2
2010

This is to certify that the
thesis entitled

MEASURING ANTHROPOGENIC DISTURBANCES IN A
HYDROGEOMORPHIC-BASED LAKE CLASSIFICATION BUILT
USING FISH ASSEMBLAGES IN 360 NORTH-TEMPERATE LAKES

presented by

BRETT MICHAEL ALGER

has been accepted towards fulfillment
of the requirements for the

M.S degree in FISHERIES AND WILDLIFE

Mary Bremigan
Major Professor's Signature

16 December 2009

Date

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

MEASURING ANTHROPOGENIC DISTURBANCES IN A HYDROGEOMORPHIC-BASED LAKE CLASSIFICATION BUILT USING FISH ASSEMBLAGES IN 360 NORTH-TEMPERATE LAKES

By

Brett Michael Alger

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Fisheries and Wildlife

2009

ABSTRACT

MEASURING ANTHROPOGENIC DISTURBANCES IN A HYDROGEOMORPHIC-BASED LAKE CLASSIFICATION BUILT USING FISH ASSEMBLAGES IN 360 NORTH-TEMPERATE LAKES

By

Brett Michael Alger

Hydrogeomorphic features and anthropogenic disturbances affect fish assemblages but the relative importance of these variables is inadequately understood. Quantifying these linkages, across multiple spatial scales, is key to developing lake ecosystem assessment tools and management plans. Accordingly, I classified lakes with similar hydrogeomorphic features and within these lake groups, identified the most important anthropogenic influences on fish species richness. I used fish assemblage, hydrogeomorphic, and anthropogenic data from 360 lakes in Maine, New Hampshire, Iowa, Michigan, and Wisconsin. To build lake classes, I used classification and regression trees; lakes having similar fish species richness were best grouped by lake area, ecological drainage unit, and watershed area. Results indicated that in addition to positive associations of lake and watershed area with species richness, regionalization frameworks were useful tools in identifying unique areas of freshwater biodiversity. Within classes of naturally similar lakes, species richness generally had a negative association with fish stocking, dams, and urban and agricultural land cover/use. Species richness was positively associated with human population density and road density. The methodology of classifying lakes appears to be a tool that lake managers can utilize to better understand fish assemblages in inland lakes. Classifications were also useful in grouping naturally similar lakes to identify the effects of human disturbances.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Thesis Introduction

The natural landscape has shaped freshwater ecosystems across the Earth; glaciers, climate, and other natural processes have affected the distributions and abundances of biota in streams, inland lakes, and other aquatic ecosystems. Because freshwater ecosystems are so complex and influenced by many features, natural and human, at multiple temporal and spatial scales, lake managers and researchers have attempted to partition the many effects of the landscape on in-lake biota by classifying ecosystems. These efforts have resulted in a strong understanding of how land use/cover types (e.g., forest, wetlands), lake morphometry (e.g., lake area, lake depth), water quality (e.g., pH, total phosphorous), hydrology, and surface connectivity affect fish assemblages, yet often all of these factors have not often been included comprehensively in the same lake classification study (Johnson et al. 1977, Dolman 1990, Schupp 1992).

Regionalization frameworks have been created to identify area of unique biodiversity; they represent sets of natural features across a large region, with subregions distinguished by finer scale variation among features (Seelbach et. al. 2002). These schemes can help to explain variation among ecosystems at different spatial scales (Omernik 1987, McMahon et al. 2001, Irz et al. 2004, Cheruvelil et al. 2008). Subsequently, I attempt to compare the ability of these various metrics of the hydrogeomorphic landscape across scales and in combination to explain patterns of fish assemblages using a lake classification approach.

Classifications have been built as a way to comprehensively explain the variation in fish species composition across the landscape, resulting in relatively similar lakes within a class and higher variability among classes Iowa (Johnson et al. 1977, Dolman

1990, Schupp 1992). Since fish management agencies for most states are often financially limited and lack the ability to gather data for *every* lake, classifications are potentially an efficient tool for tackling the pitfalls of managing thousands of lakes with limited information. While each lake classification is best developed with distinctive goals in mind (and often unique predictor variables as a result) to explain any number of characteristics of fish assemblages; the entire method (that of developing classifications and using them in a management context) is an important concept to further investigate and upon which to build.

Similar to the natural landscape, human activities also impact aquatic systems; lake managers and researchers have attempted to use anthropogenic variables to explain variation within and among lakes as well (Brunberg and Blomqvist 2001, Pierce and Tomcko 2003). While some anthropogenic disturbances positively affect valued qualities of fish assemblages through species additions (stocking), many have negative effects, such as blockage of movement (e.g., dams), and landscape alterations (e.g., land use/cover) (Radomski and Goeman 1995, Carpenter et al. 2007, Johnson et al. 2008). Because management agencies are often financially limited, they are often resigned to weigh trade-offs between preventing human-induced species invasions while at the same time supplementing faltering fisheries through stocking and habitat alterations, all the while maintaining relatively stable and healthy ecosystems. Although individual lakes are unique in many ways, there are similarities of how human disturbances affect fish assemblages. Therefore, I hypothesized that after grouping naturally-similar lakes using landscape-based classification techniques, the effects of humans on fish assemblages can be quantified in a way that is relevant to management agencies that then can implement plans and account for humans across classes of lakes.

The research presented here examines how natural features and human disturbances affect fish species richness in inland lakes. I build upon previous lake classification methods to investigate the ability of regional-scale features in addition to previously well studied local-scale features to classify 360 lakes across Maine, New Hampshire, Michigan, Wisconsin, and Iowa (Johnson et al. 1977, Dolman 1990, Schupp 1992). In chapter 1, I create a hydrogeomorphic lake classification using natural landscape metrics from four major categories: regionalization frameworks, species extinction/isolation, lake hydrology, and land use/cover to predict patterns of total and native species richness (having statistically removed the effects of disturbance on the response metric). In chapter 2, I measure the effects of anthropogenic disturbances on native species richness within classes of lakes from chapter 1 using land use/cover, dam, road density, fish stocking, and human population density as predictor metrics. Because no other study has covered such a broad geographical region and large number of lakes and examined regional-scale features in conjunction with local features, my research will contribute to the science of landscape limnology and to the further development of methods for lake classification. Additionally, this research will provide managers with helpful information and a fruitful approach for developing classes of lakes that are naturally similar, and likewise, respond similarly to human disturbances.

Chapter 1-What natural lake and landscape features best classify fish species richness in 360 north-temperate lakes using CART analysis?

Introduction

Researchers and lake managers have long attempted to explain and understand variation of in-lake features among lakes, often using local landscape features and lake morphometry to predict different metrics of water quality (e.g., total phosphorous, pH) and fish assemblages (e.g., spatial distribution, species richness). Lake classification systems are one way to group lakes according to the variation of a response metric (e.g., species richness), using the most significant predictor variables for a set of lakes (Johnson et al. 1977, Dolman 1990, Schupp 1992). Such studies provide a basic approach to classifying fish assemblages based on local landscape, lake morphometry, and in-lake water quality features. Previous studies typically were conducted in a single state or Canadian province, and their focus was to explain variation in fish assemblages using local landscape and in-lake features (Rahel 1984, Dolman 1990, Schupp 1992, Jones et al.2004). They provide a method and context for investigating how to monitor and manage statewide fish populations through the use of predictive lake classifications. State and federal agencies are primarily responsible for managing inland lake fish assemblages; they are often financially limited and must weigh trade-offs between investing in assessment and implementing decisions from gathered information (Hansen and Jones 2008). Management agencies in lake-rich regions therefore must implement their programs on all lakes, while only being able to collect data, particularly of in-lake biota, from a subset of lakes. However, based on known lake-to-landscape relationships established using data from sampled lakes, managers can now extrapolate these results from sampled lakes to unsampled lakes (using readily-available landscape data for unsampled lakes) to help make decisions. To build a natural lake classification system based on hydrogeomorhic landscape features, it is important to understand and include those features already known to shape lake fish assemblages, while still accounting for human disturbance. Therefore, in this chapter, I created a series of lake classifications based on natural hydrogeomorphic features of the landscape (a combination of natural hydrologic, geologic, and topographic factors) that range in spatial scale to explain the variation in fish species richness present among lakes.

We can learn a lot from previous research of the relationship between the natural landscape and fish assemblages conducted at the local scale (i.e., lakes within a single state). Studies have extensively measured the landscape-to-fish linkages in terms of lake surface area (Johnson et al. 1977, Tonn and Magnuson 1982, Rahel 1984, Eadie et al. 1986, Jackson and Harvey 1989, Holmgren and Appleberg 2000, Riera et al. 2000, Irz et al. 2004), lake perimeter (Olden et al. 2001, Olden and Jackson 2002, Hershey et al. 2005, Olden et al. 2006), lake depth (Marshall and Ryan 1987, Mehner et al. 2005), surface water connectivity as an isolation metric and as a proxy for landscape position (Kratz et al. 1997, Magnuson et al. 1998, Riera et al. 2000, Olden et al. 2001), and catchment area (Allan and Johnson 1997, Tonn et al. 2003, Irz et al. 2004). These studies show positive relationships between lake surface area, depth, and/or perimeter and fish species richness; likewise, lakes lower in the landscape with more surface connections have higher species richness. Conversely, lakes with relatively low habitat complexity (e.g., low values for lake area, depth or perimeter) or isolated lakes (e.g., higher in the landscape, with no surface connections) typically have relatively low species richness. By identifying the linkages between the landscape and species richness through past literature, it is possible to create a lake classification that is interpretable and applicable to lake managers.

Past lake classification systems were created for unique purposes in single states or provinces using a limited number of natural landscape, in-lake water quality, and human disturbance predictors to group inland lake fish assemblages (Johnson et al. 1977, Tonn et al. 1983, Dolman 1990, Schupp 1992). The term "fish assemblage" has included a variety of response metrics: presence or absence of individual or key fish species (Johnson et al. 1977), relative species abundance (Marshall and Johnson 1987), associations of dominant species (Dolman 1990), or species richness (SR) and composition (Eadie et al. 1986, Kratz et al. 1997). This previous research has incorporated many different lake morphometry characteristics to group inland lake fish assemblages (Tonn et al. 1983), water quality data such as pH and temperature (Dolman 1990), physical substrate (Johnson et al. 1977), or local human disturbances such as distance to nearest road (Magnuson et al. 1998) to explain among-lake differences in fish assemblage metrics. Lake ecosystems are complex and there are many landscape factors that shape fish assemblages at the local level; however, regional-scale predictors (e.g., spatial patterns of past glaciation, regional climate characteristics) may play a role as well. While unique in their purpose and spatial scale (across a single state), many past attempts at classifying lakes have disregarded regional-scale factors when attempting to better understand the distribution and variation of fish assemblages across the landscape.

Past studies have laid a strong foundation for developing lake classifications and determining what local features shape inland lake fish assemblages; however, research suggests incorporating landscape features at a much larger spatial scale can aid in explaining variation among lakes (Tonn 1990, Poff 1997, Cheruvelil et al. 2008). Although individual lakes have somewhat unique physical, chemical, and biological characteristics, there also are similarities across a state and even across broader regions, exemplified by the wide-ranging distribution of many fish species. Therefore, in developing an approach to classify lakes across large lake-rich areas for a variety of purpose, it is important to incorporate hydrogeomorphic landscape predictors at multiple spatial scales. When working at such spatial scales, lakes can be viewed as patches within a landscape that are hierarchically organized, with regional factors such as climate and soil type constraining local and in-lake impacts on lake biota (Tonn 1990, Poff 1997, Cheruvelil et al. 2008, Soranno et al. 2009).

One way to incorporate a broader-scale landscape perspective into a classification framework is to include a regionalization scheme as a classifying variable. Such schemes have been developed for both terrestrial and aquatic systems, and they can serve as strong proxies for natural features (e.g., soil types, sub-surface geology, land use; Omernik 1987). Most regionalizations represent a set of natural features across a large region, with subregions distinguished by finer scale variation among features (Seelbach et. al. 2002). These schemes can help to explain variation among ecosystems at different spatial scales (Omernik 1987, McMahon et al. 2001, Irz et al. 2004, Cheruvelil et al. 2008), and thus should be included as a potential classification variable to group lake fish assemblages.

I included in my analysis three regionalization frameworks that have been shown to help to explain variation among aquatic systems: United States Geological Survey (USGS) hydrologic units, (HUC, Seaber et al. 1987), ecological drainage units (EDU, Higgins et al. 2005), and freshwater ecoregions (Abell et al. 2000). HUC's were developed by the USGS to approximate watersheds at multiple scales across the country and provide a standardized method of locating, storing, retrieving, and exchanging hydrologic data for water-resources organizations (Seaber et al. 1987). HUC's are hierarchical in spatial scale; six-digit HUC's are further divided by eight-digit HUC's with 12-digit HUC's being an even finer spatial scale. EDU's were developed by agglomerating eight-digit HUC's using landscape features, such as climate and landform, to form contiguous, geographically dependent regions. These regions were developed to classify freshwater ecosystems to represent coarse-scale regions of freshwater biodiversity for regional conservation planning (Higgins et al. 2005). Freshwater ecoregions were created using native fish distributions (presence/absence data) to delineate sub-regions across North America; they represent likely freshwater aquatic biodiversity patterns for conservation purposes (Abell et al. 2000). All of these regionalization schemes have been delineated at different spatial scales; across the five-state region there are 97 six-digit HUC's, 28 EDU's, and five freshwater ecoregions, each allowing for an examination of the most appropriate spatial scale for grouping lake fish assemblages.

Because I was attempting to build classes based upon natural landscape features, it is imperative to recognize that human disturbances can also significantly alter lake fish assemblages at multiple spatial scales and mediate the effects of natural features. For example, land use/cover, the prevalence of roads and dams, and fish stocking rates all vary across the landscape and affect inland lake fish assemblages (Radomski and Goeman 1995, Magnuson et al. 1998, Roth et al. 2007). Human disturbances such as fish stocking or dam and reservoir creation can mask natural patterns by directly and/or indirectly altering fish assemblage patterns through species introductions and food web alterations (Radomski and Goeman 1995, Johnson et al. 2008). Because natural lake classifications are based on predictable relationships between natural landscape features and fish assemblages, human disturbances, which can mask natural patterns, must be considered and accounted for when attempting to explain natural variation in fish assemblages.

My study examines the variation of fish species richness across five states (Maine, New Hampshire, Michigan, Wisconsin, Iowa), ranging from the Atlantic Ocean to west of the Mississippi River. My work incorporates both local and regional-scale predictor metrics; the focus is to use the natural landscape across multiple states and spatial scales to identify which metrics explain the greatest amount of variation in lake fish species richness. Because previous classifications are unique to their own regions, hypotheses, and goals, my study is not necessarily intended to compare my classification with previous results, but rather to examine the approach of using regional-scale predictors along with local landscape features to explain variation in fish species richness among lakes.

In this chapter, I developed lake classifications to explain patterns of fish species richness based on natural hydrogeomorphic features of the landscape (combination of natural hydrologic, geologic, and topographic factors) that range in spatial scale. My study provides a good opportunity to compare local and regional patterns of fish assemblages, unlike past studies that only focused on a single state. Species richness, for this study, was judged to be a more reliable response metric than others (e.g., fish abundance) because fish assemblage data was collected from multiple agencies each with their own methods and purposes. Each state had a unique set of species to examine, so species counts were determined to be a more universal and comparable response metric than assemblage type. I also advanced the methodology behind building natural lake classification systems by taking advantage of advances in geographical information systems (GIS) and incorporating regional-scale predictors with previously well-studied local landscape factors and by accounting for human disturbances. I restricted the potential predictor variables to those that are map-based and hence available for all lakes. Although in-lake water quality can predict fish assemblages (Bachmann et al. 1996, Jeppesen et al. 2000), I did not include water quality metrics as predictors. This way, it is possible to extrapolate my natural lake classification results that are based on map and GIS data to 'new,' unsampled lakes. I also focused on the ability of natural hydrogeomorphic features (as opposed to human-mediated landscape features) to build lake classes because management and conservation efforts often need to know what the 'natural' state of a lake would be in the absence of anthropogenic factors. Natural features such as lake depth or lake area are relatively unchanging as compared to human disturbances; therefore, the relationship between the landscape and fish species richness may be more predictable when including only natural features.

Methods

Data Collection and Lake Selection

I collected data on hydrogeomorphic features (e.g., regionalization schemes, lake hydrology, lake morphometry) and anthropogenic disturbances (e.g., land use, road density, human population) for approximately 2300 inland lakes as part of a larger study (e.g., Cheruvelil et al. 2008, Webster et al. 2008) to determine how the natural and human landscapes shape lake water quality and fish assemblages (Tables 1-3). The initial data collection focused primarily on lakes with water quality response metrics that are typically easier to collect and more readily available than fish assemblage information. Available fish assemblage data were then obtained for New Hampshire, Maine, Michigan, Wisconsin, and Iowa (Figure1, Table 1). Because water quality and fish assemblage monitoring is often conducted by different agencies using different methods for lake selection, not every lake that existed in the original hydrogeomorphic (2,348 lakes) and landscape (2,573 lakes) databases existed in the fish assemblage database (555 lakes). Therefore, I created an overlap database that contained 360 lakes with all three data categories (hydrogeomorphic, anthropogenic, and fish assemblage; Table 1). For these 360 lakes, I obtained additional hydrogeomorphic (e.g., lake isolation metrics, number of surface connections) and anthropogenic disturbance (e.g., dam density, fish stocking metrics) data from agency sources and the use of Geographic Information System (GIS; Table 1).

Fish Assemblage Response Metrics

I gathered fish assemblage data from state and federal agencies and universities (Table 4); each entity's goal was to quantify presence/absence of fish species or estimate abundance based upon sampling effort to effectively monitor fishery resources. For lakes represented multiple times in the fish assemblage data, I used the most recent year sampled. Most fish had been identified to the species level. However, some fish were identified only to an aggregate level of genus or family, or were of unknown identity (7%-24% of fish across ME, NH, MI, WI, and IA). For example, some states reported fish as "unknown bullhead" or "minnow sp". Depending on the response metric (see below for response metric details), these aggregate terms were either included or removed based upon lake-specific fish assemblages and by using the following decision rules. If a species was present in a lake, higher taxonomic aggregates and unknown fishes related to that species were eliminated. However, the aggregate term was included in the species richness count if no other species representing that aggregate term was found in the lake. For example, if the aggregate "unknown bullhead" was found in a lake but no specific bullhead species (yellow, black, or brown) was recorded, I counted the aggregate as representing one bullhead species. On the other hand, if a specific bullhead species was captured in a lake in addition to an aggregate being recorded, the aggregate was eliminated from the total species count.

I created two metrics to use as response metrics in our natural lake classes: Total Species Richness (Total SR) and Native Species Richness (Native SR) (Table 5). To standardize across all 360 lakes in five states, species richness was used as a consistent response metric for total and native species (native by state). While metrics such as species abundance or assemblage type (e.g., bass and bluegill co-occurrence) were considered, fish assemblage data came from different agencies and were collected using different methods and amount of effort (e.g., electroshocking, fyke netting). Fish assemblages were quite different from state to state and management agencies have different social values to manage for in their lakes; these facts made species abundance or assemblage type less advantageous response metrics than species richness.

