-out-one cross validation. Predicted membership for all 4-ha and larger lakes revealed variation in the numbers, surface area, and distribution across Michigan (Table 4; Figure 4). In terms of numbers, Michigan was dominated by class 1 lakes (59%). These lakes were primarily distributed across the southern portion of the Lower Peninsula and in the coastal areas of northern Michigan, where the ice-free period is relatively long because of the moderating effects of the Great Lakes. Because class 1 lakes were relatively small (Figure 3), they only represented one-quarter of the total lake surface area in Michigan (i.e., calculated based on the 6,544 inland lakes with surface area $\geq$4 ha). The other abundant lake type was class 4 (33%), which was found across Michigan's Upper Peninsula and the northern portion of the Lower Peninsula. Because lakes in class 4 were also relatively small (Figure 3), they represented less than 10% of the lake surface area in the state. Lakes in classes 2 and 3 were relatively less abundant, but because of the large size of these lakes (Figure 3), each class comprised 22% of the state's lake surface area. Class 2 lakes were distributed across the Lower Peninsula, while class 3 lakes primarily were found in northern Michigan and were concentrated in the western portion of the Upper Peninsula. Lakes in classes 5 and 6 were also relatively rare; because of their smaller sizes, these lake classes represented 14% and 8%, respectively, of the lake surface area in Michigan. Lakes belonging to classes 5 and 6 were found in the northern Lower Peninsula and across the Upper Peninsula.

DISCUSSION

We identified six fish assemblage types that can be used to classify Michigan lakes. Our study represents the most comprehensive analysis of fish assemblages for north temperate lakes in North America. Our study lakes included a wide range of habitat types and are distributed across a large geographic region. In contrast, most of the previous studies that have classified fish assemblages in north temperate lakes were focused on a relatively low number of lakes that were primarily small, had relatively few species, and in some cases contained no large-bodied piscivorous species (e.g., northern pike or black basses; Harvey 1978, 1981; Tonn and Magnuson 1982; Tonn et al. 1983; Robinson and Tonn 1989). As a result, a common assemblage identified in these prior studies consists of small-bodied species,

