Fish Indicators of Anthropogenic Stress at Great Lakes Coastal Margins: Multimetric and Multivariate Approaches

By

Yakuta Bhagat

A Thesis
Submitted to the Faculty of Graduate Studies and Research
through the Department of Biological Sciences
in Partial Fulfillment of the Requirements for
the Degree of Master of Science
at the University of Windsor

Windsor, Ontario, Canada

2005

© 2005 Yakuta Bhagat

Fish Indicators of Anthropogenic Stress at Great Lakes Coastal Margins: Multimetric and Multivariate Approaches

by

Yakuta Bhagat

APPROVED BY:

V. C. Lakhan
Department of Earth Sciences

L.B. Johnson
NRRI, University of Minnesota, Duluth

L. D. Corkum
Department of Biological Sciences

J.J.H. Ciborowski, Advisor
Department of Biological Sciences

A. Ali, Chair of Defense
Department of Biological Sciences

24 June 2005

ABSTRACT

Two commonly applied approaches to assessing the health of an aquatic habitat are the multimetric index of biotic integrity (IBI) and the multivariate approach. Fish community composition and local environmental conditions were measured at 143 sites across the entire U.S. coastline between 2002 and 2003, using overnight-set fyke nets (a large collaborative study to develop environmental indicators at Great Lakes coastal margins (GLEI)). Of these sites, data from 55 wetlands were used in an attempt to develop IBIs for the Northern Great Lakes (NGL) ecoregion and the Erie and Ontario (EOL) ecoregion. Candidate metrics were evaluated with respect to agricultural stress measured at drainage basin and immediate watershed topographical scales. An IBI could be developed only for the NGL ecoregion, based on a criterion requiring inclusion of at least 4 metrics representing attributes of a natural habitat. Fish community metrics in the NGL ecoregion varied more strongly with stress at the larger segment-shed scale.

A further test of the IBI approach was done using independently developed IBIs for *Typha* and *Scirpus* aquatic plant zones of Great Lakes coastal wetlands (Uzarski et al. 2005). IBI scores were calculated for 32 wetlands with dominant *Typha* and *Scirpus* vegetation using data collected through the GLEI project. The results indicated that both the *Typha* and *Scirpus* IBI were selectively responsive to anthropogenic stressors. The *Typha* IBI varied most significantly as a function of population density-related stress and the *Scirpus* IBI most strongly reflected agriculture-related stress.

To test the multivariate approach of fish community assessment, cluster analysis was used to distinguish unique groupings of reference sites based on relative abundances of fish species. A discriminant function analysis model distinguished groupings on the basis of ecoregion and 7 other environmental variables. Bray-Curtis ordination was used to assess changes in fish community in all 143 sites sampled with respect to 2 classes of human activity: agriculture and population density. Stronger effects of population density-related stress than of agriculture-related stress were observed. Overall, multivariate derivations were more effective in delineating fish indicators of anthropogenic stress at Great Lakes coastal margins.

To Mom and Dad
for their love and support
and to my brother, Yusuf,
who continually inspires me
to challenge myself

ACKNOWLEDGEMENTS

I would like to acknowledge my supervisor, Dr. Jan J.H. Ciborowski for his undying support, both academic and moral. His love and devotion to research is truly inspiring and I will always be grateful to him for believing in my potential and granting me the opportunity to work and gain experience in his lab.

Drs. Lucinda B. Johnson and Valerie J. Brady provided valuable insights and critiques that helped enhance the project and Dr. Brady also provided assistance with data compilation for many of the statistical analyses. Dr. Thomas P. Hollenhorst provided much needed assistance with GIS software and maps. My committee members, Drs. Lynda D. Corkum and V. Chris Lakhan provided valuable feedback for several aspects of the research.

This work would not be possible without the tireless efforts of the following people in the field: Katherine Andriash, Jesse Baillargeon, Lucas Byrk, Rebecca Danard, Justin Duncan, Carolyn Foley, Joe Gathman, Jeff Holland, MiSun Kang, and Paige Short. I would also like to acknowledge my family, friends and lab colleagues for their immense support and helpful suggestions on several presentations and paper reviews.

This research was supported by grants from the U.S. Environmental Protection Agency's Science to Achieve Results (STAR) Estuarine and Great Lakes (EaGLe) Coastal Initiative through funding to the Great Lakes Environmental Indicators (GLEI) project, US EPA/R-8286750 and the Reference Condition Agreement US EPA/R-82877701-0. This document has not been subjected to the Agency's required peer and policy review and therefore does not necessarily reflect the views of the Agency, and no official endorsement is inferred. Other funding was provided by a grant from the Natural Sciences and Engineering Research Council of Canada to Dr. J.J.H. Ciborowski.

STATEMENT OF ORIGINALITY

I certify that this thesis, and the research to which it refers, are the product of my own work and any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged according to the standard referencing practices of the discipline. I acknowledge the helpful guidance and support of my supervisor, Dr. Jan J.H. Ciborowski. This work was not previously submitted for a degree or diploma at any university.

TABLE OF CONTENTS

ABSTRACT…….iii

DEDICATION…….iv

ACKNOWLEDGEMENTS……..v

STATEMENT OF ORIGINALITY…….vi

LIST OF TABLES……...ix

LIST OF FIGURES……..xi

CHAPTER 1 – GENERAL INTRODUCTION……1
GENERAL METHODS……..5

CHAPTER 2 – DEVELOPING A FISH INDEX OF BIOTIC INTEGRITY (IBI) FOR GREAT LAKES COASTAL WETLANDS: STRATIFIYING BY ECOREGION……9

ABSTRACT……10

INTRODUCTION…….....11

METHODS……..12

RESULTS……19

DISCUSSION……..27

SUMMARY AND CONCLUSIONS……..29

CHAPTER 3 – TESTING A FISH INDEX OF BIOTIC INTEGRITY FOR GREAT LAKES WETLANDS STRATIFIED BY PLANT ZONES……31

ABSTRACT……...32

INTRODUCTION……..33

METHODS…….34

RESULTS……...36

DISCUSSION…….48

SUMMARY AND CONCLUSIONS…… 51

CHAPTER 4 – USING MULTIVARIATE TECHNIQUES TO DEVELOP FISH INDICATORS OF STRESS FOR GREAT LAKES COASTAL MARGINS…….53

ABSTRACT……...54

INTRODUCTION…….55

METHODS……57

RESULTS……..65

DISCUSSION……86

SUMMARY AND CONCLUSIONS……91

CHAPTER 5 – GENERAL DISCUSSION……...92

GENERAL CONCLUSIONS……...95

FUTURE RESEARCH……96

BIBLIOGRAPHY……..97

APPENDIX I……104

APPENDIX II……...107

APPENDIX III……..111

APPENDIX IV……..112

APPENDIX V……...113

APPENDIX VI……..114

APPENDIX VII…….116

VITA AUCTORIS…….120

LIST OF TABLES

Table 2.1. List of candidate metrics and the expected responses for each metric in EOL and NGL ecoregions……14

Table 2.2. Spearman r values between EOL metrics and agricultural stress measured at segment-shed scale and watershed scale. Metrics significantly correlated to agricultural stress ($p < 0.05$) are indicated with an asterisk (*)……...21

Table 2.3. Spearman r values between NGL metrics and agricultural stress measured at segment-shed scale and watershed scale. Metrics were significantly correlated to agricultural stress if $p < 0.05$. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p< 0.001$……...……...24

Table 2.4. Scoring criteria for preliminary fish IBI for NGL wetlands……...….25

Table 3.1. Trophic guild (Tr_guild) and Catch per unit effort (CPUE of all fish species collected in fyke nets between 2001 and 2003 in *Typha* and *Scirpus* dominant wetlands of all 5 Great Lakes (* and ** denotes species unique to *Typha* and *Scirpus* sites respectively. Piscivores (PISC), Omnivores (OMN), Insectivores (INS), Herbivores (HERB), Zoobenthivores (ZOOBEN) and Molluscivores (MOLL)……...38

Table 3.2. Site locations and IBI scores for *Typha*- and *Scirpus*-dominant sites sampled through the GLEI project…….41

Table 3.3. Spearman correlations between *Typha* (n = 19) and *Scirpus* (n = 13) IBI scores based on GLEI fish catches and values of 6 anthropogenic stressors measured at segment shed levels. Significant correlations ($p < 0.05$) are denoted by an asterisk (*)…….……..42

Table 4.1. Correlation (factor loading) between values of 48 environmental variables measured at 133 Great Lakes coastal margin sites and 14 principal components. Variables subsequently used for discriminant function analysis of clusters of sites associated by fish species relative abundance are shown in bold face……..60

Table 4.2. Correlation (factor loading) between values of 6 anthropogenic stressor variables and 2 principal component factors. Factor 1 represents population density related pressure and factor 2 represents agriculture related pressure. Variables combined in factors 1 and 2 are shown in bold face……66

Table 4.3. Summary of observed number of sites in each cluster (columns) identified by fish relative abundances and membership predicted (rows) by discriminant function classification (Appendix VII) on the basis of environmental variables measured at those sites……69

Table 4.4. Variables determined by DFA model as significant in classifying REF site cluster membership……70

Table 4.5. Reference and non-reference indicator species assemblages for each cluster of sites based on Spearman rank correlation coefficient of species relative abundance (octave) with population pressure (Pop.) and agricultural pressure (Ag.)……74

Table 4.6. Classification of GLEI sites whose fish species relative abundances identify them as being equivalent to reference sites or not, based on threshold criteria and ordination distances from reference endpoint……..…77

Table 4.7. Classification of reserve sites as reference or non-reference based on extrapolated ordination scores…….85

Table 4.8. Relative abundances of fish species at reserve sites. Species with 0 abundances have been omitted from the table…….87

LIST OF FIGURES

Figure 1.1. Illustration of a hypothetical segment-shed as defined under the GLEI project and a hypothetical watershed as defined under the Reference Area project……...8

Figure 2.1. Relationship between total number of species and GLEI agricultural stress score at the segment shed level for NGL ($r = 0.38$, $p < 0.05$, n=28). The metric was calibrated so that sites with high number of species received a score of 1, sites with low number of species received a score of 5 and sites in between received a score of 3…….22

Figure 2.2. Relationship between percent non-indigenous species and GLEI agricultural stress score at the segment shed level for NGL ($r = 0.62$, $p < 0.001$, n=28). The metric was calibrated so that sites with high percent non-indigenous species received a score of 1, sites with low percent non-indigenous species received a score of 5 and sites in between, received a score of 3……..23

Figure 2.3. Relationship between overall IBI scores and GLEI agricultural stress at segment shed level for NGL sites ($r = -0.33$, $p > 0.05$, n=28)……...26

Figure 3.1. Map of the Great Lakes showing GLEI *Scirpus* and GLEI *Typha* dominant sites…….37

Figure 3.2. Relationship between *Typha* IBI and population density stress for GLEI ($r = -0.70$, $p < 0.05$, n = 19)……43

Figure 3.3. Relationship between *Typha* IBI and atmospheric deposition for GLEI sites ($r = 0.60$, $p < 0.05$, n = 19)……..44

Figure 3.4. Relationship between *Scirpus* IBI and agricultural stress calculated for GLEI sites ($r = -0.64$, $p < 0.05$, n = 13)…….45

Figure 3.5. Relationship between *Scirpus* IBI and point source discharge for GLEI sites ($r = -0.57$, $p < 0.05$, n = 13)…….46

Figure 3.6. Map of the Great Lakes segment sheds showing Uzarski et al. (2005) *Scirpus* and *Typha* dominant sites……47

Figure 3.7. Relationship between GLEI and Uzarski *Scirpus* dominant sites to agricultural stress ($r = -0.66$, $p < 0.001$, n = 30)…….....49

Figure 4.1. Dendogram of REF sites (n = 42) grouped according to similar species composition (Ward's method clustering city-block distances of octave-transformed relative abundances of fish species). Latitude, longitude and place names of sites corresponding to site labels are summarized in Appendix I)……..67

Figure 4.2. Relationship between Bray-Curtis fish relative abundance ordination scores and population pressure for sites in cluster 1. Low fish ordination scores are most similar to the reference score of zero. High fish ordination scores are most similar to the score most closely associated with large values of population pressure score……..71

Figure 4.3. Relationship between Bray-Curtis fish relative abundance ordination scores and agricultural pressure for sites in cluster 1. Low relative abundance scores are most similar to the reference condition centroid. High scores are most similar to the suite of fish relative abundances associated with the site having the highest agricultural pressure score……72

Figure 4.4. Relationship between BC axis 1 and BC axis 2 ordination scores for cluster 1 sites. Threshold lines indicate boundaries of reference sites along both axes……76

Figure 4.5. Relationship between BC axis 1 and BC axis 2 ordination scores for cluster 2 sites. Threshold lines indicate boundaries of reference sites along both axes……..80

Figure 4.6. Relationship between BC axis 1 and BC axis 2 ordination scores for cluster 3 sites. Threshold lines indicate boundaries of reference sites along both axes……81

Figure 4.7. Relationship between BC axis 1 and BC axis 2 ordination scores for cluster 4 sites. Threshold lines indicate boundaries of reference sites along both axes……83

Figure 4.8. Relationship between BC axis 1 and BC axis 2 ordination scores for cluster 5 sites. Threshold lines indicate boundaries of reference sites along both axes……84

GENERAL INTRODUCTION

The effects of human-induced disturbances affecting the Great Lakes basin have been of major concern to managers and researchers alike. Since the passing of the U.S. Clean Water Act of 1972 (PL 92-500), there has been much focus on developing biological indicators of anthropogenic stress to measure 'biotic integrity' of a habitat. Biotic integrity is one of the primary objectives set forth by the Clean Water Act and is defined as "the ability of a habitat to support and maintain a balanced, integrated, adaptive community of organisms having a composition, diversity and functional organization comparable to that of a natural habitat" (Frey 1977). A natural habitat as such, refers to an area with minimal levels of anthropogenic stress, also referred to as a reference condition area (Host et al. in press).

Many groups of taxa have been used to develop indicators of stress, ranging from diatoms (Ferguson et al. 2003) and benthic invertebrates (Burton et al. 1999) to aquatic plants (Galatowitsch et al. 1999), amphibians (Grabas et al. 2004) and more commonly, fish (e.g., Simon 1991). Fishes serve as especially good indicators of stress because they are often philopatric, represent a broad spectrum of community tolerances to manifestations of anthropogenic disturbance from very sensitive to highly tolerant, and respond to physical, chemical and biological degradation (Plafkin et al. 1989). Fish communities include species representing various trophic levels (piscivores, omnivores, insectivores, herbivores), and their position in relation to diatoms and benthic invertebrates helps to provide an integrative approach to habitat assessment (Karr 1981).

Two approaches are widely used in developing biological indicators of stress – multimetric and multivariate, and both have advantages and disadvantages. The index of biotic integrity (IBI) is the most commonly used multimetric approach, particularly in assessing streams and rivers in the U.S. (Karr 1981, Lyons and Wang 1996, Mundahl and Simon 1999, Emery et al. 2003). Originally formulated specifically for the evaluation of fish communities, the IBI takes into account a variety of measures or attributes in connection with the region under investigation. Each metric, or measure of an attribute of the region, is equally weighted and contributes to an overall IBI score, which signifies the "integrity" of a fish community at a site. In theory, the IBI reflects the degree to which the local environment influences the fish community. Development of an IBI depends on

the *a priori* identification of reference sites (locations that are subject to minimal levels of anthropogenic stress) typically delineated by watershed boundaries (Karr 1981, Fausch et al. 1984) or ecoregions (Hughes and Larson 1988, Omernik 1995). Thus, each ecological zone could theoretically require its own IBI consisting of metrics that represent measures of a healthy community, specific to that ecoregion or watershed. One of the main advantages of using the IBI approach is that it takes into account a variety of attributes that represent the fish community in a site (Simon 1991). The IBI produces a single score that can be compared to a sampling distribution of expected scores based on scores of minimally disturbed sites in the region. Thus, the IBI reflects the fish community response to relative degree of disturbance at a particular site (Karr 1981).

However, there are some disadvantages with using the IBI approach. Not all information that is collected is used in developing an IBI, so metrics that may seem meaningful to include, may ultimately be rejected because of their lack of significant correlation with a particular stressor of choice (Reynoldson et al. 1997). Metrics in an IBI can often be redundant; in which case, the researcher overweighs the effect of a single measurable attribute. Because IBIs are frequently used in stream assessment (Fausch et al. 1984) and more recently, in wetland assessment (Wilcox et al. 2002, Uzarski et al. 2005), there is no general consensus on the regional framework used in developing and applying an IBI. For instance, some researchers contend that hydrogeomorphic type is an important classification feature across which fish communities are segregated (Keough et al. 1999, Smith et al. 1995), yet others have shown that plant zonation is a better predictor of metric expectations (Burton et al. 1999, Uzarski et al. 2005). Such incongruent ideas are implicit in the development of an IBI and have often resulted in multiple IBIs developed for similar regions based on differences in a priori assumptions.

In contrast to the multimetric approach, the multivariate approach groups reference sites based on similarity of species composition (Reynoldson et al. 1995, Norris 1995, Fore et al. 1996, Reynoldson et al. 1997). With the use of a discriminant function analysis and a set of pre-determined environmental variables, clusters of compositionally similar reference sites are grouped, with which one can compare test sites for assessment (Reynoldson et al. 1997). One of the advantages of using multivariate approaches is that no *a priori* assumptions need be made about the similarity of fish communities at

different sites based on either physical or chemical descriptions (Reynoldson et al. 1997). The multivariate approach thus provides an objective way of grouping reference sites that possess similar fish assemblages. The main drawback to this approach is that the initial model construction requires statistical expertise. The results can also be difficult to convey to the general public, which has often deterred managers from using or applying multivariate techniques to habitat quality assessments (Yoder 1995).

Few studies to date have compared the results of multimetric and multivariate approaches to assessing aquatic habitats using fish communities. Reynoldson et al. (1997) conducted a comparison of the two approaches for river assessment using a benthic macroinvertebrate data set from the Fraser River, British Columbia. Although their results showed that higher accuracy and precision was obtained using multivariate methods compared to multimetric, Reynoldson et al., (1997) recommend using the two approaches in conjunction when assessing habitat quality. Milner and Oswood (2000) conducted a similar study for benthic invertebrates in streams in Anchorage, Alaska and found that multivariate classification of site quality had the highest overall strength compared to classifications of sites using multimetric methods.

Although fish communities are often studied in stream (Leonard and Orth 1986, Simon 1991), river (Emery et al. 2003), and wetland assessments (Wilcox et al. 2002, Uzarski et al. 2005), no studies to date have compared the multimetric and multivariate approaches using fish communities, for Great Lakes coastal margins.

My objectives in this thesis were as follows – 1) to use the multimetric (IBI) approach for assessing wetland quality in a northern and southern ecoregion of the Great Lakes, 2) to calibrate an IBI stratified by plant zones (Uzarski et al. 2005) against known stressor gradients and 3) to use the multivariate approach to assess habitat quality for sites along the coastal margins of the Great Lakes. Each of the objectives listed above is addressed in a separate chapter. This project is part of a larger effort, the Great Lakes Environmental Indicators (GLEI) project, funded by the US EPA, whose overall objective was to develop indicators of anthropogenic stress for Great Lakes coastal margins using birds, amphibian, diatoms, fish, benthic invertebrates and plants in a landscape context.

In Chapter 2, data collected from fyke net sampling in 55 wetlands across a northern and southern ecoregion of the Great Lakes, were used in an attempt to develop

an index of biotic integrity (IBI) for these two ecoregions. Published literature and the fish assemblage at reference wetlands was used to compile a suite of candidate metrics (attributes of fish communities) evaluated (using Spearman rank correlations) with respect to a measure of agricultural activity derived from the GLEI project (see Danz et al. 2005 for methods). Based on previous studies using similar approaches to assessing streams (Fausch et al. 1984), rivers (Emery et al. 2003) and wetlands (Wilcox et al. 2002), my expectations were that a separate ecoregion IBI could be developed, consisting of metrics reflective of the unique species composition in that ecoregion.

In Chapter 3 fish sampling data and stressor data (derived from the GLEI project) were used to calibrate two independently developed IBIs for two major plant zones of Great Lakes coastal wetlands: *Typha* and *Scirpus* to known stressor gradients (from Danz et al. 2005). Uzarski et al. (2005) proposed that the IBIs they developed would serve as good indicators of overall habitat integrity as long as the sites sampled and tested were dominated by either *Typha* or *Scirpus*. Uzarski et al. (2005) tested their metrics against disturbance gradients that were identified by assessing land use/cover (developed land, agriculture, forest, and wetlands) and chemical/physical data (pH, turbidity, specific conductance etc.). I thus expected to find both the *Typha* and *Scirpus* IBIs to respond most strongly to the GLEI land cover and agriculture stressor gradients

In Chapter 4, fish assemblage data and a suite of environmental variables were used to assess the quality of 133 sites along the U.S Great Lakes coastline to develop fish indicators for reference and non-reference ends of a disturbance gradient using cluster analysis, discriminant function analysis and ordination analysis. My expectations from the cluster analysis were that clear separations would be obtained between groups of sites since fish species compositions (rather than a priori assumptions) are used in selecting groupings of sites. I expected the discriminant function analysis to show that temperature, latitude (Lougheed et al. 2001) and plant cover (Uzarski et al. 2005) were important variables in separating groups of sites and consequently, fish communities. The ordination analyses were expected to reveal distinct thresholds between reference and non-reference sites based on ordination distances in stressor space (Brazner and Beals 1997). Likewise, I expected to see unique assemblages of reference and non-reference indicator species for each of the unique clusters of sites.

General Methods

Data were collected at sites along Great Lakes coastal margins that were selected according to 2 distinct sampling designs – the US EPA funded the Great Lakes Environmental Indicators study (Niemi et al. 2001) and the Reference Area project (Schuldt et al. 2000). The goal of the GLEI project was to develop and test the efficacy of a suite of indicators of anthropogenic stress at Great Lakes coastal margins. The goal of the Reference Area project was to quantitatively identify reference sites within *a priori* based classification units of Great Lake coastal ecosystems. Identical field methodology was used to sample biota and environmental characteristics. Sites identified according to the GLEI protocol are called GLEI sites. Sites sampled to evaluate the reference area concepts are called Reference Area sites.

Site Selection

Great Lakes Environmental Indicators (GLEI) project

The objective for the site selection portion of this project was to identify and sample test sites (coastal ecosystems) that were distributed across the full gradient of anthropogenic stress. Due to the large size of the Great Lakes basin, portions of the coastline that contained the coastal ecosystems were delineated and environmental variables were computed for these coastal portions also referred to as 'segment-sheds'. The segment-sheds were delineated in a two-step process using GIS. The first step involved defining segments as lengths of shoreline beginning and ending half-way between each second order or higher stream using Reach File version 3.0, (RF3) (U.S. EPA 1994). The drainage area associated with each segment was delineated using the National Elevation Dataset (Gesch et al. 2002) (Figure 1.1). The result was 762 segment-sheds for the entire U.S. Great Lakes coastal margin.

Each of the 762 segment-sheds was inventoried for the presence of one or more of 5 hydrogeomorphic types – high energy shoreline, embayment, and 3 types of coastal wetlands: open-coast wetland (lacustrine), drowned-river mouth and flooded delta wetland (river-influenced) and protected wetland (as per Keough et al. 1999). Since high energy and embayment sites sometimes crossed segment shed boundaries, a test site was then identified as the portion of the hydrogeomorphic type within a segment-shed. To

ensure that the sites ultimately selected for field sampling spanned the full range of anthropogenic disturbance, Danz et al. (2005) assessed 207 geospatially referenced indicators of human activity. Danz et al. then used principal components analysis to group the variables into independent suites. They identified (PC-factor) one category of environmental variation not reflective of stress (i.e., soils) and 6 categories of human disturbance: agriculture-related (including agricultural chemicals), atmospheric deposition, land cover, human population density and development, point source pollution and shoreline modification, at each segment-shed. Segment-sheds were grouped together in 14 clusters using k-means cluster analysis of the 7 principal component factors. The individual segment-sheds were then evaluated, using digital topographical and aerial photographs from clusters, to find sites (within segment-sheds) that were most appropriate for sampling (i.e., accessible and containing the correct hydrogeomorphic type). A total of 97 sites along the entire U.S. coastline were identified for the purpose of sampling fish communities.

Reference Area Project

The underlying objective for selecting sites in this project was to use GIS and remotely sensed data to identify sites as most representative of reference condition. Reference condition is defined as the condition that exists in ecosystems that are least impacted by anthropogenic stressors (Host et al. in press). The 8,000 km stretch of the U.S. coastline was stratified by 6 Level III ecoregions *sensu* Omernik 1987. Coastline habitats were further stratified into five hydrogeomorphic types: high energy shoreline, embayments, lacustrine wetlands, river-influenced wetlands and protected wetlands (as per Keough et al. 1999). The degree of anthropogenic disturbance in contributing areas of these hydrogeomorphic types was assessed using a sliding window (spatially-defined moving sum or average) approach for 30m x 30m blocks (pixels identified by land satellite TEM) high energy shorelines and embayments and a watershed approach for wetland types. A watershed is defined as the topographical catchment or drainage basin of a stream also containing the delineated sampling area (Figure 1.1). The watershed surrounding each wetland within the 6 ecoregions was characterized according to 5 indicators of anthropogenic stress variables: agricultural land use, urban land use,

distance from a point source of pollution, road density and population density, measured using a geographic information systems (GIS) based analysis of land use and remote sensing. Values for each stressor variable at a site were then scaled to a proportion of the maximum observed value. Each site was assigned a 'Anthropogenic Pressure (AP)'score based on the maximum stressor value, (the highest value of the 5 stressor variables for that particular site) based on the assumption that the biological community is limited by the single greatest stressor. A wetland was classified as 'reference' if its score placed within the lowest 20th percentile of a frequency distribution of wetlands within a particular ecoregion.

Since high energy or embayment sites do not have drainage areas, a sliding window approach was used to assess stress at landscape scale resolution (30m x 30m pixels). The stress on each pixel was determined as the mean of each stressor value in a 1 km^2 area surrounding each pixel. Similarly to assessing reference wetlands, each pixel was ranked based on its AP score and the pixels that placed within the lowest 20th percentile became 'reference area' pixels. Adjacent 'reference' pixels were agglomerated into polygons and a final ranking of polygons $\geq$ 2 km^2 in length was used to identify reference high energy and embayment sites. A total of 46 reference sites were identified for sampling fish communities along the entire U.S. coastline.

In summary, Reference Area and GLEI sites differ in several respects. Reference Area sites are among the least affected by human disturbance within their ecoregion as determined at the local watershed scale. GLEI sites may be subject to any level of anthropogenic stress, ranging from the minimum to the maximum of any of 6 classes of stress found on the entire Great Lakes. Reference Area stress was assessed on the basis of 5 geospatially reference variables determined at the watershed scale. GLEI stress was scored in terms of 6 multivariate classes of human activity measured at the scale of segment-sheds. Because the position of each Reference Area sampling site could be related to a segment-shed, GLEI stress scores could be assigned to them. However, relatively undisturbed wetlands could occur in a highly urbanized or agricultural segment-shed, resulting in occasionally anomalously high stress scores for what is really a minimally disturbed location.

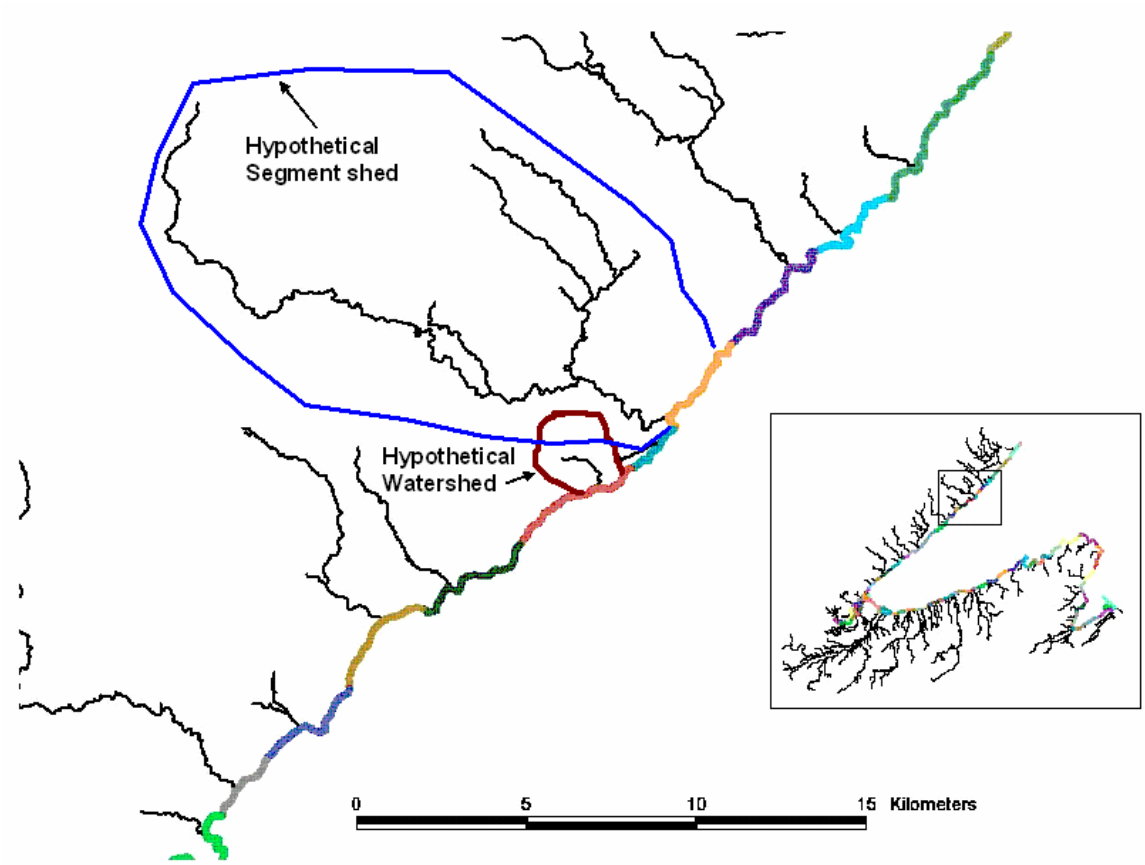

Figure 1.1. – Illustration of a hypothetical segment-shed as defined under the GLEI project and a hypothetical watershed as defined under the Reference Area project.

CHAPTER 2

Developing a fish index of biotic integrity (IBI) for Great Lakes wetlands: stratifying by ecoregions

Abstract

Simple measures of community composition such as species richness and abundance are seldom precise enough to reflect subtle differences of environmental condition. Fish communities are often used to characterize the 'health' of a water body, most commonly using multimetric analyses to generate Indices of Biotic Integrity (IBI). In 2002 and 2003 we sampled fish at over 80 wetlands along the U.S. Great Lakes coastline using overnight set fyke nets. Indices representing frequently used measures of biotic integrity were evaluated for each of 6 Great Lakes ecologically distinct regions. Metrics were evaluated across a range of environmental stress scores summarizing each of population density, land cover, atmospheric deposition, agricultural intensity, point source pollution, and shoreline modification. The scores were relativized, integrated values obtained for second order drainage basins (segment shed scale). Metrics were also evaluated against stress scores based on the wetlands sampled and their adjacent land use (watershed scale). In the Erie Ontario Lake Plain ecoregion the values of only 2 of 13 candidate metrics varied significantly across a gradient of agricultural disturbance at the segment shed scale. For the Northern Great Lakes ecoregion however, five of 16 candidate metrics varied significantly, which met the minimal criteria for IBI development. The results of the stressor scale comparison showed that fish communities respond more strongly at segment shed scale in the NGL ecoregion than the watershed scale.

Introduction

The Great Lakes represent one fifth of the world's freshwater and support the world's largest freshwater fisheries (US EPA and Government of Canada, 1995). Industrialization, urbanization and agricultural landuse subject the Great Lakes basin to varying levels of environmental stress. Some traditional approaches to determining biological responses to the level of environmental stress have included species richness measures or abundance measures of some key species of interest such as salmonids (Sonstegard and Leatherland 1984) or walleye (Schneider and Leach 1977). However, species richness measures are thought to be too simplistic and not sensitive enough to permit one to detect subtle differences in environmental condition (Karr et al. 1999).