I selected the two most prominent gears for each state (Maine and New Hampshire (gill net and trap net); Michigan, Wisconsin, and Iowa (electroshocking and fyke net)). I used only the species caught with these corresponding gears to create my response metrics to include the greatest number of fish species from each lake while minimizing gear biases across lakes and management agencies (Table 4). In other words, I did not want to include species caught with a gear that was only used in a few lakes and/or a single state. State-specific records (not lake-specific) were used to determine whether a fish was native or not for all lakes in that state (Maine (David Halliwell, personal contact), New Hampshire (Scarola 1973), Michigan (Hubbs and Lagler 2007), Wisconsin (Becker 1983), and Iowa (Harlan and Speaker 1956)). Because it is impossible to determine lake-by-lake historical assemblages, I considered a fish to be native to a state if it was at least native to a single lake in the state. Since my two response metrics were derived using the same fish assemblage data, to quantify the association between these two variables, I ran a correlation of one (Total SR) against the other (Native SR).

Hydrogeomorphic Features and Anthropogenic Disturbances

I included hydrogeomorphic features and anthropogenic disturbance data in categories of regionalization, extinction/isolation, lake hydrology, and land use/cover (see details of methods in Webster et al. 2008). I compiled maximum lake depth, mean lake depth, and watershed cumulative catchment area from existing state databases, and quantified lake surface area and perimeter from state geographic information systems (GIS) data at a 1: 24,000 resolution. A lake's runoff was calculated using mean annual runoff from 1951 to 1980 within a 500-m buffer from a GIS coverage for a lake (Webster et al. 2008). Land use/ cover data within each lake's 500-m buffer were obtained from the 1992 National Land Cover Dataset (Webster et al. 2008).

I calculated additional hydrogeomorphic features not described in Webster et al. (2008) as summarized in Table 2. The regionalization frameworks used are as described above: USGS six-digit hydrologic units (97 total regions) (HUC, Seaber et al. 1987), ecological drainage units (28 total regions) (EDU, Higgins et al. 2005), and freshwater ecoregions (5 total regions) (Abell et al. 2000). To obtain a measure of species isolation, I displayed maps showing the furthest glacial extent (> 15,000 years ago) in a GIS to determine whether a lake's current location had been covered by glaciers or not (Prest 1969).

In order to get a measure of potential immigration of new species (leading to an increase in species richness), I included measures of lake connectivity and surficial hydrology as predictor variables in my classifications. I determined the number of surface water connections of each lake, both upstream and downstream, using GIS and the National Hydrography Dataset (NHD) at a 1: 100,000 resolution (http://nhd.usgs.gov, 2007). Lakes with no surface connection were declared "seepage" while lakes with one or more surface connections were determined to be "drainage". I used GIS to calculate regional and local lake density, defined as the number of lakes within each EDU (range = 159-2856) and 12-digit HUC (range = 1-89; 10,609 total regions across five-state region). An individual lake's regional and local lake density were measured by translating all water bodies (>0.001 km^2) across the entire spatial range of EDU's and 12-digit HUC's from polygons to points; these centroids were placed within the polygon of each water body.

Each lake was assigned a region (EDU and 12-digit HUC) and the total number of water bodies within each EDU and 12-digit HUC was counted. My 360 study lakes were then assigned to a region (EDU and 12-digit HUC) with its corresponding lake density for each regionalization scheme. At a finer spatial scale, I collected measures of lake hydrology such as average annual precipitation and mean base-flow index (BFI), an estimate of sub-surface water inputs relative to surface water inputs (USGS) (http://water.usgs.gov)). Average annual precipitation for years 1971-2000 was obtained from the Spatial Climate Analysis Service at Oregon State University (www.ocs.oregonstate.edu).

I also collected anthropogenic disturbance variables at multiple spatial scales (summarized in Table 3) to account for the effects of disturbances on my response metrics before creating lake classes. Related to measures of potential fish species isolation, I calculated dam density within each EDU and 12-digit HUC using dam GIS layers from state agency databases. Just like calculating lake density at multiple spatial scales, dams were assigned a region, and each study lake was assigned a regional and local dam density using EDU's (range = 48-3,634) and 12 digit-HUC's (range = 0-59). Other anthropogenic variables related to potential fishing pressure and shoreline habitat modification were collected for each lake including human population and housing density (United States 2000 Census (http://factfinder.census.gov)), and local road density (ft/acre) (calculated within a 500 m buffer of each lake).

Lake-specific fish stocking records were collected from each state to control for the effects of fish stocking on fish species richness (Maine (http://www.pearl.maine.edu), New Hampshire (Gabe Gries, personal contact), Michigan (http://www.michigandnr.com/fishstock), Wisconsin (http://infotrek.er.usgs.gov/doc/wdnr_biology/Public_Stocking), and Iowa (http://limnoweb.eeob.iastate.edu/fishgrowth)) (Table 6). I created three different metrics of stocking pressure for each individual lake: total number of species stocked, total number of stocking events (regardless of species), and total number of stocking events divided by the number of years for which I had stocking data for each state (Table 6). These metrics likely vary some from state to state because each state has different angler pressures for particular species, varying financial resources for stocking, and many biological factors to consider when determining which species to stock, when, and how often. The third stocking metric was created to help standardize stocking across states by accounting for the different number of years of stocking data available for each state (Table 6).

Developing Natural Lake Classes

To build *natural* lake classes based on hydrogeomorphic features, I wanted to first account for the effects of human disturbances on lake fish assemblages. I first transformed all anthropogenic and fish response metrics to meet normality assumptions for regressions models including arcsine-square root transformation for land use/cover variables and natural log, squared, or square root for all other variables. Because every transformation did not necessarily result in meeting assumptions (approximately 75% of the variables did not meet assumptions); I transformed the remaining metrics using the best relative transformation based on the overall distribution of data and that with the least amount of outliers. I then ran multiple linear regressions of each response metric (Total SR and Native SR) against all human disturbance variables (Table 3), regardless of correlations among variables, to obtain residual values to be response metrics when building natural lake classes. Best model selection was used to determine which disturbance variables to include in the final model (SAS version 9.1). I used condition index (CI) and variance inflation factors (VIF) to determine if multicollinearity among predictor variables was unduly influencing our models. Variables with CI values greater than 10 and VIF values greater than 30 in the regression were removed from the model (Hair et al. 1995). The best model was significant using 0.05 alpha and had both the highest adjusted R-squared and lowest Akaike's information criterion (AIC) value (Table 7).

The residuals, related to the amount of variation unexplained by these multiple regression models of human disturbance variables and fish metrics, were tested to see if they met normality assumptions after these initial regression models. They were normally distributed; therefore, no subsequent transformations were needed. The residuals were then used as the response variables in classification and regression tree (CART) analyses (R version 2.9.0) with multiple hydrogeomorphic variables as predictors (Table 2). All hydrogeomorphic variables were left untransformed and included in the CART models (this method allows for non-normal data); a correlation matrix of highly correlated (>.6) predictors was used, however, when interpreting results. CART models were built using the recursive partitioning algorithm “rpart” in the R software system that initially splits a parent group into two daughter groups based on the value of a predictor variable that maximizes the reduction in the total residual sum of squares of the response metric (Breiman et al. 1984). Each daughter group is then split using the same algorithm; I examined the proportional reduction in error (PRE) for each split in the tree and summed them to produce an overall PRE value.

Each CART model was run using a range of complexity parameters (CP; 0.01, 0.015, 0.02, 0.025, and 0.03). Increased CP values result in models with relatively coarse, less refined tree (few splits and subsequent terminal classes); decreased CP values result in a more refined tree with more splits and terminal classes. Since my other objective for this research was to examine human disturbance on lake classes (chapter 2), I sought to find a reasonable and pragmatic final classification. Therefore, I compared the number of classes and the number of lakes within each class for the classifications that were made using each CP cutoff. To be sure that each classification resulted in classes that were significantly different from one another, I ran analyses of variance (ANOVA) and post hoc Tukey tests to compare the variation between classes within a classification (SAS version 9.1). I did not objectively select a final, "best" lake classification but rather, subjectively determined the most useful classification based on the total number of classes, the range of the number of lakes in the classes, and the ANOVA results.

Results

Controlling for Human Disturbance

The two response metrics were highly correlated (r = .97); state-by-state, the total number of species in a lake is very similar to the total number of native species (native by state) (Table 5). Human disturbances explained 16% (Total SR) and 22% (Native SR) of the variation in lake fish assemblages and both final models contained similar disturbances variables: number of fish species stocked, number of stocking events/years of data, dam density (12-digit HUC), local agricultural land use, and human population density (Table 7). Even though some predictor metrics were highly correlated, all were included in the models so that I could explain the greatest amount of variation by human disturbances. With the exception of agricultural land use metrics (Total SR -, Native SR +), both models had similar parameter estimates: number of fish species stocked (-), number of stocking events/years of data (+), dam density (12-digit HUC) (-), and human population density (+) (Table 7). Within a general landscape category, many of the land use/cover and fish stocking metrics exhibited high correlation with one another (Table 8). For example, agricultural land use metrics (crop, pasture, all agriculture, and agriculture combined with urban) had coefficients greater than 0.66as were all three stocking metrics ($r > 0.85$) (Table 8).

Analysis of Variance Across Classes

Before I present the results of my lake classifications (below), I discuss how I chose a CP value, and thus a final classification for each response metric (Total SR and Native SR). ANOVA analysis of all classifications for Total SR and Native SR revealed all models to be statistically significant (Table 9). The classification with the fewest number of classes (CP=0.03) had the greatest percentage of significantly different class-by-class comparisons; few classes were not significantly different from another. Conversely, the classification with the most classes (CP=0.01) had the lowest percentage of significantly different classes; many classes were not significantly different from another. The tradeoff seen here makes intuitive sense, as the number of classes increases, there is less in-class variation but not as great of a difference in species richness from one group of lakes to the next. Because any classification is created for different purposes (with this tradeoff likely driving how refined or unrefined final classes are), lake managers will likely determine on a case-by-case basis how many classes (and how different those classes should be from one another) they would like to include in a classification.

My resulting classifications could all be considered and further analyzed since they all had significant among-class variation. The final classifications selected for presentation in chapter 1 (CP = 0.015) were not meant to be the statistically "best" classification. Rather, I chose these classifications based on the number of final classes and the range of the number of lakes across classes; in both classifications (Total SR and Native SR), over 75% of the pair-wise comparisons were significantly different (Table 9). Since my ultimate goal was to measure the effects of human disturbances on classes of lakes (chapter 2), I selected a classification (Native SR) that did not have very few classes with many lakes (e.g., CP of 0.03) or a classification with a large number of classes, each with very few lakes (e.g., CP of 0.01).

In the first scenario (too many lakes in a class), some classes are not refined enough (split into smaller groups) to make final classes management-applicable. For example, if all 360 lakes were split into only small and large lakes (< or > 53 ha), little would be known about lakes within the two classes (e.g., land use/cover) other than large lakes have more species than small lakes. Therefore, it is likely more helpful to lake managers to have smaller classes and know more about the lakes in them by refining these larger classes. In the second scenario (too few lakes in a class), there would not be enough observations to run multiple regression analysis (chapter 2) given the number of disturbance predictors in this study. For example, a class with only five lakes would result in a non-significant regression model, making the class (and its lakes) statistically unhelpful. Given this issue, it was important to have a moderate number of lakes in an individual class. For chapter 1, I presented the results for classifications using a CP of 0.015, which appeared to fall in the middle of ranges for class number and class size (e.g., number of lakes per class). Ultimately, for chapter 2, I selected a classification that had 10 classes ranging in size from 10 to 69 lakes; this decision led to a more pragmatic and thorough examination of human disturbances across the landscape on my study lakes.

Determining the Hydrogeomorphic Variables that Classify Fish Species Richness

Both lake classification trees (Total SR and Native SR) included predictors representing multiple spatial scales and general landscape categories. Most notably, both models included the extinction metric lake area as the main split in the model, a regional scale predictor (EDU), and a lake hydrology metric (watershed area) (Figures 2 and 3). Many of the morphometry variables were highly correlated with one another, (e.g., lake area vs. lake perimeter; r = 0.92, mean depth vs. maximum depth; r = 0.88, Table 8). Other highly correlated variables that appeared in at least one CART model include: wetland (woody) which is correlated with wetland (all) (r = 0.97), and cumulative watershed area which is correlated with lake perimeter (r = 0.74), lake area (r = 0.70), and number of surface connections (r = 0.69) (Table 8). It is important to recognize how these significant hydrogeomorphic variables are correlated because it affects our ability to interpret our classifications; it is difficult to distinguish the effects of some predictors from each other (e.g., lake perimeter and lake area).

The final results presented here use a CP cutoff of 0.015 in the Total SR classification. Five natural landscape predictor metrics partitioned the 360 lakes into 11 final lake classes that included 7 to 105 lakes and explained 60% of the variation in Total SR disturbance residuals (Figure 2). Lake area was the first split that separated the entire group of lakes; the regionalization framework EDU split the two main subgroups further. In addition to lake area and EDU, wetland (woody), cumulative watershed area, and mean base flow partitioned the remaining subgroups of lakes (Figure 2). The smaller lakes (< 53 ha), classes A and B, which have the lowest total species richness, lay primarily in Maine and New Hampshire, with a few lakes in parts of Wisconsin and Iowa (Figure 2 and 4). Lake classes C, D, and E primarily lay in Michigan, Wisconsin, and Iowa and have relatively higher total species richness than Maine and New Hampshire lakes (Figure 2 and 4). Larger lakes (> 53 ha), classes F, G, H, and I (151 total lakes), lay entirely in Maine and New Hampshire and make up most of Michigan, Wisconsin, and Iowa (Figure 5). These regions have relatively higher mean total species richness than classes A through E, but have relatively lower total species richness than classes J and K (Figures 2 and 5). Classes J and K lay primarily in Southeastern Michigan, Northwestern Wisconsin, and Eastern Iowa; these lakes have the highest total species richness across all 360 lakes in the 5-state region (Figures 2 and 5).

In the Native SR classification, again using a CP cutoff of 0.015, only four natural landscape predictor metrics partitioned the 360 lakes into 10 final lake classes that included 10 to 69 lakes and explained 58% of the variation in Native SR disturbance

residuals (Figure 3). Lake area was the first split that separated the entire group of lakes; the regionalization framework EDU split the two main subgroups further. While the response metric was different in the Native SR model, lake area and EDU separated the initial group of 360 lakes similarly to what I saw with Total SR. In addition to lake area and EDU, lake elevation and cumulative watershed area partitioned the remaining subgroups of lakes (Figure 3). In the smaller lakes (< 53 ha), classes A and B had the same distribution as classes A and B from the Total SR model; they lay primarily in Maine and New Hampshire, with a few lakes in parts of Wisconsin and Iowa and these lakes had the lowest relative native species richness (Figure 3 and 6). Lake classes C and D primarily lay in Michigan, Wisconsin, and Iowa and had relatively higher total species richness than lakes in classes A and B (Figure 3 and 6). The larger lakes (> 53 ha), classes E, F, G, H, I, and J (227 total lakes), had a similar distribution to the classes in the Total SR classification. Classes E, F, and G made up all of Maine and New Hampshire and almost all of Iowa and had relatively higher mean total species richness than classes A through D, but have relatively lower total species richness than classes H, I, and J (Figures 3 and 7). Classes H, I, and J lay primarily in the entire Lower Peninsula of Michigan, the Western part of the Upper Peninsula in Michigan, and Northwestern Wisconsin; there were also some small regions (with few lakes) in Southern Wisconsin and Iowa (Figure 7). Classes H, I, and J had the highest total native species richness across all 360 lakes in the 5-state region (Figure 3).

Comparing the Total SR and Native SR Classifications

The variation explained (PRE) in disturbance residuals across the two classifications using a CP of 0.015 (Total SR = 0.602 and Native SR = 0.575) was relatively similar. The two response metrics were highly correlated (r = .97) and the hydrogeomorphic predictors and spatial scales that were included and represented in the classifications were similar. Both classifications were driven primarily by lake area (extinction), EDU (regionalization), and cumulative watershed area (hydrology). Lake area, an extinction predictor and proxy for lake habitat complexity, split all 360 lakes into two subgroups (small lakes (< 53 ha) and large lakes (> 53 ha)) and explained a majority of the variation at the initial split of each classification (Total SR = 29%, Native SR = 27%) (Figures 2 and 3). Lake area also explained an additional portion of the variation in both classifications in the subgroups of lakes (Total SR = 6%, Native SR = 4%).

In addition to lake surface area, EDU's explained a significant amount of total variation in each of the classifications (Total SR =17%, Native SR = 17%). Beyond the initial split of lake area in both classifications, EDU's split the two main subgroups of lakes into four subgroups, two subgroups of smaller lakes (<53 ha) and two subgroups of larger lakes (> 53 ha) (Figures 2 and 3). In smaller lakes for both Total SR and Native SR, EDU's primarily classified lakes into two main groups, Maine/New Hampshire lakes and Michigan/Wisconsin/Iowa lakes. Although there were a few exceptions, generally speaking, small lakes in the Northeastern United States had lower species richness than lakes in the Midwestern states (Figures 4 and 6). Similar patterns were seen in the larger lakes for both classifications. Maine and New Hampshire had relatively lower species richness, although most of Iowa and parts of Michigan and Wisconsin also shared similar

EDU patterns and corresponding fish species richness. In fact for larger lakes, the highest species richness lake classes were in EDU's in parts of Michigan and Wisconsin.

In addition to lake area and EDU, cumulative watershed area explained variation in subgroups of lakes in both classifications (Total SR = 3%, Native SR = 8%), and mean base flow (hydrology) split a subgroup of lakes in the Total SR classification. Specifically, in the Total SR classification, of the larger lakes located in classes J and K, lakes > 40 km^2 watershed area (class K) had a greater species richness than lakes < 40 km^2 watershed area (class J) (Figure 2). Similarly in the Native SR classification, greater cumulative watershed area resulted in relatively greater species richness when comparing classes F to G, H to I and J, and I to J (Figure 3).

Lake elevation, an isolation metric related to the probability of a species introduction, split a subgroup of lakes in the Native SR classification. Of the larger lakes, lakes < 165.4 m in elevation (class E) had lower species richness than lakes > 165.4 m in elevation (classes F and G), opposite of what would be predicted. For these lakes, it appeared that other landscape predictors (e.g., lake area, EDU) significantly explained the variation seen in the response metric and lake elevation merely grouped the lake according to location. Lake elevation, which explained 2% of the variation among lakes, split lakes close to the ocean in Maine and New Hampshire (lower elevation, class E) and those higher in elevation in Maine, New Hampshire, and Iowa (classes F and G). The last isolation metric included in either classification was wetland (woody), which split a subgroup of lakes in the Total SR model (Figure 2). Lakes with greater local wetland cover > 1.6% (measured within 500 m buffer) had greater species richness (classes H and I) than lakes with less wetland cover (classes F and G).

Discussion

Hydrogeomorphic Features That Explain Variation in Lake Fish Assemblages

Lake area explained over half of the variation in the disturbance residuals in both classifications (Total SR = 35% of 60%, Native SR = 31% of 58%); the literature reflects the significance of lake surface area in predicting species richness; (r = 0.43-0.81; Rahel 1984, Eadie et al. 1986, Jackson and Harvey 1989). These studies were conducted in a much smaller region than my study; it appears that lake area may be even more important when only a smaller spatial scale is being studied. My results follow the literature in that at every branch where lake area was the primary splitting variable, in both classifications, larger lakes had significantly higher species richness (less likely of species extinctions) than smaller lakes. Lake area was highly correlated with other extinction metrics (e.g., lake perimeter) so it could be broadly stated that greater habitat complexity (for my 360 lakes) was the primary hydrogeomorphic predictor in shaping either total or native species richness across a 5-state region. Lakes with larger surface areas tend to have extended littoral zones allowing for increased habitat and refuge for smaller species of fish (e.g., protection from predation). Deeper lakes typically stratify thermally, which can provide habitat for cool water species in the open water zone (pelagic) of a lake. With more complex habitat available in larger, deeper lakes, unique species are able to exist (upon introduction), which supports the potential for large populations and less vulnerability to extinction. The relationship between greater species richness and greater habitat complexity has been shown in many other studies; extinction metrics play a large role in shaping inland lake fish assemblage (Tonn and Magnuson 1982, Magnuson et al. 1998).