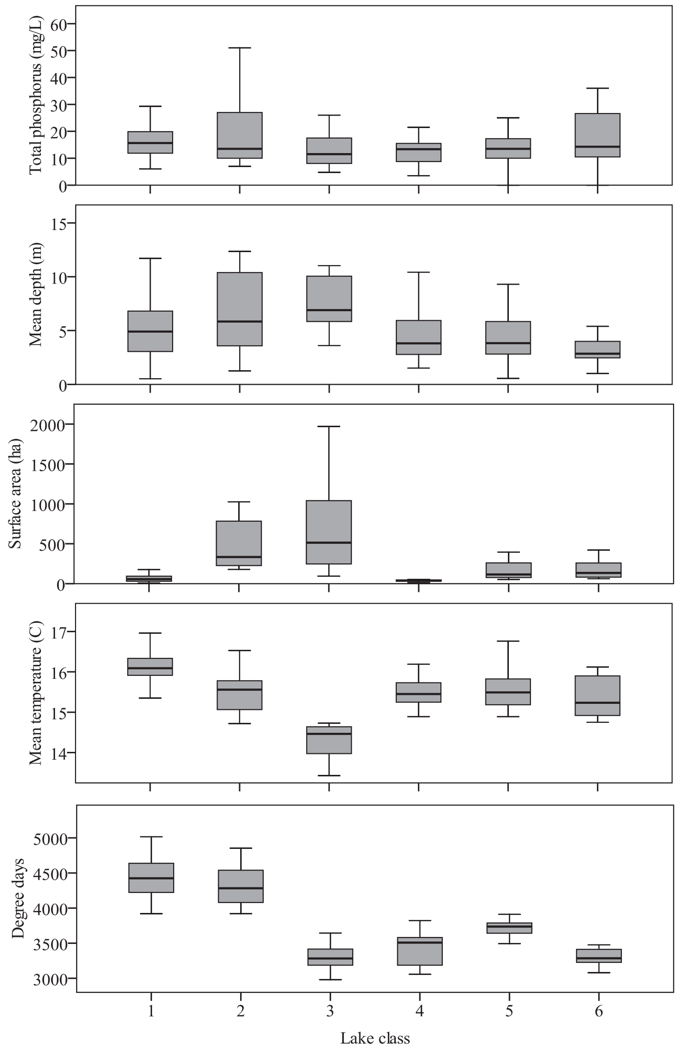

FIGURE 3. Box-and-whisker plots of nutrient, morphological, and thermal characteristics of the six lake classes identified in multivariate regression tree analysis (horizontal line within each box = median; ends of box = 25th and 75th percentiles; ends of whiskers = minimum and maximum values).

TABLE 4. Distribution of 6,544 Michigan lakes (with surface area ≥4 ha) across lake classes summarized by number and surface area of lakes.

	Number		Surface area (ha)	
Lake class	Total	Percent	Total	Percent
1	3,849	59	200,216	25
2	128	2	177,008	22
3	88	1	181,164	22
4	2,165	33	75,582	9
5	187	3	114,861	14
6	127	2	60,009	8

such as cyprinids, that are susceptible to predation but dominate lakes from which piscivores are absent. The lake types found in small north temperate lakes are poorly represented in our classification. However, our goal was to develop a classification for use in fisheries management. Most of the fisheries management effort is placed on 4-ha and larger lakes because smaller lakes have a limited ability to support exploitable populations of piscivores, such as walleyes, northern pike, largemouth bass, and smallmouth bass. Our classification complements previous classifications and together with these existing classifications represents a more complete picture of lake fish assemblages in north temperate lakes of North America.

Our results indicate that the variation in climate across Michigan plays an important role in controlling distribution and abundance of lake fishes. Mehner et al. (2007) found that latitude was related to fish assemblage structure in European lakes within the Central Plains ecoregion. They hypothesized that latitude was important in their analysis because it reflected a temperature gradient created by the climatic differences (mean annual air temperature = 6.2–9.7°C) across their study region. In Michigan, mean annual air temperature varies from 2°C in the Upper Peninsula to 11°C in the Lower Peninsula, which results in considerable among-lake variation in degree-days during the ice-free period. Not surprisingly, degree-days had the most influential effect in separating fish assemblages of southern and northern lakes. Although warmwater and coolwater species are found in lakes throughout the state, warmwater species (e.g., bluegill, black crappie, and largemouth bass) tend to dominate the southern lakes, whereas coolwater species (e.g., yellow perch, walleye, rock bass, white sucker, and northern pike) typically occur in higher abundance in the northern lakes. These patterns probably arise because many northern Michigan lakes have short growing seasons and long winters that limit growth and survival of warmwater species (Post et al. 1998), whereas many southern Michigan lakes have extended periods in summer during which the thermal optima of coolwater species are exceeded, thus limiting growth and survival, and also have short winters, which limit the egg development of coolwater species (Hokanson 1977).

Mean temperature in the epilimnion was important for separating assemblage types in our lake classification. This pattern represents a more local influence of water temperature in comparison with the regional influence of climate. In northern Michigan, large lakes with relatively cool surface temperatures (class 3) appear to be less suitable for warmwater fishes, such as the bluegill, black crappie, and largemouth bass; instead, these lakes are dominated by coolwater species, such as the white sucker, rock bass, smallmouth bass, walleye, and yellow perch. Coldwater species are also abundant in class 3 lakes, probably because of the greater volume of coldwater habitat present in these large, deep lakes. Lakes in class 3 tend to have greater surface areas and therefore greater depths, greater volumes, and smaller surface area : volume ratios than do lakes in other classes (Figure 3). Consequently, large, deep lakes have greater thermal