A more integrative approach that has been widely used in the U.S. is the Index of Biotic Integrity (IBI) approach (Karr 1981, Lyons et al. 1996, Mundahl and Simon 1999, Emery et al. 2003). The IBI, a multimetric approach, takes into account a variety of metrics or measures of attributes of the fish community in connection with the region under investigation. A metric is a calculated term or enumeration representing some aspect of biological assemblage structure, function, or other measurable characteristic that changes with increasing human disturbance, in a predictable manner (Fausch et al. 1984). In an IBI, each metric is equally weighted and contributes to an overall IBI score, which signifies the "integrity" of a fish community at a site. In theory, the IBI reflects the degree to which environment influences the fish community. Because the expected fish community composition is assumed to be adapted to the 'natural' conditions of the local habitat, development of an IBI depends on the a priori ecological classification of reference sites (locations that have minimal levels of anthropogenic stress), typically by watershed boundaries (Karr 1981, Fausch et al. 1984) or ecoregions (Hughes and Larson 1988, Omernik 1995).

The first IBI, developed for use in small warmwater streams, consisted of 12 metrics representing species richness and composition, trophic composition and abundance and condition (Karr 1981). Since then, several researchers have developed IBIs for streams and rivers in the United States (Simon and Lyons 1995, Hughes and Oberdorff 1999). However, few have attempted to develop an IBI for the Great Lakes littoral zones and wetlands (Minns et al. 1994, Wilcox et al. 2002, Uzarski et al. 2005).

The primary objective of this paper is to determine whether the IBI approach can be used to develop a set of indicators or metrics for assessing the condition of Great Lakes coastal margins. Published literature (Minns et al., 1994, Wilcox et al. 2002, Karr 1981, Karr 1999) as well as fish assemblage information at reference sites was used to compile a list of candidate metrics to test in a northern and a southern ecoregion of the Great Lakes. The second objective in this paper was to compare the scale at which the values of metrics correlate most strongly to degree of anthropogenic disturbance: segment-shed level against watershed level (see Chapter 1 for definitions). Holland et al. (in preparation) showed that most fish species in Great Lakes wetlands respond to landuse at scales larger than watersheds (i.e., 30-50 km surrounding a wetland), while few species respond to landuse at local scales (0.5-5 km^2). We thus expected to find most metrics to be most highly correlated with stress measured at the segment-shed scale rather than the more localized watershed scale.

Methods

Site Selection and Classification

A total of 46 reference sites, sampled through the Reference Area project (see Chapter 1) were identified as least affected by anthropogenic disturbance using a geospatially referenced analysis of landuse and remote sensing that measured 5 watershed scale based stressor variables: agricultural land use, urban land use, distance from a point source of pollution, road density and population density (see Chapter 1). A second set of 97 sites, selected based on a design developed by the GLEI project was sampled in concert with reference sites using identical field protocols (see Chapter 1). The sites identified the GLEI project spanned the entire gradient of 6 multivariate categories of human disturbance and one category of environmental variation (soils). The 6 categories of human disturbance were related to agriculture, population density, land cover, atmospheric deposition, point source discharge and shoreline modification (see Chapter 1).

For the purpose of this analysis, only data from wetlands in the northern Great Lakes (NGL) ecoregion and the Erie and Ontario Lake plains (EOL) ecoregion were used. Sites belonging to different hydrogeomorphic types (open lacustrine wetlands, protected

wetlands and river-influenced wetlands) were pooled for developing an ecoregion-based IBI. Eleven reference area sites and 16 GLEI sites in the EOL ecoregion, and 10 reference area sites and 18 GLEI sites in the NGL ecoregion were sampled for fish communities.

Fish sampling

Fish communities at 55 wetlands in the EOL and NGL ecoregions were sampled using 2 large fyke net arrays (1.25 cm mesh) and two small fyke net arrays (0.5 cm mesh) set overnight at each site following the methods of Brazner and Beals (1997). Each fyke net array was placed lead-to-lead (leads parallel to shore), with the wings set at 45° angles. One pair of large and one pair of small nets were placed near each of the two dominant shoreline habitats at a site. All fishes were identified to species, measured (total length to the nearest mm), examined for condition (incidence of damage or disease), and released. Specimens of uncertain identity were euthanized in clove oil, preserved in formalin-ethanol mixture, and returned to the laboratory for detailed examination.

IBI development

Based on a preliminary assessment of the native species composition of reference sites, as well as a literature review of studies conducted in coastal wetlands, streams, and rivers (Wilcox et al. 2002, Randall and Minns 2002, Minns et al. 1994, Karr 1981, Karr et al. 1986), a suite of 13 candidate metrics and 16 candidate metrics were compiled for the EOL and NGL ecoregion respectively. The candidate metrics were classified into 4 ecological categories - species richness and composition, indicator species, trophic composition, and abundance and condition (Karr 1981, Karr et al. 1986). The candidate metrics and the expected direction of change for each metric as a function of anthropogenic pressure are indicated in Table 2.1.

Metric values (based on raw values of each variable) were calculated for the reference wetlands in each ecoregion, and a matrix of Spearman rank correlation coefficients was created for the metrics proposed for each ecoregion. In the case where two metrics were significantly correlated ($p < 0.05$), the poorer discriminator of the two was discarded.

Table 2.1. List of candidate metrics and the expected responses for each metric in EOL and NGL ecoregions.

Candidate Metrics	**EOL**	**NGL**
Species richness and composition		
Number of species	Decrease[1,2]	Increase[3,4]
Number of native species	Decrease[1,2]	Increase[4]
Number of native minnows	Decrease[1,5]	Increase[3,6]
Number of sunfish species (except green sunfish)	Decrease[1]	Increase[4]
Number of sucker species (except white suckers)		Decrease[3]
Indicator species		
Percent tolerant species	Increase[1]	Increase[1]
Percent intolerant species	Decrease[1]	Decrease[1]
Percent of individuals as green sunfish	Increase[1,5]	Increase[1,5]
Percent of individuals as white suckers		Increase[4,7]
Percent of individuals as slimy sculpin		Decrease[6]
Trophic composition		
Percent of individuals as piscivores	Decrease[1]	Decrease[1]
Percent of individuals as insectivorous cyprinids	Decrease[1,5]	Decrease[1,5]
Percent of individuals as omnivores	Increase[1]	Increase[1]
Abundance and condition		
Number of individuals (CPUE standardized)	Variable[1]	Variable[1]
Percent non-indigenous species	Increase[8]	Increase[8]
Percent of individuals with DELT anomalies	Increase[1]	Increase[1]

[1]Karr 1981
[2]Simon and Lyons 1995
[3]Lyons et al. 1996
[4]Lyons 1992
[5]Karr et al. 1986
[6]Mundahl and Simon 1999
[7]Langdon 1988

The raw values of each metric were plotted against scales representing amounts of natural land cover, atmospheric deposition, shoreline modification, population density and point source discharge. If a metric was significantly correlated with the stressor ($p < 0.05$) then that metric was included as a component of the IBI. Metric-stress score correlations were determined for stress scores calculated at both the segment-shed scale and the watershed scale to test whether metrics and consequently, fish communities, more strongly reflect anthropogenic stress at large scales or more localized scales.

The basic criterion for developing a conceptually valid IBI was that it had to include at least one metric from each of the categories of species richness and composition, indicator species, trophic composition, and abundance and condition. If the criterion was met, the chosen metrics were scaled to generate an IBI scoring system. Natural breaks in the data were used to assign each test site with a score of 1, 3, or 5 for each metric that was calibrated, based on Karr's (1981) original IBI methods. The overall IBI scores of the test sites were compared to an integrity-rating table that was developed based on average IBI scores of reference and test sites. Four classes of integrity – excellent, good, fair and poor were developed based on Karr's (1981) original IBI.

As a test of IBI behaviour, it was determined how candidate metrics from both EOL and NGL varied in relation to other classes of stressors measured as part of the GLEI project. Metric scores were plotted were plotted against agricultural stress scores previously determined for each test wetland to determine if a candidate metric was sensitive to the stressor, agriculture, measured through the GLEI project. The IBI developed for NGL ecoregion was tested using fish community data from 3 randomly selected sites in the same ecoregion that were not part of the original IBI development. The IBI scores were compared against the sites' agricultural stress scores (segment-shed scale) to determine if the IBI could predict biotic integrity among the sites. If the IBI that was developed was a good measure of biotic integrity, sites with high biotic integrity would have low agricultural stress scores and vice versa.

Description of Metrics

The expected patterns of metric scores were based on literature citing the responses of certain metrics in warmwater regions as compared to coolwater regions.

Coolwater streams generally have a mean maximum daily temperature of 22- 24°C during the summer months whereas warmwater regions typically have mean maximum daily temperatures above 24°C. Temperature data recorded for EOL reference sites showed a mean ± SE maximum temperature of 26.6 ± 0.38 °C (n = 11), which conforms with the mean temperatures of most warmwater regions (Karr 1981, Karr et al. 1986, Fausch et al. 1984). The metric responses for sites in EOL were expected to be similar to those metrics noted in warmwater habitats. In contrast, the mean maximum temperature recorded in reference NGL sites was 23.7 ± 0.49 °C (n = 10), characteristic of most coolwater regions. The expected metric responses would thus be similar to those of coolwater streams (Leonard and Orth 1986, Mundahl and Simon 1995).

Species Richness and Composition

Total number of species - Total species richness is the most commonly used metric in biotic integrity assessment of aquatic habitats, and is defined as the total number of species sampled at the site under investigation (standardized by catch per unit effort) (Karr 1981). Species richness is expected to decline with increasing stress, particularly in warmwater regions (Simon and Lyons 1995). The species richness metric for EOL sites was expected to decline with increasing agricultural stress. For the NGL ecoregion however, the expected response was an increase in the number of species with increasing stress based on the rationale that cool or coldwater regions typically support fewer species that are adapted to the latitude and temperature regime of such habitats (Lyons et al. 1996). Lyons (1992) hypothesized that an increase in human related disturbance leads to an increase in temperature, which in turn can support a wider range of species that are otherwise intolerant of cold waters.

Native species richness – This metric measures the number of native species at a site and can sometimes be correlated with total number of species. Redundancy among metrics is however reduced by performing correlations between metrics and discarding one of the metrics that are highly correlated. A decline of native species richness can be indicative of degraded conditions as certain species can be intolerant to various stressors (Lyons 1992). The expected response is thus a decline in native species richness with increasing agricultural stress for both EOL and NGL ecoregions.

Number of native cyprinid species – Cyprinidae richness has often been used as a species richness metric in stream quality assessment (Karr 1981, Karr et al. 1986) and has more recently been applied to wetland quality assessment (Wilcox et al. 2002). The expected relationship in the EOL ecoregion is a decline in native minnow richness with increasing disturbance. Lyons et al. (1996) reported that environmental degradation of coldwater habitats results in increased species richness within some family of fishes. Mundahl and Simon (1999) used minnow richness as a metric in coldwater stream assessment and expected an increase in minnow richness in concert with increasing disturbance. The expected response for sites in the NGL ecoregion was thus an increase in minnow richness with increasing agricultural stress.

Number of sunfish species – Sunfish species, like minnows, are important indicators of wetland quality and are typically found in lower latitude regions (Jude and Pappas 1992). For this reason, the metric sunfish richness was unique to EOL ecoregion, and the expected response was a decline in sunfish species with increasing stress. However, green sunfish were not counted amongst sunfish richness due to its apparent tolerance to degraded water quality conditions and its dominance in impaired habitats (Karr 1981), green sunfish serve best as an indicator of impairment.

Number of sucker species - Karr (1981) proposed the use of species richness and composition metrics for most families of fish including catostomids. Lyons et al. (1996) reported significant but small differences in the number of sucker species between high quality and low quality sites. White suckers are considered to be a tolerant species and were thus not included in the sucker species richness count. The expected response of the metric was a decrease in sucker species with increasing stress.

Indicator species

Tolerance values for all fish species sampled were based on guidelines by U.S. EPA (Appendix II) (Barbour et al. 1999).

Percent tolerant species – Percent tolerant species is a measure of the percentage of species found in a site that are considered to be tolerant. A site with many tolerant species is indicative of degrading conditions (Karr 1981). The tolerant species metric was

used for both the EOL and NGL IBI and the expected response in both ecoregions is an increase in tolerant species with increasing stress.

Percent intolerant species – In contrast to percent tolerant species, the percent of intolerant species at a site is informative because a decline in intolerant species suggests that conditions of the site are deteriorating (Karr 1981). As such, the expected response in both ecoregion is a decline in intolerant species with increasing stress.

Percent green sunfish – Green sunfish is mostly abundant in mid-western streams of the United States and due to its tolerance of degraded habitats, serves as an important indicator species (Karr 1981, Karr et al. 1986). A survey of the native composition in reference condition wetlands showed that there was a low abundance of green sunfish sampled in both EOL and NGL, which led us to include percent green sunfish as an indicator metric. The expected response is an increase in green sunfish relative abundance with increasing agricultural stress.

Percent white suckers – This metric is a measure of the percent of the catch comprised of common white suckers. White suckers are considered tolerant of many different types of degradation and are typically found in high densities in impaired stream sites (Langdon 1988 and Lyons 1992). The metric, percent white suckers, was used as an indicator species metric for the NGL IBI. The expected response was an increase in white sucker relative abundance with increasing stress.

Percent of individuals as slimy sculpin– Slimy sculpins are generally found in coolwater or coldwater habitats and are considered to be an intolerant species (Mundahl and Simon 1999). As a result, a decrease in the percent of slimy sculpin individuals is expected with increasing agricultural stress in the NGL ecoregion.

Trophic Composition

Trophic guild designations for all species sampled were based on guidelines by U.S. EPA (Appendix II) (Barbour et al. 1999)

Percent of individuals as piscivores – Piscivores are particularly sensitive to impairment due to their position within the trophic structure, and their numbers tend to decline as habitat quality declines (Karr 1981). A decline in percent piscivores is expected with an increase in agricultural stress for both the EOL and NGL ecoregion.

Percent of individuals as insectivorous cyprinids – Karr (1981) and Karr et al. (1986) used this metric in stream quality assessment and typically, a decline in insectivorous cyprinids is indicative of degraded conditions.

Percent of individuals as omnivores – The dominance of omnivores tends to suggest a degradation in the trophic structure of a habitat, and as a result, indicates a degradation in the invertebrate community structure (Karr 1981). Greater relative abundance of omnivores is thus considered to be an indication of poor habitat conditions in both EOL and NGL ecoregions.

Abundance and Condition

Number of individuals – The abundance of fish in general is a relative metric and by itself, does not indicate good or bad habitat conditions (Karr 1981). However, when used in conjunction with other metrics, the number of individuals (catch per unit standardized) can serve as a good indicator of habitat structure. As such, a decline in the overall abundance of fish is expected with increasing stress at a site.

Percent non-indigenous species – The presence of non-indigenous species at a site is generally indicative of deteriorating conditions, as non-indigenous species are considered tolerant to adverse conditions and can sometime replace the native species in the habitat through competition and predation (Crumby et al. 1990). Percent non-indigenous species is a measure of the proportion of species that are non-native to the lake in which the site occurs. The expected response is an increase in relative number of non-indigenous species with increasing agricultural stress.

Percent of individuals with diseases, eroded fins, lesions and tumours (DELT) – This metric is as widely used in IBIs as is total species richness. The presence and especially abundance of fish with DELT anomalies is a sign of severe degradation at a site (Karr 1981). For the EOL and NGL ecoregion, an increase in fish with DELT is expected with increasing stress.

Results

A total of 62 species represented by 7,302 individuals were captured in the EOL ecoregion. The 13 candidate metrics were tested for redundancy with a Spearman rank correlation test. Metrics highly correlated with each other were rejected from further

analysis (Appendix III). When comparing metric response at the segment-shed level, two metrics varied significantly with respect to agricultural stress: percent intolerant species ($r = 0.45$, $p < 0.05$, $n = 27$) and percent of individuals as omnivores ($r = 0.42$, $p < 0.05$, $n = 27$) (Table 2.2). At the watershed level, again, only two metrics varied significantly as a function of agricultural stress: percent tolerant species ($r = -0.39$, $p < 0.05$, $n = 27$) and percent of individuals with deformities, eroded fins, lesions or tumours (DELT) ($r = 0.44$, $p < 0.05$, $n = 27$) (Table 2.2).

A total of 66 species represented by 19,054 individuals were collected in the NGL ecoregion. Metrics highly correlated with each other were discarded before further IBI analysis (Appendix IV). Five metrics varied significantly with agricultural stress at the segment-shed level: number of species ($r = 0.38$, $p < 0.05$, $n = 28$) (Fig. 2.1), percent intolerant species ($r = - 0.51$, $p < 0.01$, $n = 28$), percent green sunfish individuals ($r = 0.40$, $p < 0.05$, $n = 28$), percent non-indigenous species ($r = 0.62$, $p < 0.001$, $n = 28$) (Fig. 2.2) and percent of individuals as omnivores ($r = -0.51$, $p < 0.01$, $n = 28$) (Table 2.3). When comparing metric variation for the NGL ecoregion at the watershed level, 3 metrics correlated significantly to agricultural stress: percent intolerant species ($r = -0.52$, $p < 0.01$, $n = 28$), percent green sunfish individuals ($r = 0.46$, $p < 0.05$, $n = 28$) and percent insectivorous cyprinids ($r = -0.41$, $p < 0.05$, $n = 28$) (Table 2.3).

The basic criterion for IBI development was to include at least one metric from each ecological classification that correlated significantly with the intensity of the stressor against which it was tested. Since 5 metrics were found to show significant correlation at the segment-shed level for NGL ecoregion, each metric was calibrated to develop scores that would be summed to obtain an overall IBI score (Table 2.4). The following were the integrity rating classifications based on IBI score ranges indicated in parentheses– excellent (21-25), good (17-20), fair (12-16) and poor (5-11). Of the 18 GLEI sites in the NGL ecoregion, there was an equal distribution of sites classified as good, fair and poor. The total IBI scores were plotted against agricultural stress (measured at the segment-shed level. There was a negative, yet non-significant correlation between IBI scores and stress scores ($r = -0.33$, $p > 0.05$, $n=28$) (Figure 2.3.)

Table 2.2. Spearman r values between EOL metrics and agricultural stress measured at segment-shed scale and watershed scale. Metrics significantly correlated to agricultural stress ($p < 0.05$) are indicated with an asterisk (*).

Metrics	**r (segment-shed)**	**r (watershed)**
# native species	0.04	-0.04
% tolerant species	0.17	-0.39*
% intolerant species	0.45*	-0.09
% piscivores	0.03	0.25
% insectivorous cyprinids	0.30	0.20
% omnivores	0.42*	0.30
CPUE	-0.04	-0.28
% DELT	0.23	0.44*

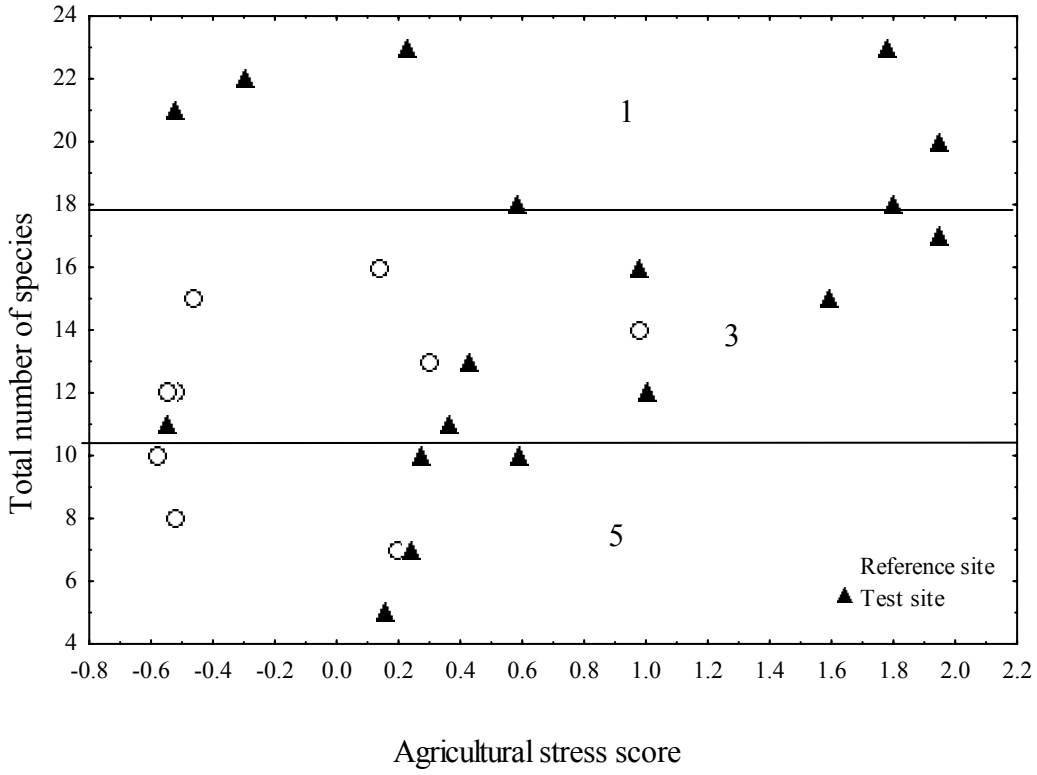

Figure 2.1. Relationship between total number of species and GLEI agricultural stress score at the segment-shed level for NGL ($r = 0.38$, $p < 0.05$, $n=28$). The metric was calibrated so that sites with high number of species received a score of 1, sites with low number of species received a score of 5 and sites in between received a score of 3.

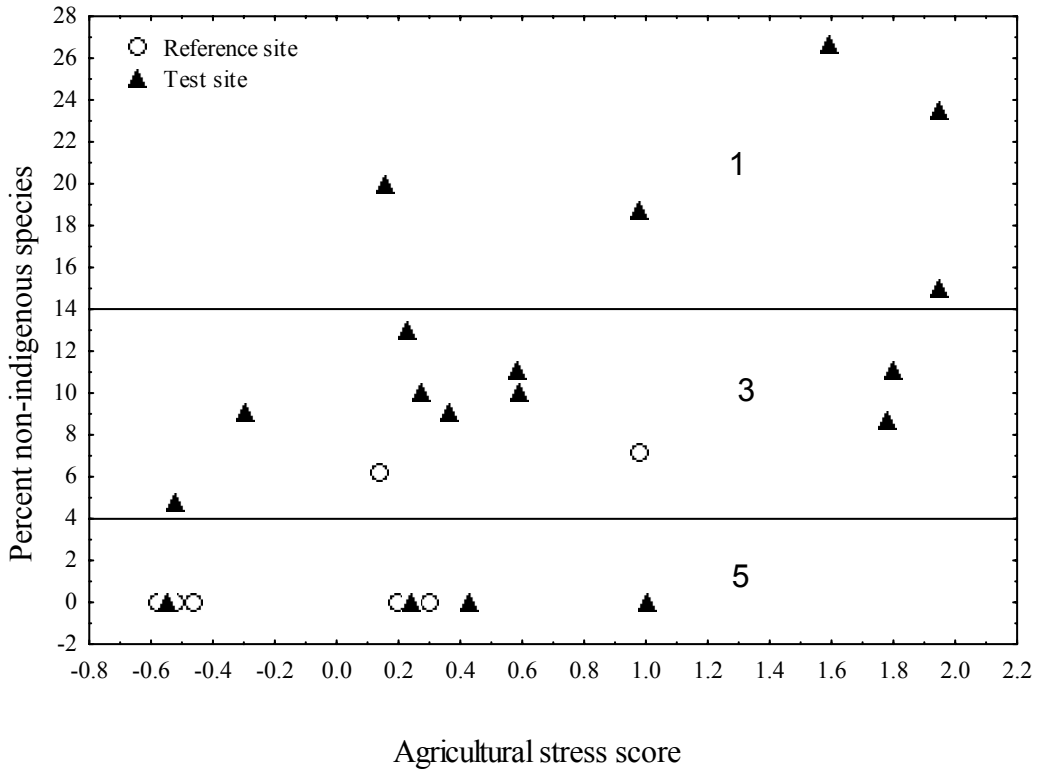

Figure 2.2. Relationship between percent non-indigenous species and GLEI agricultural stress score at the segment shed level for NGL ($r = 0.62$, $p < 0.001$, $n=28$). The metric was calibrated so that sites with high percent non-indigenous species received a score of 1, sites with low percent non-indigenous species received a score of 5 and sites in between, received a score of 3.

Table 2.3. Spearman r values between NGL metrics and agricultural stress measured at segment-shed scale and watershed scale. Metrics were significantly correlated to agricultural stress if $p < 0.05$. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$.

Metrics	**r (segment-shed)**	**r (watershed)**
# species	0.38*	0.28
# sunfish species	0.20	0.27
% tolerant species	0.04	0.04
% intolerant species	-0.51**	-0.52**
% green sunfish	0.40*	0.46*
% piscivores	-0.26	0.05
% insectivorous cyprinids	-0.09	-0.42*
% omnivores	-0.51**	-0.35
% non-indigenous	0.62***	0.36

Table 2.4. Scoring criteria for preliminary fish IBI for NGL wetlands.

Metric Scoring

1. Total number of species:

>18 = 1	11 to 18 = 3	≤10 = 5

2. Number of intolerant species:

< 2 = 1	2 to 12 = 3	>12 = 5

3. Percent of individuals as green sunfish:

>10 = 1	2 to 10 = 3	< 2 = 5

4. Percent of individuals as omnivores:

>40 = 1	10 to 40 = 3	<10 = 5

5. Percent non-indigenous species:

>14 = 1	4 to 14 = 3	< 4 = 5

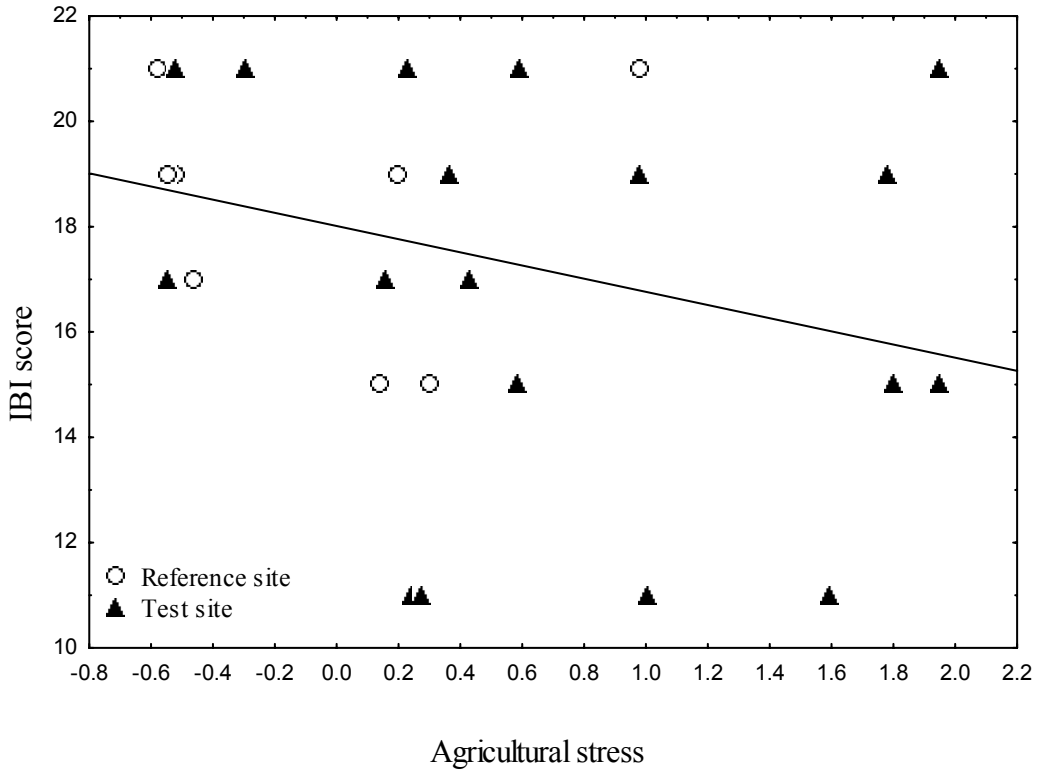

Figure 2.3. Relationship between overall IBI scores and GLEI agricultural stress at segment-shed level for NGL sites ($r = -0.33$, $p > 0.05$, $n=28$).

When testing the response of candidate metrics from EOL and NGL to other types of stressors, the results showed only one or two metrics that were significantly correlated to

the other stressors (Appendix V). However, there was no clear pattern of increasing metric response with increasing stress.

IBI Testing

Three NGL wetlands that were sampled as part of the GLEI project but not included in the original IBI development were selected for testing the NGL ecoregion IBI. All three sites were classified as excellent in terms of their integrity rating with total IBI scores of 21, 23 and 21 and their corresponding agricultural stressor scores were -0.35, -0.56 and 0.24 respectively.

There was a non-significant, yet negative correlation between the IBI scores and agricultural stress scores for the 3 NGL sites ($r = -0.70$, $p > 0.05$). The site with the highest IBI score of 23 also had the lowest agricultural stress score (-0.56), as was expected. However, with only 3 sites available to test the IBI, it could not be confirmed whether IBI scores would consistently decline with increasing stress.

Discussion

A principal components analysis of the 212 environmental stressors that were measured as part of the GLEI project (Danz et al. 2005) showed that agricultural stress explained the greatest amount of variance in the original data. Agricultural stress was chosen to test for metric robustness to clearly differentiate sites with high biotic integrity, from sites with low biotic integrity. In addition to using published literature to select the candidate metrics (Karr 1981, Minns et al. 1994, Wilcox et al. 2002), the species composition of reference sites was used based on the rationale that if a site is relatively undisturbed (i.e. similar to a reference condition site), the fish assemblage in that site should be similar to that of a reference condition site (Reynoldson et al. 1997), provided the sites belong to the same ecoregion, and are of similar hydrogeomorphic types. Minns et al. (1994) had used biomass indices for trophic composition metrics. However, since biomass data were not available for this study, percent measures were used for similar metrics tested by Minns et al. (1994).

Ecoregion was chosen as the regional framework for developing the IBIs based on the contention that sites within an ecoregion have unique fish assemblages, characteristic

of the unique physical, chemical and biological attributes of that ecoregion (Omernik 1987). Some researchers however have argued that hydrogeomorphic type (HGM) is a better predictor of metric responses and an individual IBI is thus warranted for each unique HGM type. Based on previous studies using an ecoregion framework (Hughes et al. 1986, Omernik and Griffith 1991, Simon 1991) and to determine if a unique IBI could indeed be developed for wetlands belonging to the same ecoregion, I decided to use ecoregion as a stratification tool for metric development.

The metrics evaluated for fish catches in the EOL ecoregion did not show promising results. The criterion for deriving an IBI was to have at least one metric from each ecological classification show a significant response to the stressor. Percent intolerant species from the indicator species category and percent omnivores from the trophic composition category varied significantly with respect to agricultural intensity, but since at least 1 metric was required from each ecological category, IBI development was not pursued further for the EOL ecoregion. Most of the candidate metrics did not show a significant response to the stressor, suggesting that either there was no significant difference between the metric scores of reference and stressed sites, or that the metrics were not associated with agricultural stress. It is possible that due to the extensive amount of agricultural development particularly in the Lake Erie basin (Johnson et al. 1989, Host et al. in press), the reference sites chosen for the study may not have been representative of truly unimpaired conditions, thus skewing the metric scores and producing non-significant responses. There was no difference between metric response patterns when stress was calculated at the segment-shed level than when it was estimated for the watershed level, suggesting that for the EOL ecoregion, fish assemblages do not differ in the boundaries of their response to stress.