These studies, however, did not explicitly incorporate regional-scale metrics or attempt to remove the effects of anthropogenic disturbances. By using a large geographic area, by including regionalization frameworks, and by accounting for the effects of humans, other predictor metrics (e.g., EDU) explained some of the variation seen in fish species richness. Cumulative watershed area was highly correlated ($r > .7$) with habitat complexity metrics (lake area and lake perimeter) (Table 8); it can generally be stated that a greater cumulative watershed area (relatively higher habitat complexity) may also result in greater species richness for all sizes of lakes. Cumulative watershed area was also correlated with the number of surface water connections ($r = .69$) (Table 8), a strong proxy for landscape position (isolating metric). Lakes that are higher in the landscape are less likely to experience species introductions whereas lakes lower in the landscape are less isolated (higher likelihood of an introduction). Lakes with greater cumulative watershed area (complex habitat, lower in landscape) exhibit positive correlations with species richness (Kratz et al. 1997, Magnuson et al. 1998, Riera et al. 2000, Olden et al. 2001); this was exhibited in the Total SR classification (Figure 3).

Overall, local extinction and isolation metrics explained the greatest amount of variation in both total species richness and native species richness. In every case, with the exception of the lake elevation split in the Total SR model (Figure 3), lakes lower in the landscape, less isolated, had greater species richness than lakes higher in the landscape (less likely for an introduction). Lake elevation in a local region (e.g., single state or group of counties) followed that of landscape position; lower elevations had more species than lakes at a higher elevation. However, since my study region spanned from sea level (Maine) to relatively higher elevations in the Midwest (Michigan and

Wisconsin), elevation in this study was likely not an appropriate proxy for landscape position. For example, a lake with a relatively low landscape position in a high elevation area may be at the same elevation as a lake with a relatively high landscape position in another watershed.

While local metrics explained a large portion of the natural variation, EDU's (large spatial scale predictor) explained a significant amount of variation as well (17% in both classifications). Across classifications and lake sizes, EDU's in Maine and New Hampshire generally split from those in Michigan, Wisconsin, and Iowa. However, exceptions existed for smaller lakes in the Total and Native SR classifications. Some EDU's in Michigan (Upper Peninsula), Wisconsin, and Iowa shared characteristics with those in the Maine and New Hampshire (Figures 4 and 6). Likewise, for larger lakes, EDU's in parts of Michigan, Wisconsin, and most of Iowa shared some commonalities with the Maine and New Hampshire lakes, as reflected in the distribution of lake classes (Figures 5 and 7).

Regionalization frameworks of three different spatial scales were included in analysis; however, EDU's were the only regional-scale metric to explain variation in either Total SR or Native SR. Although freshwater ecoregions were delineated for freshwater systems to identify distributions of fish species across all of North America (Abell et al. 2000), they appear to be too large in scale (five regions across five states) to significantly explain variation in my study lakes. In contrast, EDU's were delineated primarily using drainage boundaries such as watersheds (eight-digit HUC's) and identified areas with similar biotic patterns (Higgins et al. 2005). EDU's, generally between 1,000 and 10,000 km^2 in area, are at a scale where watershed boundaries, climate, and landscape features would reveal aquatic biodiversity patterns. Their consistent inclusion in my lake classifications, regardless of lake size, reflected their ability to detect broad-scale species richness patterns. The classes of lakes here, partitioned by a regional-scale predictor, support earlier findings that demonstrate the importance of including relatively broad spatial-scale features (Tonn 1990, Poff 1997, Cheruvelil et al. 2008). Past attempts at classifying lakes have not included regionalization schemes when building classifications based on fish assemblages (e.g., Dolman 1990, Schupp 1992). Research on classification schemes for fish assemblages was conducted before many of the regionalization frameworks were created. Likewise, these studies were uniquely developed within single states and may not have warranted including such broad-scale predictors to examine fish assemblages. Some of the regionalization frameworks themselves sought to identify unique biodiversity patterns nationally and internationally. Current research and conservation efforts are focused on multi-state, and even nationwide collaboration and management (e.g., the Environmental Protection Agencies' National Lakes Assessment (http://www.epa.gov/owow/lakes/lakessurvey/)). Therefore, examining variation of fish assemblages across state boundaries warrants the inclusion of regional-scale predictors in new classification schemes.

It is important to note that there are some limitations that I faced when building these lake classifications. For example, I utilized the National Land Cover Dataset from 1992, but my fish assemblage data came from many years following that, as late as 2008. It is possible that land use/cover changes that occurred between when the two sets of data were collected influenced my results. In addition, the stream layer I used to categorize stream connections was at a 1:100,000 resolution. It is possible that I did not capture all relevant hydrological connections such as small perennial streams that could be sources of immigration of new species and/or refuges from hypoxia (extinction metric). In addition, by not completely and accurately identify all hydrological connections, I was likely not able to properly identify seepage or drainage status for every lake or determine for all lakes which ones were natural and man-made (reservoirs). Reservoirs have different hydrological characteristics than natural lakes, including a short water retention time and typically larger watersheds and surface areas. Lake area and cumulative watershed area explained a great amount of variation in my lakes; however, specific classes had both natural lakes and reservoirs (e.g., Iowa reservoirs shared classes with natural lakes from other states). Although different types of lakes (e.g., natural vs. reservoir) can share fish assemblage characteristics (e.g., same number of species), a rigorous search into the background of all 360 lakes could have identified whether it was natural or man-made. Using continuous data I already have, a metric such as the ratio of cumulative watershed area to lake area may identify hydrological difference among lake types.

Differences in Species Richness Between the Northeastern and Midwestern States

The glacier predictor metric I included in CART analysis was not significant in any tree splits. However, the effects from glaciers and local glacial refugia appeared to play a role in my results, and might have been captured in part by EDU. During the furthest glacial extent thousands of years ago, fish in the Midwest (Michigan, Wisconsin, and Iowa) and parts of Canada had the Mississippi drainage system as a freshwater retreat from lakes being covered in ice. Once the glaciers retreated, fish were able to colonize or disperse to previously glaciated areas, resulting in the upper Midwestern states and Canada becoming relatively speciose (Rempel and Smith 1997). Fish in Maine and New Hampshire did not have as accessible of a freshwater refuge from glaciers. In fact, many freshwater species were likely displaced into the Atlantic Ocean (saltwater) and extirpated due to the inability to quickly evolve with the rapid changes, resulting in the Northeastern inland lakes having fewer fish species (Table 5). Once the glaciers revealed freshwater ecosystems upon their retreat, there were fewer species left behind to colonize these lake and streams.

As discussed in the introduction, I selected fish species richness (a measure of biodiversity) as my response metric over other potential fish assemblage metrics (e.g., fish abundance, assemblage types). Although other studies have used species associations in streams (Brenden et al. 2008) or species abundance in lakes (Marshall and Ryan 1987), I felt that it was more pragmatic to use general species counts; this metric has been successfully used in other classification studies (Magnuson et al. 1998). There were differences in the biodiversity patterns of the five states; as discussed above, glaciations patterns result in Maine and New Hampshire naturally having many less species than Michigan, Wisconsin, and Iowa. Additionally, there are differences in the types of species found in different areas; the Northeast has multiple species that are not present in the Midwest and vice versa. I also used general species counts rather than species abundance because my fish assemblage data was from different management agencies that for the most part, attempt to examine the status and trends of their own lakes and are typically concerned mainly with recreational species. Therefore, I could not be certain that agencies were not targeting recreational species when selecting lakes to study or at least specifically targeting and capturing particular species while sampling in a lake. Although species richness as a response metric has its drawbacks, species abundance, in this case, potentially could have had even more biases associated with it. My decision to use species richness was also based on my inability to control for having data from five different states (four different agencies). I used species richness based on considerations of what would best integrate all five states and their fish assemblages. Species richness is useful because even though management agencies do not always specifically concern themselves with biodiversity, they certainly take precautions to avoid species loss across the state. Some may not be greatly concerned if they stock a species into a lake that results in the loss of a minnow or sunfish species for that specific lake. However, state-wide, there are divisions within state management agencies that are concerned with biodiversity, maintain a conservative approach, and care for all public resources, not just those that have recreational value.

In the future, consideration could be given to using fish abundance or co-occurrence of species as a response metric. For example, a lake with seven species in New Hampshire could have a totally different composition than a lake with seven species in Wisconsin. Tolman Pond in New Hampshire shared four species with Friendship Lake in Wisconsin, both of which had seven total species and were in the same class. However, both Pine Lake and Squaw Lake in Wisconsin, also with seven total species and in the same class, only shared two species with Tolam Pond. By using other response metrics such as co-occurrence or abundance of fish species, natural features may identify different lake classes across such a broad region. However, species richness warrants as much consideration by management agencies, resulted in interesting and useful classifications in my study, and has been shown to reveal significant features shaping inland lake fishes (Magnuson et al. 1998)

Although I attempted to look at differences of how well the natural landscape shaped total species richness versus total native species richness, my response metrics were likely too similar (r = 0.97) to confidently discern differences between one another , which may indicate that my count of native species for each lake was too coarse. I considered any species that was native to a state to be native to any lake; however, there are cases where a species may only be native to a few watersheds in a state and non-native in any other lakes outside those watersheds. Given the way I calculated the number of native species in a lake across a state, I was not able to identify which species were truly non-native to a lake. Most of my lakes were estimated as having zero or one non-native species (~300 lakes out of 360); this is likely not what is present in most lakes that are affected by some form of human disturbance. A more refined method of identifying native and non-native species in each lake may reveal a greater presence of non-natives in my data than what I present here. Ultimately, my two metrics were likely too similar; more work is required to best identify what natural features specifically affect all or just native fish species.

Conclusions

Despite the shortcomings of my research and some classification frameworks in general, my predictive lake classification using different categories of hyrdogeomorphic and anthropogenic landscape features at multiple spatial scales explained variation in different metrics of fish assemblages. Given the recent advances in technology, GIS data can be combined with easily obtainable descriptive lake information to form landscape-based classes with a variety of predictor variables such as regionalizations, isolation/extinction metrics, and lake hydrology. Based on known lake-to-landscape relationships tested using data from sampled lakes, managers can then apply these results to unsampled lakes to help guide their management decisions. Even though individual lakes are unique in some respects, they have many local and regional characteristics in common that help to shape fish communities.

By examining these shared characteristics across lakes within and even among states, and identifying the similarities and differences in the fish communities, managers can better understand the landscape's effects on lakes in order to conserve and protect unique aquatic resources while providing recreational opportunities in inland lakes. More credence should be given to species richness (i.e., biodiversity) and conserving by management agencies. Sport fisheries can be created through fish stocking in lakes with poor biodiversity, but recreational opportunities are more stable in lakes with high biodiversity and strong ecosystem-level interactions. Since lake managers are required to implement fishing regulations and restrictions on all public waters while explicitly having fish assemblage data on only a small portion of them, classifications provide a tool that can aid those charged with the task of managing inland lakes. Additionally, by dividing lakes into classes using map-based characteristics, lake managers may be able to appropriate resources to other projects otherwise used for rigorous and expensive sampling programs. Classifications can also help to indentify regional and national biodiversity patterns by grouping lakes of relatively lower species richness and higher species richness. Choosing lakes to monitor across very broad regions using a classification approach could also reveal the loss or gain of species due to human impacts and allow us to better-understand the natural state of aquatic biota in the absence of humans.

My research explored previously overlooked predictor metrics and their ability to explain variation in fish assemblages using a predictive classification framework. My study was unique in that I incorporated many states, each with a wide ranging set of natural features and human disturbances, as well as regional-scale landscape predictors such as regionalization frameworks and the number of dams per EDU. My classification results highlight the importance of regional-scale predictors and suggest that they be considered when attempting to explain the variation seen in fish assemblages, especially when using a classification approach and a large geographic area.

Chapter 2-Using a hydrogeomorphic-based lake classification to compare effects of anthropogenic disturbances on fish species richness in 360 north temperate lakes.

Introduction

Lake managers and researchers seek to understand and explain the variation among fish assemblages in inland lakes, often trying to distinguish between effects stemming from natural features and human disturbances. In the absence of human disturbances, hydrogeomorphic predictors shape in-lake biota. For example, lakes with greater habitat complexity (e.g., larger lake area), increased potential for species invasion (e.g., greater number of surface connections past and/or present), and lower hydrologic position in the landscape (e.g., greater cumulative watershed area) generally have greater species richness (Tonn and Magnuson 1982, Rahel 1984, Kratz et al. 1997, Hershey et al. 2000, and Irz et al. 2004). However, because humans can impact inland lake fish species through species additions (e.g., fish stocking), blockage of movement (e.g., dams), and landscape alterations (e.g., land use/cover), among other disturbances, the relationship between the natural landscape and fish assemblages can be disrupted (Radomski and Goeman 1995, Carpenter et al. 2007, Johnson et al. 2008). Therefore, understanding how the hydrogeomorphic landscape structures fish assemblages must be augmented with an understanding of how anthropogenic disturbances affect fish assemblages in inland lakes.

With advances in technology such as Geographic Information Systems (GIS), researchers have powerful tools to investigate the effects of land use/cover (e.g., agriculture, urban) and human population densities on all aquatic systems. To date, much attention has been paid to the impacts of humans on water quality and/or fish assemblages in lotic systems (Gergel et al. 2002, Buck et al. 2003, Jones et al. 2004), whereas research on lentic systems has lagged behind. However, recent research has demonstrated that agricultural and urban land within catchments can negatively impact fish communities in lakes (Carpenter et al. 2007). Riparian alterations around a lake have also caused negative impacts to fish assemblages such as the loss of forested land due to shoreline development (i.e., increased agriculture land use, higher human populations) (Jennings et al. 1999, Carpenter et al. 2007). Increased shoreline development has resulted in the loss of coarse woody debris (CWD) within a lake and directly caused negative changes to the fish community (Roth et al. 2007). GIS has been useful in identifying these lake catchment and riparian alterations which may aid in measuring the negative effects on fish assemblages.

In addition to land use/cover, inland lake fish assemblages can be affected by roads. Roads are large, flat, impervious surfaces that alter local hydrology and pH (Rahel 1984); additionally, lakes near roads (areas of relatively higher human populations) can have higher fishing pressure and fish stocking rates than remote lakes (Kratz et al. 1997, and Magnuson et al. 1998). Roads have been viewed as a direct source of species immigration (e.g., lake to nearest road distance), but also as a catchment-wide influence from hydrological alterations that may ultimately influence a lake's fish assemblage.

Another disturbance that can affect fish movement is dams. These structures fragment aquatic systems and have been found to change species composition of vegetation and fish communities in streams (Hill et al. 1998, Brunberg and Blomqvist 2001). In this sense, for lakes connected by streams, movement of species between lakes is more likely in unfragmented systems than in those with dams blocking movement. Not only do dams fragment stream connections between lakes, they have also been used to create large impoundments where otherwise a lake would not be present. Recent research suggests that impoundments have higher likelihood of non-native species invasions than natural lakes (Johnson et al. 2008). These researchers found that impoundments are likely to be more accessible to humans than natural lakes (e.g., greater number of boat landings in impoundments). In addition, impoundments are more likely to be connected to stream networks and have more surface connections than natural lakes. Therefore, it appears that dams can both fragment stream systems (i.e., limit movement of native species) and, through the creation of man-made systems (e.g., impoundments), increase non-native species (Johnson et al. 2008).

Perhaps the most profound and well-established anthropogenic disturbance to affect fish assemblages in inland lakes is fish stocking. Research on fish stocking is quite extensive both at the individual species level and fish assemblage levels (Anderson and Schupp 1986, Radomski and Goeman 1995, Holmlund and Hammer 2004, Irz et al. 2004). Adding a species to a lake (stocking or unintended release) can result in altered food web dynamics, the elimination of a species from a system (through competition or predation), indirect changes in water clarity, and changes in fish growth rates (Anderson and Schupp 1986, Radomski and Goeman 1995, Holmlund and Hammer 2004, Irz et al. 2004). For example, Irz et al. (2004) found native fish species richness and non-native species richness to be negatively correlated, indicating that introductions have led to saturated communities (limited potential of fish species richness to increase). Although fish stocking across the landscape is primarily intended to help manage faltering ecosystems and/or create new fisheries, stocking can result in unintended negative consequences that alter fish assemblages in inland lakes.

Some human disturbances cause an increase to total species richness (e.g., fish stocking, road proximity), while other disturbances result in a decrease in species richness by fragmenting lake connectivity (e.g., dams). Since anthropogenic disturbances vary across the landscape (e.g., dam densities, stocking rates within and across states), and because features of the natural landscape may mediate effects of human disturbances on fish assemblages, impacts from human disturbances on inland lake fish assemblages should be measured more broadly than in a single lake or even single state and should measure the effects of disturbances on groups of naturally similar lakes (i.e. fish assemblages). However, to date, research investigating effects of human disturbances on fish assemblages has typically looked at a particular species, only in-lake features (such as CWD), or a small set of water bodies in a single state (Rahel 1984, Roth et al. 2007). Therefore, I hypothesize that measuring effects of anthropogenic disturbances on fish assemblages in classes of naturally similar lakes across a broad region will reveal unique patterns otherwise masked by lake heterogeneity.

In order to measure the effects of anthropogenic disturbances on subgroups of naturally similar lakes (lake classes), I first developed a hydrogeomorphic-based lake classification system (that accounted for anthropogenic disturbances) using only the native fish species within each lake (see Chapter 1 for details). I chose native species richness (as opposed to species abundance or assemblage type) to make a more straightforward comparison between different regions, each with different fish species composition. In this effort, I identified the most significant natural predictor variables across 360 lakes, subdividing the original group of lakes into 10 classes, ranging in size from 10 to 69 lakes. The classification system was driven primarily by lake area, cumulative catchment area, and the ecoregion (ecological drainage unit, EDU; Higgins et al. 2005) in which each lake was located (Figure 3). The mean native species richness across classes ranged from a low of 5.1 in Class A to a high of 20.4 in Class J (Figure 3 described in Chapter 1). In this chapter, I then sought to compare the effects of anthropogenic disturbances on total native species richness within each lake class. Because lake area was the primary variable in splitting all lakes into two subgroups, small lakes (< 53 ha, Classes A, B, C, and D) and large lakes (> 53 ha, Classes E, F, G, H, and I) (Figure 3 described in chapter 1), I also compared the effects of anthropogenic disturbances between lakes in these broader size-based groupings.

Lastly, to look at fish assemblage responses in more detail than total native species richness as affected by human disturbances across sizes or classes of lakes, I determined what disturbance variables affected specific groups. To do so, I assigned fish species to one of two groups: recreational or non-recreational. These two categories of fish likely function somewhat differently in ecosystems, and they often are associated with contrasting social values. Recreational species often are highly sought after by humans, hold relatively high economic value, and are typically at higher trophic levels. Non-recreational species have little economic value and hold lower trophic status; in some cases, they are conserved only to maintain biodiversity in a lake. Because each fish type has different ecosystem and social value, I hypothesized that recreational and non-recreational species would differ in their response to a given disturbance. For example, walleye (recreational) and emerald shiner (non-recreational) may both be native to a given state, but each holds unique values within that state. In this sense, a specific anthropogenic disturbance (e.g., fish stocking) may affect recreational native species differently than non-recreational native species in a set of lakes within or across states. Disturbance metrics such as road density or human population density, which have been attributed to fishing pressure (Magnuson et al. 1998), is likely attributable to the presence of recreational species in a lake rather than non-recreational species. Therefore, these species disturbances may affect recreational species differently than non-recreational. Therefore, I measured the effects of human disturbances on recreational native species and non-recreational native species within each of my 10 hydrogeomorphic-based lake classes.