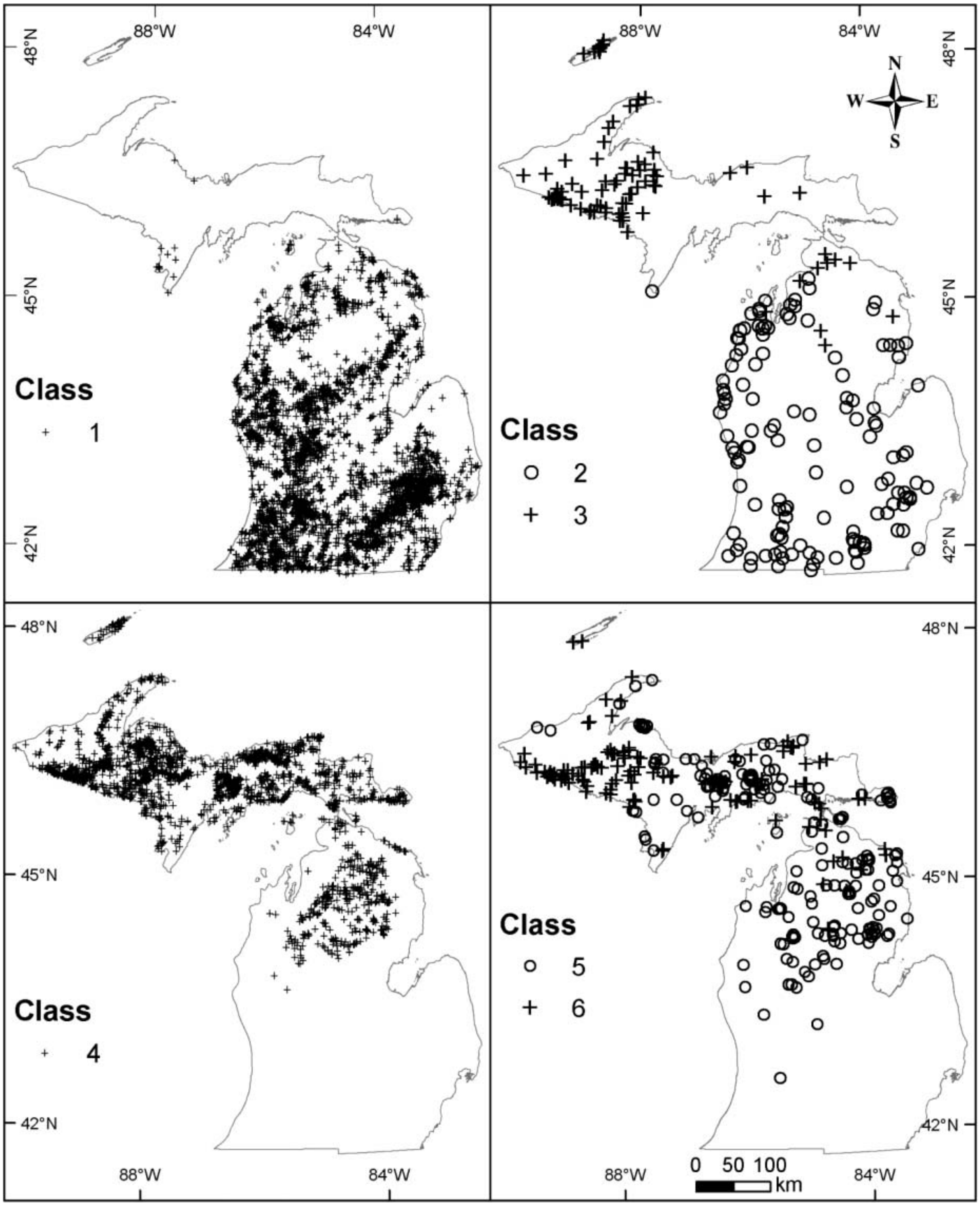

FIGURE 4. All 4-ha and larger lakes in Michigan ($N = 6{,}544$ lakes), classified by potential fish assemblage. See Figure 2 for definition of the six lake classes.

inertia, are more buffered against direct solar insolation, and do not get as warm as smaller, shallower lakes.