Five metrics varied significantly in response to stress at the segment-shed level for the NGL ecoregion, meeting the criteria for IBI development, and thus allowing for calibration of metric scores to establish integrity ratings for the test sites. Based on the integrity rating table, there was an equal distribution of sites that were classified as good, fair and poor. However, the overall correlation between IBI scores and agricultural stress was weak and the noise in the scatter plot could likely be attributed to the fact that sites were pooled together from different hydrogeomorphic types leading to a high amount of

variance in the data. When comparing the scale of response, variation in fish community IBI in the NGL ecoregion was more highly correlated to stress measured at larger segment-shed scales than at the more localized watershed scales. The lack of any significant metric response to the other 5 stressors shows that fish communities in both EOL and NGL ecoregion respond most strongly to agriculture related stress.

Most researchers to date have used disturbance gradients related to percentage of land in agriculture, percentage of land having urban development, and combinations of land uses at the watershed level (Wilcox et al. 2002, Mundahl and Simon 1995) and few have evaluated disturbance measures at two different scales (Uzarski et al. 2005). The scales of metric response were compared to determine whether stress at local scales (i.e. watershed scales) is more important in affecting fish communities than stress measured across larger (segment-shed) scales. Holland et al. (in preparation) showed that correlations between fish abundance and intensity of landuse for most fish species of the Great Lakes reach their greatest values at large scales. This study showed in the NGL ecoregion, more metrics varied significantly against stress at the segment shed scale when compared to the watershed scale of response, consistent with results by Holland et al. (in preparation).

A further step toward developing an effective IBI involves testing the IBI on sites that are independently sampled and classified, and were not used in the IBI development (Simon 1995). Since each site that was part of the GLEI project was sampled independently of other sites, 3 wetlands in the NGL ecoregion that were not included in the original IBI development were reserved for IBI testing. The IBI scores for the 3 wetlands were negatively correlated with their corresponding agricultural stress scores, which is consistent with expectations of the IBI's efficacy in assessing biotic integrity in the NGL ecoregion. However, a greater number of sites would be required to make the observed relationship statistically significant, and the reliability of the IBI developed for NGL cannot be ascertained until it is tested on a set of independently sampled sites in the same ecoregion.

Summary and Conclusions

The objectives for the paper were to test the applicability of the classical IBI approach (Karr 1981) in assessing wetland condition in the Great Lakes. An IBI for the EOL ecoregion could not be developed using the reference condition approach, possibly due to the large amount of agricultural stress present even in the segment-sheds of EOL reference sites. This may not have allowed for clear boundary separation of sites with high biotic integrity sites from sites with low biotic integrity. There is potential to develop an IBI for the NGL ecoregion in that sufficient metrics were available to permit IBI development, and calibrate the IBI in NGL wetlands. Detailed stressor information obtained at both segment-shed scales and watershed scales allowed us to test the metrics' sensitivity to a range of anthropogenic stressors as well as determine the scale at which fish communities are most responsive to that stress. The recommendations based on the results of this IBI development exercise are to test the IBI further by using fish data collected at sites with similar hydrogeomorphic types and in the same ecoregion.

CHAPTER 3

Testing a Fish Index of Biotic Integrity for Great Lakes Wetlands Stratified by Plant Zones

Abstract

Fish community composition is often segregated along ecoregions, lakes or hydrogeomorphic types. However, attempts to develop an Index of Biotic Integrity (IBI) for environmentally homogeneous sites of the Great Lakes have had only limited success. Recently, Uzarski et al. (2005) used correspondence analysis to determine that the primary correlate in coastal wetland fish community composition is emergent plant zonation. Consequently, they developed an IBI for sites dominated by (>50% cover) *Typha* (cattail) vegetation and a separate IBI for sites dominated by *Scirpus* (bulrush). We calculated Uzarski et al. IBI scores for 33 Great Lakes wetland sites with dominant *Typha* and *Scirpus* vegetation, using overnight sets of fyke nets. The sites had been selected to fall across gradients of population density, road density, urban development, point source pollution, and agriculture measured using a GIS-based analysis of land use. Sites subject to low levels of disturbance had high IBI scores. The *Typha*-specific IBI was most highly negatively correlated with a disturbance variable that combined population density, road density and urban development, whereas the *Scirpus*-specific IBI negatively correlated most strongly with agricultural intensity. The Uzarski IBI appears to be an effective indicator of some but not all classes of anthropogenic disturbance at Great Lakes coastal margins.

Introduction

Wetlands, an integral link between lakes and surrounding land areas, provide unique habitats for birds, amphibians, fish and benthic invertebrates. Yet, wetlands continue to be lost at an alarming rate. The United States alone loses 60,000 acres of wetlands each year (U.S. Environmental Protection Agency 2004). Almost two-thirds of southern Ontario wetlands have been lost or severely degraded as a result of agricultural run off, development, toxic runoff, exotic species invasion, and artificial changes in water levels (Environment Canada 2002). Measuring the biotic integrity of wetlands has been a primary focus for many researchers since the passing of the Great Lakes Water Quality Agreement in 1972 and the recent partnership of US EPA and Environment Canada through biannual State of the Lake Ecosystem Conference (SOLEC) meetings held since 1994. Natural ecosystems such as wetlands are exposed to a multitude of stress at varying levels, making it difficult to attribute degradation to specific causes. It is therefore important to identify the various types of stress present at a site before developing indicators to assess the condition of a site (Danz et al. 2005).

The Index of Biotic Integrity (IBI) approach has been widely used in assessing the ecological condition of streams and rivers across the United States (Karr 1981, Karr et al. 1986, Fausch et al. 1984, Fausch et al. 1990) and has been successfully modified for use in assessing coldwater and coolwater streams (Leonard and Orth 1986, Mundahl and Simon 1999). However, the applicability of the IBI approach to wetlands has been questioned and despite a few successful attempts (Thoma 1999, Randall and Minns 2002), Wilcox et al. (2002) concluded that due to constant water level fluctuations experienced in wetlands, a wetland IBI developed in one year would not necessarily be applicable in subsequent years. Using multivariate techniques such as correspondence analysis and nonmetric multidimensional scaling, Uzarski et al. (2005) determined that fish community composition in Great Lakes wetlands was more strongly correlated with plant zones rather than ecoregion, Great Lake, or wetland type, as was previously believed and applied in IBI development (Hughes and Gammon 1987, Lyons 1989, Simon 1991). Consequently, Uzarski et al. (2005) developed an IBI for each of 2 major emergent plant zones of Great Lakes wetlands. They argued that an IBI derived from plant zonations would be most applicable to areas with similar plant cover and thus,

would only be valid for testing similar wetlands. Using data collected from wetlands varying in chemical/physical parameters and surrounding landuse/cover measurements, Uzarski et al. (2005) ultimately described distinct candidate metrics for two separate IBIs – one pertaining to *Typha*-dominated areas and one relating to *Scirpus*-dominated sites.

Previous attempts to develop IBIs for the Great Lakes wetlands (see Chapter 2) were only partially successful in that an ecoregion-based IBI could be derived to assess ecological condition in the northern ecoregion of the Great Lakes, but not for the Lake Erie and Lake Ontario wetlands, possibly due to the lack of undisturbed reference sites particularly in the lower Lake Erie basin. However, the data collected during that study are suitable to evaluate Uzarski et al.'s (2005) plant-zone based IBIs. Furthermore, the stratified-random site-selection design of our study permitted us to determine which of 6 classes of anthropogenic disturbance were most strongly reflected by variation in IBI scores. Most researchers to date have quantified primarily agricultural and land use as disturbance measures affecting fish communities (Brazner and Beals, 1997, Crosbie and Chow-Fraser 1999). The overall objective in this paper was to examine how values of the Uzarski et al. (2005) *Typha* and *Scirpus* IBI varied across a broad suite of anthropogenic stressors derived from GIS-recorded data measured at the segment-shed scale (Danz et al. 2005). The objective was met by firstly applying the Uzarski et al. (2005) *Typha* and *Scirpus* IBIs to fish data recorded through the GLEI project and to relate the IBI scores to the GLEI stressor scores from 6 categories of stress (see General Methods, Chapter 1). Secondly, we used Uzarski et al.'s (2005) *Scirpus* sites data (site locations and IBI scores) to evaluate the overall pattern of stress response and to determine how consistent the two studies were in the response to various GLEI stressors.

Methods

Site Selection

A total of 82 coastal wetlands were sampled along U.S. Great Lakes coastal margins in 2002 and 2003, according to a design that balanced effort across Lakes, hydrogeomorphic wetland type, and degree of local anthropogenic disturbance (Danz et al. 2005) (See Chapter 1). To account for the hydrogeomorphic differences with respect to the influence of the watershed versus the lake, coastal wetlands were classified as

lacustrine, protected or river-influenced (as per Keough et al. 1999). All 3 types of coastal wetlands sampled across the entire U.S. coastline (as per GLEI and Reference project sampling design; see Chapter 1) were assessed in this study.

Fish Sampling

Fish communities were sampled using 2 large fyke net arrays (1.25 cm mesh) and 2 small fyke net arrays (0.5 cm mesh) set overnight at each site. Each fyke net array was placed lead-to-lead (leads parallel to shore), with the wings set at 45° angles (Brazner and Beals 1997). One set of large and one set of small nets were placed near each of the two dominant shoreline habitats at a site. Fish community composition (numbers of individuals of each species) and condition of up to 25 fish per species (total length, incidence of damage or disease) were measured at each fyke net the next day. Unidentifiable fish were euthanized in clove oil, preserved in 9:1 v/v ethanol: formalin mixture, and taken to the lab for identification. Physicochemical variables (temperature, dissolved oxygen concentration, conductivity pH) were measured at each net using a multi-probe meter (YSI 556 MPS). Water clarity at each net was measured using a Secchi disk and turbidity tube. Dominant and subdominant genera of emergent, submergent and floating vegetation (cover and density) were also noted at each net per site.

IBI Application (using GLEI fish data)

To test the IBI developed by Uzarski et al. (2005), 36 sites in which *Typha* and *Scirpus* were dominant genera were used in this analysis. Uzarski et al. (2005) included sites that they identified visually as having >50% of the wetland area covered by emergent vegetation at the whole site level. In this study, vegetation density and cover were noted at the net level rather than the site level; therefore, the criterion for dominant cover was lowered to 30% or greater. In doing so, a greater number of sites was included for the IBI testing. In all, 23 sites were identified as having dominant *Typha* cover, and 13 sites met the criterion for dominant *Scirpus* vegetation across the entire U.S. Great Lakes shoreline (Figure 3.1). All fish species counts were standardized by catch per unit effort (CPUE). Uzarski et al.'s (2005) *Typha* IBI and *Scirpus* IBI metrics were

calculated for data collected from the GLEI sites; to maintain consistency, we followed Uzarski et al.'s (2005) trophic guild designations for fish species. For species not listed in Uzarski et al. (2005), we followed the trophic guild designations of the U.S. EPA (Barbour et al. 1999) (Appendix II).

Spearman's rank correlations were performed to determine if there was any relationship between IBI scores and each of the 6 stressor variables measured at segment-shed scales. Comparisons were also done between IBI scores and watershed-based stressor scores; however, there were no significant correlations, nor was there any clear pattern in the distribution of metric scores with increasing stress.

If the Uzarski et al. (2005) *Typha* and *Scirpus* IBI contained robust metrics that reflected all types of anthropogenic disturbance (i.e., agriculture, shoreline modification, land cover, atmospheric deposition), then one would expect to find a significant negative correlation between the IBI scores and the stressor scores for each of the stressor variables. To assess the relationship between GLEI stressors and the fish data collected by Uzarski et al. (2005), we overlaid geographic co-ordinates for each site sampled by Uzarski et al. (2005) on a map of our segment-shed delineations. Once we had ascertained the segment-shed in which each Uzarski site fell, we could determine its corresponding GLEI stressor score. Spearman rank correlation were calculated and graphs of the combined distribution of IBI scores from Uzarski et al. (2005) *Scirpus*-dominant sites and GLEI *Scirpus*-dominant sites were correlated with the 6 GLEI stressor scores. Since the stressor information was only available for sites located on the U.S. coastline, we were only able to compare Uzarski et al. (2005) data with GLEI data for Scirpus-dominant sites.

All statistical analyses and graphical interpretations were performed using Statistica ® software package, version 6.0 (StatSoft Inc. 2001).

Results

A total of 53 fish species represented by 3,045 individuals in *Typha*-dominant sites, and 45 fish species represented by 2,026 individuals in *Scirpus*-dominant sites were collected and identified (Table 3.1).

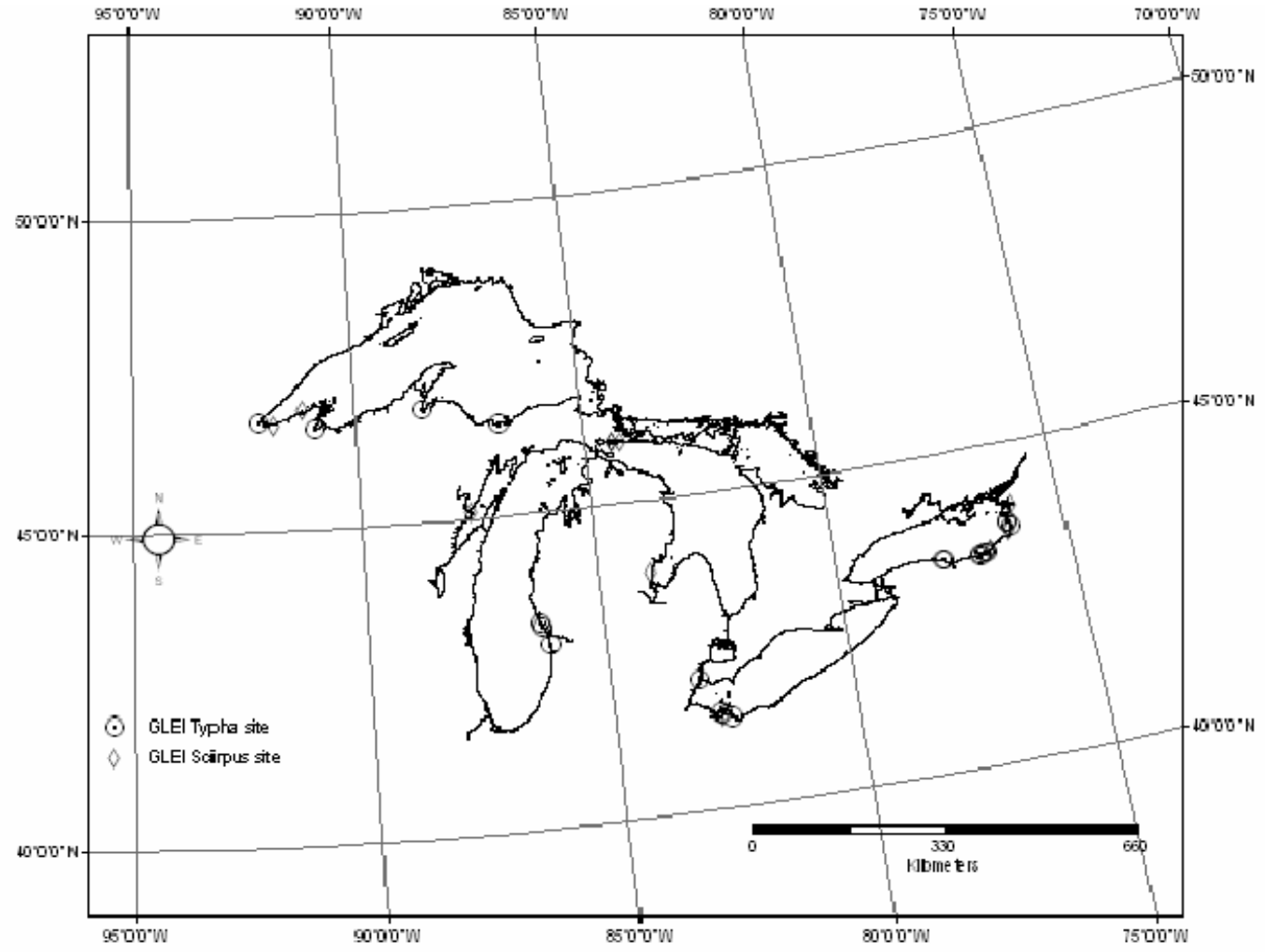

Figure 3.1 Map of the Great Lakes showing GLEI *Scirpus* and GLEI *Typha* dominant sites.

Table 3.1. Trophic guild (Tr_guild) and Catch per unit effort (CPUE of all fish species collected in fyke nets between 2001 and 2003 in *Typha* and *Scirpus* dominant wetlands of all 5 Great Lakes (* and ** denotes species unique to *Typha* and *Scirpus* sites respectively. Piscivores (PISC), Omnivores (OMN), Insectivores (INS), Herbivores (HERB), Zoobenthivores (ZOOBEN) and Molluscivores (MOLL).

Common name	**Family name**	**Species**	**Tr_ Guild**	***Typha* sites CPUE**	***Scirpus* sites CPUE**
Bowfin	Amiidae	*Amia calva*	PISC	78	5
Carpsucker	Catostomidae	*Carpiodes carpio*	OMN	0	6
White Sucker	Catostomidae	*Catostomus commersoni*	OMN	254.75	662.5
Bigmouth Buffalo*	Catostomidae	*Ictiobus cyprinellus*	INS	0.5	0
Spotted Sucker*	Catostomidae	*Minytrema melanops*	INS	0.50	0
Silver Redhorse*	Catostomidae	*Moxostoma anisurum*	INS	0.75	0
Shorthead Redhorse**	Catostomidae	*Moxostoma macrolepidotum*	INS	0	0.5
N. Rock Bass	Centrarchidae	*Ambloplites rupestris*	PISC	19.5	85.7
Warmouth*	Centrarchidae	*Chaenobryttus gulosus*	PISC	6.5	0
Green Sunfish	Centrarchidae	*Lepomis cyanellus*	INS	4.5	3.5
Pumpkinseed Sunfish	Centrarchidae	*Lepomis gibbosus*	INS	617.1	79.75
Bluegill Sunfish	Centrarchidae	*Lepomis macrochirus*	INS	926.7	90.75
Smallmouth Bass	Centrarchidae	*Micropterus dolomieui*	PISC	0.75	17.6
Largemouth Bass	Centrarchidae	*Micropterus salmoides*	PISC	172.7	16.75
White Crappie	Centrarchidae	*Pomoxis annularis*	PISC	13.4	1.5
Black Crappie	Centrarchidae	*Pomoxis nigromaculatus*	PISC	24.75	4
Alewife*	Clupeidae	*Alosa pseudoharengus*	OMN	1.5	0
Gizzard Shad*	Clupeidae	*Dorosoma cepedianum*	OMN	18.25	0
Eastern Slimy Sculpin*	Cottidae	*Cottus cognatus*	INS	0.5	0
Common Stoneroller*	Cyprinidae	*Campostoma anomalum*	HERB	0.5	0
Goldfish	Cyprinidae	*Carassius auratus*	OMN	12.25	2
Finescale Dace*	Cyprinidae	*Chrosomus neogaeus*	INS	0.5	0
European Carp	Cyprinidae	*Cyprinus carpio*	OMN	7.25	38.9
Brassy Minnow**	Cyprinidae	*Hybognathus hankinsoni*	OMN	0	36.7
Silvery Minnow**	Cyprinidae	*Hybognathus nuchalis*	HERB	0	0.5
Horneyhead Chub	Cyprinidae	*Nocomis biguttatus*	INS	1	3.5

Table 3.1 Continued

Common name	**Family name**	**Species**	**Tr_ Guild**	***Typha* sites CPUE**	***Scirpus* sites CPUE**
Common Shiner**	Cyprinidae	*Notropis cornutus*	INS	0	4
Blackchin Shiner*	Cyprinidae	*Notropis heterodon*	OMN	2.5	0
Blacknose Shiner	Cyprinidae	*Notropis heterolepis*	INS	9	45
Spottail Shiner	Cyprinidae	*Notropis hudsonius*	INS	12.25	5.5
Spotfin Shiner*	Cyprinidae	*Notropis spilopterus*	ZOOBEN	4.25	0
Sand Shiner**	Cyprinidae	*Notropis stramineus*	INS	0	0.5
Northern Mimic Shiner**	Cyprinidae	*Notropis volucellus*	INS	0	98
Bluntnose Minnow	Cyprinidae	*Pimephales notatus*	OMN	25.5	11
Northern River Shiner*	Cyprinidae	*Notropis blennius*	INS	0.5	0
Golden Shiner	Cyprinidae	*Notemigonus crysoleucas*	OMN	101.7	10.5
Emerald Shiner	Cyprinidae	*Notropis atherinoides*	INS	56	9
Bridled Shiner**	Cyprinidae	*Notropis bifrenatus*	INS	0	7

Although most of the taxa sampled were common among GLEI sites and Uzarski et al. (2005) sites, we did not find any pugnose shiner (*Notropis anogenus*) or longnose gar (*Lepisosteus osseus*) at any of our sites and as a result, the metrics scores for these two metrics remained 0. Longnose gar, although not very abundant in the GLEI samples, were captured at non-*Typha* dominant sites in northern Lake Huron and Lake Erie.

However, pugnose shiners were not captured at any GLEI sites. The most abundant species captured in both *Typha*- and *Scirpus*-dominant sites were white sucker, yellow perch, bluegill sunfish and pumpkinseed sunfish (Table 3.1). Some species were unique to each dominant plant zone type. For instance, alewife (*Alosa pseudoharengus*) were found only in *Typha*-dominant sites, whereas common shiner (*Notropis cornutus*) and bridled shiner (*Notropis bifrenatus*) were restricted to *Scirpus*-dominant sites (Table 3.1). The total IBI scores for *Typha*-dominant sites ranged from 6-42, with the highest possible score being 61 (Table 3.2). The total IBI scores for *Scirpus*-dominant sites ranged from 29-58, with the highest possible IBI score being 72 (Table 3.2).

Spearman rank correlations revealed that values of the *Typha* IBI was significantly correlated with the population density variable ($r = -0.70$, $p < 0.01$, $n = 19$) (Figure 3.2) and atmospheric deposition ($r = 0.60$, $p < 0.01$, $n = 19$) (Figure 3.3), whereas the *Scirpus* IBI was significantly negatively correlated to agriculture ($r = -0.64$, $p < 0.05$, $n = 13$) (Figure 3.4) and point source discharge ($r = -0.57$, $p < 0.05$, $n = 13$) (Figure 3.5) (Table 3.3). The scatter plots of IBI scores against stressor scores were also examined for patterns of linearity. Values appeared to vary linearly between *Typha* IBI scores and population density and *Typha* IBI scores atmospheric desposition respectively (Figures 3.2 and 3.3). In contrast, *Scirpus* IBI scores appeared to reflect a threshold effect (Figure 3.4) at agricultural stress scores of -0.5 or less, whereas at stress scores > 0, there were no IBI score greater than 45. Similarly, the pattern was more consistent with a threshold effect rather than a linear response (Figure 3.5). At a point source discharge score of -0.4 or greater, there were no IBI scores > 45.

As a further step in the IBI analysis, the IBI values obtained from using GLEI data were combined with the dataset of Uzarski et al. (2005) for *Scirpus* dominant sites (Figure 3.6).

Table 3.2. Site locations and IBI scores for *Typha*- and *Scirpus*-dominant sites sampled through the GLEI project.

Site name	**Latitude**	**Longitude**	**Plant zone**	**Lake**	**IBI score**
Dwights Point	46.7	-92.17	*Typha*	Superior	20
Prentice Park	46.58	-90.91	*Typha*	Superior	19
Baraga	46.76	-88.48	*Typha*	Superior	19
Au Train	46.44	-86.82	*Typha*	Superior	18
Port Shelden	42.88	-86.21	*Typha*	Michigan	19
Mona lake	43.17	-86.3	*Typha*	Michigan	26
Muskegon Lake	43.24	-86.35	*Typha*	Michigan	25
Little Muddy Creek	41.5	-82.8	*Typha*	Erie	37
Plum Brook	41.42	-82.62	*Typha*	Erie	19
Sodus Creek	43.27	-76.93	*Typha*	Ontario	37
Mudge Creek	43.29	-76.89	*Typha*	Ontario	31
South Sandy Creek	43.71	-76.2	*Typha*	Ontario	40
Cherry Island	42.06	-83.19	*Typha*	Erie	6
Wolcott Creek	43.3	-76.84	*Typha*	Ontario	22
Red Creek	43.31	-76.79	*Typha*	Ontario	25
Braddock Bay	43.31	-77.72	*Typha*	Ontario	26
Buttonwood Creek	43.3	-77.71	*Typha*	Ontario	22
Blind Creek	43.65	-76.16	*Typha*	Ontario	29
South Sandy Creek	43.71	-76.2	*Typha*	Ontario	42
Middle River	46.68	-91.82	*Scirpus*	Superior	58
Clover	46.88	-91.17	*Scirpus*	Superior	35
McKay Creek	45.99	-84.34	*Scirpus*	Huron	46
Menominee River	45.09	-87.59	*Scirpus*	Michigan	42
Pinconning	43.85	-83.92	*Scirpus*	Huron	32
Little Pickerel Creek	41.46	-82.79	*Scirpus*	Erie	36
Sterling Creek	43.35	-76.68	*Scirpus*	Ontario	44
Deer Tick Creek	43.61	-76.19	*Scirpus*	Ontario	58
Skinner Creek	43.67	-76.18	*Scirpus*	Ontario	46
Black River	43.99	-76.06	*Scirpus*	Ontario	29
Sterling Creek	43.35	-76.68	*Scirpus*	Ontario	31
Bear Lake	45.97	-84.16	*Scirpus*	Huron	46
McKay Creek	45.99	-84.34	*Scirpus*	Huron	46

Table 3.3. Spearman correlations between *Typha* (n = 19) and *Scirpus* (n= 13) IBI scores based on GLEI fish catches and values of 6 anthropogenic stressors measured at segment shed levels. Significant correlations ($p < 0.05$) are denoted by an asterisk (*).

Stressor	**Typha IBI**	**Scirpus IBI**
Agriculture	0.02	-0.64*
Land cover	-0.38	0.15
Population density	-0.70*	-0.21
Point source discharge	-0.09	-0.57 *
Atmospheric deposition	0.60 *	0.04
Shoreline modification	-0.24	-0.13

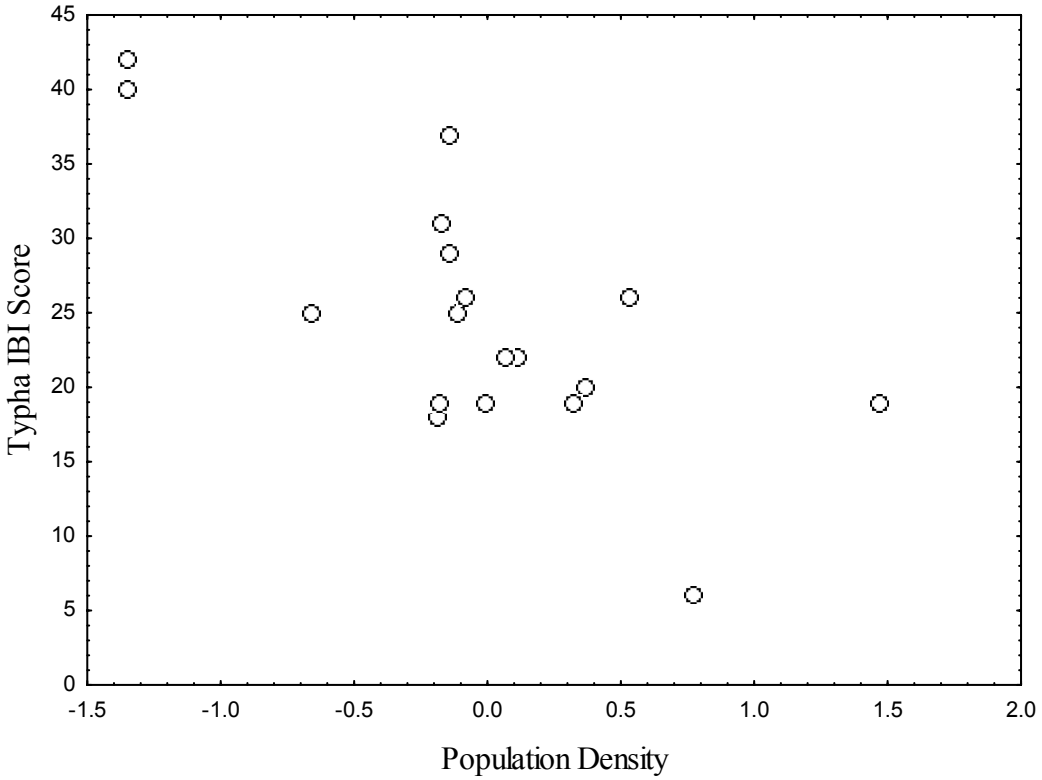

Figure 3.2. Relationship between *Typha* IBI and population density stress for GLEI ($r = -0.70$, $p < 0.05$)

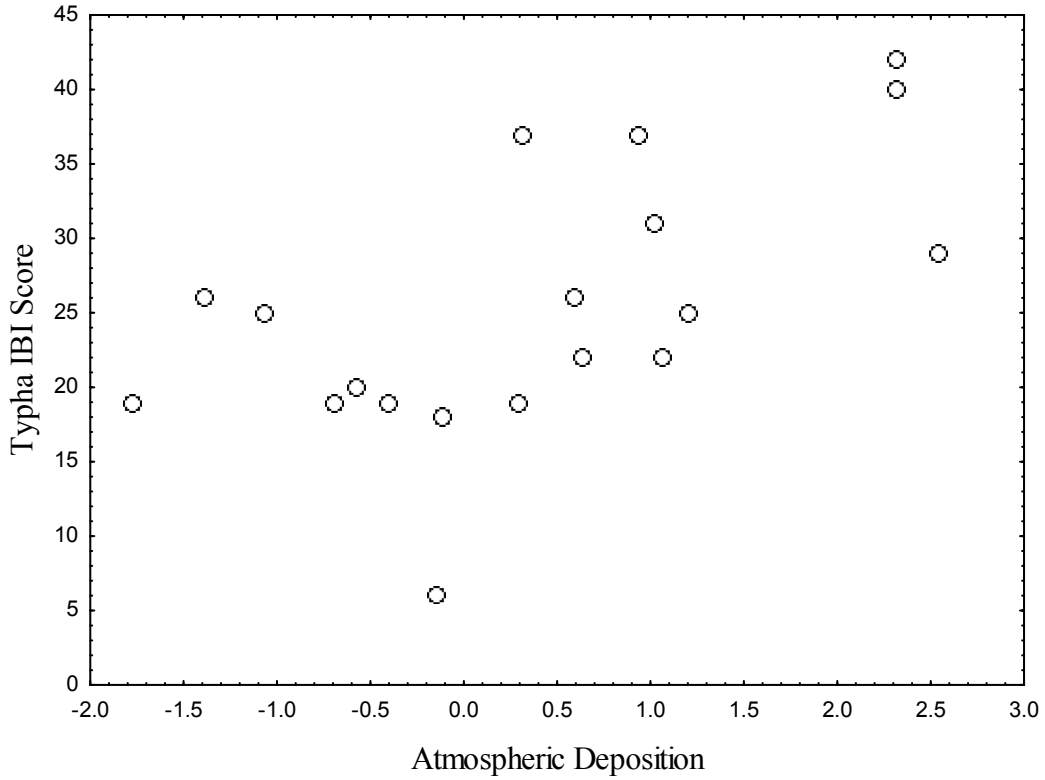

Figure 3.3. Relationship between *Typha* IBI and atmospheric deposition for GLEI sites ($r = 0.60$, $p < 0.05$)

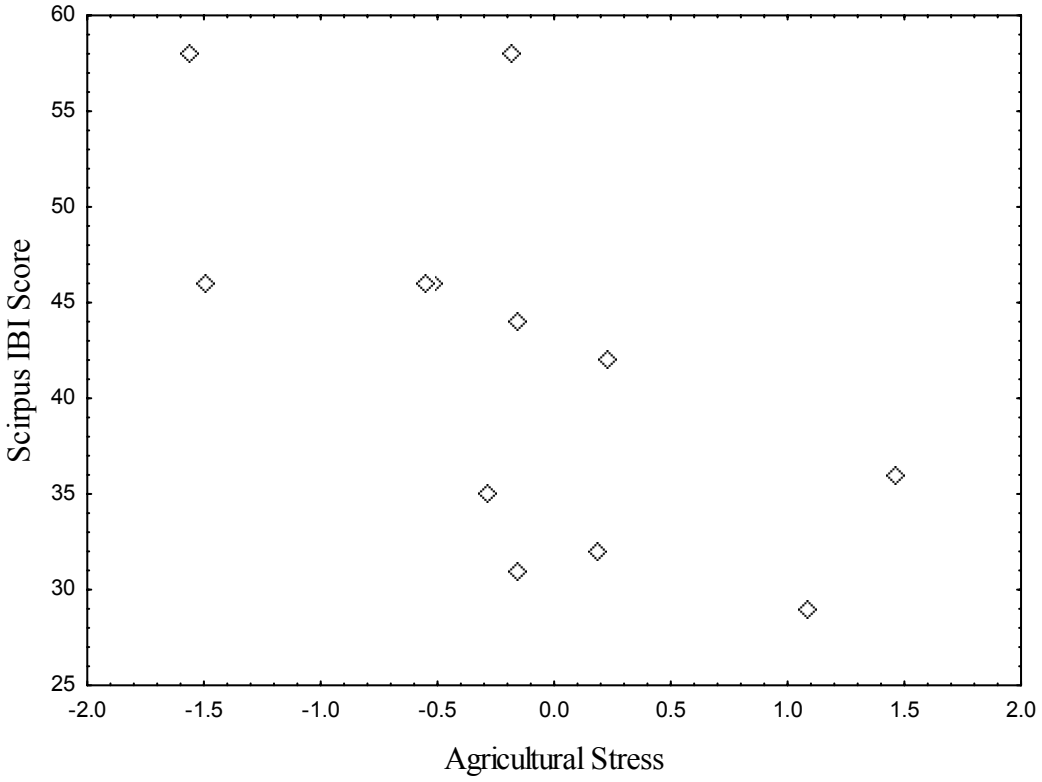

Figure 3.4. Relationship between *Scirpus* IBI and agricultural stress calculated for GLEI sites ($r = -0.64$, $p < 0.05$, $n = 13$).