Methods

Selecting a Lake Classification from Chapter 1

In Chapter 1, I created two classifications of lakes based on the hydrogeomorphic landscape using total species richness (Total SR) and native species richness (Native SR) as response variables, while accounting for anthropogenic disturbances (Chapter 1: Figures 2 and 3). Both classifications explained 67% of the total variation in species richness (combined variation explained by multiple regression and CART analysis), had a similar number of lake classes (Total SR, 11 classes; Native SR, 10 classes), and each included very similar natural landscape predictors (e.g., lake area, EDU). Because the Total SR and Native SR classifications were driven by similar predictors, contained lake classes located in relatively similar areas of the 5-state region (Chapter 1: Figures 4, 5, 6, and 7), and explained variation equally well, neither classification was objectively deemed "better" than the other. However, since Chapter 2 is primarily focused on studying how humans have impacted "natural" lake fish assemblages, I used the Native

SR classification in Chapter 2. The Native SR classification of 360 lakes is made up of 10 classes with membership ranging from 10 to 69 lakes across Maine, New Hampshire, Michigan, Wisconsin, and Iowa (Chapter 1: Figures 3, 6, and 7).

Fish Assemblages of the 10 Lake Classes

Each of the 10 final classes from the Native SR classification from Chapter 1 (Figure 3) was comprised of a relatively unique range and mean value of total native species richness. Because all lakes spanned across a wide spatial range and contained such a variable assemblage of fish species from state to state, I first wanted to determine if classes differed in terms of their predominant fish species. To do so, I determined the most prevalent species in each class, and if the most prevalent species were common for the vast majority of the lakes in each class. I conducted this analysis to determine if most of the lakes in a class contained the same three or four most common species (indicating some similarity in species composition in that class) or whether species composition, even of the most prevalent species, was highly variable among lakes in a class. Lastly, I compared the most prevalent species across classes to see if similar species were most prevalent in several classes, or whether the identity of predominant species could be used to distinguish among lake classes. For all species assemblage comparisons within and across classes, I created frequency histograms and displayed the percentage of lakes in each class for which each species was present.

Native Species Richness Response Metrics

By state, I classified each fish species as either recreational or non-recreational. I used each state's recreational fishing guide (published by the state's fish management agency) as a standard, (Maine, Department of Inland Fisheries and Wildlife 2008; New Hampshire, Fish and Game Department 2009; Michigan, Department of Natural Resources 2009; Wisconsin, Department of Natural Resources 2009; and Iowa, Department of Natural Resources 2009). If a species had a bag/size limit or seasonal/gear restriction associated with it, it was declared recreational; fish with no restrictions were deemed non-recreational (Table 10).

In total, three fish species response metrics were created, total native species richness (SR), native recreational SR, and native non-recreational SR (Table 10). Total native SR was used as the response variable to analyze the effects of disturbances between small and large lakes (Table 11) and within each of the 10 classes of lakes created in Chapter 1. Within an individual lake class (Chapter 1: Figure 3; Classes A through J), native recreational SR and native non-recreational SR also were used to compare the effects of human disturbances on these different fish types (Tables 12 and 13).

Effects of Anthropogenic Disturbances on Fish Species Richness

I transformed all anthropogenic disturbance and fish response metrics to meet normality assumptions for regression models including arcsine-square root transformation for land use/cover variables and natural log, squared, or square root for all other variables. For metrics that did not meet normality assumptions, even after transformation (approximately 75% of the variables did not meet the assumptions for any of the transformations), I selected the data transformation that resulted in the distribution most approaching normality and with the fewest number of outliers. In short, because for some metrics there was not a transformation that completely worked, I used the transformation that did the "best relative job".

In order to measure the effects of human disturbances on fish species richness, I first ran multiple linear regressions of total native SR against non-highly-correlated ($r < 0.6$) human disturbance variables that included metrics of fish stocking, dam density, land use/cover, road density, and human population density (Tables 14 and 15). Best model selection was used to determine which disturbance variables were significant ($p < 0.05$) and to include in the final model (SAS version 9.1). I used condition index (CI) and variance inflation factors (VIF) to determine whether multicolinearity among predictor variables was unduly influencing my models. Variables with CI values greater than 10 and VIF values greater than 30 in the regression were removed from the model (Hair et al. 1995). The best model had both the highest adjusted R-squared and lowest Akaike's information criterion (AIC) value.

I ran multiple linear regression models using disturbance variables with total native SR as the response metric to determine any effects (positive or negative) of human disturbances. Models were run for each of the Native SR lake classes from Chapter 1 (Figure 3) and for each of the two broader groups of small lakes (< 53 ha, 133 lakes) and large lakes (> 53 ha, 227 lakes). Lastly, to explore how different species vary in their response to disturbances, I conducted the same analysis for recreational and non-recreational species within each of the 10 classes. To compare the relative ability of the models to explain variation in species richness, weighted R-squared values were calculated by multiplying the number of lakes in a class (or size group) by the total R-squared value in that specific class (or size group), adding those values all together, and then dividing by all lakes (360).

Results

Fish Species Assemblages in 10 Lake Classes

Across all 10 classes, sunfishes such as bluegill and pumpkinseeds were "dominant" species (at least one was in the top three species by frequency in every class); largemouth bass and black crappie were consistently "dominant" as well (Figures 8 through 17). Brown bullhead, yellow perch, golden shiner, walleye, chain pickerel and northern pike were also some of the more prevalent species; each of these species was one of the top three "dominant" species in at least one class (Figures 8 through 17). Eight classes of the 10 were comprised of dominant species in a vast majority of the lakes (Classes B, C, D, E, F, H, I and J), at least 83 % of the lakes (Class F) and as high as 100 % of the lakes in two other classes (Classes D and J). Meaning, for a given class, the most dominant species were typically the same three or four species in most of the lakes rather than high variability in assemblages lake to lake in a class. Because the most prevalent species in these eight classes were common to so many lakes in the class, it indicates a lack of within-class variability of assemblage type (as defined by the most prevalent species). In fact, Classes B, C, D, and F had the same three species as their most frequent: largemouth bass, bluegill, and black crappie (Figures 9, 10, 11, and 13). Therefore, what distinguishes these classes is not their most common species (as determined by presence/absence), but rather the presence/absence of rarer taxa.

Classes A and G were more variable in species composition (Figures 8 and 14). The most prevalent (frequently observed) species (Class A, brown bullhead, golden shiner, pumpkinseed, yellow perch; Class G, bluegill, largemouth bass, yellow perch) were in fewer than half the lakes for Class A and approximately 63 % of the lakes in Class G (Figures 8 and 14). This result suggests that compared to the other eight classes, these two classes, A and G, have more variability in species composition across their lakes. Because these classes contain lakes from all five states; there are many assemblage types within the class with the dominant species being highly variable. The other eight classes have lakes from only four states (Classes C and F), three states (Classes B, D, H, I, and J) and two states (Class E), indicating less variability in assemblage types and therefore more consistent dominant species for a given class.

Comparing Disturbance Effects Among Lake Classes

My exploration of how anthropogenic disturbances were associated with total native species richness across lake classes included ten class-specific models. Of these models, eight had one to three significant predictor variables resulting in adjusted R-squared values ranging from 0.17 to 0.38 across the classes (Table 16). Two class-specific models (B and J) had no significant predictor variables, four classes (A, D, E, and F) had a single significant predictor, two classes (H and I) had two significant predictors, and two classes (C and G) had three significant predictors (Table 16). The total weighted adjusted R-squared across all 360 lakes (10 classes) was 0.20; 20% of the variation of total native SR in the natural lake classes was explained by anthropogenic disturbances (Table 16). Human population density was positively associated with six different lake classes (small and large lakes) and road density was the only disturbance metric to have no significant association with any lake classes (Table 16). The number of species stocked (Class G), dam density per EDU (Class E), dam density per 12-digit HUC (Class C), agricultural land use (Class H), urban land use (Class I) and agricultural and urban combined land use (Class E, F, and G) all had a negative association with total native SR in some lake classes (Table 16).

Comparing Disturbance Effects Between Small and Large Lakes

My exploration of how anthropogenic disturbances were associated with total native species richness across a large range of lake sizes included two main regression models based on lake size. The small lake (< 53 ha) regression model had three significant predictors; number of dams per 12-digit HUC (negative), road density (positive), and human population density (positive) together explained 21 % of the variation in total native species richness across 133 lakes (Table 17). The large lake (> 53 ha) regression model had five significant predictors. The number of dams per 12-digit HUC (negative association), road density (positive association), and human population density (positive association) had the same impacts on large lakes as small lakes and the number of species stocked (negative association) and agricultural land use (negative association) additionally affected large lakes (Table 17). However, even though there were two additional human disturbance metrics in the large lakes model, the five significant metrics collectively explained only 12 % of the variation across 227 lakes (Table 17). The total weighted adjusted R-squared across the two lake size groups is

0.15; 15% of the variation of total native SR is explained by anthropogenic disturbances (Table 17).

Comparing Disturbance Effects on Recreational and Non-Recreational Species Richness

My exploration of how anthropogenic disturbances have affected native recreational and non-recreational species richness included ten class-specific models. Of the native recreational species richness models, eight of them had one to three significant predictor variables resulting in adjusted R-squared values ranging from 0.08 to 0.71 across the classes (Table 18). Two class-specific models (I and J) had no significant predictor variables, two classes (F and H) had only a single significant predictor, three classes (A, B, and C) had two significant predictors, and three classes (D, E, and G) had three significant predictors (Table 18). The total weighted adjusted R-squared across all 360 lakes (10 classes) was 0.25; 25 % of the variation of native recreational SR in the natural lake classes was explained by anthropogenic disturbances.

In the native non-recreational species richness models, eight of them had one to three significant predictor variables resulting in adjusted R-squared values ranging from 0.14 to 0.81 across the classes (Table 19). Two class-specific models (B and J) had no significant predictor variables, one class (C) had only a single significant predictor, five classes (A, D, F, H, and I) had two significant predictors, and two classes (E and G) had three significant predictors (Table 19). The total weighted adjusted R-squared across all 360 lakes (10 classes) was 0.33; 33 % of the variation of native non-recreational SR in the natural lake classes was explained by anthropogenic disturbances.

In smaller lakes (Classes A, B, C, and D), human disturbances affected native recreational species differently than native non-recreational species; these differences were seen in the significant predictors in the final regression models. Native recreational SR was influenced primarily by the negative associations with dam density (exception Class D) and positive associations with urban land, agriculture and urban land combined, road density, and human population density (Table 18). Native non-recreational SR was also positively associated with road density and human population density; however, land use/cover metrics (e.g., agriculture, urban, and agriculture and urban combined) had negative associations (Table 19).

In large lakes (Classes E, F, G, H, I, and J), native recreational SR was affected by negative associations with fish stocking and the combined land cover/use of agriculture and urban (Table 18). Native recreational SR was positively associated with dam density, agriculture, road density, and human population density. Native non-recreational SR was negatively associated with dam density for some classes (F and G) but positively associated with Class H (Table 19). Combined agriculture and urban land use/cover also had differing associations depending on lake class for native non-recreational SR, positive in Class E, but negative in Classes F and H (Table 19). Agriculture and urban land cover/use as individual disturbances were negatively associated with native non-recreational species; human population density was a positive association. Overall, within a given class, recreational species and non-recreational species were affected differently by anthropogenic disturbances. Because these two types of fish hold such different social values and ecosystem functions, humans appear to have varying associations with different native species rather than consistent impacts across all fish in inland lakes.

Discussion

Measuring Anthropogenic Disturbance Effects on Native Species in Lake Classes

Almost all of the anthropogenic disturbances included in my models (all variables but road density and human population density) were negatively related to total native SR (for the eight classes with significant models). An increase in the number of fish species stocked was associated with a decrease in native fishes (Class G); literature has shown that an introduction of non-native species into lakes has caused declines in native species abundance (Irz et al. 2004). My metric of stocking (number of species) included all native and non-native species, therefore, it can not be concluded that the negative impact on native fishes was due solely to stocking non-native species. However, because stocking can result in altered food web dynamics, and the loss of a species from a system (through competition or predation), among many other changes, fish stocking can certainly have the effect of lowering species richness in lakes (Anderson and Schupp 1986, Radomski and Goeman 1995, Holmlund and Hammer 2004).

Dam densities had similar associations as fish stocking (negative) in Classes C (per 12-digit HUC) and E (per EDU) (Table 16). Johnson et al. (2008) found that non-indigenous species were more likely to invade impoundments (dam created) than natural lakes; this certainly has repercussions on native species. Dam density, either scaled in 12-digit HUC's or EDU's, had a negative association with total native SR. Dams fragment systems that would otherwise be connected in their absence, causing changes to the species composition (Hill et al. 1998, Brunberg and Blomqvist 2001).

Individual land use/cover metrics of agriculture (Class H) and urban (Class I), as well as their combined use/cover (Classes C, F, and G) all impacted total native SR negatively (Table 16). Agriculture was highly correlated with agriculture and urban combined ($r = 0.96$). Further, forest land cover was highly negatively correlated with agriculture ($r = -0.85$) and agriculture and urban combined ($r = -0.89$). Literature has shown that loss of forested land in the riparian of a lake (i.e., loss of CWD) can result in changes to the fish assemblage in a lake (Roth et al. 2007). In Classes C, F, and G, higher agriculture and urban combined land use (i.e., lower forested land cover) had negative associations with total native SR. Lakes with the highest relative total native SR (by class) were mostly found in Michigan and Wisconsin (i.e., higher forest, lower agriculture) whereas less species-rich lakes were mostly in Iowa (i.e., lower forest, high agriculture) (Chapter 1: Figures 6 and 7). However, it can be generally stated that Iowa had lower species richness overall than Michigan and Wisconsin, regardless of lake class, so agricultural and urban land were likely not the only factors, natural or human, causing differences between the two broad regions.

Human population density appeared positively associated with total native SR across the entire five state region; higher human population densities in six classes (small lakes, A, C, D; large lakes, G, H, and I) correspond to higher native SR. While road density did not significantly affect any class, lakes near roads (i.e., areas of relatively higher human populations) can have impacts on fish assemblages through higher fishing pressure (boat movement and bait buckets) and intentional fish stocking rates as compared to remote lakes (Rahel 1984, Kratz et al. 1997, Magnuson et al. 1998). While species from these introductions were often non-native to the specific lake, it was likely the species was native to the state. Given that we determined native/non-native status for each species by state, an increase in total native SR in areas of higher human population densities could actually be a species native to the state (non-native to the lake) that has been moved from one lake to another, increasing total native SR in lakes near human population centers. In addition, humans may choose to live near lakes with higher species richness; therefore, it is not necessarily that higher SR is an effect from humans but an association with higher human populations.

Measuring Anthropogenic Disturbance Effects on Native Species in Small and Large Lakes

The number of species stocked was negatively associated with total native SR in large lakes while dam density (per 12-digit HUC) was negatively associated with total native SR in both sizes of lakes. These results corroborate the literature in that stocking can have negative impacts on a lake's fish community while dams fragment systems and often are associated with impoundments (i.e., higher human access) (Irz et al. 2004, Johnson et al. 2008) (Table 17). Agricultural land use had a negative association with total native SR in larger lakes, potentially due to the inverse relationship agriculture had with forested land cover (r = -0.85) (data not shown). Many of the lakes with relatively lower total native SR were located in Iowa (i.e., high agriculture, low forest) while the lakes with the highest relative native SR were in Michigan and Wisconsin, where there is less agricultural land (Chapter 1: Figure 6). These results reflect the literature; areas of lower forested cover (i.e., higher agriculture, increased shoreline development) have led to changes in fish community structure (Jennings et al. 1999, Carpenter et al. 2007). This does not imply that Iowa was once vastly covered by forest and now is completely altered by humans to be covered with agriculture, but broad-scale patterns of species richness (e.g., low versus high native species richness) may reflect the broad land cover types that currently exist in our study region.

Road density and human population density were positively associated with total native SR in small and large lakes (Table 17). A lake's proximity to roads has previously been shown to negatively impact species richness (Rahel 1984, Magnuson et al. 1998). However, since road density was moderately correlated with human population density in our study (r = 0.52, data not shown), it is entirely possible that road density is exhibiting the same positive effects on total native SR as human population density (Tables 16 and 17). Additionally, our metric for roads was calculated as the road density around the lake (within a 500-m buffer), not distance to the nearest road (Rahel 1984, Magnuson et al. 1998). Our measure of road density around a lake may actually better-reflect the likelihood of human access to lakes because as road density increases it is likely that access is easier, potentially leading to a greater likelihood of species richness to increase through introductions. Similarly to what we found with road density, as human populations increased around small and large lakes, total native SR increased (Table 17). Regardless of lake size, native species richness was higher in areas with increased road and human population density, increased accessibility of humans to lakes was positively associated with greater biodiversity.

Examining Recreational and Non-Recreational Species Within Lake Classes

The effects of anthropogenic disturbances explained the variation in native recreational species ($R^2 = 0.25$) and native non-recreational species ($R^2 = 0.33$) across all classes of lakes (Tables 18 and 19), the variation explained in total native SR was 20 % (Table 16). Human disturbances explained over 50% of the variation in recreational native SR for Classes D, E, and G and over 50% of the variation in non-recreational native SR for Classes A, E, F, and G. Conversely, the most variation human disturbances were able to explain in total native SR was only 38% (Class F). More variation was explained by human disturbances in the Native SR lake classes when the specific type of species was identified for all native fish (recreational or non-recreational), rather than total native SR.

Specifically comparing recreational and non-recreational species within classes, the effects of human disturbances explain recreational native SR better in classes B ($R^2 = 0.43$) and D ($R^2 = 0.71$) than non-recreational native SR (Class B, R^2 = Non significant model; Class D, $R^2 = 0.34$) (Tables 18 and 19). Conversely, the effects of human disturbances explained a greater amount of variation in non-recreational native SR (Class E, $R^2 = 0.81$; Class F, $R^2 = 0.59$) than recreational native SR (Class E, $R^2 = 0.52$; Class F, $R^2 = 0.08$). The same trend could be seen in Class A where native non-recreational SR ($R^2 = 0.55$) was explained much better by disturbances than was native recreational ($R^2 = 0.22$). Because non-recreational species were more prevalent in lake classes than recreational species (Figures 8 through 17), variability was higher within a class for non-recreational species. In this sense, the greater within-class variability of non-recreational species likely made them a better response metric that total native SR or native recreational SR for measuring the effects of human disturbances.

Recreational species dominated individual lakes in almost all classes by frequency; for the most part, recreational species comprised the top three or four most frequent species in a lake. However, non-recreational species far outnumbered recreational species in every state (Table 10), therefore, it was reasonable to expect that more variation would be explained using non-recreational species as a response metric. Non-recreational species were more frequent for an entire class, however, unique taxa of families such as darters, minnows, madtoms, sculpins and shiners were only present in a few lakes; whereas, recreational species were typically the dominant species in almost all lakes, (Figures 8 through 17). In fact, many non-recreational species occurred in only a few lakes within a class, many of which were likely intolerant of human disturbances. Other researchers have shown that human disturbances are likely to affect species intolerant of human-induced stress; these species are therefore strong biotic indicators of anthropogenic disturbance (Drake and Pereira 2002). The presence of these unique species appeared to identify lakes with high biodiversity, a value many lake managers strongly consider when making decisions. Non-recreational species appear to be a key group of fish that managers and researchers could examine within specific lake classes for management in the face of anthropogenic disturbances. Ultimately, they may help to indicate what lakes human-induced stressors are affecting intolerant species, therefore warranting conservation efforts to maintain biodiversity.