We also found that surface area was an important predictor of lake fish assemblages. Surface area is correlated with many aspects of lake morphology, and the importance of lake area to fish assemblage structure has been well documented (Johnson et al. 1977; Harvey 1978; Tonn and Magnuson 1982; Marshall and Ryan 1987; Jackson and Harvey 1993; Magnuson et al. 1998; Diekmann et al. 2005; van Zyll de Jong et al. 2005; Bertolo and Magnan 2006; Mehner et al. 2007). The diversity of depths,

TABLE 5. Description of temperature characteristics, morphology, fish species, and distribution of the six classes of Michigan inland lakes.

Class	Description
1	High degree-days, high mean temperature, small surface area, and intermediate depth; indicator species are warmwater centrarchids, yellow bullhead, grass pickerel, spotted gar, and lake chubsucker; these lakes are found primarily in the Lower Peninsula.
2	High degree-days, high mean temperature, large surface area, and deep; indicator species are the bowfin, longnose gar, common carp, channel catfish, brook silverside, sand shiner, and logperch; these lakes are found primarily in the Lower Peninsula.
3	Low degree-days, low mean temperature, large surface area, and deep; indicator species are coolwater centrarchids, percids, coldwater species, and the common shiner; these lakes are concentrated in the western Upper Peninsula, with limited distribution in the northern Lower Peninsula.
4	Low degree-days, intermediate mean temperature, small surface area, and intermediate depth; associated species include a mixture of warmwater centrarchids and coolwater species, such as the yellow perch and white sucker; these lakes are very common in the Upper Peninsula and northern Lower Peninsula.
5	Intermediate degree-days, intermediate mean temperature, intermediate surface area, and intermediate depth; indicator species are the brown bullhead, bluntnose minnow, and other species tolerant of low oxygen concentrations; these lakes are found in the Upper Peninsula and northern Lower Peninsula.
6	Low degree-days, intermediate mean temperature, intermediate surface area, and shallow; the indicator species is the white sucker, but percids and northern pike are also common; these lakes are found primarily in the Upper Peninsula.

temperatures, and littoral habitats typically found in larger lakes may promote the existence of species that cannot persist in smaller, more homogeneous lakes (Jackson et al. 2001). For example, larger lakes tend to be deeper and are more likely to stratify, thereby providing coolwater habitat (Figure 3) for species such as the walleye. Similar to Jackson and Harvey (1993), we found a higher abundance of warmwater species (e.g., bluegill and largemouth bass) in smaller lakes (classes 1 and 4) and a higher abundance of coolwater fishes in larger lakes (classes 2, 5, and 6). This pattern suggests that a difference in availability of coolwater habitat between large and small lakes is the mechanism linking surface area to fish assemblage patterns. Lake size can also mediate the role of anoxia in structuring fish assemblages (Tonn and Magnuson 1982). Smaller lakes are more susceptible to oxygen depletion and can become unsuitable for species that are intolerant of low-oxygen conditions (Magnuson et al. 1998; Stefan et al. 2001). Bluegills and largemouth bass are intolerant of low-oxygen conditions (Jackson and Harvey 1993), and their prominence in smaller lakes (classes 1 and 4) suggests that winterkill was not a dominant driver of assemblage structure across the range of lake sizes considered in this study.

The amount of variation explained by our MRT solution (25%) and the additional variation (16%) explained by unconstrained clustering suggest that other factors are also important in structuring lake fish assemblages. A number of factors, including species interactions (Robinson and Tonn 1989; Olive et al. 2005), extinction–colonization dynamics (Magnuson et al. 1998), isolation (Olden et al. 2001), productivity (Jeppesen et al. 2000), anoxia (Stefan et al. 2001), and shoreline development (Jennings et al. 1999), can influence lake fish assemblages. Additional research is needed to determine the relative influence of these factors on fish assemblages in Michigan lakes. However, the importance of these factors may only be evident when lakes are viewed across smaller spatial scales. Given the hierarchical nature of aquatic ecosystems (Tonn 1990) and the pervasive role of temperature in determining habitat suitability, we suggest that the evaluation of additional mechanisms should take into account the lake classes we have identified.