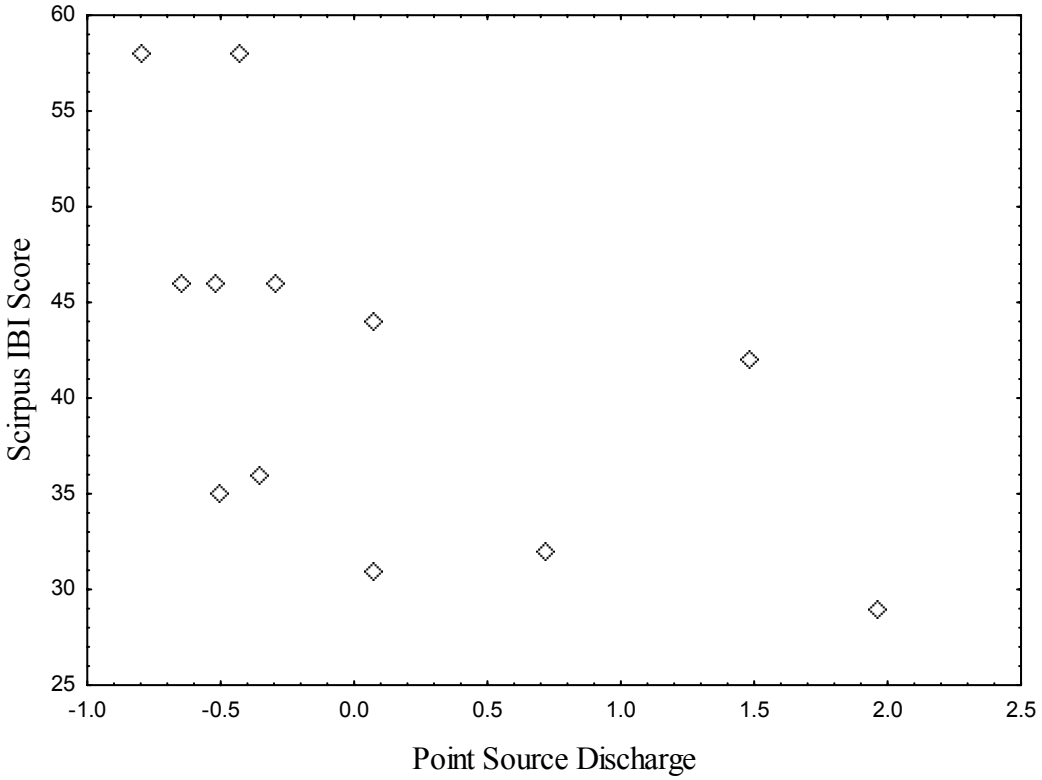

Figure 3.5. Relationship between *Scirpus* IBI and point source discharge for GLEI sites ($r = -0.57$, $p < 0.05$)

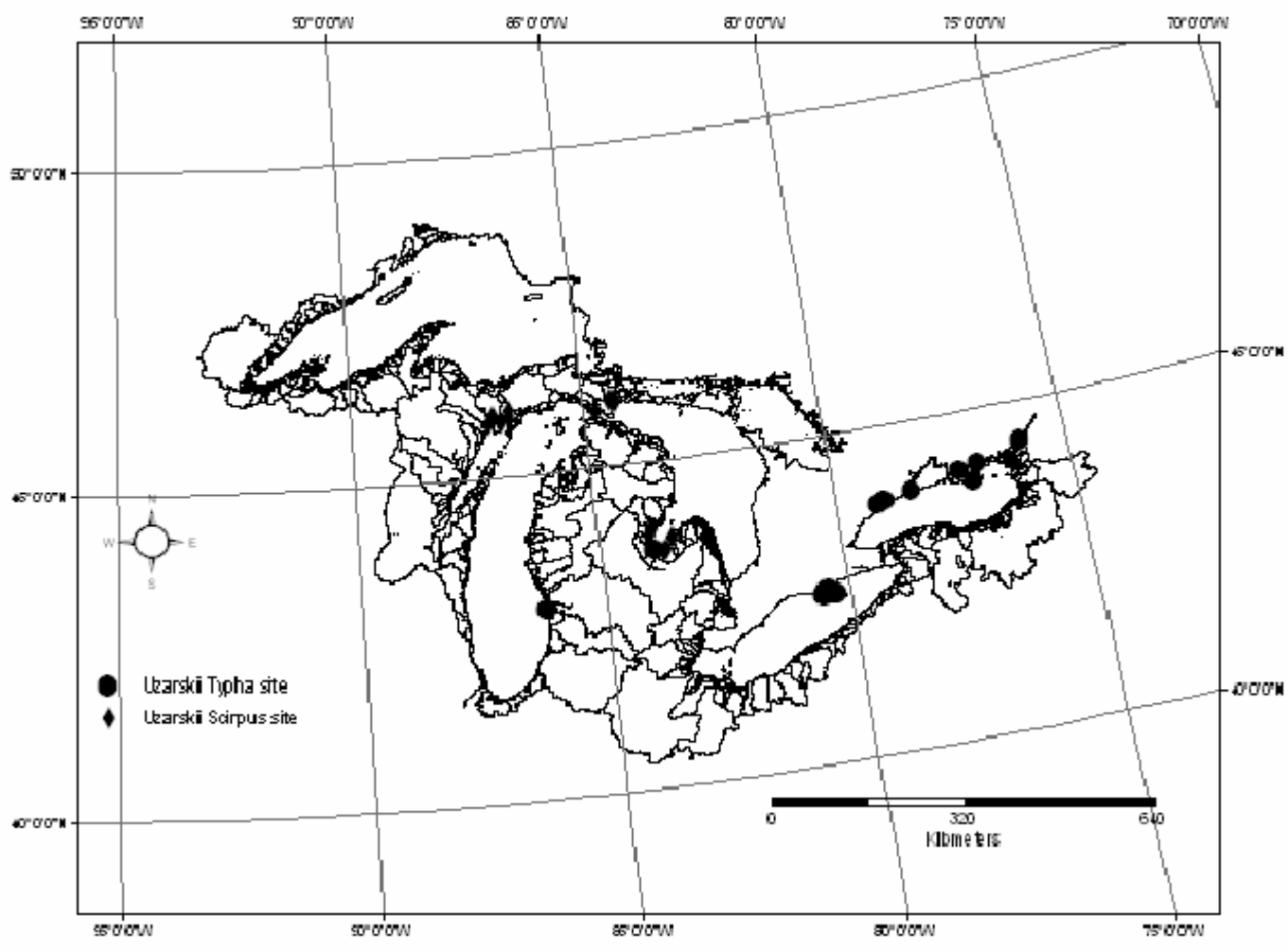

Figure 3.6. Map of the Great Lakes segment-sheds showing Uzarski et al. (2005) *Scirpus* and *Typha* dominant sites.

Since *Scirpus* IBI responded most to the agriculture stressor, we performed a Spearman correlation between the IBI scores for both sets of wetland sites against the agriculture stressor scores. There was a significant negative correlation ($r = -0.66$, $p < 0.001$), and the pattern of variation with respect to agricultural stress was consistent between Uzarski's sites and GLEI sites (Figure 3.7). However, the GLEI sites covered a broader range of the stress gradient. The combined data rendered the threshold nature of the IBI score-stress gradient relationship even more strongly.

Discussion

Minns et al. (1994) developed a littoral fish IBI for the areas of concern (AOC) of the Great Lakes, based on Karr's original IBI approach (Karr 1981). The IBI exhibited low variability between repeated measures and significant positive correlations with submergent vegetation density at the sites sampled. Thoma (1999) used a similar approach to develop an IBI for near-shore sites of Lake Erie. Yet, when applying a similar methodology to assessing fish communities, Wilcox et al. (2002) concluded that the IBI approach does not work for Great Lakes coastal wetlands, due to the inherent hydrological variability, which would in turn affect the reproducibility of the IBI in similar wetlands. Despite Wilcox et al.'s (2002) contention that interannual water level fluctuations obscure fish-habitat associations, Uzarski et al. (2005) were able to find consistent associations between fish communities and macrophyte zones.

Emergent plant community composition changes quickly as a result of water level fluctuation (Uzarski et al., 2005). Fishes seek habitats characterized by preferred ranges of water levels. Thus, an IBI that consists of metrics based on the fish assemblage attuned to a particular plant zone (i.e., *Typha* or *Scirpus*) should be applicable in the face of hydrological variability.

The *Typha*-dominant sites sampled by Uzarski et al. (2005) were located mostly in Lakes Erie and Ontario, which conforms with expectations since *Typha* sp. tend to dominate emergent marshes in lower latitudes (Lougheed et al. 2001). Using slightly different criteria to assess dominance of emergent plant species, were able to locate *Typha*-dominant sites along the entire Great Lakes coastline, which retrospectively allowed us to test the Uzarski (2005) IBI across the Great Lakes basin.

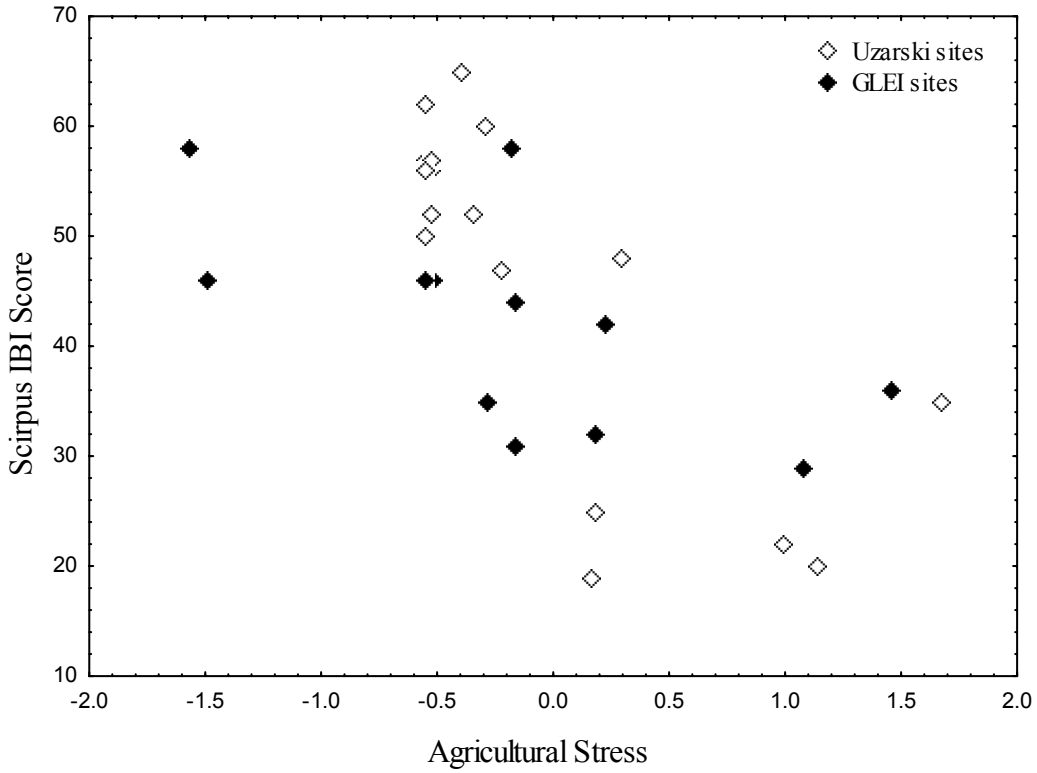

Figure 3.7. Relationship between GLEI and Uzarski *Scirpus* dominant sites to agricultural stress ($r = -0.66$, $p < 0.001$, $n = 30$).

The IBI scores derived from their data and this study, showed similar trends of declining IBI scores as a function of increasing anthropogenic stress.

However, our sampling program captured a broader range of overall human activity, especially at the reference end of the scale than did Uzarski et al. (2005), whose sites were concentrated in just three areas (Fig. 3.1). The IBI-agricultural stress relationship resulting from the combined dataset (GLEI *Scirpus* and Uzarski *Scirpus* sites) suggests that threshold relationships rather than a linear relationship best describe how the fish IBI changes as agricultural stress increases.

Several studies have shown that *Typha* sp. can tolerate hydrological fluctuations (Galatowitsch et al. 2000, Waters and Shay 1990, 1992) and remain abundant despite nutrient enrichment as a result of agriculture and urbanization (Galatowitsch et al. 1999, 2000). Janisch and Molstad (2004), when comparing the density of hydrophytic vegetation between disturbed (agricultural and non-agricultural landscape modification) and undisturbed lands, found *Typha* to be among the dominant species in over 60% of data points in disturbed lands. One would thus expect fish assemblages associated with *Typha*-dominant wetlands to be tolerant of agricultural stress as shown by the lack of any correlation between *Typha* IBI and agriculture. The *Typha* IBI did however decline in relation to increasing population-related stress. Fish communities in *Typha* wetlands appear to be most sensitive to shoreline changes resulting from population density, possibly because of a resulting decline in other species of emergent vegetation. The spatial distribution of IBI scores across the sampling area indicated that sites with high biotic integrity were primarily in Lake Ontario and Lake Erie, which typically had high agricultural stress scores, but relatively lower population density related stress.

Atmospheric deposition had a positive significant correlation with *Typha* IBI. This is contrary to what was expected and likely attributable to the lack of concentrated input from the individual parameters making up atmospheric deposition, such as calcium, chloride, magnesium and sulphate. Moreover, atmospheric deposition stress included measures of nitrogen input from the atmosphere directly into streams. Since *Typha* sp. remain abundant in wetlands disturbed by agricultural stress (Galatowitsch et al. 1999, 2000), the impact of indirect nutrient inputs from the atmosphere would not adversely affect *Typha* sp. cover or the fish community composition in such wetlands.

While Uzarski et al.'s (2005) *Typha* sites were primarily restricted to the lower Great Lakes; their *Scirpus*-dominant sites were located largely in the upper Lake Michigan and Lake Huron region. Wetlands dominated by *Scirpus* vegetation are usually found in higher latitude sites and in forested watersheds (Lougheed et al. 2001). Although Minc (1997) reported that *Scirpus*, along with *Eleocharis* and *Isoetes* are among the emergent species that are largely absent from the lower lakes, we were able to include 6 of 13 *Scirpus*-dominant sites from Lakes Erie and Ontario. This in turn, allowed testing of the applicability of the *Scirpus* IBI to lower Great Lakes wetlands as well. In contrast to the *Typha* IBI, which did not vary across the agricultural stress gradient, the *Scirpus* IBI scores were significantly negatively correlated with agricultural intensity and point source discharge. Day et al. (1988) and Minc (1997) have observed that species such as *Scirpus americanus* and *Eleocharis smalii* are better adapted to shorter growing seasons and lower substrate fertility, compared to *Typha latifolia* and *Sparganium eurycarpum* that tend to dominate highly fertile areas. The agricultural stressor score measured by Danz et al. (2005) was largely comprised of nitrogen (N) and phosphorus (P) inputs from fertilizers and pesticides. The significant negative correlation between *Scirpus* IBI and agricultural stress suggests that fish species in *Scirpus*-dominant wetlands are most sensitive to N and P loadings. The point source pollution stressor, was a combined measure of 79 variables, including the number of facilities discharging polyaromatic hydrocarbons (PAHs) into streams as well as a measure of N and P inputs from non-agricultural sources (Danz et al. 2005), again indicating that fish communities in *Scirpus* wetlands are sensitive to direct pollutant discharges.

Sites with high biotic integrity tend to have many species, high relative abundance of insectivores cyprinids and piscivores; and relatively few omnivores (Simon and Lyons 1995), all of which were included as metrics in the Uzarski *Scirpus* and *Typha* IBI. The GLEI *Typha*- and *Scirpus*-dominant sites, although slightly distinctive in their fish assemblages, were noted to have a relatively higher abundance of piscivores and insectivorous cyprinids compared to omnivores or non-indigenous species, consistent with most IBI studies (Karr 1981, Simon 1991).

Summary and Conclusions

Uzarski et al. (2005) developed an IBI based on plant zone inundations for Great Lakes coastal wetlands, and calibrated their fish IBIs against a generalized anthropogenic disturbance characterized by urbanization and agriculture. The results in this analysis verify that the Uzarski et al. (2005) IBI scores do reflect anthropogenic stress in coastal wetlands having dominant *Typha* and *Scirpus* cover. The *Typha* and *Scirpus* IBIs however, only vary with respect to certain types of anthropogenic stressors and would thus be most suitable in assessing wetland condition in response to specific stressors such as agriculture or population density rather than generalized stress. In particular, the *Typha* IBI tends to decline linearly with increasing population density, whereas the *Scirpus* IBI appears to be high below certain levels of agriculture and point source pollution stress but declines at a threshold scale of each of these classes of stress. This study emphasizes the need for carefully measured stressor data as developed by Danz et al. (2005) to use an IBI in effectively assessing biotic integrity in a habitat.

CHAPTER 4

Using multivariate techniques to develop fish indicators of stress for Great Lakes coastal margins

Abstract

Fish community composition has been widely used as an indicator of stress for wetland management and restoration. Typically, the Index of Biotic Integrity (IBI) approach has been used, combining metrics that are based on species composition, abundance and trophic guild. However, the need for a priori classification of sites as well as the inherent hydrological variability among regions affects the applicability of such metrics across systems. In contrast, multivariate methods require few prior assumptions in creating groups of reference sites against which to compare test sites. To compare the utility of multivariate approaches 143 sites spanning the entire U.S. Great Lakes shoreline were sampled using overnight-set fyke nets, during the summers of 2002 and 2003. A suite of environmental variables was concurrently measured and used in a discriminant function analysis model to separate distinct groups of reference sites clustered on the basis of fish species relative abundance. The model showed that ecoregion along with 7 other environmental variables is an important factor along which fish communities are segregated. Bray-Curtis ordination techniques were used to assess changes in fish community composition at the 143 sites sampled in relation to two main pressure variables: agriculture and population density. Stronger effects of population density-related stress than of agriculture-related stress were observed. The ordination techniques also allowed for unique groupings of reference and non-reference indicator species along both pressure axes. Multivariate derivations were effective and produced results consistent with expectations, in delineating fish indicators of anthropogenic stress at Great Lakes coastal margins.

Introduction

Within the last three decades, there has been a shift in the approaches to assessing aquatic habitats from studying changes in the abundance of indicator or keystone species (Reynoldson et al. 1989) such as walleye (Schneider and Leach 1977) and salmonids (Sonstegard and Leatherland 1984, Ryder and Edwards 1986), to the widespread use of indices of biotic integrity (IBIs) comprised of several measurable, characteristic attributes of a habitat (Karr 1981, Fausch et al. 1984, Lyons 1992, Wilcox et al. 2002). However, fish IBIs are usually based on a priori assumptions made about expected fish community composition in a particular ecological zone or region. Researchers develop IBIs from the perspective that co-occurring fish species represent unique communities that are closely adapted to local habitat characteristics. Some researchers argue that community-specific IBIs should be developed at the ecoregion level (Hughes et al. 1986, Omernik and Griffith 1991) since ecoregions are areas of similar biogeographic, physical and chemical attributes to which fish communities uniquely respond (Larsen et al. 1986). Yet other researchers believe that fish communities differ primarily with respect to the hydrogeomorphic (HGM) characteristics of the habitat, and that a separate IBI is thus warranted for each unique HGM type (Keough et al. 1999). The issue that arises then, particularly from a diagnostic point of view, is to determine where the line should be drawn, to properly assess an aquatic habitat. In the Great Lakes, do we require an IBI for each of the 6 ecoregions, or do we need 30 unique IBIs, representing 5 HGM types for each of the 6 ecoregions? The need to develop an integrated set of indicators for Great Lakes has been the forefront of many discussions sponsored by management organizations such as the State of the Lakes Environmental Conference (SOLEC), the Great Lakes National Program Office of the US Environmental Protection Agency (US EPA GLNPO) and the Great Lakes Fish Commission (GLFC). Multivariate approaches are being increasingly used to empirically determine the associations among fish species and the distribution of fish communities among habitats before determining the type and scale of response of fish communities to various anthropogenic stressors (Wang et al. 2004, Wei et al. 2004).

Many studies have been conducted to assess fish species assemblage structure along environmental gradients and with respect to shoreline features, particularly in

streams (Waite and Carpenter 2000, Maret and MacCoy 2002, Wang et al. 2003) and more recently in coastal wetlands and littoral zones of the Great Lakes (Brazner and Beals 1997, Wei et al. 2004). Maret and MacCoy (2002) using multivariate analyses determined that dater species (Cottidae) are more severely affected by elevated cadmium, lead and zinc concentrations in streams than are salmonids, and that fewer native species were occur in areas with high metal enrichment. Waite and Carpenter (2000) using multivariate classification and ordination, determined that water temperature, dissolved oxygen and physical habitat features (e.g. riparian canopy) were associated with patterns of fish assemblages in streams across all ecoregions. Wei et al. (2004) conducted a similar study on fish distribution patterns in the Great Lakes and found significant correlations between certain taxa (e.g. common carp, emerald shiner, largemouth bass) and wetland type (e.g. protected wetland versus open-shoreline wetland).

Such studies are important from a management perspective as they allow for monitoring areas of concern or sites where certain species that are expected in those habitats are declining in number. However, fish species are also affected by the consequences of anthropogenic stress (Weaver and Garman 1994, Wichert 1995). Although some researchers have used stressor information to elucidate patterns in fish assemblages, they have relied heavily on estimates of land-use related stress obtained mostly from outdated maps and government databases (Brazner and Beals 1997). A geographic information system approach permits one to automate quantification of variables such as road density and percentages of land use. Such data may assist one to more objectively assess fish assemblages along anthropogenic stressor gradients (Detenback et al. 1993).

The relative merits of multivariate vs. multimetric approaches to classification of the reference condition in aquatic systems have been argued vigorously (Reynoldson et al. 1997, Norris 1995) (See Chapter 1). Multivariate approaches tend to classify communities by the presence/absence of species (e.g. AUStralian RIVer Assessment Scheme: AusRiVAS, Wright 1995) or by absolute abundance (e.g. Benthic Assessment of SedimenT: BEAST, Reynoldson et al. 1995). Dufrene and Legendre (1997) developed 'indicator species analysis' to find the species whose presence or absence best distinguished reference and degraded sites. The bimodal nature of presence/absence data

however, does not allow for clear interpretation of the effects of different types of stressors on the community structure as a whole. Similarly, absolute abundance data is of limited value in assessing fish community response to disturbance, as the overall species richness of a habitat is not taken into account. As such, certain biases about expected abundances of species can be introduced in the interpretation of results. In this paper, we combined multivariate techniques used for benthic invertebrate assessment (e.g. BEAST model) and applied it to fish communities to assess the condition of a site relative to reference sites and their assemblages. We used the relative abundance of taxa of individual species to identify characteristic communities and indicator taxa that reflect natural environmental variation among reference condition habitats. However, we also used an additional ordination step to reveal changes in fish communities that best correlate with particular types of human-related disturbance at Great Lakes coastal margins.

The goals of this paper were 1) to use fish community composition observed at reference locations rather than a priori assumptions to guide the grouping of sites for environmental quality assessments, 2) to determine the environmental variables along which fish communities of reference areas are segregated, 3) to establish fish community composition criteria for assessing the quality of sites within a group and 4) to identify fish taxa that could serve as indicator species at the reference end and most degraded end of anthropogenic stress gradients.

Methods

Site Selection and Classification

A total of 46 reference sites, sampled through the Reference Area project (Schuldt et al. 2000) (see Chapter 1) were identified as least affected by anthropogenic disturbance using a GIS-based analysis of landuse and remote sensing that measured 5 watershed scale based stressor variables: agricultural land use, urban land use, distance from a point source, road density and population density (see Chapter 1).

A second set of 97 sites, selected based on a design developed by the Great Lakes Environmental Indicators (GLEI) (Niemi et al. 2001) project was sampled in concert with reference sites using identical field protocols (see Chapter 1). The sites identified the

GLEI project spanned the entire gradient of 6 categories of human disturbance and one category of environmental variation (i.e. soils). The 6 categories of human disturbance were agriculture, population density, land cover, atmospheric deposition, point source discharge and shoreline modification (Danz et al. 2005) (see Chapter 1)

Fish Sampling

A total of 143 sites (approximately 35 from each hydrogeomorphic type) distributed across all 5 Great Lakes were sampled during the summers of 2002 and 2003. Fish communities were sampled using 2 large fyke net arrays (1.25 cm mesh) and two small fyke net arrays (0.5 cm mesh) set overnight at each site following the methods of Brazner and Beals (1997). Each fyke net array was placed lead-to-lead (leads parallel to shore), with the wings set at 45° angles. One pair of large and one pair of small nets were placed near each of the two dominant shoreline habitats at a site. All fishes were identified to species, measured (total length to the nearest mm), examined for condition (incidence of damage or disease) and released. Specimens of uncertain identity were euthanized in clove oil, preserved in formalin-ethanol mixture, and returned to the laboratory for detailed examination.

Environmental variables

A suite of environmental variables was measured at each site sampled, to help identify environmental similarities and differences among sites in concert with land use variables derived from GIS analysis. Physicochemical variables (water temperature, dissolved oxygen concentration, conductivity, pH and oxidation-reduction potential (ORP)) were measured at each fyke net at a sampled site using a YSI (556 MPS) multi-probe meter (calibrated daily for dissolved oxygen and pH, and every 10 days for ORP and conductivity). Water clarity measurements were also taken at each net using a Secchi disk and a turbidity tube. To perform sediment particle size analysis and organic content analysis, we took and combined sediment samples from 2-4 locations at each of the 'emergent' (20-50 cm deep) and 'submergent' (typically 50-75 cm deep) macrophyte zones and at 5 m and 10 m depths contours where they occurred. Only three zones were sampled along each transect (emergent, submergent and deepest point) if the open water

depth was less than 5 m, as was the case for embayments, river-influenced wetlands and protected wetlands. Sediments from emergent and submergent zones were sampled using 10-cm diameter coring tubes, and open water or 5 m and 10 m zones were sampled using ponar grabs. Habitat-associated variables such as the extent and density of dominant and subdominant emergent, submergent and floating plant genera were also noted at each net, in addition to shoreline features and adjacent land use.

Statistical Analyses

Forty-two reference sites and 91 GLEI test sites (Chapter 1) were used for developing the initial multivariate models. Ten sites (4 reference and 6 GLEI) were reserved to be used later for testing the multivariate models. The delineated sampling polygons for these 10 sites overlapped between the GLEI project and the Reference Condition Project (i.e., the GLEI sites fell within the lowest 20th percentile of RelMax scores; see Chapter 1 for details) and thus could not be used for developing either the multimetric or multivariate models. However, using these sites as 'reserve' for testing the model would allow for comparisons of classifications based on the model and expected results based on the projects' sampling design.

Fish species counts were standardized by effort (i.e. time fished) and by net size before measuring relative abundances of each species at each site. For the purpose of statistical analysis, relative abundances of fish species were expressed on an octave scale ($\log_2$ [100 x (proportion+0.01)], Gauch 1977). Species rarely captured (i.e., present at <15 % of sites) were eliminated from subsequent analysis.

All environmental variables were $\log_{10}$ (Y+1) transformed, to improve homogeneity of variances and normality of the data, except for latitude, longitude, pH, median particle size and those variables based on a categorical scale. Principal component analysis (PCA) was conducted on a correlation matrix of the 48 environmental variables (Table 4.1) to identify redundancy in the data. Using varimax factor rotation, 14 factors were extracted, explaining 77 % of total variance (Table 4.1). Variables with the highest factor loadings within each factor were chosen for further analyses. In the case where two variables were grouped into one factor, but were not intuitively related to each other, both variables were selected for further analyses.

Table 4.1. Correlation (factor loading) between values of 48 environmental variables measured at 133 Great Lakes coastal margin sites and 14 principal components. Variables subsequently used for discriminant function analysis of clusters of sites associated by fish species relative abundance are shown in bold face.