Conclusions

The effects of human disturbances on my study lakes were generally consistent with the literature, although they were exhibited at a much broader scale than previous studies. The disturbances that had generally positive associations with fish species richness were road density and human population density. With greater accessibility to lakes (i.e., more roads and people), fishing pressure and stocking rates have been shown to increase, therefore, likely more accidental fish introductions through bait buckets (i.e., fishing) (Kratz et al. 1997, and Magnuson et al. 1998). Dams (local scale and regional scale), agricultural and urban land use/cover, and fish stocking were negatively associated with fish species richness in my study lakes. Dams fragment connectivity of freshwater systems, restricting the movement of fish to naturally colonize a new lake, and cuting off access to stream refuge for fish during winter (hypoxic) conditions. Shoreline development (i.e., increased agriculture and urban land and corresponding loss of forest) results in the loss of habitat (CWD) for fish and likely causes negative impacts through increased nutrient inputs and alterations to the near shore zones (e.g., removal of macrophytes).Stocking of non-native species (either to a lake or even entire state (e.g., brown trout)) may initially increases species richness, although long-term food web dynamics and loss of niche space can end in the loss of other species (i.e., decrease in species richness). The metrics I created using the best, readily-available stocking data do not account for all purposeful species introductions into lakes. For example, I acquired state agency stocking information, and for Michigan and Wisconsin, these data were post-1970. However, stocking is known to have been conducted long before then, and in most states, private stocking programs likely result in introductions for which agencies

have no records. Besides intentional stocking, unintentional introductions (e.g., bait buckets, boat movement) also affect the distribution of species and surely influence the overall distribution of fish species across my five states. Altogether, my estimates of fish stocking were likely conservative and did not capture all forms of human-induced introductions of fish species (native or non-native) into inland lakes.

Many different fish response metrics were considered for analysis, ultimately, I decided on using species richness. Because of the large spatial scale of this study and how much fish distributions vary across all five states, there was high variability in the fish assemblages among all lakes and classes. Because of this, it seemed most practical to use species richness rather than abundance of each species or assemblage type. Species richness ultimately drives conservation approaches to management (e.g., maintaining biodiversity) and helps to account for the distributional differences of fish species that is inherent when researching lakes across 5 states. Non-recreational species appear to vary across and within classes making them more useful for identifying disturbances than recreational species. Species abundance in lakes where recreational fishing is prevalent is often driven by management decisions to maintain opportunities to catch game species. This fact likely compromises the use of recreational species richness as a response metric because each agency has a different fish assemblage to manage and constituents for which to manage. Different recreational values such as bass fishing versus trout fishing makes comparing lake-to-lake species abundances more difficult than a general species count. Managing with biodiversity (species richness) in mind crosses political boundaries and somewhat mitigates differences in species distributions and abundances across such a large region like in this study, therefore, species richness was selected as the sole response metric.

Future research of relationships between anthropogenic disturbances and fish species richness might also benefit from analyses such as general additive models (GAM) to identify such non-linear relationships. For example, I would expect that non-linear relationships might be revealed in fish stocking because adding a new species into a lake has a positive effect on species richness. However, continued addition of new species can alter ecosystem dynamics leading to an overall species richness decrease. Because human disturbances can likely have non-linear relationships with fish assemblages, future research should consider additional analysis to linear regression.

In conclusion, my research highlighted the ability to identify specific effects of humans on inland lake fish assemblages; species richness was positively and negatively associated with a wide range of anthropogenic disturbances. Lake managers could use this information to help mitigate the loss of species and conserve biodiversity; additionally, managers may able to create and improve species richness using other management tools. For the most part, despite the many varying effects (positive and negative) on fish within a class, lakes were dominated primarily by the same three or four species. Therefore, future research and management may be able to identify these dominant species and make inferences as to how speciose a lake might be given the presence or absence of a few key species. This may help to reduce the amount of in-lake sampling is needed to make decisions across a large number of lakes. It was more effective and informative to measure human disturbances in classes of lakes rather than across all lakes; this would also help to make decisions for a large number of lakes while only have resources to collect in-lake data on a subset.

Thesis Conclusions

In Chapter 1, I successfully developed lake classifications for total and native species richness using hyrdogeomorphic features of the landscape at multiple spatial scales. In the future, managers can use classifications such as these to assign lakes for which only hydrogeomorphic data are available to classes. Class-specific management options can then be pursued by managers, allowing them to incorporate the natural variation among lakes even when it is difficult to obtain in-lake data. Even though individual lakes are unique in some respects, they share similar characteristics due to having similar local and even regional characteristics with other lakes which collectively shape fish communities. My research and lake classifications were not necessarily intended to compare results with other studies (indeed the variety of approaches and variables used across studies challenges statistical comparisons of classification systems), but rather to explore previously overlooked predictor metrics and their ability to explain variation in fish assemblages in a classification.

My study was unique in that I incorporated multiple states, each with a wide ranging set of natural features and human disturbances. Regional-scale predictors, which have been incorporated into classification efforts relatively recently and lack thorough evaluation, were useful in explaining variation in species richness, and valuable for creating disturbance metrics such as the number of dams per EDU. Even though the primary splits in all classifications were local features, regional-scale predictors also were included in several of the classifications, highlighting the importance of regional-scale predictors.

To investigate the effects of human disturbances on inland lake fish assemblages, it was more useful to quantify these relations within classes than across all lakes. I found that the effects of human disturbances such as dams (local scale and regional scale), agricultural land use and urban land use, and stocking all are associated with lower species richness. In contrast, road density and human population density generally were associated with higher species richness. Species richness was a useful response metric for studying lakes across five states and native non-recreational species richness, in particular, appeared to be the best indicator of human disturbances on lakes of those I considered. Non-recreational species were the most prevalent by frequency in all classes; presence and absence of unique taxa therefore helped to identify patterns of biodiversity. Species richness should be a concern of management agencies for conservation goals, even while working to ensure the presence of recreational species for sport fisheries. In conclusion, lake classifications should be considered by state agencies as an approach to managing a large number of lakes while having information on a subset. By dividing lakes into classes using map-based characteristics, lake managers may be able to: (a) appropriate resources otherwise used for rigorous and expensive sampling programs to other projects, and (b) make somewhat informed decisions about the management of individual lakes, even when in-lake data are lacking.

My research highlights the importance of within and among management agency collaboration, or lack there of. I was able to collect natural landscape and human disturbance data on more than 2,300 lakes. In contrast, fish assemblage data were only available for approximately 550 lakes across all five states, reflecting the fact that fish assemblage data are more difficult to gather than map-based features. Even within a single state, agencies had non-overlapping monitoring programs, meaning that agencies collecting water quality data did not coordinate with agencies collecting fish data in order to sample the same lakes. As a result, nearly 200 lakes of my 550 lakes with fish assemblage data lacked the other data that I needed for my analysis, ultimately because they had not been sampled by the water quality management agency. From the standpoint of developing classification models such as those described here, coordinated monitoring efforts across agencies would be extremely beneficial. Efforts such as the Environmental Protection Agencies' National Lakes Assessment help align multi-state and even nationwide collaboration and management, which I view as a step in the right direction.

Appendix

Tables

Table 1. The number of lakes in each state for which data were obtained according to variable category. 'Overlap' indicates the lakes used in analysis because complete data were available.

Variable Category	ME	NH	MI	WI	IA	TOTAL
Hydrogeomorphic	569	740	486	426	127	2348
Anthropogenic	658	725	637	426	127	2573
Fish Assemblage	65	31	125	219	115	555
Overlap	**28**	**27**	**105**	**90**	**110**	**360**

Table 2. List of hydrogeomorphic variables (regionalization, extinction/isolation, hydrology, morphometry, land use/cover) and their measurement units arranged according to category. NA indicates a non-continuous variable and therefore does not have a median, minimum, or maximum. Land use/cover coverage's were calculated within a 500-m buffer of the lake.

Category/Name	Median	Minimum	Maximum
Regionalization			
6-Digit Hydrologic Unit Code (HUC)	97 Total Regions		
Ecological Drainage Unit (EDU)	28 Total Regions		
Freshwater Ecoregion	5 Total Regions		
Isolation			
Lake Glaciated or Not Glaciated	NA	NA	NA
Lake Seepage or Drainage	NA	NA	NA
# of Lakes within an EDU	1,549	159	2,856
# of Lakes within a 12-digit HUC	9	1	89
# of Surface Water Connections	2	0	37
Lake Elevation (m)	303	17	570
Extinction			
Lake Area (ha)	84.4	3.3	6,039.1
Lake Maximum Depth (m)	8.5	1.2	40.8
Lake Mean Depth (m)	3.1	0.8	12.2
Lake Perimeter (m)	6,444	722	185,962
Hydrology			
Mean Base Flow (%)	57.0	14.0	89.0
Mean Precipitation (cm/yr)	84.7	65.4	124.3
Mean Runoff (cm/yr)	27.9	5.1	76.2
Watershed Cumulative Catchment Area (km^2)	11	0.1	31,080
Land Use/Cover (%)			
Water	2.4	0	35.2
Forest	49.8	.6	99.7
Grassland/Field	1.0	0.0	36.4
Wetland Woody	3.9	0.0	71.8
Wetland Herbaceous	1.4	0.0	16.3
Wetland Combined	5.7	0.0	74.9
All Other	0.0	0.0	11.0

Table 3. List of anthropogenic disturbance variables and their measurement units (land use/cover, dams, roads, fish stocking, human populations) arranged according to category. Land use/cover coverage's were calculated within a 500-m buffer of the lake.

Category/Name	Median	Minimum	Maximum
Land Use/Cover (%)			
Urban	1.2	0.0	66.1
Agriculture Pasture	6.0	0.0	72.9
Agriculture Row Crop	10.2	0.0	89.8
Agriculture Other	0.0	0.0	33.5
Agriculture Combined	20.1	0.0	94.7
Agriculture and Urban	25.6	0.0	95.9
Dams			
# of Dams within an EDU	558	48	3,634
# of Dams within a 12-digit HUC	2	0	59
Roads			
Road Area in a 500 m Buffer (ft/acre)	38.7	0.0	141.7
Fish Stocking			
# of Fish Species Stocked per Lake	1	0	22
# of Fish Stocking Events per Lake	14.5	0	393
# of Fish Stocking Events per Lake/Years of Data	0.3	0	5.1
Human Population			
Human Population Density in 2000	1,083	18	68,636
Housing Density in 2000	816	34	29,749

Table 4. Summary of fish assemblage data by agency, collection gear, range of years sampled across all lakes, purpose of agency's fish data collection, total number of lakes in fish assemblage data (555), and number of lakes to be used in analysis (360). Agencies include the EPA's Ecological Monitoring and Assessment Program (EMAP), the Michigan and Wisconsin Departments of Natural Resources (DNR), and Iowa State University.

State	Agency	Gears Used	Sampling Years	Purpose and Goals of Fish Data Collection	Total Lakes	Analysis Lakes
ME	EMAP	Gill and Trap Nets	1991-1996	Monitoring and assessing the status and trends of national ecological resources	65	28
NH	EMAP	Gill and Trap Nets	1991-1996	Monitoring and assessing the status and trends of national ecological resources	31	27
MI	State DNR	Fyke Nets and Electroshocking	2002-2006	Baseline Monitoring Program	125	105
WI	State DNR	Fyke Nets and Electroshocking	2001-2005	Baseline Monitoring Program	219	90
IA	Iowa State University	Fyke Nets and Electroshocking	2001-2004	Monitoring state-wide lake resources	115	110
Sum					**555**	**360**

Table 5. List of fish assemblage response metrics used in building natural lake classes (Total SR and Native SR). Values of each metric are for individual lakes but displayed for the 5-state region and each individual state.

Response Metric	Median	Minimum	Maximum
5-State			
Total SR	11	1	26
Native SR	11	1	25
Maine			
Total SR	9	1	16
Native SR	9	1	16
New Hampshire			
Total SR	8	3	13
Native SR	6	2	10
Michigan			
Total SR	15	3	26
Native SR	14	2	24
Wisconsin			
Total SR	14	5	25
Native SR	14	5	25
Iowa			
Total SR	8	4	21
Native SR	7	3	19

Table 6. Summary of fish stocking by state agency (Maine-Department of Inland Fisheries and Wildlife (DIFW), New Hampshire-Fish and Game Department (FGD), Michigan-Department of Natural Resources (DNR), Wisconsin-Department of Natural Resources (DNR), and Iowa-Department of Natural Resources (DNR)), range of years across all lakes, statistics (median, minimum, and maximum) of the number of species stocked per lake, and the total number of species stocked across all lakes in a state.

			No. Sp. Stocked per Lake			Total No. Sp.
State	**Agency**	**Years in Data**	**Median**	**Min.**	**Max.**	**Stocked in All Lakes**
ME	DIFW	1937-1996	2	0	9	10
NH	FGD	1926-1996	2	0	8	13
MI	DNR	1979-2006	1	0	7	20
WI	DNR	1972-2005	2	0	5	16
IA	DNR	1928-2004	1	0	22	27

Table 7. Final multiple regression model results for each response metric versus human disturbance variables (n = 360 lakes for all analyses). All variables were significant at p < .05.

Response Variable	Adjusted R-Squared	Parameter Estimate	Standard Error
Total SR	0.16		
Number of Fish Species Stocked		-0.19	0.10
Number of Fish Stocking Events/Years in Data		0.21	0.10
Number of Dams per 12-Digit HUC		-0.10	0.03
Agricultural Land Use		-0.29	0.10
Human Population Density		0.40	0.06
Native SR	0.22		
Number of Fish Species Stocked		-0.18	0.10
Number of Stocking Events/Years in Data		0.22	0.10
Number of Dams per 12-Digit HUC		-0.14	0.03
Agricultural Land Use (Pasture)		1.14	0.30
Agricultural Land Use (Crop)		0.58	0.30
All Agricultural and Urban Land Use Combined		-1.30	0.28
Human Population Density		0.52	0.07

Table 8. Pearson correlation coefficients for those predictor variables with an r > 0.60. Correlations are grouped by major category; anthropogenic disturbances only, hydrogeomorphic features only, and anthropogenic disturbances and hydrogeomorphic features combined.

Variable	r
Anthropogenic Disturbances	
Agriculture (All) Land Use vs Agriculture and Urban Land Use	0.96
Number of Stocking Events vs Number of Stocking Events/Years	0.95
Agriculture (Crop) Land Use vs Agriculture (All) Land Use	0.92
Number of Species Stocked vs Number of Stocking Events	0.90
Agriculture (Pasture) Land Use vs Agriculture (All) Land Use	0.89
Human Population Density 2000 vs Housing Density 2000	0.88
Agriculture (Crop) Land Use vs Agriculture and Urban Land Use	0.87
Number of Species Stocked vs Number of Stocking Events/Years	0.85
Agriculture (Pasture) Land Use vs Agriculture and Urban Land Use	0.82
Agriculture (Pasture) Land Use vs Agriculture (Crop) Land Use	0.67
Urban Land Use vs Agriculture (Other) Land Use	0.63
Hydrogeomorphic Features	
Wetland (Woody) Cover vs Wetland (All) Cover	0.97
Lake Area vs Lake Perimeter	0.92
Lake Max Depth vs Lake Mean Depth	0.88
Forest Land Cover vs Mean Runoff	0.79
Lake Perimeter vs Cumulative Watershed Area	0.74
Lake Area vs Cumulative Watershed Area	0.70
Number of Surface Connections vs. Cumulative Watershed Area	0.69
Anthropogenic Disturbances and Hydrogeomorphic Features	
Forest Land Cover vs Agriculture and Urban Land Use	-0.89
Forest Land Cover vs Agriculture (All) Land Use	-0.85
Mean Runoff vs Agriculture (All) Land Use	-0.79
Mean Runoff vs Agriculture and Urban Land Use	-0.78
Forest Land Cover vs Agriculture (Crop) Land Use	-0.76
Mean Runoff vs Agriculture (Pasture) Land Use	-0.73
Forest Land Cover vs Agriculture (Pasture) Land Use	-0.72
Mean Runoff vs Agriculture (Crop) Land Use	-0.68

Table 9. Summary of Total SR and Native SR lake classifications developed using a range of CP values; listed is the number of classes, the range of lakes per class, total PRE value, ANOVA F value, and ANOVA P-value. The last column indicates the percentage of significant class-by-class pair-wise comparisons.

Total SR

CP	# of Classes	# of Lakes	PRE	ANOVA F	ANOVA P	Pair-wise Comparisons (%)
0.030	6	20-151	0.50	72.4	<.0001	93.3
0.025	7	20-117	0.53	64.8	<.0001	90.4
0.020	9	12-105	0.57	57.1	<.0001	75.0
0.015	11	7-105	0.60	51.3	<.0001	74.5
0.010	16	7-66	0.67	43.2	<.0001	62.5

Native SR

CP	# of Classes	# of Lakes	PRE	ANOVA F	ANOVA P	Pair-wise Comparisons (%)
0.030	6	27-98	0.49	86.6	<.0001	86.6
0.025	6	27-98	0.49	86.6	<.0001	86.6
0.020	8	10-82	0.54	70.6	<.0001	85.7
0.015	10	10-69	0.58	64.8	<.0001	77.7
0.010	14	10-46	0.63	48.7	<.0001	67.0

Table 10. Summary of the total number of native fish species by state with all native separated into species that are native recreational and native non-recreational.

State	Total Native	Native Recreational	Native Non-Recreational
Maine	34	6	28
New Hampshire	19	7	12
Michigan	78	18	60
Wisconsin	86	17	69
Iowa	32	17	15

Table 11. List of fish assemblage response metrics used in measuring anthropogenic disturbance across natural lake classes (Native SR). Values of each metric are for individual lakes but displayed across all lakes, small lakes only (< 53 ha), and large lakes only (> 53 ha).

Response Metric	Median	Minimum	Maximum
All Lakes (360)			
Total Native SR	11	1	26
Small Lakes (133)			
Total Native SR	7	1	21
Large Lakes (227)			
Total Native SR	13	3	25

Table 12. List of fish assemblage response metrics used in measuring anthropogenic disturbance across natural lake classes (Native SR). Values of each metric are for individual lakes but displayed across classes (A, B, C, and D) of small lakes only (< 53 ha).

Response Metric	Median	Minimum	Maximum
Class A Lakes (31)			
Total Native SR	5	1	10
Native Recreational SR	2	1	8
Native Non-Recreational SR	2	0	6
Class B Lakes (27)			
Total Native SR	5	3	11
Native Recreational SR	5	1	8
Native Non-Recreational SR	0	0	6
Class C Lakes (57)			
Total Native SR	8	2	16
Native Recreational SR	6	1	10
Native Non-Recreational SR	2	0	9
Class D Lakes (18)			
Total Native SR	13	7	21
Native Recreational SR	7	5	12
Native Non-Recreational SR	5	0	14

Table 13. List of fish assemblage response metrics used in measuring anthropogenic disturbance across natural lake classes (Native SR). Values of each metric are for individual lakes but displayed across classes (E, F, G, H, I, and J) of large lakes only (> 53 ha).