In contrast to most lake classifications, our classification was based on spatially extensive variables that were readily available or that could be predicted from landscape-level data. Our approach can thus be used to predict species assemblages in unsampled lakes and thereby classify all lakes across large spatial extents. By classifying all lakes in the state, we were able to describe the number and spatial location of lakes in each class and thereby provide a more complete characterization of Michigan's fishery resources. Fisheries management typically occurs on a species-by-species and lake-by-lake basis (Tonn et al. 1983). Our classification provides information about entire fish assemblages and can facilitate development of statewide management and monitoring programs based on lake types and their availability. For example, management of walleyes and coldwater species could be focused on lakes in class 3, thereby reducing the likelihood of inappropriate stockings in less-suitable lake classes. Using the classification, managers can make defensible decisions and communicate to the public the ecological reasons that underlie management plans. Monitoring efforts could be allocated equally across the six lake classes or could be allocated in proportion to the number or surface area of lakes within each class. Because lake fish assemblages are strongly influenced by lake thermal regime, our approach could be useful in suggesting potential consequences of climate change (Stefan et al. 2001).

Finally, when linked with other types of data, our classification can be used by managers to determine the value of specific lake types and the environmental threats (Wang et al. 2010) to those lake types. This information can be used to prioritize management efforts across all available lake fisheries in the state by identifying the highest-value lake classes that are most at risk from environmental change.

Our classification could be applied to other north temperate lakes, especially those in Wisconsin, Minnesota, and Ontario, where the regional fish fauna is similar. A fruitful area of research would be to investigate whether fish assemblages in other north temperate lakes follow patterns similar to those observed in Michigan. If successful, the application of our lake classification across multiple states and provinces would provide a framework for conducting broad-scale comparative studies and for developing regional conservation and management strategies.

ACKNOWLEDGMENTS

This project was partially supported by Federal Aid in Sport Fish Restoration Program Project F-80-R-10 through the MDNR Fisheries Division. This study would not have been possible without the effort and dedication of the MDNR Fisheries Division employees responsible for monitoring inland lake resources across Michigan. We thank Paul Seelbach for discussions on ecological classification and three anonymous reviewers for their helpful comments.

REFERENCES

Bertolo, A., and P. Magnan. 2006. Spatial and environmental correlates of fish community structure in Canadian Shield lakes. Canadian Journal of Fisheries and Aquatic Sciences 63:2780–2792.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. G. Stone. 1984. Classification and regression trees. Wadsworth, Belmont, California.

Brenden, T. O., L. Wang, and P. W. Seelbach. 2008. A river valley segment classification of Michigan streams based on fish and physical attributes. Transactions of the American Fisheries Society 137:1621–1636.

De'ath, G. 2002. Multivariate regression trees: a new technique for modeling species-environment relationships. Ecology 83:1105–1117.

De'ath, G., and K. E. Fabricus. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81:3178–3192.

Diekmann, M., U. Brämick, R. Lemcke, and T. Mehner. 2005. Habitat-specific fishing revealed distinct indicator species in German lowland lake fish communities. Journal of Applied Ecology 42:901–909.

Dolman, W. B. 1990. Classification of Texas reservoirs in relation to limnology and fish community associations. Transactions of the American Fisheries Society 119:511–520.

Dufrêne, M., and P. Legendre. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecological Monographs 67:345–366.

Efron, B., and G. Gong. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. American Statistician 37:36–48.

Emmons, E. E., M. J. Jennings, and C. Edwards. 1999. An alternative classification method for northern Wisconsin lakes. Canadian Journal of Fisheries and Aquatic Sciences 56:661–669.

ESRI (Environmental Systems Research Institute). 2002. PC ARC/GIS version 8.2. ESRI, Redlands, California.

Gassner, H., J. Wanzenböck, D. Zick, G. Tischler, and B. Pamminger-Lahnsteiner. 2005. Development of a fish based lake typology for natural Austrian lakes >50 ha based on the reconstructed historical fish communities. International Review of Hydrobiology 90:422–432.

Håkanson, L. 1996. Predicting important lake habitat variables from maps using modern modelling tools. Canadian Journal of Fisheries and Aquatic Sciences 53(Supplement 1):364–382.