	Factors													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Category 1-Habitat														
Depth of fines	**0.57**	-0.07	0.07	0.11	0.07	-0.01	0.04	-0.04	-0.06	-0.24	-0.07	0.05	-0.06	-0.21
LOI	**0.64**	-0.16	-0.21	-0.01	-0.03	0.09	0.08	-0.03	-0.01	0.24	-0.25	0.15	0.34	0.05
Growth form richness	**0.77**	-0.12	-0.06	0.02	0.09	0.05	0.12	0.02	0.08	0.16	-0.22	0.00	0.37	-0.12
Mean EM density	**0.89**	0.02	-0.05	-0.05	0.01	0.02	-0.04	0.02	0.14	0.02	-0.03	0.03	0.03	-0.01
Mean EM cover	**0.85**	-0.01	-0.04	-0.02	-0.03	-0.04	-0.08	-0.08	0.08	0.00	0.00	0.02	-0.05	-0.02
Mean FLT density	**0.81**	-0.27	0.07	0.06	0.03	0.02	0.06	0.08	0.26	0.01	-0.02	-0.06	-0.08	-0.02
Mean FLT cover	**0.48**	-0.16	0.19	-0.12	0.05	-0.15	-0.10	0.04	0.48	0.14	-0.15	-0.03	0.01	0.08
Mean OW cover	**-0.56**	0.14	0.24	0.18	-0.05	-0.12	-0.17	-0.11	-0.20	0.00	0.04	-0.13	-0.08	0.08
Mean SUB density	**0.77**	-0.25	0.12	0.09	0.12	0.02	0.24	0.05	0.14	0.02	-0.11	-0.09	0.13	-0.04
Mean SUB cover	**0.75**	-0.25	0.18	0.15	0.10	0.01	0.18	0.07	0.16	-0.07	-0.07	-0.09	0.11	-0.05
Flow influenced by lake	**-0.46**	-0.18	0.23	0.08	0.02	-0.46	0.28	0.10	-0.13	0.11	0.01	0.04	-0.27	0.18
Category 2-Ecol. region/zone														
Province	0.24	**-0.76**	-0.18	0.02	0.20	-0.29	-0.14	0.08	0.00	-0.13	-0.04	-0.08	-0.04	0.08
Longitude	0.34	**-0.76**	0.22	0.23	-0.17	0.05	-0.03	0.00	0.06	0.04	0.08	-0.06	0.17	-0.13
Latitude	-0.07	**0.87**	0.18	0.21	-0.05	0.14	0.11	-0.07	-0.01	0.15	0.08	-0.01	-0.14	-0.11
Superior	-0.04	**0.70**	-0.06	0.33	0.43	0.08	0.06	-0.07	0.06	0.06	-0.07	0.01	-0.29	-0.03
SSU	0.05	**0.63**	-0.07	0.08	0.57	0.03	0.08	-0.01	0.00	-0.12	-0.08	0.09	0.23	0.04
EOL	0.26	**-0.78**	-0.06	0.27	0.24	0.24	0.07	0.00	0.16	0.12	-0.06	0.07	0.06	-0.05
Erie	-0.26	**-0.62**	-0.35	0.19	0.24	0.20	0.06	0.00	0.11	0.02	-0.19	0.18	0.11	0.08
ORP	0.16	**-0.44**	0.13	0.03	0.24	0.11	0.00	0.33	0.03	0.09	0.20	0.18	-0.10	-0.42

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Category 3-Water clarity														
Secchi depth	0.05	0.04	**0.83**	-0.05	-0.04	0.00	-0.12	-0.01	-0.03	-0.07	0.02	0.08	-0.10	0.01
Turbidity Tube depth	0.00	0.19	**0.79**	0.05	-0.10	0.08	-0.15	0.06	0.06	-0.01	0.06	0.05	-0.07	0.02
Category 4-Lake/sampling date														
Lake Michigan	-0.10	0.09	0.08	**-0.92**	-0.05	-0.07	0.01	0.04	-0.04	-0.11	0.03	-0.04	0.06	0.01
Julian day	0.02	0.23	0.33	**0.57**	0.11	-0.10	0.16	0.13	-0.03	0.01	-0.10	-0.20	-0.01	0.09
Category 5-Lake/region														
Lake Huron	-0.13	0.06	0.04	0.40	**-0.67**	-0.28	-0.15	0.03	-0.20	-0.07	0.11	-0.04	0.17	0.08
NGL	-0.21	0.24	0.17	-0.25	**-0.73**	0.31	0.07	-0.14	0.02	-0.03	0.06	-0.02	0.12	-0.06
Conductivity	0.31	-0.06	-0.24	-0.38	**-0.52**	-0.15	-0.10	0.17	0.11	-0.03	-0.09	0.03	0.08	0.06
Category 6-Habitat/region														
Median particle size	0.22	-0.30	0.00	-0.23	-0.01	**0.43**	0.04	0.10	-0.02	-0.08	-0.13	0.00	-0.24	0.12
SCG	0.00	-0.11	-0.08	-0.12	-0.04	**-0.80**	-0.23	0.13	-0.16	-0.15	0.05	-0.13	0.03	0.00
Category 7-Adjacent landuse														
Natural/forest /marsh	0.09	0.17	-0.17	0.01	0.07	0.07	**0.79**	0.08	0.04	0.04	-0.04	-0.31	0.10	-0.01
Dyke/low dens resid /beach	-0.21	-0.06	0.11	0.00	-0.01	-0.13	**-0.83**	-0.02	0.09	0.00	0.07	-0.21	-0.10	0.03
Category 8-HGM/flowregime														
Embayment	-0.16	-0.02	0.09	0.06	0.01	0.10	-0.05	**-0.67**	0.00	0.14	0.23	0.37	0.23	0.03
Coastal wetland	0.06	0.02	-0.15	0.23	-0.26	-0.07	0.13	**0.50**	-0.04	-0.01	0.38	-0.02	0.41	0.27
No flow	0.23	-0.10	-0.23	-0.08	0.01	0.20	-0.30	**0.38**	0.37	0.00	0.01	0.24	0.14	-0.16
Flow influenced by backwater	0.13	0.27	-0.26	0.04	-0.12	0.11	-0.07	**-0.70**	0.04	-0.04	0.07	-0.11	0.03	0.03

	1	2	3	4	5	6	7	**8**	9	10	11	12	13	14
Category 9-Dissolved oxygen														
Dissolved O2%	-0.42	0.06	-0.01	-0.01	-0.03	-0.10	0.04	**0.01**	-0.79	0.09	0.04	0.08	0.03	-0.05
Dissolved O2 mg/l	-0.40	0.07	0.03	-0.06	-0.05	-0.13	0.00	**0.03**	-0.75	0.00	0.06	0.09	-0.09	-0.01
Category 10-Shoreline structure														
Rocky shoreline/cliff/riprap	-0.40	0.03	0.19	0.15	0.11	0.01	-0.09	-0.14	-0.08	**0.62**	0.22	0.02	-0.25	0.10
Sandbeach/dyke	-0.32	-0.08	0.16	-0.09	0.01	-0.11	-0.08	0.00	0.02	**-0.80**	0.14	0.11	-0.10	0.03
Category 11-Protected area														
Protected wetland	0.19	-0.07	-0.02	0.08	0.13	0.00	0.06	0.04	0.03	0.02	**-0.85**	0.05	0.06	-0.01
Hydrological Connection	0.48	0.01	-0.15	-0.02	-0.10	0.15	0.13	0.18	0.17	0.04	**-0.61**	0.03	0.25	-0.11
Category 12-Commercial landuse														
Highway/road /comm./industry	0.01	0.04	0.08	-0.01	0.06	0.08	-0.02	-0.04	-0.11	-0.08	-0.07	**0.86**	-0.07	0.04
Category 13-Physical variables														
Temperature	0.24	-0.40	-0.14	-0.06	0.01	-0.07	0.14	0.00	0.11	0.03	-0.20	-0.24	**0.65**	-0.06
NSU	-0.11	0.27	0.00	0.21	0.06	0.02	0.04	0.09	0.02	0.34	-0.01	0.06	**-0.65**	0.03
pH	-0.45	-0.29	0.04	-0.22	-0.05	-0.13	0.04	0.06	-0.22	0.13	0.01	0.02	**0.51**	0.15
High energy	-0.43	0.03	0.16	-0.12	0.17	-0.11	-0.20	0.13	-0.09	-0.15	0.17	-0.10	**-0.65**	0.02
Maximum depth	-0.33	0.06	0.36	-0.04	0.15	-0.40	-0.06	0.01	-0.02	0.05	0.04	0.01	**-0.46**	0.11
Category 14-Shore. struc./flow regime														
Vegetated bank/bog/forest	0.27	0.14	-0.07	-0.04	-0.04	0.08	-0.02	0.02	-0.18	-0.04	-0.12	-0.19	0.08	**-0.71**
Channel	0.05	0.02	0.05	-0.07	-0.01	0.36	-0.36	0.12	-0.33	0.00	0.02	-0.26	0.04	**0.51**
Explained Variance	7.61	5.05	2.51	2.26	2.27	2.03	2.11	1.72	2.16	1.54	1.81	1.52	2.88	1.30
Proportion of total variance	0.16	0.11	0.05	0.05	0.05	0.04	0.04	0.04	0.05	0.03	0.04	0.03	0.06	0.03
Cum. Proportion	0.16	0.27	0.32	0.36	0.41	0.45	0.50	0.53	0.58	0.61	0.65	0.68	0.74	0.77

Consequently, 23 independent environmental variables were retained and used to classify fish assemblage groups (bold faced variables in Table 4.1).

To identify groups of reference sites (hereafter referred to as 'REF' sites) with similar species composition, we used Ward's method of cluster analysis with the Sorensen (city-block) distance measure. Although, the cluster analysis yields similar results when using Euclidean distance measure, the use of city-block distance measure dampens the effect of single large differences or outliers and is often a preferred distance measure when working with ecological data (McCune and Grace 2002). The species contributing most to the distinctiveness of each cluster were identified by examining the F-ratio of among group: within group variances (Green and Voscotto 1978) as identified by the K-means option of the Statistica® cluster analysis subroutine.

Once groups of compositionally similar REF sites were determined through cluster analysis, a forward step-wise discriminant function analysis (DFA) was performed to identify the set of environmental variables that would best separate individual clusters of reference sites. The DFA model was then used to determine which reference cluster a particular test (GLEI) site should belong to, based on the environmental variables that were identified as important in separating the REF clusters.

If anthropogenic stress has no effect on fish community composition, then all sites classified as having similar environmental conditions should have homogeneous fish community composition. In contrast, if anthropogenic stress affects fish community composition, the most stressed sites within a cluster containing REF and GLEI sites should have different relative abundances of certain (indicator) taxa than the least stressed (REF) sites of that cluster. The final step in determining the position of fish species-ordinated sites (both reference and 'GLEI sites') along stressor axes was to use an ordination technique that would allow us to assess the environmental condition of test sites along those axes. Site-specific anthropogenic stress scores for each of the 6 stressor category measured as part of the GLEI site selection process were used to identify axis endpoints for the ordinations.

To identify fish species assemblages that would best distinguish the species within a cluster characteristic of the minimally disturbed from the most stressed sites, we used Bray-Curtis (BC) ordination with subjective endpoint selection (McCune and Grace

2002). This allowed us to select an end-point at the reference end (the reference area within a cluster possessing the lowest urban or agricultural stress score) and an end-point at the degraded end of a stressor axis (the GLEI site within a cluster possessing the highest stress score). The ordination then calculated the relative position of each site member of the cluster between the two end-points. The Sorensen distance measure (= city-block distance) was used as the distance metric (McCune and Grace 2002). This ordination produced a 'fish quality' scale bounded by the two endpoints. The theoretical limits of the scale were 0 (the reference site selected as an end point) and 1.0 (the position of the site selected to represent the most stressed conditions).

To evaluate the strength of the relationship between each fish quality index (Bray-Curtis axis score) and anthropogenic stress (PCA-derived urban pressure or agricultural pressure values), Spearman rank correlations were determined between ordination scores and pressure scores for the sites within each cluster.

Lastly, we wished to assess variability in fish community composition among REF sites as identified by the Bray-Curtis (BC) ordinations. Bray-Curtis axis 1 scores were plotted against BC axis 2 scores. The bounds of the family of equivalent-to-reference sites was delineated by graphically identifying and excluding all sites with fish quality index scores equivalent to or greater than the scores of the poorest-ranking REF sites.

To validate the multivariate models, we classified 10 'reserve' sites that were part of the original sampling program, but were not used in developing the models. We used the DFA model to determine the cluster to which each of these reserve sites belonged. The Bray-Curtis ordination factor loading coefficients of the cluster that matched the reserve site were then used to calculate the fish quality index score of the reserve site. For example, if a reserve site belonged to cluster 2, we used the site's fish community composition to estimate the fish quality score. Once scores were determined for each reserve site, similar threshold criteria (used to assess GLEI sites) were applied to assess the quality of the reserve site as reference or non-reference.

We also used the expected species composition at the reference and degraded ends of the pressure axes to compare against the species composition at the reserve site for the

corresponding cluster. The reproducibility of the model would thus be confirmed if we obtained similar species composition matches at the reserve sites.

Bray-Curtis ordinations were performed using PC-ORD®, version 4 (MjM Software Design 1999). All other statistical analyses were performed using Statistica® software package, version 6.0 (StatSoft Inc. 2001).

Results

A total of 108 species were recorded from the 133 sites sampled across the entire U.S. coastline. The most abundant and the most frequently recorded species were northern rock bass (*Ambloplites rupestris*), pumpkinseed sunfish (*Lepomis gibbosus*) and yellow perch (*Perca flavescens*). Among the species that were eliminated from the analysis due to their rare occurrences were American eel (*Anguilla rostrata*), silver redhorse (*Moxostoma anisurum*), Great Lakes muskellunge (*Esox masquinongy*), and threespine stickleback (*Gasterosteus aculeatus*), all of which have also been reported as rare or absent from all 5 Great Lakes by Jude and Pappas (1992).

The 48 environmental variables measured were categorized into 14 factors using a principal components analysis (PCA). Based on factor loadings, we identified 23 independent variables to be used for the discriminant function analysis (Table 4.1). A PCA of the 6 stressor variables (performed for all 762 segment-sheds) resulted in 2 categories of stress: the first component, which we term, 'population pressure' was strongly correlated with land cover, population density, point source discharge and extent of shoreline modification (Table 4.2). The second component termed, 'agricultural pressure' combined agriculture-related stress and atmospheric deposition (Table 4.2)

Species Assemblages and Environmental Influences

Based on the cluster analysis of fish species relative abundances, (30 species) at REF sites we identified 5 groups of sites (Figure 4.1). Cluster 1 was the largest group, consisting of 12 sites. Cluster 1 was characterized by a dominance of spottail shiner (*Notropis hudsonius*) and yellow perch (*Perca flavescens*). Cluster 2 consisted of 9 sites that typically had greater relative abundance of northern rockbass (*Ambloplites rupestris*) and bluntnose minnow (*Pimephales notatus*).

Table 4.2. Correlation (factor loading) between values of 6 anthropogenic stressor variables and 2 principal component factors. Factor 1 represents population density related pressure and factor 2 represents population density related pressure. Variable combined in factors 1 and 2 are shown in bold face.

Stressor variables	**Factor 1**	**Factor 2**
Agriculture/chemical	0.47	**0.36**
Atmospheric deposition	0.13	**0.86**
Land cover	**0.85**	0.03
Population density	**0.86**	0.20
Point source discharge	**0.71**	-0.10
Shoreline modification	**0.58**	-0.53
Explained variance	2.55	1.20
Proportion of total	0.42	0.20
Cum. Proportion	0.42	0.62

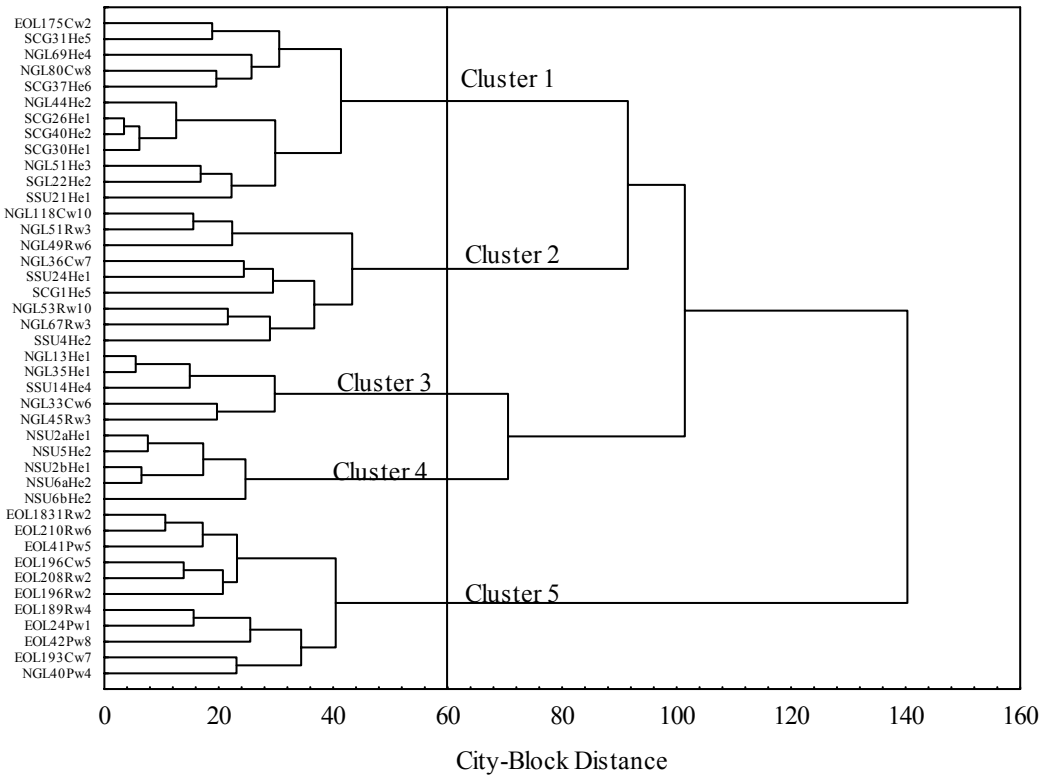

Figure 4.1. Dendogram of REF sites (n = 42) grouped according to similar species composition (Ward's method clustering city-block distances of octave-transformed relative abundances of fish species). Latitude, longitude and place names of sites corresponding to site labels are summarized in Appendix I).

Cluster 3 consisted of 5 sites that were dominated by bluegill sunfish (*Lepomis macrochirus*) and pumpkinseed sunfish (*Lepomis gibbosus*), whereas Cluster 4, also consisting of 5 sites, was dominated by longnose dace (*Rhinichthys cataractae*) and white sucker (*Catostomus commersonii*). Cluster 5 consisted of 11 sites and was characterized by an abundance of burbot (*Lota lota*) and eastern longnose sucker (*Catostomus catostomus*).

The discriminant analyses classified 41 of the 42 REF sites correctly (Table 4.3). Eight variables were identified as important in separating the clusters of REF sites (Table 4.4). The DFA model thus generated was used to classify the GLEI sites into corresponding clusters based on those 8 variables.

Stressor Influences

Two data matrices were used for ordinations on each of the 5 clusters of sites (REF and GLEI). A matrix of sites (rows) x species (columns) was used to identify which types of species were associated with particular types of sites. An artificial site, which we called, 'Centroid' was added to each species matrix, representing the hyperdimensional centre of the cluster of the reference sites. Using the endpoint selection method of Bray-Curtis ordination, the 'Centroid' site would thus represent the reference end point for each cluster. The environmental data matrix consisted of sites (rows) by 'pressure type' (columns - the two pressure axes, population and agriculture), for all sites in the cluster. The 'Centroid' site was assigned population and agriculture pressure values representing the mean of the pressure scores of the reference sites comprising the cluster. The endpoint chosen for the degraded end of the axis however was the site exhibiting the largest pressure score for each of the pressure axes.

Cluster 1 – Spottail shiner, Yellow perch

Cluster 1 consisted of 44 sites (12 reference sites, 31 GLEI sites, 1 'centroid' site) and 24 species. The two BC axes extracted 37.69% of the original distance matrix: 22.92% for axis 1 (population pressure) and 14.77% for axis 2 (agricultural pressure). There was a positive, yet non-significant, correlation between site ordination scores and the two pressure axes, population pressure ($r = 0.26$, $p > 0.05$; one-tailed) (Figure 4.2) and agricultural pressure ($r = 0.26$, $p > 0.05$; one-tailed) (Figure 4.3).

Table 4.3. Summary of observed number of sites in each cluster (columns) identified by fish relative abundances and membership predicted (rows) by discriminant function classification (Appendix VII) on the basis of environmental variables measured at those sites.

		Observed				
% Correct		Group 1	Group 2	Group 3	Group 4	Group 5
100	Group 1	12	0	0	0	0
100	Group 2	0	9	0	0	0
80	Group 3	0	1	4	0	0
100	Group 4	0	0	0	5	0
100	Group 5	0	0	0	0	11
Total 97.6		12	10	4	5	11

Table 4.4. Variables determined by DFA model as significant in classifying REF site cluster membership.

Variables	**Significance level**
EOL	$p < 0.05$
SSU	$p < 0.001$
NGL	$p < 0.001$
SCG	$p < 0.05$
Latitude	$p < 0.001$
Julian Day	$p < 0.01$
Protected wetland	$p < 0.01$
Mean Emergent cover	$p < 0.01$

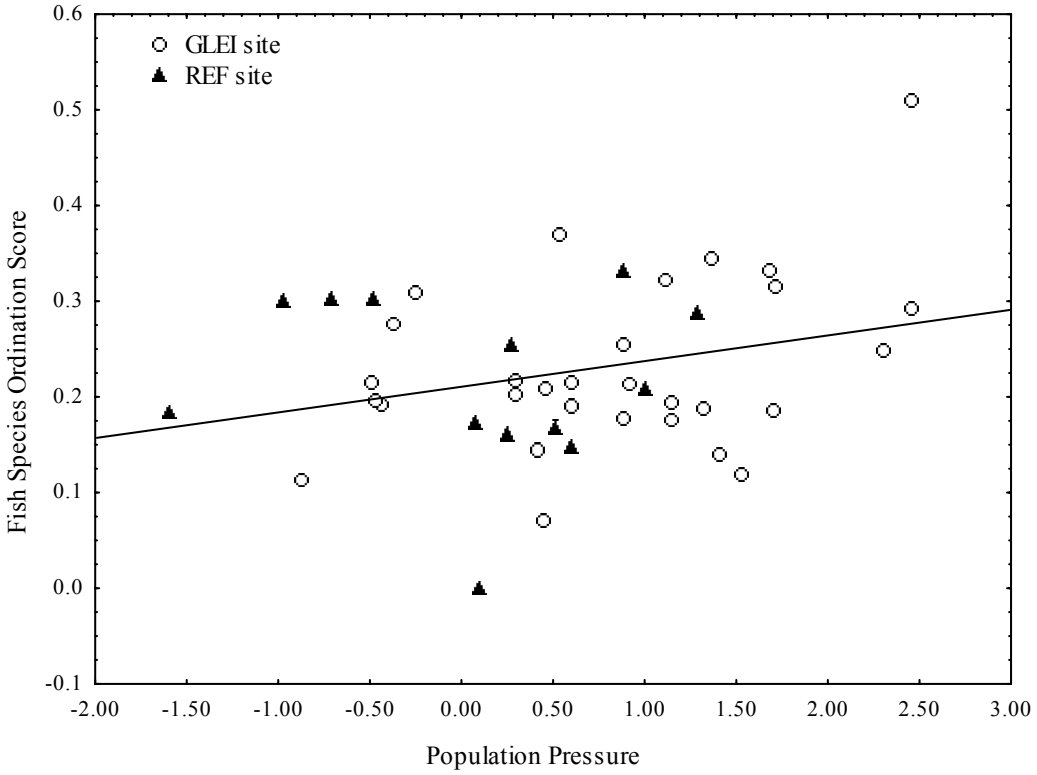

Figure 4.2. Relationship between Bray-Curtis fish relative abundance ordination scores and population pressure for sites in cluster 1. Low fish ordination scores are most similar to the reference score of zero. High fish ordination scores are most similar to the score most closely associated with large values of population pressure score.

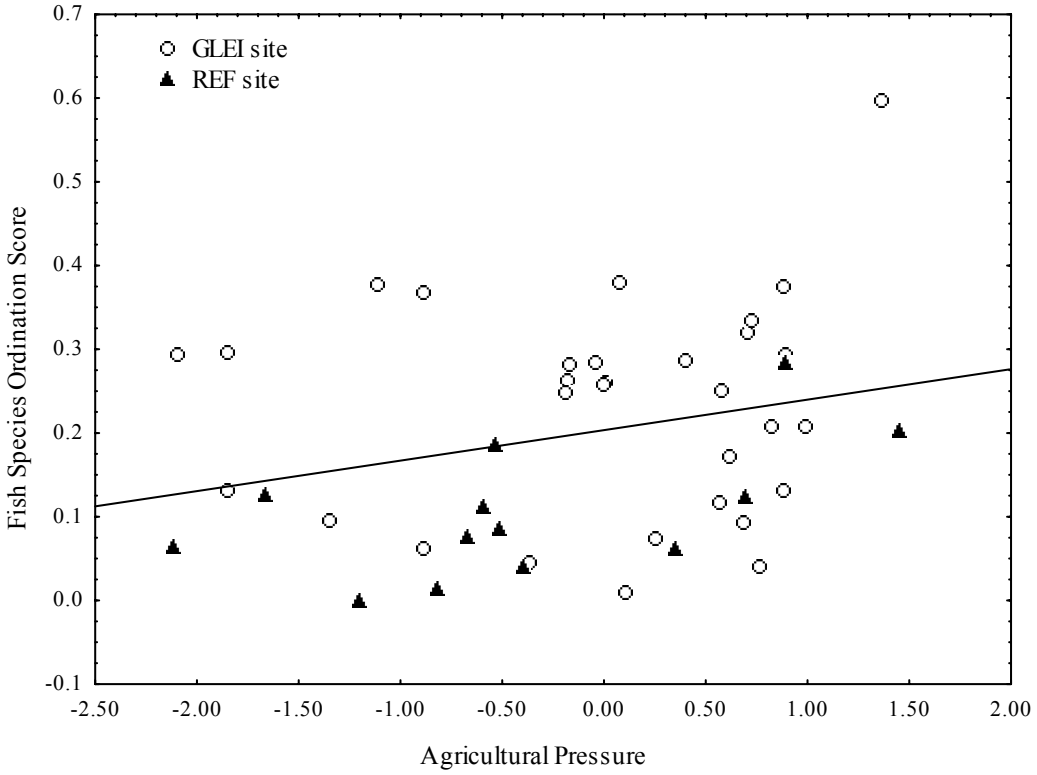

Figure 4.3. Relationship between Bray-Curtis fish relative abundance ordination scores and agricultural pressure for sites in cluster 1. Low relative abundance scores are most similar to the reference condition centroid. High scores are most similar to the suite of fish relative abundances associated with the site having the highest agricultural pressure score.

Plots of fish relative abundance at each site vs. the BC ordination axis score for that site were inspected for each species to determine the species that were most abundant at the reference end and degraded end of the pressure axes, thus representing indicator species (Table 4.5). The ordination scores for both axes were plotted against each other to determine the threshold limits (i.e., reference sites exhibiting the highest (poorest) BC-axis score) for assessing a site as similar to reference condition or non-reference condition.

The threshold level for BC axis 1 (fish community relating to population pressure) was 0.34 and for BC axis 2 (fish community condition relating to agricultural pressure) was 0.3 (Figure. 4.4). The ordination distances along with the threshold criteria were used to determine if the GLEI sites that had fish community composition within the range of variability observed among reference sites or whether they were outside that range (Table 4.6).

Cluster 2 – Rockbass, Bluntnose minnow

Cluster 2 was represented by 42 sites (9 reference sites, 32 GLEI sites and 1 'centroid' site) and 31 species. The two BC axes extracted 42.79% of the original distance matrix: 19.82% for axis 1 and 22.96% for axis 2. There was a significant positive correlation between site ordination scores and population pressure ($r = 0.59$, $p < 0.001$). However, a non-significant positive correlation between ordination scores and agricultural pressure ($r = 0.37$, $p > 0.05$) was observed. Indicator species were established for reference and degraded ends of both pressure axes (Table 4.5). The threshold value established for classifying reference and non-reference sites was 0.2 along BC axis 1 and 0.3 along BC axis 2 (Figure 4.5) (Table 4.6).

Cluster 3 – Bluegill sunfish, Pumpkinseed sunfish

Cluster 3 was relatively smaller consisting of only 9 sites (5 reference sites, 3 GLEI sites and 1 'centroid' site) and 13 species. The two BC axes extracted 86.3% of the original distance matrix: 80.1% for axis 1 and 6.2% for axis 2. The ordination scores for both axes were positively, but insignificantly correlated with population pressure ($r = 0.54$, $p > 0.05$) and agricultural pressure ($r = 0.49$, $p > 0.05$). Table 4.5 contains the list of indicator species for both reference and non-reference ends of the two pressure axes. The reference threshold line for BC axis 1 was 0.22 and for BC axis 2 was 0.27 (Figure 4.6).

Table 4.5. Reference and non-reference indicator species assemblages for each cluster of sites based on Spearman rank correlation coefficient of species relative abundance (octave) with population pressure (Pop.) and agricultural pressure (Ag.).

Cluster #	Reference species	Non-reference species	Pop. r	Ag. r
1	Northern rockbass		-0.31	
	Banded killifish		-0.37	
	Yellow perch		-0.40	
	Spottail shiner			-0.82
	Longnose dace			-0.33
		Alewife	0.52	
		Spottail shiner	0.54	
		European carp		0.42
		Pumpkinseed sunfish		0.42
		Green sunfish		0.65
		White Perch		0.34
2	Blacknose shiner		-0.27	
	Smallmouth bass		-0.30	
	Sand shiner		-0.44	-0.27
	Bluntnose minnow		-0.48	-0.28
	White sucker			-0.68
	Johnny darter			-0.29
	Eurasian ruffe			-0.33
	Spottail shiner			-0.29
	Troutperch			-0.35
		European carp	0.32	
		Yellow perch	0.68	
		Bowfin		0.44
		Northern pike		0.42
		Banded killifish		0.47
		Northern longnose gar		0.39
		Bluegill sunfish		0.76
		Pumpkinseed sunfish		0.51
3	White sucker		-0.85	-0.59
	Burbot		-0.38	-0.44
	Troutperch		-0.66	
	Longnose dace		-0.83	
		Northern rockbass	0.91	
		Largemouth bass	0.91	
		Spottail shiner	0.76	0.74
		Alewife		0.76
		Smallmouth bass		0.93

Table 4.5 Continued

Cluster #	Reference species	Non-reference species	Pop. r	Ag. r
4	Slimy sculpin		-0.36	
	Burbot		-0.81	
		Northern rockbass	0.85	
		White sucker	0.83	
		Eurasian ruffe	0.70	
5	Brown bullhead		-0.30	
	Bowfin		-0.47	
	Bluegill sunfish			-0.49
		Emerald shiner	0.62	
		Spottail shiner	0.70	
		Northern rockbass	0.51	0.57
		White sucker		0.59
		Johnny darter		0.41
		Smallmouth bass		0.58
		Spotfin shiner		0.73

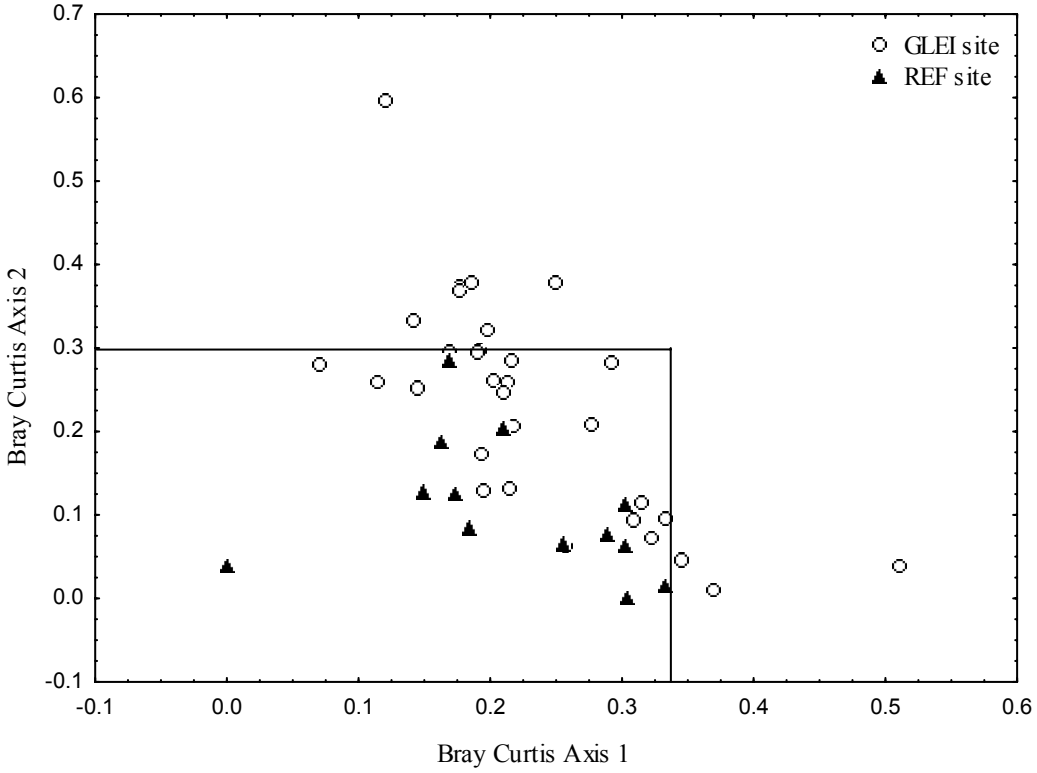

Figure 4.4. Relationship between BC axis 1 and BC axis 2 ordination scores for cluster 1 sites. Threshold lines indicate boundaries of reference sites along both axes.

Table 4.6. Classification of GLEI sites whose fish species relative abundances identify them as being equivalent to reference sites or not, based on threshold criteria and ordination distances from reference endpoint (see Appendix I for site locations).