Response Metric	Median	Minimum	Maximum
Class E Lakes (16)			
Total Native SR	8	3	12
Native Recreational SR	5	1	5
Native Non-Recreational SR	2	2	9
Class F Lakes (36)			
Total Native SR	10	6	15
Native Recreational SR	7	5	9
Native Non-Recreational SR	3	0	8
Class G Lakes (46)			
Total Native SR	11	6	20
Native Recreational SR	7	1	13
Native Non-Recreational SR	6	0	13
Class H Lakes (50)			
Total Native SR	13	5	20
Native Recreational SR	7	3	11
Native Non-Recreational SR	6	0	12
Class I Lakes (69)			
Total Native SR	17	7	25
Native Recreational SR	8	4	12
Native Non-Recreational SR	8	2	16
Class J Lakes (10)			
Total Native SR	21	16	24
Native Recreational SR	8	7	12
Native Non-Recreational SR	11	8	16

Table 14. List of anthropogenic disturbance variables and their measurement units (land use/cover, dams, roads, fish stocking, human populations) used in multiple regression analysis, data displayed here are for small lakes only (<53 ha), Native SR classes A, B, C, and D. Land use/cover coverage's were calculated within a 500-m buffer of the lake.

Category/Name	Median	Minimum	Maximum
Land Use/Cover (%)			
Urban	1.3	0.0	66.1
All Agriculture	33.8	0.0	94.7
Agriculture and Urban	47.1	0.0	95.9
Dams			
# of Dams within an EDU	615	48	3,634
# of Dams within a 12-digit HUC	2	0	59
Roads			
Road Area in a 500 m Buffer (ft/acre)	35.8	0.0	141.7
Fish Stocking			
# of Fish Species Stocked per Lake	1	0	6
Human Population			
Human Population Density in 2000	1,075	25	68,636

Table 15. List of anthropogenic disturbance variables and their measurement units (land use/cover, dams, roads, fish stocking, human populations) used in multiple regression analysis, data displayed here are for large lakes only (>53 ha), Native SR classes E, F, G, H, I, and J. Land use/cover coverage's were calculated within a 500-m buffer of the lake.

Category/Name	Median	Minimum	Maximum
Land Use/Cover (%)			
Urban	1.1	0.0	60.0
All Agriculture	15.0	0.0	88.8
Agriculture and Urban	17.5	0.0	89.7
Dams			
# of Dams within an EDU	558	51	3,634
# of Dams within a 12-digit HUC	2	0	46
Roads			
Road Area in a 500 m Buffer (ft/acre)	41.3	0.3	131.9
Fish Stocking			
# of Fish Species Stocked per Lake	1	0	22
Human Population			
Human Population Density in 2000	1,104	18	64,804

Table 16. Final multiple regression model results for total native species richness versus human disturbance variables within Native SR lake classes (Chapter 1) (n = 360 lakes for all analyses). Displayed within each class (column) is the parameter estimate and standard error in parenthesis. All variables were significant at $p < .05$, total weighted $R^2 = 0.20$.

Total Native Species Richness

	A(31)	B(27)	C(57)	D(18)	E(16)	F(36)	G(46)	H(50)	I(69)	J(10)
Number of Species Stocked							-.019 (0.09)			
Dams per EDU					-0.18 (0.06)					
Dams per 12-digit HUC			-0.11 (0.03)							
All Agriculture								-0.49 (0.18)		
Urban									-1.02 (0.36)	
Agriculture & Urban Combined			-0.37 (0.14)			-0.61 (0.13)	-0.51 (0.15)			
Road Area										
Human Population Density	0.46 (0.12)		0.22 (0.10)	0.43 (0.18)			0.43 (0.13)	0.44 (0.13)	0.48 (0.12)	
R^2	**0.31**	**XXX**	**0.26**	**0.21**	**0.35**	**0.38**	**0.37**	**0.22**	**0.17**	**XXX**

Table 17. Final multiple regression model results for total native species richness versus human disturbance variables within small lakes (< 53 ha) and large lakes (> 53 ha) (n = 360 lakes for all analyses). Displayed within each class (column) is the parameter estimate and standard error in parenthesis. All variables were significant at $p < .05$, total weighted $R^2 = 0.15$.

Total Native Species Richness

	Small Lakes (133)	Large Lakes (227)
Number of Species Stocked		-0.11 (0.06)
Dams per EDU		
Dams per 12-digit HUC	-0.12 (0.03)	-0.08 (0.03)
All Agriculture	-0.25 (0.12)	
Urban		
Agriculture & Urban Combined		
Road Area	0.11 (0.05)	0.14 (0.05)
Human Population Density	0.28 (0.09)	0.19 (0.09)
R^2	**0.21**	**0.12**

Table 18. Final multiple regression model results for total native recreational species richness versus human disturbance variables within Native SR lake classes (Chapter 1) (n = 360 lakes for all analyses). Displayed within each class (column) is the parameter estimate and standard error in parenthesis. All variables were significant at $p < .05$, total weighted $R^2 = 0.25$.

Native Recreational Species Richness

	A(31)	B(27)	C(57)	D(18)	E(16)	F(36)	G(46)	H(50)	I(69)	J(10)
Number of Species Stocked							-.035 (0.10)			
Dams per EDU	-0.18 (0.06			0.46 (0.11)						
Dams per 12-digit HUC		-0.13 (0.04)	-0.08 (0.03)	-0.11 (0.05)	0.20 (0.05)					
All Agriculture						0.21 (0.10)	0.82 (0.18)			
Urban			0.70 (0.27)							
Agriculture & Urban Combined		0.48 (0.19)			-2.43 (0.75)					
Road Area				0.17 (0.04)	0.39 (0.14)					
Human Population Density	0.51 (0.18)						0.45 (0.12)	0.24 (0.09)		
R^2	**0.22**	**0.43**	**0.18**	**0.71**	**0.52**	**0.08**	**0.63**	**0.10**	**XXX**	**XXX**

Table 19. Final multiple regression model results for total native non-recreational species richness versus human disturbance variables within Native SR lake classes (Chapter 1) (n = 360 lakes for all analyses). Displayed within each class (column) is the parameter estimate and standard error in parenthesis. All variables were significant at $p < .05$, total weighted $R^2 = 0.33$.

Native Non-Recreational Species Richness

	A(31)	B(27)	C(57)	D(18)	E(16)	F(36)	G(46)	H(50)	I(69)	J(10)
Number of Species Stocked										
Dams per EDU					-0.36 (0.06)					
Dams per 12-digit HUC						-0.20 (0.09)		0.31 (0.08)		
All Agriculture	-1.08 (0.56)						-1.57 (0.23)			
Urban				-7.94 (2.42)	-1.60 (0.61)		-1.51 (0.62)		-1.28 (0.44)	
Agriculture & Urban Combined			-1.25 (0.23)		1.74 (0.56)	-1.65 (0.24)		-0.95 (0.22)		
Road Area				1.21 (0.38)						
Human Population Density	0.56 (0.12)						0.42 (0.16)		0.53 (0.16)	
R^2	**0.55**	**XXX**	**0.35**	**0.34**	**0.81**	**0.59**	**0.63**	**0.36**	**0.14**	**XXX**

Figures

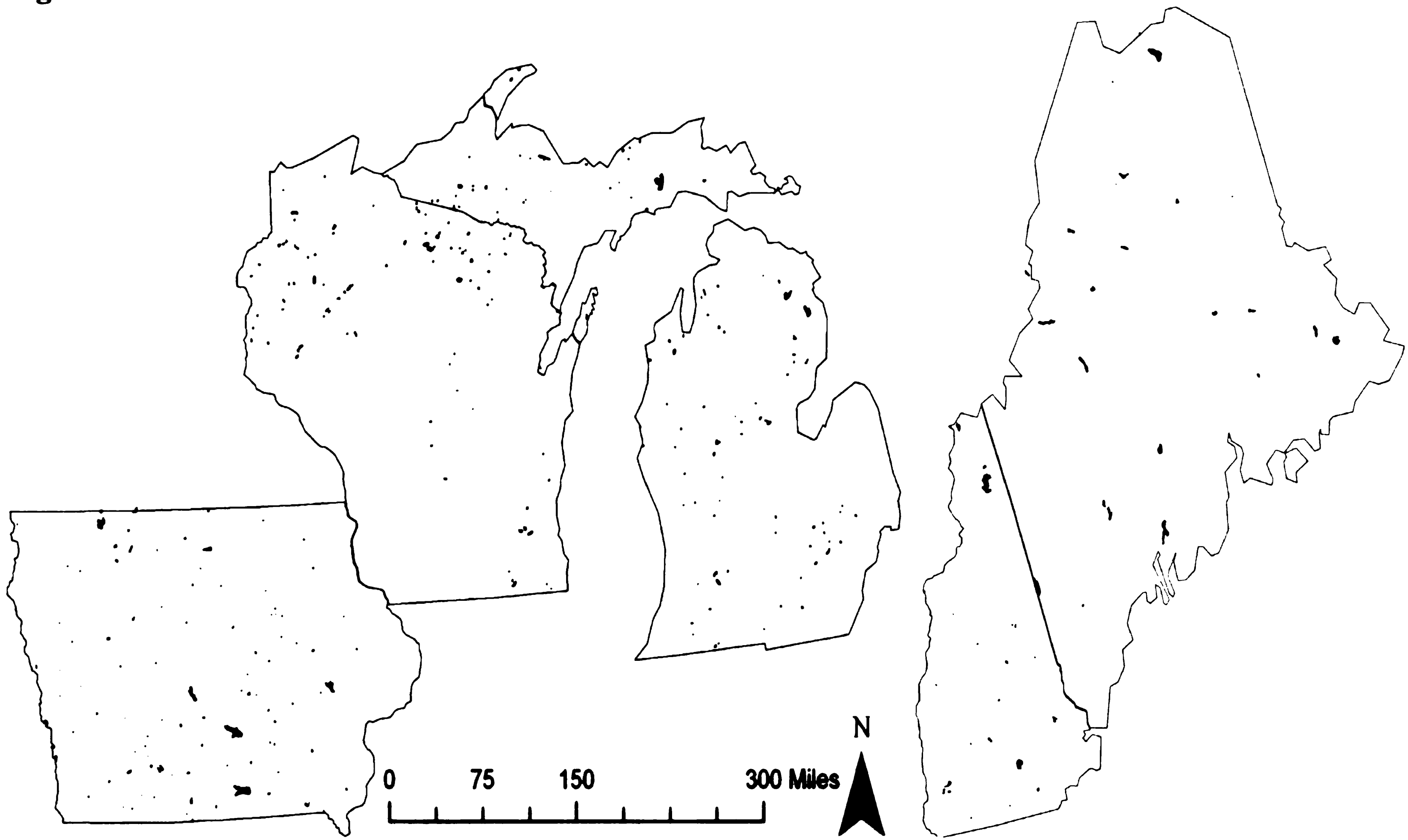

Figure 1. Map of the 360 lakes in Maine (28) New Hampshire (27), Michigan (105), Wisconsin (90), and Iowa (110) used to classify fish assemblage metrics using natural hydrogeomorphic features.

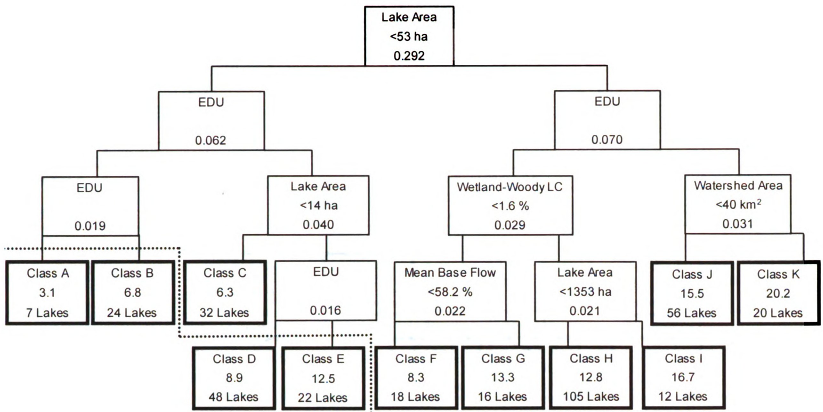

Figure 2. Total SR classification tree with 11 final classes (bold boxes) using a CP of 0.015, resulting in a PRE value of 0.602. In each subgroup (non-bold boxes), the splitting variable, cutoff value, and individual PRE value are listed and in each class type (bold boxes), the mean response value and number of lakes is listed. Boxes above dashed line indicate final classes (9) using a CP of 0.02, resulting in a PRE value of 0.568.

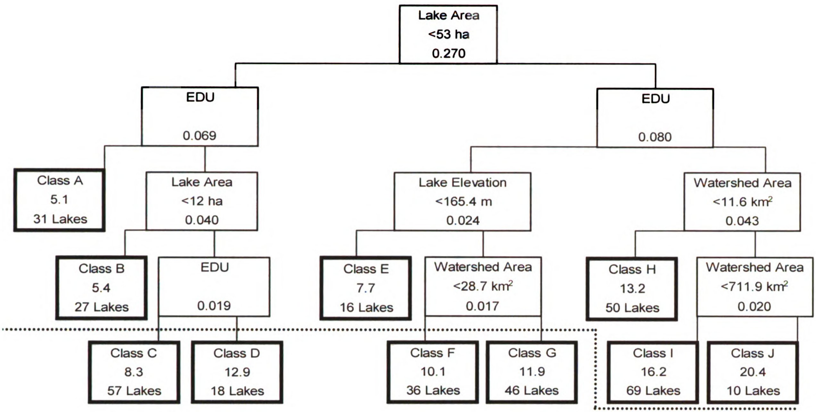

Figure 3. Native SR classification tree with 10 final classes (bold boxes) using a CP of 0.015, resulting in a PRE value of 0.575. In each subgroup (non-bold boxes), the splitting variable, cutoff value, and individual PRE value are listed and in each class type (bold boxes), the mean response value and number of lakes is listed. Boxes above dashed line indicate final classes (7) using a CP of 0.02, resulting in a PRE value of 0.538.

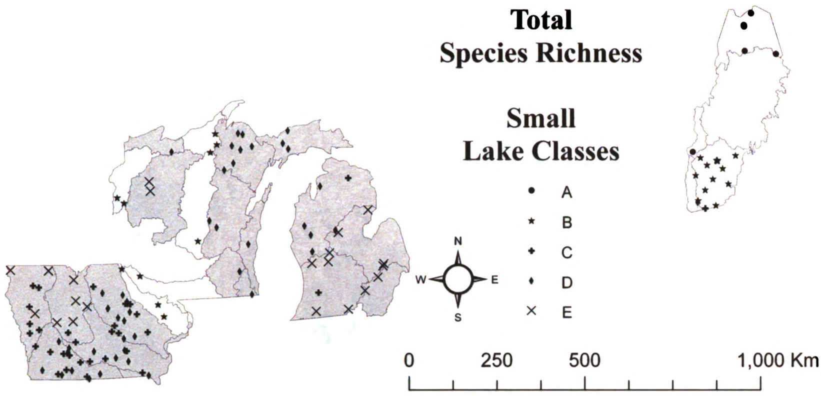

Figure 4. Spatial distribution of Total SR small lake classes (< 53 ha) (A-7 lakes, B-24 lakes, C-32 lakes, D-48 lakes, E-22 lakes) shown with the secondary split (EDU) of the small lakes subgroup indicated by the white (9 regions, classes A and B) and grey (18 regions, classes C, D, and E). See Figure 2 for Total SR mean values and lake class descriptions.

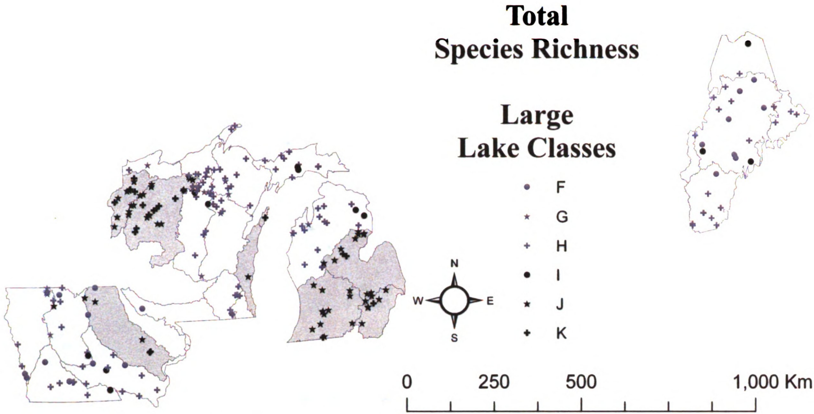

Figure 5. Spatial distribution of Total SR large lake classes (>53 ha) (F-18 lakes, G-16 lakes, H-105 lakes, I-12 lakes, J-56 lakes, and K-20 lakes) shown with the secondary split (EDU) of the large lakes subgroup indicated by the white (19 regions, classes F, G, H, and I) and grey (7 regions, classes J and K). See Figure 2 for Total SR mean values and lake class descriptions.

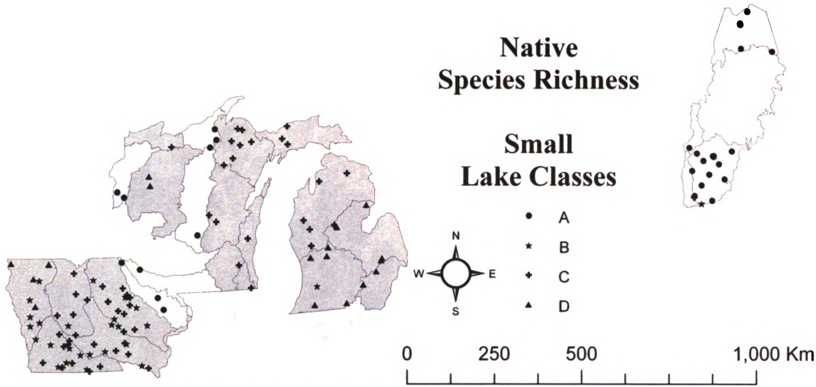

Figure 6. Spatial distribution of Native SR small lake classes (<53 ha) (A-31 lakes, B-27 lakes, C-57 lakes, D-18 lakes) shown with the secondary split (EDU) of the small lakes subgroup indicated by the white (9 regions, class A) and grey (18 regions, classes B, C, and D). See Figure 3 for Native SR mean values and lake class descriptions.

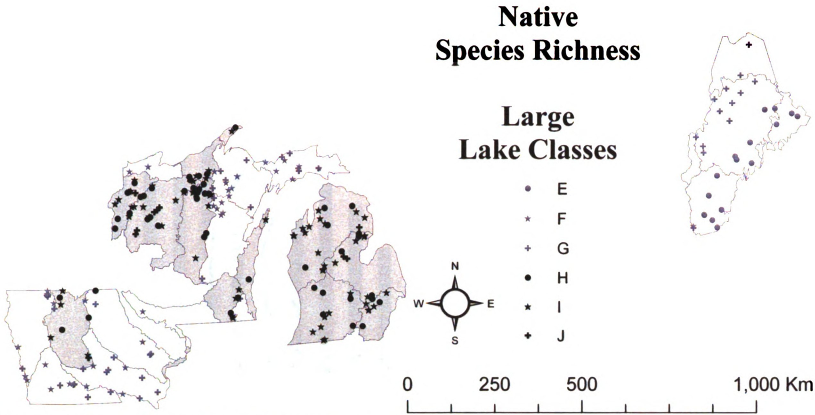

Figure 7. Spatial distribution of Native SR large lake classes (>53 ha) (E-16 lakes, F-10 lakes, G-46 lakes, H-50 lakes, I-69 lakes, J-10) shown with the secondary split (EDU) of the larges lake subgroup indicated by the white (15 regions, classes E, F, and G) and grey (11 regions, classes H, I, and J). See Figure 3 for Native SR mean values and lake class descriptions.