Harvey, H. H. 1978. Fish communities of the Manitoulin Island lakes. Internationale Vereinigung für theoretische und angewandte Limnologie Verhandlungen 20:2031–2038.

Harvey, H. H. 1981. Fish communities of the lakes of the Bruce Peninsula. Internationale Vereinigung für theoretische und angewandte Limnologie Verhandlungen 21:1222–1230.

Hastie, T., R. Tibshirani, and J. H. Friedman. 2001. The elements of statistical learning: data mining, inference, and prediction. Springer-Verlag, New York.

Hinch, S. G., N. C. Collins, and H. H. Harvey. 1991. Relative abundance of littoral zone fishes: biotic interactions, abiotic factors, and postglacial colonization. Ecology 72:1314–1324.

Hokanson, K. E. F. 1977. Temperature requirements of some percids and adaptations to the seasonal temperature cycle. Journal of the Fisheries Research Board of Canada 34:1524–1550.

Jackson, D. A., and H. H. Harvey. 1989. Biogeographic associations in fish assemblages: local vs. regional processes. Ecology 70:1472–1484.

Jackson, D. A., and H. H. Harvey. 1993. Fish and benthic invertebrates: community concordance and community-environment relationships. Canadian Journal of Fisheries and Aquatic Sciences 50:2641–2651.

Jackson, D. A., P. R. Peres-Neto, and J. D. Olden. 2001. What controls who is where in freshwater fish communities: the roles of biotic, abiotic, and spatial factors. Canadian Journal of Fisheries and Aquatic Sciences 58:158–170.

Jennings, M. J., M. A. Bozek, G. R. Hatzenbeler, E. E. Emmons, and M. D. Staggs. 1999. Cumulative effects of incremental shoreline habitat modification on fish assemblages in north temperate lakes. North American Journal of Fisheries Management 19:18–27.

Jeppesen, E., J. P. Jensen, M. Søndergaard, T. Lauridsen, and F. Landkildehus. 2000. Trophic structure, species richness, and biodiversity in Danish lakes: changes along a phosphorus gradient. Freshwater Biology 45:201–218.

Johnson, M. G., J. H. Leach, C. K. Minns, and C. H. Olver. 1977. Limnological characteristics of Ontario lakes in relation to associations of walleye (*Stizostedion vitreum*), northern pike (*Esox lucius*), lake trout (*Salvelinus namaycush*), and smallmouth bass (*Micropterus dolomieui*). Journal of the Fisheries Research Board of Canada 34:1592–1601.

Kalff, J. 2002. Limnology: inland water ecosystems. Prentice-Hall, Upper Saddle River, New Jersey.

Leach, J. H., and R. C. Herron. 1992. A review of lake habitat classification. Pages 27–73 *in* W. D. N. Busch and P. G. Sly, editors. The development of an aquatic habitat classification system for lakes. CRC Press, Boca Raton, Florida.

Legendre, P., and E. D. Gallagher. 2001. Ecologically meaningful transformations for ordination of species data. Oecologia 129:271–280.

Magnuson, J. J., W. M. Tonn, A. Banerjee, J. Toivonen, O. Sanchez, and M. Rask. 1998. Isolation vs. extinction in the assembly of fishes in small northern lakes. Ecology 79:2941–2956.

Marshall, T. R., and P. A. Ryan. 1987. Abundance patterns and community attributes of fishes relative to environmental gradients. Canadian Journal of Fisheries and Aquatic Sciences. 44(Supplement 2):198–215.

McCune, B., J. B. Grace, and D. L. Urban. 2002. Analysis of ecological communities. MjM Software Design, Gleneden Beach, Oregon.

McCune, B., and M. J. Mefford. 2006. PC-ORD^{TM}: multivariate analysis of ecological data, version 5.10. MjM Software Design, Gleneden Beach, Oregon.

Mehner, T., M. Diekmann, U. Bramick, and R. Lemcke. 2005. Composition of fish communities in German lakes as related to lake morphology, trophic state, shore structure and human-use intensity. Freshwater Biology 50: 70–85.