Site	Reference	Non-reference	BC Axis 1	BC Axis 2
Cluster 1				
273 Em	X		0.113	0.260
291 Rw	X		0.145	0.252
294 Cw	X		0.169	0.294
299 Rw		X	0.120	0.596
300 Cw		X	0.176	0.374
300 Em	X		0.195	0.130
302 Cw		X	0.249	0.379
313 Em	X		0.315	0.116
321 Rw	X		0.292	0.284
325 He		X	0.511	0.040
350 He		X	0.332	0.096
359 He		X	0.344	0.046
361 He		X	0.256	0.063
361 Rw		X	0.177	0.369
364 Pw		X	0.185	0.378
374 He		X	0.215	0.133
374 Rw	X		0.191	0.297
378 Rw	X		0.189	0.294
388 He		X	0.322	0.073
393 He		X	0.369	0.009
420 Em	X		0.216	0.286
421 Cw	X		0.193	0.172
425 He		X	0.308	0.094
573 Rw	X		0.070	0.282
578 He		X	0.276	0.209
585 Em	X		0.213	0.259
586 He		X	0.202	0.262
590 Rw	X		0.209	0.247
599 Cw		X	0.197	0.321
606 Cw	X		0.217	0.207
646 Em		X	0.141	0.334

Table 4.6 Continued

Site	Reference	Non-reference	BC Axis 1	BC Axis 2
Cluster 2				
59 Rw		X	0.247	0.241
63 Em	X		0.285	0.154
68 Rw	X		0.175	0.142
86 Pw		X	0.248	0.211
104 Rw	X		0.249	0.090
105 Pw		X	0.283	0.259
107 Cw		X	0.271	0.254
109 Cw	X		0.236	0.197
110 Pw		X	0.240	0.211
192 Em	X		0.282	0.077
192 Pw	X		0.275	0.036
200 Pw		X	0.289	0.230
217 Em		X	0.313	0.154
247 Em		X	0.155	0.257
248 Cw	X		0.194	0.144
275 Cw	X		0.243	0.197
281 Rw		X	0.242	0.243
289 He		X	0.253	0.107
289 Pw		X	0.431	0.162
312 Em		X	0.501	0.109
379 Rw		X	0.250	0.307
389 Rw		X	0.223	0.365
419 Em	X		0.228	0.186
422 Rw		X	0.222	0.364
442 Cw		X	0.279	0.244
446 Cw		X	0.251	0.212
447 Cw		X	0.272	0.277
448 Cw		X	0.256	0.534
453 Cw		X	0.374	0.224
458 Cw		X	0.263	0.262
470 He		X	0.272	0.195
486 Cw	X		0.196	0.119
Cluster 3				
247 Cw		X	0.925	0.086
257 He		X	0.296	0.227
417 Em		X	0.727	0.103
Cluster 4				
42 He		X	0.451	0.056
51 He		X	0.724	0.044
55 Pw		X	0.541	0.335

Table 4.6 Continued

Site	Reference	Non-reference	BC Axis 1	BC Axis 2
Cluster 5				
372 Pw	X		0.155	0.025
393 Rw	X		0.225	0.110
420 Pw		X	0.303	0.124
578 Pw	X		0.225	0.042
580 Pw	X		0.222	0.176
581 Em		X	0.396	0.079
588 Pw	X		0.158	0.120
593 Pw	X		0.167	0.088
602 Pw	X		0.270	0
645 He		X	0.573	0.099
718 Pw	X		0.161	0.093
719 Rw	X		0.211	0.042
723 Pw		X	0.172	0.205
739 Rw	X		0.198	0.089
743 Cw	X		0.194	0.142
744 Pw	X		0.190	0.051
746 Em		X	0.271	0.231
748 Em		X	0.256	0.223
748 Cw		X	0.291	0.282
752 Em		X	0.242	0.321
755 Em		X	0.307	0.120

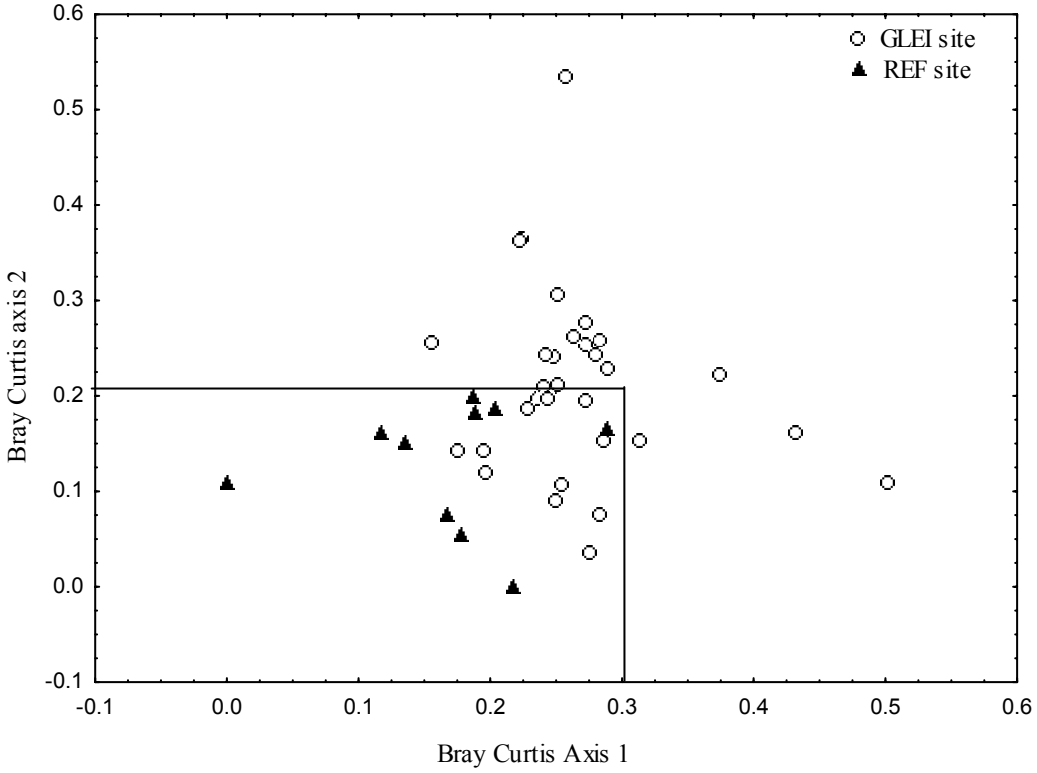

Figure 4.5. Relationship between BC axis 1 and BC axis 2 ordination scores for cluster 2 sites. Threshold lines indicate boundaries of reference sites along both axes.

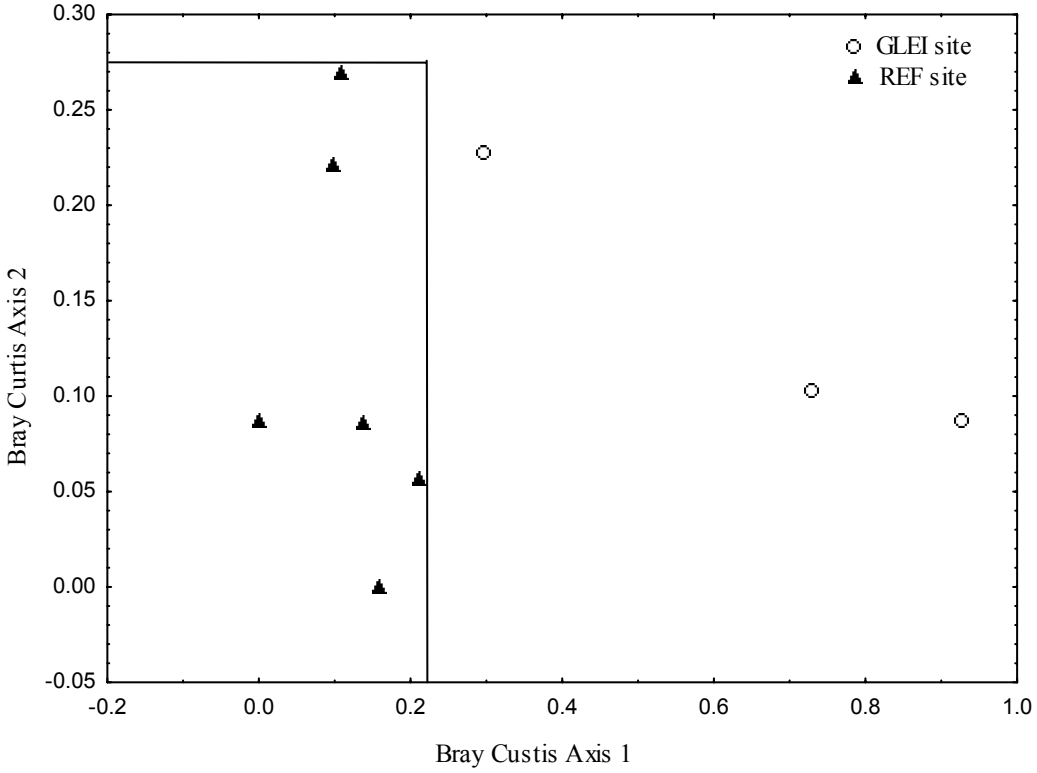

Figure 4.6. Relationship between BC axis 1 and BC axis 2 ordination scores for cluster 3 sites. Threshold lines indicate boundaries of reference sites along both axes.

Reference and non-reference GLEI sites were classified based on the threshold criteria (Table 4.6).

Cluster 4 – Longnose dace, White sucker

Cluster 4 consisted of 9 sites (5 reference sites, 3 test sites and 1 'centroid' site and 11 species. BC Axis 1 extracted 60.55% of the original distance matrix while BC axis 2 extracted 12.19% of the original distance matrix for a cumulative extraction of 72.74%.

The ordination scores for both axes were significantly correlated with both population pressure ($r = 0.94$, $p < 0.001$) and agricultural pressure ($r = 0.87$, $p < 0.01$). We determined indicator species at both reference and non-reference ends of both pressure axes (Table 4.5). The threshold line for reference GLEI sites along axis 1 was drawn at 0.24 and for axis 2 at 0.12 (Figure 4.7). GLEI sites were classified as reference or non-reference based on the threshold criteria (Table 4.6).

Cluster 5 – Burbot, Eastern longnose sucker

Cluster 5 represented 33 sites (11 references sites, 21 test sites and 1 'centroid' site) and 24 species. The two BC axes extracted 44.20% of the original distance matrix: 25.92% for axis 1 and 18.28% for axis 2. The ordination scores for both BC axes were significantly positively correlated with population pressure ($r = 0.45$, $p < 0.01$) and agricultural pressure ($r = 0.52$, $p < 0.01$). Indicator species for the reference and non-reference ends of both pressure axes are listed in Table 4.5. Threshold lines were drawn at 0.29 along BC axis 1 and at 0.19 along BC axis 2 (Figure. 4.8). GLEI sites were assessed according to the reference threshold criteria (Table 4.6).

Model Testing

Of the 10 reserve sites, 4 sites belonged to cluster 2, 1 site belonged to cluster 3 and 5 sites belonged to cluster 5. We extrapolated BC axis 1 and BC axis 2 scores for each site based on the graphs of ordination scores against pressure scores (Table 4.7). Using the threshold criteria (already set in place for GLEI sites); we determined that 8 of the 10 reserve site were classified as similar to reference sites. One site from cluster 2 and 1 site from cluster 5 was assessed as non-reference based on their BC axes scores (Table 4.7).

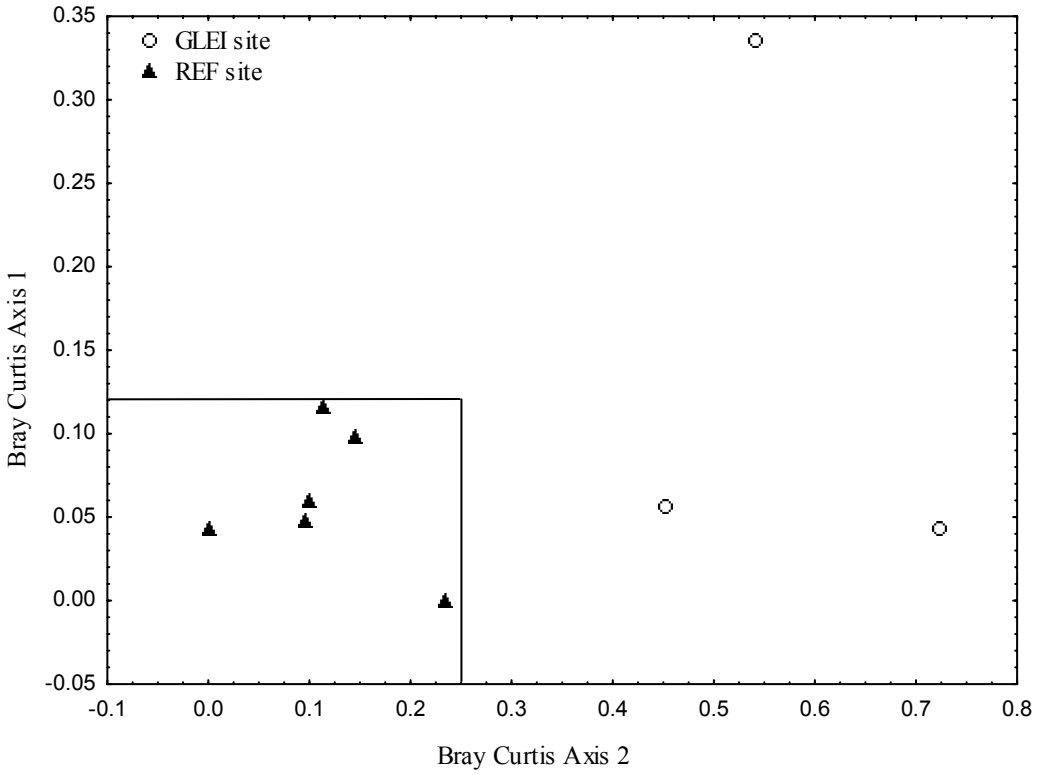

Figure 4.7. Relationship between BC axis 1 and BC axis 2 ordination scores for cluster 4 sites. Threshold lines indicate boundaries of reference sites along both axes.

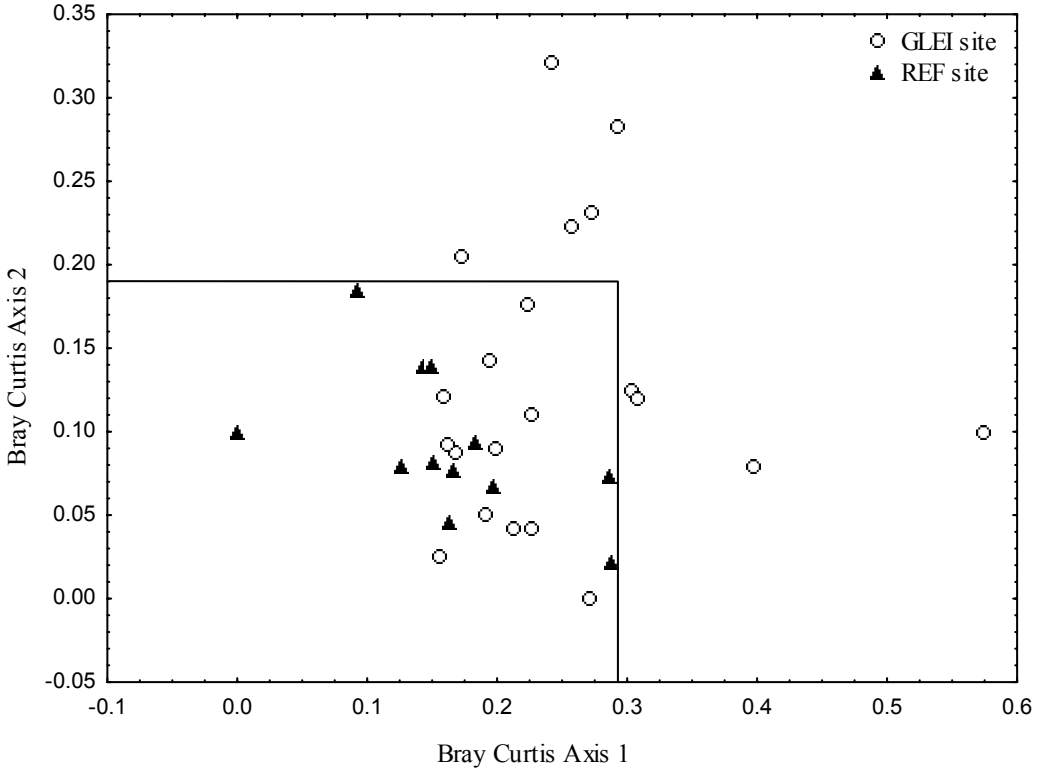

Figure 4.8. Relationship between BC axis 1 and BC axis 2 ordination scores for cluster 5 sites. Threshold lines indicate boundaries of reference sites along both axes.

Table 4.7. Classification of reserve sites as reference or non-reference based on extrapolated ordination scores.

Reserve sites	Reference	Non-reference	BC Axis 1	BC Axis 2
Cluster 2				
180 He	X		-0.05	-0.05
213 Rw	X		-0.02	0.17
215 Rw	X		0.1	0.1
407 Cw		X	0.19	0.26
Cluster 3				
213 He	X		0.1	0.2
Cluster 5				
739 He		X	< -0.1	0.2
586 Pw	X		0.3	< -0.05
588 Pw	X		0.29	0.06
720 Rw	X			

When comparing the species assemblages at the reference and non-reference end of the axes with the indicator species list (Table 4.5), we found that most of the reserve sites in Cluster 2 had a mixture of high relative abundances of both reference and non-reference indicator species (Table 4.8). Reserve sites belonging to clusters 3 and 5 however, showed results more consistent with expectations, in that there was a high abundance of reference indicator species at reference sites and a high abundance of non-reference indicator species at non-reference sites.

Discussion

Species Assemblages

The cluster analysis of REF sites revealed unique assemblages of fish species among groups of sites. Samples from sites in Cluster 5 for instance, were well represented by burbot (*Lota lota*) and eastern longnose sucker (*Catostomus catostomus*), both cold-cool water species (Coker et al. 2001), whereas sites making up Cluster 3 tended to be dominated by bluegill sunfish (*Lepomis macrochirus*) and pumpkinseed sunfish (*Lepomis gibbosus*), both belonging to warm water thermal groups (Coker et al. 2001). Spottail shiner (*Notropis hudsonius*) and yellow perch (*Perca flavescens*), which were the prevalent species at sites making up Cluster 1 also both belong to the same thermal group (cool-water species), but are among the most frequently recorded species in wetlands throughout the Great Lakes regions (Jude and Pappas 1992). Since the dominant species found in most clusters belonged to the same thermal groups, we had expected mean temperature to be an important variable in separating the clusters of sites. Furthermore, species such as longnose dace (*Rhinichthys cataractae*) are rare in Lake Ontario and Lake Erie, but are commonly found in northern Lake Michigan (Jude and Pappas 1992).

Pumpkinseed sunfish on the other hand have been reported as abundant in Lake Ontario marshes (Jude and Pappas 1992). We had thus expected lake and correspondingly, latitude, to also be important variables separating groups of sites and fish species. However, neither lake nor temperature proved to be important diagnostic variables relative to the others identified by the discriminant function model.

Table 4.8. Relative abundances of fish species at reserve sites. Species with 0 abundances in all sites have been omitted from the table.

Cluster 2	Blacknose shiner	Bluegill sunfish	Bluntnose minnow	Brown bullhead	Johnny darter	Northern pike
180 He	0	0	0	0	0	0
213 Rw	5.27	0.59	0	2.83	1.01	0.33
215 Rw	1.90	0.38	0	0	0	0.69
407 Cw	0	0	2.43	0	0	0

Cluster 2	Smallmouth bass	Spottail shiner	White sucker	Yellow perch	Pumpkinseed sunfish	Sand shiner
180 He	0	6.20	3.11	0.42	0	0
213 Rw	3.07	0	3.35	3.07	3.35	0
215 Rw	0.38	0	0.38	5.67	2.69	0
407 Cw	3.83	2.70	0	0	0	4.72

Cluster 3	Burbot	Longnose Dace	Troutperch	White Sucker
213 He	1.69	5.33	5.06	2.04

Cluster 5	Bluegill Sunfish	Bowfin	Brown Bullhead	Emerald Shiner	Johnny Darter	Northern Rock Bass	Smallmouth Bass
739 He	0.93	0	0	0.75	0	2.68	0.93
586 Pw	3.34	1.99	0.44	0.22	0	0.00	0
588 Pw	4.49	2.43	0.32	0	0	0	0
720 Rw	1.77	0.93	0.38	0	0.38	0	0
695 Cw	3.60	2.33	2.82	0	0	1.59	0

Cluster 5	Spottail shiner	White sucker
739 He	0.29	0.29
586 Pw	0	0
588 Pw	0	0
720 Rw	0	0
695 Cw	0	0

Environmental Influences

The discriminant function analysis revealed that ecoregion was perhaps the most important variable in the model, distinguishing among the 5 groups of sites with unique fish assemblages. Four of the six ecoregion variables had very low Wilk's lambda and p values, indicating that at least 5 of the ecoregions were significantly associated with particular clusters of sites. This finding corroborates studies that have developed ecoregion-specific IBIs based on the contention that fish communities are structured along ecoregions (Simon 1991, Lyons 1992).

Latitude, as we predicted was another important variable in the DFA model. Using canonical correspondence analysis, Wang et al. (2003) determined that latitude (among other variables) was a significant correlate of fish species abundance and presence-absence data. Latitude has also been implicated as a primary explanation factor describing the limits of macrophyte communities in the Great Lakes (Crowder and Painter 1991, Lougheed et al. 2001). Among vegetation-related variables, we had used percent emergent, submergent and floating vegetation cover, as well as growth form richness in the DFA model. We also used dominant *Typha* and *Scirpus* percent cover in our model, based on the report by Uzarski et al. (2005) that fish species composition within coastal wetlands is primarily correlated with plant zones. Percent emergent cover was the only vegetation variable that was statistically significant in classifying sites with different fish communities. This is consistent with other studies that have shown that emergent and submergent plant cover, are important variables in fish species associations (Wang et al. 2003). Brazner and Beals (1997) noted that macrophyte cover and diversity can be used as a predictor variable for bluegill sunfish and golden shiner abundance in Saginaw Bay of Lake Huron.

Sites protected from direct lake influences provide unique habitats to certain species, which may not thrive in open water or embayment sites that lack macrophytes. Protected wetlands, by definition, are habitats sheltered from high energy waves by a permanent sand bar or dyke and typically have significant macrophyte coverage (Keough et al. 1999). Species that favour vegetated areas include white bass, spotfin shiner, spottail shiner and bluntnose minnow (Jude and Pappas 1992). Wei et al. (2004) found that wetland type, especially open-shoreline wetland versus protected embayments

influenced the distribution of certain taxa. The protected wetland variable was the only one that was significant in segregating the clusters of sites. Sampling day, although often recorded is not considered in analyses involving fish associations with environmental variables (e.g. Wang et al. 2003, Waite and Carpenter 2000, Maret and MacCoy 2002). The DFA model showed that sampling day was also a significant variable in separating clusters of sites, indicating that fish community composition follows basin-wide phenology. Although our fish sampling was conducted between May and August, during which time fish species tend to remain in the same area or habitat, we suspect that there could have been some variability in the catch between months when temperatures would rise and also when macrophyte cover may decline due to receding late-summer water level fluctuations.

Stressor Influences

We used Bray-Curtis ordination techniques to develop criteria for assessing the quality of sites (in terms of fish community composition) along the stressor gradients previously determined from GIS-based measures of human activity. Our results showed clear distinctions and strong relationships between the fish community ordination scores and both anthropogenic pressure axes only for sites classified as belonging in clusters 4 and 5. Cluster 2 sites showed a strong relationship between ordination scores and population pressure, but not agricultural pressure, which suggests that species present in cluster 2 sites are more strongly affected by the range of population density-related stress than by the range of agriculture-related stress. Sites associated with clusters 1 and 3, however, did not show any statistically significant relationships along either stressor axis, and most of the reference area sites within these clusters were found to have both high ordination scores and high pressure scores, contrary to what was expected.

When assessing the distribution of sites in Cluster 1, we noted that most of the sites belonged to three different ecoregions (Northern Great Lakes, South Central Great Lakes and Southwestern Great Lakes) but more importantly, the REF area sites, which were expected to be seen in the lower quadrant of the graph (Figures 4.2 and 4.3), were found to be distributed across the entire stress gradient. The same was true for cluster 3 sites, which suggests that either the REF sites were not representative of good quality at

the basin-wide scale, or the GLEI sites within those clusters were of higher quality than the reference sites. One possible reason for this result is attributable to the fact that REF sites were selected at finer scales than the GLEI sites (Host et al. in press). Reference area sites were locations whose maximum exposure to disturbance was within the lowest 20th percent of all locations within their ecoregion, as determined at the most local watershed scale. Other stresses that may have been significant within the bounds of the second-order or higher segments-shed but which weren't linked directly with the wetland evaluated would contribute to a segment-shed level analysis but not the watershed level assessment of disturbance. Furthermore, the best wetlands that could be found within an ecoregion such as the EOL may have been subject to significantly more disturbance than non-reference wetlands in ecoregions such as South Superior Uplands (SSU), where the overall level of human activity is minimal by comparison. Similarly, GLEI test sites determined to have low stress scores at the segment-shed scale could potentially be subject to high stress scores at the watershed level. The scale mismatch, likely contributes to the relatively weak relationship observed between fish community and stress. A comparison of GLEI watershed-level stressor scores to fish community index scores would be more appropriate.

When we compared the species assemblages from sites at the reference and non-reference ends of the axes with the indicator species list (Table 4.5), Site 407 Cw, which was classified as a non-reference site in Cluster 2, had a high relative abundance of sand shiner and smallmouth bass, both species noted as reference indicator species (Table 4.5). Three other reserve sites in cluster 2 were classified as reference sites but had a high relative abundance of reference and non-reference species, contrary to expectations. The GLEI sites present in cluster 2 belonged to 3 different ecoregions, and although the cluster analysis grouped the sites based on species composition, some species that would normally be found in a northern ecoregion, could potentially be very abundant in another ecoregion due to their broad environmental tolerance, and thus result in a common grouping of sites that are really disparate across latitudes. We obtained results more consistent with expectations when we examined the reserve site classified as belonging to Cluster 3. The site was classified as a reference site and as expected, had relatively high abundances of longnose dace and troutperch, two species that were recognized as

reference indicator species for Cluster 3. The results from Cluster 5 also showed expected results. Site 739 He, which was classified as a non-reference site had a high abundance of northern rock bass (a non-reference indicator species), whereas the remaining 4 reserve sites in Cluster 5 had high relative abundances of the reference indicator species, bowfin, brown bullhead and bluegill sunfish, in accordance with expectations.

Summary and Conclusions

The multivariate approach to identifying biological indicators of environmental condition is unique in that faunal data are used to group together reference sites that have similar species composition, thus providing an objective strategy for assessing habitat quality based on species assemblages. In this paper, we were able to develop models of fish species relative abundances that can be used to evaluate the quality of sites in response to specific anthropogenic stressors. Results from testing the model with reserve sites were consistent with expectations for sites belonging to 2 of the clusters in that sites classified as 'reference' based on extrapolated ordination scores also had a high abundance of the reference indicators species (determined through the Ordination analysis). Overall, the multivariate model seems to be an effective method of determining indicators species assemblages as well as assessing fish community composition along Great Lakes coastal margins.

GENERAL DISCUSSION

The purpose of my research has been to evaluate the multimetric index of biotic integrity (IBI) approach and the multivariate approach of assessing habitat quality for Great Lakes sites. Few studies have directly compared the efficacy of the IBI approach relative to multivariate approaches in assessing water quality and impairment as a result of anthropogenic stressors. Reynoldson et al. (1997) tested two multivariate methods of habitat assessment (AUStralian RIVer Assessment Scheme: AusRiVAS, and Benthic Assessment of SedimenT: BEAST) against the multimetric method more commonly applied in the United States, and found that the multivariate methods applied to stream macroinvertebrate community composition yielded greater precision and accuracy in classifying sites as reference versus non-reference than multimetric methods. Milner and Oswood (2000) tested the TWINSPAN multivariate classification technique (Hill 1979) against the multimetric IBI approach and a hybrid multimetric/multivariate technique (which used rapid bioassessment metrics as input variables for the TWINSPAN classification technique). They concluded that the multivariate technique produced the classification with the highest overall strength, while the multimetric approaches resulted in the lowest strength of classification of reference sites. Both these studies however, were conducted for stream assessment using benthic invertebrates. Although fish communities are largely favoured for stream and wetland assessments particularly for IBI development (Plafkin et al. (1989), no studies to date have been conducted on comparing the robustness of the two approaches using fish communities, especially for Great Lakes coastal margins.

We chose ecoregion as the geographic framework for developing an IBI and setting expectations based on the observation that fish community composition differs among ecoregions (Lyons 1989). Fish community composition is further believed to be unique in wetlands compared to littoral zones (Jude and Pappas 1999). Although we evaluated wetland IBIs only for two ecoregions, EOL and NGL, we did pool sites across different wetland types: river-influenced wetlands, lacustrine wetlands and protected wetlands, based on previous studies that have followed similar methodology. Uzarski et al. (2004) developed a single invertebrate IBI for open lacustrine marshes (synonymous to our definition of lacustrine wetlands) and protected-fringing lacustrine marshes

(synonymous to our definition of protected wetlands) of lakes Huron and Michigan. Uzarski et al. (2005) developed two fish-based IBIs for coastal wetlands (open lacustrine, protected lacustrine, barrier-beach, and drowned river mouth) of all 5 Great Lakes.

The results from Chapter 2 showed that very few metrics varied across stressor gradients for the Erie and Ontario (EOL) ecoregion and thus there were insufficient variables to meet our criterion for IBI development. The metrics chosen for testing were based on published literature and species composition at reference sites. One explanation for the failure of significant metric score differences between reference and non-reference sites is that the species composition at reference sites may not have been representative of true reference condition in that the reference sites may not have been minimally disturbed. Johnson et al. (1989) noted that fish species composition in Lake Erie wetlands is becoming increasingly dominated by more tolerant species that are able to withstand higher temperatures and periods of low dissolved oxygen. With the continual increase in agricultural pressure in the lower Lake Erie basin, it is possible that our reference condition sites, although relatively less disturbed than other Lake Erie test sites, were not different enough from degraded sites to have had dissimilar fish species composition and hence, higher overall integrity, than the more degraded sites present in the same ecoregion.

The results for the Northern Great Lakes (NGL) ecoregion indicated that an IBI can indeed be developed for this ecoregion, but we were unable to test the IBI with a sufficient number of sites to properly evaluate our results. Further testing of the IBI was recommended with data from at least 15 wetlands of the same hydrogeomorphic type and containing varying degrees of disturbance.

Most studies to date have tested metric response using stressor information from land use and agriculture measured at the watershed scale (Burton et al. 1999, Wilcox et al. 2002). The comparison of the scale of metric response (segment-shed versus watershed) indicated that more metrics varied significantly as a function of stress at the segment-shed level than at the finer, watershed scale. This further suggests that fish species in Great Lakes wetlands reflect degree of stress at broader geographical scales than previously believed, concurrent with findings of Holland et al. (in preparation). It is important to not only detect anthropogenic disturbance, but also to identify the nature of the disturbance

or suite of disturbances that is likely causing observed changes in metric scores (Burton et al. 1999). The detailed GLEI stressor data reported by Danz et al. (2005) allowed us to not only compare the scale of metric response, but also distinguish differences in metric response to different types of anthropogenic stressors. With the availability of detailed stressor data, we were further able to test two independently developed IBIs (Uzarski et al. 2005) for two major plant zones in Great Lakes coastal wetlands, *Typha* and *Scirpus* (Chapter 3). Our results conclusively showed that the *Typha* and *Scirpus* IBI's selectively reflected specific, unique classes of stress and thus were not necessarily good overall indicators of generalized stress in wetlands dominated by either plant zone.