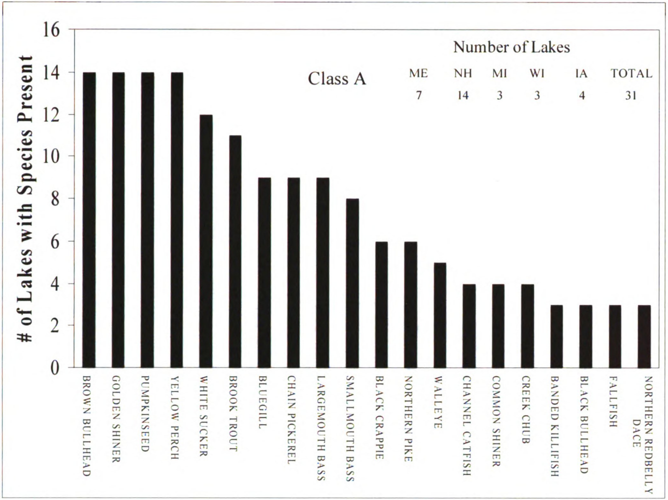

Figure 8. Number of species present in each class across all lakes and states displayed to show the predominate fish assemblages in a class. Four species were present in two lakes and 10 species were present in one lake.

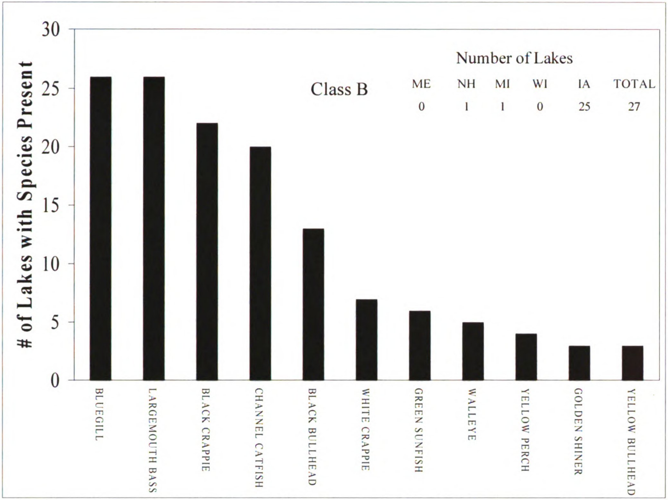

Figure 9. Number of species present in each class across all lakes and states displayed to show the predominate assemblages in a class. One species were present in two lakes and 11 species were present in one lake.

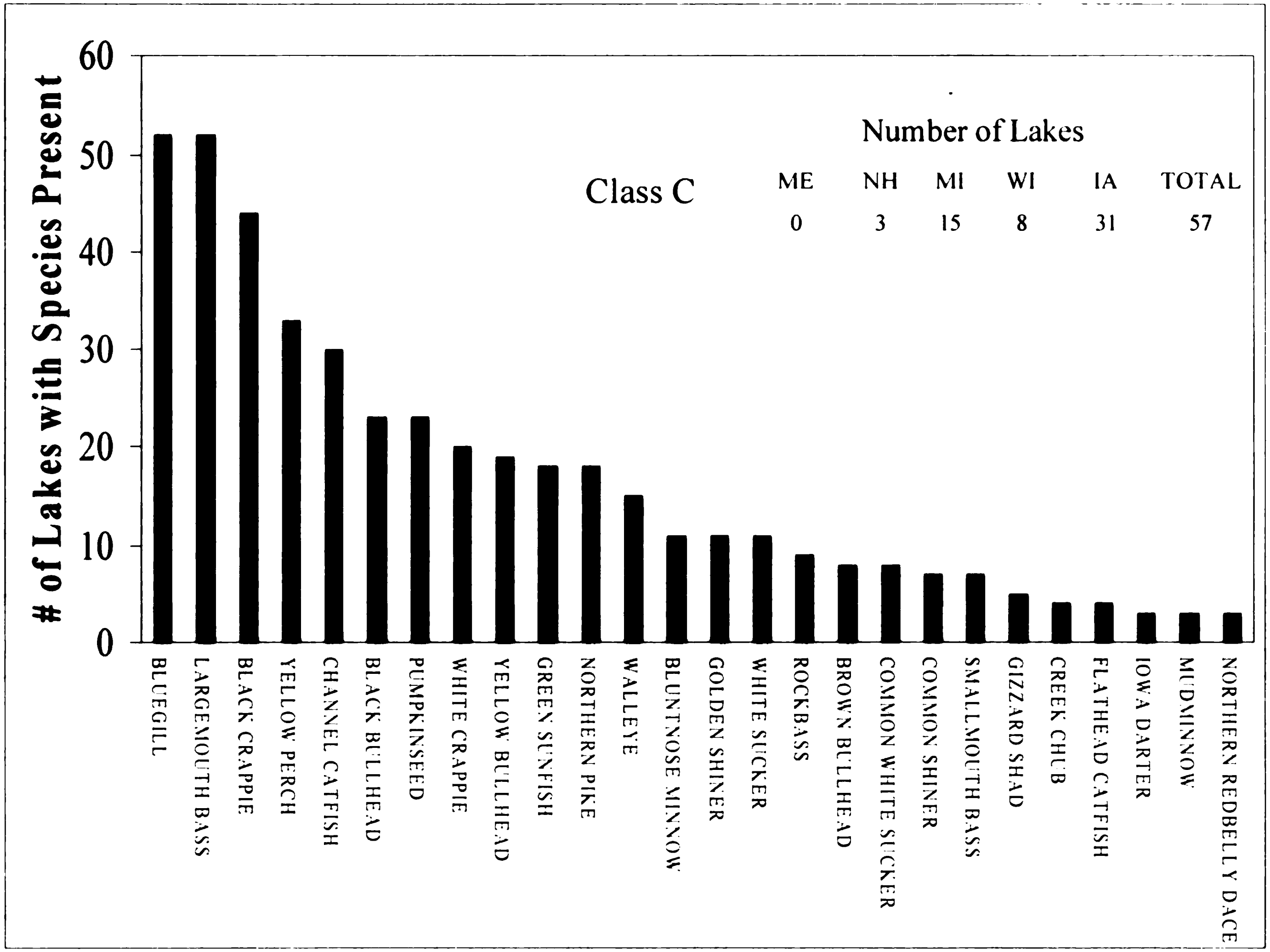

Figure 10. Number of species present in each class across all lakes and states displayed to show the predominate assemblages in a class. 10 species were present in two lakes and 20 species were present in one lake.

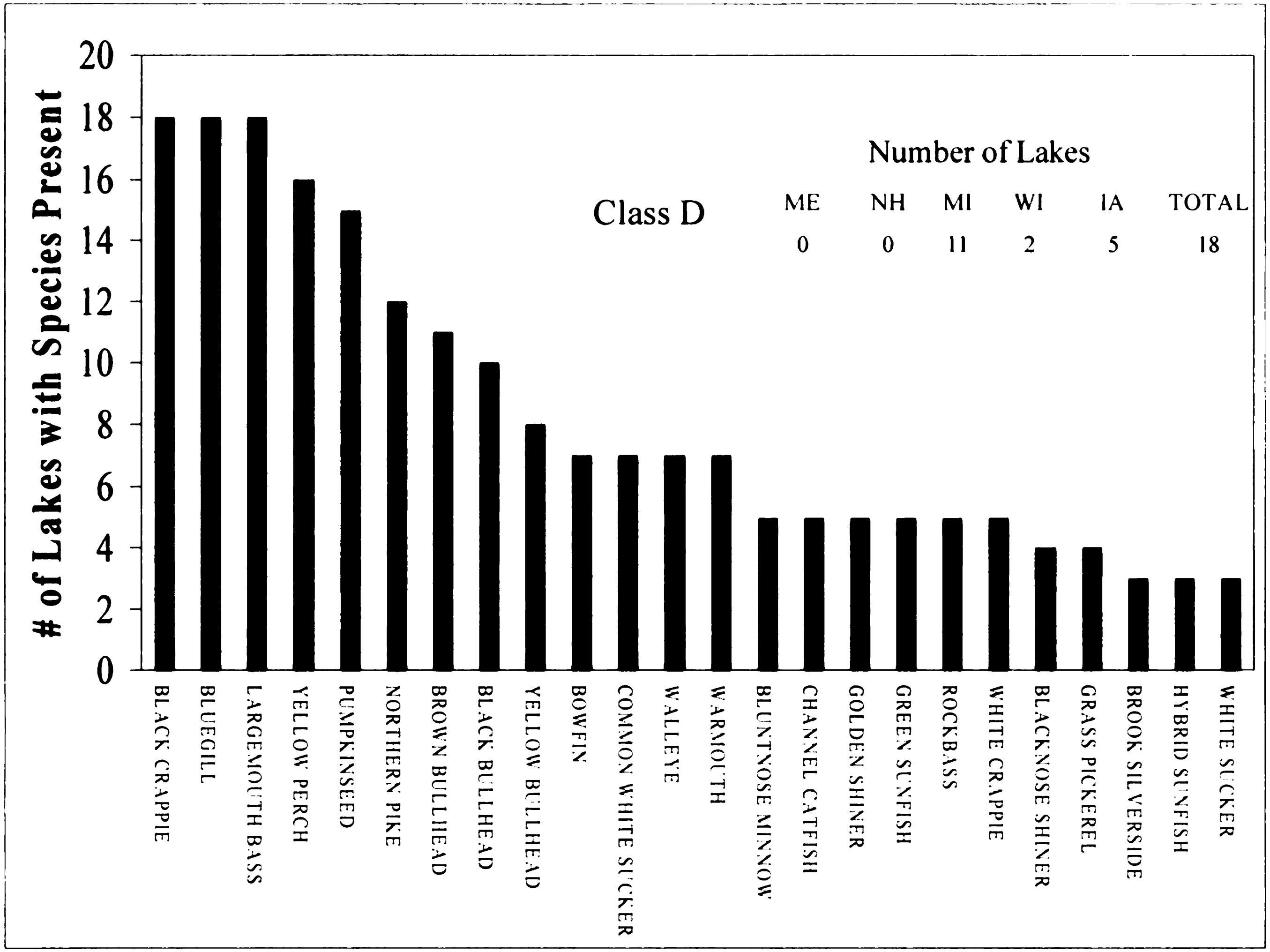

Figure 11. Number of species present in each class across all lakes and states displayed to show the predominate assemblages in a class. Nine species were present in two lakes and 13 species were present in one lake.

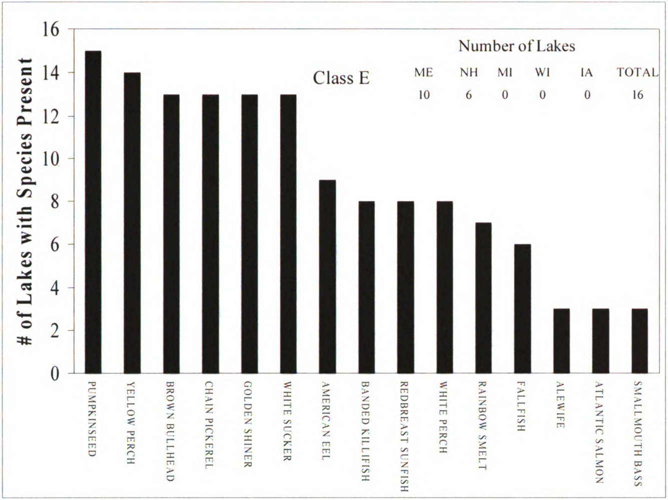

Figure 12. Number of species present in each class across all lakes and states displayed to show the predominate assemblages in a class. Two species were present in two lakes and 10 species were present in one lake.

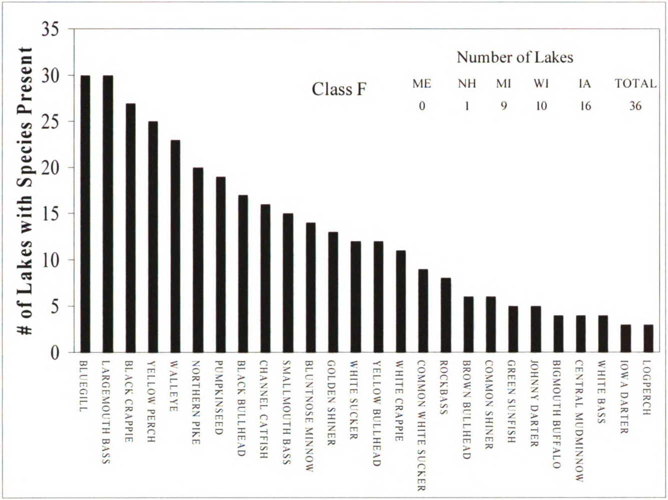

Figure 13. Number of species present in each class across all lakes and states displayed to show the predominate assemblages in a class. Six species were present in two lakes and 17 species were present in one lake.

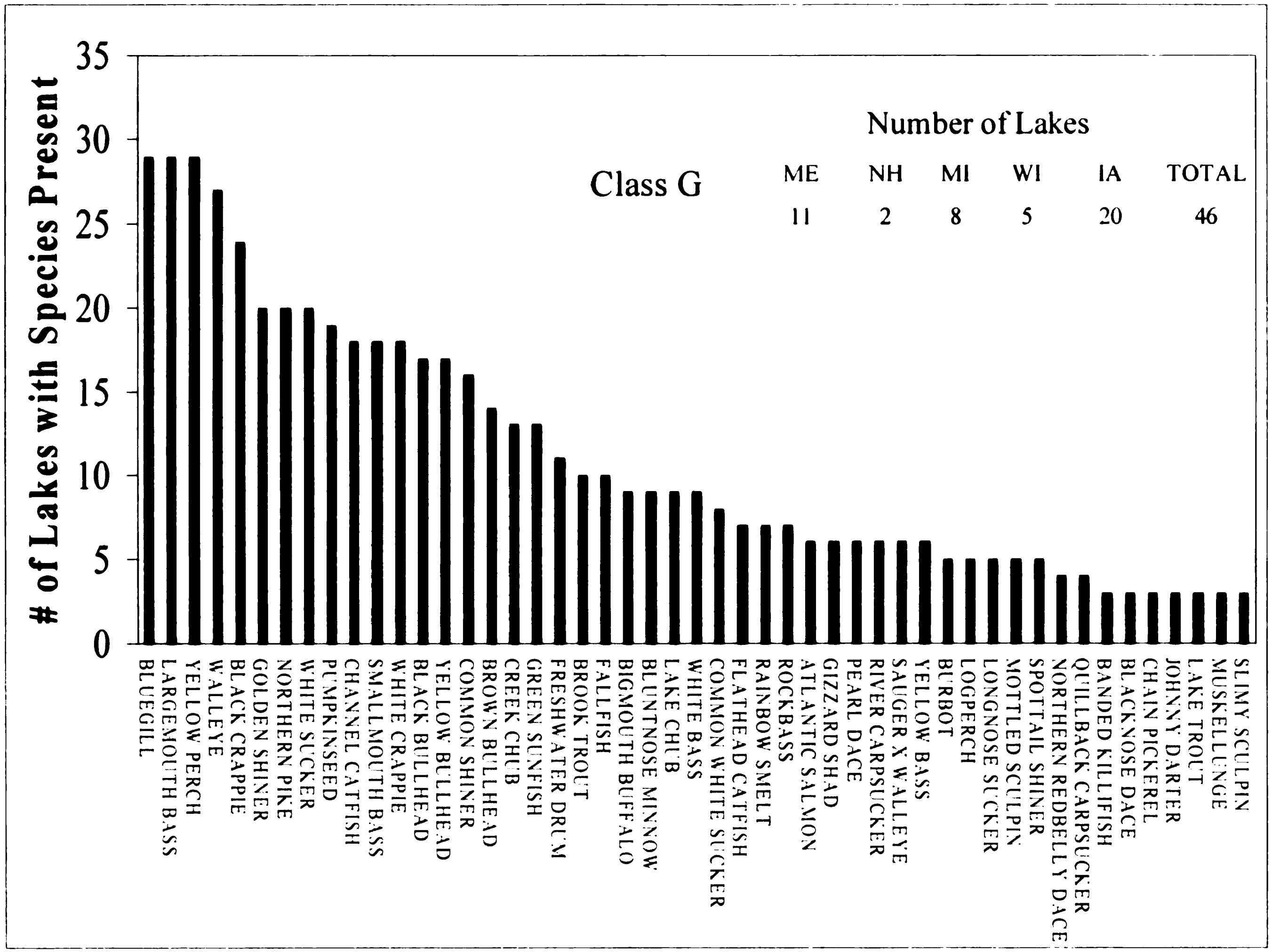

Figure 14. Number of species present in each class across all lakes and states displayed to show the predominate assemblages in a class. 11 species were present in two lakes and 18 species were present in one lake.

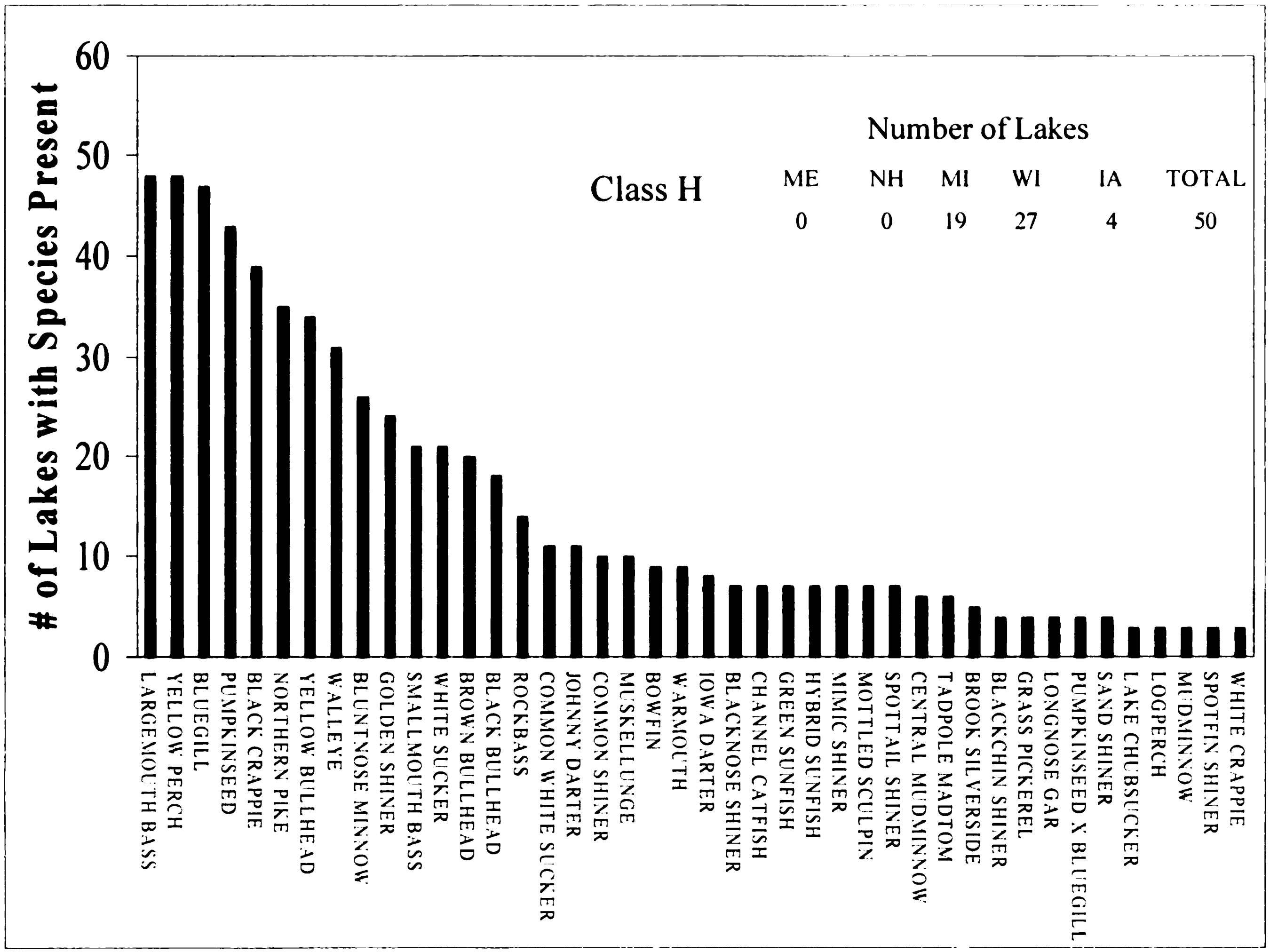

Figure 15. Number of species present in each class across all lakes and states displayed to show the predominate assemblages in a class. Six species were present in two lakes and 12 species were present in one lake.