Mehner, T., K. Holmgren, T. L. Lauridsen, E. Jeppesen, and M. Diekmann. 2007. Lake depth and geographical position modify lake fish assemblages of the European “central plains” ecoregion. Freshwater Biology 52:2285–2297.

Olden, J. D., D. A. Jackson, and P. R. Peres-Neto. 2001. Spatial isolation and fish communities in drainage lakes. Oecologia 127:572–585.

Olden, J. D., D. A. Jackson, and P. R. Peres-Neto. 2002. Predictive models of fish species distributions: a note on proper validation and chance predictions. Transactions of the American Fisheries Society 131:329–336.

Olive, J. A., L. E. Miranda, and W. D. Hubbard. 2005. Centrarchid assemblages in Mississippi state-operated fishing lakes. North American Journal of Fisheries Management 25:7–15.

Post, D. M., J. F. Kitchell, and J. R. Hodgson. 1998. Interactions among adult demography, spawning data, growth rate, predation, overwinter mortality, and the recruitment of largemouth bass in a northern lake. Canadian Journal of Fisheries and Aquatic Sciences 55:2588–2600.

R Development Core Team. 2010. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Available: www.Rproject.org. (May 2011).

Robinson, C. L. K., and W. M. Tonn. 1989. Influence of environmental factors and piscivory in structuring fish assemblages of small Alberta lakes. Canadian Journal of Fisheries and Aquatic Sciences 46:81–98.

Schneider, J. C. 1981. Fish communities in warmwater lakes. Michigan Department of Natural Resources, Fisheries Research Report 1980, Ann Arbor.

Schupp, D. H. 1992. An ecological classification of Minnesota lakes with associated fish communities. Minnesota Department of Natural Resources, Investigational Report 417, St. Paul.

Shuter, B. J., D. A. Schlesinger, and A. P. Zimmerman. 1983. Empirical predictors of annual surface water temperature cycles in North American lakes. Canadian Journal of Fisheries and Aquatic Sciences 40:1838–1845.

Søndergaard, M., E. Jeppesen, J. P. Jensen, and S. L. Amsinck. 2005. Water framework directive: ecological classification of Danish lakes. Journal of Applied Ecology 42:616–629.

Stefan, H. G., X. Fang, and J. G. Eaton. 2001. Simulated fish habitat changes in North American lakes in response to climate warming. Transactions of the American Fisheries Society 130:459–477.

Strahler, A. N. 1957. Quantitative analysis of watershed geomorphology. Transactions, American Geophysical Union 38:913–920.

Tonn, W. M. 1990. Climate change and fish communities: a conceptual framework. Transactions of the American Fisheries Society 119: 337–352.

Tonn, W. M., and J. J. Magnuson. 1982. Patterns in the species composition and richness of fish assemblages in northern Wisconsin lakes. Ecology 63:1149–1166.

Tonn, W. M., J. J. Magnuson, and A. M. Forbes. 1983. Community analysis in fishery management: an application with northern Wisconsin lakes. Transactions of the American Fisheries Society 112:368–377.

van Zyll de Jong, M. C., I. G. Cowx, and D. A. Scruton. 2005. Association between biogeographical factors and boreal lake fish assemblages. Fisheries Management and Ecology 12:189–199.

Vehanen, T., and J. Aspi. 1996. Classification of northern Finnish lakes and the suitability for the stocking for brown trout (*Salmo trutta* (L.) m. *lacustris*). Fisheries Research 27:37–49.

Wang, L., K. E. Wehrly, J. E. Breck, and L. Kraft. 2010. Landscape based assessment of human disturbance for Michigan lakes. Environmental Management 46:471–483.

Wehrly, K. E., G. S. Carter, and J. E. Breck. 2012. Standardized sampling methods for the inland lakes status and trends program. Michigan Department of Natural Resources, Fisheries Special Report, Ann Arbor.

Wilson, M. A., and S. R. Carpenter. 1999. Economic valuation of freshwater ecosystem services in the United States: 1971–1997. Ecological Applications 9:772–783.