The results derived from testing multivariate techniques (Chapter 4) showed that ecoregion is an important variable across which a distinct association of fishes exist in reference area wetlands, and this corroborates our decision to use ecoregion as a classification framework for IBI development in Chapter 2. Cluster analysis grouped together reference sites belonging to different hydrogeomorphic types, and apart from the protected wetlands categorical variable, HGM type did not seem to be one of the important variables in segregating fish communities. Although the ordination analyses indicated the presence of strong relationships between fish community ordination scores and stressor scores for 3 clusters of sites, sites from two clusters (clusters 1 and 3) did not show significant correlations with stressor scores. The sites in these latter clusters belonged to different ecoregions and hydrogeomorphic types. However, throughout Furthermore, in our entire study, reference sites were compared to test sites that were ranked at a different scale in terms of their stressor scores, and this may have obscured the patterns obtained with the ordinations. Nonetheless, the method of ordinating sites in stressor space allowed us to develop scales permitting us to assess the overall condition of sites along the entire U.S. coastline. We were further able to identify a set of reference and non-reference indicator species for each cluster along each pressure axis.

General Conclusions

Numerous studies have been conducted on assessing habitat quality in streams, rivers, wetlands and lakes and these studies have often shown that fish species serve as good indicators of stress especially in comparison to taxa such as diatoms or benthic invertebrates. My research suggests that although fish species can be used in indicator development, the methodology behind developing indicators may be inherently flawed. For instance, the IBI approach relies heavily on a priori assumptions that are used to set metric expectations. The metrics are usually selected based on already published literature (e.g. Minns et al. 1994, Wilcox et al. 2002) which in retrospect may bias results if those metrics were originally developed for different systems (such as streams), different regions, latitudes or even at different time periods. Temporal variability is not often addressed in assessment studies mainly because of the difficulty associated with continual data collections of fish sampling and more importantly stressor information. Furthermore, IBIs that are developed for specific regions are validated by the researchers using land-use data obtained from outdated maps and government data bases, which ultimately requires further testing and validation as often suggested by the authors themselves (e.g. Wilcox et al. 2002, Burton et al. 1999).

The multivariate approach showed promising results. Relative abundance data were used to determine the factors important in grouping sites for assessment, which further allowed for individual group assessment of sites (i.e. clusters) against combinations of anthropogenic stressors. A more affective approach in habitat assessment would be to use GIS and remote sensing data in conjunction with in situ sampling data to develop robust indicators of stress.

The multimetric IBI approach is favoured by managers due to the relatively easily model construction compared to multivariate approaches. However, the rationale behind using one type of regional framework in developing an IBI over another is often questioned. A more adaptable approach to indicator development would be to integrate multimetric and multivariate techniques such that cluster analysis is used to determine groupings of sites based on unique fish associations followed by the development of IBIs for each cluster of reference sites, distinguished from other clusters based on variables identified through a discriminant function analysis.

Future Research

In attempting to develop wetland IBIs, we were able to compare evaluate the association between fish biotic integrity and anthropogenic stress measured at the scale of segment-shed and watershed. Since the results from the multivariate analyses produced clusters of sites based on fish species composition, we would have ideally wanted to compare the associations between segment-shed stressor and watershed stressors for each of the 5 clusters already identified. Watershed scale stressor data were not available for high energy and embayment sites, but are currently being compiled. Once available, I plan to do a scale comparison (as done in Chapter 2) using grouping of sites from the multivariate results. Evaluating fish association between reference and non-reference sites, whose stress scores are measured at the same scale, would lead to more consistent results.

Although we were able to develop an IBI for the NGL ecoregion, we had insufficient sites to effectively test and validate the IBI. An additional component to the study would thus be to sample at least 15 wetlands along the Great Lakes coastline of varying anthropogenic disturbance to see if the IBI is responsive to different types of stressors and capable of assessing overall integrity of the sites. Multiple sampling events at the 15 sites would allow us to evaluate temporal variability of IBI scores between replicates, which would give us a measure of model precision. Ultimately, the precision between both multimetric and multivariate models should be objectively evaluated to determine which approach is most effective in assessing habitat quality using fish communities.

BIBLIOGRAPHY

Barbour, M. T., J. Gerritsen, B. D. Snyder, and J. B. Stribling. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish. Second edition. U.S. Environmental Protection Agency. Office of Water, Washington, D.C.

Brazner, J. C., and E. W. Beals. 1997. Patterns in fish assemblages from coastal wetland and beach habitats in Green Bay, Lake Michigan: a multivariate analysis of abiotic and biotic forcing factors. Canadian Journal of Fisheries and Aquatic Sciences 54:1743-1761.

Coker, G. A., C. B. Portt, and C. K. Minns. 2001. Morphological and ecological characteristics of Canadian freshwater fishes.

Crowder, A., and D. S. Painter. 1991. Submerged macrophytes in Lake Ontario: current knowledge, importance, threats to stability, and needed studies. Canadian Journal of Fisheries and Aquatic Sciences 48:1539-1545.

Danz, N. P., R. R. Regal, G. J. Niemi, V. J. Brady, T. Hollenhorst, L. B. Johnson, G. E. Host, J. M. Hanowski, C. A. Johnston, T. Brown, J. Kingston, and J. R. Kelly. 2005. Environmentally stratified sampling design for the development of great lakes environmental indicators. Environmental Monitoring and Assessment 102:41-65.

Day, R. T., P. A. Keddy, J. McNeill, and T. Carleton. 1988. Fertility and disturbance gradients: a summary model for riverine marsh vegetation. Ecology 69:1044-1054.

Detenback, N. E., C. A. Johnston, and G. J. Niemi. 1993. Wetland effects on lake water quality in the Minneapolis/St. Paul metropolitan area. Landscape Ecology 8:39-61.

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

Emery, E. B., T. P. Simon, F. H. McCormick, P. L. Angermeier, J. E. Deshion, C. O. Yoder, R. E. Sanders, W. D. Pearson, G. D. Hickman, R. J. Reash, and J. A. Thomas. 2003. Development of a multimetric index for assessing the biological condition of the Ohio river. Transactions of the American Fisheries Society 132:791-808.

Environment Canada. 2002. Great Lakes facts sheet. Catalogue No. En 40-222/11-2002E.

Fausch, K. D., J. R. Karr, and P. R. Yant. 1984. Regional Application of an Index of Biotic Integrity Based on Stream Fish Communities. Transactions of the American Fisheries Society 113:39-55.

Fausch, K. D., and J. Lyons. 1990. Fish communities as indicators of environmental degradation. Pages 123-144 in S. M. Adams, editor. Biological Indicators of Stress in Fish. American Fisheries Society Symposium 8, Bethesda, M.D.

Fore, L. S., J. R. Karr, and R. W. Wisseman. 1996. Assessing invertebrate responses to human activities: evaluating alternative approaches. Journal of the North American Benthological Society 15:212-231.

Frey, D. G. 1977. Biological integrity of water - an historic approach. Pages 127-140 in R. K. Ballantine and L. J. Guarraia, editors. The Integrity of Water. Proceedings of a Symposium, U.S. Environmental Protection Agency, Washington, D.C.

Galatowitsch, S. M., N. O. Anderson, and P. D. Ascher. 1999. Invasiveness in wetland plants in temperate North America. Wetlands 19:733-755.

Galatowitsch, S. M., D. C. Whited, R. Lehtinen, J. Husveth, and K. Schik. 2000. The vegetation of wet meadows in relation to their land-use. Environmental Monitoring and Assessment 60:121-144.

Gauch, H. G. 1982. Sampling methods. Pages 43-71 in E. Beck, H. J. B. Birks, and E. F. Connor, editors. Multivariate Analysis in Community Ecology. Cambridge University Press, New York, USA.

Gesch, D., M. Oimoen, S. Greenlee, C. Nelson, M. Steuck, and D. Tyler. 2002. The National Elevation Dataset. Photogrammatic Engineering and Remote Sensing 68:5-12.

Government of Canada and U.S. Environmental Protection Agency. 1995. The Great Lakes: an environmental atlas and resource book. Third edition.

Grabas, G., A. D. Pemanen, M. Galloway, and K. Holmes. 2004. Durham Region Coastal Wetland Monitoring Project: Year 2 Technical Report. ECB-OR. Environmental Canada and Central Lake Ontario Conservation Authority, Downsview, Ontario, Canada.

Green, R. H., and G. L. Vascotto. 1978. A method for the analysis of environmental factors controlling patterns of species composition in aquatic communities. Water Research 12:583-590.

Hill, M. O. 1979. TWINSPAN: a FORTRAN program for arranging multivariate data in an ordered two-way table by classification on the individuals and the attributes. In Ecology and Systematics, Cornell University, Ithaca, NY.

Holland, J. D., J. J. H. Ciborowski, T. P. Hollenhorst, and L. B. Johnson. In preparation. The spatial scale at which fish indicator species indicate. For submission to Ecological Indicators.

Host, G. E., J. Schuldt, J. J. H. Ciborowski, L. B. Johnson, T. Hollenhorst, and C. Richards. 2005. Use of GIS and remotely sensed data for a priori identification of reference areas for Great Lakes coastal ecosystems. International Journal of remote sensing. In press.

Hughes, R. M., D. P. Larsen, and J. M. Omernik. 1986. Regional Reference Sites - a Method for Assessing Stream Potentials. Environmental Management 10:629-635.

Hughes, R. M., and D. P. Larson. 1988. Ecoregions: an approach to surface water protection. Journal of the Water Pollution Control Federation 60:486-493.

Hughes, R. M., and T. Oberdorff. 1999. Applications of IBI concepts and metrics to waters outside the United States and Canada. Pages 79-93 in T. P. Simon, editor. Assessing the sustainability and biological integrity of water resources using fish communities. CRC Press, Boca Raton, Florida.

Hughes, R. M., E. Rexstad, and C. E. Bond. 1987. The Relationship of Aquatic Ecoregions, River Basins and Physiographic Provinces to the Ichthyogeographic Regions of Oregon. Copeia:423-432.

Janisch, J. E., and N. E. Molstad. 2004. Disturbance and the three parameters of wetland delineation. Wetlands 24:820-827.

Johnson, D. L. 1989. Lake Erie wetlands: fisheries considerations. Pages 257-274 in K. A. Krieger, editor. Lake Erie Estuarine Systems: Issues, resources, status, and management. NOAA Estuary-of-the-Month Seminar Series No. 14. Estuarine Program Office, Washington, D.C.

Jude, D. J., and J. Pappas. 1992. Fish utilization of Great Lakes coastal wetlands. Journal of Great Lakes Research 18:651-672.

Karr, J. R. 1981. Assessment of Biotic Integrity Using Fish Communities. Fisheries 6:21-27.

Karr, J. R., and E. W. Chu. 1999. Restoring life in running waters: better biological monitoring. Island Press. Covelo, California.

Karr, J. R., K. D. Fausch, P. L. Angermeier, P. R. Yant, and I. J. Schlosser. 1986. Assessing biological integrity in running waters: a method and its rationale. Champaign, Illinois.

Keough, J. R., T. A. Thompson, G. R. Guntenspergen, and D. A. Wilcox. 1999. Hydrogeomorphic factors and ecosystem responses in coastal wetlands of the Great Lakes. Wetlands 19:821-834.

Larsen, D. P., J. M. Omernik, R. M. Hughes, C. M. Rohm, T. R. Whittier, A. J. Kinney, A. L. Gallant, and D. R. Dudley. 1986. Correspondence between Spatial Patterns

in Fish Assemblages in Ohio Streams and Aquatic Ecoregions. Environmental Management 10:815-828.

Leonard, P. M., and D. J. Orth. 1986. Application and Testing of an Index of Biotic Integrity in Small, Coolwater Streams. Transactions of the American Fisheries Society 115:401-414.

Lougheed, V. L., B. Crosbie, and P. Chow-Fraser. 2001. Primary determinants of macrophyte community structure in 62 marshes across the Great Lakes basin: latitude, land use, and water quality effects. Canadian Journal of Fisheries and Aquatic Sciences 58:1603-1612.

Lyons, J. 1989. Correspondence between the Distribution of Fish Assemblages in Wisconsin Streams and Omernik Ecoregions. American Midland Naturalist 122:163-182.

Lyons, J. 1992. Using the index of biotic integrity (IBI) to measure environmental quality in warmwater streams of Wisconsin. General Technical Report NC-149. North Central Forest Experiment Station. U.S. Dept. of Agriculture, Forest Service, St. Paul, MN.

Lyons, J., L. Wang, and T. D. Simonson. 1996. Development and validation of an index of biotic integrity for coldwater streams in Wisconsin. North American Journal of Fisheries Management 16:241-256.

Maret, T. R., and D. E. Maccoy. 2002. Fish assemblages and environmental variables associated with hard-rock mining in the Coeur d'Alene River Basin, Idaho. Transactions of the American Fisheries Society 131:865-884.

McCune, B., and J. B. Grace. 2002. Analysis of Ecological Communities. MjM Software Design, Gleneden Beach, Oregon.

McCune, B., and M. J. Mefford. 1999. Multivariate Analysis of Ecological Data. Gleneden Beach, Oregon, U.S.A.

Minc, L. D. 1997. Great Lakes coastal wetlands: an overview of controlling abiotic factors, regional distribution and species composition. Michigan Natural Features Inventory, Lansing, MI.

Minns, C. K., V. W. Cairns, R. G. Randall, and J. E. Moore. 1994. An Index of Biotic Integrity (IBI) for Fish Assemblages in the Littoral-Zone of Great-Lakes Areas of Concern. Canadian Journal of Fisheries and Aquatic Sciences 51:1804-1822.

Mundahl, N. D., and T. P. Simon. 1999. Development and application of an index of biotic integrity for coldwater streams of the upper midwestern United States. Pages 383-415 in T. P. Simon, editor. Assessing the Sustainability and Biological

Integrity of Water Resources Using Fish Communities. CRC Press, Boca Raton, Florida.

Niemi, G.L., R. Axler, B. Bedford, J. Ciborowski, G. Host, L. Johnson, C. Johnston, J. Kingston, R. Regal, C. Richards, D. Swackhamer, and R. Howe. 2001. Development of environmental indicators of condition, integrity and sustainability in the Great Lakes basin. Research proposal submitted to US EPA Star Program. US EPA/R- 8286750. 25 p.

Norris, R. H. 1995. Biological monitoring: the dilemma of data analysis. Journal of the North American Benthological Society 14:440-450.

Omernik, J. M. 1987. Ecoregions of the Conterminous United-States. Annals of the Association of American Geographers 77:118-125.

Omernik, J. M. 1995. Ecoregions: a spatial framework for environmental management. Pages 49-62 in W. S. Davis and T. P. Simon, editors. Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. CRC Press, Boca Raton, Florida.

Omernik, J. M., and G. E. Griffith. 1991. Ecological Regions Versus Hydrologic Units - Frameworks for Managing Water-Quality. Journal of Soil and Water Conservation 46:334-340.

Plafkin, J. L., M. T. Barbour, K. D. Porter, S. K. Gross, and R. M. Hughes. 1989. Rapid Bioassessment Protocols for Use in Streams and Rivers. Office of Water Regulations and Standards, U.S. Environmental Protection Agency, Washington, D.C.

Randall, R. G., and C. K. Minns. 2002. Comparison of a habitat productivity index (HPI) and an index of biotic integrity (IBI) for measuring the productive capacity of fish habitat in nearshore areas of the Great Lakes. Journal of Great Lakes Research 28:240-255.

Reynoldson, T. B., R. C. Bailey, K. E. Day, and R. H. Norris. 1995. Biological guidelines for freshwater sediment based on BEnthic Assessment of SedimenT (the BEAST) using a multivariate approach for predicting biological state. Australian Journal of Ecology 20:198-219.

Reynoldson, T. B., R. H. Norris, V. H. Resh, K. E. Day, and D. M. Rosenberg. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water quality impairment using benthic macroinvertebrates. Journal of the North American Benthological Society 16:833-852.

Reynoldson, T. B., D. W. Schloesser, and B. A. Manny. 1989. Development of a benthic invertebrate objective for mesotrophic Great Lakes waters. Journal of Great Lakes Research 15:669-689.

Ryder, R. A., and C. J. Edwards. 1985. A conceptual approach for the application of biological indicators of ecosystem quality in the Great Lakes basin. International Joint Commission, Windsor, Ontario.

Schneider, J. C., and J. H. Leach. 1977. Walleye fluctuations in the Great Lakes and possible causes, 1800-1975. Journal of the Fisheries Research Board of Canada 34:1878-1889.

Schuldt, J., L. Johnson, G. Host, C. Richards, J. Ciborowski. 2000. Protocols for selection of classification systems and reference conditions: A comparison of methods. Research proposal submitted to US EPA. US EPA/R-82877701-0. 15p.

Shannon, C. E., and W. Weaver. 1949. The Mathematical Theory of Communication. Pages 125p. University of Illinois Press. Urbana, Illinois.

Simon, T. P. 1991. Development of Ecoregion Expectations for the Index of Biotic Integrity. I. Central Corn Belt Plain. U.S. Environmental Protection Agency, Region 5, Chicago, Illinois.

Simon, T. P., and J. Lyons. 1995. Application of the index of biotic integrity to evaluate water resources integrity in freshwater ecosystems. Pages 245-262 in W. S. Davis and T. P. Simon, editors. Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. CRC Press, Boca Raton, Florida.

Smith, R. D., A. Ammann, C. Bartoldus, and M. M. Brinson. 1995. An approach for assessing wetland functions using hydrogeomorphic classification, reference wetlands and functional indices. U.S. Army Corps of Engineers, Waterways Experiment Station,, Vicksburg, MS, USA.

Sonstegard, R. A., and J. F. Leatherland. 1984. Great-Lakes Coho Salmon as an Indicator Organism for Ecosystem Health. Marine Environmental Research 14:480-480.

StatSoft Inc. 2001. STATISTICA (data analysis software system). Version 6.

Thoma, R. F. 1999. Biological monitoring and an index of biotic integrity for Lake Erie's nearshore waters. Pages 417-461 in T. P. Simon, editor. Assessing the Sustainability and Biological Integrity of Water Resources Using Fish Communities. CRC Press, Boca Raton, Florida.

U.S. Environmental Protection Agency. 1994. The U.S. EPA Reach File Version 3.0 Alpha (RF3-Alpha) Technical Reference. Office of Water, U.S. Environmental Protection Agency, Washington, D.C.

U.S. Environmental Protection Agency. 2004. Wetlands overview. U.S. Environmental Protection Agency, Office of Water., Washington, D.C.

Uzarski, D. G., T. M. Burton, M. J. Cooper, J. W. Ingram, and S. Timmermans. 2005. Fish habitat use within and across wetland classes in coastal wetlands of the five Great Lakes: development of a fish-based index of biotic integrity. Journal of Great Lakes Research. In press.

Waite, I. R., and K. D. Carpenter. 2000. Associations among fish assemblage structure and environmental variables in Willamette Basin streams, Oregon. Transactions of the American Fisheries Society 129:754-770.

Wang, L., J. Lyons, P. Rasmussen, P. Seelbach, T. Simon, M. Wiley, P. Kanehl, E. Baker, S. Niemela, and P. M. Stewart. 2003. Watershed, reach, and riparian influences in stream fish assemblages in the northern lakes and forest ecoregion, U.S.A. Canadian Journal of Fisheries and Aquatic Sciences 60:491-505.

Waters, I., and J. M. Shay. 1990. A Field-Study of the Morphometric Response of Typha-Glauca Shoots to a Water Depth Gradient. Canadian Journal of Botany-Revue Canadienne De Botanique 68:2339-2343.

Waters, I., and J. M. Shay. 1992. Effect of Water Depth on Population Parameters of a Typha-Glauca Stand. Canadian Journal of Botany-Revue Canadienne De Botanique 70:349-351.

Weaver, L. A., and G. C. Garman. 1994. Urbanization of a Watershed and Historical Changes in a Stream Fish Assemblage. Transactions of the American Fisheries Society 123:162-172.

Wei, A., P. Chow-Fraser, and D. Albert. 2004. Influence of shoreline features on fish distribution in the Laurentian Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 61:1113-1123.

Wichert, G. 1995. Effects of improved sewage effluent management and urbanization on fish associations of Toronto streams. North American Journal of Fisheries Management 15:440-456.

Wilcox, D. A., J. E. Meeker, P. L. Hudson, B. J. Armitage, M. G. Black, and D. G. Uzarski. 2002. Hydrologic variability and the application of index of biotic integrity metrics to wetlands: A Great Lakes evaluation. Wetlands 22:588-615.

Yoder, C. O. 1995. Policy issues and management applications of biological criteria. Pages 327-343 in W. S. Davis and T. P. Simon, editors. Biological Assessment and Criteria: Tools for Water Resource Planning and Decision Making. CRC Press, Boca Raton, Florida.

Appendix I. Site locations sampled under GLEI project (Site id beginning with 'GLE') and Reference Condition Project (Site id beginning with 'REF'). Hydrogeomorphic types are: high energy (He), embayment (Em), lacustrine wetland (lw), river-influenced wetland (rw) and protected wetland (pw). Ecoregions (Omernik 1987) are: Northern Superior Uplands (NSU), Southern Superior Uplands (SSU), Northern Great Lakes (NGL), Southern Great Lakes Morainal (SGL), South Central Great Lakes (SCG), and Erie and Ontario Lake Plains (EOL)

Site Name	Site ID	HGM	Ecoregion	Lake	Latitude	Longitude
Two Harbors	GLE_042	He	NSU	Superior	47.00	-91.69
Duluth	GLE_051	He	NSU	Superior	46.79	-92.08
Fond du Lac	GLE_055	Pw	NSU	Superior	46.66	-92.27
Dwights Point	GLE_059	Rw	SSU	Superior	46.68	-92.15
Allouez Bay	GLE_063	Em	SSU	Superior	46.69	-91.99
Middle River	GLE_068	Rw	SSU	Superior	46.69	-91.83
Clover	GLE_086	Pw	SSU	Superior	46.85	-91.20
Bark Bay	GLE_087	Em	SSU	Superior	46.86	-91.17
South Fish Creek	GLE_104	Rw	SSU	Superior	46.58	-90.94
Prentice Park	GLE_105	Pw	SSU	Superior	46.58	-90.93
Wood Creek	GLE_107	Cw	SSU	Superior	46.66	-90.74
Sand Cut Slough	GLE_109	Cw	SSU	Superior	46.66	-90.70
Bad River	GLE_110	Pw	SSU	Superior	46.62	-90.63
Agate Harbor	GLE_180	He	SSU	Superior	47.47	-88.12
Baraga	GLE_192	Em	SSU	Superior	46.75	-88.49
Baraga	GLE_192	Pw	SSU	Superior	46.75	-88.49
Peterson Creek	GLE_200	Pw	SSU	Superior	46.90	-88.19
Laughing Whitefish	GLE_213	He	NGL	Superior	46.54	-87.01
Laughing Whitefish	GLE_213	Rw	NGL	Superior	46.52	-87.03
Au Train	GLE_215	Rw	NGL	Superior	46.43	-86.83
Munising	GLE_217	Em	NGL	Superior	46.44	-86.63
McKay Creek	GLE_247	Cw	NGL	Huron	46.00	-84.30
McKay Creek	GLE_247	Em	NGL	Huron	46.00	-84.40
Pine River 1	GLE_248	Cw	NGL	Huron	46.01	-84.47
Hog Island Creek	GLE_257	He	NGL	Michigan	46.07	-85.30
Fishdom River	GLE_273	Em	NGL	Michigan	45.78	-86.57
Ogantz River	GLE_275	Cw	NGL	Michigan	45.86	-86.77
Ford River	GLE_281	Rw	NGL	Michigan	45.68	-87.14
Menominee River	GLE_289	He	SSU	Michigan	45.08	-87.59
Menominee River	GLE_289	Pw	SSU	Michigan	45.11	-87.59
Peshtigo River	GLE_291	Rw	SSU	Michigan	44.98	-87.66
Oconto River	GLE_294	Cw	SSU	Michigan	44.95	-87.71
Little Suamico River	GLE_299	Rw	NGL	Michigan	44.71	-87.99
Little Tail Point	GLE_300	Cw	NGL	Michigan	44.68	-88.00
Little Tail Point	GLE_300	Em	NGL	Michigan	44.66	-87.99
Saumico River	GLE_302	Cw	NGL	Michigan	44.62	-88.01
Sturgeon Bay	GLE_312	Em	NGL	Michigan	44.88	-87.44
Three Springs Creek	GLE_313	Em	NGL	Michigan	45.22	-87.02
Kewaunee River	GLE_321	Rw	NGL	Michigan	44.47	-87.51
West Twin River	GLE_325	He	NGL	Michigan	44.13	-87.58
Pike River	GLE_350	He	SGL	Michigan	42.64	-87.80
Trail Creek	GLE_359	He	SGL	Michigan	41.72	-86.94
Galien River	GLE_361	He	SCG	Michigan	41.80	-86.73

Galien River	GLE_361	Rw	SCG	Michigan	41.82	-86.73
Dunes State Park	GLE_364	Pw	SCG	Michigan	42.00	-86.55
Kalamazoo River	GLE_372	Pw	SCG	Michigan	42.67	-86.21
Port Shelden	GLE_374	He	SCG	Michigan	42.88	-86.22
Port Shelden	GLE_374	Rw	SCG	Michigan	42.90	-86.18
Mona lake	GLE_378	Rw	SCG	Michigan	43.18	-86.23
Muskegon Lake	GLE_379	Rw	NGL	Michigan	43.27	-86.22
Lincoln Lake	GLE_388	He	NGL	Michigan	43.99	-86.49
Big Sable	GLE_389	Rw	NGL	Michigan	44.08	-86.38
Arcadia Lake	GLE_393	He	NGL	Michigan	44.49	-86.22
Arcadia Lake	GLE_393	Rw	NGL	Michigan	44.45	-86.26
Crossville	GLE_407	Cw	NGL	Michigan	45.75	-84.98
Swan River	GLE_417	Em	NGL	Huron	45.41	-83.77
Ground Lake Outlet	GLE_419	Em	NGL	Huron	45.33	-83.48
Bell River	GLE_420	Em	NGL	Huron	45.25	-83.41
Bell River	GLE_420	Pw	NGL	Huron	45.33	-83.47
Long Luke Creek	GLE_421	Cw	NGL	Huron	45.08	-83.31
Thunder Bay River	GLE_422	Rw	NGL	Huron	45.07	-83.46
Butternut Creek	GLE_425	He	NGL	Huron	44.77	-83.28
Great Lakes Marina	GLE_442	Cw	NGL	Huron	43.94	-83.88
Cooper Drain	GLE_446	Cw	SCG	Huron	43.86	-83.91
Pinconning	GLE_447	Cw	SCG	Huron	43.84	-83.92
Johnson Drain	GLE_448	Cw	SCG	Huron	43.82	-83.91
Saginaw River	GLE_453	Cw	SCG	Huron	43.64	-83.81
Shebeon Creek	GLE_458	Cw	SCG	Huron	43.82	-83.42
Willow Creek	GLE_470	He	SCG	Huron	44.03	-82.78
Purdy Bay	GLE_486	Cw	SCG	Huron	43.76	-82.61
Sandy Creek	GLE_573	Rw	EOL	Erie	41.92	-83.33
Toledo Beach	GLE_578	He	EOL	Erie	41.82	-83.41
Toledo Beach	GLE_578	Pw	EOL	Erie	41.83	-83.42
Yetter Drain Creek	GLE_580	Pw	EOL	Erie	41.79	-83.46
Little Lake Creek	GLE_581	Em	EOL	Erie	41.75	-83.46
Maumee River	GLE_585	Em	EOL	Erie	41.72	-83.45
Cedar Point	GLE_586	He	EOL	Erie	41.68	-83.26
Crane Creek	GLE_588	Pw	EOL	Erie	41.64	-83.23
Toussaine River	GLE_590	Rw	EOL	Erie	41.58	-83.11
Little Muddy Creek	GLE_593	Pw	EOL	Erie	41.47	-82.99
Scherz Ditch	GLE_599	Cw	EOL	Erie	41.42	-82.94
Little Pickerel Creek	GLE_602	Pw	EOL	Erie	41.47	-82.81
Plum Brook	GLE_606	Cw	EOL	Erie	41.43	-82.64
Walnut Creek	GLE_645	He	EOL	Erie	42.10	-80.20
Erie	GLE_646	Em	EOL	Erie	42.13	-80.11
Sodus Creek	GLE_718	Pw	EOL	Ontario	43.26	-76.94
Mudge Creek	GLE_719	Rw	EOL	Ontario	43.28	-76.90
Sterling Creek	GLE_723	Pw	EOL	Ontario	43.34	-76.69
Deer Tick Creek	GLE_739	He	EOL	Ontario	43.60	-76.19
Deer Tick Creek	GLE_739	Rw	EOL	Ontario	43.61	-76.21
Skinner Creek	GLE_743	Cw	EOL	Ontario	43.68	-76.19
South Sandy Creek	GLE_744	Pw	EOL	Ontario	43.70	-76.19
Sackets Harbor	GLE_746	Em	EOL	Ontario	43.87	-76.21

Black River	GLE_748	Cw	EOL	Ontario	43.97	-76.12
Black River	GLE_748	Em	EOL	Ontario	43.98	-76.06
Guffin Creek	GLE_752	Em	EOL	Ontario	44.02	-76.15
Barnes Bay	GLE_755	Em	EOL	Ontario	44.05	-76.17
Cedar Point	REF_EOL-11	Pw	EOL	Erie	41.69	-83.31
Crane Creek	REF_EOL-14	Pw	EOL	Erie	41.62	-83.17
Cherry Island	REF_EOL-175	Cw	EOL	Erie	42.07	-83.19
2nd Creek	REFEOL-1831	Rw	EOL	Ontario	43.26	-77.00
Wolcott Creek	REF_EOL-187	Rw	EOL	Ontario	43.29	-76.85
Red Creek	REF_EOL-189	Rw	EOL	Ontario	43.30	-76.82
Erie	REF_EOL-193	Cw	EOL	Erie	42.16	-80.08
Braddock Bay	REF_EOL-195	Cw	EOL	Ontario	43.30	-77.72
Buttonwood Creek	REF_EOL-196	Cw	EOL	Ontario	43.30	-77.72
Rice Creek	REF_EOL-196	Rw	EOL	Ontario	43.44	-76.57
Blind Creek	REF_EOL-208	Rw	EOL	Ontario	43.66	-76.15
South Sandy Creek	REF_EOL-210	Rw	EOL	Ontario	43.72	-76.19
Springbrook	REF_EOL-24	Pw	EOL	Erie	41.44	-82.88
Red Creek	REF_EOL-41	Pw	EOL	Ontario	43.32	-76.75
Sterling Creek	REF_EOL-42	Pw	EOL	Ontario	43.36	-76.67
Cheboygan River	REF_NGL-118	Cw	NGL	Huron	45.66	-84.43
Grand Sable Dunes	REF_NGL-13	He	NGL	Superior	46.67	-86.03
Pine River 1	REF_NGL-33	Cw	NGL	Huron	45.97	-84.51
Greene Creek	REF_NGL-35	He	NGL	Huron	45.63	-84.19
Pine River 1	REF_NGL-36	Cw	NGL	Huron	46.03	-84.55
Platte River	REF_NGL-40	Pw	NGL	Michigan	44.84	-86.06
Cedar River	REF_NGL-44	He	NGL	Michigan	45.43	-87.32
Sucker River	REF_NGL-45	Rw	NGL	Superior	46.67	-85.94
Tahquamenon R.	REF_NGL-49	Rw	NGL	Superior	46.56	-85.04
Lake Leelanau	REF_NGL-51	He	NGL	Michigan	44.94	-85.97
Bear Lake	REF_NGL-51	Rw	NGL	Huron	45.98	-84.20
McKay Creek	REF_NGL-53	Rw	NGL	Huron	45.98	-84.36
Rapid River	REF_NGL-67	Rw	NGL	Michigan	45.92	-86.95
Sugar Creek	REF_NGL-69	He	NGL	Michigan	44.81	-87.65
Oconto River	REF_NGL-80	Cw	SSU	Michigan	44.88	-87.83
Paradise Beach	REF_NSU-2a	He	NSU	Superior	47.82	-90.02
Swamp Lake	REF_NSU-2b	He	NSU	Superior	47.87	-89.86
Spruce Creek	REF_NSU-5	He	NSU	Superior	47.68	-90.57
Manitou River	REF_NSU-6a	He	NSU	Superior	47.41	-91.10
Caribou River	REF_NSU-6b	He	NSU	Superior	47.46	-91.03
Eagle Bay	REF_SCG-1	He	SCG	Huron	44.07	-82.89
Sandpoint	REF_SCG-26	He	SCG	Huron	43.35	-82.52
Lakeport	REF_SCG-30	He	SCG	Huron	43.10	-82.46
Muskegon Lake	REF_SCG-31	He	NGL	Michigan	43.26	-86.37
Glenn Creek	REF_SCG-37	He	SCG	Michigan	42.53	-86.25
Turner Shores	REF_SCG-40	He	SCG	Michigan	41.87	-86.67
Crestview	REF_SGL-22	He	SGL	Michigan	42.79	-87.77
Conway Bay	REF_SSU-14	He	SSU	Superior	46.85	-87.73
Kenabeek Creek	REF_SSU-21	He	SSU	Superior	46.74	-89.94
Smith Creek 1	REF_SSU-24	He	SSU	Superior	46.75	-91.64
Bete Grise	REF_SSU-4	He	SSU	Superior	47.34	-87.93