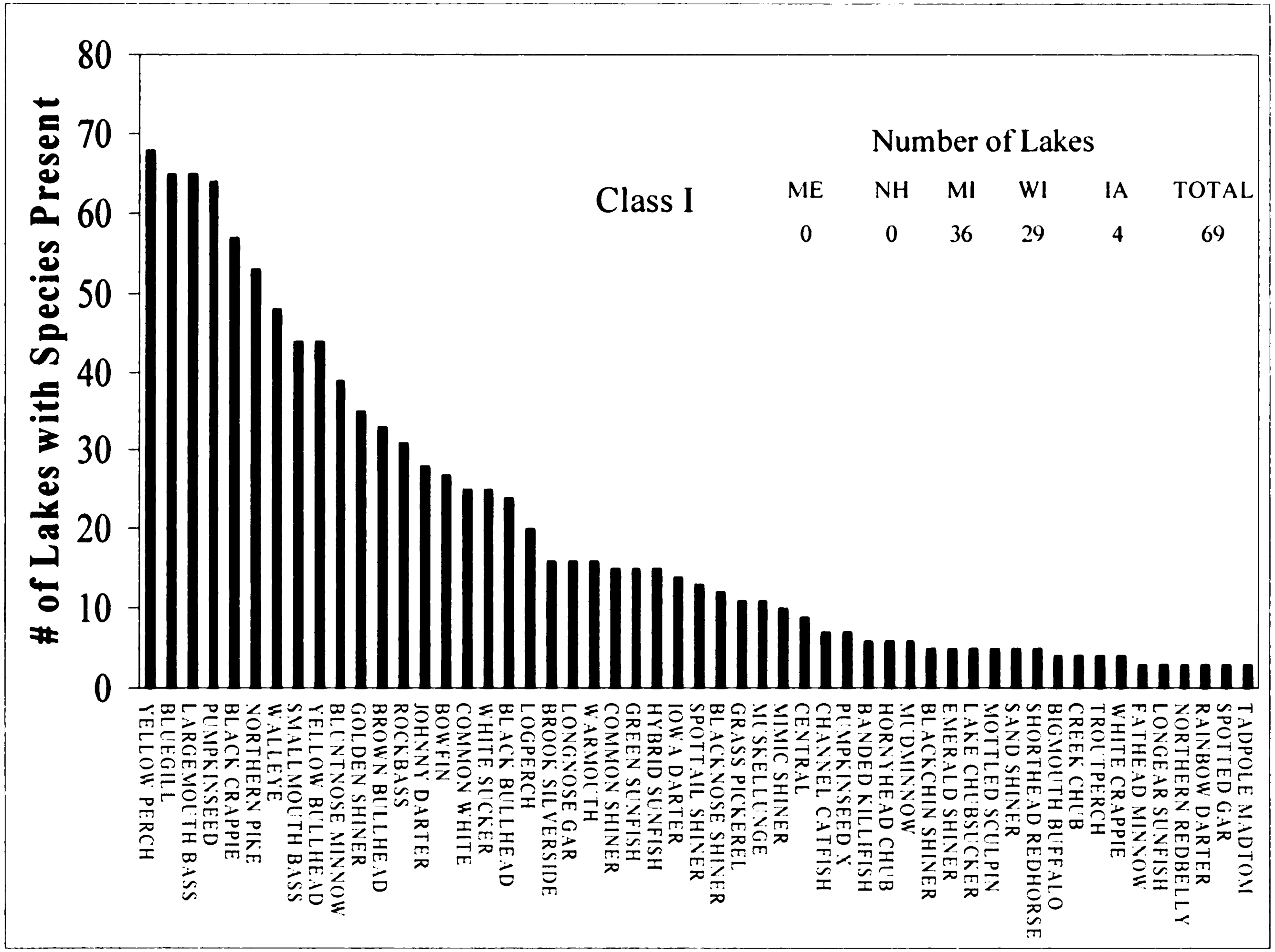

Figure 16. Number of species present in each class across all lakes and states displayed to show the predominate assemblages in a class. Five species were present in two lakes and 22 species were present in one lake.

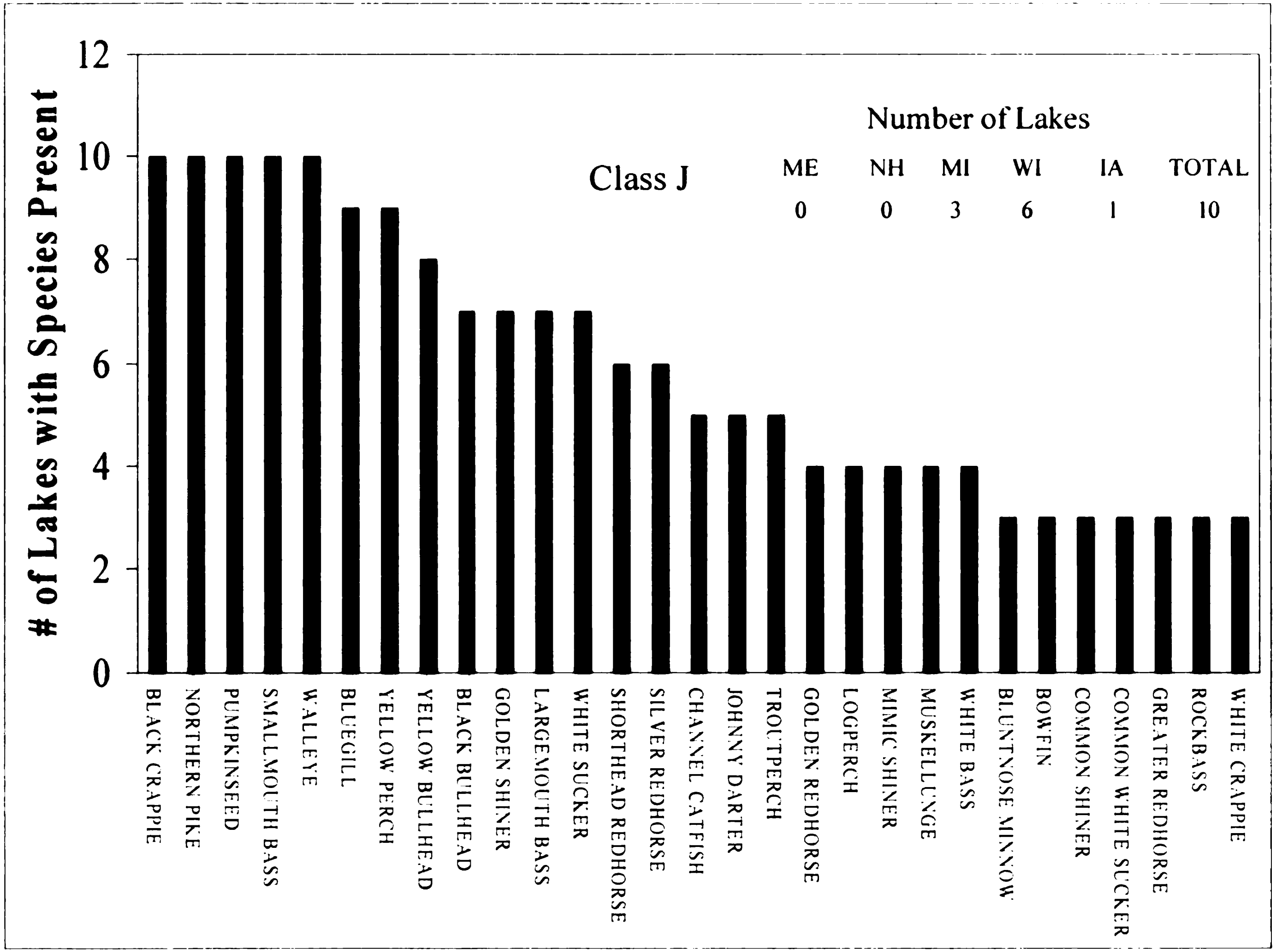

Figure 17. Number of species present in each class across all lakes and states displayed to show the predominate assemblages in a class. Seven species were present in two lakes and 17 species were present in one lake.

Bibliography

Abell, R.A., D.M. Olson, E. Dinerstein, P.T. Hurley, J.T. Diggs, W. Eichbaum, S. Walters, W. Wettengel, T. Allnutt, C.J. Loucks, and P. Hedao. Freshwater Ecoregions of North America. Island Press. Washington, D.C. 2000.

Allan, J.D. and L.B. Johnson. 1997. Catchment-scale analysis of aquatic ecosystems. *Freshwater Biology*. 37: 107-111.

Anderson, D.W. and D.H Schupp. 1986. Fish community responses to northern pike stocking in Horseshoe Lake, Minnesota. *Minnesota Department of Natural Resources Report.* No. 387

Bachmann, R.W., B.L. Jones, D.D. Fox, M. Hoyer, L.A. Bull, and D.E. Canfield Jr. 1996. Relations between trophic state indicators and fish in Florida (U.S.A.) lakes. *Canadian Journal of Fisheries and Aquatic Sciences.* 53: 842-855.

Becker, G.C. Fishes of Wisconsin. University of Wisconsin Press. Madison, Wisconsin. 1983.

Breiman, L., J.H. Friedman, R. Olshen, and C. Stone. Classification and Regression Trees. Wadsworth, Belmont. 1984.

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.

Brunberg, A.K. and P. Blomqvist. 2001. Quantification of anthropogenic threats to lakes in a lowland county of Central Sweden. *Ambio*. 30(3): 127-134.

Buck, O., D.K. Niyogi, and C.R. Townsend. 2003. Scale-dependence of land use effects on water quality of streams in agricultural catchments. *Environmental Pollution*. 130: 287- 299.

Carpenter, S.R., B.J. Benson, R. Biggs, J.W. Chipman, J.A. Foley, S.A. Golding, R.B. Hammer, P.C. Hanson, P.T.J. Johnson, A.M. Kamarainen, T.K. Kratz, R.C. Lathrop, K.D. McMahon, B. Provencher, J.A. Rusak, C.T. Solomon, E.H. Stanley, M.G. Turner, M.J. Vander Zanden, C. Wu, and H. Yuan. 2007. Understanding regional change: A comparison of two lake districts. *BioScience.* 57(4): 323-335.

Cheruvelil, K.S., P.A. Soranno, M.T. Bremigan, T. Wagner, and S.L. Martin. 2008. Grouping lakes for water quality assessment and monitoring: the roles of regionalization and spatial scale. *Environmental Management.* 41: 425-440.

Dolman, W.B. 1990. Classification of Texas reservoirs in relation to limnology and fish community associations. *Transactions of the American Fisheries Society*. 19: 511-520.

Drake, M.T. and D.L. Pereira. 2002. Development of a fish-based index of biotic integrity for small inland lakes in Central Minnesota. *North American Journal of Fisheries Management.* 22: 1105-1123.

Eadie, J. McA., T.A. Hurley, R.D. Montgomerie and K.L. Teather. 1986. Lakes and rivers as islands: species-area relationships in the fish faunas of Ontario. *Environmental Biology of Fishes.* 15(2): 81-89.

Gergel, S.E., M.G. Turner, J.R. Miller, J.M. Melack and E.H. Stanley. 2002. Landscape indicators of human impacts to riverine systems. *Aquatic Sciences.* 64: 118-128.

Hair J.F. Jr., R.E. Anderson, R.L. Tatham, and W.C. Black. Multivariate Data Analysis, 3rd edn. New York, Macmillan. 1995.

Hansen, G.A. and M.L. Jones. 2008. The value of information in fishery management. *Fisheries* 33: 340-348.

Harlan, J.R. and E.B. Speaker. Iowa Fish and Fishing. State of Iowa. 1956.

Hershey, A.E., S. Beaty, K. Fortino, M. Keyse, P.P. Mou, W.J. O'Brien, A.J. Ulseth, G.A. Gettel, P.W. Lienesch, C. Luecke, M.E. McDonald, C.H. Mayer, M.C. Miller, C. Richards, J.A. Schuldt and S.C. Whalen. 2005. Effect of landscape factors on fish distribution in arctic Alaskan lakes. *Freshwater Biology.* 51: 39–55.

Hill, N.M., P.A. Keddy, and I.C. Wisheu. 1998. A hydrological model for predicting the effects of dams on the shoreline vegetation of lakes and reservoirs. *Environmental Management.* 22(5): 723–736.

Higgins, J.V., M.T. Bryer, M.L. Khoury, and T.W. Fitzhugh. 2005. A freshwater classification approach for biodiversity conservation planning. *Conservation Biology.* 19(2): 432-445.

Holmgren, K. and M. Appelberg. 2000. Size structure of benthic freshwater fish communities in relation to environmental gradients. *Journal of Fish Biology.* 57: 1312-1330.

Holmlund, C.M. and M. Hammer. 2004. Effects of fish stocking on ecosystem services: An overview and case study using the Stockholm Archipelago. *Environmental Management*. 33(6): 799–820.

Hubbs, C.L. and K.F. Lagler. Fishes of the Great Lakes Region. University of Michigan Press. Ann Arbor, MI. 2007

Iowa Department of Natural Resources. 2009. Iowa Fishing Guide.

Irz, P., C. Argillier and J.P. Proteau. 2004. Contribution of native and non-native species to fish communities in French reservoirs. *Fisheries Management and Ecology.* 11: 165–172.

Jackson, D.A. and H.H. Harvey. 1989. Biogeographic associations in fish assemblages: local vs. regional processes. *Ecology.* 70(5): 1472-1484.

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

Jeppesen, E.J., J.P. Jensen, M. Sondergaard, T. Lauridsen, and F. Landkildehus. 2000. *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 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.

Johnson, P.TJ., J.D. Olden, M.J. Vander Zanden. 2008. Dam invaders: impoundments facilitate biological invasions into freshwaters. *Frontiers in Ecology and the Environment.* 6(7): 357-363.

Jones, J.R., M.F. Knowlton, D.V. Obrecht, and E.A. Cook. 2004. Importance of landscape variables and morphology on nutrients in Missouri reservoirs. *Canadian Journal of Fisheries and Aquatic Sciences.* 61: 1503-1512.

Kratz, T.K., K.E. Webster, C.J. Bowser, J.J. Magnuson, and B.J. Benson. 1997. The influence of landscape position on lakes in northern Wisconsin. *Freshwater Biology.* 37: 209-217.

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(8): 2941-2956.

Maine Department of Inland Fisheries and Wildlife. 2008. Open Water Fishing Regulations.

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 Science*. 44(2): 198-215.

McMahon G., S.M. Gregonis, S.W. Waltman, J.M. Omernik, T.D. Thorson, J.A. Freeouf, A.H. Rorick, and J.E. Keys. 2001. Developing a spatial framework of common ecological regions for the conterminous United States. *Environmental Management.* 28(3): 293–316.

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

Michigan Department of Natural Resources. 2009. Michigan Fishing Guide.

New Hampshire Fish and Game Department. 2009. Freshwater Fishing Digest.

Olden, J.D. and D.A. Jackson. 2002. A comparison of statistical approaches for modeling fish species distributions. *Freshwater Biology*. 47: 1976–1995.

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., O.P. Jensen and M.J. Vander Zanden. 2006. Implications of long-term dynamics of fish and zooplankton communities for among-lake comparisons. *Canadian Journal of Fisheries and Aquatic Sciences.* 63: 1812-1821.

Omernik, J. M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers.* 77(1): 188-125.

Pierce, R.B. and C.M. Tomcko. 2003. Effects of walleye stocking on the fish community of Lake Thirteen. *Minnesota Department of Natural Resources Investigational Report.* No. 505.

Poff, N.L. 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. *Journal of the North American Benthological Society.* 16(2): 391-409.

Prest, V.K. 1969. Retreat of Wisconsin and Recent ice in North America. Geological Survey of Canada Map, 1257A, scale 1:5,000,000.

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3900051-07-0, URL http://www.R-project.org.

Radomski, P.J. and T.J. Goeman. 1995. The homogenizing of Minnesota lake fish assemblages. *Fisheries*. 20(7): 20-23.

Rahel, F.J. 1984. Factors structuring fish assemblages along a bog lake successional gradient. *Ecology*. 65(4): 1276-1289.

Rempel, L.L. and D.G. Smith. 1998. Postglacial fish dispersal from the Mississippi refuge to the Mackenzie River basin. *Canadian Journal of Fisheries and Aquatic Sciences*. 55: 893-898.

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

Roth, B.M., I.C. Kaplan, G.G. Sass, P.T. Johnson, A.E. Marburg, A.C. Yannarell, T.D. Havlicek, T.V. Willis, M.G. Turner, and S.R. Carpenter. 2007. Linking terrestrial and aquatic ecosystems: The role of woody habitat in lake food webs. *Ecological Modeling*. 203: 439–452.

The data analysis for this paper was generated using *SAS software*, Version 9.1 of the *SAS* System for Windows. Copyright © 2004 *SAS* Institute Inc. *SAS* and all other *SAS* Institute Inc. product or service names are registered trademarks or trademarks of *SAS* Institute Inc., Cary, NC, USA.

Scarola, J.F. Freshwater Fishes of New Hampshire. New Hampshire Fish and Game Department. Division of Inland and Marine Fisheries. 1973.

Schupp, D.H. 1992. An ecological classification of Minnesota lakes with associated fish communities. *Minnesota Department of Natural Resources Report*. No. 417.

Seaber, P.R., F.P Kapinos, and G.L. Knapp, 1987. Hydrologic Unit Maps. U.S. Geological Survey Water-Supply Paper 2294, United States Department of the Interior, Geological Survey, Denver, Colorado.

Seelbach, P.W., M.J. Wiley, P.A. Soranno, and M.T. Bremigan. 2002. Aquatic conservation planning: Using landscape to maps to predict ecological reference conditions for specific waters. Pages 454-478 *in* K. Gutzwiller, editor. Concepts and Applications of Landscape Ecology in Biological Conservation. Springer-Verlag.

Soranno, P.A., K.E. Webster, K.S. Cheruvelil and M.T. Bremigan. 2009. The lake landscape-context framework: linking aquatic connections, terrestrial features and human effects at multiple spatial scales. *Verh. Internat. Verein. Limnol.* 30(5): 695-700.

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(4): 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.

Tonn, W.M., C.A. Paszkowski, G.J. Scrimgeour, P.K.M. Aku, M. Lange, E.E. Prepas and K. Westcott. 2003. Effects of forest harvesting and fire on fish assemblages in boreal plains lakes: A reference condition approach. *Transactions of the American Fisheries Society.* 132: 514-523.

Wisconsin Department of Natural Resources. Bureau of Fisheries Management. 2009. Guide to Wisconsin Hook and Line Fishing Regulations.

Webster, K.E., P.A. Soranno, K.S. Cheruvelil, M.T. Bremigan, J.A. Downing, P.D. Vaux, T.R. Asplund, L.C. Bacon, and J. Connor. 2008. An empirical evaluation of the nutrient-color paradigm for lakes. *Limnology and Oceanography.* 53(3): 1137-1148.