Appendix II. Trophic guild (tguild) and tolerance designations for all fish species sampled (as per Barbour et al. 1999)

Family	Common name	Genus	Species	tguild	tolerance
Petromyzontidae	Sea Lamprey	*Petromyzon*	*marinus*	Piscivore	Moderate
Lepisosteidae	Spotted Gar	*Lepisosteus*	*oculatus*	Piscivore	Moderate
Lepisosteidae	Northern Longnose Gar	*Lepisosteus*	*osseus*	Piscivore	Moderate
Lepisosteidae	Shortnose Gar	*Lepisosteus*	*platostomus*	Piscivore	Moderate
Amiidae	Bowfin	*Amia*	*calva*	Piscivore	Moderate
Hiodontidae	Mooneye	*Hiodon*	*tergisus*	Insectivore	Intolerant
Anguillidae	American Eel	*Anguilla*	*rostrata*	Piscivore	Moderate
Clupeidae	Alewife	*Alosa*	*pseudoharengus*	Filterer	Moderate
Clupeidae	Gizzard Shad	*Dorosoma*	*cepedianum*	Omnivore	Moderate
Cyprinidae	Central Stoneroller	*Campostoma*	*anomalum*	Herbivore	Moderate
Cyprinidae	Lake Chub	*Couesius*	*plumbeus*	Insectivore	Moderate
Cyprinidae	Spotfin Shiner	*Cyprinella*	*spiloptera*	Insectivore	Moderate
Cyprinidae	Common Carp	*Cyprinus*	*carpio*	Omnivore	Tolerant
Cyprinidae	Streamline chub	*Erimystax*	*dissimilis*	Insectivore	Intolerant
Cyprinidae	Bigeye Chub	*Hybopsis*	*amblops*	Insectivore	Intolerant
Cyprinidae	Common Shiner	*Luxilus*	*cornutus*	Insectivore	Moderate
Cyprinidae	Horneyhead Chub	*Nocomis*	*biguttatus*	Insectivore	Intolerant
Cyprinidae	River Chub	*Nocomis*	*micropogon*	Insectivore	Intolerant
Cyprinidae	Golden Shiner	*Notemigonus*	*crysoleucas*	Omnivore	Tolerant
Cyprinidae	Pugnose Shiner	*Notropis*	*anogenus*	Insectivore	Intolerant
Cyprinidae	Bigeye Shiner	*Notropis*	*boops*	Insectivore	Intolerant
Cyprinidae	Blackchin Shiner	*Notropis*	*heterodon*	Insectivore	Intolerant
Cyprinidae	Blacknose Dace	*Rhinichthys*	*atratulus*	Generalist	Tolerant
Cyprinidae	Blacknose Shiner	*Notropis*	*heterolepis*	Insectivore	Intolerant
Cyprinidae	Brassy Minnow	*Hybognathus*	*hankinsoni*	Omnivore	Moderate
Cyprinidae	Bridled Shiner	*Notropis*	*bifrenatus*	Insectivore	Intolerant
Cyprinidae	Creek Chub	*Semotilus*	*atromaculatus*	Generalist	Tolerant
Cyprinidae	Emerald Shiner	*Notropis*	*atherinoides*	Insectivore	Moderate
Cyprinidae	Bluntnose Minnow	*Pimephales*	*notatus*	Omnivore	Tolerant
Cyprinidae	Fathead Minnow	*Pimephales*	*promelas*	Omnivore	Tolerant
Cyprinidae	Finescale Dace	*Chrosomus*	*neogaeus*	Insectivore	Moderate
Cyprinidae	Silver Shiner	*Notropis*	*photogenis*	Insectivore	Intolerant

Cyprinidae	Silvery Minnow	*Hybognathus*	*nuchalis*	Herbivore	Moderate
Cyprinidae	Goldfish	*Carassius*	*auratus*	Omnivore	Tolerant
Cyprinidae	Northern Mimic Shiner	*Notropis*	*volucellus*	Insectivore	Intolerant
Cyprinidae	Northern Redbelly Dace	*Chrosomus*	*eos*	Herbivore	Moderate
Cyprinidae	Northern River Shiner	*Notropis*	*blennius*	Insectivore	Moderate
Cyprinidae	Sand Shiner	*Notropis*	*stramineus*	Insectivore	Moderate
Cyprinidae	Spottail Shiner	*Notropis*	*hudsonius*	Insectivore	Moderate
Cyprinidae	Longnose Dace	*Rhinichthys*	*cataractae*	Insectivore	Intolerant
Cyprinidae	Pugnose Minnow	*Notropis*	*emilliae*	Insectivore	Intolerant
Cyprinidae	Rosyface Shiner	*Notropis*	*rubellus*	Insectivore	Intolerant
Catostomidae	Bigmouth Buffalo	*Ictiobus*	*cyprinellus*	Insectivore	Moderate
Catostomidae	Black Buffalo	*Ictiobus*	*niger*	Insectivore	Moderate
Catostomidae	Carpsucker	*Carpiodes*	*carpio*	Omnivore	Moderate
Catostomidae	E. Longnose Sucker	*Catostomus*	*catostomus*	Insectivore	Moderate
Catostomidae	Highfin Carpsucker	*Carpiodes*	*velifer*	Omnivore	Intolerant
Catostomidae	Northern Hogsucker	*Hypentelium*	*nigricans*	Insectivore	Intolerant
Catostomidae	Black Redhorse	*Moxostoma*	*duquesnii*	Insectivore	Intolerant
Catostomidae	Quillback Carpsucker	*Carpiodes*	*cyprinus*	Omnivore	Moderate
Catostomidae	River Redhorse	*Moxostoma*	*carinatum*	Insectivore	Intolerant
Catostomidae	Shorthead Redhorse	*Moxostoma*	*macrolepidotum*	Insectivore	Moderate
Catostomidae	Silver Redhorse	Moxostoma	anisurum	Insectivore	Moderate
Catostomidae	Spotted Sucker	Minytrema	melanops	Insectivore	Moderate
Catostomidae	White Sucker	Catostomus	commersoni	Omnivore	Tolerant
Ictaluridae	Black Bullhead	Ameriurus	melas	Insectivore	Moderate
Ictaluridae	Brown Bullhead	Ameriurus	nebulosas	Insectivore	Tolerant
Ictaluridae	Channel Catfish	Ameriurus	punctatus	Piscivore	Moderate
Ictaluridae	Stonecat	Noturus	flavus	Insectivore	Intolerant
Ictaluridae	Tadpole Madtom	Noturus	gyrinus	Insectivore	Moderate
Ictaluridae	Yellow Bullhead	Ameriurus	natalis	Insectivore	Tolerant
Esocidae	Grass Pickerel	Esox	americanus	Piscivore	Moderate
Esocidae	Great Lakes Muskellunge	Esox	masquinongy	Piscivore	Moderate
Esocidae	Northern Pike	Esox	lucius	Piscivore	Moderate
Umbridae	Central	Umbra	limi	Insectivore	Tolerant

	Mudminnow				
Osmeridae	Rainbow Smelt	Osmerus	mordax	Invertivore	Moderate
Salmonidae	Brook Trout	Salvelinus	fontinalis	Piscivore	Moderate
Salmonidae	Brown Trout	Salmo	trutta	Piscivore	Moderate
Salmonidae	Chinook Salmon	Oncorhynchus	tshawytscha	Piscivore	Moderate
Salmonidae	Coho Salmon	Oncorhynchus	kisutch	Piscivore	Moderate
Salmonidae	Rainbow Trout	Oncorhynchus	mykiss	Piscivore	Moderate
Percopsidae	Troutperch	Percopsis	omisomaycus	Insectivore	Moderate
Atherinidae	Northern Brook Silverside	Labidesthes	sicculus	Insectivore	Moderate
Fundulidae	Banded Killifish	Fundulus	diaphanus	Insectivore	Tolerant
Gasterosteidae	Brook Stickleback	Culaea	inconstans	Insectivore	Moderate
Gasterosteidae	Ninespine Stickleback	Pungitius	pungitius	Insectivore	Moderate
Gasterosteidae	Threespine Stickleback	Gasterosteus	aculeatus	Insectivore	Moderate
Cottidae	Slimy Sculpin	Cottus	cognatus	Insectivore	Intolerant
Cottidae	Mottled Sculpin	Cottus	bairdi	Insectivore	Intolerant
Serranidae	White Perch	Morone	americana	Piscivore	Moderate
Serranidae	White bass	Roccus	chrysops	Piscivore	Moderate
Serranidae	Yellow Bass	Morone	interrupta	Piscivore	Moderate
Centrarchidae	Black Crappie	Pomoxis	nigromaculatus	Piscivore	Moderate
Centrarchidae	Bluegill Sunfish	Lepomis	macrochirus	Insectivore	Moderate
Centrarchidae	Bluespotted Sunfish	Enneacanthus	gloriosus	Insectivore	Moderate
Centrarchidae	Green Sunfish	Lepomis	cyanellus	Insectivore	Tolerant
Centrarchidae	Largemouth Bass	Micropterus	salmoides	Piscivore	Moderate
Centrarchidae	Northern Rock Bass	Ambloplites	rupestris	Piscivore	Moderate
Centrarchidae	Orangespotted Sunfish	Lepomis	humilis	Insectivore	Moderate
Centrarchidae	Pumpkinseed Sunfish	Lepomis	gibbosus	Insectivore	Moderate
Centrarchidae	Smallmouth Bass	Micropterus	dolomieui	Piscivore	Moderate
Centrarchidae	Warmouth	Chaenobryttus	gulosus	Piscivore	Moderate
Centrarchidae	Western Redear Sunfish	Lepomis	microlophus	Insectivore	Moderate
Centrarchidae	White Crappie	Pomoxis	annularis	Piscivore	Moderate
Percidae	Blackside Darter	Percina	maculata	Insectivore	Moderate
Percidae	Fantail Darter	Etheostoma	flabellare	Insectivore	Moderate
Percidae	Iowa Darter	Etheostoma	exile	Insectivore	Moderate
Percidae	Johnny Darter	Etheostoma	nigrum	Insectivore	Moderate
Percidae	Least Darter	Etheostoma	micropera	Insectivore	Moderate

Percidae	Logperch	Percina	caprodes	Insectivore	Moderate
Percidae	Orangethroat Darter	Etheostoma	spectabile	Insectivore	Moderate
Percidae	River Darter	Percina	shumardi	Insectivore	Moderate
Percidae	Walleye	Stizostedion	vitreum	Piscivore	Moderate
Percidae	Yellow Perch	Perca	flavescens	Insectivore	Moderate
Percidae	Eurasian Ruffe	Gymnocephalus	cernuus	Insectivore	Moderate
Sciaenidae	Freshwater Drum	Aplodinotus	grunniens	Invertivore	Moderate
Gadidae	Burbot	Lota	lota	Piscivore	Moderate

Appendix III – Spearman rank correlations between candidate metrics for the EOL ecoregion

	#sp	#native species	#native minnows	# sunfish sp	%tol_sp	%intol sp	%green sunfish	% pisc	%insect cypr	% omn	CPUE	%non-indig	% DELT
#species	1.00	0.95	0.56	0.42	-0.03	0.33	0.29	0.20	0.25	0.42	0.26	0.46	0.02
#native species	0.95	1.00	0.52	0.52	-0.10	0.31	0.22	0.17	0.13	0.27	0.27	0.23	-0.03
#native minnows	0.56	0.52	1.00	-0.04	-0.15	0.35	0.03	-0.23	0.65	0.70	0.34	0.18	0.09
# sunfish sp	0.42	0.52	-0.04	1.00	-0.29	0.04	-0.10	0.35	-0.43	-0.14	0.21	0.03	-0.20
%tol sp	-0.03	-0.10	-0.15	-0.29	1.00	0.25	0.41	-0.36	-0.17	0.06	0.10	0.20	-0.22
%intol_sp	0.33	0.31	0.35	0.04	0.25	1.00	0.20	-0.20	0.18	0.19	0.14	0.12	0.06
%green sunfish	0.29	0.22	0.03	-0.10	0.41	0.20	1.00	-0.04	-0.06	0.08	0.22	0.17	0.21
%pisc	0.20	0.17	-0.23	0.35	-0.36	-0.20	-0.04	1.00	-0.28	-0.06	-0.07	0.30	0.07
%insect cypr	0.25	0.13	0.65	-0.43	-0.17	0.18	-0.06	-0.28	1.00	0.49	0.03	0.21	0.26
%omn	0.42	0.27	0.70	-0.14	0.06	0.19	0.08	-0.06	0.49	1.00	0.28	0.53	-0.15
CPUE	0.26	0.27	0.34	0.21	0.10	0.14	0.22	-0.07	0.03	0.28	1.00	0.08	-0.32
%non-indig	0.46	0.23	0.18	0.03	0.20	0.12	0.17	0.30	0.21	0.53	0.08	1.00	0.08
%DELT	0.02	-0.03	0.09	-0.20	-0.22	0.06	0.21	0.07	0.26	-0.15	-0.32	0.08	1.00

Appendix IV. Spearman rank correlations between candidate metrics for the NGL ecoregion.

	#sp	#nativ sp	#nativ cyp	#sunfis h	#sucker	% tol sp	%intol sp	%gr sunfish	%white suckers	%slimy sculpin	% pisc	%insec t cypr	% omn	CPU E	%non-indig	% DELT
#species	1.00	0.97	0.63	0.20	0.57	0.35	-0.12	0.61	0.26	-0.15	-0.27	0.13	0.13	0.58	0.42	-0.45
#native sp	0.97	1.00	0.70	0.21	0.63	0.36	-0.02	0.54	0.36	-0.07	-0.28	0.14	0.16	0.58	0.23	-0.38
#native minnows	0.63	0.70	1.00	0.01	0.56	0.23	0.43	0.21	0.37	0.16	-0.28	0.38	0.25	0.42	0.00	-0.10
#sunfish	0.20	0.21	0.01	1.00	0.09	-0.36	-0.22	-0.18	-0.09	-0.40	0.02	-0.05	-0.25	0.27	0.02	0.07
#suckers	0.57	0.63	0.56	0.09	1.00	0.33	0.09	0.21	0.71	0.15	-0.18	0.24	0.40	0.43	0.06	-0.03
% tol sp	0.35	0.36	0.23	-0.36	0.33	1.00	-0.21	0.31	0.11	-0.02	-0.31	-0.25	0.31	0.43	0.39	-0.48
%intol sp	-0.12	-0.02	0.43	-0.22	0.09	-0.21	1.00	-0.41	0.46	0.61	-0.02	0.26	0.46	-0.23	-0.47	0.38
%green sunfish	0.61	0.54	0.21	-0.18	0.21	0.31	-0.41	1.00	-0.06	-0.32	-0.17	-0.14	-0.18	0.27	0.46	-0.49
%white suckers	0.26	0.36	0.37	-0.09	0.71	0.11	0.46	-0.06	1.00	0.61	-0.21	0.13	0.63	0.16	-0.33	0.16
%slimy sculpins	-0.15	-0.07	0.16	-0.40	0.15	-0.02	0.61	-0.32	0.61	1.00	0.02	0.09	0.51	-0.17	-0.47	0.26
%pisc	-0.27	-0.28	-0.28	0.02	-0.18	-0.31	-0.02	-0.17	-0.21	0.02	1.00	0.21	-0.19	-0.37	-0.21	0.27
%insect cypr	0.13	0.14	0.38	-0.05	0.24	-0.25	0.26	-0.14	0.13	0.09	0.21	1.00	0.07	-0.10	-0.19	0.14
%omn	0.13	0.16	0.25	-0.25	0.40	0.31	0.46	-0.18	0.63	0.51	-0.19	0.07	1.00	-0.05	-0.02	-0.01
CPUE	0.58	0.58	0.42	0.27	0.43	0.43	-0.23	0.27	0.16	-0.17	-0.37	-0.10	-0.05	1.00	0.33	-0.48
%non-indig	0.42	0.23	0.00	0.02	0.06	0.39	-0.47	0.46	-0.33	-0.47	-0.21	-0.19	-0.02	0.33	1.00	-0.59
%DELT	-0.45	-0.38	-0.10	0.07	-0.03	-0.48	0.38	-0.49	0.16	0.26	0.27	0.14	-0.01	-0.48	-0.59	1.00

Appendix V. Spearman rank correlations between EOL and NGL metrics (after discarding highly correlated metrics) and 5 stressor gradients (measured at segment-shed scale).

EOL metrics	Atmospheric deposition	Land cover	Population density	Point source discharge	Shoreline modification
# native species	-0.16	0.17	0.14	0.09	0.04
CPUE	-0.12	0.28	-0.18	0.35	0.41
% tolerant species	-0.17	0.05	0.18	-0.06	0.06
% intolerant species	-0.27	0.21	-0.06	0.07	0.31
% piscivores	0.06	-0.14	-0.03	0.12	0.06
% insectivorous cyprinids	-0.39	0.24	0.16	0.12	0.07
% omnivores	-0.64	0.45	0.40	0.16	0.23
% DELT	-0.11	-0.23	0.05	0.03	-0.08

NGL metrics	Atmospheric deposition	Land cover	Population density	Point source discharge	Shoreline modification
# species	0.09	0.10	0.17	0.21	-0.01
% tolerant species	0.11	-0.12	0.09	0.05	0.25
% intolerant species	-0.64	-0.41	-0.23	-0.20	-0.22
# sunfish species	0.25	0.27	0.04	0.31	0.04
% green sunfish individuals	0.30	0.25	0.37	0.13	0.06
% non-indigenous species	0.57	0.44	0.43	0.26	0.31
% piscivores	-0.23	0.06	-0.14	-0.13	-0.11
% insectivorous cyprinids	-0.31	-0.22	-0.14	-0.40	-0.54
% omnivores	-0.47	-0.34	-0.24	-0.21	0.20

Appendix VI. Stressor scores (segment-shed scale) for Typha dominant and Scirpus dominant sites

Site name	Lake	Plant zone	Agriculture	Population density	Land cover	Atmospheric deposition	Point source discharge	Shoreline modification
Dwights Point	Superior	Typha	-0.28711	0.36848	0.36414	-0.57486	1.98732	-0.58421
Prentice Park	Superior	Typha	-0.28692	1.4688	1.89579	-0.4096	-0.50291	1.07126
Baraga	Superior	Typha	-0.7882	-0.00781	-0.44238	-0.6921	-0.24349	2.30629
Au Train	Superior	Typha	-0.56373	-0.18574	-0.38643	-0.1198	0.79456	-0.16875
Port Sheldon	Michigan	Typha	0.55292	-0.17985	0.90774	-1.77161	0.15604	1.43134
Mona Lake	Michigan	Typha	-1.21833	0.53207	2.29399	-1.38519	0.92869	1.43134
Muskegon Lake	Michigan	Typha	-1.43862	-0.66214	1.09559	-1.06629	0.60644	1.43134
Little Muddy Creek	Erie	Typha	1.49446	-0.14312	0.65316	0.309	1.26445	0.39335
Plum Brook	Erie	Typha	1.4608	0.32131	0.18125	0.2959	0.9356	-1.1728
Sodus Creek	Ontario	Typha	-0.24516	-0.14197	-0.43335	0.93701	-0.7734	-0.6873
Mudge Creek	Ontario	Typha	-0.25086	-0.17552	-0.40297	1.01767	-0.56332	-1.1728
South Sandy Creek	Ontario	Typha	-1.48922	-1.3472	-0.15008	2.31619	-0.07742	-0.85265
Cherry Island	Erie	Typha	-2.17459	0.76986	1.89508	-0.14689	-0.46634	-0.37283
Wolcott Creek	Ontario	Typha	-0.15711	0.11344	-0.23236	1.05718	0.5029	-1.1728
Red Creek	Ontario	Typha	-0.22984	-0.11061	-0.23488	1.19918	-0.6508	-1.1728
Braddock Bay	Ontario	Typha	0.01757	-0.08085	-0.2745	0.59642	-0.11401	0.24809
Buttonwood Creek	Ontario	Typha	0.00724	0.06696	-0.29792	0.63922	-0.7726	-1.1728
Blind Creek	Ontario	Typha	-1.56427	-0.14077	-0.05622	2.53928	-0.79514	-1.1728
South Sandy Creek	Ontario	Typha	-1.48922	-1.3472	-0.15008	2.31619	-0.07742	-0.85265

Continuation of Appendix VI.

Site name	Lake	Plant zone	Agriculture	Population density	Land cover	Atmospheric deposition	Point source discharge	Shoreline modification
Middle River	Superior	Scirpus	-0.18207	0.28872	0.17369	-0.4269	-0.42905	-0.58421
Clover	Superior	Scirpus	-0.28764	0.15501	-0.75654	-0.86189	-0.50292	-0.58421
McKay Creek	Huron	Scirpus	-0.54924	0.05255	-0.25338	-0.32316	-0.29702	-0.06912
Menominee River	Michigan	Scirpus	0.22765	0.34574	-0.54506	0.212	1.47964	2.16716
Pinconning	Huron	Scirpus	0.18094	-0.47618	-0.10284	-1.02872	0.71715	-1.1728
Little Pickerel Creek	Erie	Scirpus	1.4608	0.45902	0.80495	0.31669	-0.35352	0.74119
Sterling Creek	Ontario	Scirpus	-0.16053	-0.28429	-0.34078	1.35744	0.06977	-0.30475
Deer Tick Creek	Ontario	Scirpus	-1.56431	-0.69619	-0.03364	2.59221	-0.79514	-1.1728
Skinner Creek	Ontario	Scirpus	-1.49408	-0.96364	-0.17531	2.5004	-0.6442	-1.1728
Black River	Ontario	Scirpus	1.07982	0.15348	-0.32383	2.69027	1.9635	-0.58421
Sterling Creek	Ontario	Scirpus	-0.16053	-0.28429	-0.34078	1.35744	0.06977	-0.30475
Bear Lake	Huron	Scirpus	-0.51999	-0.81758	-0.58838	-0.2529	-0.52217	-0.58421
McKay Creek	Huron	Scirpus	-0.54924	0.05255	-0.25338	-0.32316	-0.29702	-0.06912

Appendix VII – Classifications of REF and GLEI sites based on discriminant function analysis (DFA) ('G' refers to group/cluster #). The 'Observed' column contains the grouping of the REF sites (beginning with ecoregion acronym) based on fish species relative abundance data. Column '1' contains the groupings of GLEI test sites based on the DFA analysis with 23 environmental variables

Site ID	Observed	1	2	3	4	5
EOL 175cw	G_1:1	G_1:1	G_5:5	G_3:3	G_2:2	G_4:4
EOL 1831rw	G_5:5	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
EOL 189rw	G_5:5	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
EOL 193cw	G_5:5	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
EOL 196cw	G_5:5	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
EOL 196rw	G_5:5	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
EOL 208rw	G_5:5	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
EOL 210rw	G_5:5	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
EOL 24pw	G_5:5	G_5:5	G_1:1	G_2:2	G_3:3	G_4:4
EOL 41pw	G_5:5	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
EOL 42pw	G_5:5	G_5:5	G_3:3	G_1:1	G_2:2	G_4:4
NGL 118cw	G_2:2	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
NGL 13he	G_3:3	G_3:3	G_2:2	G_1:1	G_5:5	G_4:4
* NGL 33cw	G_3:3	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
NGL 35he	G_3:3	G_3:3	G_1:1	G_2:2	G_5:5	G_4:4
NGL 36cw	G_2:2	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
NGL 40pw	G_5:5	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
NGL 44he	G_1:1	G_1:1	G_3:3	G_2:2	G_5:5	G_4:4
NGL 45rw	G_3:3	G_3:3	G_2:2	G_1:1	G_5:5	G_4:4
NGL 49rw	G_2:2	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
NGL 51he	G_1:1	G_1:1	G_3:3	G_2:2	G_5:5	G_4:4
NGL 51rw	G_2:2	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
NGL 53rw	G_2:2	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
NGL 67rw	G_2:2	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
NGL 69he	G_1:1	G_1:1	G_3:3	G_2:2	G_5:5	G_4:4
NGL 80cw	G_1:1	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4
NSU 2a He	G_4:4	G_4:4	G_3:3	G_2:2	G_1:1	G_5:5
NSU 2b He	G_4:4	G_4:4	G_3:3	G_2:2	G_1:1	G_5:5
NSU 5he	G_4:4	G_4:4	G_3:3	G_2:2	G_1:1	G_5:5
NSU 6a He	G_4:4	G_4:4	G_3:3	G_2:2	G_1:1	G_5:5
NSU 6b He	G_4:4	G_4:4	G_3:3	G_2:2	G_1:1	G_5:5
SCG 1he	G_2:2	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
SCG 26he	G_1:1	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4

Appendix VII Continued.

Site ID	Observed	1	2	3	4	5
SCG 30he	G_1:1	G_1:1	G_3:3	G_2:2	G_5:5	G_4:4
SCG 31he	G_1:1	G_1:1	G_3:3	G_2:2	G_5:5	G_4:4
SCG 37he	G_1:1	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4
SCG 40he	G_1:1	G_1:1	G_3:3	G_5:5	G_2:2	G_4:4
SGL 22he	G_1:1	G_1:1	G_3:3	G_2:2	G_5:5	G_4:4
SSU 14he	G_3:3	G_3:3	G_2:2	G_1:1	G_5:5	G_4:4
SSU 21he	G_1:1	G_1:1	G_3:3	G_2:2	G_5:5	G_4:4
SSU 24he	G_2:2	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
SSU 4he	G_2:2	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
042He1	---	G_4:4	G_3:3	G_2:2	G_1:1	G_5:5
051He1	---	G_4:4	G_3:3	G_2:2	G_1:1	G_5:5
055Pw1	---	G_4:4	G_5:5	G_2:2	G_3:3	G_1:1
059Rw1	---	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
063Em7	---	G_2:2	G_1:1	G_3:3	G_5:5	G_4:4
068Rw7	---	G_2:2	G_1:1	G_3:3	G_5:5	G_4:4
086Pw6	---	G_2:2	G_3:3	G_5:5	G_1:1	G_4:4
087Em2	---	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
104Rw3	---	G_2:2	G_1:1	G_3:3	G_5:5	G_4:4
105Pw5	---	G_2:2	G_5:5	G_3:3	G_1:1	G_4:4
107Cw4	---	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
109Cw1	---	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
110Pw1	---	G_2:2	G_5:5	G_3:3	G_1:1	G_4:4
192Em8	---	G_2:2	G_1:1	G_3:3	G_5:5	G_4:4
192Pw11	---	G_2:2	G_5:5	G_1:1	G_3:3	G_4:4
200Pw2	---	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
217Em3	---	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
247Cw13	---	G_3:3	G_2:2	G_1:1	G_5:5	G_4:4
247Em11	---	G_2:2	G_1:1	G_3:3	G_5:5	G_4:4
248Cw2	---	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
257He1	---	G_3:3	G_2:2	G_1:1	G_5:5	G_4:4
273Em3	---	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4
275Cw1	---	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
281Rw2	---	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
289He1	---	G_2:2	G_1:1	G_3:3	G_5:5	G_4:4
289Pw10	---	G_2:2	G_1:1	G_5:5	G_3:3	G_4:4
291Rw10	---	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4
294Cw2	---	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4

Appendix VII Continued

Site ID	1	2	3	4	5
299Rw2	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4
300Cw2	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4
300Em8	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4
302Cw2	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4
312Em3	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
313Em6	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4
321Rw1	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4
325He2	G_1:1	G_3:3	G_2:2	G_5:5	G_4:4
350He1	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4
359He1	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4
361He1	G_1:1	G_3:3	G_2:2	G_5:5	G_4:4
361Rw2	G_1:1	G_2:2	G_5:5	G_3:3	G_4:4
364Pw3	G_1:1	G_5:5	G_2:2	G_3:3	G_4:4
372Pw3	G_5:5	G_1:1	G_2:2	G_3:3	G_4:4
374He12	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4
374Rw2	G_1:1	G_5:5	G_2:2	G_3:3	G_4:4
378Rw2	G_1:1	G_5:5	G_2:2	G_3:3	G_4:4
379Rw1	G_2:2	G_1:1	G_3:3	G_5:5	G_4:4
388He1	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4
389Rw11	G_2:2	G_1:1	G_3:3	G_5:5	G_4:4
393He1	G_1:1	G_3:3	G_2:2	G_5:5	G_4:4
393Rw2	G_5:5	G_2:2	G_1:1	G_3:3	G_4:4
417Em4	G_3:3	G_2:2	G_1:1	G_5:5	G_4:4
419Em6	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
420Em9	G_1:1	G_3:3	G_2:2	G_5:5	G_4:4
420Pw1	G_5:5	G_2:2	G_1:1	G_3:3	G_4:4
421Cw2	G_1:1	G_3:3	G_2:2	G_5:5	G_4:4
422Rw2	G_2:2	G_1:1	G_3:3	G_5:5	G_4:4
425He1	G_1:1	G_2:2	G_3:3	G_5:5	G_4:4
442Cw1	G_2:2	G_1:1	G_3:3	G_5:5	G_4:4
446Cw1	G_2:2	G_1:1	G_3:3	G_5:5	G_4:4
447Cw2	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4

Appendix VII Continued

Site ID	1	2	3	4	5
448Cw1	G_2:2	G_1:1	G_3:3	G_5:5	G_4:4
453Cw2	G_2:2	G_1:1	G_3:3	G_5:5	G_4:4
458Cw2	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
470He1	G_2:2	G_3:3	G_1:1	G_5:5	G_4:4
486Cw2	G_2:2	G_1:1	G_3:3	G_5:5	G_4:4
573Rw7	G_1:1	G_5:5	G_3:3	G_2:2	G_4:4
578He2	G_1:1	G_5:5	G_3:3	G_2:2	G_4:4
578Pw3	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
580Pw5	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
581Em3	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
585Em2	G_1:1	G_5:5	G_3:3	G_2:2	G_4:4
586He1	G_1:1	G_5:5	G_3:3	G_2:2	G_4:4
588Pw6	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
590Rw1	G_1:1	G_5:5	G_3:3	G_2:2	G_4:4
593Pw6	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
599Cw3	G_1:1	G_5:5	G_3:3	G_2:2	G_4:4
602Pw1	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
606Cw1	G_1:1	G_5:5	G_3:3	G_2:2	G_4:4
645He1	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
646Em2	G_1:1	G_5:5	G_3:3	G_2:2	G_4:4
718Pw1	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
719Rw1	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
723Pw1	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
739Rw4	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
743Cw3	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
744Pw1	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
746Em9	G_5:5	G_3:3	G_1:1	G_2:2	G_4:4
748Cw7	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
748Em5	G_5:5	G_1:1	G_3:3	G_2:2	G_4:4
752Em2	G_5:5	G_3:3	G_1:1	G_2:2	G_4:4
755Em3	G_5:5	G_3:3	G_1:1	G_2:2	G_4:4

VITA AUCTORIS

NAME:	Yakuta Bhagat
PLACE OF BIRTH:	Gujarat, India
YEAR OF BIRTH:	1978
EDUCATION:	Marc Garneau Collegiate Institute, Toronto, Ontario 1995-1996
	University of Guelph, Guelph, Ontario 1997-2000
	University of Windsor, Windsor, Ontario 2003-2005